

Université  
de Toulouse

# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

---

**Présentée et soutenue par :**

**Sarah ROUSSEAU**

Le vendredi 13 décembre 2013

**Titre :**

CLONAGE POSITIONNEL DE GENES PREDISPOSANT  
AUX ANOMALIES CONGÉNITALES MAJEURES DANS L'ESPECE PORCINE

---

**École doctorale et discipline ou spécialité :**

ED BSB : Génétique moléculaire

**Unité de recherche :**

Laboratoire de Génétique Cellulaire, INRA, Castanet-Tolosan

**Directeur(s) de Thèse :**

Mr Denis MILAN - Directeur de Recherche, INRA, Castanet-Tolosan

Mme Juliette RIQUET - Directrice de Recherche, INRA, Castanet-Tolosan

**Rapporteurs :**

Mr Didier BOICHARD - Directeur de Recherche, INRA, Jouy-en-Josas

Mr Philippe MONGET - Directeur de Recherche, INRA, Nouzilly

**Autre(s) membre(s) du jury :**

Mme Catherine SOULA - Professeur, Université P. Sabatier, Toulouse (Présidente du jury)

Mr Alain DUCOS - Directeur de Recherche, ENVT, Toulouse (Examineur)

Mr Aurélien CAPITAN - Docteur-Ingénieur, UNCEIA, Jouy-en-Josas (Examineur)

Cette thèse CIFRE a été financée par le laboratoire d'analyses génétiques pour les espèces animales LABOGENA, Jouy-en-Josas, et réalisée au laboratoire de génétique cellulaire LGC, INRA Toulouse.



Σε σένα αγάπη μου,  
σ'ευχαριστώ που υπάρχεις κάθε μέρα στην καρδιά μου, σ'ευχαριστώ για την  
υποστήριξή σου και την αγάπη σου.

To M. and D.,  
Thanks for your support and your love for nearly 30 years.  
This work is yours.  
K&L



A la mémoire de Papou

*J'ai mis mon képi dans la cage  
et je suis sorti avec l'oiseau sur la tête*

*Alors*

*on ne salue plus  
a demandé le commandant*

*Non*

*on ne salue plus  
a répondu l'oiseau*

*Ah bon*

*excusez moi je croyais qu'on saluait  
a dit le commandant*

*Vous êtes tout excusé tout le monde peut se tromper  
a dit l'oiseau.*



# Remerciements

---

Juliette,

Tout d'abord, merci pour la chance que tu m'as donnée de travailler sur ce projet, si complet et si passionnant, alors que je ne soupçonnais même pas il y a trois ans l'existence du microsattellite. J'ai beaucoup appris à tes côtés, avec passion et enthousiasme. Tu n'as pas laissé échapper ton scepticisme quant à mes hypothèses scientifiques souvent farfelues, mais tu m'as encouragé et aiguillé lors de mes doutes et mes déceptions. Enfin, merci pour ton aide et ton implication ces derniers mois.

Thomas,

Thomas est le genre de collègue qui donne envie d'aller au travail le matin. Brillant et d'une grande humilité, pragmatique et plein d'humour, tu as toujours été de très bon conseil, tant au point professionnel que personnel. Je te remercie pour tout ce que tu m'as apporté, et pour m'avoir aidé à passer les moments difficiles.

Emmanuelle,

Merci pour ton écoute, ton réconfort, ton coaching, et ta joie de vivre. Quelle chance nous avons eu de t'avoir à nos cotés. Bon courage pour la suite, et merci pour tout.

Eric, Annie, Nath, Katia, Sylvain, Flo, Julie, Patrice, Olivier,

Je vous remercie du fond du cœur pour votre aide et votre soutien dans l'aboutissement de ce projet. Ce fut un plaisir de travailler à votre coté, et vous m'avez apporté bien plus qu'une collaboration professionnelle. Merci plus particulièrement à Annie, pour ton aide dans la finalisation du manuscrit.

A vous tous membres du LGC,

Il semble difficile de pouvoir réaliser une thèse dans des conditions aussi favorables: joie et bonne humeur, petites attentions et café chaud au quotidien, toujours à l'écoute et prêt a rendre service, merci pour cette belle preuve d'humanisme. Maria-Inès, Laure F, Laure G, Manu, Annabelle, Maxime, Hervé, David, je compte bien me rattraper sur les sorties LGCiennes auxquelles je n'ai pues participer.



Merci également à Jean-Michel, Marie-Yvonne, Eric et Julie à nouveau, pour vos conseils et vos encouragements en tant que membres de mon comité de thèse.

Merci à Marie-Pierre et Jean-Benoît pour m'avoir aidée et soutenue dans ma demande de passerelle, et pour m'avoir apporté vos connaissances en génétique humaine et dépistage néo-natal.

Je remercie très sincèrement le Pr Galinier et le Dr Juricic, avec qui j'ai eu la chance d'assister aux activités chirurgicales et pédiatriques. Ces activités m'ont plus que motivé dans mon projet futur, et ont enrichi mon travail de thèse.

Merci à vous les copains, Bertrand, Pierre, Fée, Val, Juice, Chloé, Seb, Guitz, Pierrot, Emilie, Erwan, à mes chères voisines Jeanne et Agnès, pour m'avoir soutenue durant ces trois années. Vos pensées et attentions ont beaucoup compté pour moi, et je vous en remercie du fond du cœur. Mille mercis à Marie pour la correction du manuscrit.

Enfin et surtout, merci à ma famille et à ma belle-famille. Merci à mes sista lova Mathilde et Ingrid, dad et mum, Joelle et Dimitri, pour avoir cru en moi et m'avoir soutenue dans les moments difficiles, et encore une fois merci à mon M. pour ta patience, ton amour et ton soutien.

Pour finir, quelles phrases de grands philosophes français du XXI siècle, rencontrés durant ces trois années :

« RTFM »

S. F.

« C'est quoi la vraisemblance ?! »

B. S.

« [...], C'est pas si pire »

M. F.

« Ah oui, il est plutôt bayésien lui non ? »

D. R.

« Tu sais Sarah, on a tous survécus [à la thèse] »

J. D., S. F., H. A., T. F., A. R., J. R., ...



# Table des matières

---

<b>INTRODUCTION .....</b>	<b>23</b>
<b>DONNEES BIBLIOGRAPHIQUES.....</b>	<b>27</b>
<b>I. IMPACT DES ANOMALIES CONGENITALES EN PRODUCTION PORCINE.....</b>	<b>25</b>
I.A. QUELQUES ELEMENTS DE CONTEXTE AGRO-ECONOMIQUE .....	25
<i>I.A.1. La consommation de viande de porc dans le monde.....</i>	<i>25</i>
<i>I.A.2. La production porcine à l'échelle mondiale.....</i>	<i>27</i>
I.B. LES ELEVAGES PORCINS ET LA SELECTION GENETIQUE.....	29
I.C. LES MALADIES GENETIQUES DANS L'ESPECE PORCINE.....	37
<i>I.C.1. Impact des maladies congénitales sur la production porcine.....</i>	<i>37</i>
<i>I.C.2. Les maladies congénitales chez le porc .....</i>	<i>39</i>
I.D. DETERMINISME GENETIQUE DES ANOMALIES CONGENITALES ETUDIEES.....	41
<i>I.D.1. Etudes de ségrégation et estimations des paramètres génétiques.....</i>	<i>41</i>
<i>I.D.2. Causes moléculaires des anomalies congénitales .....</i>	<i>43</i>
<b>II. LES OUTILS DE CARTOGRAPHIE DU GENOME DU PORC.....</b>	<b>45</b>
II.A LA CARTOGRAPHIE PHYSIQUE .....	45
<i>II.A.1. Cartographie physique chromosomique.....</i>	<i>45</i>
<i>II.A.2 Cartographie physique moléculaire.....</i>	<i>49</i>
II.B LES VARIATIONS DE L'ADN .....	55
<i>II.B.1.Les macro-remaniements chromosomiques.....</i>	<i>55</i>
<i>II.B.2. Les variations du nombre de copies d'une séquence .....</i>	<i>57</i>
<i>II.B.3. Les séquences répétées du génome.....</i>	<i>61</i>
<i>II.B.4. Les SNP et la puce 60K porcine .....</i>	<i>65</i>
II.C LA CARTOGRAPHIE GENETIQUE .....	69
<i>II.C.1. La liaison génétique.....</i>	<i>69</i>
<i>II.C.2. Le déséquilibre de liaison chez le porc .....</i>	<i>73</i>
<i>II.C.3. Recherche de mutations causales.....</i>	<i>77</i>
<b>III. PHYSIOPATHOLOGIE DE L'INTERSEXUALITE, DE LA CRYPTORCHIDIE ET DE LA HERNIE SCROTALE CHEZ L'HOMME ET CHEZ D'AUTRES MODELES .....</b>	<b>85</b>
III.A. INTRODUCTION AUX ANOMALIES CONGENITALES HUMAINES.....	85
III.B. DIFFERENCIATION SEXUELLE CHEZ LES MAMMIFERES .....	91
<i>III.B.1. Mise en place du sexe chromosomique : .....</i>	<i>91</i>

III.B.2. Déterminisme du sexe gonadique : .....	93
III.B.3. Différenciation du sexe phénotypique : .....	109
III.C. ETIOPATHOGENIE DES ANOMALIES ETUDIEES .....	113
III.C.1. DSD 46,XX – désordre du développement ovarien .....	113
III.C.2. La cryptorchidie .....	127
III.C.3. La hernie inguino-scrotale .....	139
<b>MATERIEL ET METHODES.....</b>	<b>155</b>
<b>I. GENOTYPAGE .....</b>	<b>151</b>
I.A. TEST DIAGNOSTIQUE "SRY" .....	151
I.B. CONTROLE DE FILIATION.....	151
I.C. GENOTYPAGE A L'AIDE LA PUCE PORCINE HAUTE DENSITE D'ILLUMINA .....	153
I.D. GENOTYPAGE A L'AIDE DE LA TECHNOLOGIE FLUIDIGM.....	158
I.E. GENOTYPAGE DU PANEL PORCIN D'HYBRIDES D'IRRATION .....	162
<b>II. SEQUENÇAGE .....</b>	<b>166</b>
II.A. SEQUENÇAGE AUTOMATISE SELON LA METHODE DE SANGER .....	166
II.B. SEQUENÇAGE HAUT DEBIT A L'AIDE DU SEQUENCEUR HISeq 2000 .....	168
II.B.1. Contrôle de la qualité globale du séquençage.....	171
II.B.2. Alignement des séquences et critères de sélection lecture-spécifique : .....	173
<b>III. ANALYSES DES DONNEES DE GENOTYPAGE .....</b>	<b>175</b>
III.A. RECHERCHE DE MOTIFS REPETES .....	175
III.B. HAPLOTYPES ET CONSTRUCTION DE CARTE GENETIQUE .....	177
III.C. ANALYSES TDT .....	179
III.D. CONSTRUCTION DE CARTE PHYSIQUE .....	181
<b>IV. ANALYSES DES DONNEES DE SEQUENÇAGE.....</b>	<b>187</b>
IV.A. RECHERCHE D'ELEMENTS REPETES <i>IN SILICO</i> .....	187
IV.A.1. Recherche de motifs répétés en tandem.....	187
IV.A.2. Recherche d'éléments dispersés connus et de séquences de faible complexité.....	189
IV.B. RECHERCHE DE POLYMORPHISMES PAR COMPARAISON DE SEQUENCE .....	189
 <b>CHAPITRE 1 : GWAS pour la détection de QTL associés aux principales anomalies congénitales chez le porc.....</b>	 <b>195</b>
<b>I. PRESENTATION DU DISPOSITIF FAMILIAL MIS EN PLACE POUR LES ETUDES TDT .....</b>	<b>196</b>
I.A. MISE EN PLACE D'UNE COLLECTION D'ECHANTILLONS D'ANIMAUX PHENOTYPES.....	196

*I.A.1. La collection ANOPORC* ..... 196

I.A.2. Premier bilan de la collection ANOPORC : .....	198
I.A.3. Le phénotypage des animaux collectés .....	200
I.B. ANIMAUX GENOTYPES DANS LE CADRE DU PROJET SWAN .....	202
I.B.1. Critères de sélection des animaux génotypés .....	202
I.B.2. Bilan des génotypages réalisés dans le cadre de SwAn .....	206
<b>II. ESTIMATION DE LA QUALITE ET DE L'INFORMATIVITE DE LA PUCE PORCINE .....</b>	<b>210</b>
II.A. ESTIMATION DE LA COUVERTURE DU GENOME .....	210
II.B. INFORMATIVITE DE LA PUCE POUR LES ANALYSES TDT .....	212
II.B.1. Informativité des marqueurs (MAF) .....	212
II.B.2. Informativité du dispositif TDT - proportion de parents hétérozygotes .....	212
II.C. ESTIMATION DU DESEQUILIBRE DE LIAISON ENTRE MARQUEURS: .....	216
<b>III. RESULTATS DES ANALYSES TDT POUR LES TROIS ANOMALIES CONGENITALES .....</b>	<b>220</b>
III.A. STRATEGIE D'ANALYSE ADOPTEE .....	220
III.A.1. Hypothèses d'étude et modèles utilisés.....	220
III.A.2. Procédure suivie et mise en forme des résultats .....	222
III.B. RESULTATS OBTENUS POUR CHAQUE ANOMALIE CONGENITALE .....	224
III.B.1. L'intersexualité .....	224
III.B.2. La hernie inguino-scrotale .....	228
III.B.3. La Cryptorchidie.....	238
III.C. DISCUSSION .....	242

## **CHAPITRE II : Cartographie fine de la région candidate à l'intersexualité chez le porc.....269**

<b>I. ARTICLE 1 : A GENOME-WIDE ASSOCIATION STUDY POINTS OUT THE CAUSAL IMPLICATION OF SOX9 IN THE SEX-REVERSAL PHENOTYPE IN XX PIGS.....</b>	<b>270</b>
I.A. RESUME DE L'ARTICLE.....	272
I.B. CONCLUSION .....	274
<b>II. RESEQUENÇAGE COMPLET.....</b>	<b>278</b>
II.A. STRATEGIE DE SEQUENÇAGE UTILISEE.....	278
II.B. RESULTATS .....	282
II.B.1. Nombre de séquences obtenues.....	282
II.B.2. Contrôle qualité .....	284
II.B.3. Alignement des lectures .....	286

II.C. ETUDE DE LA COUVERTURE DANS LA REGION CANDIDATE .....	290
---	-----

<b>III. RECHERCHE DE POLYMORPHISMES DANS LA REGION CANDIDATE A L'INTERSEXUALITE.....</b>	<b>294</b>
III.A. RECHERCHE EXHAUSTIVE DE POLYMORPHISMES .....	294
III.B. VALIDATION DES POLYMORPHISMES IDENTIFIES .....	297
III.C. IDENTIFICATION DE POLYMORPHISMES ASSOCIES A L'INTERSEXUALITE .....	299
<b>IV. RECHERCHE DE SOURCES DE VARIABILITE AUTRES QUE LES POLYMORPHISMES .....</b>	<b>303</b>
IV.A. RECHERCHE DE CNV.....	303
IV.B. ETUDE DES SITES DE FIXATION .....	305
<b>V. DENSIFICATION DE LA REGION CANDIDATE EN MARQUEURS SNP .....</b>	<b>309</b>
V.A. CHOIX DES MARQUEURS ET DES ANIMAUX .....	309
V.B. ANALYSES TDT A L'AIDE DES DONNEES DE GENOTYPAGE FLUIDIGM .....	313
V.C. ETUDE HAPLOTYPIQUE.....	315
<b>VI. DISCUSSION .....</b>	<b>321</b>

## **CHAPITRE III : Cartographies physique et génétique du chromosome X porcin.....397**

<b>I. ETUDE DE L'IMPLICATION DE LA REGION DU CHROMOSOME X DANS LE DETERMINISME DE LA HERNIE INGUINO-SCROTALE, DE LA CRYPTORCHIDIE ET DE L'INTERSEXUALITE.....</b>	<b>337</b>
I.A. LOCALISATION DE LA REGION PSEUDO-AUTOSOMALE (PAR) PORCINE .....	337
I.B. TRANSMISSION DES ALLELES PARX ET PARY .....	339
I.C. TAUX DE RECOMBINAISON SUR PAR CHEZ LES MALES ET LES FEMELLES .....	341
I.D. FREQUENCES ALLELIQUES ET HAPLOTYPIQUES DES MARQUEURS PARX ET PARY.....	343
I.E. CONCLUSION .....	347
<b>II. ARTICLE 2: A GENETIC AND PHYSICAL FINE CHARACTERIZATION OF PIG SSCX .....</b>	<b>351</b>

## **DISCUSSION GENERALE ET CONCLUSION.....491**

## **REFERENCES BIBLIOGRAPHIQUES.....515**

## **ANNEXES.....553**



# Liste des figures

FIGURE 1 : CONSOMMATION MOYENNE DE VIANDE PORCINE EN KILOGRAMMES PAR HABITANT DANS LE MONDE EN 2010.....	24
FIGURE 2 : EVOLUTION DE LA CONSOMMATION DE VIANDE DANS LE MONDE ET PAR TYPE DE PAYS EN KG/HABITANT/AN - SOURCE: FRANCEAGRIMER D'APRES FAO, OCDE ESTIMATIONS .....	24
FIGURE 3 : PRODUCTION, CONSOMMATION, EXPORTATION ET RATIO ENTRE LA PRODUCTION ET LA CONSOMMATION POUR LES PRINCIPAUX PAYS CONSOMMATEURS DE VIANDE DE PORC AU MONDE .....	26
FIGURE 4 : PRINCIPAUX COURANTS D'ECHANGES INTERNATIONAUX DE VIANDE DE PORC (COMMERCE MONDIAL 2009, OBSERVATOIRE INTERNATIONAL DE L'IFIP). .....	26
FIGURE 5 : STRUCTURE EN TROIS NIVEAUX DU SCHEMA DE PRODUCTION PORCINE FRANÇAISE (SOURCES: AGENCE DE LA SELECTION PORCINE 2009, IFIP 2009, SPP AGRESTE ENQUETE STRUCTURES 2007) .....	30
FIGURE 6 : EVOLUTION DES PERFORMANCES POUR LES QUATRE RACES PRINCIPALEMENT UTILISEES EN FRANCE. LF : LANDRACE FRANÇAIS, LWF/M : LARGE WHITE FRANÇAIS FEMELLE/MALE, PP : PIETRAIN. ....	32
FIGURE 7 : FREQUENCE DES PRINCIPALES ANOMALIES CONGENITALES DANS L'ESPECE PORCINE SELON UNE ETUDE EUROPEENNE REALISEE ENTRE 2006 ET 2010 (WALTERS, 2010). ....	40
FIGURE 8 : PRINCIPE DE LA CREATION DES PANELS D'HYBRIDES D'IRRADIATION.....	46
FIGURE 9 : PRINCIPE DE L'ESTIMATION DE L'ORDRE DES LOCI LES UNS PAR RAPPORT AUX AUTRES. EXEMPLE AVEC 3 MARQUEURS ET 4 HYBRIDES (LA PRESENCE D'UN MARQUEUR DANS UN HYBRIDE EST NOTE "+", SON ABSENCE EST NOTE "-").....	46
FIGURE 10 : CONSTRUCTION DES BANQUES DE DONNEES DE SEQUENCES CHEZ LE PORC .....	48
FIGURE 11 : CONSTRUCTION DU DRAFT PORCIN .....	48
FIGURE 12 : NOMBRE DE CONTIGS PAR CLONE AVANT ( <i>POST SHOTGUN</i> ) ET APRES L'ETAPE DE <i>PRIMER WALKING (=AUTOPREFIN)</i> ....	50
FIGURE 13 : VALEURS STATISTIQUES N50, N75 ET N90 POUR LES SCAFFOLDS DES DRAFTS V9 ET V10.2 .....	50
FIGURE 14 : LONGUEUR TOTALE (EN PB) ET HORS GAPS PAR CHROMOSOME SUR LES DRAFTS 9 ET 10.2 .....	52
FIGURE 15 : MODIFICATIONS DE STRUCTURE AFFECTANT UN CHROMOSOME.....	54
FIGURE 16 : MODIFICATIONS DE STRUCTURE IMPLIQUANT DEUX CHROMOSOMES .....	54
FIGURE 17 : REPRESENTATION DE CNV : A GAUCHE LA DELETION D'UNE COPIE DU LOCUS C, A DROITE LA DUPLICATION D'UNE COPIE DU LOCUS C. ....	56
FIGURE 18 : EXEMPLE DE LA VARIATION DU NOMBRE DE COPIES DU GENE KIT CODANT POUR LA COULEUR DE LA PEAU (LEE AND JEON, 2008).....	58
FIGURE 19 : LES DIFFERENTES CATEGORIES D'ELEMENTS TRANPOSES ET LEUR TAILLE (WILSON ET AL., 2007). A : REPETITIONS SIMPLES, B : TRANSPOSON, C : LTR, D : LINE, E : SINE. LES BARRES AU BAS DE L'IMAGE INDIQUENT LA TAILLE DES DIFFERENTS ELEMENTS. ....	60
FIGURE 20 : COMPARAISON DES FONCTIONS DE HALDANE ET KOSAMBI .....	68
FIGURE 21 : DISTANCE GENETIQUE EN FONCTION DU TAUX DE RECOMBINAISON .....	70

FIGURE 22 : REGIONS CHROMOSOMIQUES ETUDIEES PAR AMARAL ET AL ET REPARTITION DES SNP. LES POSITIONS DES SNP SONT INDIQUEES PAR DES TRAIT VERTICAUX, ET LES NOMBRES INDICENT LE NOMBRE TOTAL DE SNP PAR REGION. LES BARRES HORIZONTALES GRISES INDICENT LES BAC. TAILLE DE REGION A : 0.5 Mb, REGION B : 3 Mb, REGION C : 3.4 Mb. .... 74

FIGURE 23 : DL EN FONCTION DE LA DISTANCE GENOMIQUE (PB) PAR POPULATION ET PAR REGION ETUDIEE .....	74
FIGURE 24 : DIFFERENTES APPROCHES DE CARTOGRAPHIE GENETIQUE UTILISEES DANS LA RECHERCHE DE GENES A EFFET QUANTITATIF. .....	76
FIGURE 25 : NOMBRE DE GWAS DEPUIS 2005 .....	80
FIGURE 26 : ETAPES DE L'ORGANOGENESE CHEZ L'HOMME .....	84
FIGURE 27 : ETAPES DE LA DIFFERENCIATION SEXUELLE, DE LA FECONDATION A LA 38 <sup>EME</sup> SEMAINE DE LA VIE EMBRYONNAIRE CHEZ L'HOMME. LES SEMAINES DE GESTATION (NOTEES SG EN NOIR) SONT INDIQUEES CHEZ L'HOMME .....	91
FIGURE 28: A : EBAUCHE DES CRETES GENITALES B : MIGRATION DES CELLULES GERMINALES. ....	92
FIGURE 29 : GENES IMPLIQUES DANS LA DIFFERENCIATION GONADIQUE, DE LA VIE FOETALE A LA VIE ADULTE. ....	94
FIGURE 30 : REGULATION DE LA VOIE DE SIGNALISATION WNT/B-CATENINE (HAN AND NA, 2011) .....	104
FIGURE 31 : SYSTEME SEXUEL INDIFFERENCIE JUSQU'A LA 7EME SEMAINE .....	108
FIGURE 32 : DESCENTE TESTICULAIRE .....	110
FIGURE 33 : TRACTUS DU SYSTEME REPRODUCTEUR ET GONADES DE DEUX PORCELETS INTERSEXUES, 38 XX (PAILHOX ET AL., 2001A) .....	112
FIGURE 34 : PHENOTYPES OBSERVES CHEZ DES SOURIS XX ET XY NORMALES, ET XY <i>SOX9</i> <sup>-/-</sup> RESPECTIVEMENT (LAVERY ET AL., 2011). A, B, C : APPAREIL GENITAL EXTERNE, D, E, F : APPAREIL REPRODUCTEUR INTERNE. TESTICULE (T), VESICULE SEMINALE (SV), VESSIE (B) OVAIRE (OV), OVIDUCTE (OD), URETRE (U) OVAIRE DROIT (R), OVAIRE GAUCHE (L).....	116
FIGURE 35 : PHENOTYPES OBSERVES CHEZ UNE SOURIS FEMELLE NORMALE A GAUCHE, .....	116
FIGURE 36 : REARRANGEMENTS CHROMOSOMIQUES IDENTIFIES EN AMONT DE <i>SOX9</i> , .....	118
FIGURE 37 : SCHEMATISATION DES VARIATIONS D'EXPRESSION OBSERVEES CHEZ L'ANIMAL XX INTERSEXUE. LES CERCLES VERTS INDIQUENT UNE SUR-EXPRESSION PAR RAPPORT A LA FEMELLE XX, ET LES CERCLES ROUGES UNE SOUS-EXPRESSION.....	124
FIGURE 38 : TESTICULE NON DESCENDU CHEZ LE JEUNE ENFANT .....	126
FIGURE 39 : CLASSIFICATION DES CAS DE CRYPTORCHIDIES SELON LA POSITION DU TESTICULE NON DESCENDU .....	126
FIGURE 40 : TESTICULE EN POSITION ABDOMINALE CHEZ UN PORCELET CRYPTORCHIDE. ....	128
FIGURE 41 : CRYPTORCHIDIE CHEZ LA SOURIS KO POUR <i>INSL3</i> (NEF AND PARADA, 1999). ....	134
FIGURE 42 : CRYPTORCHIDIE CHEZ LA SOURIS KO POUR <i>RXFP2</i> (GORLOV ET AL., 2002) A DROITE, COMPARE A LA SOURIS NORMALE A GAUCHE. LES TESTICULES CHEZ LE MALE MUTANT <i>RXFP2</i> <sup>-/-</sup> SONT EN POSITION INTRA-ABDOMINALE, PRES DES REINS. LES TESTICULES SONT INDIGUES PAR (T).....	135
FIGURE 43 : HERNIE INGUINALE CHEZ L'HOMME .....	138
FIGURE 44 : HERNIE INGUINALE GAUCHE CHEZ UN PORCELET .....	138
FIGURE 45 : CONCENTRATION SERIQUE DE <i>MMP2</i> CHEZ DES CONTROLES .....	141
FIGURE 46 : CONCENTRATIONS CALCMIQUES CHEZ DES PORCELETS .....	141
FIGURE 47 : LOCALISATION DES QTL IDENTIFIES CHEZ LE PORC D'APRES QUATRE ETUDES DE CARTOGRAPHIE .....	146
FIGURE 48 : IDENTIFICATION DU GENOTYPE PAR ELECTROPHORESE CAPILLAIRE. EXEMPLE AVEC 5 ALLELES D'UN MICROSATELLITE, 4 ECHANTILLONS ET UNE INCOHERENCE CHEZ UN DESCENDANT .....	152
FIGURE 49 : PUCE PORCINE D'ILLUMINA ET SES MICROBILLES RECOUVERTES DE SONDAS .....	154
FIGURE 50 : SONDAS DE GENOTYPE UTILISEES RESPECTIVEMENT EN TECHNOLOGIE INFINIUM I ET II.....	154

FIGURE 51 : ÉTAPES DU GENOTYPAGE A L'AIDE DE LA PUCE PORCINE D'ILLUMINA..... 156

FIGURE 52 : IDENTIFICATION DES GENOTYPES A PARTIR DES EMISSIONS DE FLUORESCENCE DES BILLES DE LA PUCE PORCINE D'ILLUMINA	157
FIGURE 53 : PRINCIPE DU GENOTYPAGE FLUIDIGM. STA = SPECIFIC TARGET AMPLIFICATION PRIMER, LSP = LOCUS SPECIFIC PRIMER ET ASP = ALLELE SPECIFIC PRIMERS	159
FIGURE 54 : PLAQUE BIOMARK, HEATMAP ET ATTRIBUTION DES GENOTYPES	159
FIGURE 55 : SCHEMA DES ETAPES DE SEQUENÇAGE AUTOMATISE SELON LA METHODE DE SANGER	165
FIGURE 56 : ETAPES DU SEQUENÇAGE HISeq 2000	167
FIGURE 57 : EXEMPLE DE "PHRED SCORE" ESTIME POUR LES BASES D'UNE LECTURE	167
FIGURE 58 : CONTROLES QUALITE DES SEQUENCES PAR BASE	169
FIGURE 59 : CONTROLES QUALITE DES SEQUENCES PAR LECTURE	170
FIGURE 60 : ETAPES D'ANALYSES DES SEQUENCES HISeq	172
FIGURE 61 : BAF ET LRR EN ABSENCE DE VARIATION DU NOMBRE DE COPIES (2 COPIES PAR LOCUS)	174
FIGURE 62 : BAF ET LRR LORS DE CNV (1, 2 ET 3 COPIES PAR LOCUS)	175
FIGURE 63 : DISTRIBUTION DU $\chi^2$	178
FIGURE 64 : BILAN DU NOMBRE D'ECHANTILLONS ATTEINTS COLLECTES SELON LES DIFFERENTES RACES ET LES PRINCIPALES ANOMALIES EN 2008	197
FIGURE 65 : PEDIGREE INRA	199
FIGURE 66 : AUTOPSIE D'UN ANIMAL CRYPTORCHIDE	199
FIGURE 67 : AVANT ET APRES AUTOPSIE D'UN ANIMAL ATTEINT DE HERNIE INGUINALE GAUCHE	199
FIGURE 68 : DESCRIPTION DES ANIMAUX INTERSEXUES ISSUS DU PROJET DE PAILHOX ET AL. (PAILHOX ET AL., 1997). LE NOMBRE D'ANIMAUX PAR CATEGORIE EST EGALEMENT REPARTI	201
FIGURE 69 : BILAN DU NOMBRE D'ECHANTILLONS ATTEINTS COLLECTES SELON LES DIFFERENTES RACES ET LES PRINCIPALES ANOMALIES EN 2011	201
FIGURE 70 : CHOIX DES FAMILLES POUR LESQUELLES LE MEME PERE A ETE UTILISE DANS PLUSIEURS FAMILLES ET DANS DIFFERENTS OSP	203
FIGURE 71 : NOMBRE TOTALE D'ECHANTILLONS GENOTYPES, NOMBRE D'ATTEINTS ET NOMBRE DE FAMILLES PAR ANOMALIE ET PAR RACE	205
FIGURE 72 : NOMBRE DE MARQUEURS PAR CHROMOSOME ET SELON LA TAILLE DU CHROMOSOME	209
FIGURE 73 : DISTRIBUTION DES DISTANCES ENTRE PAIRES DE MARQUEURS ADJACENTS	209
FIGURE 74 : NOMBRE DE MARQUEURS PAR INTERVALLES DE MAF POUR LES RACES LARGE WHITE (LW), PIETRAIN (Pi) ET LANDRACE (LR)	211
FIGURE 75 : FREQUENCE D'HETEROZYGOTIE CHEZ LES PARENTS DES DESCENDANTS ATTEINTS PAR RACE	211
FIGURE 76 : NOMBRE DE PARENTS HETEROZYGOTES DANS LES FAMILLES PIETRAIN SUR LE CHROMOSOME 8	215
FIGURE 77 : FREQUENCE CUMULEE DE SNP SELON LE R2 MAXIMUM OBTENU ENTRE CHAQUE COUPLE DE SNP DANS DES INTERVALLES DE 2 Mb	215
FIGURE 78 : DISTANCE ENTRE DEUX MARQUEURS SELON LE R2 MAX	217
FIGURE 79 : RESULTATS TDT EN TOUT GENOME POUR L'INTERSEXUALITE - FAMILLES LARGE WHITE	223

FIGURE 80 : RESULTATS TDT EN TOUT GENOME POUR A. LES FAMILLES INRA, ET B. LES FAMILLES DES OSP ..... 223

FIGURE 81 : RESULTATS TDT EN TOUT GENOME POUR LA HERNIE SCROTALE - FAMILLES LARGE WHITE+ LANDRACE (A), FAMILLES LARGE WHITE SEULES (B) ET FAMILLES LANDRACE SEULES (C). .....	227
FIGURE 82 : RESULTATS TDT EN TOUT GENOME POUR LA HERNIE SCROTALE SELON L'ORIGINE PARENTALE. A. BRIN PATERNEL, B. BRIN MATERNEL.....	229
FIGURE 83 : RESULTATS TDT SELON L'ORIGINE MATERNELLE SUR LE CHROMOSOME 9 .....	231
FIGURE 84 : COMPARAISON ENTRE LES SIGNAUX OBTENUS DANS CETTE ETUDE ET LES RESULTATS DES ETUDES PRECEDENTES.....	233
FIGURE 85 : RESULTATS TDT EN TOUT GENOME POUR LA CRYPTORCHIDIE – FAMILLES PIETRAIN+LARGE WHITE (A), FAMILLES PIETRAIN (B) ET FAMILLES LARGE WHITE (C) .....	237
FIGURE 86 : RESULTATS TDT EN TOUT GENOME POUR LA CRYPTORCHIDIE SELON LE BRIN PATERNEL CHEZ LES PIETRAIN .....	239
FIGURE 87 : RESULTATS TDT SIGNIFICATIFS SUR L'EXTREMITÉ DU CHROMOSOME X POUR LES DIFFÉRENTES POPULATIONS, ET LES GENES ANNOTES DANS LA REGION .....	243
FIGURE 88 : DISTRIBUTION DES VARIANTS RESPONSABLES DE MALADIES (McCARTHY ET AL., 2008).....	253
FIGURE 89 : HAPLOTYPES DANS LA REGION IS POUR LES DEUX ANIMAUX ATTEINTS ET LES DEUX TÉMOINS SÉLECTIONNÉS POUR LE RESEQUENÇAGE.....	279
FIGURE 90 : SEQUENÇAGE HiSEQ DES FRAGMENTS D'ADN .....	281
FIGURE 91 : NOMBRE TOTAL DE SEQUENCES OBTENUES PAR ANIMAL .....	281
FIGURE 92 : DEMARCHE GÉNÉRALE DE L'ALIGNEMENT DES SEQUENCES ISSUES DU RESEQUENÇAGE HiSEQ .....	285
FIGURE 93 : RESULTATS D'ALIGNEMENT DES LECTURES SUR LE DRAFT 9.....	285
FIGURE 94 : TAILLE DES INSERTS ENTRE LES LECTURES APPARIEES .....	287
FIGURE 95 : COUVERTURE DES LECTURES PAR CHROMOSOMES SUR LE DRAFT 9, LE DRAFT 10.2 ET APRES CONTRÔLE QUALITÉ Q30 SUR LE DRAFT 10.2 .....	287
FIGURE 96 : ALIGNEMENT DES LECTURES DE L'INDIVIDU TEM_44(2) SUR LES DRAFTS 9 ET 10.2 .....	289
FIGURE 97 : ALIGNEMENT DE LA REGION CANDIDATE DE V10.2 APRES RECONSTRUCTION SUR LA V10.2 .....	291
FIGURE 98 : ALIGNEMENT DES LECTURES DE L'INDIVIDU TEM_44(2) SUR LE DRAFT 10.2 APRES RECONSTRUCTION .....	291
FIGURE 99 : ALIGNEMENT DES LECTURES DE L'INDIVIDU IS_35 SUR LE DRAFT 10.2 APRES RECONSTRUCTION .....	291
FIGURE 100 : NOMBRE DE POLYMORPHISMES (SNP ET INDELS) ET NOMBRE TOTAL DE LECTURES CHEZ LES QUATRE ANIMAUX SEQUENCES EN TOUT GENOME. L'INTERSECTION CORRESPOND AUX POLYMORPHISMES IDENTIFIES CHEZ LES DEUX ANIMAUX D'UN MEME GROUPE.....	293
FIGURE 101: COUVERTURE ET POLYMORPHISMES DANS LA SEQUENCE DE SOX9. SOX9 EST INDIQUE PAR UN RECTANGLE VERT .....	296
FIGURE 102 : GENOTYPES POUR CHACUN DES 6801 POLYMORPHISMES OBSERVES DANS LA REGION IS, ET POUR CHACUN DES QUATRE ANIMAUX. BLANC 0/0 : PAS DE GENOTYPE ATTRIBUE, BLEU : HOMOZYGOTE, VERT : HETEROZYGOTE .....	299
FIGURE 103 : NOMBRE DE LECTURES OBSERVEES ET VALEUR DE QUALITE POUR LES 98 SNP HOMOZYGOTES POUR CHAQUE ALLELE RESPECTIVEMENT POUR LES DEUX INDIVIDUS IS (IS_35 : 1/1 ET IS_49 : 2/2 OU IS_35 : 2/2 ET IS_49 : 1/1) .....	300
FIGURE 104 : NOMBRE DE LECTURES OBSERVEES POUR CHAQUE ALLELE DES 66 SNP HETEROZYGOTES POUR L'INDIVIDU IS_49 ....	300
FIGURE 105 : PROFONDEURS DES LECTURES PAR BASE ENTRE LES TÉMOINS (ROUGE) ET LES IS (VERT) DANS LA REGION DU CHROMOSOME 12. L'AXE DES ORDONNÉES CORRESPOND A LA VARIATION DE LA PROFONDEUR DE SEQUENÇAGE PAR RAPPORT A LA MOYENNE DE CHACUN DES LOTS TÉMOINS / IS. LE SEUIL A ÉTÉ FIXÉ A DIX AFIN DE MIEUX VOIR LES PETITES FLUCTUATIONS	302

FIGURE 106 : NOMBRE DE VARIATIONS (AXE DES ORDONNEES) EN FONCTION DE LA TAILLE DE LA VARIATION ..... 302

FIGURE 107 : GENES IMPLIQUES DANS LE DIFFERENCIATION GONADIQUE ET ROLE DE KLF4 .....	304
FIGURE 108 : RESULTATS TDT A PARTIR DES DONNEES DE GENOTYPAGE FLUIDIGM. LA COURBE BLEUE CORRESPOND AU $-\log$ DE LA P-VALUE DU CHI2, ET LES POINTS ROUGES INDIQUENT LE NOMBRE DE PARENTS HETEROZYGOTES PAR MARQUEUR (AXE DE DROITE), ET DONC LEUR INFORMATIVITE. ....	310
FIGURE 109 : NOMBRE D'ECHANTILLONS SELON LE STATUT A LA MALADIE .....	310
FIGURE 110 : RESULTATS TDT AVEC LES MARQUEURS COMPLEMENTAIRES FLUIDIGM, ET LES MARQUEURS DE LA PUCE PORCINE, AINSI QUE LE NOMBRE DE PARENTS HETEROZYGOTES PAR MARQUEURS FLUIDIGM .....	312
FIGURE 111 : DL DANS LA REGION CANDIDATE A L'INTERSEXUALITE AVEC LES MARQUEURS DE LA PUCE PORCINE. LES CARRS ROUGES INDIQUENT UN DL PROCHE DE 1 .....	312
FIGURE 112 : DL DANS LA REGION CANDIDATE A L'INTERSEXUALITE AVEC LES MARQUEURS COMPLEMENTAIRES FLUIDIGM.....	312
FIGURE 113 : COMPARAISON DES HAPLOTYPES OBTENUS AVEC LES MARQUEURS DE LA PUCE PORCINE (HAPLOTYP_1), ET APRES AJOUT DE NOUVEAUX MARQUEURS DANS LA REGION 8,6 A 9 Mb (HAPLOTYP_2) .....	314
FIGURE 114 : FREQUENCE HAPLOTYPIQUE CHEZ DES ANIMAUX INTERSEXUES, LEURS SCEURS SAINES ET DES MALES D'UN AUTRE PROJET .....	316
FIGURE 115 : FREQUENCE GENOTYPIQUE DES HAPLOTYPES .....	316
FIGURE 116 : EXEMPLES DE FAMILLES HOMOZYGOTES POUR L'HAPLOTYP 1, HETEROZYGOTE POUR L'HAPLOTYP 5 OU SANS HAPLOTYP 1 OU 5. L'ORIGINE DU BRIN DES DESCENDANTS EST INDIQUEE EN ROSE POUR L'ORIGINE MATERNELLE ET BLEU POUR L'ORIGINE PATERNELLE. LES NUMEROS DES HAPLOTYPES SONT INDIQUES A DROITE.....	318
FIGURE 117 : RESULTATS TDT EN RETIRANT LES DESCENDANTS ATTEINTS HOMOZYGOTES POUR L'HAPLOTYP 1 ET HETEROZYGOTES POUR L'HAPLOTYP 5.....	318
FIGURE 118 : LOCALISATION DE LA REGION PSEUDO-AUTOSOMALE PORCINE. LE SIGNAL TDT EST INDIQUE PAR LES POINTS ORANGE, ET LES CROIX BLEUES CORRESPONDENT A LA FREQUENCE D'HETEROZYGOTIE. LA PAR EST LOCALISEE EN BLEU SUR LES CHROMOSOMES SEXUELS EN BAS DE LA FIGURE. LES RESULTATS TDT PRESENTE ICI PROVIENNENT DES ANALYSES CHEZ LES LARGE WHITE ATTEINTS DE HERNIE SCROTALE, MAIS LA LOCALISATION DU SIGNAL SIGNIFICATIF EST IDENTIQUE POUR TOUS. ....	336
FIGURE 119 : PROBABILITE DE TRANSMISSION SUR LA PAR. PROBABILITE DE TRANSMISSION DES ALLELES PARY DES PERES VERS LES DESCENDANTS MALES (COURBE BLEUE), DES ALLELES PARX DES PERES VERS LES DESCENDANTES FEMELLES (COURBE ROSE), ET DES DEUX ALLELES DES MARQUEURS DE LA PAR VERS TOUS LES DESCENDANTS (COURBE GRISE).....	338
FIGURE 120 : TAUX DE RECOMBINAISON SUR LA REGION PSEUDO-AUTOSOMALE PORCINE. LA COURBE BLEU CLAIR REPRESENT LE TAUX DE RECOMBINAISON CHEZ LES MALES (BLEU FONCE POUR LES VALEURS LISSEES), ET LES COURBES ROSES (DONNEES BRUTES ET DONNEES LISSEES) LE TAUX DE RECOMBINAISON CHEZ LES FEMELLES.....	340
FIGURE 121 : VARIABILITE HAPLOTYPIQUE ESTIMEE RESPECTIVEMENT A PARTIR DE POOLS DE REGIONS CHROMOSOMIQUES PARY ET PARX. LA RECONSTRUCTION HAPLOTYPIQUE S'EST FAIT DE LA DROITE (MARQUEUR « N ») VERS LA GAUCHE (MARQUEUR « N-10 »), SOIT DU MARQUEUR LE PLUS PROCHE DE LA LIMITE DE LA PAR VERS LE TELOMERE (POUR LES 11 DERNIERS MARQUEURS DE LA PAR). LES HAPLOTYPES, DONT LA SEQUENCE EST INDIQUEE, SONT COMMUNS AUX PARX ET AU PARY.....	342
FIGURE 122 : DIFFERENCE DE FREQUENCE ENTRE LES ALLELES PARX ET PARY. L'AXE DES ABCISSES CORRESPOND A L'ORDRE DES MARQUEURS SUR LA PAR, ET L'AXE DES ORDONNEES A LA DIFFERENCE DE FREQUENCES DES ALLELES PARX-PARY.....	344
FIGURE 123 : FREQUENCE DES ALLELES SUR PARX ET PARY POUR LE MARQUEUR LE PLUS PROCHE DE LA LIMITE DE LA PAR.....	346



FIGURE 124 : SIGNAL TDT OBSERVE SUR LA PAR POUR LES LARGE WHITE ATTEINTS DE HERNIE INGUINO-SCROTALE, AVANT ET APRES CORRECTION DU SIGNAL. LA CORRECTION A ETE APPOREE EN DEFINISSANT UNE NOUVELLE HYPOTHESE NULLE, SPECIFIQUE DE LA PAR, QUI TIENT COMPTE DE LA DISTORSION DE SEGREGATION PRESENTE SUR LA PAR. .... 346

FIGURE 125 : DIFFERENCE DE FREQUENCE GENOTYPIQUE ENTRE LES LARGE WHITE ATTEINTS DE HERNIE INGUINO-SCROTALE ET DES TEMOINS LARGE WHITE SUR LA PAR ..... 348



# Liste des tableaux

TABLEAU 1 : NOUVEAUX OBJECTIFS DE SELECTION ANIMALE (FAO, 2008).....	32
TABLEAU 2 : ANOMALIES HEREDITAIRES REFERENCEES CHEZ LE PORC (OLLIVIER, 1979) .....	38
TABLEAU 3 : HERITABILITES DE LA HERNIE SCROTALE, LA CRYPTORCHIDIE ET L'INTERSEXUALITE SELON DIFFERENTES ETUDES REALISEES ENTRE 1951 ET 2011 .....	40
TABLEAU 4 : CARACTERES PORCINS POUR LESQUELS LE GENE IMPLIQUE A ETE IDENTIFIE - SOURCE : OMIA.....	42
TABLEAU 5 : TAILLE TOTALE, NOMBRE DE CONTIGS ET NOMBRE DE SCAFFOLDS DES DRAFTS V9 ET V10.2 .....	50
TABLEAU 6 : COMPARAISON DE LA TAILLE DU GENOME ET DU CONTENU EN ELEMENTS REPETES CHEZ PLUSIEURS MAMMIFERES. LES DONNEES DES ELEMENTS CHEZ LE PORC SONT ISSUES DE L'ETUDE DE WERNERSSON (ROTHSCHILD AND RUVINSKY, 2011). ....	60
TABLEAU 7 : GENES UTILISES EN SELECTION PORCINE (DEKKERS, 2004).....	78
TABLEAU 8 : ANOMALIES CONGENITALES LES PLUS FREQUENCES CHEZ L'HOMME .....	86
TABLEAU 9 : NOUVELLE CLASSIFICATION DES DESORDRES DU DEVELOPPEMENT SEXUEL .....	88
TABLEAU 10 : GENES ASSOCIES AU DEVELOPPEMENT DE LA GONADE DIFFERENCIEE CHEZ LA SOURIS ET L'HOMME (BLEU : GENES DE LA DIFFERENCIATION MALE, ROSE : GENES DE LA DIFFERENCIATION FEMELLE). LOF : LOSS OF FUNCION, GOF : GAIN OF FUNCTION.....	114
TABLEAU 11 : COMPOSITION DU MELANGE REACTIONNEL ET CONDITIONS D'AMPLIFICATION PCR UTILISES POUR DIAGNOSTIQUER LA PRESENCE OU NON DU LOCUS SRY .....	150
TABLEAU 12 : MICROSATELLITES ET AMORCES UTILISEES POUR LES TESTS DE PARENTE. CHR : CHROMOSOME, ALL : NOMBRE D'ALLÈLES IDENTIFIÉS, FLUO : FLUOROCHROME, TEMP X : TEMPÉRATURE D'HYBRIDATION .....	150
TABLEAU 13 : COMPOSITION DU MIX ET CONDITIONS PCR DU GENOTYPAGE POUR LES TESTS DE PARENTE .....	152
TABLEAU 14 : CONDITIONS D'AMPLIFICATION POUR LE GENOTYPAGE FLUIDIGM .....	161
TABLEAU 15 : AMORCES UTILISEES POUR L'AMPLIFICATION SUR LE PANEL D'HYBRIDES D'IRRADIATION.....	163
TABLEAU 16 : AMORCES POUR LE SEQUENÇAGE DE <i>SOX9</i> ET <i>TESCO</i> , ET CONDITIONS PCR CORRESPONDANTES.....	165
TABLEAU 17 : EXEMPLE DES VALEURS OBSERVEES UTILISEES POUR LE TEST DU $\chi^2$ DANS UNE ANALYSE TDT .....	178
TABLEAU 18 : BILAN DES MARQUEURS SIGNIFICATIFS POUR LA HERNIE INGUINO-SCROTALE. LE FOND VERT INDIQUE QUE LES MARQUEURS SONT SIGNIFICATIFS POUR LES ANALYSES COMBINANT LES FAMILLES DES DEUX RACES, LE FOND ROSE NE CONCERNE QUE LES FAMILLES LW, ET LE FOND ORANGE QUE LES FAMILLES LR. HS : HERNIE INGUINO-SCROTALE, LW : LARGE WHITE, LR : LANDRACE, CHR : CHROMOSOME, SSC : CHROMOSOME PORCIN, MRK : MARQUEUR, POS : POSITION EN PB SUR LE GENOME DE REFERENCE PORCIN. ....	231
TABLEAU 19 : REGLE DE DECISION POUR LE CHOIX DES POLYMORPHISMES. LES VALEURS CORRESPONDENT AUX FREQUENCES ALLELIQUES DE L'ALLELE 1. LES VALEURS 1 ET 0 SIGNIFIENT QUE L'ANIMAL EST HOMOZYGOTE, LA VALEUR 0.5 EST ATTRIBUEE A UN ANIMAL HETEROZYGOTE .....	298



# Introduction

---

Les anomalies congénitales sont définies comme des défauts de structure ou de fonction présents à la naissance. Chez le porc, les anomalies les plus fréquentes sont de type urogénital (hernie inguino-scrotale congénitale, cryptorchidie et intersexualité) et représentent un problème majeur pour l'industrie porcine. Elles sont responsables de pertes économiques importantes et portent atteinte au bien-être des animaux. Au delà de la mise en évidence de facteurs environnementaux et sanitaires, les études ont montré qu'une part de leur déterminisme était génétique. Cependant, les analyses de liaison en tout génome et les approches gène candidat réalisées jusqu'à présent n'ont pas permis d'identifier de variants génétiques associés aux principales anomalies rencontrées chez le porc.

Depuis 2006, la compréhension des mécanismes moléculaires impliqués dans les anomalies congénitales est une priorité pour l'industrie porcine. Dans ce cadre, un projet de recherche, appelé SwAn pour *Swine Anomalies*, a été établi entre les organismes privés français de sélection porcine OSP (association BIOPORC), l'IFIP (Institut de la Filière Porcine) et l'INRA (Institut National de la recherche agronomique). Les outils aujourd'hui disponibles en génomique porcine, tels que la puce de génotypage haute densité porcine d'Illumina®, permettent une recherche rapide à l'échelle du génome de marqueurs associés à un phénotype. Ces études d'association, appelées GWAS (*Genome-wide Association Studies*), ont fait leur preuve chez d'autres espèces (Charlier et al., 2008; Risch and Merikangas, 1996; Shearin and Ostrander, 2010). Dans le cadre de SwAn, l'identification de gènes ou mutations causales permettra de comprendre les mécanismes moléculaires associés aux anomalies, et de mettre en place des schémas de sélection pour réduire la prévalence des défauts congénitaux dans les élevages.

**L'objectif de la thèse est d'identifier les gènes impliqués dans le déterminisme des principales anomalies congénitales porcines par une approche de clonage positionnel.**



**La première partie** de la thèse est une introduction bibliographique sur la filière porcine et les défauts génétiques affectant les élevages porcins, ainsi que les outils génétiques disponibles chez le porc permettant la recherche et l'identification de mécanismes moléculaires responsables de défauts génétiques, et les connaissances actuelles sur la physiopathologie et les causes génétiques des principales anomalies congénitales porcines.

**La deuxième partie** du manuscrit est une présentation des outils et les méthodes utilisés dans ces travaux. Les résultats seront ensuite répartis en trois chapitres dans la **troisième partie** du manuscrit. Le **chapitre I** présente le dispositif familial mis en place pour les études TDT, et les résultats des GWAS pour la hernie inguino-scrotale, la cryptorchidie et l'intersexualité. Ces travaux ont permis de mettre en évidence une région candidate commune aux trois anomalies, localisée sur le chromosome X, et une région sur le chromosome 12 associée à l'intersexualité.

Le seul gène annoté dans la région du chromosome 12 est un excellent candidat à l'intersexualité car il est un acteur majeur de la différenciation sexuelle chez le mâle. Des défauts de sa régulation chez la souris (Bishop et al., 2000a; Lavery et al., 2011) et chez l'homme (Benko et al., 2009; Cox et al., 2011; Vetro et al., 2011) sont impliqués dans des désordres du développement sexuel. Les travaux de cartographie fine de la région sont traités dans le **chapitre II**. Ces travaux ont permis de valider et réduire la région candidate, et ont donné lieu à un article accepté dans *PlosOne*.

La région mise en évidence sur le chromosome X pour les trois anomalies a fait l'objet des travaux présentés dans le **chapitre III**. Bien que des corrélations aient été établies entre la hernie inguino-scrotale et la cryptorchidie (anomalies de la migration testiculaire), l'implication d'un même locus dans les trois maladies a motivé des travaux plus approfondis de la région. Peu de données étant disponibles concernant le chromosome X porcine, des cartographies physique et génétique haute densité ont été réalisés pour la région spécifique du chromosome X et la région pseudo-autosomale porcine. Les travaux de cartographie ont fait l'objet d'un article qui sera soumis à *BMC genomics*.

Chaque chapitre fera l'objet d'une discussion qui lui est propre, et l'ensemble des travaux sera discuté plus globalement dans une **quatrième partie**.

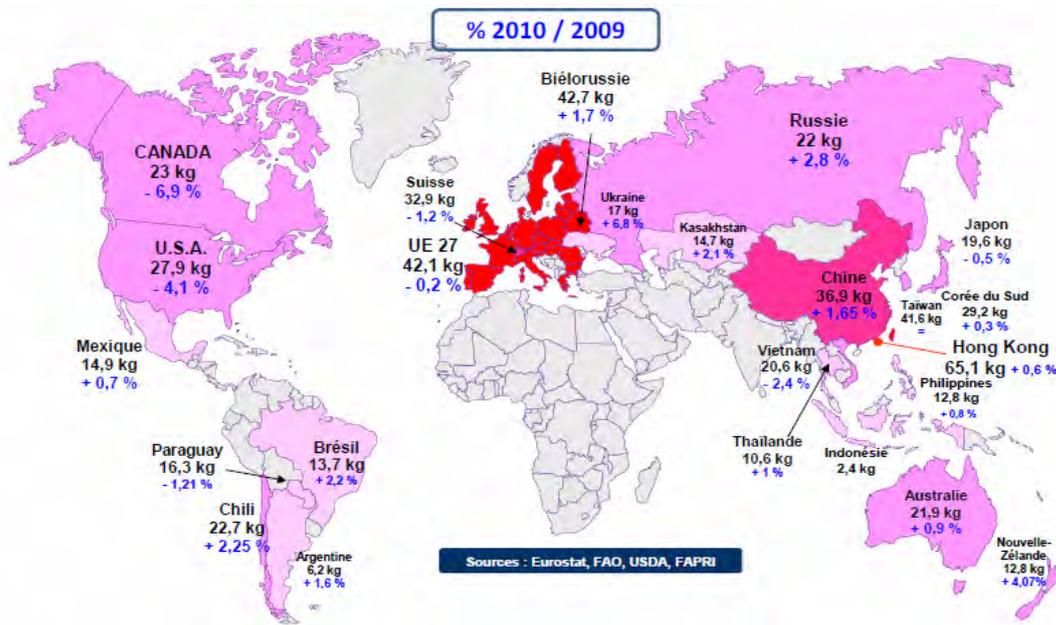


Figure 1 : Consommation moyenne de viande porcine en kilogrammes par habitant dans le monde en 2010

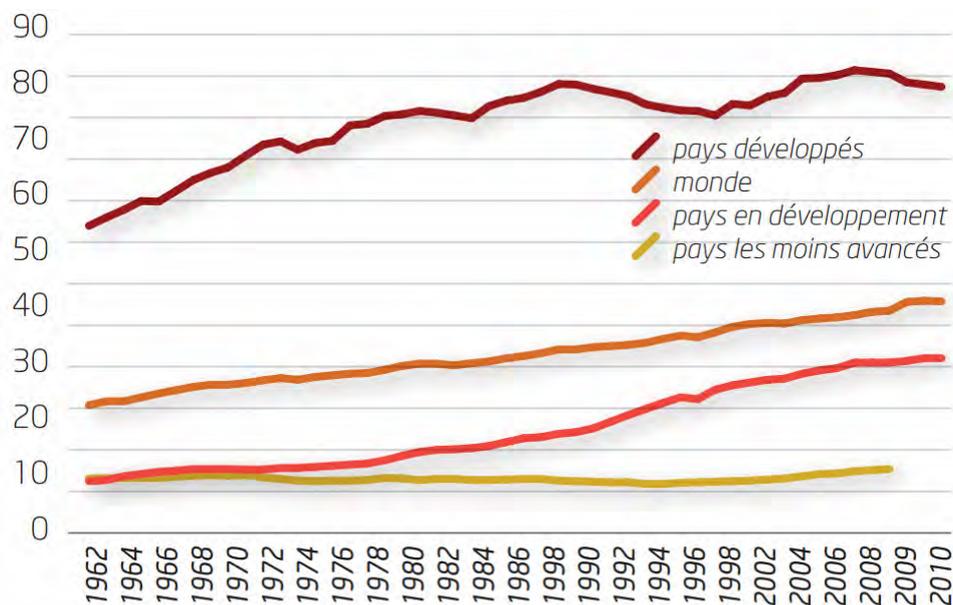


Figure 2 : Evolution de la consommation de viande dans le monde et par type de pays en kg/habitant/an - source: FranceAgriMer d'après FAO, OCDE estimations

# Données bibliographiques

---

## I. Impact des anomalies congénitales en production porcine

### I.A. Quelques éléments de contexte agro-économique

#### I.A.1. La consommation de viande de porc dans le monde

Les animaux domestiques représentent une des principales sources d'apports en protéines de l'alimentation humaine (Pereira and Vicente, 2013), et la consommation de viande a considérablement augmenté au cours des cinquante dernières années. Chaque année, 42,5 kilogrammes de viande sont consommés en moyenne dans le monde par habitant (Worldwatch Institute, 2011). A titre comparatif, la consommation de poisson est de 17 kg/habitant/an. Le porc représente la source majoritaire de viande dans le monde (Food and agriculture organization of the United Nations, 2012; Smith et al., 2012), faisant de l'élevage porcin une filière majeure en agroalimentaire. La viande porcine représente 40% de la consommation totale en viande, contre 29% pour les volailles et 24% pour le bœuf.

A l'échelle mondiale, 17 kg de viande porcine sont consommés en moyenne par habitant et par an. De fortes différences existent cependant selon les pays et s'expliquent par des différences culturelles et économiques (Figure 1). La consommation la plus importante est observée dans les pays d'Europe et en Chine où elle est supérieure à 35 kg/habitant/an, et aux Etats-Unis et en Corée du Sud, où elle avoisine les 30 kg/habitant/an. A l'inverse, la consommation de viande porcine est quasi nulle dans les pays musulmans, en Afrique du Nord et au Proche-Orient, ainsi qu'en Inde.

Il y a cinquante ans, la consommation dans les pays développés était cinq fois supérieure aux autres pays. Depuis les années 70, la consommation a doublé à l'échelle mondiale (Figure 2), et triplé dans les pays en voie de développement, où les revenus et le niveau d'urbanisation ont augmenté. Elle s'est stabilisée dans les pays développés en raison de la prise en compte, par les consommateurs, de nouvelles notions ayant trait à la nutrition et la santé, au bien-être animal, et à l'impact environnemental de l'élevage.

D'après l'organisation des nations unies pour l'alimentation et l'agriculture, la consommation totale devrait à nouveau doubler d'ici 2050 (Speedy, 2003).

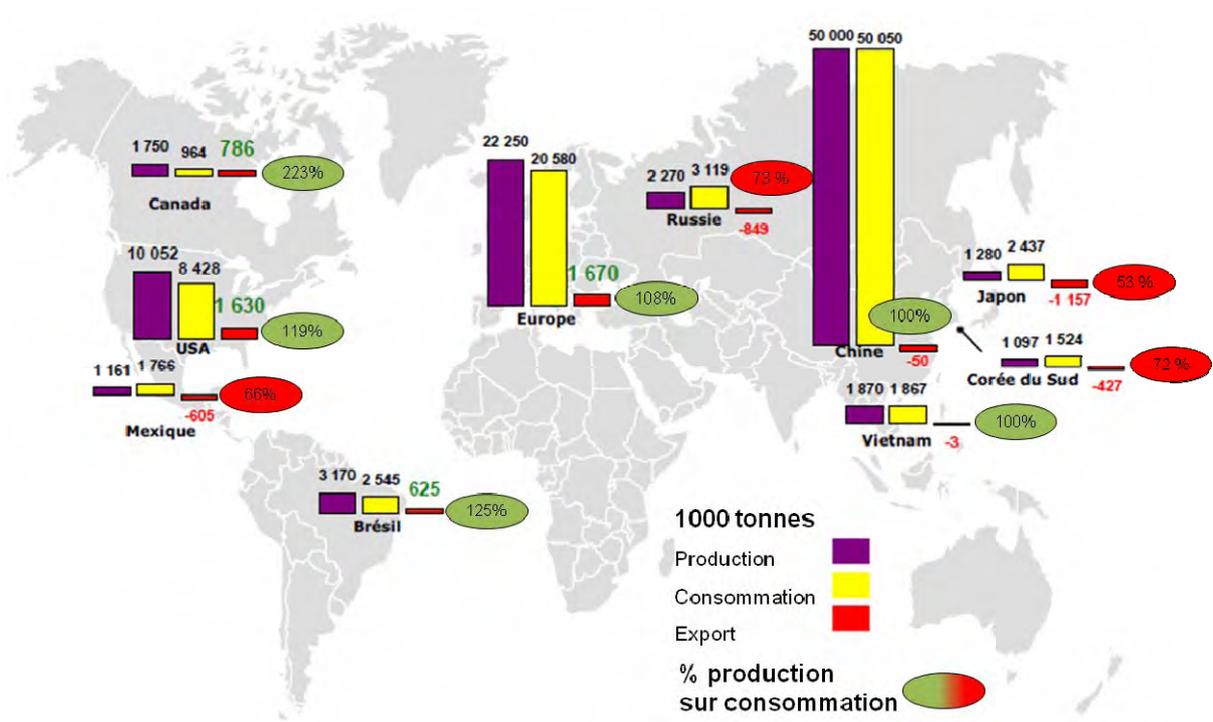


Figure 3 : Production, consommation, exportation et ratio entre la production et la consommation pour les principaux pays consommateurs de viande de porc au monde

Principaux flux 2009 en milliers de tonnes  
(échanges de plus de 100 000 tonnes)

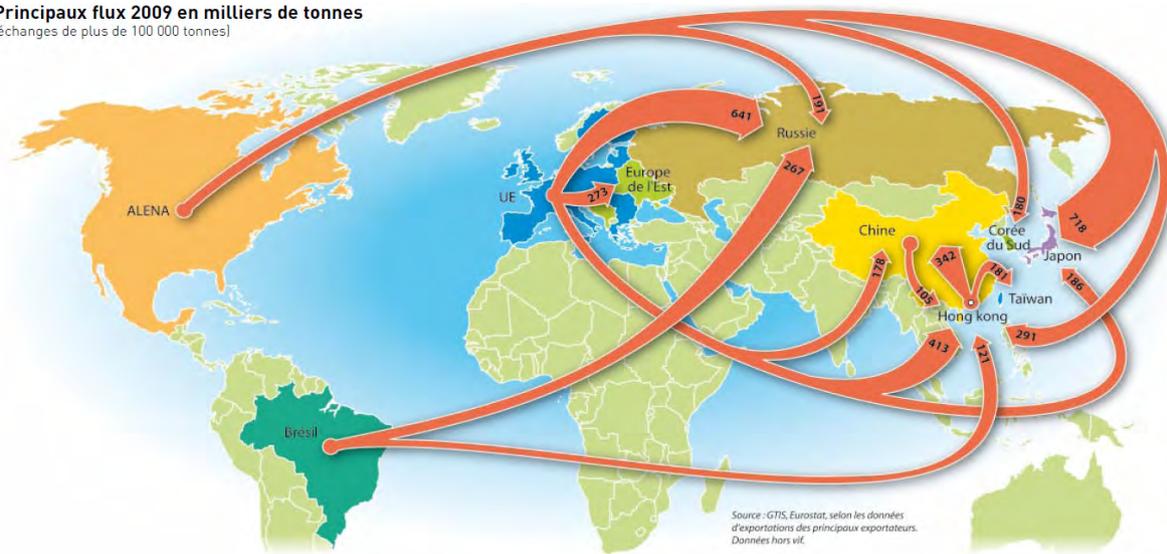


Figure 4 : Principaux courants d'échanges internationaux de viande de porc (Commerce mondial 2009, Observatoire international de l'IFIP).

## I.A.2. La production porcine à l'échelle mondiale

Globalement, les pays dont la consommation est la plus importante sont également ceux qui produisent le plus de viande porcine.

La Chine est le premier producteur du monde, avec 50 millions de tonnes produites en 2010 (Figure 3). Sa production a été multipliée par trente en cinquante ans (2012). Les deuxième et troisième plus grands producteurs de viande porcine sont l'Europe avec 22,25 millions et les Etats-Unis avec 10 millions de tonnes. Au sein de l'Europe, la France est le troisième producteur de porc, avec 2,3 millions de tonnes équivalent carcasses produites chaque année, derrière l'Allemagne et l'Espagne, qui produisent 4,9 et 3,5 tonnes respectivement par an. La production de la France est stable depuis dix ans, alors qu'elle a augmenté de 26% en Allemagne et 17% en Espagne. Le principal bassin de production en France est situé en Bretagne, et correspond à 58% de la production porcine française.

Dans certains pays, des différences importantes entre la production et la consommation sont observées (Figure 3) : Le Canada produit deux fois plus de porc qu'il n'en consomme (Canada Pork International, 2013) et son exportation est principalement tournée vers l'Asie (Van Ferneij et al., 2012). A l'inverse, le Japon a une consommation deux fois supérieure à sa production annuelle. Il est important de souligner que les échanges de viande de porc ne se limitent pas aux pays qui ont un excédent ou un déficit de production. Le commerce mondial de la viande de porc évolue sans cesse et gagne en importance, en volume et en valeur (Figure 4). Les exportations vers diverses destinations s'ajustent selon le niveau de production, la situation économique et les changements de consommation. La moitié des échanges mondiaux est constituée de « viandes fraîches, réfrigérées et congelées ». L'ensemble des exportations mondiales, tous produits et toutes destinations confondus totalise près de 7 millions de tonnes en 2010, soit 6% à 7% de la production porcine mondiale et la valeur de ces échanges dépasse désormais 11 milliards d'euros (Van Ferneij et al., 2012).

Cette augmentation de la demande a induit d'importants changements dans les méthodes de production. Les élevages porcins, de type « fermiers » (Rainelli, 2001), associant polycultures et élevage jusque dans les années 70, ont été progressivement industrialisés via la spécialisation et l'augmentation de la taille des élevages dans les années 80 (Teffene et al., 1998). Parallèlement, la mise en place de schémas de sélection spécialisés a permis de répondre progressivement aux besoins de la filière.



## I.B. Les élevages porcins et la sélection génétique

Jusqu'aux années 70, le porc est traditionnellement produit dans des fermes familiales (Rainelli, 2001). L'élevage est structuré selon un modèle polyculture-élevage à côté de productions telles que les céréales, les pommes de terre et le lait. Les animaux sont dépendants des productions végétales de la ferme, car elles sont leurs principales sources d'alimentation, et les rejets des animaux servent d'engrais pour les cultures. Les élevages sont alors de petite taille et indépendants les uns des autres. Alors que la période après-guerre est marquée par une augmentation de la demande (Observatoire Bruxellois de la Consommation Durable, 2007), l'essor de la France dans les années 50-60 entraîne une hausse des revenus, et un exode rural qui affecte la production porcine. Des méthodes d'élevage industriel se mettent alors en place pour relancer la production, et la « Loi sur l'élevage » est instaurée (1966). Cette restructuration de la production porcine porte sur la mise en place de systèmes de production spécifiques (élevages hors sols) et la mise en place d'un système d'évaluation génétique des reproducteurs (Teffene et al., 1998).

La sélection génétique vise à obtenir les « meilleurs » géniteurs, selon leurs valeurs génétiques. A partir des années 90, l'utilisation du modèle BLUP-modèle animal, pour *Best Linear Unbiased Prediction* (Bidanel et al., 1990) se généralise. Le BLUP consiste à prendre en compte la totalité de l'information disponible sur les candidats à la sélection et l'ensemble de leurs apparentés connus, et à estimer simultanément les effets du milieu et les valeurs génétiques.

Les performances utilisées pour estimer les valeurs génétiques sont définies en fonction des objectifs de sélection, qui concernent notamment des caractères de production et de qualité de la carcasse, et des caractères de reproduction. Il s'agit d'obtenir un porc charcutier qui répond à la fois aux attentes (1) du consommateur par la qualité de la viande, (2) de l'abatteur par le rendement et l'homogénéité des carcasses, ainsi que les qualités de découpe et de conservation, et (3) du producteur, qui souhaite une croissance rapide et efficace des animaux, notamment en termes de gain de poids moyen et d'indice de consommation, ainsi que du nombre de porcelets par truie. Ses attentes sont aussi celles de l'abatteur, le rendement des carcasses fixant le prix du porc.

Afin de maîtriser la sélection animale, des systèmes bien définis ont été mis en place.

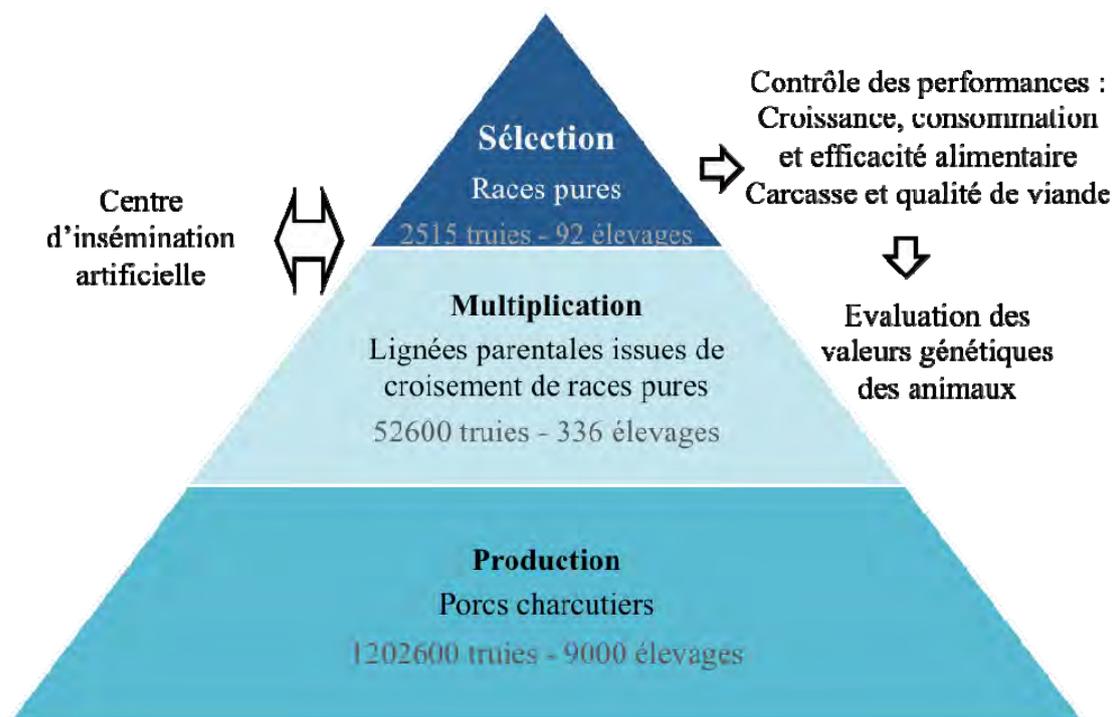


Figure 5 : Structure en trois niveaux du schéma de production porcine française (Sources: Agence de la sélection porcine 2009, IFIP 2009, SPP Agreste enquête structures 2007)

En France, le schéma de production répond à une structure pyramidale à trois niveaux (Figure 5):

- 1. La sélection de races pures :** La production et la sélection d'animaux de race pure ou de lignées est réalisée par les organismes de sélection porcine (OSP). Les races les plus importantes sont le Large White, le Landrace, et le Piétrain. Ces races sont évaluées génétiquement afin de contribuer au schéma d'amélioration collectif encadré par l'INRA et l'Institut technique du Porc. La sélection de races pures permet de conserver la variabilité génétique dans l'espèce, chaque race possédant des avantages bien particuliers (Reiner et al., 2002). Disposer de plusieurs races pures permet d'adapter les croisements à la demande. Les truies Large White ou Landrace présentent des qualités hyperprolifiques alors que le mâle Large White possède de bonnes qualités de croissance. Le Piétrain est sélectionné pour sa teneur en muscle et le rendement de sa carcasse, et le Duroc pour sa rusticité et la qualité de sa viande. Enfin, des lignées composites sino-européennes ont été constituées à partir des années 90 et sélectionnées afin de tirer parti des qualités maternelles et de la prolificité des races chinoises utilisées dans le croisement.
- 2. La multiplication :** L'étage de multiplication permet de produire des verrats et des truies croisés destinées aux élevages de production, issus de croisements d'animaux de race pure sélectionnés en amont. L'utilisation du croisement permet de combiner les avantages de plusieurs races et de bénéficier d'un premier effet d'hétérosis (supériorité du produit croisé par rapport à la moyenne des valeurs des populations parentales sur les qualités maternelles). L'effet d'hétérosis varie selon les caractères et la distance génétique entre les races. A titre d'exemple le croisement d'animaux Meishan et d'animaux des races européennes pures a permis d'augmenter la taille des portées de 20% par rapport à la moyenne des parents (Haley and Lee, 1993).
- 3. La production de porcs charcutiers :** Le porc charcutier (produit terminal) est issu du croisement de truies et de verrats parentaux produits à l'étage de multiplication et bénéficie d'un second effet d'hétérosis sur les caractères bouchers. Aujourd'hui, la moitié des porcs charcutiers est issue de truies croisées Large White femelles et Landrace mâles, pour leurs qualités maternelles, et de mâles Piétrain, aux qualités de carcasse importantes. 24% des cochettes utilisées sont issus de femelles chinoises, et 20% possèdent des gènes Duroc. Dans les élevages français, 25 millions de porcs sont produits chaque année.

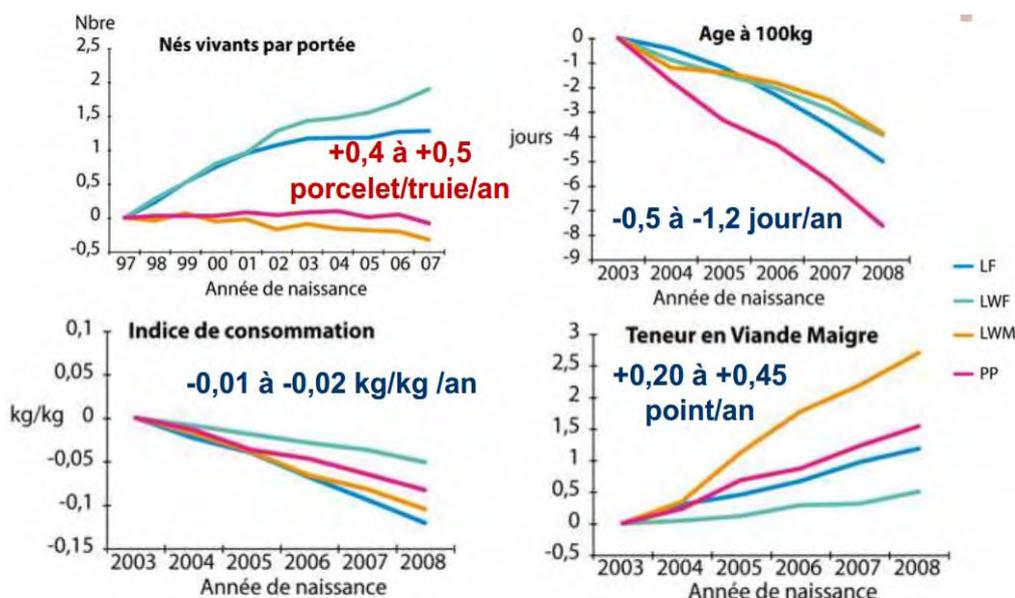


Figure 6 : Evolution des performances pour les quatre races principalement utilisées en France. LF : Landrace Français, LWF/M : Large White Français Femelle/Mâle, PP : Piétrain.

Objectifs	Critères	Remarques
<b>Caractères de production</b>	Taux de croissance	A différents âges
	Poids de la carcasse	
	Qualité de la carcasse	Uniformité, maigreur de la carcasse
	Qualité de la viande	Capacité de rétention hydrique, couleur, goût
<b>Caractères fonctionnels</b>	<b>Santé et bien être</b>	
	Résistance générale	Robustesse
	Vitalité des porcelets, survie des porcs	Capacité maternelle, nombres de trayons
	Stress	Gène (halothane) de l'élimination du stress dans les lignées maternelles et, si possible, dans les lignées mâles
	Effets congénitaux	Exemples : atrésie anale, cryptorchidie, dysplasie, hermaphrodisme et hernie
	Problème des pattes	Faiblesse des pattes et claudication
	<b>Efficacité</b>	
	Taille de portée	Nombre de porcs abattus par truie et par an
	Efficacité de conversion alimentaire	
	Longévité	
Vie de troupeau fonctionnelle	Rendement total avec problèmes sanitaires minimaux	

Tableau 1 : Nouveaux objectifs de sélection animale (FAO, 2008)

La sélection génétique des animaux a permis d'améliorer significativement les performances globales (Figure 6) pour des caractères telles que l'efficacité alimentaire et la taille des portées. Les animaux atteignent aujourd'hui le poids de 100kg une semaine plus tôt en moyenne qu'en 2003, et l'indice de consommation a diminué de 10g environ par kg et par an. Leur teneur en viande maigre a augmenté, notamment pour le Large White Français. Enfin, les femelles Large White et Landrace donnent naissance à un à deux porcelets de plus depuis 1997.

Ces améliorations génétiques ont permis aux éleveurs français d'abaisser leur coût de production depuis vingt ans. Cependant, le premier poste de dépense pour les élevages est l'aliment (67% du coût de revient en 2008, source: IFIP-GTE-TB), et la viabilité des exploitations est donc très sensible au prix des matières premières. Des progrès génétiques sont encore attendus concernant les caractères héréditaires, responsables des pertes économiques les plus importantes.

Aujourd'hui, les notions de bien-être animal et d'amélioration durable sont primordiales pour la sélection animale. En 2008, la FAO a publié un rapport sur l'état des ressources zoogénétiques et les nouveaux objectifs de sélection (Food and Agriculture Organization of the United Nations, 2008). Les objectifs sont orientés à la fois vers la sélection d'animaux plus robustes et plus efficaces face aux variations environnementales, mais aussi vers des caractères dit « secondaires » (Tableau 1) tels que la survie des porcelets, la réponse au stress et les défauts de naissance.

La prévalence parfois importante de certaines anomalies peut s'expliquer par des effets positifs sur certains caractères qui compensent les effets défavorables sur la santé de l'animal (Ollivier, 1979). Le syndrome de stress aigu, ou hyperthermie maligne (Fujii et al., 1991), pouvant entraîner la mort de l'animal, en est un exemple. Cette maladie, caractérisée par d'intenses contractures musculaires et une hyperthermie, est déclenchée lors d'un état de stress ou par l'inhalation d'halothane (Mormède and Mercat, 2002). A l'abattage, les viandes sont dites PSE (*Pale, Soft and Exsudative*). La maladie est à transmission autosomique récessive, seuls les individus homozygotes pour la mutation responsable présentent ce syndrome. Cependant les animaux hétérozygotes et homozygotes mutés ont une carcasse dont la qualité est supérieure aux homozygotes non porteurs, c'est pourquoi la fréquence de l'allèle délétère a progressivement augmenté dans certaines races et lignées utilisées en lignées mâles.



Il est donc primordial de comprendre les mécanismes génétiques impliqués dans les maladies héréditaires « fréquentes » et les phénotypes associés aux génotypes, afin de pouvoir contre sélectionner les allèles responsables ou gérer leur fréquence au sein des schémas de sélection.



## I.C. Les maladies génétiques dans l'espèce porcine

### I.C.1. Impact des maladies congénitales sur la production porcine

Les maladies congénitales, dites de naissance, sont responsables de pertes économiques majeures dans l'industrie porcine. Le terme « congénital » n'est pas synonyme de « héréditaire » car ces maladies peuvent être causées par des facteurs environnementaux. Cependant, nous nous intéresserons ici aux causes génétiques associées aux défauts congénitaux, car elles peuvent être maîtrisées en sélection génétique.

Les défauts génétiques ont un impact économique aux différents stades de la chaîne de production :

**Entre la fécondation et la naissance**, certaines anomalies chromosomiques peuvent entraîner la mortalité préimplantatoire. De plus, on estime que 1,1 à 2 % des porcelets n'arrivent pas à terme à cause d'un défaut de naissance.

**A la naissance**, les malformations diminuent la viabilité du porcelet. Le temps entre la mise bas et l'arrivée aux mamelles est crucial pour la survie du porcelet. Des malformations des membres ou des anomalies musculaires, telles que l'hypoplasie myofibrillaire et les tremblements congénitaux, handicapent le porcelet dans sa capacité à se déplacer jusqu'aux tétines. Toutes anomalies confondues, elles touchent 2,07 à 2,9% des porcelets dans les élevages (Mulley and Edwards, 1984; Thaller et al., 1996).

**Pendant le sevrage**, on estime que 5,6% des animaux ne survivent pas (BPEX, 2010). Les principaux problèmes létaux rencontrés sont l'atrésie anale, et les complications suite à l'étranglement de l'intestin dans les cas de hernies inguino-scrotales (Straw et al., 2009a). Il est important de souligner que le bien-être des animaux est affecté par ces malformations. Le meilleur moyen d'éviter la souffrance aux porcelets présentant des malformations sévères est parfois l'euthanasie, selon les méthodes recommandées par l'IFIP (Chevillon et al., 2004). Enfin, à terme les anomalies peuvent avoir un impact sur la croissance des animaux ou la qualité de la viande et induire des pertes économiques au moment de l'abattage des animaux.

**Durant la période de reproduction**, la fertilité des animaux peut être réduite à cause d'anomalies du tractus génital femelle. Celles-ci expliquent environ un quart des échecs de reproduction (Ollivier, 1979) et correspondent principalement à des obstructions de l'oviducte et des follicules kystiques.

Type d'anomalie	Anomalie
<b>Anomalies de la peau</b>	Epithrliogenesis imperfecta
	Pityriasis rosea
	Dermatosis vegetans
<b>Anomalies du squelette</b>	Hydroencéphalie
	Fissure palatine
	Anomalies de la colonne vertébrale
	Anomalies des membres
	Anomalies des doigts (syndactylie, polydactylie, absence)
	Faiblesse d'aplombs (Ostéochondrose)
<b>Anomalies des systèmes nerveux et musculaire</b>	Tremblements congénitaux
	Epilepsie
	Ataxie congénitale
	Paralyse du train postérieur
	Rigidité articulaire congénitale
	Hypoplasie myofibrillaire
	Autres myopathies
	Hyperthermie maligne
<b>Anomalies sanguines</b>	Porphyrie
	Hémophilie
	Leucémie
	Maladie hémolytique
	Thrombocytopénie
<b>Anomalies hormonales et métaboliques</b>	Nanisme
	Œdème du nouveau né
	Détresse respiratoire aigüe
	Rachitisme
	Obésité
<b>Anomalies du tractus digestif</b>	Hernie ombilicale
	Hernie Inguino-scrotale
	Anus imperforé
<b>Anomalies urogénitales</b>	Hypoplasie rénale
	Obstruction des urètres
	Cryptorchidie
	Stérilité du verrat
	Anomalies du tractus femelle
	Intersexualité
	Féminisation testiculaire
	Freemartinisme

Tableau 2 : Anomalies héréditaires référencées chez le porc (Ollivier, 1979)

La fonction reproductrice mâle peut également être affectée, soit par des défauts chromosomiques (Ducos et al., 2002), soit via des mutations affectant les gènes de la spermatogénèse (Sironen et al., 2006). Enfin, les anomalies du développement sexuel, tels que l'intersexualité ou la cryptorchidie ont un impact négatif sur la fertilité.

On estime aujourd'hui que les pertes financières dues aux maladies génétiques s'élèvent à 17,9 millions de dollars par an dans le monde (Walters, 2010). Ces estimations ne prennent en compte que les pertes dues à la mortalité et aux sacrifices des animaux malades. Au total, les coûts associés aux défauts génétiques pourraient être multipliés par dix.

### **I.C.2. Les maladies congénitales chez le porc**

Les différentes maladies héréditaires rapportées chez le porc sont nombreuses. Un rapport très complet a été proposé par Ollivier en 1979 sur toutes les anomalies héréditaires observées au moins une fois dans les élevages (Ollivier, 1979). Elles sont présentées dans le Tableau 2.

Depuis les travaux d'Ollivier il y a trente ans, aucune liste aussi exhaustive n'a été réalisée sur les défauts présents dans les élevages porcins, mais l'incidence des anomalies congénitales les plus fréquentes a été estimée dans plusieurs études (Bampton, 1994; Partlow et al., 1993). La dernière d'entre elles est une étude européenne réalisée entre 2006 et 2010, portant sur l'estimation de la fréquence actuelle de principales anomalies congénitales porcines (Walters, 2010). Elle a été réalisée sur près de 180.000 animaux Large White et Landrace et les résultats sont présentés dans la Figure 7. Les principales anomalies observées dans les élevages sont la hernie inguino-scrotale, la cryptorchidie et l'intersexualité.

Les résultats de l'étude européenne sont cohérents avec les principales anomalies congénitales répertoriées dans une étude française récente (Larzul et al., 2008). Cette étude porte sur la fréquence des principaux défauts de naissance estimée entre 1996 et 2006 dans trois populations sélectionnées françaises. D'après ces résultats, 1,09 % des animaux sont atteints de hernie scrotale, 0,37% de cryptorchidie, et 0,36% d'intersexualité. La hernie ombilicale a également été identifiée chez 1,26% des porcelets.

Outre des facteurs environnementaux et sanitaires (Mattsson, 2011), il est désormais connu que des effets génétiques sont impliqués dans le déterminisme de ces anomalies.

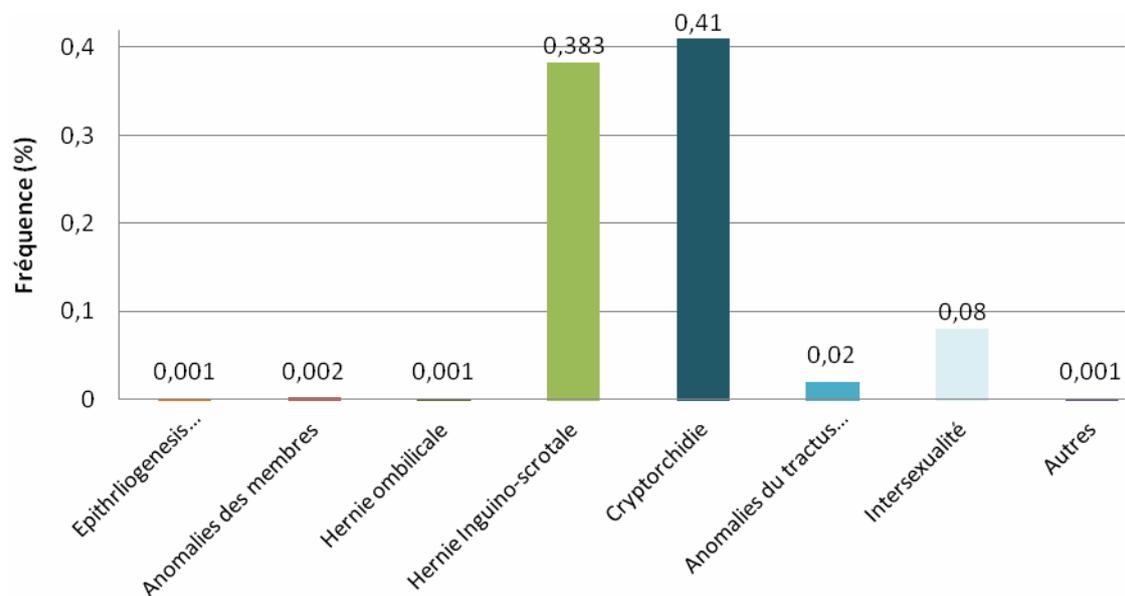


Figure 7 : Fréquence des principales anomalies congénitales dans l'espèce porcine selon une étude européenne réalisée entre 2006 et 2010 (Walters, 2010).

Auteurs	Année	Races	Hernie scrotale	Cryptorchidie	Intersexualité
Magee	1951	Lignées de l'Iowa, USA	0,15		
Mikami	1979	Populations A et B	0,65 et 0,86	0,5	
Sittmann	1980	Landrace hollandais			0,78
Knap	1986	Landrace, Yorshire allemands	0,35	0,16	
Vogt	1990	Duroc, Landrace, Yorshire	0,29, 0,34 et 0,34		
Ranberg	2007	Duroc, Landrace norvégiens	0,026	0,029	
Larzul	2008	Populations françaises A et B	A : 0,31 - B : 0,06 A+B : 0,27	A : 0,11 A+B : 0,07	A : 0,81 A+B : 0,72
Mattsson	2011	Hampshire, Landrace, Yorshire suédois	0,09 - 0,20 0,02	0,07 - 0,005 0,23	

Tableau 3 : Héritabilités de la hernie scrotale, la cryptorchidie et l'intersexualité selon différentes études réalisées entre 1951 et 2011

## **I.D. Déterminisme génétique des anomalies congénitales étudiées**

### **I.D.1. Etudes de ségrégation et estimations des paramètres génétiques**

La première étude génétique sur la hernie scrotale chez le cochon date de 1926 (Warwick, 1926). Warwick a produit trois générations d'animaux à partir d'un mâle Duroc-Jersey présentant une hernie inguinale bilatérale. Dès la 2<sup>ème</sup> génération, près de la moitié des porcelets présentaient une hernie. Ce type de croisement a également été réalisé pour l'intersexualité et la cryptorchidie (Pailhoux et al., 1997; Rothschild et al., 1988). Des descendants intersexués ont été obtenus à partir du croisement d'animaux ayant déjà donné des descendants intersexués dans des portées précédentes (Pailhoux et al., 2001a). Le croisement de mâles cryptorchides et de femelles dont les pleins frères présentaient des anomalies de descente testiculaire a permis de mettre en évidence le caractère héréditaire de la cryptorchidie.

Des effets race ont été mis en évidence pour la hernie scrotale : 0,39 % des porcelets Landrace allemands sont atteints contre 0,64% des porcelets Piétains d'après l'étude de Thaller (Thaller et al., 1996). D'autres travaux montrent des différences de fréquence entre Landrace et Large White : La fréquence pour la hernie scrotale est de 0,67 pour les Landrace contre 0,44 pour les Large White, celle pour la cryptorchidie de 0,13 pour les Large White et 0,31 pour les Landrace. Concernant le pourcentage d'atteints intersexués, il est similaire dans les deux races (0,05-0,06) (Walters, 2010)

Enfin des estimations d'héritabilité ont été réalisées pour la hernie scrotale, la cryptorchidie et l'intersexualité. Le Tableau 3 reprend les résultats obtenus dans de la cadre de 8 études menées de 1951 à 2011, pour des races et des pays différents. L'héritabilité la plus importante est enregistrée pour l'intersexualité, avec des valeurs comprises entre 0,72 et 0,81. Les héritabilités estimées pour la hernie scrotale et la cryptorchidie sont très variables – entre 0,02 et 0,86 pour la hernie scrotale, et 0,005 à 0,5 pour la cryptorchidie – ce qui semble souligner le caractère polygénique et complexe des deux maladies. Enfin, une étude a montré une corrélation génétique de 0,20 entre la hernie scrotale et la cryptorchidie (Mikami and Fredeen, 1979).

N° OMIA	Caractère	Gène identifié
001089-9825	Système du groupe sanguin ABO	GGTA1
001249-9825	Couleur de la peau : brune	TYRP1
000209-9825	Couleur de la peau : blanc dominant	KIT
001199-9825	Couleur de la peau : extension de la pigmentation	MC1R
001743-9825	Couleur de la peau : apparition de taches	KIT
001216-9825	Couleur de la peau : rouan	KIT
001745-9825	Couleur de la peau : ceinture blanche	KIT
000259-9825	Surdit�	MITF
001718-9825	Nanisme	COL10A1
001579-9825	Taille des oreilles	PPARD
000499-9825	Hypercholest�rol�mie	LDLR
000621-9825	Hyperthermie maligne	RYR1
001085-9825	Qualit� de la viande (Rendement Napole)	PRKAG3
000636-9825	glom�rulon�phrite membrano-prolif�rative type II	CFH
000683-9825	Hypertrophie musculaire	MSTN
001436-9825	Absence de thermor�gulation non associ�e aux frissons	UCP1
000862-9825	R�sistance � l'�d�me du nouveau n� (r�cepteur F18)	FUT1
001334-9825	Flagelle spermatique raccourci	SPEF2
001673-9825	Arr�t de la spermatog�nese	TEX14
001685-9825	Syndrome de Stress	DMD
001200-9825	Tremblements de haute fr�quence (Campus syndrome)	MYH7
000837-9825	Rachitisme type I par carence en vitamine D	CYP27B1

Tableau 4 : Caract res porcins pour lesquels le g ne impliqu  a  t  identifi  - source : OMIA

## I.D.2. Causes moléculaires des anomalies congénitales

### *Anomalies chromosomiques :*

Près de 150 anomalies chromosomiques ont été rapportées dans l'espèce porcine (Ducos et al., 2002; Pinton et al., 2012), et concernent principalement des translocations réciproques. Bien qu'il existe peu d'études sur les conséquences de ces anomalies, certains travaux ont montré leur association à l'hypoproliféricité. Les anomalies chromosomiques peuvent être associées à une baisse de la fertilité chez le porc comme chez d'autres espèces (Henricson and Backstrom, 1964). La diminution de la taille de portée issue d'un verrat avec une anomalie chromosomique peut atteindre -68%, comme c'est le cas pour la translocation réciproque entre les chromosomes 6(p1.5) et 13(q4.1). Deux remaniements chromosomiques ont été associés à l'intersexualité : une duplication du chromosome 9, et une inversion sur le chromosome 9 de type p12;p22 (Ducos et al., 1997, 2002; Pinton et al., 2011, 2012). Trois cas de caryotypes 39,XXY et un cas de mosaïcisme 38,XY/39,XXY ont été rapportés chez le porc. Chez l'homme, la duplication d'un chromosome X, soit un caryotype 47,XXY, correspond au syndrome de Klinefelter et représente l'anomalie chromosomique sexuelle la plus fréquente chez l'homme, touchant 15 naissances sur 10000 (Groth et al., 2013).

Aujourd'hui le caryotype de 90% des verrats candidats à la sélection est contrôlé afin de limiter la propagation d'anomalies chromosomiques. A posteriori, une attention particulière est portée aux animaux géniteurs de portées hypoprolifériques. L'hypoproliféricité est le premier caractère phénotypique affecté par ce type de défaut génétique, des contrôles complémentaires sont donc réalisés chez ces géniteurs.

### *Mutations de l'ADN :*

Les causes moléculaires de la plupart des anomalies ne sont pas connues. Une banque de données l'OMIA (pour Online Mendelian Inheritance in Animals) référence les caractères pour lesquels la transmission héréditaire a été décrite chez les animaux. Chez le porc, bien que 221 caractères soient enregistrés dans cette base, le déterminisme génétique sous-jacent a été caractérisé dans 22 cas seulement. Ils sont présentés dans le Tableau 4.

L'ensemble de ces données montre qu'une part du déterminisme des anomalies les plus fréquentes chez le porc est due à des facteurs génétiques. Les outils aujourd'hui disponibles en génétique porcine apportent des moyens nouveaux pour la recherche et l'identification de

mécanismes moléculaires associés. Ces outils sont présentés dans la partie suivante de ce chapitre bibliographique.

## II. Les outils de cartographie du génome du porc

### II.A La cartographie physique

#### II.A.1. Cartographie physique chromosomique

L'objectif de la cartographie physique est de déterminer, à l'aide de mesures physiques (exprimées en nombre de bases) la position et les distances séparant des loci le long du génome. Les premières cartes physiques ont été réalisées par l'hybridation *in situ* de sondes marquées du locus d'intérêt, sur des préparations métaphasiques de chromosomes. Les premières sondes utilisées étaient marquées via un isotope radioactif (Gall and Pardue, 1969) puis par fluorescence (on parle alors de *Fluorescence In Situ Hybridization* ou FISH) (Rudkin and Stollar, 1977). Cette technique permet de localiser les loci sur les chromosomes, et d'estimer leurs positions relatives. Néanmoins la résolution est relativement faible, de l'ordre de 5Mb. A partir des années 70 le développement de panels d'hybrides somatiques chez certaines espèces a permis d'obtenir une information équivalente voire, avec certains panels ou dans certaines régions chromosomiques, d'obtenir une meilleure résolution (Weiss and Green, 1967). Les hybrides somatiques sont des clones de cellules, issus de la fusion d'une lignée établie receveuse (le plus souvent de hamster ou de souris) et d'une lignée primaire d'intérêt. Au cours de leur multiplication *in vitro*, les cellules fusionnées perdent aléatoirement des chromosomes de la lignée primaire et de nombreux remaniements surviennent entre les chromosomes des deux espèces. La "ségrégation" non équiprobable des chromosomes de la lignée primaire permet d'obtenir une série de clones indépendants contenant une portion du génome d'intérêt. Dans les années 80, l'évolution des techniques de cytogénétique a permis de caractériser finement le contenu chromosomique de l'espèce d'intérêt, dans chaque clone. Dès lors, l'estimation de la présence ou l'absence de marqueurs spécifiques du type cellulaire d'intérêt dans chacun des hybrides permettait d'obtenir sa position chromosomique. L'avantage majeur de cet outil est sa facilité d'utilisation : la localisation d'un marqueur est obtenue via l'amplification PCR du locus d'intérêt à partir de l'ADN de chaque clone hybride du panel. Le second avantage majeur de cet outil de cartographie est qu'il ne nécessite pas que le marqueur cartographié soit polymorphe. Les premiers hybrides somatiques contenant de l'ADN humain ont été mis au point en 1967 par Weiss (Weiss and Green, 1967). En 1996, différents panels ont été développés chez le porc (Rettenberger et al., 1996; Yerle et al., 1996). Un d'entre eux est utilisé comme référence et comprend 27 clones d'hybrides.

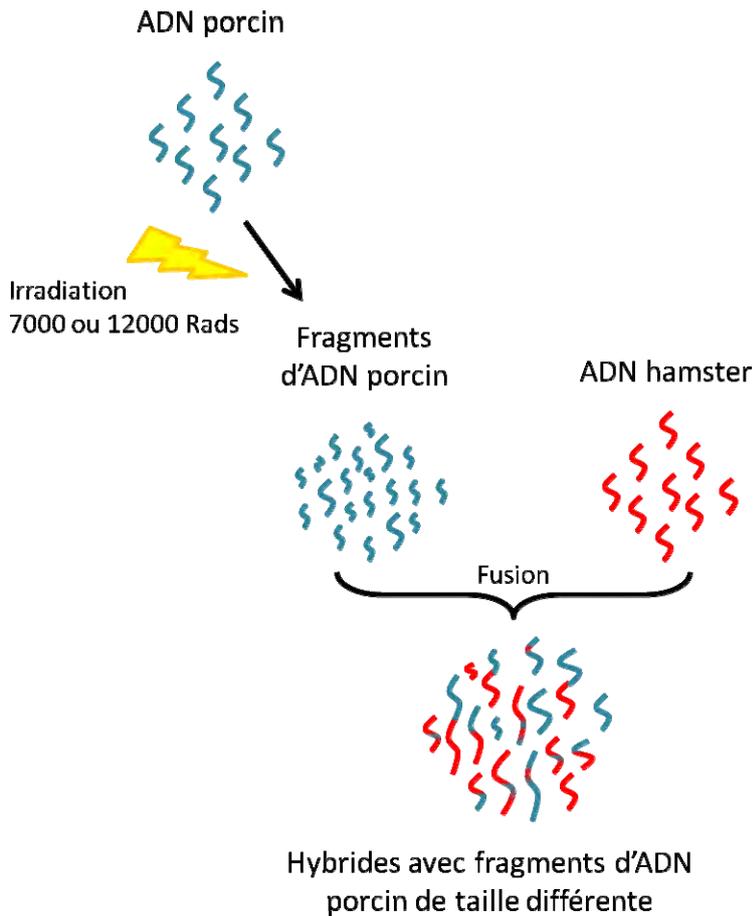


Figure 8 : Principe de la création des panels d'hybrides d'irradiation

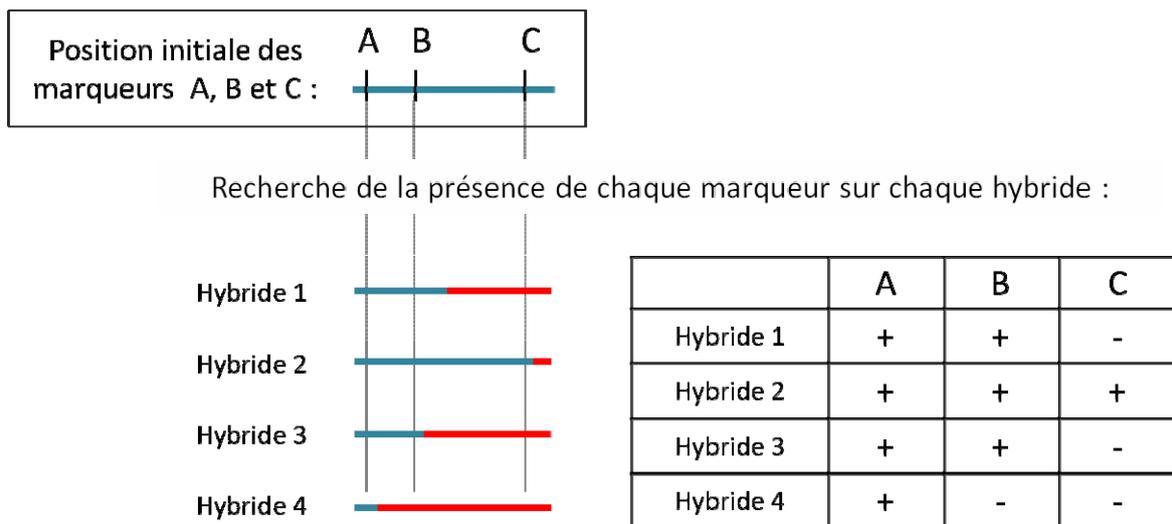


Figure 9 : Principe de l'estimation de l'ordre des loci les uns par rapport aux autres. Exemple avec 3 marqueurs et 4 hybrides (la présence d'un marqueur dans un hybride est noté "+", son absence est noté "-")

La résolution des cartes d'hybrides somatiques est de l'ordre du remaniement chromosomique. Afin d'améliorer cette résolution, de nouveaux panels d'hybrides ont été développés, les panels d'hybride d'irradiation ou RH (pour *Radiation Hybrid*). Le principe de cet outil repose sur l'irradiation de la lignée primaire d'intérêt afin de fragmenter le génome avant la fusion. Plus la dose d'irradiation est élevée, plus les fragments d'ADN sont petits et plus la cartographie sera fine. A partir du nombre de fois où les marqueurs sont présents ensemble dans les mêmes hybrides par rapport au nombre de fois où ils sont présents sur des hybrides différents, on estime la probabilité qu'ils soient proches, et la distance relative qui les sépare, exprimée en centiRays (cR). Par définition 1 cR correspond à une fréquence de cassure de 1% entre deux loci (Farce, 2000).

Plusieurs panels d'hybrides irradiés existent chez l'homme. Les trois panels aujourd'hui utilisés sont GB4 (Gyapay et al., 1996) irradié à 3000 rads, G3 irradié à 10000 rads et TNG4 50000 rads (Stewart et al., 1997).

Chez les animaux d'élevage, un panel bovin de 100 hybrides, irradiés à 5000 rads, est disponible (Womack et al., 1997). Deux panels sont utilisés chez le porc. Le panel IMpRH (pour *INRA-Minnesota Porcine Radiation Hybrid*) est composé de 118 ou 90 hybrides irradiés à 7000 rads (Yerle et al., 1998), et le panel IMNpRH2, composé de 90 hybrides, a été irradié à 12000 rads (Yerle et al., 2002).

Le principe de la création des hybrides d'irradiation porcins est présenté en Figure 8. Les lignées primaires utilisées sont des lymphocytes ou des fibroblastes porcins. Seuls les clones possédant la séquence du gène HPRT (*Hypoxanthine-guanine Phosphoribosyl Transferase*) sont sélectionnés: Ce gène est présent dans le génome porcine, mais est inactivé dans les lignées murines utilisées; ainsi seuls les clones contenant de l'ADN porcine sont sélectionnés. Comme pour les panels d'hybrides somatiques l'utilisation de cet outil de cartographie est simple ; chaque locus d'intérêt est amplifié par PCR (*Polymerase Chain Reaction*) à partir de l'ADN de chacun des hybrides (Figure 9). La première carte RH porcine a été réalisée à partir du panel IMpRH avec essentiellement les marqueurs microsatellites disponibles sur la carte génétique. Une première carte de 757 marqueurs localisés sur l'ensemble du génome, comprenant 19 EST (*Expressed Sequence Tag*) et 39 gènes, a été publiée (Hawken et al., 1999). Afin de densifier cette carte, une base de données a été mise en place pour centraliser l'ensemble des localisations obtenues à l'aide des panels hybrides d'irradiation IMpRH et IMNpRH2 (Milan et al., 2000a).

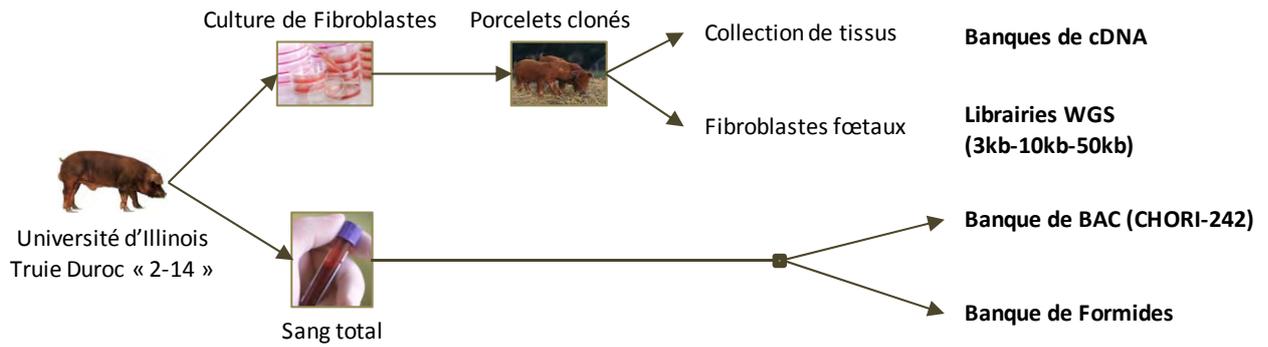


Figure 10 : Construction des banques de donn es de s quences chez le porc

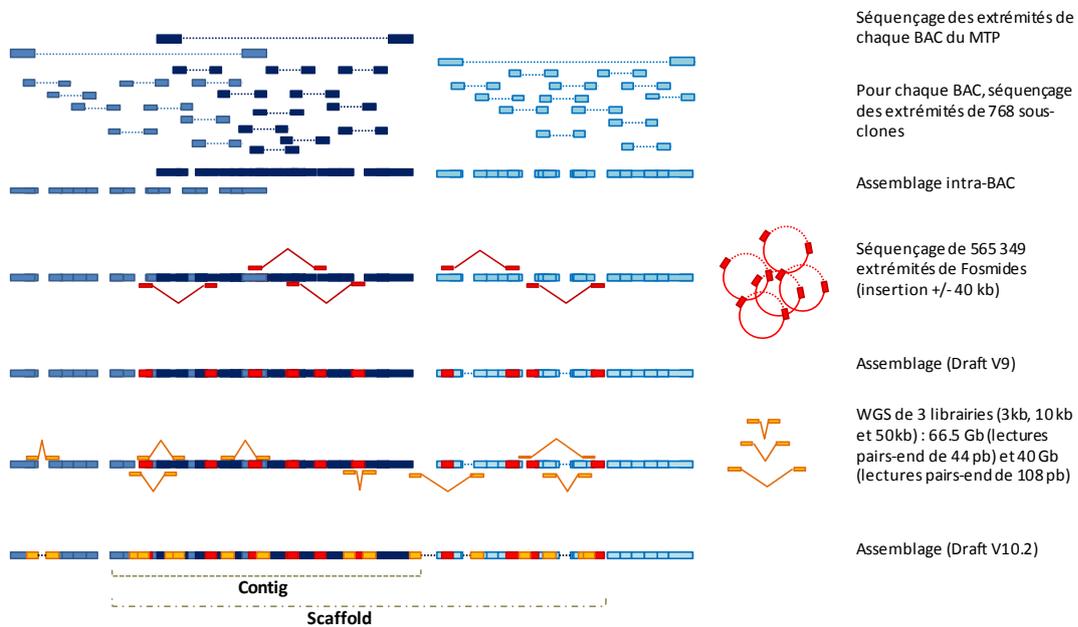


Figure 11 : Construction du draft porcine

Une seule carte de référence mise à jour au fur et à mesure de l'accumulation des données a ainsi été produite. A partir de ces panels de nombreux travaux de cartographie ont été publiés et sont destinés (1) à caractériser finement la localisation d'un maximum de gènes sur le génome porcin, (2) à préciser les homologies entre le génome du porc et le génomes d'autres espèces (essentiellement l'homme) et (3) à réaliser une cartographie haute résolution de régions QTL d'intérêt.

Récemment, une cartographie RH haute résolution a été réalisée à l'aide de ces mêmes panels et de 38.000 marqueurs moléculaires, issus de la puce porcine de génotypage haut débit (Servin et al., 2012). Cette dernière carte présente une densité en marqueurs multipliée par 50 par rapport à la toute première carte RH porcine publiée (Hawken et al., 1999).

### **II.A.2 Cartographie physique moléculaire**

Au fur et à mesure du temps, de nouvelles technologies ont été mises en place afin de parvenir à une résolution de cartographie de plus en plus fine. Des résolutions de quelques dizaines de kb ont été obtenues à l'aide de banques de grands fragments d'ADN comme les banques BAC (*Bacterial Artificial Chromosome*). Des banques de BAC ont été construites chez de nombreuses espèces telles que l'homme (Asakawa et al., 1997), le bovin (Cai et al., 1995), la chèvre (Schibler et al., 1998) et le porc (Suzuki et al., 2000). Chez le porc, il existe en tout 5 banques de BAC et la couverture totale du génome porcin à l'aide de ces clones est estimée à 38X (Rothschild and Ruvinsky, 2011). La banque choisie comme référence pour l'obtention d'un séquençage complet du génome est une librairie américaine générée à partir d'une femelle Duroc (Figure 10). A partir des années 2000, un programme de grande ampleur destiné à obtenir un premier draft du génome du porc a été initié dans le cadre d'un consortium international.

La stratégie choisie pour le séquençage du génome de cette espèce est un séquençage par ordonnancement hiérarchique à partir des clones BAC constituant un *Minimum Tiling Path* (MTP) de l'ensemble du génome. Le MTP a permis d'assurer une couverture de 98.3% de la carte physique. Pour chaque BAC, les deux extrémités de l'insertion et les extrémités de 768 sous-clones ont été séquencées (Figure 11). Afin de compléter la couverture du génome apportée par le MTP, une banque de fosmidés (40 kb) a été construite à partir du même animal et près de 600.000 extrémités de fosmidés ont été ajoutées à ce premier assemblage, offrant une couverture de 4 X.

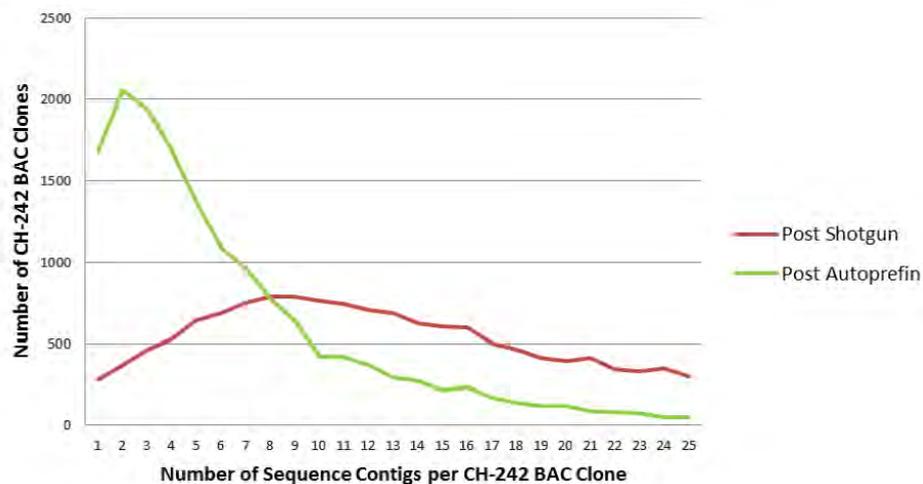


Figure 12 : Nombre de contigs par clone avant (*Post Shotgun*) et après l'étape de *primer walking* (=Autoprefin)

	Draft V9	Draft V10.2
Taille totale	2 262 484 801	2 808 525 991
Nombre de contigs	100 198	241 883
Nombre de scaffold (dont non-cartographiés)	3135	9905 (4562)

Tableau 5 : Taille totale, nombre de contigs et nombre de scaffolds des Drafts v9 et v10.2

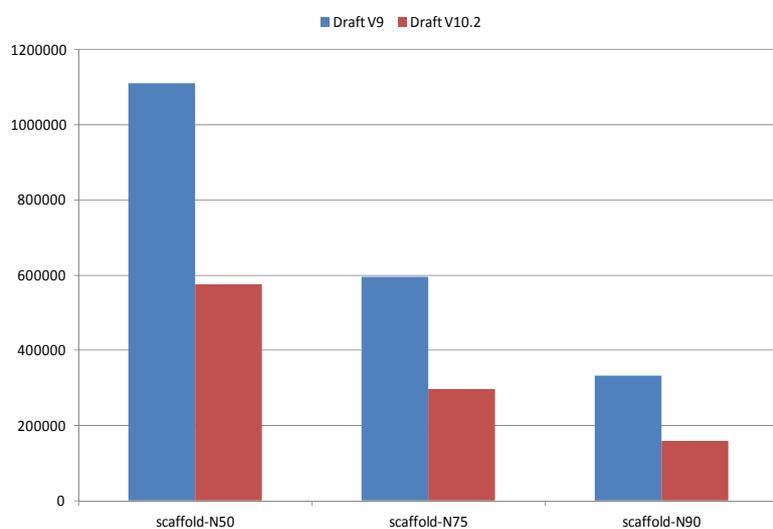


Figure 13 : Valeurs statistiques N50, N75 et N90 pour les scaffolds des drafts v9 et v10.2

Enfin une étape de *primer walking* (amplification de séquence à partir du design d'amorces des extrémités des contigs) a permis d'améliorer la couverture d'un facteur 2 X environ (Figure 12). Ce premier assemblage a donné lieu au draft 9 porcine.

Dans un second temps, suite aux évolutions des technologies de séquençage haut débit (NGS : *Next Generation Sequencing*) des fragments d'ADN obtenus aléatoirement ont été séquencés puis ordonnés (Groenen et al., 2012). L'ADN utilisé est issu de la même femelle Duroc (Figure 10). Cette deuxième méthode est appelée globale, ou *Whole-Genome Shotgun* (WGS). Elle a permis d'obtenir 66.5 Gb de lectures en paired-end de 44pb et 40 Gb de lectures de 108pb (Figure 11). L'avantage du WGS est la rapidité et un coût moindre, par rapport à la méthode par ordonnancement hiérarchique. L'inconvénient est l'ordonnancement des séquences répétées, fréquentes chez les mammifères, qui est difficile à réaliser. Parallèlement, le séquençage systématique de banque de cDNA a permis de recueillir une information complémentaire et spécifique des régions transcrites du génome. L'ensemble des données de séquençage obtenues a été combiné et les chromosomes et gènes ont été annotés.

Des méthodes de cartographie complémentaires ont permis de corriger d'éventuelles erreurs d'assemblage. Servin et al. ont ainsi étudié l'ordonnancement des marqueurs moléculaires de la puce porcine par méthode de cartographie RH (Servin et al., 2012). Ils ont montré que l'ordre des marqueurs était globalement cohérent entre les résultats de cartographie RH et leur ordre dans l'assemblage de référence, à l'exception de quelques régions qui ont été rectifiées. La cartographie RH a également été réalisée avec les marqueurs identifiés sur les scaffolds non placés dans le génome. Cette méthode a permis d'intégrer 72 mégabases de séquence génomique sans position initiale dans l'assemblage. L'ensemble de ces nouvelles informations a donné lieu à une mise à jour de la séquence de référence (version 10.2).

Si l'on compare les deux drafts porcins, la taille totale du génome de la v10.2 (2.80 Gb) est plus élevée que celle de la v9 (2.26 Gb) (Tableau 5), ce qui s'explique par le nombre plus important de scaffolds et de contigs (2.5 à 3 fois plus élevé entre les deux versions). Cependant, la taille des scaffolds est moins longue entre la v9 et la v10.2 comme l'indique les valeurs de N\* (Figure 13) (la valeur du paramètre N50 signifie que plus de 50% des scaffolds sont d'une longueur égale ou supérieure à cette valeur). Ces résultats montrent que la v10.2 est composée de scaffolds deux fois plus petits que la v9, mais deux fois plus nombreux.

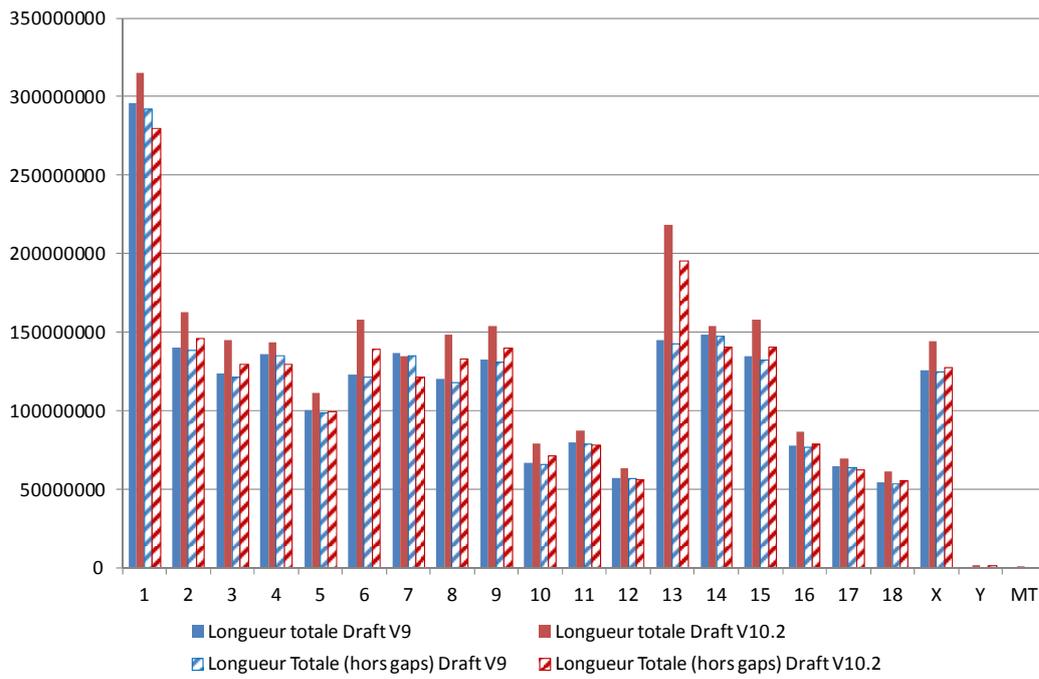


Figure 14 : Longueur totale (en pb) et hors gaps par chromosome sur les drafts 9 et 10.2

Le "gain" de séquence entre les deux versions n'est pas homogène pour tous les chromosomes : à l'extrême, la séquence disponible est plus petite suite à la mise à jour, comme pour les chromosomes 1 et 7, alors que la taille de la séquence disponible pour le chromosome 13 a augmenté de 25% (Figure 14).

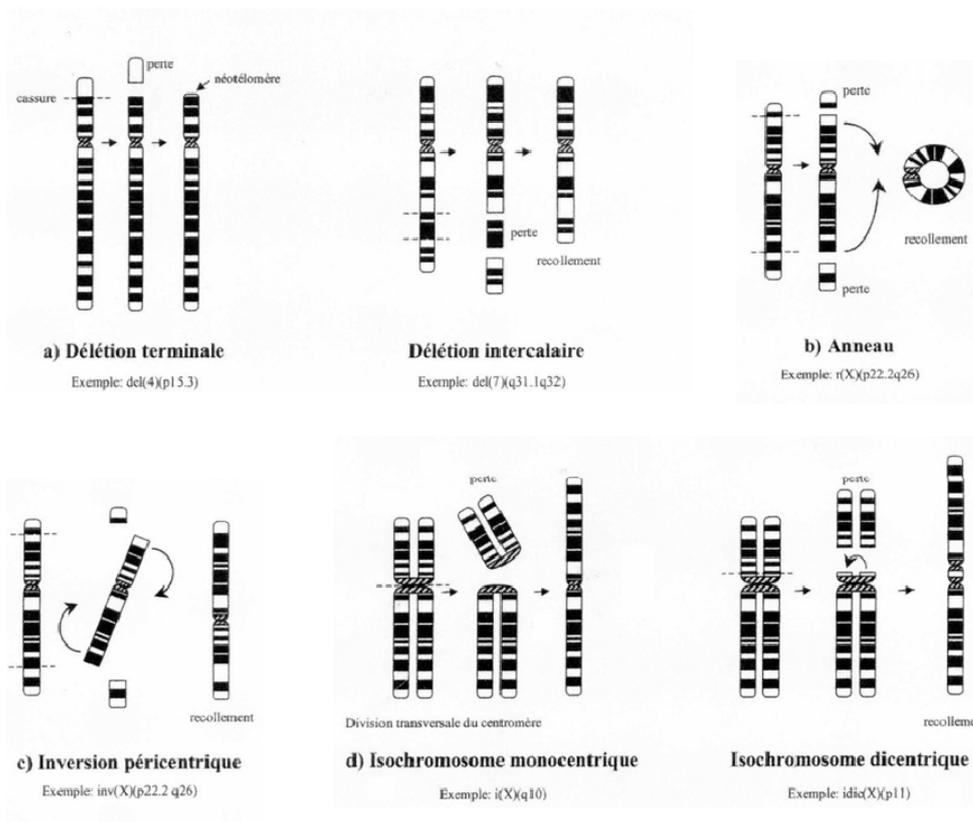


Figure 15 : Modifications de structure affectant un chromosome

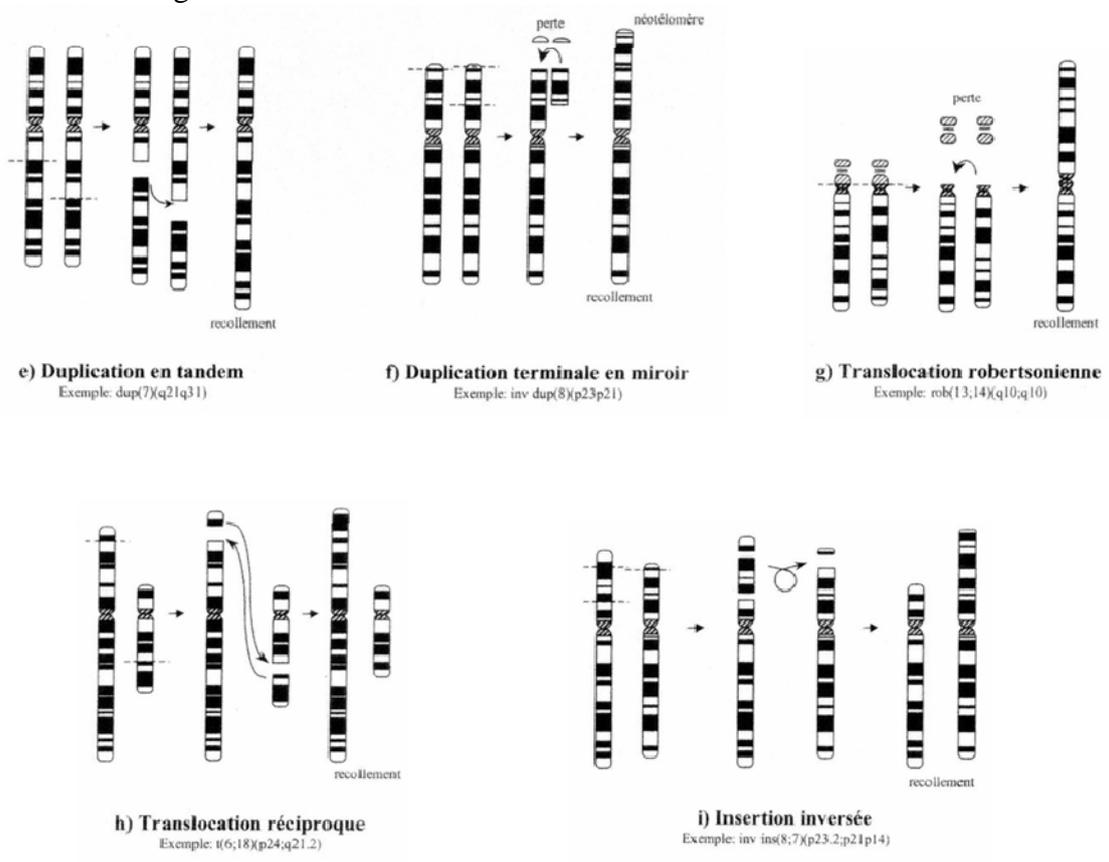


Figure 16 : Modifications de structure impliquant deux chromosomes

## **II.B Les variations de l'ADN**

Pour deux individus d'une même espèce, environ 1% du génome diffère (Venter et al., 2001). Ce 1% de variation peut expliquer certaines différences phénotypiques observées, ou être sans conséquence connue. Ces variations peuvent être des macro ou des micro-remaniements, des variations du nombre de copies de segments d'ADN (CNV), des séquences particulières répétées dans le génome, ou des variations d'un seul nucléotide (SNP). En cartographie génétique ces variations peuvent être utilisées comme marqueurs moléculaires.

### **II.B.1. Les macro-remaniements chromosomiques**

Les macro-remaniements chromosomiques correspondent à des modifications dans la structure d'un ou de plusieurs chromosomes. Des exemples de remaniements sont présentés sur les Figure 15 et Figure 16 (Turleau and Prieur, 2000). Les modifications d'un chromosome peuvent être des délétions, en extrémité de chromosome, ou dans un bras chromosomique, des inversions autour du centromère, ou des structures particulières telles que la formation d'anneaux ou d'isochromosomes. Les isochromosomes correspondent à des chromosomes composés de deux bras courts ou deux bras longs, avec perte de l'autre bras (Figure 15), la délétion pouvant survenir au niveau du centromère, pour les isochromosomes monocentriques, ou après la région centromérique, pour les isochromosomes dicentriques. Lorsque plusieurs chromosomes sont impliqués, les modifications peuvent être des duplications en tandem (le fragment se duplique dans le même sens) ou en miroir (le fragment se duplique en sens inverse), des translocations robertsonniennes (au niveau des centromères) ou réciproques, et des inversions (Figure 16).

Près de 150 anomalies chromosomiques porcines ont été rapportées (Ducos et al., 2002; Pinton et al., 2012). Les plus fréquentes sont des translocations robertsoniennes (0,9 naissances pour 1000) et réciproques (0,9 à 1,4 pour 1000). Certaines de ces anomalies sont associées à une hypoproliférite (diminution de la taille de portée jusqu'à -68% pour la translocation réciproque entre SSC6 (p1.5) et SSC13 (q4.1) (Henricson and Backstrom, 1964)).

Les caryotypes ont d'abord été réalisés par coloration des chromosomes bloqués en métaphase, puis via l'hybridation de sondes fluorescentes. De nouvelles méthodes d'hybridation sur puce permettent aujourd'hui de mettre en évidence des réarrangements chromosomiques d'une Mb à quelques kb (Andrieux, 2008).

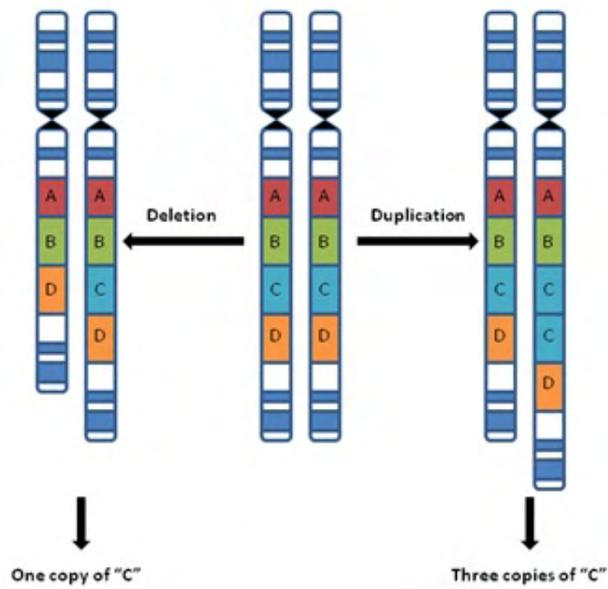


Figure 17 : Représentation de CNV : à gauche la délétion d'une copie du locus C, à droite la duplication d'une copie du locus C.

## II.B.2. Les variations du nombre de copies d'une séquence

Les CNV (pour *Copy Number Variation*) correspondent à la variation du nombre de copies d'un gène ou d'un segment d'ADN de plus de 1kb (Figure 17). Dans les années 20, une duplication d'un segment chromosomique, visible à l'échelle cytologique, a été mise en évidence chez la drosophile. Cette duplication était associée à la diminution de la taille de l'œil (Sutton, 1943). Depuis, il est établi que ce type de variation peut expliquer des différences phénotypiques entre les individus (Perry et al., 2007). Chez l'homme, les CNV couvrent environ 3,7% du génome (Conrad et al., 2012). Ils concerneraient 4,6 % du génome bovin (Hou et al., 2011) et 5% du génome murin (She et al., 2008).

Le nombre de CNV chez le porc a été estimé selon deux types de méthodes. La première méthode est l'hybridation génomique comparative (CGH pour *Comparative Genomic Hybridation*) qui consiste à comparer la fluorescence de deux échantillons d'ADN marqués différemment l'un de l'autre. La deuxième méthode est l'utilisation des puces de génotypage où, pour chaque SNP donné, la fluorescence de l'allèle 1 est comparée à la fluorescence de l'allèle 2 (la somme des deux fluorescences est normalisée et doit être égale à 1 en absence de CNV, inférieure à 1 dans le cas de délétion et supérieure à 1 lors de duplication).

Les premiers travaux ont été réalisés par CGH (Fadista et al., 2008) sur quatre chromosomes porcins (SSC4, 7, 14 et 17) et douze animaux Duroc. Trente sept CNV de taille supérieure à 1kb (1,74 à 61,92kb) ont été identifiés, correspondant d'après leur estimation à 0,18% du génome porcin. Par la même approche, une seconde étude en tout génome sur douze porcs (trois lignées chinoises, cinq lignées européennes (Large White, Piétrain, White Duroc and deux Landrace), deux lignées synthétiques et deux lignées croisées) a permis d'estimer que le pourcentage de CNV dans le génome serait de 0,74% (Li et al., 2012), la taille des régions variant de 2,30 kb à 1,55 Mb (moyenne de 65,07 kb).

La puce porcine a été utilisée dans plusieurs études pour la recherche de CNV. Ramayo-Caldas et al ont identifié 49 régions CNV, à partir d'un échantillon de 55 porcs ibériques x Landrace (Ramayo-Caldas et al., 2010). Wang et al ont estimé la fréquence de CNV (taille moyenne de 250,7kb) à 4,23% (382 régions) à partir de 474 porcs (Yorkshire, Landrace, Songliao Black, et Duroc×Erhualian) (Wang et al., 2012). Enfin, Chen et al. ont travaillé sur 1693 animaux issus de 18 populations différentes (Chen et al., 2012). Ils ont identifié 565 régions (5,84% du génome) de 50,39 kb à 8,10 Mb (moyenne de 252,71 kb).



Les résultats des études avec la puce porcine sont significativement différents des données de CGH. Cette différence peut s'expliquer par les méthodes utilisées et le dispositif sélectionné. La détection de CNV avec la puce haute densité est limitée aux séquences d'ADN de 150 kb environ, car la densité en marqueurs de la puce est de un SNP tous les 50kb environ, et que des fenêtres de trois SNP sont utilisées dans la détection de CNV. Le nombre de CNV est donc potentiellement supérieur si l'on considère les segments de moins de 150kb. Cependant les résultats obtenus avec les données de la puce sont plus cohérents avec les résultats chez les autres espèces que ceux obtenus avec les données de CGH, dont le seuil de détection de CNV est plus faible. Des différences entre les races peuvent également expliquer ces variations. L'étude de Chen rapporte par exemple 23 CNV chez les Duroc blancs contre 683 chez les F2 Duroc blanc x Erhualian. Enfin, le nombre de CNV est très variable d'un chromosome à l'autre, et pourrait expliquer les variations observées entre les deux études utilisant la CGH.

Les CNV sont préférentiellement localisés dans des régions non géniques (Bejerano et al., 2004). Cependant certains gènes peuvent avoir été impliqués dans des événements de duplication et être présents en nombre variable de copies. La duplication du nombre de copies d'un gène peut être un avantage pour l'évolution, car les copies fournissent une « matière première » aux mutations, les nouvelles copies pouvant muter ou prendre de nouvelles fonctions. A contrario, ces duplications peuvent parfois avoir des effets délétères lorsque les copies supplémentaires d'un gène entraînent un surdosage de la protéine dont l'impact est néfaste pour la fonction physiologique où elle intervient (Clancy et al., 2008).

Chez le porc, près de 1500 gènes ont été identifiés dans des CNV (Chen et al., 2012). Certains d'entre eux, comme les gènes associés au système immunitaire et olfactif, ou à la croissance et la reproduction, ont également été identifiés chez l'homme et la souris (Du et al., 2010). Deux grandes fonctions semblent particulièrement corrélées aux CNV identifiés chez le porc. La famille des ORF (*Olfactory Receptor Activity*) est surreprésentée dans l'espèce porcine, soulignant l'importance de l'olfaction pour le porc dans sa quête de nourriture. Une grande partie des gènes en nombre de copies variable est également associée au métabolisme lipidique. Cette expansion est spécifique au porc, et peut s'expliquer par l'histoire évolutive de cette espèce (Du et al., 2010). Enfin, le gène Kit de coloration a fait l'objet de nombreuses études. La version normale du gène code pour la couleur foncée, alors que le gène muté et dupliqué est responsable de la couleur blanche comme détaillé dans la Figure 18. Selon le génotype la couleur de l'animal sera différente.

	Dog	Mouse	Pig	Rat	Human	Cow
SINEs	7.96	7.63	11.3	7.78	13.14	17.66
LINEs	19.54	16.46	16.14	20.1	20.42	23.29
LTR elements	10.39	8.72	2.8	10.28	8.29	3.2
DNA elements	0.88	0.36	1.51	0.86	2.84	1.96
Unclassified	0.32	0.37	0	0.37	0.14	na
Small RNA	0.06	0.04	0.02	0.03	na	na
Satellite repeats	0.04	na	1.47	0.31	na	na
Simple repeats	2.39	na	0.62	2.41	na	2.27
Low complexity repetitive sequences	0.73	na	0.53	na	na	na
Total	42.31	33.58	34.39	42.14	44.83	48.38
Genome size	2.45	2.5	2.56	2.75	2.85	2.87

Tableau 6 : Comparaison de la taille du génome et du contenu en éléments répétés chez plusieurs mammifères. Les données des éléments chez le porc sont issues de l'étude de Wernersson (Rothschild and Ruvinsky, 2011).

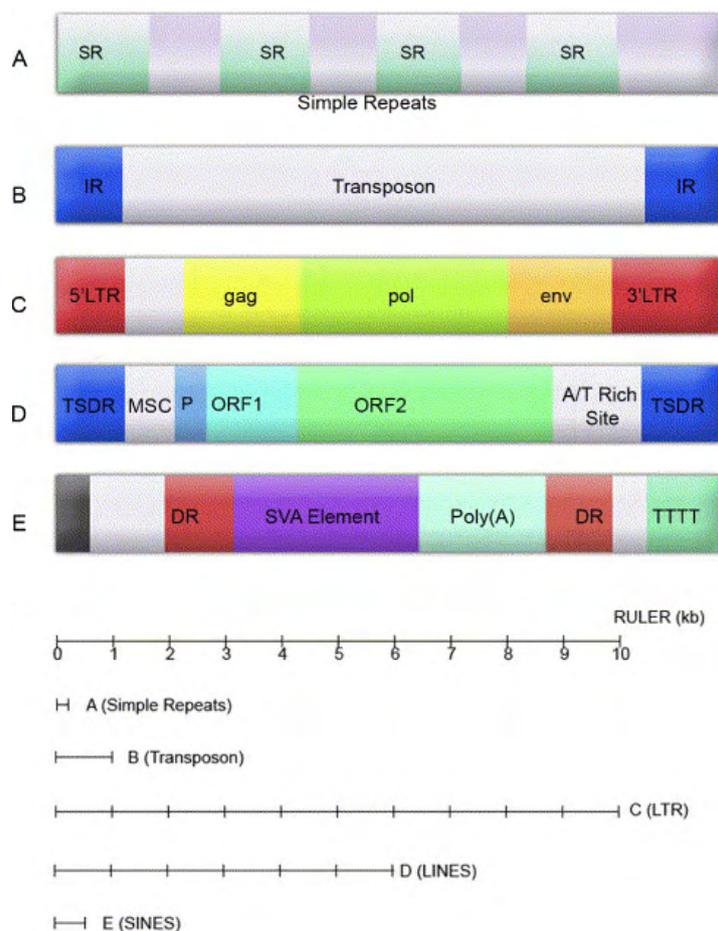


Figure 19 : Les différentes catégories d'éléments transposés et leur taille (Wilson et al., 2007). A : répétitions simples, B : transposon, C : LTR, D : LINE, E : SINE. Les barres au bas de l'image indiquent la taille des différents éléments.

### II.B.3. Les séquences répétées du génome

Les génomes des Eucaryotes supérieurs sont constitués de 40-45% d'éléments répétés (Kramerov and Vassetzky, 2005) (Tableau 6). Cet ADN répétitif peut être composé de séquences fonctionnelles, telles que les familles de gènes dispersés (histones, hémoglobine), répétés en tandem (ARN ribosomal) (Griffiths et al., 2001), ou de séquences sans fonction connue (environ 20% du génome), comme au niveau des centromères ou des télomères. Le génome du porc est composé de 40% de séquences répétées. Ce pourcentage est similaire aux estimations faites chez la souris et inférieure aux autres mammifères (Groenen et al., 2012; Rothschild and Ruvinsky, 2011). Parmi ces séquences répétées, certaines sont en tandem alors que d'autres correspondent à des éléments dispersés.

#### *Les séquences répétées dispersées*

Encore appelées éléments transposés, ce sont des séquences d'ADN capables de se déplacer dans le génome. Ils ont été découverts en 1983 par Barbara Mc Clintock (McClintock, 1983). Leur homologie de séquence avec les virus laisse supposer une origine commune. Ils représentent un tiers des génomes des mammifères (Sela et al., 2010) et semblent jouer un rôle important dans l'évolution et la spécialisation des espèces. Les éléments répétés peuvent être divisés en deux classes selon leur nature : les éléments à ADN, ou transposons, qui se déplacent de façon directe dans le génome, et les éléments à ARN, ou retrotransposons, qui se propagent par transcription inverse d'un ARN intermédiaire. Parmi les éléments à ARN, on distingue les éléments encadrés par des séquences terminales longues répétées ou LTR pour *long-terminal repeats*, et ceux qui ne possèdent pas ces répétitions terminales (Figure 19).

La famille des non LTR est la catégorie la plus représentée dans les génomes (Tableau 6). On distingue deux types de non LTR : les éléments courts, les SINE pour *Short Interspersed Repeat Elements*, ou longs : LINE (*Long Interspersed Repeat Elements*). Certains de ces éléments sont encore capables de se déplacer dans le génome. Chez l'homme ils appartiennent aux catégories SINE-Alu, spécifiques des primates, et aux LINE 1, ou L1 (Mills et al., 2007). La présence d'éléments homologues aux L1 retrouvés chez les plantes et les protistes laisse supposer que l'origine des L1 est très ancienne (Rothschild and Ruvinsky, 2011). Chez l'homme et chez la souris, les L1 sont associés à la chromatine condensée (Bailey et al., 2000).



### *Les séquences en tandem:*

Les séquences en tandem sont des répétitions de taille variable. On distingue les minisatellites ou VNTR (*Variable Number Tandem Repeats*) qui sont des répétitions de dix à quelques centaines de nucléotides, et les microsatellites, plus courts, de deux à dix nucléotides. La taille totale de la répétition peut atteindre 5kb pour les minisatellites, celle des microsatellites quelques dizaines de pb.

Les minisatellites ont été identifiés en 1980 (Wyman and White, 1980). Ils sont retrouvés préférentiellement au niveau des centromères et des télomères. Jusqu'à la mise au point de la technologie de PCR, les minisatellites ont été utilisés comme marqueurs génétiques pour différentes espèces. Les méthodes de biologie moléculaire (southern blot) utilisées à cette époque pour l'analyse des marqueurs génétiques ne permettaient pas de tirer parti du polymorphisme des microsatellites. Mais à partir des années 90, ces motifs répétés de petite taille sont devenus les marqueurs de référence de nombreuses cartes génétiques. Ils présentent l'avantage d'être (1) fréquents (1/50000bp), (2) bien répartis tout au long du génome, (3) polymorphes et co-dominants et (4) faciles et peu coûteux à génotyper. Les microsatellites chez le porc ont permis la construction des premières cartes génétiques balisant la totalité du génome (Archibald et al., 1995; Rohrer et al., 1996). Ils ont été utilisés pour la mesure de distances génétiques entre lignées ou populations, et l'identification de gènes d'intérêt économique (Behl et al., 1999).

Bien que les microsatellites soient encore couramment utilisés comme marqueurs génétiques, pour les analyses de structures de population par exemple (Rothschild and Ruvinsky, 2011), les SNP sont désormais préférés pour la cartographie génétique haute densité, car ils sont plus faciles à génotyper avec les outils actuels, et sont présents en plus grand nombre.



#### II.B.4. Les SNP et la puce 60K porcine

Le SNP (pour *Single Nucleotide Polymorphism*) correspond à la variation d'un seul nucléotide et représente la principale source de variation des génomes (90% du nombre total de polymorphismes humains) (Venter et al., 2001). L'estimation du nombre de SNP par génome varie de 1,4 million (Sachidanandam et al., 2001) à 10 millions (Botstein and Risch, 2003) chez l'homme. Chez le porc, pour l'heure, 2,5 millions de SNP ont été répertoriés (Groenen et al., 2012). Le taux de mutation des SNP est faible, il est estimé à  $10^{-8}$  par génération chez l'homme (Kondrashov, 2003). Leur abondance et la stabilité en font des marqueurs de choix pour les études d'association en tout génome et des analyses de structure et d'évolution des génomes.

Bien que les SNP soient répartis sur l'ensemble du génome (1 SNP tous les kb environ (Erichsen and Chanock, 2004)), ils sont plus fréquemment retrouvés dans les régions non codantes des gènes, et les régions intergéniques (Barreiro et al., 2008). Leur utilisation pour l'étude des traces de sélection a apporté des informations précieuses sur l'histoire évolutive des espèces et des gènes soumis à sélection (Ai et al., 2013; Wilkinson et al., 2013).

Le génotypage de milliers de SNP est aujourd'hui possible en quelques jours, avec le développement des puces de génotypage haute densité. Développées chez de nombreuses espèces, elles permettent de génotyper de quelques milliers de SNP à 4,3 millions chez l'homme. Chez le porc, la puce commercialisée par Illumina® a été réalisée dans le cadre du consortium international *Porcine SNP Chip*, et comprend 60 000 marqueurs.

Les SNP de la puce sont issus de deux sources de données : d'une part, une recherche de SNP *de novo* à partir de banques de séquences de quatre races de porcs (Large White, Landrace, Piétrain, Duroc) et d'un sanglier, et d'autre part de SNP disponibles dans les banques de données publiques. Par séquençage *de novo*, 372K SNP ont été identifiés, et 549K SNP au total furent testés pour la réalisation d'une puce commerciale. Les marqueurs ont été sélectionnés pour leur informativité multirace, la fréquence de l'allèle minoritaire (MAF pour *Minor Allele Frequency* supérieure à 5%) et leur qualité d'alignement sur la séquence de référence (Ramos et al., 2009). Plus de 60.000 marqueurs furent finalement choisis pour assurer une couverture homogène du génome (Fan et al., 2010) et génotypés chez 158 individus.

La puce porcine a été validée pour sept races porcines d'intérêt économique : la MAF moyenne observée est de 0,29, et plus de 55.000 marqueurs ont une MAF supérieure à 0,05.



La puce est donc très informative, mais les allèles dont la fréquence dans la population est faible (inférieure à 0,05%) peuvent ne pas être détectés.

Grâce aux puces haute densité, il est aujourd'hui possible d'obtenir en quelques jours le génotype de plusieurs individus pour des dizaines de milliers de marqueurs. Ce génotypage en tout génome a de nombreuses applications, comme le diagnostic ou le contrôle d'apparentés des animaux. Il permet également la recherche de loci associés à des caractères. Ces études sont appelées GWAS (pour *Genome Wide Association Study*) et seront détaillées dans la partie suivante.

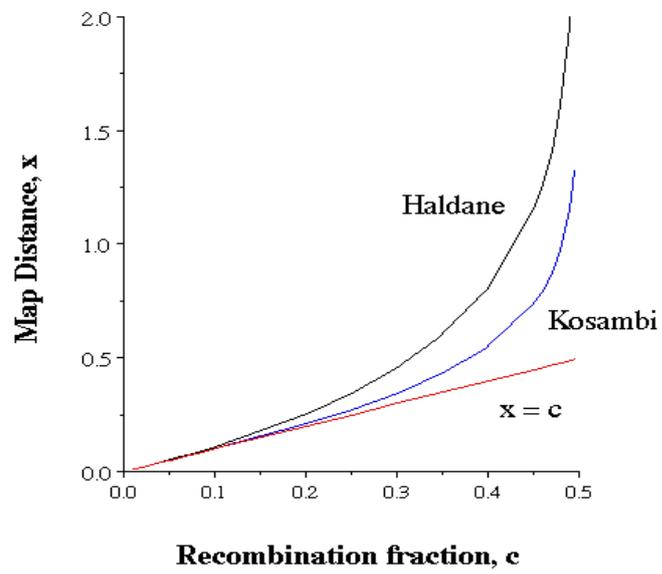


Figure 20 : Comparaison des fonctions de Haldane et Kosambi

## II.C La cartographie génétique

### II.C.1. La liaison génétique

L'étude de la liaison entre marqueurs moléculaires au sein d'une famille permet de construire des cartes génétiques. Plus deux loci sont proches l'un de l'autre, plus la probabilité qu'un événement de recombinaison au cours de la méiose survienne entre eux est faible. Morgan est le premier à proposer une estimation de la distance entre deux loci sur un chromosome en fonction de leur probabilité de recombinaison : 1% de recombinaison entre deux loci correspond alors à une distance génétique de 1 cM (Sturtevant, 1913). Deux loci non liés (très éloignés sur un chromosome ou présents sur deux chromosomes différents) ont une chance sur deux de recombiner, soit un taux de recombinaison de 50%, qui correspond à une distance de 50cM (Figure 20 courbe rouge). Ce modèle suppose que les crossing-overs sont indépendants. Cependant, la relation entre le taux de recombinaison et la distance génétique n'est plus linéaire lorsque les fréquences sont supérieures ou égales à 10%. Quand deux loci sont très éloignés, la probabilité d'avoir plusieurs crossing-overs augmente proportionnellement au carré de la distance. Afin de corriger ce biais et de définir une unité de distance additive, des fonctions de cartographie ont été proposées. La fonction de Haldane considère une indépendance entre crossing-overs multiples sur un même chromosome. La fonction de Kosambi cherche à prendre en compte l'existence de phénomènes d'interférences : la présence d'un crossing-over dans une région diminue l'occurrence d'un crossing-over dans une région adjacente (Haldane, 1919; Kosambi, 1944) (Figure 20). Les distances (en Morgans) sont calculées ainsi :

Haldane :

$$d = -\frac{1}{2} \log(1 - 2r) \quad \forall r \in [0, 1/2[$$

$r$  la probabilité de recombinaison entre deux loci distants de  $d$ .

Kosambi :

$$d = \frac{1}{4} \log \left( \frac{1 + 2r}{1 - 2r} \right) \quad \forall r \in [0, 1/2[$$

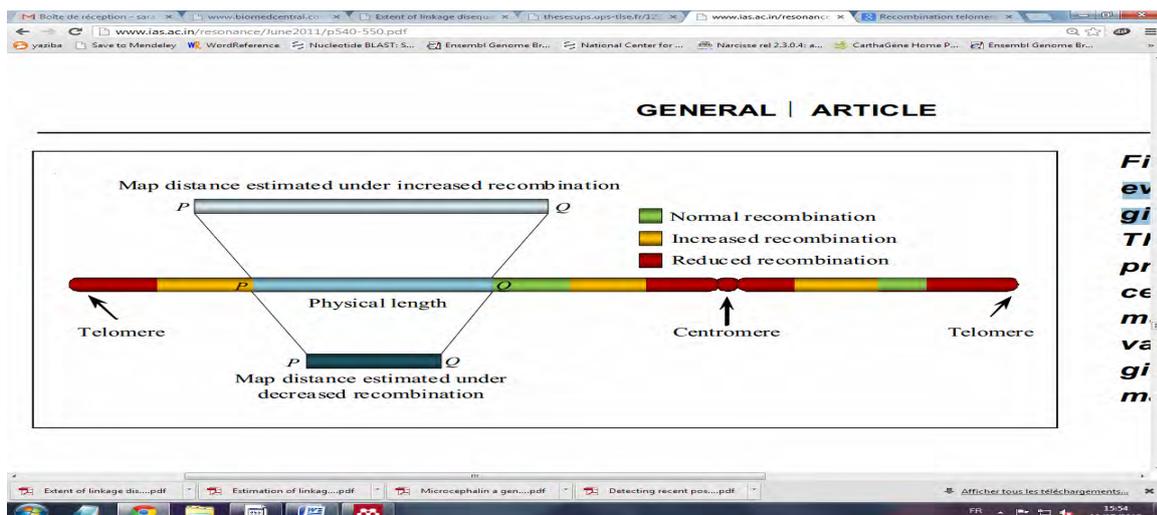


Figure 21 : Distance génétique en fonction du taux de recombinaison

Chez les mammifères, les premières cartes génétiques ont été réalisées chez la souris en 1915 (Haldane et al., 1915), et chez l'homme en 1932 (Wiener, 1932).

Chez le porc, la première carte génétique date des années 60, et a été réalisée avec les seuls marqueurs moléculaires disponibles à l'époque : les variants protéiques des groupes sanguins (Andresen, 1964). Le développement des marqueurs microsatellites, ainsi que les efforts groupés dans le cadre de différents projets internationaux (consortium PiGMap (Archibald et al., 1995), et de travaux suédois (Andersson et al., 1994; Marklund et al., 1993) et américains (Rohrer et al., 1996)), ont permis la localisation de 1500 marqueurs. Suite à la publication des cartes dans les années 90, des marqueurs génétiques ont été progressivement ajoutés (Rothschild, 2004; Vingborg et al., 2009). Aujourd'hui, plus de 5000 loci, et des centaines de gènes sont positionnées sur la carte génétique porcine ([www.thearkdb.org](http://www.thearkdb.org)).

En 2012, les données de séquençage du porc et de la puce porcine 60KSNP ont rendu possible la construction d'une carte génétique haute densité (Tortereau et al., 2012), composée de 38599 SNP recouvrant les autosomes porcins. Le ratio entre les tailles génétique et physique du génome du porc est de 0,7 à 0,8 cM/Mb (Tortereau et al., 2012). Cette valeur diffère entre les différentes espèces : en moyenne, 1 cM est égal à 1 Mb chez l'homme (Schneider et al., 2002), 1,25 cM pour 1 Mb chez le bovin (Arias et al., 2009) et 0,6 cM/Mb pour la souris (Shifman et al., 2006).

Il existe des variations du taux de recombinaison au sein même d'un génome en fonction des chromosomes (ou régions chromosomiques) (Figure 21) ou du sexe. Les régions centromériques et télomériques présentent par exemple un faible taux de recombinaison par rapport à la moyenne du génome. Les régions où le taux de recombinaison est élevé sont appelées points chauds de recombinaison (ou *hotspots*), et points froids (ou *coldspots*) lorsqu'il est faible. Il a été montré que des corrélations entre le taux de recombinaison et le contenu en séquence existaient, notamment chez les mammifères. La corrélation positive la plus significative concerne le pourcentage en GC (Groenen et al., 2009; Kong et al., 2002; Shifman et al., 2006; Wong et al., 2010). Des corrélations existent également entre la densité d'éléments transposables et le taux de recombinaison (chez l'homme : corrélation de 0,29 pour les éléments SINE/MIR et DNA/hAT-Charlie, et de -0,2 pour les LINE/L1 (Lee et al., 2011)). Enfin, comme chez d'autres espèces une différence de taille de carte génétique entre le sexe hétérogamétique et le sexe homogamétique a également été observée chez le porc.



La carte génétique femelle est plus importante que la carte mâle (21,5 Morgans vs 16,5 Morgans), excepté pour le chromosome 1.

Les analyses de liaison sont basées sur la comptabilisation des évènements de recombinaisons méiotiques, au sein d'une ou de deux générations de ségrégation familiale. Ce dispositif est donc limitant. En 1986, Bodmer puis Xiong en 1997 (Xiong and Guo, 1997) ont proposé une approche de cartographie basée sur le nombre de méioses au sein d'une population. Ces cartes utilisent alors l'association non aléatoire entre allèles observée dans une population : on parle alors de déséquilibre de liaison ou DL.

### **II.C.2. Le déséquilibre de liaison chez le porc**

Le déséquilibre de liaison dans une population, ou déséquilibre de phase gamétique, se définit comme l'association non aléatoire d'allèles à des locus différents. Il est créé par différentes forces évolutives dont, entre autres, la sélection, les mutations, le mélange des populations, la dérive génétique et les goulets d'étranglement. Plusieurs mesures du déséquilibre de liaison ont été proposées.

Soit deux loci présentant chacun 2 allèles (A/a et B/b) :

$D = f(AB) - f(A)*f(B)$  avec  $f(x)$  la fréquence de l'allèle x

$D'$  est le rapport entre D et sa valeur maximale  $D_{max}$  :

$D' = D/D_{max}$  avec  $D_{max} = f(A)*f(B)$

$D' = 1$  indique un déséquilibre de liaison complet entre deux allèles. Enfin, le coefficient  $r^2$  permet d'estimer la corrélation entre deux allèles :

$r^2 = (D')^2 * [f(a)*f(B)/f(A)*f(b)]$

Chez l'homme, les études récentes ont montré des traces de sélection positive impliquant des gènes liés à la malaria (Sabeti et al., 2002), à la digestion du lactose (Bersaglieri et al., 2004) ou au développement cérébral (Evans et al., 2005). L'exploitation du DL dans le cadre de travaux de clonage positionnel est donc devenue une priorité pour différentes espèces. Au cours des dernières années, les travaux d'estimation de la structure du DL chez les différentes espèces d'animaux domestiques ont fortement bénéficié du développement des outils de génotypage haut débit.

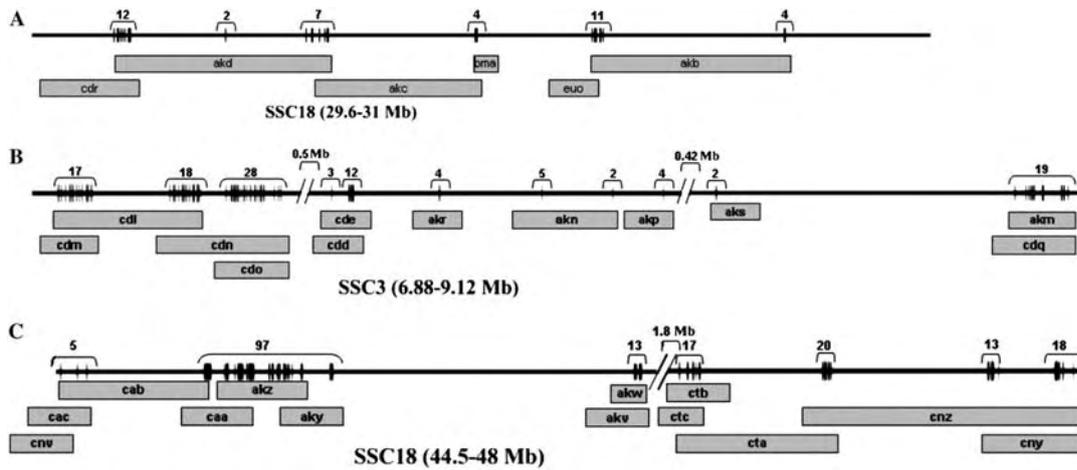


Figure 22 : Régions chromosomiques étudiées par Amaral et al et répartition des SNP. Les positions des SNP sont indiquées par des traits verticaux, et les nombres indiquent le nombre total de SNP par région. Les barres horizontales grises indiquent les BAC. Taille de région A : 0.5 Mb, région B : 3 Mb, région C : 3.4 Mb.

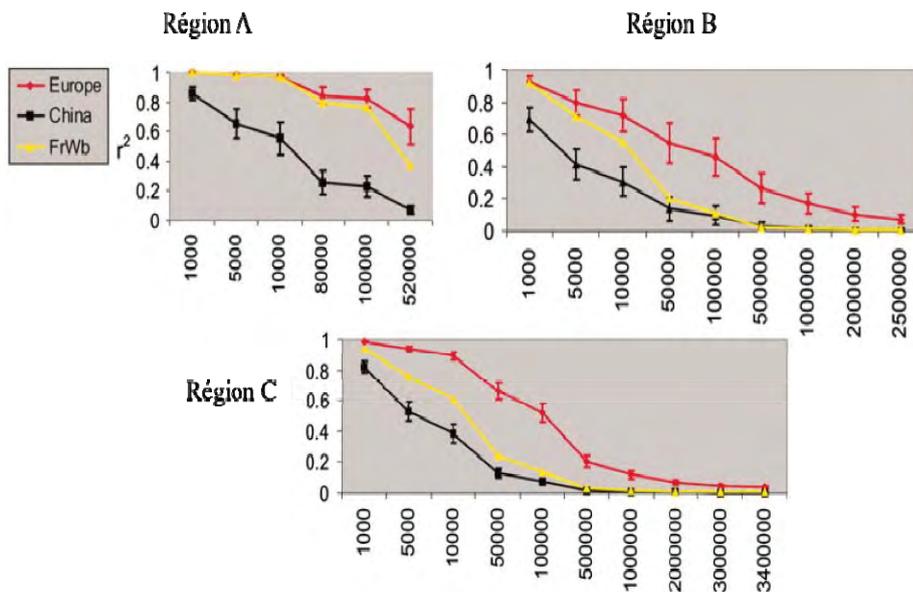


Figure 23 : DL en fonction de la distance génomique (pb) par population et par région étudiée

Chez le porc, plusieurs études ont été réalisées pour estimer le DL. Les premiers travaux ont été faits avec des microsatellites (Harmegnies et al., 2006; Nsengimana et al., 2004). Nsengimana et al ont montré (1) la présence d'un DL non nul à grande distance (jusqu'à 40 cM, voire plus) et même entre loci non synténiques et (2) une décroissance rapide du niveau de DL avec l'augmentation de la distance entre les marqueurs : Le  $r^2$  varie entre 0.15 et 0.50 pour les marqueurs distants de moins de 1 cM, et diminue rapidement lorsque les marqueurs sont plus distants.

D'autres estimations ont été réalisés à partir de marqueurs SNP (Ai et al., 2013; Amaral et al., 2008; Du et al., 2007a; Vingborg et al., 2009). Les résultats des travaux d'Amaral et al sont présentés en Figure 22 et Figure 23. Ils portent sur l'étude du DL et la diversité haplotypique chez différentes races chinoises et européennes, à l'aide de 371 SNP localisés sur 3 régions chromosomiques (Figure 22). Ils illustrent bien la structure du DL observée chez le porc. Le DL est très différent entre les races européennes (2 cM) et chinoises (0,05 cM) (Figure 23). La taille des blocks d'haplotypes peut atteindre 400 kb chez les races européennes, alors qu'elle n'excède généralement pas 10 kb chez les races chinoises. Les races chinoises ont une diversité haplotypique importante, et les haplotypes les plus fréquents sont également retrouvés chez les Large White, Landrace, et Duroc. Ces données ont été confirmées dans le cadre d'études réalisées à partir de la puce 60K chez des races de porcs finlandaises (Uimari and Tapio, 2011) et américaines (Badke et al., 2012).

Outre les études de diversité génétique et de l'histoire évolutive des races, les concepts de liaison et de déséquilibre de liaison ont été utilisés pour l'identification de mutation responsable d'un caractère phénotypique. En effet, lorsqu'un marqueur moléculaire est en DL fort avec une mutation causale (à l'échelle de la population ou recréé au sein de familles), les allèles de deux loci co-ségrègent ensemble chez des individus atteints. L'utilisation des marqueurs moléculaires a permis une recherche sans a priori, et est devenue, depuis les années 90, l'alternative à l'étude de gènes candidats utilisée initialement en cartographie génétique.

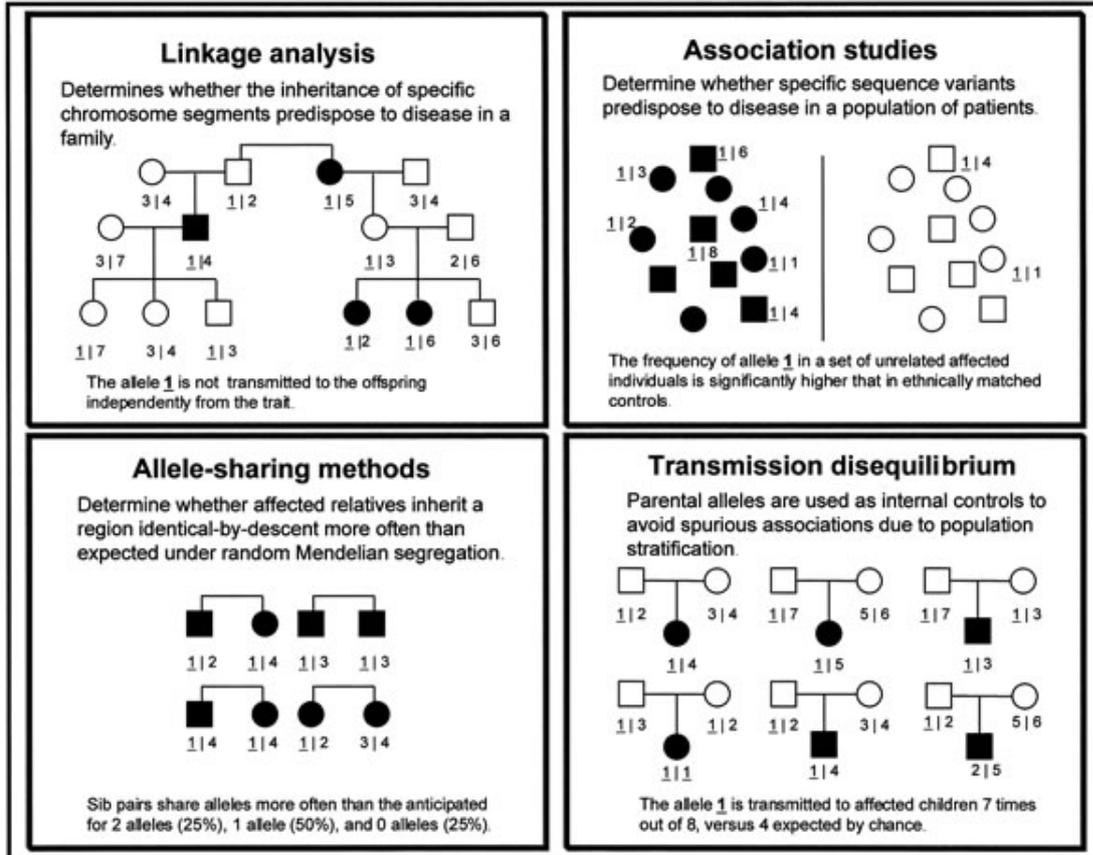


Figure 24 : Différentes approches de cartographie génétique utilisées dans la recherche de gènes à effet quantitatif.

### II.C.3. Recherche de mutations causales

La recherche de mutation responsable d'un phénotype a d'abord été réalisée par des approches gène candidat, c'est-à-dire avec des a priori sur les mécanismes moléculaires impliqués dans la maladie. Les « pistes » moléculaires peuvent par exemple être une protéine défectueuse connue, d'autres gènes de la même famille associés à des phénotypes voisins, ou un gène associé à la maladie chez d'autres espèces. Cependant, trop de pistes biologiques potentielles existent pour un grand nombre de caractères. Les approches de clonage positionnel ont permis de palier cette contrainte: il s'agit d'identifier des allèles de marqueurs qui ségrégent préférentiellement chez les individus atteints. La position connue de ces marqueurs moléculaires permet d'identifier les régions chromosomiques impliquées dans les caractères d'intérêt.

Le développement des cartes de microsatellites a été une aide précieuse dans ces études. Cependant, leur densité relativement faible à l'échelle du génome nécessitait (1) la mise en place de dispositifs familiaux dans lesquels ségrège le caractère d'intérêt et (2) la réalisation de travaux de cartographie plus fine des régions mises en évidence. Aujourd'hui l'utilisation des puces de génotypage haute densité (1) permet la mise en place de dispositif populationnel et (2) offre une résolution qui permet de s'acquitter d'une partie des étapes de cartographie (Figure 24).

#### *Etudes familiales :*

L'étude de liaison permet d'identifier des marqueurs qui ségrégent avec un phénotype au sein d'une famille. Il s'agit d'évaluer la probabilité globale des familles sous deux hypothèses :  $H_0$  : le marqueur n'est pas lié au locus responsable de la maladie (hypothèse d'indépendance avec  $\theta = 0,5$ ),  $H_1$  : ils sont liés (hypothèse de liaison avec une valeur de vraisemblance à maximiser) (Read and Donnai, 2008). Le rapport arithmétique (*odd ratio*) entre les 2 vraisemblances donne une information sur la probabilité d'une liaison. En cas de liaison, le taux de recombinaison le plus probable entre le marqueur et le locus responsable est celui qui maximise la valeur de ce rapport. Le logarithme du rapport de vraisemblance est le *lod score* (*logarithm of odds score*) (Ott, 1999). Un *lod score* de 3 signifie que l'hypothèse d'une liaison est 1000 fois plus vraisemblable que l'indépendance entre marqueur et gène.

Le premier gène majeur identifié chez le porc est le gène de la sensibilité à l'halothane, *Hal*, responsable de l'hyperthermie maligne (Fujii et al., 1991).

<b>Caractère</b>	<b>Gène</b>	<b>Auteurs, date</b>
<b>Susceptibilité au stress</b>	RYR	(Fujii et al., 1991)
<b>Couleur de la peau</b>	cKIT	(Marklund et al., 1998)
<b>Qualité de la viande</b>	RYR	(Fujii et al., 1991)
	PRKAG3	(Milan et al., 2000b)
<b>Prise alimentaire</b>	MC4R	(Kim et al., 2000)
<b>Composition et croissance</b>	MC4R	(Kim et al., 2000)
	IGF2	(Van Laere et al., 2003)
	CAST	(Ciobanu et al., 2004)
<b>Résistance aux maladies</b>	F18	(Meijerink et al., 2000)
	K88	(Jørgensen et al., 2003)
<b>Reproduction</b>	ESR	(Rothschild et al., 1996)
	PRLR	(Vincent et al., 1998)
	RBP4	(Rothschild et al., 2000)

Tableau 7 : Gènes utilisés en sélection porcine (Dekkers, 2004)

Les progrès en matière de cartographie génétique et de panels de marqueurs microsatellites ont ensuite permis de localiser le gène RN, locus à effet majeur sur la qualité de la viande, puis d'identifier le gène sous-jacent PRKAG3 (Milan et al., 2000b).

Certains caractères sont contrôlés par de nombreuses régions chromosomiques, chacune d'entre elles ayant un effet moyen à faible : on parle de QTL (*Quantitative Trait Loci*). Le premier QTL identifié par analyse de liaison chez le cochon concerne la quantité de gras dorsal et abdominal. Ce locus est localisé sur le chromosome 4 porcine (Andersson et al., 1994). A l'heure actuelle la mutation causale de ce gène n'a pas été identifiée. Aujourd'hui, ce sont 8014 QTL qui ont été primo-localisés et sont répertoriés dans la base de données PigQTLdb (Hu et al., 2013). Ces QTL concernent principalement la qualité de la viande et de la carcasse (5406 QTL identifiés), la santé des animaux (784), la reproduction (782), les caractéristiques externes (764) et la production (675). Pour la plupart de ces QTL, peu de mutations ont été caractérisées. Cependant, parmi les quelques mutations identifiées, certaines sont utilisées en sélection assistée par marqueurs (Tableau 7).

La majorité de ces loci a été cartographiée à partir de grands dispositifs familiaux. Depuis la commercialisation de la puce 60K, des variants génétiques qui ségrégent dans une ou plusieurs populations sont recherchés à partir de dispositifs populationnels.

### *Etudes de population*

Les études d'association sont basées sur la comparaison des fréquences alléliques entre deux populations. Dans le cas d'une maladie, on recherche si certains allèles ségrégent plus que d'autres dans une population de sujets atteints par rapport à une population de sujets sains. Le test statistique du KHI2 est utilisé pour savoir si les deux fréquences sont significativement différentes. La méthode qui consiste à comparer des sujets atteints et sains est le dispositif cas/contrôles. Il est facile à mettre en place et le test est puissant. Cependant, des effets de stratification peuvent induire des faux positifs, lorsque le fond génétique des deux groupes étudiés est différent, et ce indépendamment du statut à la maladie. Des outils statistiques existent désormais pour corriger ces biais (Zhou et al., 2013).

D'autres stratégies permettent de pallier au biais de stratification en utilisant des familles. Le test TDT (*Transmission Disequilibrium Test*) est le plus connu et permet de combiner, au sein d'une même analyse, un test de liaison et un test d'association. Ce test permet de mettre en évidence une distorsion de ségrégation entre des parents hétérozygotes et des descendants atteints.

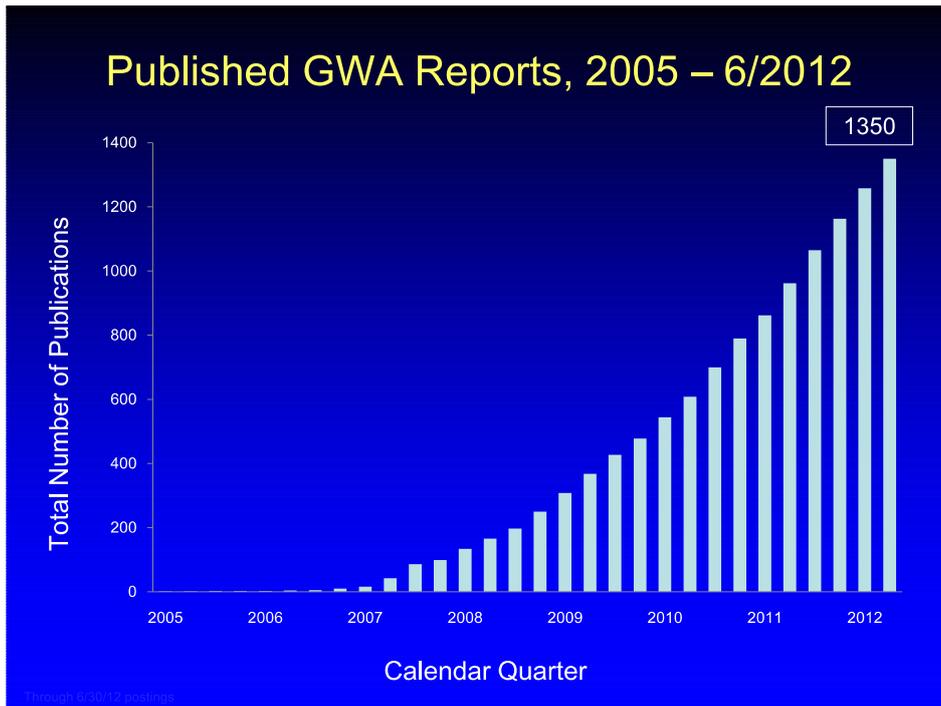


Figure 25 : Nombre de GWAS depuis 2005

L'inconvénient du TDT est la perte de puissance due au fait que seuls les parents hétérozygotes sont utilisés. Cependant, ce test est extrêmement robuste, car le test de liaison permet de s'abstraire des problèmes de stratification. Plusieurs travaux chez le porc ont été réalisés selon cette approche (Du et al., 2009; Grindflek et al., 2006).

Depuis quelques années, les études d'association sont réalisées à l'échelle du génome. Elles sont nommées GWAS, pour *Genome Wide Association Study*. Ces études ont permis d'identifier de nombreux variants génétiques chez l'homme notamment, associés à des maladies pour lesquelles les approches gène-candidat n'avaient pas été concluantes. Les maladies concernées sont principalement multifactorielles (diabète de type 2, sclérose en plaques, maladie de Crohn, etc.) (Yeo, 2011).

La première étude publiée date de 2005, et concerne la DMLA (Dégénérescence maculaire liée à l'âge). Dans cette étude, deux SNP d'un facteur du système immunitaire inné furent identifiés par étude cas/contrôles (Klein et al., 2005). Depuis les GWAS se sont multipliées (Figure 25). Aujourd'hui, les 1350 études pan-génomiques sont rassemblées dans une base de données publique : <http://www.genome.gov/gwastudies/>. En parallèle, les loci et gènes identifiés sont en accès libre sur le site de l'OMIM (*Online Mendelian Inheritance in Man*) (<http://www.omim.org>). En Janvier 2013, 130 gènes connus associés à un phénotype étaient accessibles, et 1765 maladies étaient répertoriées. Dans la grande majorité des cas, seul le locus responsable a été identifié mais les mécanismes moléculaires sous-jacents restent inexpliqués.

Chez les animaux domestiques, les GWAS se sont développées à partir du moment où les premières séquences de référence des génomes ont été publiées et que le nombre de SNP répertoriés a été suffisant pour permettre le développement d'une puce de génotypage. Les études réalisées portent principalement sur des pathologies et des caractères d'intérêt économique (Zhang et al., 2012). Chez les bovins, la première étude publiée portait sur l'identification de trois gènes contenant les mutations causales de la dystonie musculaire congénitale (types 1 et 2) et de l'ichtyose fœtale (Charlier et al., 2008). Le chien est également une espèce chez laquelle plusieurs mutations ont été identifiées par des approches GWAS. L'histoire évolutive des races canines permet de disposer d'un large DL intra race et d'un faible DL lorsque plusieurs races présentant la même anomalie sont combinées au sein une même analyse (Shearin and Ostrander, 2010).



Chez le porc, les principales études portent sur la reproduction (Grindflek et al., 2011a; Onteru et al., 2011; Sironen et al., 2010a), les odeurs sexuelles mâles dans la viande (Duijvesteijn et al., 2010; Grindflek et al., 2011b; Ramos et al., 2011), les caractères de production (Fan et al., 2011; Ponsuksili et al., 2011), les paramètres sanguins (Luo et al., 2012) et la couleur de la peau (Ren et al., 2011). Plusieurs gènes ont ainsi été identifiés, tels que TYRP1 responsable de la couleur brune de races asiatiques (Ren et al., 2011), ou les gènes CHCHD3 et BMP2 associés à des caractères physiques (Fan et al., 2011).

Cependant, pour la plupart des études chez les différentes espèces, de nombreux SNP ont été identifiés mais peu de mécanismes moléculaires ont été caractérisés. Dans le cas de caractères polygéniques, les loci identifiés ne permettent en général d'expliquer qu'une part très faible de la variation phénotypique. De plus, beaucoup de discordances sont observées entre les résultats des différentes études portant sur un même caractère. Dans certains cas, ces différences peuvent être expliquées par la taille et des problèmes de stratification dans la structure des populations (Campbell et al., 2005). Enfin, des différences de méthodes de phénotypage peuvent induire un biais à l'identification des variants.

Le chapitre suivant fera l'état des lieux des connaissances physiopathologiques et génétiques des anomalies congénitales.

Conceptus		Embryonic development (weeks)						Fetal period (weeks)			
1	2	3	4	5	6	7	8	9	16	20-36	38
											
		<b>Neural</b>									
		<b>Heart</b>									
		<b>Upper limbs</b>									
		<b>Lower limbs</b>									
		<b>Ear</b>									
		<b>Eye</b>									
			<b>Palate</b>								
			<b>Teeth</b>								
				<b>External genitalia</b>							
<b>Loss</b>		<b>Major abnormalities</b>					<b>Functional and Minor abnormalities</b>				

Figure 26 : Etapes de l'organogénèse chez l'homme

### III. Physiopathologie de l'intersexualité, de la cryptorchidie et de la hernie scrotale chez l'homme et chez d'autres modèles

#### III.A. Introduction aux anomalies congénitales humaines

Les anomalies congénitales sont définies comme des défauts de structure ou de fonction présents à la naissance (OMS, 2012).

Ces défauts surviennent principalement entre la 3<sup>e</sup> et la 8<sup>e</sup> semaine suivant la conception. L'embryon est particulièrement vulnérable à cette période de la vie embryonnaire (Schlotz & Phillips, 2009). On estime en effet que 90% des 4500 structures de l'organisme se mettent en place à cette période (Irving, 2006) (Figure 26). Avant la 3<sup>ème</sup> semaine, un défaut pré-embryonnaire entraînera un avortement spontané s'il n'est pas réparé par les cellules embryonnaires (Jacquet, 2004; Schull and Otake, 1999). Après la 8<sup>ème</sup> semaine, la majorité des organes est formée, l'apparition d'anomalies est donc moins fréquente. Cependant, certains organes dont la formation est plus tardive, comme le cerveau, restent sensibles après la période embryonnaire, durant la vie fœtale (Akhtar and Karim, 2012).

Littéralement, le terme congénital vient du latin cum : avec, et genitus: né de. Il signifie « présent à la naissance », et non héréditaire. Ainsi, les causes des anomalies congénitales peuvent être environnementales, génétiques ou multifactorielles. Bien que la moitié des défauts de naissance n'ait pas d'origine connue, certains facteurs ont été identifiés (Emile et al., 2012).

Les aberrations chromosomiques concernent 10 à 15% des causes identifiées, les défauts génétiques à transmission mendélienne 2 à 10%. Actuellement, 5000 maladies génétiques environ sont recensées, et plus de 1000 gènes ont été identifiés (OMIM : *Online Mendelian Inheritance in Man* - <http://www.omim.org>).

Un dixième des anomalies congénitales est dû à des facteurs environnementaux, tels que les infections et pathologies maternelles, ou les substances ou rayonnements auxquels la mère serait soumise durant sa grossesse. Les infections maternelles et placentaires concernent 2 à 3% des causes et sont principalement la rubéole, la toxoplasmose, la syphilis, le cytomégalovirus, et le HIV. Les pathologies maternelles comme le diabète, la phénylcétonurie et les endocrinopathies sont responsables de 6-8 % des défauts congénitaux.

Anomalies	Fréquence (par nombre de Symptômes naissances)		Causes
Cardiopathies congénitales	1/125-150	Défaut de formation du système cardiaque, difficultés à respirer, peau bleutée	Génétiques ou environnementales, parfois associées à une médication ou une pathologie chez la mère
Autisme	2-6/1000 (3-4 fois plus chez les garçons)	Interaction sociale anormale, mouvements répétés, troubles d'apprentissage	Peu connues, causes génétiques ou environnementales responsables du développement cérébral anormal
Surdité congénitale	3/1000	Absence de réaction aux bruits, parfois difficile à détecter chez l'enfant	Génétique ou infections maternelles durant la grossesse
Drépanocytose	1/500 + (populations africaines subsahariennes)	Douleurs, infections, infarctus, troubles de la vision, croissance ralentie	Génétique, hémoglobine anormale responsable d'hématies falciformes qui bouchent les vaisseaux sanguins
Syndrome de Down	1 /800-1000	Variables, incluant un retard mental, faciès caractéristique, cardiopathies, infections, troubles de la vue et de l'audition	Chromosome 21 surnuméraire
Fente labio-palatine	Fente labiale : 1/1000 Fente palatine : 1/2000	Fissure de la lèvre et du palais	Mal connues, génétiques ou environnementales
Spina bifida	1/2000	Allant de l'absence de symptômes aux malformations de la colonne vertébrale, infections, paralysies et incontinences urinaires	Multifactorielles, responsables d'anomalies de fermeture du tube neural

Tableau 8 : Anomalies congénitales les plus fréquentes chez l'homme

Enfin, 1% des défauts sont dus à l'irradiation, et 1% à des substances tels que l'alcool, l'antagoniste de l'acide folique, les androgènes, la phénytoïne, le thalidomide, la warfarine et l'acide 13-cis-rétinoïque.

Malgré ces facteurs connus, la plupart des anomalies congénitales est d'origine inconnue (40 à 60% des défauts), ou multi factorielle (20 à 25 %). Par exemple, la fente labio-palatine et le spina-bifida, qui sont pourtant des maladies fréquentes, demeurent de cause inconnue. D'autres défauts, comme la luxation congénitale de hanche, associent à la fois un déterminisme génétique (Jiang et al., 2003; Rouault et al., 2010) et des facteurs environnementaux (Emile et al., 2012). Le Tableau 8 présente certaines causes identifiées comme les défauts de naissance les plus fréquents.

La fréquence globale des malformations congénitales est difficile à estimer car le défaut peut être détecté très tardivement. On estime que 3% des nouveau-nés présentent un défaut à la naissance (Canfield et al., 2006), soit près de 4 millions d'enfants atteints sur les 130 millions de naissances annuelles dans le monde. Le degré de sévérité est très variable, allant de la déformation sans gravité au décès pouvant survenir avant ou après la naissance. Globalement, 14,5% des enfants atteints de maladie congénitale (580 000 enfants) n'atteignent pas l'âge de 20 ans (Tennant et al., 2010).

Les principaux défauts congénitaux sont présentés dans le Tableau 8 (Cheung, 2005). Les défauts les plus fréquents sont les cardiopathies congénitales, touchant 1 enfant sur 125 à 150, et l'autisme, notamment chez le garçon (près d'un garçon atteint sur 150). Le syndrome de Down, ou trisomie 21 est relativement fréquent, car il représente une naissance sur 800. Il est à noter que 92% des grossesses sont interrompues lorsque la maladie est diagnostiquée chez le fœtus (Natoli et al., 2012).

L'Organisation Mondiale de la Santé a établi une classification codifiée des maladies intitulée « Classification statistique internationale des Maladies et des Problèmes de Santé connexes » ou CIM, composée de 22 catégories de maladies et affections (OMS, 2009). Les anomalies congénitales sont répertoriées dans la catégorie « Malformations congénitales et anomalies chromosomiques » (Annexe 1). Cette catégorie est composée de plus de 600 défauts présents à la naissance, et subdivisée en 11 sous-catégories, classées selon les systèmes physiologiques (systèmes nerveux, circulatoire, respiratoire, digestif, urinaire et musculo-squelettique) ou les organes impliqués (œil, Oreille, visage et cou, organes génitaux).

### **DSD liés aux chromosomes sexuels**

---

45,X

syndrome de Turner et variants

47,XXY

syndrome de Klinefelter et variants

46 XX

dysgénésie gonadique complète

45,X/46,XX

dysgénésie gonadique mixte

46,XX/46,XY

DSD ovotestis chimérique

### **DSD avec caryotype 46,XY**

---

DSD testiculaire

Dysgénésie gonadique complète (Syndrome de Swyer)

Dysgénésie gonadique partielle

Régression gonadique

DSD ovotestis

Anomalies de la synthèse ou de l'action des androgènes

Déficit en biosynthèse (déficience en 17 $\beta$ -HSD III ou 5 $\alpha$ RD2, mutations de StAR)

Défaut de l'action des androgènes

Défauts du récepteur de l'hormone lutéinisante (aplasie, hypoplasie des cellules de Leydig)

Anomalies de l'hormone antimüllérienne ou de son récepteur

### **DSD avec caryotype 46,XX**

---

DSD ovarien

Dysgénésie gonadique (Sry, Sf1, Wt1, Sox9,...)

DSD testiculaire (Sry+, duplication de Sox9)

Régression gonadique

DSD ovotestis

Excès d'androgènes

d'origine fœtale (21-hydroxylase, 11-hydroxylase, 3  $\beta$ -HSD)

d'origine fœtoplacentaire (déficit en aromatasase, anomalies de POR=P450 oxidoréductase)

d'origine maternelle (tumeurs virilisantes (lutéome), iatrogènes,..)

Tableau 9 : Nouvelle classification des désordres du développement sexuel

La dernière sous-catégorie correspond aux anomalies chromosomiques : trisomie 21, trisomies 18 et 13, autres trisomies des autosomes, monosomies et délétions des autosomes, réarrangements équilibrés, syndrome de Turner (45, X), anomalies des chromosomes sexuels avec phénotype féminin, des chromosomes sexuels avec phénotype masculin, ou anomalies chromosomiques non classées ailleurs.

Les trois anomalies congénitales étudiées dans cette thèse sont des défauts du développement sexuel. Leur classification selon la CIM est la suivante :

- Le type d'intersexualité étudié chez le porc dans le cadre de cette thèse est classé dans la catégorie des anomalies congénitales des organes génitaux,
- La cryptorchidie, ou testicule non descendu, est également définie comme une anomalie congénitale des organes génitaux,
- La hernie inguino-scrotale est répertoriée dans la catégorie des maladies du système digestif.

Certains termes tels que « pseudohermaphrodisme » et « intersexualité » ont été sujets à controverse, perçus comme péjoratifs par les patients, et confus pour les praticiens (Lee et al., 2006). C'est pourquoi depuis quelques années, une nouvelle classification spécifique aux désordres du développement sexuel est en place : elles regroupent les DSD, pour *Disorders of Sexual development* (Tableau 9). Ces défauts peuvent être causés par une anomalie du nombre de chromosomes sexuels, comme le syndrome de Turner dû à la présence d'un seul chromosome X, ou le syndrome de Klinefelter, causé par la présence de trois chromosomes sexuels XXY. D'autres DSD sont dus à des défauts hormonaux, ou des défauts génétiques.

Selon cette nouvelle classification, le type d'intersexualité étudié dans ces travaux de thèse est défini comme DSD 46,XX - DSD ovarien avec DSD testiculaire ou DSD ovotestis.

Afin de bien comprendre la physiopathologie de ces maladies, il est essentiel de faire un rappel des "acteurs" connus de la différenciation sexuelle.

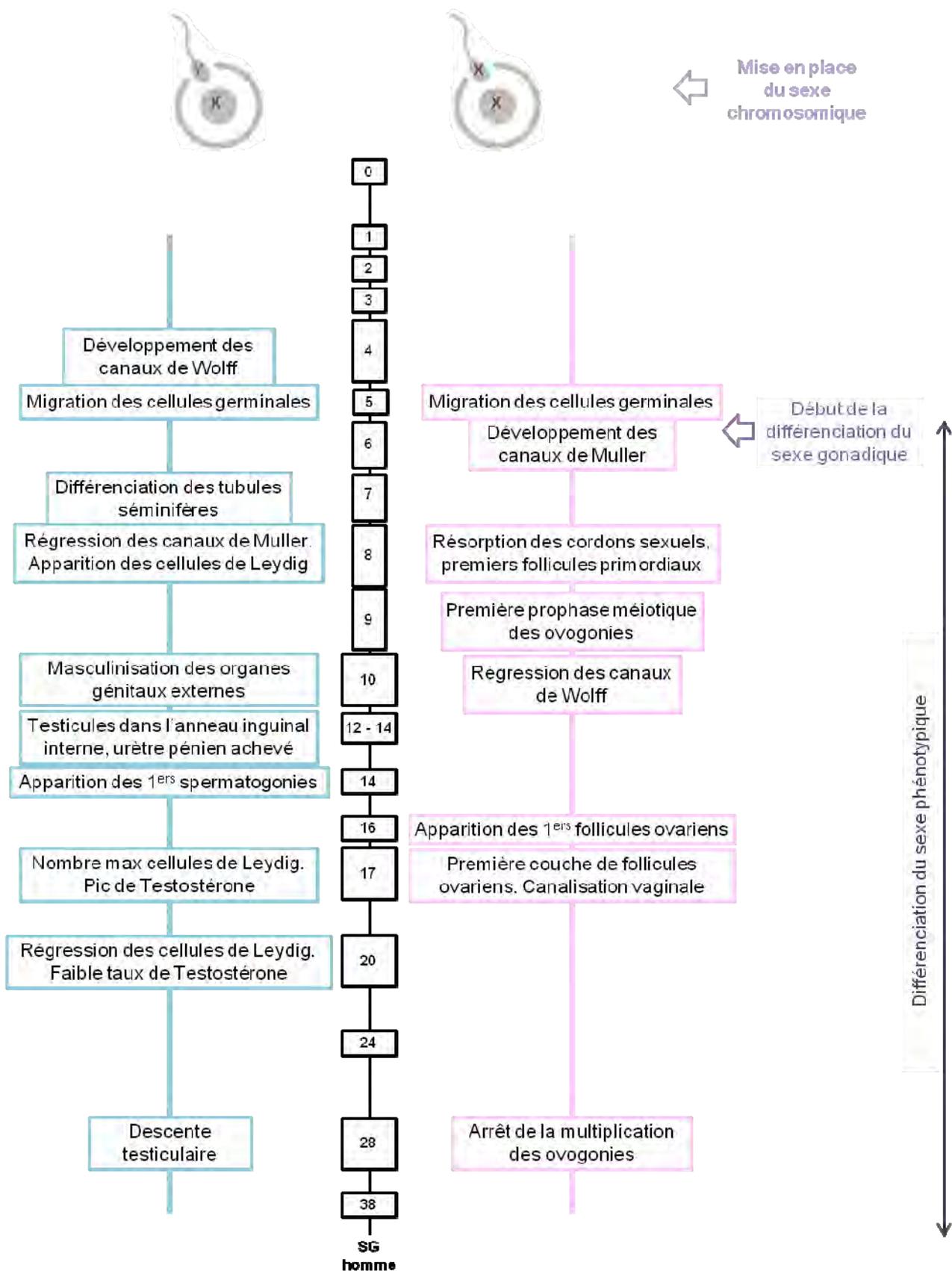


Figure 27 : Etapes de la différenciation sexuelle, de la fécondation à la 38<sup>ème</sup> semaine de la vie embryonnaire chez l'homme. Les semaines de gestation (notées SG en noir) sont indiquées chez l'homme.

### **III.B. Différenciation sexuelle chez les mammifères**

La différenciation sexuelle est un processus spatio-temporel complexe, impliquant des mécanismes génétiques, hormonaux et de différenciation tissulaire et cellulaire (Camerino et al., 2006; Öçal, 2011). Elle est comparable chez tous les mammifères, mais la durée de chaque étape varie, le temps de gestation étant propre à chaque espèce (à titre d'exemple, il est de 21 jours chez la souris contre 266 chez l'homme et 115 chez le porc). Le sexe est d'abord défini au niveau chromosomique lors la fécondation (en violet sur la Figure 27), puis au niveau gonadique (Sherwood, 2006) à partir de la 7<sup>ème</sup> semaine chez l'homme. La différenciation sexuelle se fait ensuite au niveau phénotypique, et comprend l'ensemble des caractères sexuels internes et externes. Elle se termine à la puberté, lorsque tous les caractères sexuels secondaires sont mis en place.

#### **III.B.1. Mise en place du sexe chromosomique :**

La première étape de la différenciation sexuelle est la transmission paternelle lors de la fécondation soit du chromosome Y, chez le mâle, soit du chromosome X, chez la femelle (de la Chapelle, 1987). En 1959, la présence du chromosome Y est identifiée comme le principal facteur qui détermine le sexe masculin chez l'homme et la souris, à l'inverse de la situation observée chez la drosophile, où la détermination du sexe est dépendante du dosage du chromosome X (Wilhelm et al., 2007). Le locus déterminant le sexe mâle ne fut identifié que trente années plus tard. Des travaux de cytogénétique permirent de mettre en évidence la présence du bras court du chromosome Y transloqué sur le chromosome X chez des mâles XX (Magenis et al., 1982). Des analyses d'hybridation montrèrent qu'un fragment du chromosome Y était absent chez certaines femmes XY (Affara et al., 1987). Plusieurs gènes localisés dans la région furent suspectés (MIC2, ZFY (Weissenbach and Rouyer, 1989)) puis exclus, car absents chez certains sujets masculins XX. La comparaison de séquence d'une femme au caryotype XY et de trois hommes au caryotype XX permit finalement d'identifier une région de 35kb contenant Sry (*Sex-determining Region of Y chromosome*), gène majeur de la différenciation sexuelle (Berta et al., 1990a; Jäger et al., 1990; Sinclair et al., 1990). En 1991, le transfert du fragment génomique murin de 14kb portant Sry chez une souris XX montra que Sry était suffisant pour le développement de testicules et de caractéristiques mâles (Koopman et al., 1991). Bien que le chromosome Y soit transmis lors de la fécondation, Sry

n'intervient qu'à la 7<sup>ème</sup> semaine de la vie embryonnaire chez l'homme, activant le développement du testicule chez le mâle.

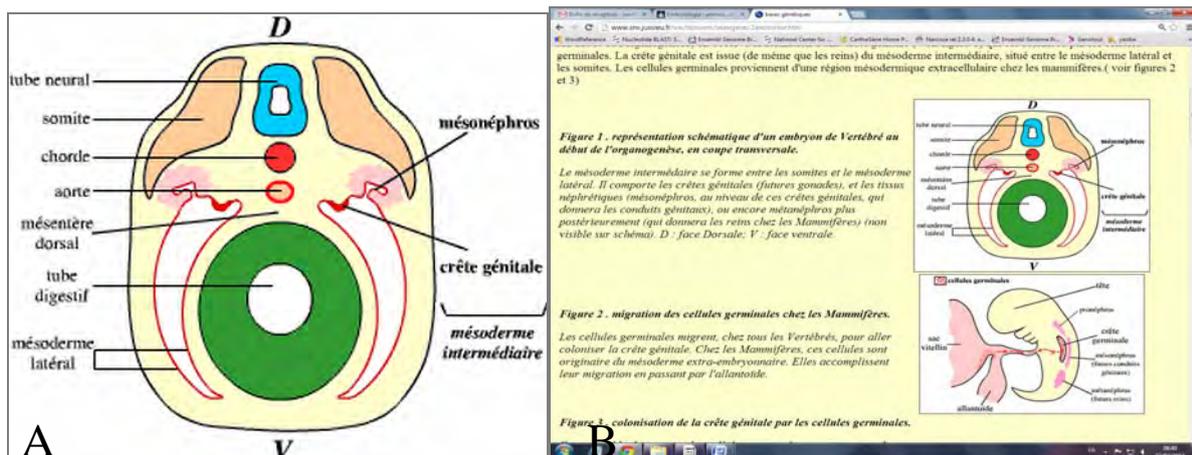


Figure 28: A : Ebauche des crêtes génitales B : Migration des cellules germinales.

### III.B.2. Déterminisme du sexe gonadique :

Le développement des gonades peut être divisé en deux phases. La première phase est caractérisée par le développement des gonades indifférenciées, identiques chez les mâles et les femelles. A ce stade, les lignées cellulaires ne sont pas encore prédestinées vers l'une ou l'autre des voies sexuelles. La deuxième phase, déclenchée par Sry, est le développement du testicule, ou de l'ovaire en absence de Sry. Des réseaux complexes de gènes sont impliqués dans les différentes étapes du développement gonadique, identiques durant la première phase, puis spécifiques de chaque lignée sexuelle durant la seconde phase. Les facteurs interviennent de manière directe ou par l'intermédiaire de voies de signalisation. Une véritable compétition entre les facteurs de la voie mâle et de la voie femelle s'opère jusqu'à la vie adulte.

#### *Le développement de la gonade bipotentielle :*

Vers la 4<sup>ème</sup> semaine après la fécondation chez l'homme, la crête génitale se développe sur la face interne du mésonephros par épaissement de l'épithélium coelomique (Figure 28. A). Les cellules germinales primordiales (Figure 28. B), issues du mésenchyme extracellulaire autour du diverticule allantoidien, viennent coloniser la crête génitale, puis se multiplient pour former l'ébauche génitale (McLaren, 2003).

La gonade précoce des mammifères est un primordium indifférencié composé de cellules précurseurs bipotentiels, capables de suivre l'une ou l'autre des voies sexuelles. Deux types cellulaires sont présents : les cellules germinales primordiales (CGP) et les cellules somatiques. Les CGP deviendront les gamètes mâles (spermatozoïdes) ou femelles (ovocytes). Les cellules somatiques ou nourricières viennent se positionner aux alentours des CGP, voire entourer les gamètes chez le mâle (Figure 29). Elles se différencieront en cellules de Sertoli (cellules de soutien) et de Leydig (cellules interstitielles) chez le mâle, ou en cellules folliculaires et thécales chez la femelle (McLaren, 1998).

Le développement de la gonade bipotentielle est programmé génétiquement (Eggers and Sinclair, 2012; Ono and Harley, 2013; Park and Jameson, 2005). Des études de cas cliniques chez l'homme (désordres du développement sexuel) et des expériences de transgénèse chez la souris ont permis d'identifier plusieurs gènes nécessaires au bon développement de la gonade indifférenciée. Les principales études portent donc sur des modèles murins et humains.

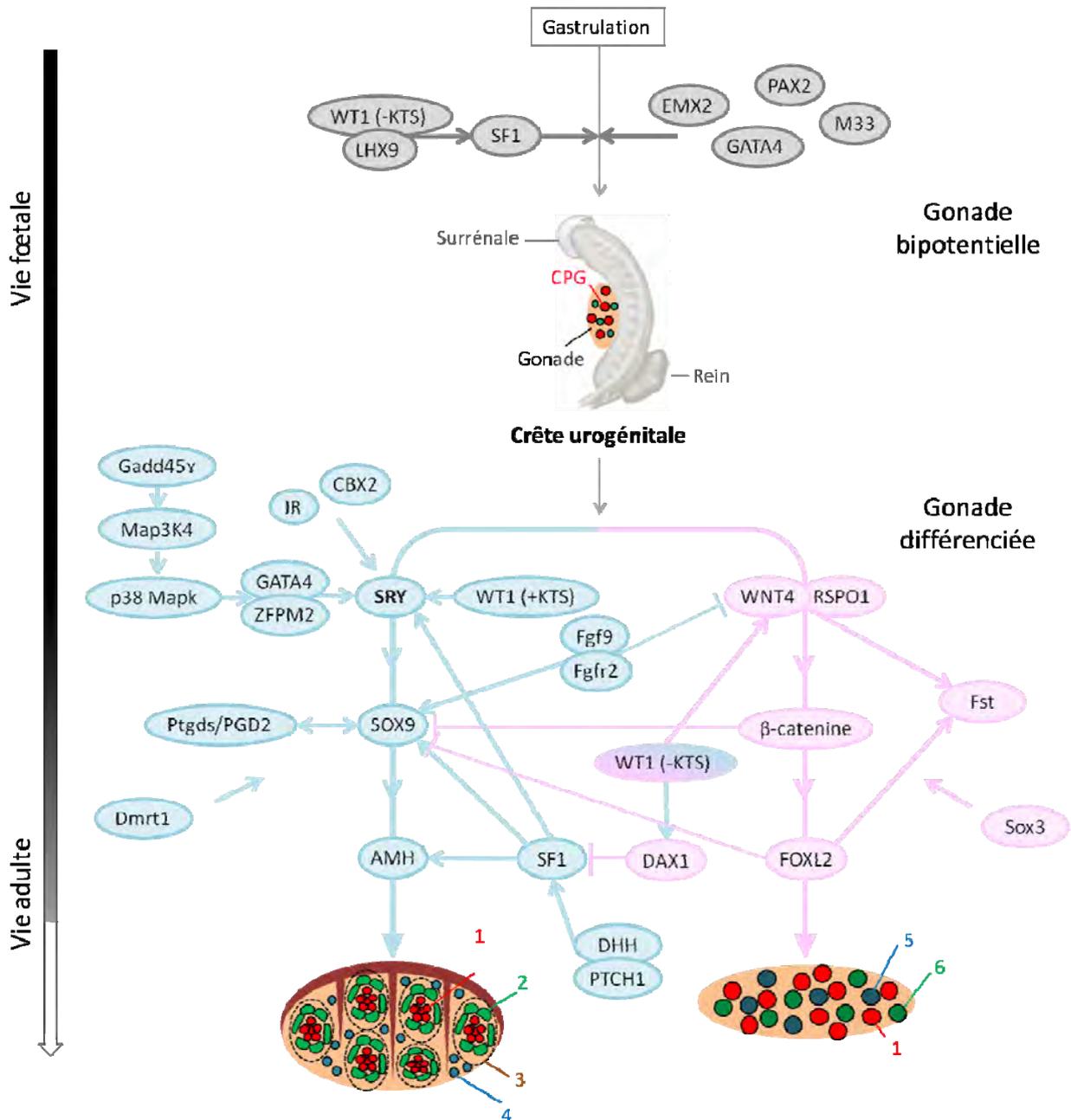


Figure 29 : Gènes impliqués dans la différenciation gonadique, de la vie foetale à la vie adulte.

Légende des gonades : CPG : cellules germinales primordiales ; 1 : cellules germinales (spermatozoïdes ou ovocytes); 2 : cellules de Sertoli; 3 : cellules péritubulaires myoïdes; 4 : cellules de Leydig; 5 : cellules thécales; 6 : follicules/cellules de granulosa.

Les gènes identifiés sont *Wt1* (*Wilms tumor protein*) (Wagner, 2003), *Sf1* (*Steroidogenic Factor 1*) (Chen et al., 2013; Val and Swain, 2010), *Lhx9* (*LIM homeobox 9*) (Birk et al., 2000), *Emx2* (*Empty spiracles homeobox 2*) (Miyamoto et al., 1997), *M33* (Katoh-Fukui et al., 2005), *Gata4* (Biason-Lauber, 2010) et *Pax2* (*Paired box gene 2*) (Torres et al., 1995) (Figure 29).

- **Wt1** code pour un facteur de transcription, qui possède un domaine de fixation à l'ADN composé de quatre structures en doigt de zinc. Il est exprimé précocement dans les tissus du mésoderme embryonnaire, tels que la crête urogénitale, les gonades et le mésonephros (Armstrong et al., 1993). Son invalidation chez la souris conduit à une absence de développement des reins, des surrénales et des gonades (Kreidberg et al., 1993).

Chez l'homme, WT1 fut d'abord identifié comme un anti-oncogène, car des mutations dans sa séquence furent observées chez des patients atteints de tumeur de Wilms, la forme de cancer du rein chez l'enfant la plus fréquente (95% des cas) (Davidoff, 2012). Cette tumeur fut ensuite associée au syndrome WAGR (*Wilms tumor, Aniridia, Genitourinary anomalies, mental Retardation*), un syndrome complexe incluant des caractéristiques cliniques comme des anomalies génito-urinaires, une absence totale d'iris, et un retard mental.

Deux isoformes de WT1, + KTS et -KTS, ont été particulièrement étudiés : ils diffèrent par la présence ou l'absence de trois acides aminés KTS, entre le troisième et le quatrième doigt de zinc du domaine de liaison à l'ADN. La présence du motif affecte la liaison de WT1 à l'ADN (Laity et al., 2000). Les deux isoformes semblent avoir des fonctions distinctes mais le ratio d'expression des deux isoformes est très important (Wagner, 2003).

- **Sf1** code pour un récepteur nucléaire, jouant un rôle dans le développement des gonades et des surrénales (Chen et al., 2013; Val and Swain, 2010). Chez la souris, l'invalidation de *Sf1* entraîne la régression des gonades des deux sexes au stade E12.5 (Luo et al., 1994) Il semble être un gène cible de *Lhx9* (Birk et al., 2000).

- **Lhx9** est impliqué chez la souris dans la prolifération des cellules somatiques de la crête génitale et la formation des gonades (Birk et al., 2000). Des analyses biochimiques in vitro ont montré que LHX9 pouvait potentiellement se fixer et activer le promoteur de *Sf1*, en coopération avec *Wt1* (Wilhelm and Englert, 2002).

- **Emx2** joue un rôle essentiel dans le développement du télencéphale dorsal chez la souris (Yoshida et al., 1997), et semble impliqué dans le développement urogénital (reins, urètres, gonades et tractus génitaux) (Miyamoto et al., 1997).

Chez les deux sexes, il est exprimé dans les lignées cellulaires somatiques (Sertoli dans le testicule et granulosa dans l'ovaire), suggérant une implication dans le développement gonadique précoce et un possible dimorphisme sexuel (Manuylov et al., 2011).

- **M33** correspond à la version murine du gène *Polycomb* de la famille Hox chez la drosophile. Il est capable de remodeler la chromatine et entraîne le silencing épigénétique des gènes durant le développement embryonnaire chez la drosophile. M33 est suspecté jouer un rôle dans le développement de la gonade précoce (Kato-Fukui et al., 2005).
- **Gata4**, facteur de transcription, contenant deux domaines de fixation à l'ADN en doigt de zinc, est requis pour le développement du cœur (Lourenço et al., 2011). Il est exprimé dans les gonades et pourrait être un régulateur de l'expression génétique gonadique (Biaison-Lauber, 2010). Ce gène est impliqué dans la différenciation gonadique mais il n'est pas considéré comme un facteur crucial de la formation des crêtes et des gonades indifférenciées.

A ce stade, plusieurs gènes qui seront par la suite déterminants dans la formation des testicules chez le mâle, sont exprimés chez les mâles et les femelles au niveau des crêtes génitales. Il s'agit notamment de **SOX9** (*Sry-bOX gene 9*), impliqué dans la chondrogenèse et la neurogénèse (Mead et al., 2013) et **Wnt4** (*wingless-type MMTV integration site family, member 4*) impliqué dans la prolifération cellulaire (Chassot et al., 2012; Miang-Lon Ng and Lufkin, 2012). Malgré leur expression précoce dans la gonade indifférenciée, leur rôle à ce stade n'est pas encore établi.

### **Différenciation gonadique chez le mâle :**

A partir de la 7<sup>ème</sup> semaine de gestation chez l'homme, la différenciation des gonades débute chez les individus XY, par le développement des tubes séminifères, sièges de la formation des spermatozoïdes. Cette différenciation est activée par SRY. Chez la souris, l'expression de SRY débute le 10<sup>ème</sup> jour et demi après la fécondation (soit à la fin de la 4<sup>ème</sup> semaine chez l'homme) au centre des crêtes génitales XY, pour s'étendre à toute la gonade 1 jour plus tard (11.5 jpc) et s'éteindre le lendemain (12.5 jpc ; 7<sup>ème</sup> semaine chez l'homme) (Bullejos and Koopman, 2001).

L'expression de SRY persiste toute la vie chez l'homme (Hanley et al., 2000), chez le cochon (Daneau et al., 1996; Parma et al., 1999) et chez la chèvre (Montazer-Torbati et al., 2010) alors qu'elle s'arrête le 12,5<sup>ème</sup> jour chez la souris (Hacker et al., 1995).



Cette expression est régulée par des facteurs de transcription, tels que WT1, SF1, GATA4 et son cofacteur ZPFM2 (Tevosian et al., 2002), CBX2 (Pitetti et al., 2013) et la famille des récepteurs à l'insuline (Nef et al., 2003).

Les isoformes de **WT1**, + KTS et -KTS, semblent avoir un rôle paradoxal dans la différenciation gonadique (Wagner, 2003). WT1 (+KTS) active Sry, et joue donc un rôle dans la différenciation mâle (Hammes et al., 2001). WT1 (-KTS) pourrait également être impliqué dans la masculinisation de la gonade, car il est capable d'activer directement l'AMH, ou indirectement en interaction avec Sf1. Cependant, il est également capable d'activer *Dax1* (*Dosage-sensitive sex reversal, Adrenal hypoplasia critical region, on chromosome X, gene 1*), un répresseur de l'AMH, ou Wnt-4, spécifique de la voie femelle (Wagner, 2003).

**Sf1** possède deux domaines de liaison à l'ADN en doigt de zinc, et un domaine de fixation à un ligand formant une structure AF-2. Il a été montré in vitro que SF1 se fixait et activait les promoteurs de Sry chez l'homme (de Santa Barbara et al., 2001), et chez le cochon (Pilon et al., 2003). De plus, des études fonctionnelles suggèrent que Sf1 pourrait agir comme médiateur de la sur-régulation de Sry (de Santa Barbara et al., 2001).

**Gata4** chez la souris est surexprimé dans les cellules de Sertoli à 13,5 jours (équivalent à la 7<sup>ème</sup> semaine chez l'homme) (Kyrönlähti et al., 2012) et maintenu jusqu'après la naissance (Eggers and Sinclair, 2012). Il est sous régulé chez les souris XX. Gata4 est requis pour la différenciation testiculaire et pour la morphogénèse des cordons testiculaires (Manuylov et al., 2011). Il agit avec ZFPM2 (*Zinc Finger Protein Multitype 2*) pour réguler l'expression de Sry. Chez la souris *Gadd45 $\alpha$*  interagit avec Mekk4 pour activer p38-Mpk, qui phosphoryle GATA-4, induisant l'expression de Sry (Gierl et al., 2012; Warr et al., 2012).

**Cbx2** est essentiel pour le développement normal de la gonade mâle (Katoh-Fukui et al., 2012). Cependant, des études fonctionnelles montrent qu'il ne régule pas directement l'expression des gènes tels que SF1 connus pour être impliqués dans la différenciation sexuelle (Biaison-Lauber, 2010). Cbx2 pourrait agir en amont de Sry dans le développement sexuel chez l'homme (Biaison-Lauber et al., 2009).



Les membres de la famille des récepteurs à l'insuline (**IR**), INSR (*insulin receptor*), et IGF1R (*IGF type I receptor*), régulent négativement des centaines de gènes impliqués dans le développement des surrénales, des testicules et des ovaires (Pitetti et al., 2013), comme Wt1, Lhx9, Sf1, et 134 autres gènes. SF1 pourrait être régulé directement par IR, ou par l'intermédiaire des voies de signalisation PI3K/AKT (Pitetti et al., 2013) et/ou les membres de la voie MAPK, notamment **MAP3K1** et **MAP3K4** (Bashamboo et al., 2010; Pearlman et al., 2010).

Enfin, Ross et al. ont montré que la séquence en 5' de Sry était très conservée entre l'homme, le bovin, le cochon et la chèvre (Ross et al., 2008) et ont identifié 38 familles de facteurs de transcription pouvant potentiellement s'y lier pour le réguler. L'homologie avec la souris est moins importante, d'autres facteurs de transcription régulateurs de Sry pourraient donc intervenir chez l'espèce murine.

SRY agit comme un facteur de transcription dans la voie sexuelle. Ce gène appartient à la famille de facteurs de transcription SOX (*Sry-related HMG bOX protein*), composée de 20 membres. Chacun des membres SOX possède un motif HMG (*High Mobility Group*) capable de se lier à une séquence d'ADN consensus (A/T)ACAA(T/A) (Polanco and Koopman, 2007). Chez la souris, la répression de Sry après le 12,5<sup>ème</sup> jour pourrait être régulée par SOX9 (Sekido and Lovell-Badge, 2008).

Le principal gène cible de Sry est SOX9, appartenant également à la famille SOX (Eggers and Sinclair, 2012). SOX9 fut le premier à être associé au développement testiculaire (Kwok et al., 1995; Ninomiya et al., 1996) lorsque des mutations furent identifiées dans et autour de SOX9 chez des patients atteints de dysplasie campomélique, caractérisée par des désordres osseux et du développement sexuel (Jakobsen et al., 2007).

Chez la souris, SOX9 est d'abord exprimé dans la crête génitale à bas niveau juste après l'expression de Sry dans les embryons XY au jour 10,5 (Morais da Silva et al., 1996a). Son niveau d'expression croit ensuite considérablement au jour 11,5 dans la gonade XY, alors qu'elle décroît dans la gonade XX. SOX9 est principalement impliqué dans la chondrogenèse, et le développement des cellules de Sertoli chez le mâle (Jiang et al., 2012). Il possède différents sites de régulation (Benko et al., 2009; Mead et al., 2013), et deux de ces sites semblent être spécifiques du testicule (Fonseca et al., 2013).

Il s'agit de TESCO (*TES*tis -*specific enhancer of SOX9 including a 1.3 kb CO*re region) (Sekido and Lovell-Badge, 2008), et une région à 550kb environ en amont de *SOX9* (Fonseca et al., 2013). TESCO est composé de 1.3 kb, et possède une région ECR (pour *Evolutionarily Conserved Region*), très conservée entre les espèces. Cette région de 180 pb est présente chez les euthériens, les oiseaux, les reptiles et les amphibiens, et contient des modules hautement conservés, sites de régulation potentiels des protéines SOX, TCF/LEF, *Forkhead*, DMRT, et GATA (Bagheri-Fam et al., 2010).

Sry agit en synergie avec Sf1 pour réguler positivement l'expression de *SOX9* par fixation sur TESCO. De plus, *SOX9* et Sf1 se lient ensemble sur TESCO pour maintenir l'expression de *SOX9* (Knower et al., 2011; Sekido and Lovell-Badge, 2008). *SOX9* est également régulé par FGF9-FGFR2 et PTGDS-Prostaglandine D2.

**FGF9** (*Fibroblast Growth Factor 9*) code pour un facteur de croissance impliqué dans de nombreux processus tels que la prolifération cellulaire, la survie cellulaire, la migration et la différenciation cellulaire (Eggers and Sinclair, 2012). Il induit l'expression de *SOX9* et bloque l'expression de *Wnt4* (Kim et al., 2006).

**Ptgds** (*Prostaglandin D synthase*) est impliqué dans la synthèse de Prostaglandine 2 ou PGD2. *SOX9* peut se fixer et activer directement le promoteur de *Ptgds*. De plus, PGD2 active la transcription de *SOX9* et sa translocation dans le noyau (Moniot et al., 2009).

Une seule des cibles avérées de *SOX9* dans les gonades est le gène codant pour l'hormone anti Müllérienne, l'AMH (*Anti-Müllerian Hormone*) impliquée dans la régression des canaux de Müller chez le mâle (Arango et al., 1999). La différence de phénotypes observés chez la souris XY entre un KO de *SOX9* (inversion sexuelle (Lavery et al., 2011)) et un KO de *Amh* (canal de Müller persistant et cryptorchidie bilatérale (Hughes et al., 2006)) laisse supposer que *SOX9* régule d'autres gènes.

D'autres gènes jouent également un rôle dans la formation de la gonade :

**Dhh** (*Desert hedgehog*) est exprimé dans la gonade XY murine en développement à partir du 11<sup>ème</sup> jour et demi (semaine 5 chez l'homme) puis dans les cellules de Sertoli. Aucune expression n'est détectée dans la gonade XX (Beverdam and Koopman, 2006).



DHH se fixe à son récepteur PTCH1 (*Patched 1*), qui est exprimé rapidement après DHH. PTCH1 est présent à la membrane des cellules de Leydig, et semble nécessaire à l'activation de Sf1 dans les cellules de Leydig (Yao et al., 2002).

**Dmrt1** : (*Double-sex and mab-3 related, transcription factor 1*) est régulé positivement dans les crêtes génitales en développement dans les embryons XY. Il est essentiel pour le maintien de la détermination testiculaire chez les mammifères, et est conservé chez les métazoaires (Matson et al., 2012). Des copies de ce gène correspondent au gène majeur de la différenciation sexuelle chez le poulet, le médaka et le xénope (Kikuchi and Hamaguchi, 2013).

Enfin, un autre membre de la famille SOX, **Sox10** semble également impliqué dans le développement gonadique mâle. Il est exprimé dans la gonade et des études chez la souris ont montré qu'il était capable de mimer l'activité transcriptionnelle de *SOX9* (Polanco et al., 2010).

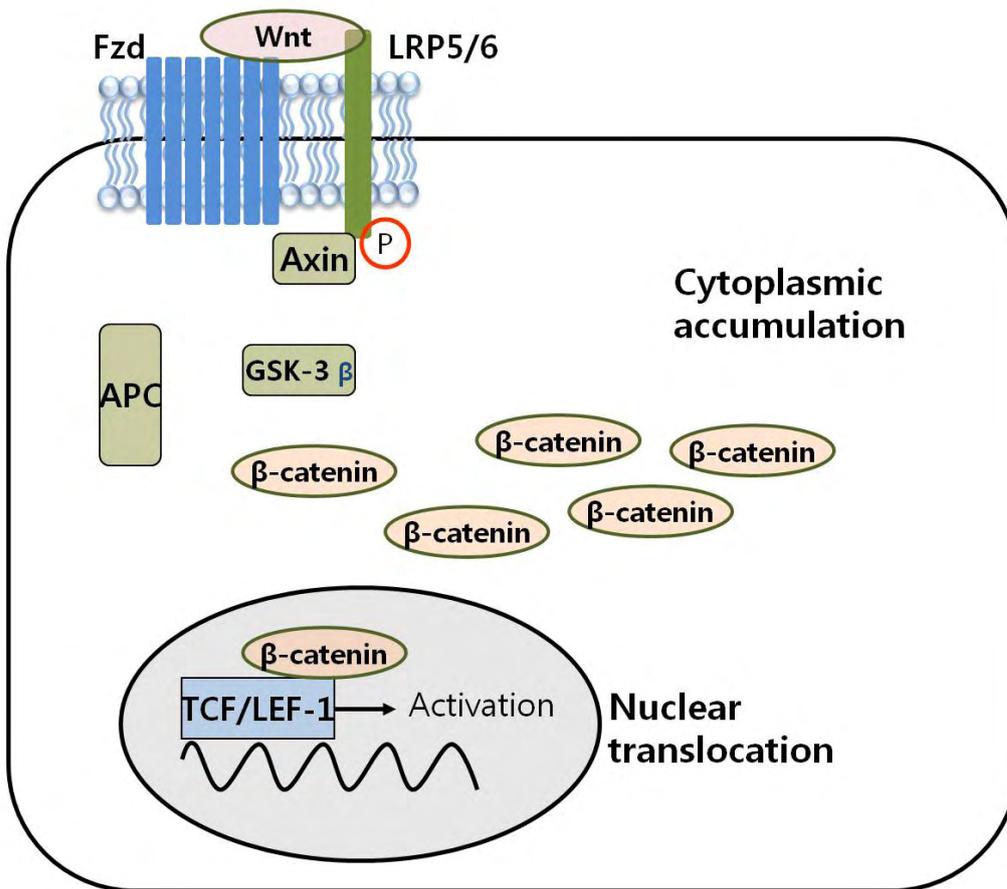


Figure 30 : Régulation de la voie de signalisation Wnt/β-caténine (Han and Na, 2011)

### *Différenciation gonadique chez la femelle :*

La différenciation des gonades femelles débute à la 8<sup>ème</sup> semaine de gestation. La féminisation des gonades indifférenciées a été considérée comme le déterminisme sexuel par défaut, car en l'absence de Sry chez la souris XY, les gonades se féminisent (Jost, 1952). Cependant, des défauts d'activation de Foxl2 ou Wnt4, deux gènes majeurs du déterminisme sexuel chez la femelle, sont responsables de masculinisation des gonades. Il est indispensable pour le bon développement de la gonade femelle que SOX9 soit réprimé (Ottolenghi et al., 2005; Vainio et al., 1999). Cette répression dure tout au long de la vie chez la femelle, selon différents processus.

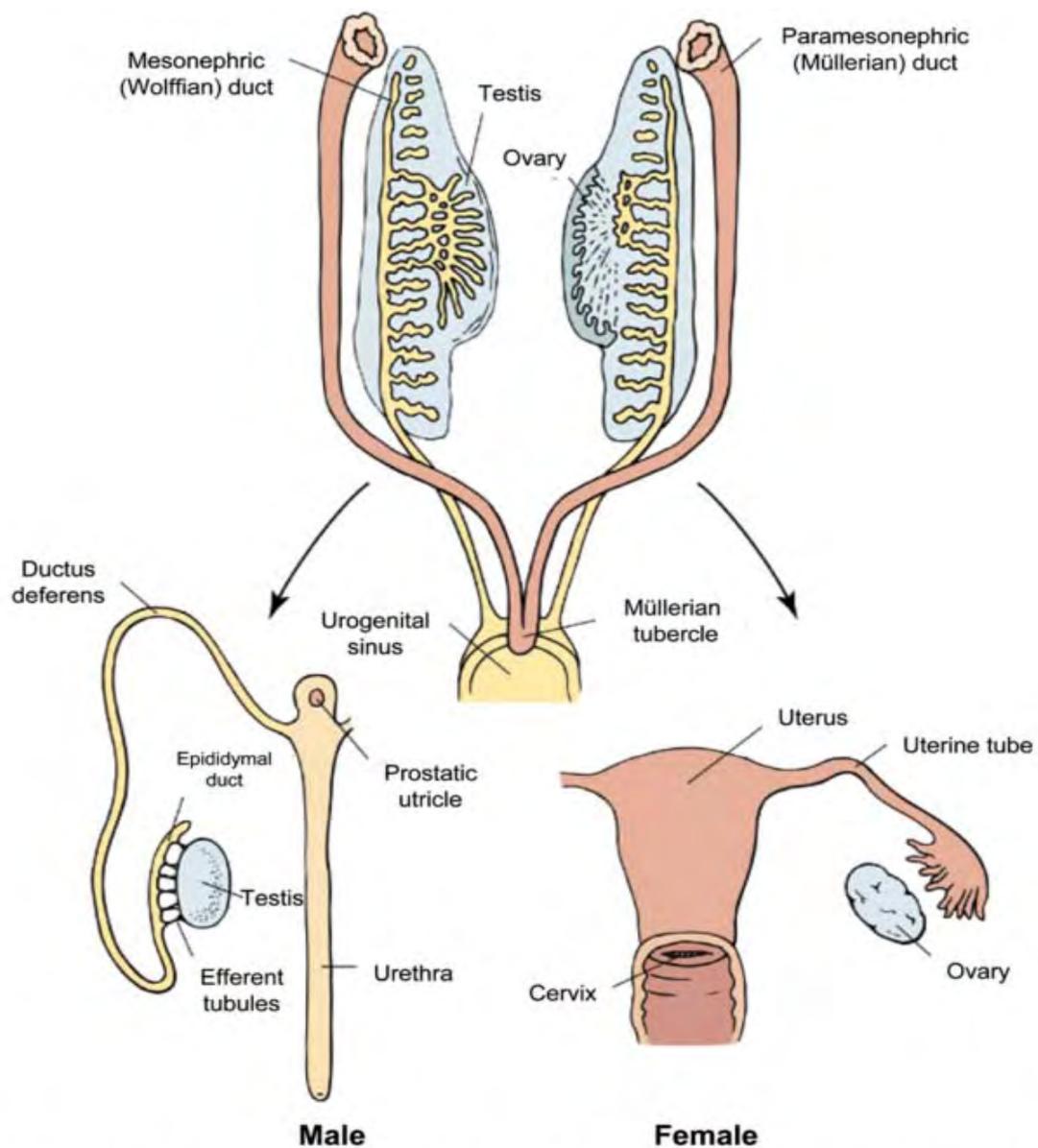
La première étape de la répression dans la gonade femelle induit la différenciation des cellules de soutien en cellules de la granulosa. Le gène **Wnt4** (*Wingless-type MMTV integration site family, member 4*) joue un rôle primordial dans cette première étape (Jakob and Lovell-Badge, 2011). WNT4 appartient à la famille WNT, impliquée dans plusieurs processus du développement (Eggers and Sinclair, 2012). Au 11<sup>ème</sup> jour après la fécondation chez la souris, il est exprimé dans le mésonephros et la gonade bipotentielle, puis est réprimé au niveau du testicule en développement. Son expression persiste dans le mésonephros de l'ovaire en développement ainsi que dans le mésenchyme entourant les canaux de Müller (Jääskeläinen et al., 2010). Les résultats de Jääskeläinen et al. montrent le rôle potentiel de Wnt4 dans la maturation de l'ovaire. Il agit en amont de la follistatine (Fst) dans l'ovaire en développement (Yao et al., 2004). Il bloque la fixation de Sf1 sur TESCO, inhibant ainsi l'expression de SOX9 (Bernard et al., 2012). Wnt4 agit avec Rspo1 (*R-spondin1*) pour stabiliser la  $\beta$ -caténine (Chassot et al., 2012; Parma et al., 2006; Tomizuka et al., 2008). La voie Wnt/ $\beta$ -caténine est présentée en Figure 30. Les protéines Wnt jouent essentiellement un rôle via leur liaison à un récepteur protéique à sept domaines transmembranaires Fzd (pour *Frizzled*) et les protéines LRP (pour *LDL-Related Protein*) 5 et 6. Leur recrutement extracellulaire entraîne la phosphorylation intracellulaire des LRP, puis la déphosphorylation de la  $\beta$ -caténine et son entrée dans le noyau de la cellule. L'accumulation de la  $\beta$ -caténine dans le noyau entraîne l'activation des gènes cibles (Han and Na, 2011), par liaison avec les complexes TCF (*T-cell Factor*) et LEF (*Lymphoid Enhancer Factor*) (Maatouk et al., 2008).



**Fst** code pour une glycoprotéine à chaîne simple qui se lie aux protéines de la famille des TGFβ comme l'activine pour neutraliser leur activité. Elle fut d'abord identifiée comme un inhibiteur de la sécrétion de FSH, *Follicle-Stimulating Hormone* qui stimule divers processus du développement et de la reproduction chez les mâles et les femelles (George et al., 2011). Chez la souris XX, la répression de Fst est associée à une formation du tissu testiculaire (Yao et al., 2004), soulignant son rôle dans le développement ovarien (Kashimada et al., 2011)

Après la naissance, la répression de SOX9 est assurée par Foxl2 (*Forkhead box protein L2*). **Foxl2** code pour un facteur de transcription avec un domaine de liaison à l'ADN dit à « tête de fourchette ». Son rôle dans le développement ovarien fut d'abord suggéré lors de l'identification d'une délétion près de FOXL2 chez la chèvre PIS (pour *Polled/intersex syndrome*) (Pailhoux 2001) présentant une inversion sexuelle et une absence de cornes. La répression de SOX9 par FOXL2 passe par la fixation sur TESCO du complexe FOXL2/récepteurs aux estrogènes ESR (*Estrogen Receptor*) 1 et 2 (Uhlenhaut et al. 2009 (Cell 139 : 1130-42) ; Bagheri 2010). FOXL2 agit également avec BMP2 pour assurer l'expression de Fst dans l'ovaire (Jakob, 2010).

A partir de la 8<sup>ème</sup> semaine, les cordons sexuels se résorbent et deviennent des follicules primordiaux. Après plusieurs stades de maturation, ils deviendront des follicules ovariens dont le stock est établi de manière définitive vers la 28<sup>ème</sup> semaine.



**Fig. 2.** The undifferentiated sexual system at 6–7 week of gestation. Precursor structure and their mature counterparts have the same color. Fossum G. Atlas of Clinical Gynecology: Pediatric and Adolescent Gynecology. Edited by Morton Stenchever (series editor), Alvin F. Goldfarb.

Figure 31 : Système sexuel indifférencié jusqu'à la 7ème semaine

### III.B.3. Différenciation du sexe phénotypique :

La différenciation de la gonade mâle induit la sécrétion d'hormones responsables de la masculinisation du tractus génital interne et des organes génitaux externes.

Jusqu'à la 7<sup>ème</sup> semaine, mâles et femelles possèdent les mêmes systèmes de canaux : les canaux de Wolff, qui se développent à partir de la 4<sup>ème</sup> semaine de gestation (Figure 27 et Figure 31) et les canaux de Müller, dont la mise en place se fait au cours de la 6<sup>ème</sup> semaine.

Chez la femelle, les canaux de Müller persistent pour former les trompes de Fallope, l'utérus et le tiers supérieur du vagin (Figure 31). L'absence de testostérone entraîne la dégénérescence des canaux de Wolff durant la 10<sup>ème</sup> semaine. Le sinus urogénital se développe et donne l'urètre et les deux tiers inférieurs du vagin (Hughes et al., 2006)

Chez le mâle, les cellules de Leydig (Bernard et al., 2012) se différencient dans l'espace interstitiel du testicule, entre les cordons testiculaires. Le rôle majeur des cellules de Leydig est la sécrétion de testostérone, qui agit sur les organes génitaux internes et externes : elle induit la différenciation des canaux de Wolff en épididyme, canaux déférents et vésicules séminales, et le développement du pénis et du scrotum, après conversion en dihydrotestostérone (DHT) par la 5- $\alpha$ -réductase, produite par les tissus cibles (Hughes, 2001). Les cellules de Leydig sécrètent également l'*Insulin like-3* (Insl3) impliquée dans la descente du testicule (Adham and Agoulnik, 2004), dont la position finale dans le scrotum permet d'assurer une température optimale à la spermatogénèse. Les cellules de Sertoli sécrètent l'AMH, dont l'action entraîne la régression des canaux de Müller.

La descente testiculaire (Figure 32) se déroule entre le 4<sup>ème</sup> mois et le terme de la grossesse, selon deux mécanismes principaux : Le rétrécissement du gubernaculum testis, un cordon reliant le testicule et le futur scrotum, et l'évagination du péritoine, ou *processus vaginal* qui aide le testicule dans sa traversée du canal inguinal. Ces mécanismes font interagir des facteurs génétiques et hormonaux tels que la testostérone et l'*Insulin like-3* (Insl3).

Le testicule se met en place dans le scrotum selon deux phases (Hutson, 1997, Baumns, 1983): la phase trans-abdominale et la phase inguino-scrotale. Deux ligaments jouent un rôle majeur dans la descente testiculaire : le ligament suspenseur cranial (CSL), et le gubernaculum ou ligament génito-inguinal.

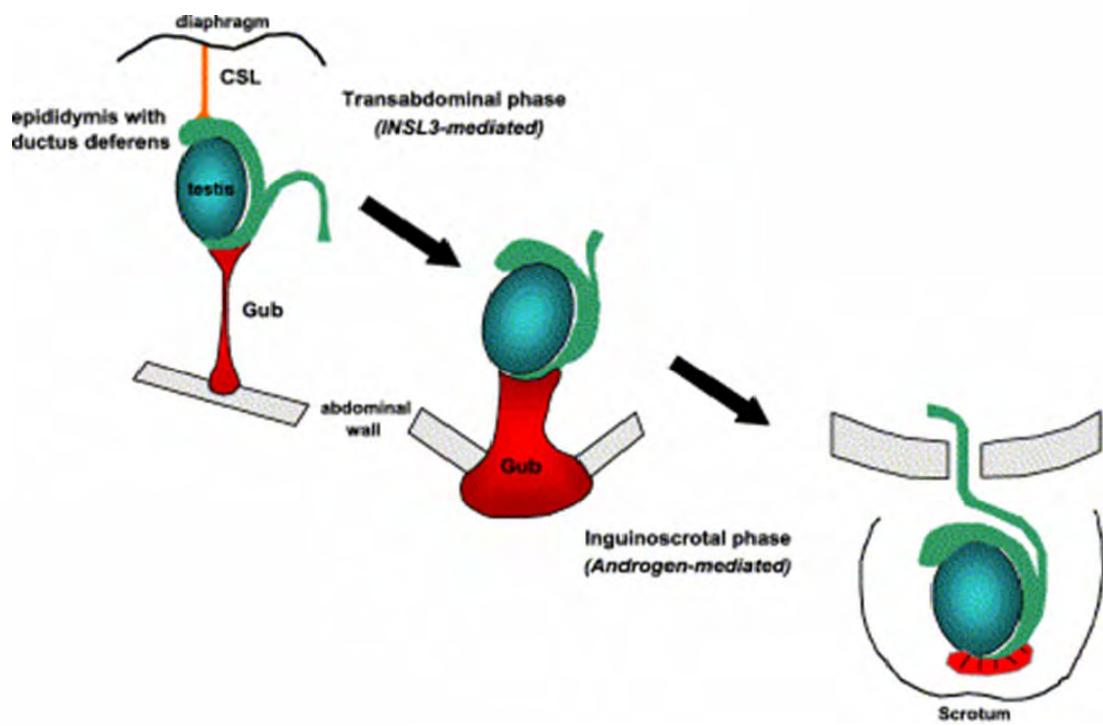


Figure 32 : Descente testiculaire

Initialement, le testicule est en position pararénale, reliée à la paroi abdominale par le ligament suspenseur cranial, et à l'anneau inguinal par le gubernaculum.

La phase trans-abdominale, entre la 10<sup>ème</sup> et la 23<sup>ème</sup> semaine de vie fœtale, est caractérisée par l'élargissement du gubernaculum au niveau caudal, sous l'action de l'InsI3 sécrétée par les cellules de Leydig (Adham and Agoulnik, 2004; Nef and Parada, 1999). Le ligament suspenseur cranial régresse sous contrôle de la testostérone (Husmann and Levy, 1995). Le bulbe gubernaculaire, ou futur scrotum, se développe au niveau de la région caudale de l'anneau inguinal (Figure 32). La phase inguino-scrotale débute la 26<sup>ème</sup> semaine de la vie fœtale environ et se termine entre la 28<sup>ème</sup> semaine et le terme de la grossesse. Le testicule migre de la région inguinale vers le scrotum, par la régression du gubernaculum. Le bulbe gubernaculaire continue à se développer. Le *calcitonin gene-related peptide* (CGRP) contrôle la migration du testicule. Ce neurotransmetteur est relargué par le nerf genitofémoral, sous l'action de la testostérone.

Trois nouveau-nés sur cent ont un testicule non descendu à la naissance. Dans 70% des cas, la descente testiculaire sera post natale, entre le 4<sup>ème</sup> et le 6<sup>ème</sup> mois après la naissance. Ce phénomène de descente testiculaire spontanée est associé à des forts taux de LH et de testostérone entre le 2<sup>ème</sup> et le 3<sup>ème</sup> mois après la naissance. Cependant la descente post natale ne survient pas au-delà de 4 à 6 mois.

Après la mise en place du testicule dans le scrotum et jusqu'à la première année suivant la naissance, le canal inguinal s'oblitére par des phénomènes d'apoptose cellulaires, pour ne devenir qu'un ligament péritonéo-vaginal.

Comme la spécialisation des gonades indifférenciées et des canaux génitaux, les ébauches des organes génitaux externes se différencient à partir d'une structure initialement identique, sous l'action hormonale. D'autres caractéristiques sexuelles, telles que la modification de la poitrine ou la pilosité se mettent en place plus tardivement, à la puberté. Ce n'est qu'au moment de la vie adulte que l'appareil reproducteur devient fonctionnel.

La différenciation sexuelle est un processus long et complexe, impliquant de nombreux systèmes génétiques et hormonaux. Des défauts peuvent subvenir aux différents stades de la différenciation, et induire des anomalies du développement sexuel, voire des inversions sexuelles.

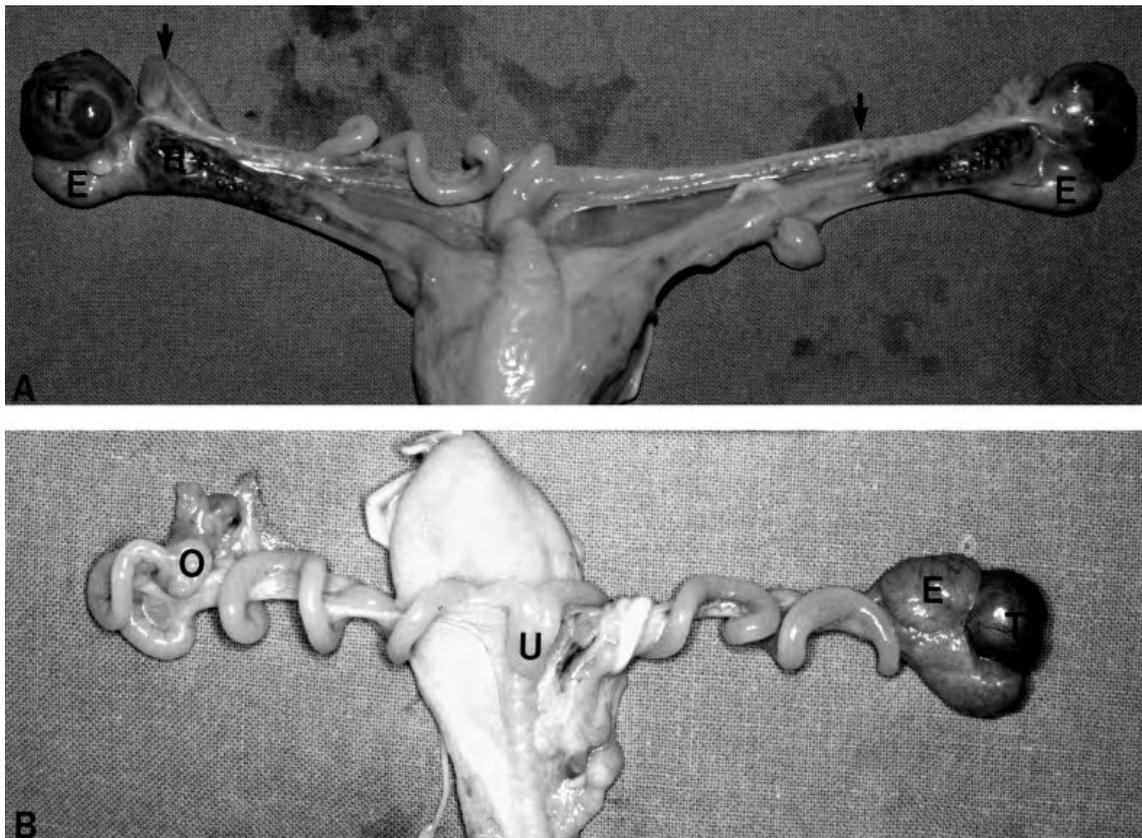


Figure 33 : Tractus du système reproducteur et gonades de deux porcelets intersexués, 38 XX (Pailhoux et al., 2001a)

- A. Mâle XX avec des gonades similaires à des testicules (T). L'épididyme (E) s'étend à partir des gonades, les canaux de Muller et l'utérus sont présents. Les cornes utérines s'arrêtent avant les gonades (flèches).
- B. Vrai hermaphrodite avec un ovaire (O) à gauche et une gonade testiculaire (T) à droite. L'utérus (U) est présent et relié aux gonades.

### III.C. Etiopathogénie des anomalies étudiées

#### III.C.1. DSD 46,XX – désordre du développement ovarien

Les DSD 46, XX sont répartis en 3 catégories :

- les désordres du développement ovarien,
- les défauts par excès d'androgènes aux niveaux fœtal (hyperplasie congénitales des surrénales, mutations du récepteur aux glucocorticoïdes), fœto-placentaire (déficience en aromatase) ou maternel (lutéome, exogène),
- les défauts autres tels que l'atrésie vaginale, l'extrophie du cloaque, les anomalies utérines.

Nous nous intéressons dans cette thèse aux défauts entrant dans la catégorie des désordres du développement ovarien.

#### *Physiopathologie*

Le DSD 46,XX ovarien correspond à une masculinisation des organes génitaux chez un individu XX. Au niveau gonadique, l'individu atteint possède des tissus ovariens (DSD ovarien), testiculaires (DSD testiculaire) ou testiculaires et ovariens dans la même gonade, on parle alors d'ovotestis (DSD ovotestis). Cette nomenclature a remplacé respectivement les termes de pseudohermaphrodisme femelle, inversion sexuelle XX et hermaphrodisme vrai (Lee et al., 2006).

Les organes génitaux externes sont virilisés à des degrés variables, avec présence de clitoris péniforme et hypospadias, scrotum vide, et parfois oblitération vaginale. Dans certains cas, la virilisation est complète et le phénotype mâle est normal. La pathologie DSD 46,XX peut alors passer inaperçue, jusqu'à ce que les individus consultent car ils ne peuvent pas avoir d'enfant (Cox et al., 2011; Vetro et al., 2011). Les testicules sont généralement hypotrophiés et les individus sont infertiles.

Chez les animaux d'élevage, des cas d'intersexualité DSD XX sont retrouvés chez le porc et la chèvre (Pailhoux et al., 1998). Ces animaux intersexués sont un problème dans les élevages car ils sont stériles, et des concentrations anormales d'hormones mâles peuvent entraîner une dépréciation des carcasses.

Gène	Phénotype gonadique chez la souris	Phénotype gonadique chez l'homme
Sry	Translocation de Sry chez XX: inversion sexuelle	GOF/translocation : DSD testiculaire chez XX
SOX9	Ods (délétion en amont de SOX9) : inversion sexuelle chez XX	GOF : inversion sexuelle chez XX Duplication incluant SOX9 : DSD testiculaire chez XX Duplication/triplication en amont de SOX9 : DSD testiculaire chez XX
	Knockout conditionnel : développement testiculaire chez XX	Délétion en amont de SOX9 : DSD ovarien chez XY avec dysplasie campomélique acampomélique, dysgénésie gonadique, organes génitaux externes femelles ou ambigus
	Pas de phénotype gonadique associé	Duplications (incluant Sox3) et délétion en amont de Sox3 : DSD testiculaire chez XX
Sox10	Pas de phénotype gonadique associé	Duplication environnant Sox10 (incluant d'autres gènes) : féminisation incomplète ou masculinisation chez XX
Wnt4	Wt4 -/- : agénésie du canal de Muller, synthèse de testostérone, formation de vessie cœlomique	LOF : agénésie du canal de Muller, synthèse de testostérone, et formation de vessie cœlomique
Rspo1	Inversion sexuelle partielle chez XX, similaire à Wt4 -/-	DSD testiculaire et ovotestis chez XX
$\beta$ -caténine	Knockout conditionnel dans les cellules de Sertoli chez XX, similaire à Wt4 -/- et Rspo1 -/-	Pas de mutation décrite
Foxl2	Foxl2 -/- : insuffisance ovarienne prématurée	BPES (Blepharophemosis/Ptois/Epicanthus inversus syndrome) et insuffisance ovarienne prématurée
	Foxl2 -/-, Wt4 -/- : inversion sexuelle chez XX	
Fst	Inversion sexuelle partielle chez XX, formation de vessie cœlomique	Pas de mutation décrite
Mamld1		GOF : DSD avec dysgénésie gonadique chez XX

Tableau 10 : Gènes associés au développement de la gonade différenciée chez la souris et l'homme (bleu : gènes de la différenciation mâle, rose : gènes de la différenciation femelle).  
LOF : Loss Of Function, GOF : Gain Of Function.

Dans l'espèce porcine, deux types phénotypiques sont observés : et répertoriés comme: vrais hermaphrodites (50% des intersexués), et mâles XX avec ambiguïtés (50% des animaux).

Les vrais hermaphrodites (Pailhoux et al., 2001b) possèdent des ovotestis ou un ovaire et un testicule (Pailhoux et al., 2001a). Les canaux de Muller ne sont pas totalement résorbés lors de la présence de testicules. L'utérus est présent, mais les trompes utérines s'arrêtent avant les gonades. Le plus souvent, les organes génitaux externes sont femelles.

Les mâles XX sont des animaux intersexués correspondant à des cas de masculinisation chez des femelles (caryotype XX). Les animaux possèdent des gonades similaires à des testicules, retenues dans la cavité abdominale. Les organes génitaux internes et externes sont masculins, avec parfois un clitoris péniforme. Après la puberté, les testicules apparaissent plus petits que des testicules normaux, dû à l'absence de cellules germinales (Pailhoux et al., 2001b). Les animaux sont stériles, malgré la présence de tubes séminifères et de cellules de Sertoli et de Leydig.

Chez la chèvre, les cas d'intersexués sont observés chez les femelles sans cornes (plutôt dans les races Saanen, Toggenburg, Alpine ou Damascus). L'absence de cornes est un caractère intéressant pour les sélectionneurs caprins car les cornes représentent un risque de blessures pour les éleveurs et entre les animaux lors de combat. Cependant, l'absence de cornes chez la chèvre peut s'accompagner d'une inversion sexuelle chez les animaux XX, responsable de stérilité (Pailhoux et al., 2001b). Deux types phénotypiques et gonadiques sont observés : 50% des animaux possèdent des testicules, et des organes génitaux externes mâles, et 50% présentent des gonades ambiguës. Chez les animaux avec gonades ambiguës, les cordons spermatiques sont peu développés, et les cellules germinales sont en arrêt de méiose. Un clitoris proéminent avec allongement de la distance ano-génitale peut également être observé.

### *Les causes génétiques identifiées*

Des défauts génétiques impliquant les gènes de la détermination sexuelle ont été identifiés chez l'homme et chez la souris notamment. Le Tableau 10 présente les gènes et mutations identifiés dans les cas de DSD XX.

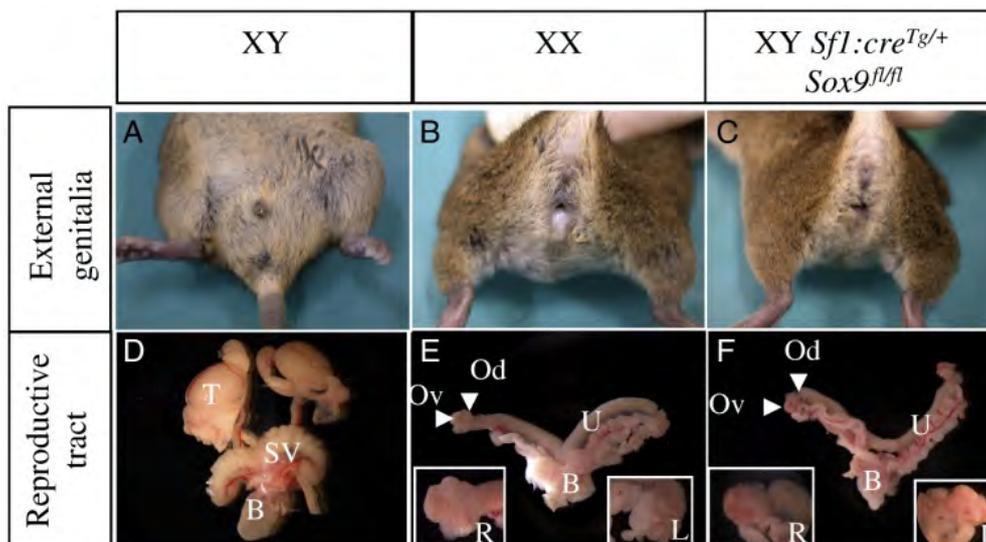


Figure 34 : Phénotypes observés chez des souris XX et XY normales, et XY *SOX9*<sup>-/-</sup> respectivement (Lavery et al., 2011). A, B, C : appareil génital externe, D, E, F : appareil reproducteur interne. Testicule (T), Vésicule séminale (SV), vessie (B) Ovaire (Ov), oviducte (Od), urètre (U) ovaire droit (R), ovaire gauche (L).



Figure 35 : Phénotypes observés chez une souris femelle normale à gauche,

et une souris *Ods*, à droite (Bishop et al., 2000a).

**Sry** : Sry est le premier gène causal à avoir été identifié (Sinclair et al., 1990). Certains cas de DSD s'expliquent par l'activation de ce gène chez des individus XX ou, *a contrario*, son inactivation chez des individus XY.

La translocation du locus de Sry chez la souris XX induit le développement d'organes génitaux mâles (Koopman et al., 1991). Aujourd'hui, 90% des cas de DSD XX chez l'homme s'expliquent par une translocation de SRY. Chez le porc, seuls 6% des cas s'expliquent par la présence de SRY (Pailhoux et al., 1997).

Chez l'homme XY, des mutations ou la perte de fonction de SRY causent une inversion sexuelle (Berta et al., 1990a). 15% des cas de DSD XY sont associés à des mutations perte-de-fonction de SRY (Cameron and Sinclair, 1997).

**SOX9** : *SOX9* est l'un des gènes les plus étudiés dans les cas de DSD XX, chez la souris (transgénèse) et chez l'homme (mutations, délétions, insertions).

Chez la souris XY, l'inactivation par transgénèse de *SOX9* au niveau des gonades entraîne le développement d'ovaires fertiles et de caractéristiques externes femelles (Lavery et al., 2011) (Figure 35). A l'inverse, la surexpression de *SOX9* dans la gonade XX induit la formation de testicule, et un phénotype mâle (Vidal et al., 2001). La lignée transgénique de souris *Ods* (*Odd Sex*) est un excellent modèle murin de DSD XX. Cette lignée a une mutation insertionnelle dominante *Ods* qui entraîne une délétion de 150kb à 1Mb en amont de *SOX9* (Bishop et al., 2000a). Le gène *SOX9* est sur-exprimé chez les souris *Ods* XX, qui se développent comme des mâles malgré l'absence de Sry (Figure 35). Ces animaux sont néanmoins stériles. L'hypothèse avancée est que la région délétée, bien que très éloignée du gène *SOX9*, comprendrait un (ou des) site(s) de fixation d'éléments répresseurs de *SOX9*.

Chez l'homme, des mutations au niveau de l'exon 3 de *SOX9* ont été identifiées comme responsables de dysplasie campomélique. Cette pathologie associe des malformations du squelette et des ambiguïtés des organes sexuels avec un phénotype plutôt féminin chez les individus XY. Plusieurs délétions et duplications dans la séquence encadrant *SOX9* ont également été identifiées chez l'homme.

La première duplication concerne le fragment (17)(q23.1q24.3) contenant *SOX9*, identifiée chez un enfant intersexué XX (Huang et al., 1999). Les duplications ou triplications identifiées par la suite sont présentées dans la Figure 36 (Fonseca et al., 2013).

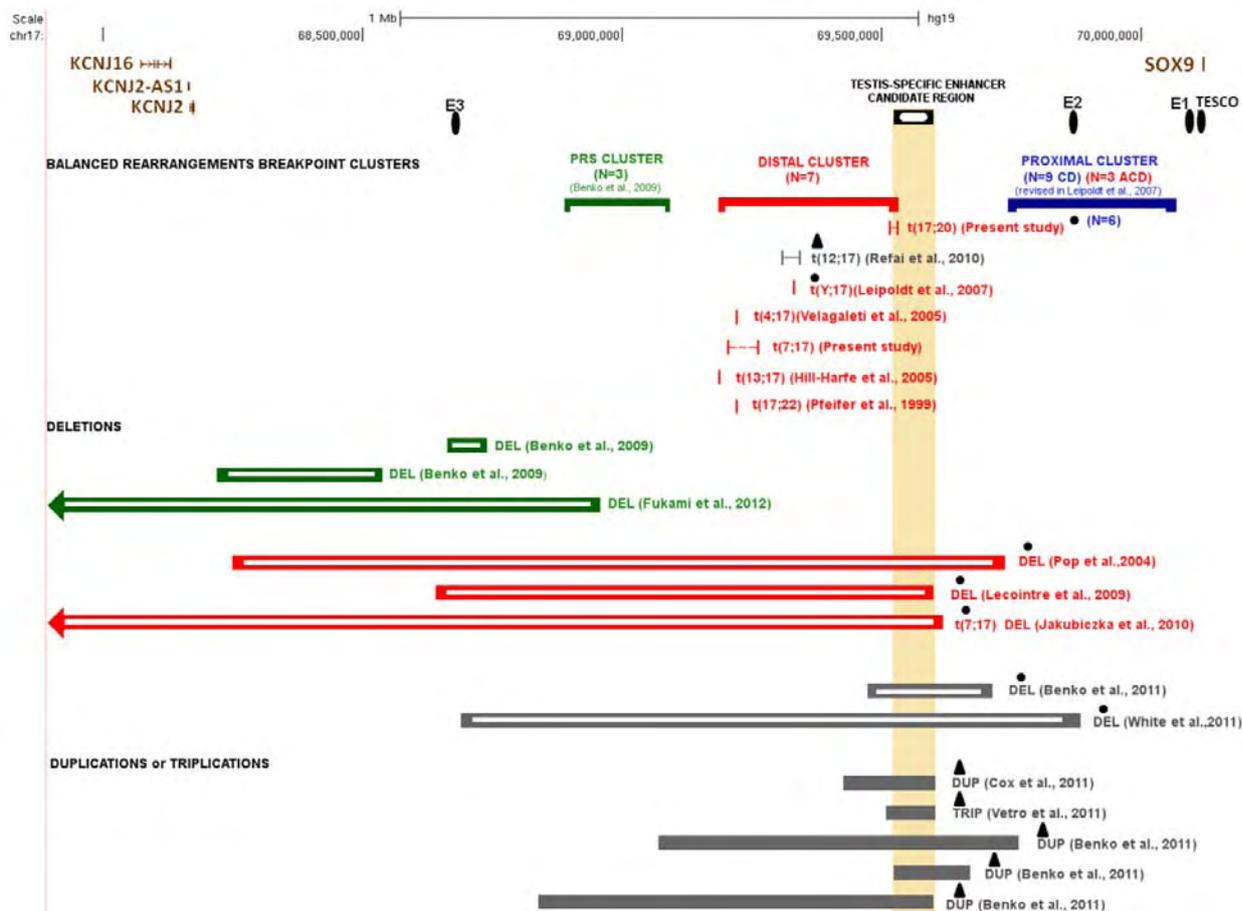


Figure 36 : Réarrangements chromosomiques identifiés en amont de *SOX9*, potentiellement associés à des dérégulations de *SOX9*. (Fonseca et al., 2013)

Réarrangements associés à des cas de dysplasie campomélique (AC, bleu), dysplasie acampomélique campomélique (ACD, rouge), Pierre Robin (PRS, vert) défauts squelettiques (gris), brachydactylie avec anonychie (noir). Les DSD XX sont indiqués avec des triangles noirs, et les DSD XY avec des ronds noirs. La région candidate testicule spécifique autre que TESCO est indiquée avec une barre verticale jaune/orange

Les régions dupliquées ne sont pas similaires entre elles, mais couvrent globalement une même région en amont de *SOX9* (Benko et al., 2009; Cox et al., 2011; Vetro et al., 2011). Ces différentes études laissent supposer qu'un enhancer spécifique des gonades (TESTIS-SPECIFIC ENHANCER sur la Figure 36) pourrait être localisé dans un intervalle de 100kb environ, à 500-600 kb en amont de *SOX9*.

**Sox3 :** Certains arguments sont en faveur de l'implication de Sox3 dans les désordres du développement sexuel. D'une part, la surexpression de Sox3 dans la gonade bipotentielle chez des souris XX transgéniques entraîne une inversion sexuelle complète et une surexpression de *SOX9* (Sutton et al., 2011). D'autre part, l'analyse génomique de trois patients XX avec DSD testiculaire a permis de mettre en évidence des réarrangements au niveau et à proximité de la séquence de *SOX3*. Ces études laissent supposer que *SOX3* est capable de mimer l'action de *SRY* en activant *SOX9*. Cependant, d'autres études montrent que le rôle de Sox3 dans la gonade n'est pas primordial. Chez l'homme et chez la souris, l'expression de *SOX3* dans les gonades en développement est très faible voire nulle (Houmard et al., 2009; Sutton et al., 2011), et des mutations perte-de-fonction chez la souris n'affectent pas le développement sexuel.

**Sox10 :** La surexpression par transgénèse de Sox10 chez la souris XX induit une inversion sexuelle (Polanco et al., 2010) et le degré d'inversion semble corrélé avec son niveau d'expression dans la gonade. *SOX10* serait capable d'activer les cibles transcriptionnelles de *SOX9*. Une duplication du fragment d'ADN du chromosome 22 humain contenant *SOX10* a été identifiée chez un patient DSD XX, et pourrait expliquer la masculinisation des organes génitaux externes observée (Seeherunvong et al., 2004).

**Wnt4 :** Des souris *Wnt4* <sup>-/-</sup> présentent un développement de tissus testiculaires dans la gonade femelle et une régression des canaux de Muller (Vainio et al., 1999). Les canaux de Wolff sont développés, indiquant des taux de testostérone suffisamment importants pour induire leur développement, mais les organes génitaux externes ne sont pas virilisés. Une mutation hétérozygote de *Wnt4* a également été identifiée chez des patients atteints du syndrome de Mayer-Rokitansky-Küster-Hauser, caractérisé par une absence totale ou partielle de vagin et d'utérus, avec des trompes et des ovaires normaux (Biason-Lauber et al., 2004). Certains de ces patients présentent aussi des anomalies rénales et squelettiques.



Une mutation non-sens a également été identifiée chez des individus XX présentant des DSD testiculaires ou ovotestis, avec dysgénésie des reins, des surrénales, et des poumons (SERKAL syndrome) (Mandel et al., 2008)

**Rspo1** : Une mutation homozygote de *Rspo1* a été décrite dans une famille dont certains individus présentaient des DSD testiculaires XX, ainsi que des défauts de la peau (Parma et al., 2006). Une autre mutation de *Rspo1* a été identifiée chez une patiente XX avec DSD ovotestis (Tomaselli et al., 2008). *Rspo1* active la voie de signalisation de la  $\beta$ -caténine par ou en synergie avec *Wnt4* (Chassot et al., 2008; Tomaselli et al., 2011). Un Knock out de *Rspo1* chez la souris n'induit pas une inversion sexuelle complète, mais les souris présentent des ovaires masculinisés, dans lesquelles se sont formés des vaisseaux coelomiques spécifiques des mâles. Les taux de testostérone sont similaires à ceux observés chez le mâle, et les oocytes sont déplétées. Ce phénotype est semblable à celui observé lors de KO de *Wnt4*, supposant une co-dépendance entre *Wnt4* et *Rsp1* (Tomizuka et al., 2008).

**$\beta$ -caténine**: le KO de la  $\beta$ -caténine induit les mêmes caractères d'inversion sexuelle XX que ceux observés lors de KO de *Rspo1* ou *Wnt4* (Liu et al., 2009).

**Fst** : L'inactivation spécifique de la *Fst* ou Follistatine dans les cellules de granulosa induit des cas d'infertilité chez les souris XX (Jorgez et al., 2004). De plus, l'inactivation de *Fst* entraîne une inversion sexuelle partielle chez la souris XX (Yao et al., 2004)

**Pirst1 et Foxl2** : Chez la chèvre, la mutation responsable de l'intersexualité associée au caractère sans corne a été identifiée (Pailhoux et al., 2001c): La mutation PIS, situé sur le chromosome 1, correspond à une délétion de 11,7 kilobases. Cette mutation dominante affecte l'expression d'au moins deux gènes : *Pirst1* (*PIS-regulated transcript1*), et *Foxl2*, situés respectivement à 20 et 200kb de la délétion. Cette délétion réprime l'expression de *Pirst1* et de *Foxl2*, répresseurs naturels de *SOX9* et donc de la voie mâle (Pailhoux et al., 2001b). L'étude des profils d'expression chez les animaux intersexués a montré une surexpression de *SOX9* et de l'AMH caractéristiques des gonades mâles.

**Mamld1** : Le gène *Mamld1* (*mastermind-like domain 1*) a été mis en évidence chez deux cas de DSD (Brandão et al., 2011). Le premier cas est une patiente avec DSD XX présentant une dysgénésie gonadique. Une mutation non-sens gain de fonction de *Mamld1* à l'état homozygote été identifiée chez cette patiente. Le deuxième cas est un patient avec DSD YX et dysgénésie gonadique, hétérozygote pour la même mutation.



Les causes génétiques identifiées chez le porc :

Il existe peu de données génétiques concernant l'intersexualité chez le porc. La présence de SRY chez les individus DSD XX, principale cause d'intersexualité chez l'homme, n'explique que 6% des cas de DSD chez le porc (Pailhoux et al., 1997). Quelques macro-remaniements chromosomiques ont été identifiés chez des porcelets intersexués, et une étude a été réalisée sur les profils d'expression de gènes connus pour être impliqués dans le déterminisme sexuel.

#### *Analyses caryotypiques*

Dans le cadre de l'étude du caryotype de 33 porcs avec DSD, Tambasco et al ont identifié un cas de cassure sur le chromosome 9, ainsi qu'un cas de mosaïcisme 38,XY/38,XX (Tambasco, 1990).

Deux remaniements ont également été mis en évidence par Ducos et al dans le cadre des contrôles chromosomiques réalisés chez les populations d'élevage (Ducos et al., 2002) : Une duplication du bras court du chromosome 9 a été observée chez un animal intersexué de race Large White, et une inversion de type (9) (p12;p22) chez un animal intersexué d'une lignée synthétique.

Enfin, trois cas de caryotypes 39,XXY (Breeuwsma, 1967; Hancock and Daker, 1981; Mäkinen et al., 1998) et un cas de mosaïcisme 38,XY/39,XXY (Ducos et al., 2008) ont été rapportés. Chez l'homme, la duplication d'un chromosome X (caryotype 47,XXY) correspond au syndrome de Klinefelter (Groth et al., 2013). Chez le porc, ces anomalies chromosomiques sont associées à une hypoplasie testiculaire et un défaut de spermatogénèse (Pinton et al., 2011).

#### *Etude des profils d'expression de gènes clefs de la différenciation sexuelle*

Le niveau d'expression de certains gènes impliqués dans le déterminisme sexuel a été étudié par Pailhoux et al (Pailhoux et al., 2001a). L'étude a été faite dans les tissus gonadiques de mâles et de femelles, ainsi que chez des vrais hermaphrodites et des animaux XX présentant une inversion sexuelle. Dès 50dpc, il apparaît que les gonades de fœtus DSD sont affectés et diffèrent d'ovaires normaux. L'expression de *SOX9* et de l'AMH augmente de manière corrélée au nombre de cellules de Sertoli. Le gène *FOXL2*, pourtant décrit comme spécifique des cellules folliculaires, est également exprimé dans les structures testicule-like d'animaux DSD.

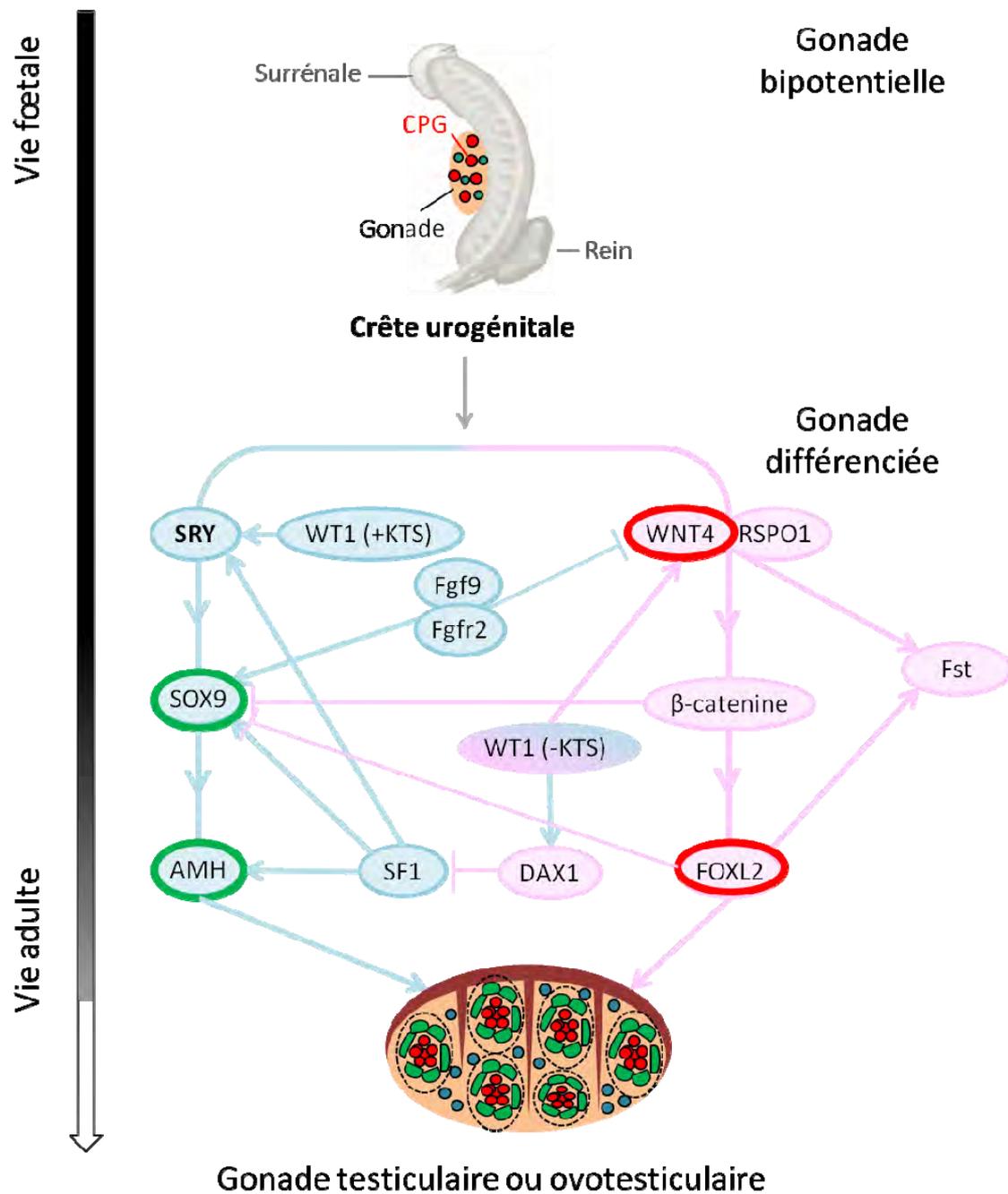


Figure 37 : Schématisation des variations d'expression observées chez l'animal XX intersexué. Les cercles verts indiquent une sur-expression par rapport à la femelle XX, et les cercles rouges une sous-expression.

Ces données suggèrent que les voies de la différenciation du testicule et de l'ovaire sont indépendantes et co-existent au sein de la même cellule, ou que des populations de cellules différentes sont présentes au sein d'une même gonade (Figure 37).

Malgré ces données, aucun gène permettant d'expliquer la cause de 94% des cas d'animaux DSD n'a pour l'heure été identifiée chez le porc. Un des objectifs de cette thèse est de rechercher via une approche génétique le ou les gènes qui seraient impliqués précocement dans la masculinisation de la gonade chez le porc intersexué.

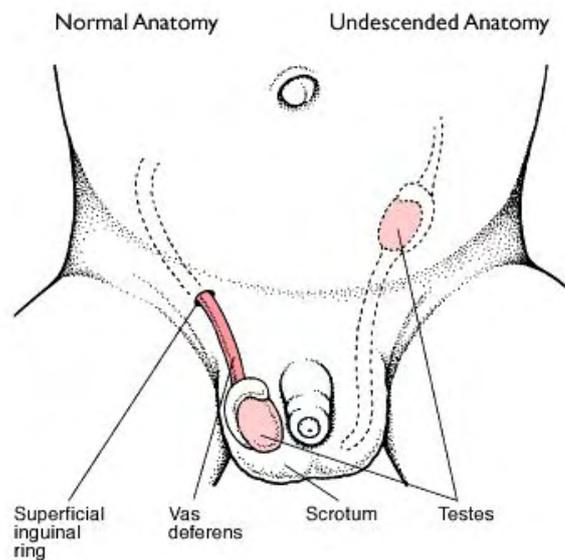


Figure 38 : Testicule non descendu chez le jeune enfant

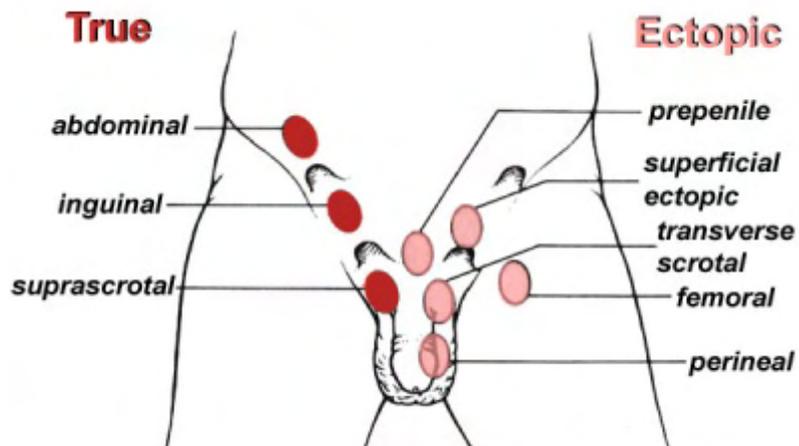


Figure 39 : Classification des cas de cryptorchidies selon la position du testicule non descendu

### III.C.2. La cryptorchidie

Littéralement, le terme « cryptorchidie » vient du grec *kruptos* : caché et *orkhis* : testicule. Il correspond à un défaut de migration vers le scrotum d'un ou des deux testicules (Figure 28). La cryptorchidie est surtout fréquente chez les espèces où la descente testiculaire est tardive, et pour lesquelles la gonade mâle se trouve dans un scrotum externe, tels que chez l'homme, le cheval, le porc et le chien. Chez l'homme, le terme de testicule non descendu ou ectopie testiculaire est préféré à celui de cryptorchidie.

#### *La fréquence chez l'homme et quelques espèces animales*

L'ectopie testiculaire représente le défaut le plus fréquent observé en pédiatrie chez le garçon (Care Effective health program, 2012a). Elle concerne 30% des prématurés, 3% des nouveau-nés à terme, et est maintenue chez 1% des nourrissons entre six mois et un an.

Chez les espèces domestiques, les fréquences les plus élevées sont observées chez le chien (Mattos et al., 2000). La fréquence varie de 1% chez les animaux croisés (Miller et al., 2004) à 10% chez certaines races pures (Hayes, 1986). Les races de petits chiens semblent plus affectées que les autres. Chez le porc, les fréquences sont de 0,37 et 0,41 selon les lignées étudiées d'après les dernières études publiées (Larzul et al., 2008; Walters, 2010).

#### *Les différentes formes cliniques*

La cryptorchidie peut être unilatérale ou bilatérale. La forme unilatérale est la plus fréquente. Elle survient dans 100% des cas de cryptorchidie chez le lapin, 80-90% chez l'homme, le chat, le bovin et le cheval, et dans 60% des cas chez le cochon et le mouton (Amann and Veeramachaneni, 2007). Une exception existe chez une population de chevreuils à queue noire, sur l'île de Kidiak en Alaska, où la fréquence de cryptorchides unilatéraux n'est que de 30% (Veeramachaneni et al., 2005), les 2 tiers des cas de cryptorchidie étant bilatéraux.

Chez l'homme, le testicule non descendu est plus fréquemment observé du côté gauche, et concerne 80% des cas de cryptorchidie unilatérale environ (Ghazzal, 2006). Chez le cochon (Warwick, 1926) et le cheval (Cox et al., 1979), la cryptorchidie unilatérale gauche semble également être la forme la plus fréquente. Chez les autres espèces, aucune différence n'a été observée. Le testicule non descendu peut se trouver sur son trajet normal de migration, on parlera de cryptorchidies vraies (Figure 39), ou en dehors du trajet de la migration, il s'agit de cryptorchidies ectopiques.

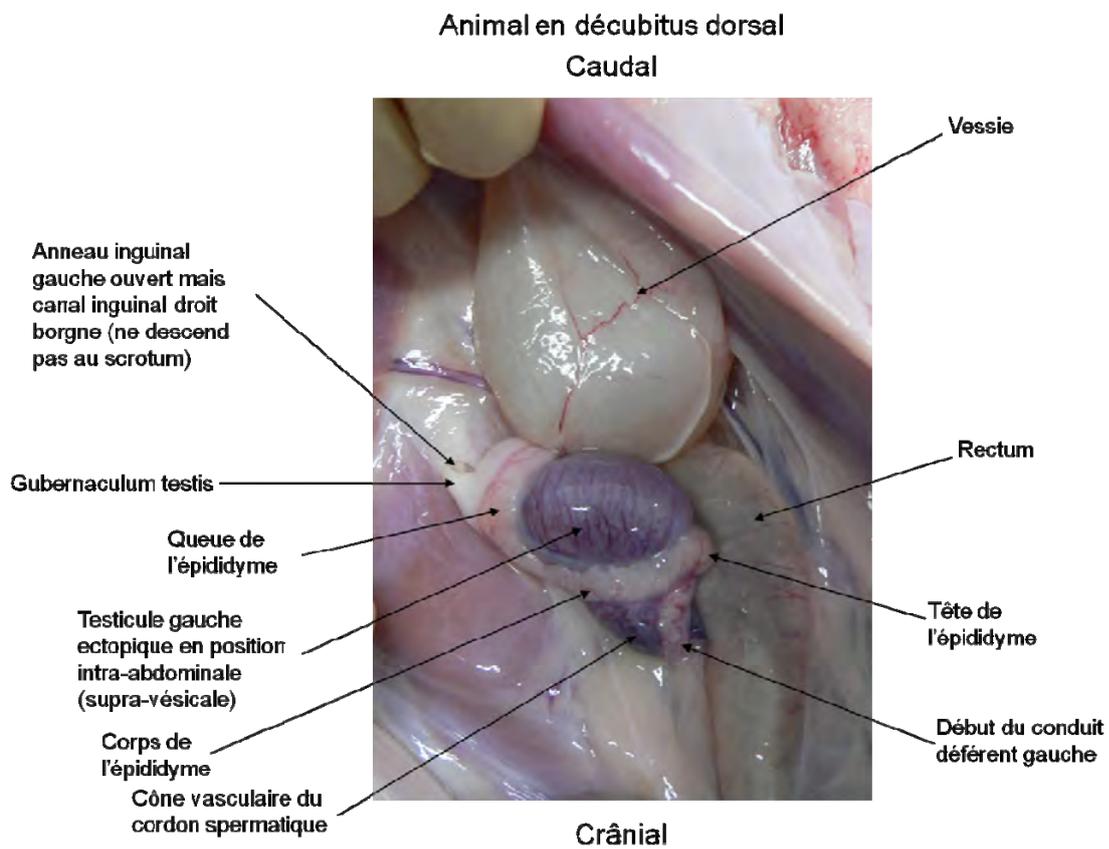


Figure 40 : Testicule en position abdominale chez un porcelet cryptorchide.

Dans les cas de cryptorchidies vraies, trois positions du testicule sont observées : au niveau de la cavité abdominale, du canal inguinal ou dans la partie supra scrotale (Figure 39). Chez l'homme, il est en position inguinale dans 60% des cas, la position abdominale du testicule étant la moins fréquente (8%) (Care Effective health program, 2012b). Chez le porc le testicule est plus souvent positionné au niveau abdominal (Figure 40). Chez le chien, 92% des cas de cryptorchidie sont abdominaux (Amann and Veeramachaneni, 2006). Chez le cheval, la position du testicule non descendu varie en fonction du coté considéré. Dans les cas de cryptorchidie droite, le testicule est en position abdominale dans la moitié de cas, avec une queue épидидymale descendue, alors que cette position ne concerne que 20% des cas à gauche (Cox et al., 1979). Chez le chien, la cryptorchidie droite est incomplète (Bladon, 2002). Chez le chevreuil à queue noire, les testicules sont dans la cavité abdominale, entre le rein et l'anneau inguinal interne et le gubernaculum n'est pas développé.

Le testicule peut être palpable (80% des cas chez l'homme) ou non palpable (20%). Dans les cas où le testicule est non palpable, l'exploration chirurgicale chez l'homme révèle un testicule absent pour 20 à 40% des patients. Lorsque le testicule est palpable, le traitement consiste à libérer le cordon testiculaire, à ligaturer le processus peritoneovaginal et à fixer le testicule dans le scrotum.

La cryptorchidie n'est pas une anomalie grave, cependant le risque de développer une tumeur testiculaire chez l'enfant cryptorchide est cinq à dix fois plus important que chez l'enfant sain. De plus, elle peut être responsable d'infertilité. La position du testicule dans le scrotum est primordiale car elle permet au testicule d'être à une température optimale pour assurer la spermatogénèse. 50% des individus atteints de cryptorchidie unilatérale présentent des défauts de spermatogénèse, contre 80% dans les cas de cryptorchidie bilatérale (Hadziselimovic et al., 2007). Une élévation de la température induirait une déplétion des cellules germinales. En effet, la température corporelle est de 37°C, comparée à une température optimale de 32°C pour les cellules germinales dans le scrotum (Ferguson and Agoulnik, 2013).



## *Les facteurs de risque connus et suspectés*

### *La prématurité*

De nombreuses observations indiquent que la prévalence de la cryptorchidie est augmentée lorsque la durée de la grossesse diminue (inférieure ou égale à 37 semaines chez l'homme) (Jensen et al., 2012). Elle est multipliée par dix lorsque le poids à la naissance est inférieur ou égal à 2,75 kg. La naissance de jumeaux (Zakaria et al., 2012) et la fécondation *in vitro*, facteurs de risque de la prénatalité, induiraient ainsi des défauts de descente testiculaire (Ghirri et al., 2002). L'hypothèse émise est qu'une diminution de la durée de la vie foetale réduit le temps disponible pour que le testicule se mette en place.

### *Les perturbations endocriniennes*

L'exposition des parents ou du fœtus à différentes substances augmenterait la prévalence de la maladie, via une action directe ou indirecte sur le système endocrinien.

Il peut s'agir d'un traitement médicamenteux de la mère, tel que les stéroïdes (Wang and Wang, 2002) ou l'antihistaminique cimétidine (Takeshi, Kai, & Suita, 2002).

Des substances comme l'alcool ou le glucose pourraient avoir un effet indirect sur la prévalence de la cryptorchidie. La consommation d'alcool par la mère enceinte multiplie le risque de cryptorchidie par trois (Damgaard et al., 2007). Des dysfonctionnements du métabolisme glucidique chez la mère, tels que le diabète gestationnel ou l'intolérance au glucose, triple également le risque de cryptorchidie chez le nouveau né par rapport à la normale (Virtanen et al., 2006). Des travaux ont montré que ces deux substances étaient associées à des perturbations endocriniennes (Fraietta et al., 2013; Kozakowski and Zgliczyński, 2013).

Androgènes et estrogènes semblent impliqués dans la prévalence de la maladie. Le nombre de cas de porcelets cryptorchides augmente lorsque les truies gestantes sont exposées aux estrogènes (Husmann and Levy, 1995). L'exposition chez le rat (van der Schoot, 1992) et chez le porc (Kopera et al., 2010) à un antagoniste des androgènes, le flutamide, est également associée à l'augmentation du nombre de cas.

Enfin, l'exposition des parents à certains pesticides, qui pourraient avec une action antagoniste des androgènes (Orton et al., 2012), multiplie le risque du testicule non descendu chez l'enfant par treize (Wang and Wang, 2002).



Globalement, il a été constaté que la fréquence de la cryptorchidie était en augmentation chez l'homme (Ferguson and Agoulnik, 2013) et chez le chien (Mattos et al., 2000). D'après les estimations, elle a été multipliée par deux à trois depuis cinquante ans.

L'exposition des parents à de nouveaux facteurs environnementaux, tels que les pesticides, pourrait expliquer cette augmentation. En 2001, une étude française a été menée suite aux observations de l'augmentation du nombre de cas de cryptorchidies et d'hypospadias (anomalies du pénis et de l'urètre) depuis deux ans chez les familles d'agriculteurs ou chez des individus potentiellement exposés aux pesticides (Pascal et al., 2001). Les résultats n'ont pas été concluants concernant l'association entre la prévalence des cas et l'exposition à des substances particulières. D'une part, il est difficile de lister exhaustivement les substances auxquelles les individus sont exposés, et d'évaluer le niveau d'exposition. D'autre part, l'augmentation du nombre de cas observés peut être expliquée par l'augmentation du nombre d'opérations (Paty et al., 2012). A ce jour, les études sont toujours en cours.

Au delà des facteurs environnementaux, les études ont montré que l'étiologie de la cryptorchidie était multifactorielle, et certains gènes pourraient être impliqués dans le déterminisme de la maladie.

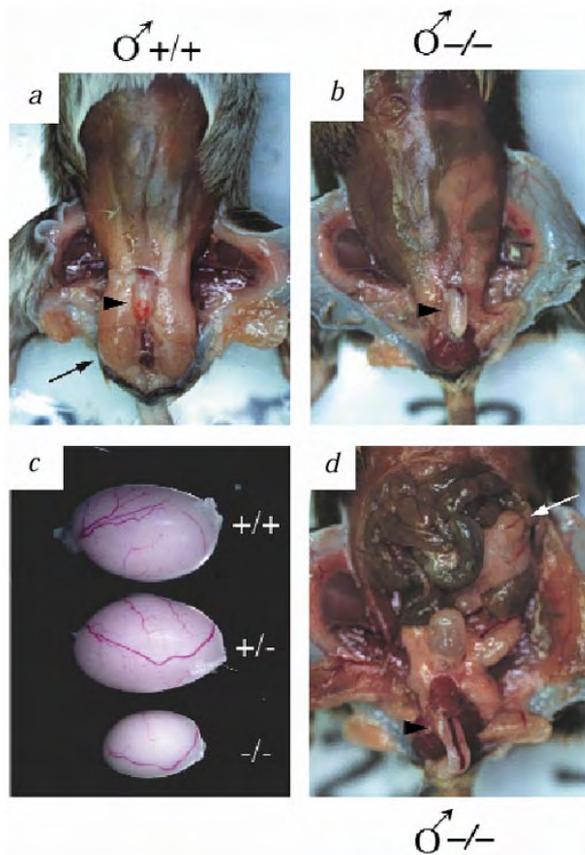


Figure 41 : Cryptorchidie chez la souris KO pour Ins13 (Nef and Parada, 1999).  
 a. localisation externe des testicules (flèches) chez le mâle normal ; b. Absence de testicules en position externe chez le mâles Ins13<sup>-/-</sup> ; c. Dissection testiculaire chez le mâle normal Ins13<sup>+/+</sup>, le mâle hétérozygote Ins13<sup>+/-</sup> et le mâle homozygote Ins13<sup>-/-</sup> ; d. localisation intra-abdominale des testicules (flèche blanche) chez l'animal Ins13<sup>-/-</sup>. Le triangle noir indique le pénis.

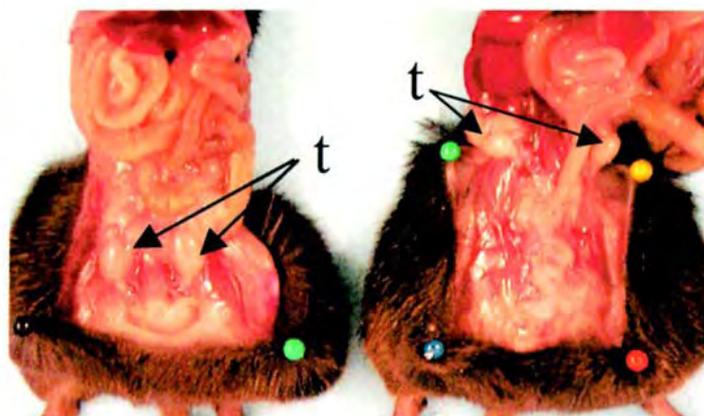


Figure 42 : Cryptorchidie chez la souris KO pour Rxfp2 (Gorlov et al., 2002) à droite, comparé à la souris normale à gauche. Les testicules chez le mâle mutant Rxfp2<sup>-/-</sup> sont en position intra-abdominale, près des reins. Les testicules sont indiqués par (t).

### *La pathogénie*

Physiologiquement, la descente testiculaire est régie par deux phases : la phase trans-abdominale, où le gubernaculum rétrécit sous l'action de l'Insl3, et la phase inguino-scrotale, contrôlée par la testostérone. Des défauts de l'une ou l'autre de ces voies, aux niveaux biochimiques et génétiques, ont été identifiés comme associés à un défaut de migration testiculaire.

#### *La phase trans-abdominale et la voie INSL3*

Des études biochimiques et génétiques du gène INSL3, gène majeur de la phase trans-abdominale de la descente testiculaire, ainsi que de son récepteur RXFP2 (*relaxin/insulin-like family peptide receptor 2*) ont été réalisées.

Au niveau biochimique, la concentration de INSL3 est diminuée dans les cas de cryptorchidie bilatérale chez le chien (Pathirana et al., 2012). Sur le plan génétique, l'inactivation de Insl3 chez la souris induit une cryptorchidie bilatérale et une infertilité chez le mâle, tandis que sa surexpression chez la femelle entraîne la descente des ovaires (Nef and Parada, 1999) (Figure 41). Lorsque Insl3 est surexprimé au niveau du pancréas chez ces mêmes souris KO Insl3, aucun défaut de descente testiculaire n'est observé (Adham and Agoulnik, 2004).

Le récepteur de INSL3, RXFP2 (*relaxin/insulin-like family peptide receptor 2*) est fortement exprimé au niveau du gubernaculum (Bogatcheva et al., 2003). Le KO chez la souris du gène codant pour ce récepteur est responsable de cryptorchidie et d'absence de spermatogénèse (Gorlov et al., 2002) (Figure 42).

Chez l'homme, bien que des mutations dans la séquence des gènes INSL3 et RXFP2 aient été identifiées (Bogatcheva et al., 2007), elles ne sont retrouvées que dans 5% des cas de cryptorchidie (1,8% pour INSL3, et 2,9% pour RXFP2 (Ferlin et al., 2006; El Houate et al., 2007)).

D'autres mécanismes, impactant sur le gubernaculum peuvent également être associés à la maladie. En effet, il est établi que chez les patients atteints de cryptorchidie, le gubernaculum n'est pas correctement fixé au scrotum (Fiegel et al., 2011). Chez ces animaux cryptorchides, des altérations du gubernaculum, notamment des défauts de la régression

gubernaculaire peuvent expliquer les anomalies de descente testiculaire (McMahon et al., 1995).

### *La phase inguino-scrotale et la voie des androgènes*

Au niveau biochimique, les concentrations en testostérone sont diminuées dans les cas de cryptorchidie bilatérale chez le chien (Pathirana et al., 2012). Chez l'homme, l'ectopie testiculaire est fréquemment observée dans des pathologies liées à des défauts de la voie des androgènes, tels que le syndrome d'hypogonadisme hypogonadotrope, ou l'insensibilité complète aux androgènes (Giannopoulos et al., 2001; Quigley et al., 1995).

Au niveau génétique, des mutations du gène codant pour le récepteur aux androgènes induisent des anomalies du tractus génital et une cryptorchidie bilatérale (Chang et al., 2013). Chez l'homme, plus de 400 mutations du gène du récepteur aux androgènes ont été identifiées dans les cas de syndrome d'insensibilité aux androgènes, et la cryptorchidie y est fréquemment observée (Galani et al., 2008). Une cryptorchidie peut être induite chez la souris par invalidation du récepteur à la LH, et corrigée par injection de testostérone (Yuan et al., 2006). Enfin, des ectopies testiculaires ont été observées en présence de mutations de la gonadolibérine GnRH (Grocock et al., 1988).

### *Les autres voies suspectées*

Des cas d'ectopie testiculaire ont été relevés dans le cadre d'études biochimiques ou d'analyse de KO chez la souris. Cependant, aucune mutation candidate à la cryptorchidie chez l'homme ou d'autres espèces n'a été identifiée jusqu'à présent. Foresta et al. ont rapporté qu'une déficience en hormone anti müllérienne (Foresta et al., 2008) induirait un défaut de descente testiculaire. Le facteur de transcription Hoxa10, impliqué dans le développement de la fonction reproductive (Ivell, 2003), semble également impliqué dans cette anomalie. Des KO chez la souris du gène codant pour HOXA10 entraînent une ectopie testiculaire bilatérale et un gubernaculum vestigial. Enfin, des KO du récepteur 1 aux estrogènes (ESR1) chez la souris induisent une rétraction des gonades.

Tous ces résultats soulignent la complexité de ce caractère multifactoriel. Chez le porc, aucune donnée biochimique ou génétique n'est disponible. Seules des concentrations anormales de calcium ont été enregistrées chez les porcelets atteints de cryptorchidie (Beuermann et al., 2009) et sont abordés dans le paragraphe suivant. Il est en effet établi que cryptorchidie et hernie scrotale pourraient être corrélés et induits par des voies similaires, comme la voie calcique.

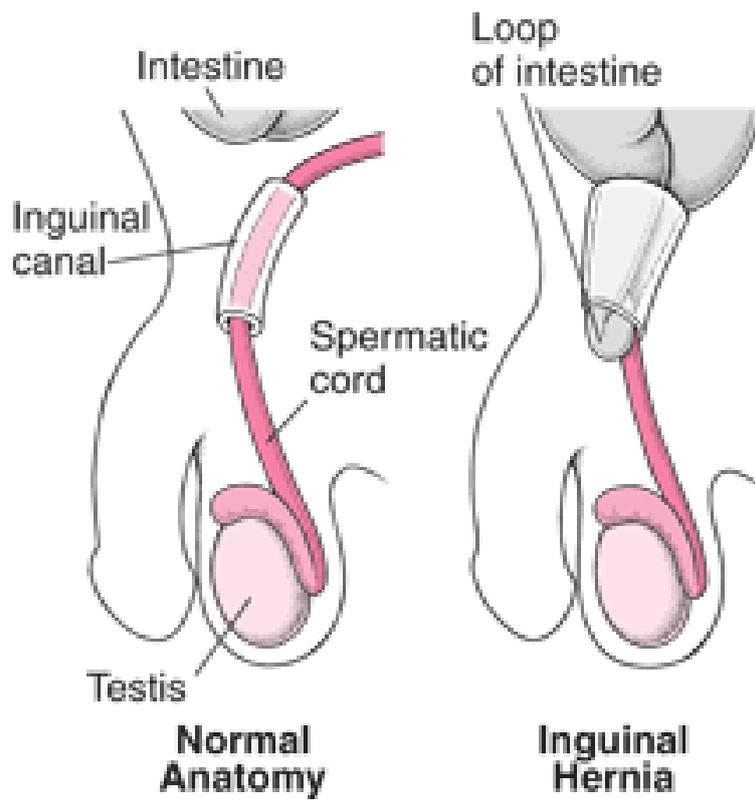


Figure 43 : Hernie inguinale chez l'homme

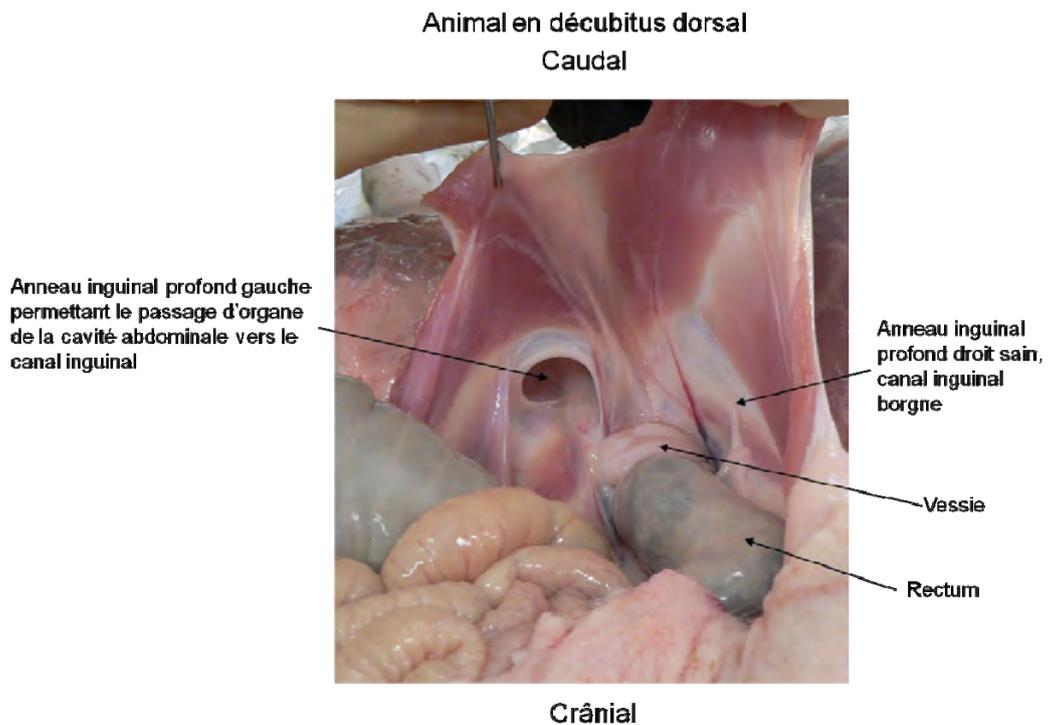


Figure 44 : Hernie inguinale gauche chez un porcelet

### III.C.3. La hernie inguino-scrotale

La hernie inguino-scrotale correspond au passage d'organes abdominaux (principalement omentum et intestin grêle) le long du cordon spermatique dans le scrotum via l'anneau inguinal (Chow et al., 2008) (Figure 43). Lorsque le contenu abdominal est situé dans le scrotum, on parle de hernie scrotale. Lorsqu'il est positionné au niveau du canal inguinal, la hernie est inguinale. La hernie scrotale n'est observée que chez les mâles, alors que la hernie inguinale peut survenir chez les mâles et les femelles. En effet il existe une zone de faiblesse dans la paroi abdominale qui est constituée par l'anneau inguinal profond. Chez la femelle, le péritoine et le *fascia transversalis* passent tangentiellement à l'anneau inguinal profond, amenant ainsi à son recouvrement. En revanche, chez le mâle, l'anneau inguinal profond n'est recouvert ni par le péritoine ni par le *fascia transversalis* du fait de la descente testiculaire au cours du développement. Ceci explique pourquoi la fréquence des hernies inguinales est beaucoup plus importante chez les mâles que chez les femelles.

La fréquence des hernies chez l'homme varie de 4 à 27% selon les études. Il est important de distinguer les hernies congénitales des hernies acquises, qui surviennent à l'âge adulte. La hernie congénitale concerne 1 à 5% des nouveaux nés, et jusqu'à 10% des prématurés. Chez le porc, la fréquence de la hernie est de 0,4 à 1%.

#### *La physiopathologie*

La hernie peut être uni ou bilatérale. La Figure 44 présente le cas d'une hernie inguinale unilatérale gauche chez le porc. Chez l'homme, la hernie est localisée à droite dans 51,6% des cas, à gauche dans 34,4% des cas. Elle est bilatérale pour 14% des patients atteints de hernie. Il a été montré que le testicule droit descendait plus tardivement que le gauche, le processus vaginalis est donc oblitéré plus tard à droite qu'à gauche, ce qui pourrait expliquer la différence de fréquence entre hernies droites et gauches.

On distingue deux types d'hernies inguinales, selon le passage dans la zone médiale ou latérale des organes herniés aux vaisseaux épigastriques inférieurs (Jenkins and O'Dwyer, 2008). Ces hernies sont qualifiées de directes ou indirectes, selon la zone médiale ou latérale, respectivement. Enfin, la hernie peut être réductible ou non, selon la possibilité de repousser les organes herniés par palpation ou non. Dans le cas de la hernie irréductible, on observe une adhérence des organes herniés au sac herniaire (Desarda, 2003).

La hernie peut conduire à des complications parfois sévères, telles que l'étranglement de l'intestin, les dysfonctions gonadiques, les nécroses et perforations intestinales, les occlusions intestinales, l'atrophie ou la nécrose testiculaire (Jenkins and O'Dwyer, 2008). Le risque d'incarcération est de plus de 60% chez les prématurés avec hernie.

Les facteurs de risque identifiés sont la prématurité (durée de gestation et poids à la naissance) et les anomalies urogénitales (cryptorchidie, l'hypospadias, l'ambigüité sexuelle) (Ghazzal, 2006). Lau et al. ont montré que le risque pour qu'un individu développe une hernie inguinale était multiplié par huit s'il avait dans sa famille un cas de hernie (Laus, 1984).

### *La pathogénie*

Il existe peu de données sur la pathogénie de la hernie inguino-scrotale congénitale, d'une part car elle est moins fréquente que la hernie acquise, d'autre part car la hernie acquise est traitée sans difficulté. En 2003, deux tiers des hernies traitées aux Etats-Unis (un million d'opérations de hernies au total) correspondaient à des hernies inguino-scrotales (Matthews and Neumayer, 2008). C'est pourquoi la majorité des travaux sur la hernie inguino-scrotale concerne le traitement de la hernie acquise.

Les quelques études biochimiques et génétiques sur la hernie inguino-scrotale congénitale portent sur la persistance du processus peritoneovaginal, impliqué dans la migration testiculaire, et/ou des faiblesses de la paroi abdominale constituée par l'anneau inguinal profond (Rodrigues Junior et al., 2002).

### *La paroi abdominale et la voie du collagène*

La voie du collagène joue un rôle majeur dans le maintien des tissus et la matrice extracellulaire. Des dysfonctionnements de la matrice du collagène pourraient être impliqués dans le déterminisme de la hernie inguinale. Chez l'homme, le ratio entre le collagène de type I, résistant et élastique, et le collagène de type III, impliqué principalement dans la cicatrisation et moins élastique, est important pour le maintien des tissus. Il a été montré, chez les patients atteints de hernie inguinale, que la quantité de collagène de type III était augmentée (Meyer et al., 2007; Rosch et al., 2003).

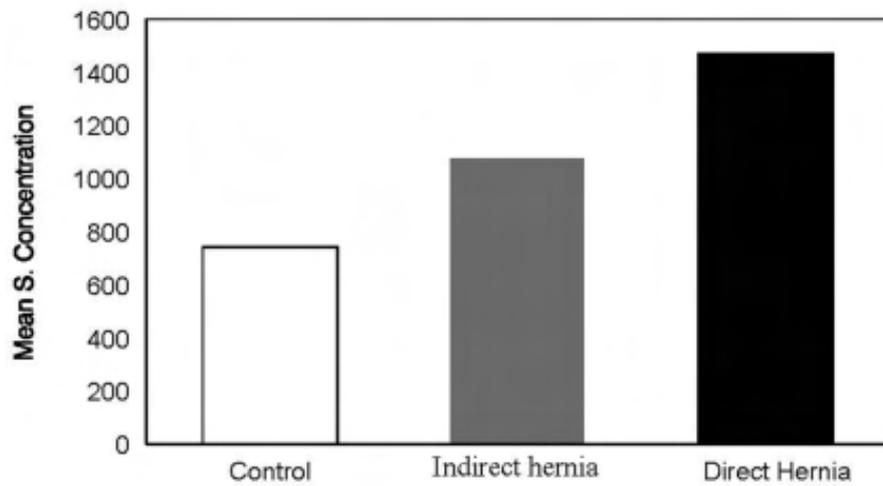


Figure 45 : Concentration sérique de MMP2 chez des contrôles et des cas de hernie inguinale indirecte et directe (Jain et al., 2009)

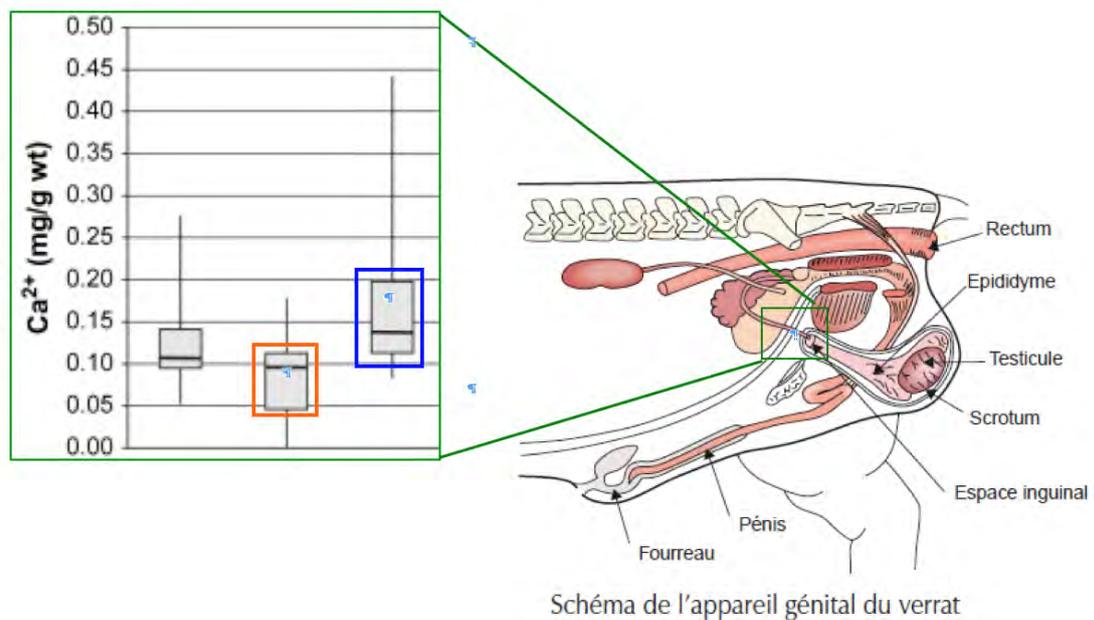


Figure 46 : Concentrations calciques chez des porcelets

normaux, atteints de hernie scrotale (orange) et de cryptorchidie (bleu)

Les enzymes de dégradation du collagène, les *Matrix metallo-proteinases* ou MMP, pourraient également être associées à la faiblesse des tissus. Le niveau sérique de MMP2 est augmenté chez les individus atteints de hernie inguinale (Jain et al., 2009) (Figure 45). L'étude du niveau d'expression de MMP2 a permis de confirmer ces résultats (Aren & al., 2011), et de montrer que MMP1 et MMP9 étaient également surexprimées dans les cas de hernies inguinales.

L'activité des MMP est contrôlée par les TIMP (*tissue inhibitors of metalloproteinases*), inhibiteurs physiologiques qui contribuent à l'homéostasie tissulaire. Une étude récente a montré que les concentrations de TIMP1 et TIMP2 étaient réduites dans les tissus herniaires (Antonioni et al., 2011). Dans cette même étude, les concentrations tissulaires de MMP2 et MMP9 sont augmentées chez des malades par rapport aux contrôles, ce qui est en accord avec les résultats des études précédentes.

Enfin, une étude a montré que la forme active de TGF- $\beta$ 1 (pour *Tumor Growth Factor- $\beta$* ) était surexprimée chez des cas de hernie inguinale (Pascual et al., 2007). Cette cytokine, impliquée dans le remodelage de la matrice extracellulaire, est capable d'activer MMP2 et MMP9 et d'inhiber MMP1 et MMP3.

#### *Le processus peritoneovaginal et la voie calcique*

Le calcium est impliqué dans l'apoptose cellulaire. C'est par ce mécanisme que le canal inguinal s'oblitére pour ne devenir qu'un ligament péritonéo-vaginal après le passage du testicule lors de la descente testiculaire.

Deux études publiées ont porté sur les concentrations calciques chez les individus atteints de hernie inguino-scrotale. La première étude, réalisée chez l'homme, a permis de montrer que les concentrations calciques étaient diminuées dans le sac herniaire d'individus atteints de hernie par rapport à ceux atteints de cryptorchidie ou hydrocèle (Tanyel et al., 2003).

La deuxième étude, chez le porc, montre que les concentrations calciques sont également diminuées au niveau du sac herniaire et du muscle crémaster chez les porcelets atteints de hernies (Beuermann et al., 2009) (Figure 46). A l'inverse, les concentrations calciques chez les animaux cryptorchides sont supérieures à la normale.



Ces travaux montrent que les gènes candidats à la hernie inguino-scrotale sont nombreux. Au niveau biochimique, il n'existe pas d'autres études publiées chez le cochon. Cependant, l'impact économique majeur de la hernie scrotale dans les élevages porcins a motivé la recherche de QTL influençant le déterminisme de la hernie. L'objectif est de diminuer la prévalence de la maladie dans les élevages par sélection assistée par marqueurs.

### *Les données génétiques disponibles chez le porc*

Deux études de primo-localisation de QTL ont été réalisées par des équipes norvégienne (Grindflek et al., 2006) et chinoise (Ding et al., 2009). Les animaux utilisés pour les études sont des familles composées de paires de germains atteints (ou ASP pour *Affected Sib Pair*), et les marqueurs sont des microsatellites balisant l'ensemble du génome porcine. Deux tests statistiques ont été combinés pour chacune des analyses : des tests de liaison non paramétriques, et des analyses TDT.

L'étude norvégienne a porté sur 103 ASP de race Landrace et d'origine danoise et Norvégienne, et 137 microsatellites (Grindflek et al., 2006). Elle a permis de mettre en évidence huit QTL significatifs communs à l'analyse de liaison et au TDT ( $p < 0.05$ ), ou très significatifs pour l'une ou l'autre des méthodes ( $p < 0.006$ ). Ces QTL sont localisés sur les chromosomes 1, 2, 5, 6, 15, 17 et X. Les intervalles de ces QTL comprennent un à cinq marqueurs significatifs, et leur taille peut atteindre 40 cM.

L'étude chinoise a été réalisée avec 10 ASP sino-européennes (Erhualian x duroc blanc) et 194 marqueurs (microsatellites et SNP) (Ding et al., 2009). Onze régions chromosomiques de un à quinze marqueurs ont été mises en évidence avec l'analyse de liaison (SSC2, 4, 8, 13, et 16). Quatre marqueurs isolés sont significatifs d'après l'analyse TDT (SSC2, 8, 10, et 18). La plus grande région est celle du chromosome 13, et mesure 50cM. Un seul marqueur localisé sur le SSC8 est significatif pour les deux analyses statistiques.

Seules les régions identifiées sur les chromosomes 2 et 13 sont communes aux études norvégienne et chinoise. Les différences obtenues peuvent être expliquées par un déterminisme génétique race-spécifique, et/ou une densité de marqueurs trop faible n'ayant pas permis de trouver d'autres QTL en commun. De plus, le nombre de familles utilisées dans l'étude chinoise est très faible.

Pour chacune des études, les résultats obtenus entre les deux tests statistiques diffèrent. Sur l'ensemble des deux études, 40 marqueurs ont été trouvés significatifs avec l'étude de liaison, contre une dizaine seulement avec le TDT. L'analyse de liaison teste uniquement la liaison entre les marqueurs et l'anomalie au sein des familles, alors que le TDT teste la liaison et l'association à l'échelle de la population ; si le DL entre la mutation et un des allèles au(x) marqueur(s) microsatellite(s) est faible voir nul, aucun résultat significatif ne pourra être obtenu par TDT malgré l'existence d'une liaison génétique.

Une autre étude, basée sur la recherche de segments chromosomiques homozygotes et IBD, a été réalisée par une équipe allemande (Knorr et al., 2006). La première étape de ce travail a été la recherche des fragments chromosomiques identiques entre descendants atteints via une approche de biologie moléculaire (GMS pour *Genome Mismatch Scanning*). Des fragments candidats localisés sur les chromosomes 1, 3, 4, 7, 10, 12, 14 et 15 ont ainsi été obtenus. Dans un second temps, une analyse de liaison non-paramétrique a été réalisée à partir d'une population de porcs allemands (Piétrain, Large White, Landrace et Hampshire), et 139 marqueurs (microsatellites et SNP), répartis sur les 18 autosomes porcins. Des QTL sur les chromosomes 3, 6, 7, 12 et 15 ont été identifiés, mais seule la région du chromosome 3 était commune avec quelques fragments identifiés par GMS. Suite à ces résultats de primo-localisation, une étude fonctionnelle de deux gènes candidats, localisés dans l'intervalle du QTL du chromosome 3, a été réalisée (Beck et al., 2006a, 2006b). Les deux gènes étudiés sont GUSB, associé à des défauts de matrice extracellulaire, et SERCA1, impliqué dans l'apoptose des cellules musculaires lisses du gubernaculum. Cependant, aucune association significative entre des polymorphismes de ces gènes et le statut à la maladie n'a pu être identifiée.

Une étude plus complète comprenant les étapes de primo-localisation, validation des régions candidates et identification de gènes candidats associés à la hernie scrotale ont permis d'identifier certains gènes : dans un premier temps une étude de liaison avec 34 familles de Duroc blancs et des études d'associations cas/contrôles ont permis d'identifier des QTL sur SSC2 et SSC12 (Du et al., 2004). Dans un second temps, ces régions ont été confirmées grâce à une étude comparée de 421 atteints et de 100 témoins de Duroc, Landrace, Large White, et des animaux croisés (comprenant du piétrain) (Du et al., 2009). Enfin des SNP ont été identifiés dans 15 gènes localisés dans les intervalles des QTL de SSC2 et SSC12.

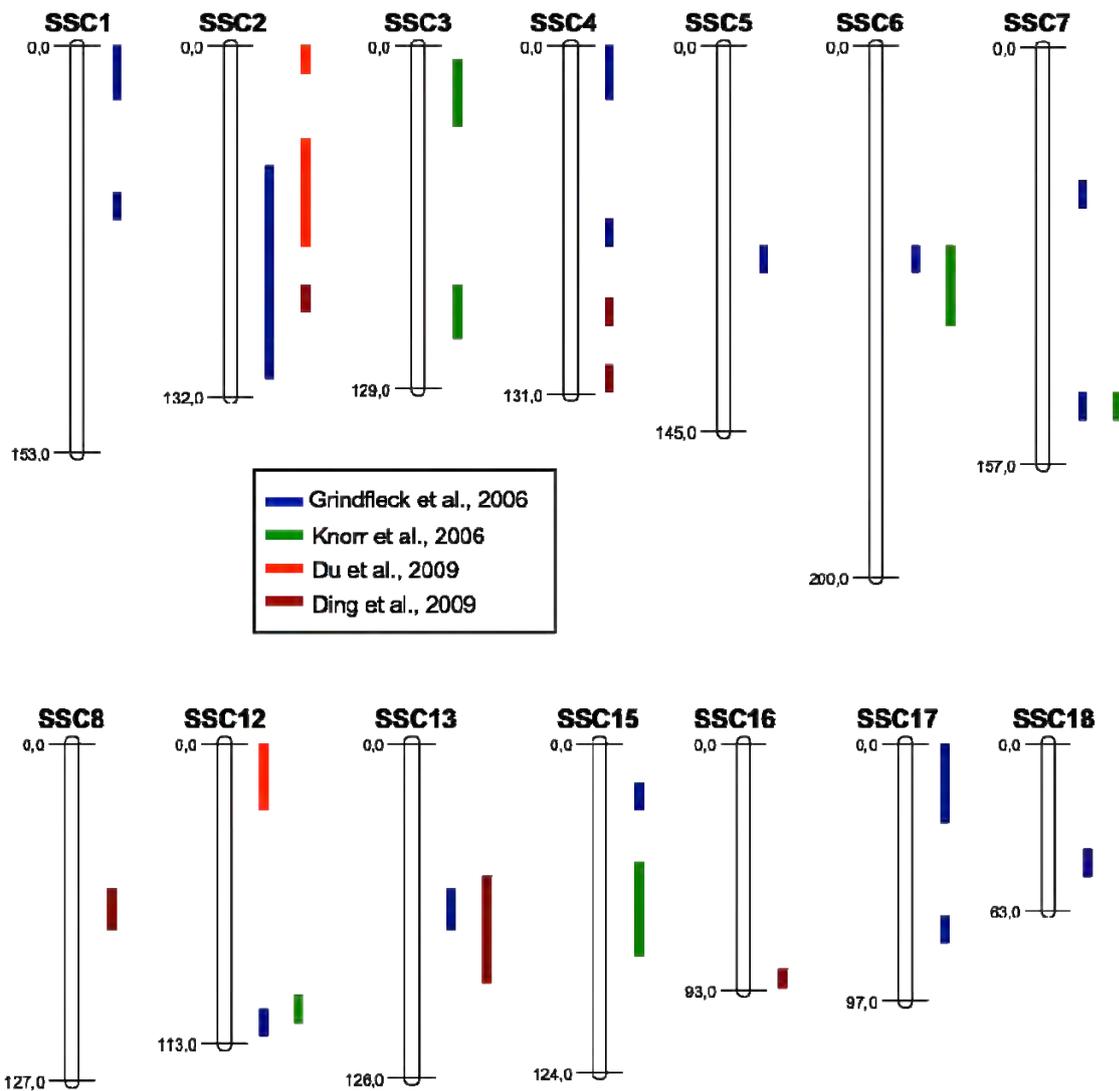


Figure 47 : Localisation des QTL identifiés chez le porc d'après quatre études de cartographie

Trois gènes sur le chromosome 2 et un gène sur le chromosome 12 ont été mis en évidence par tests d'association cas-contrôles et TDT ( $p$ -value  $< 0.05$ ). Il s'agit d'ELF5 (transition épithélio-mésenchymateuse), KIF18A et COL23A1 (métabolisme du collagène) et NPTX1 (voie de signalisation des récepteurs d'œstrogènes). (Du et al., 2009).

Enfin, une étude basée sur une approche gène-candidat a également été publiée. Des tests d'association entre des marqueurs identifiés dans des gènes choisis comme candidats et le statut à la maladie ont été réalisés avec 1534 animaux Piétrain. Trois gènes ont ainsi été retenus comme candidats : HOXA10 et MMP2 (métabolisme du collagène) et ZFPM2 (important dans le développement de la hernie diaphragmatique congénitale) ( $p$ -value  $< 0.01$ ) (Zhao et al., 2009). Aucun des gènes identifiés dans l'étude de Du et al n'a été retrouvé dans cette étude, bien que la même race ait été étudiée.

L'ensemble des QTL mis en évidence dans les différentes études est présenté dans la Figure 47. Tous ces résultats soulignent la complexité et le caractère polygénique de la maladie. Malgré l'identification de plusieurs gènes candidats, aucun mécanisme moléculaire n'a été caractérisé jusqu'à présent.

L'utilisation de la puce porcine haute densité ouvre de nouvelles perspectives de recherche de QTL pour ce type de maladie complexe. Un des objectifs de la thèse est d'identifier des variants génétiques associés à la hernie inguino-scrotale dans les populations d'élevage en France.



# Matériel et Méthodes

Mix réactionnel	Conditions PCR
<b>Mix PCR (12 µl) :</b> - 1 µM 5X Green GoTaq® Reaction Buffer1 - 0.5 µM de chaque amorce - 200 µM de dNTP - 0.05 U de l'enzyme GoTaq® DNA Polymerase - Promega	<b>Dénaturation initiale :</b> 5 minutes à 94°C
	<b>35 cycles de :</b> - dénaturation : 30 secondes à 94°C - hybridation : 30 secondes à 60°C - élongation : 30 secondes à 72°C
<b>ADN :</b> 3 µl à 10 ng/µl	<b>Elongation finale :</b> 10 minutes à 72°C

Tableau 11 : Composition du mélange réactionnel et conditions d'amplification PCR utilisés pour diagnostiquer la présence ou non du locus SRY

Marqueur	Chr	All	Tailles Min./Max.	Fluo	Amorce sens	Amorce antisens	Temp x
BMP5	7	9	227,0/252,4	6Fam	ACAGCAAAAAGCAAAGACCA	GTAAGTAGGTTGTCAGGCTTGC	58°C
MCS10F15SP6	7	16	210,6/255,4	Hex	GGTCCACACTGAATGGGTTT	GTGCTTCTCTGCCAAGATT	55°C
MCS455C8A	1	11	144,3/192,1	Ned	TCTGAAGATTTGGTGGCATAAC	CAACACCCTGAAGCTGGAA	55°C
MCSE264C2A	1	9	335,1/382,1	Hex	TCCTTTGCCCTGAGCTCTC	CACCCATCACTCTGTTGCTC	58°C
MCSI234B9B	2	12	316,2/353,0	6Fam	TAAGTCACAGTTGCCGCTTG	GGAAAGCATAGTTTTTCCATCAA	58°C
TNFB	7	11	172,4/221,3	6Fam	CTGGTCAGCCACCAAGATTT	GGAAATGAGAAGTGTGGAGACC	58°C
S0121	6	12	163,3/193,1	Hex	TTGTACAATCCCAGTGAATCC	AATAGGGCATGAGGGTGTGTTGA	55°C
S0141	2	12	208,8/233,9	Ned	GATCTGGTCTGTCTGTGTCTCCT	AGACCCCAACTCTTGGTCTCAT	50°C
SW492	6	13	111,6/146,0	Hex	TCCATCAGCTCACATAGTTAGC	ACCATGACAGGAAGTCCGAG	55°C
SW71	6	9	88,8/119,4	Ned	GATCACCTTATCCCCATTC	TAGAAACACCATCATCCATTCA	55°C
SW969	4	12	105,4/139,2	6Fam	AGCCTGGAACATTTTTGAGTG	TTTCAATTGGTTCCTGTGTCC	55°C

Tableau 12 : Microsatellites et amorces utilisées pour les tests de parenté. Chr : chromosome, All : nombre d'allèles identifiés, Fluo : fluorochrome, Temp x : température d'hybridation

# Matériel et Méthodes

---

## I. Génotypage

### I.A. Test diagnostique "Sry"

**Contexte :** Les mâles sont porteurs d'un chromosome X et d'un chromosome Y, alors que les femelles possèdent deux copies de chromosome X. Le gène SRY, localisé sur le bras court du chromosome Y est un des gènes majeurs de la différenciation sexuelle mâle et sa présence est spécifique des mâles. La translocation d'un fragment chromosomique portant SRY peut être la cause de certains cas d'intersexualité chez des individus XX. Un des objectifs de cette thèse étant d'identifier d'autres mécanismes moléculaires impliqués dans l'intersexualité chez le porc, un test diagnostique a été utilisé afin de contrôler l'absence de SRY chez les femelles intersexuées de la collection.

**Protocole :** Le contrôle de l'absence (ou de la présence) de SRY a été réalisé par amplification PCR d'un fragment du gène (références : Sscrofa10.2:Y: (279587-280297, NC\_010462.2). Un couple d'amorces (Sens : CTGTAGCCTCTGTGCCTCCT / Antisens : TTTCATTGTGTGGTCTCGTG) a été défini afin de cibler et d'amplifier spécifiquement 150 pb du gène. Les conditions d'amplification PCR sont rapportées dans le Tableau 11.

**Résultats :** Les produits d'amplification PCR sont analysés par électrophorèse sur gel d'agarose à 3%, et visualisés par marquage au bromure d'éthidium. La taille du fragment est vérifiée grâce au marqueur de taille (100pb ladder Eurogentec). La présence ou l'absence d'un fragment de 150pb est interprétée comme reflétant la présence ou l'absence de SRY. Lorsque le gène SRY est identifié chez un animal noté femelle intersexuée par l'éleveur, l'animal n'est pas retenu pour les analyses ultérieures.

### I.B. Contrôle de filiation

**Contexte :** Les études génétiques réalisées dans cette thèse sont basées sur les déséquilibres de transmission allélique au sein de familles. Les informations généalogiques transmises par les éleveurs ont donc été contrôlées avant de sélectionner les familles pour le GWAS : un contrôle de filiation a été réalisé pour toutes les familles d'animaux enregistrées dans la base de données, quelque soit la maladie.

Mix réactionnel	Conditions PCR
<b>Mix PCR (12 µl) :</b> - 1 µM 5X Green GoTaq® Reaction Buffer1 - 0.25 µM de chaque amorce - 200 µM de dNTP - 0.25 U de l'enzyme GoTaq® DNA Polymerase - Promega	<b>Dénaturation initiale :</b> 5 minutes à 94°C
	<b>x cycles de :</b> - dénaturation : 30 secondes à 94°C - hybridation : 30 secondes à 60°C - élongation : 30 secondes à 72°C
<b>ADN :</b> 3 µl à 10 ng/µl	<b>Elongation finale :</b> 3 minutes à 72°C

Tableau 13 : Composition du mix et conditions PCR du génotypage pour les tests de parenté

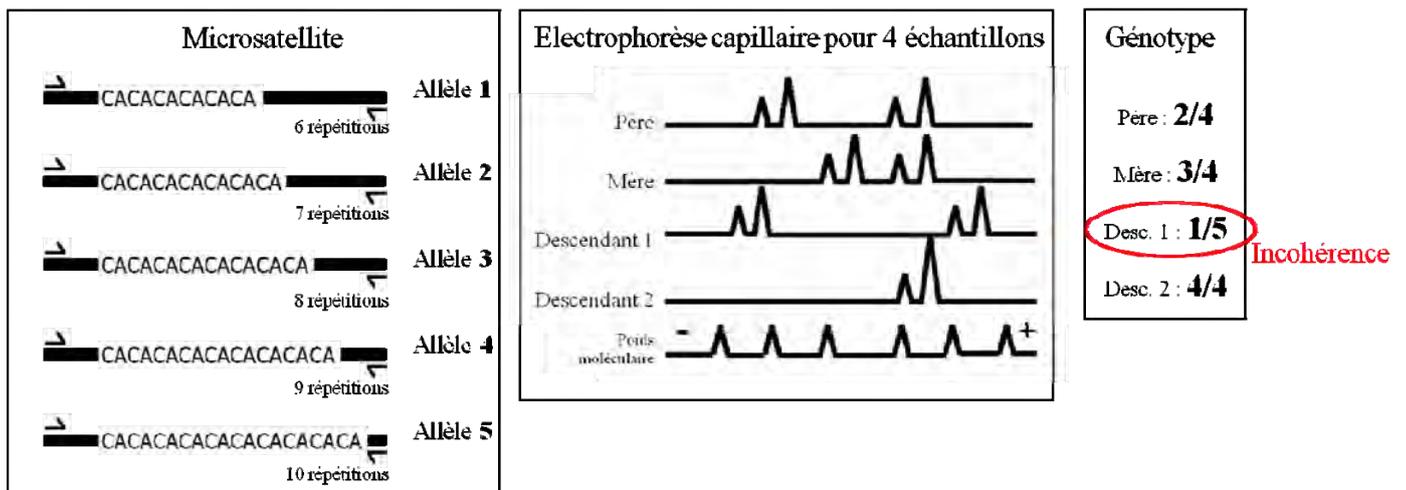


Figure 48 : Identification du génotype par électrophorèse capillaire. Exemple avec 5 allèles d'un microsatellite, 4 échantillons et une incohérence chez un descendant

**Principe :** Un jeu de 11 microsatellites (Tableau 12), présentant une forte informativité (9 à 16 allèles) et localisés sur des chromosomes différents a été constitué. Chaque marqueur est amplifié individuellement sur l'ADN de chaque animal, à partir d'un couple d'amorce dont une est marquée avec un fluorophore (6-FAM, HEX ou NED).

Pour chaque animal, les produits d'amplifications des 11 microsatellites sont poolés (*multi-loading*) avant migration sur gel ce qui permet de définir précisément la taille des amplicons. En fonction de la taille du fragment analysé et de la fluorescence émise, chaque fragment est assigné à un marqueur et un numéro d'allèle lui est associé.

**Protocole :** Pour chacun microsatellite, une réaction PCR a été réalisée dans les conditions rapportées dans le Tableau 13. Deux microlitres de produits PCR sont ensuite poolés et additionnés à 8µl de mix, lui-même composé de 9.85µl de formamide et 0.15µl de marqueur de taille (GENESCAN™ 400HD [ROX™]). L'automate 3730 Genetic Analyzer (Applied Biosystems) a été utilisé afin de séparer par électrophorèse microcapillaire les produits PCR marqués.

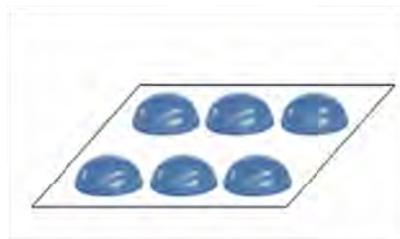
**Résultats :** Le signal de fluorescence détecté pour chacun des fragments est analysé grâce au logiciel GeneMapper V4.0 (Applied Biosystems) : chaque marqueur est identifié par la taille de ses amplicons et sa couleur de fluorescence puis, pour chaque individu, le ou les produits d'amplification sont interprétés comme "allèles" (Figure 48). Les génotypes sont enregistrés dans la base de donnée GEMMA afin de contrôler la cohérence mendélienne entre les parents et les descendants (fonction "compatibilité mendélienne"). Les règles d'exclusion des familles présentant des incohérences ont été les suivantes : (1) Si au moins deux allèles sont incohérents entre un parent et les descendants, la famille est retirée de l'analyse. Si deux incohérences concernent un seul des descendants atteints, cet individu est retiré des analyses.

### **I.C. Génotypage à l'aide la puce porcine haute densité d'Illumina**

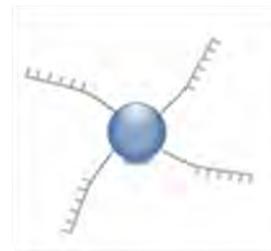
**Contexte :** Les puces de génotypage haute densité sont les outils de référence pour les GWAS, car elles permettent le génotypage de dizaine de milliers de SNP couvrant tout le génome, ceci en quelques jours, et pour plusieurs échantillons en parallèle (24 pour la puce porcine). La puce porcine 60KSNP d'Illumina a donc été choisie pour la première étape d'identification de loci impliqués dans les anomalies congénitales porcines.



Puce Illumina  
*Porcine*SNP60



Lame de verre recouverte  
de microbilles



Microbille et ses sondes  
spécifiques

Figure 49 : Puce porcine d'Illumina et ses microbilles recouvertes de sondes

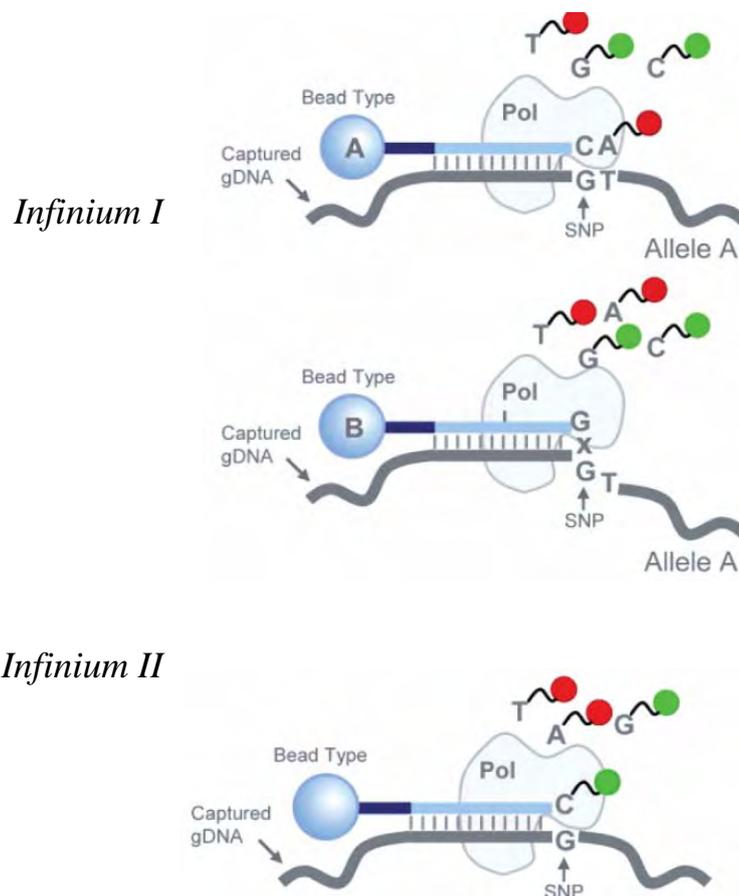


Figure 50 : Sondes de génotype utilisées respectivement en technologie Infinium I et II

**Principe :** Le principe du génotypage sur puces repose sur l'hybridation de l'ADN de l'individu analysé à des sondes spécifiques de marqueurs polymorphes de type SNP (Single Nucleotide Polymorphism). La fluorescence émise en chaque sonde spécifique de l'allèle (des allèles) des SNP est ensuite analysée pour l'attribution du génotype.

La puce 60KSNP est composée d'une lame de verre sur laquelle sont liées des microbilles, elles-mêmes recouvertes de sondes spécifiques de plus de 60000 SNP (Figure 49). Deux types de sondes sont présents sur la puce (Steemers and Gunderson, 2007) :

- Les sondes *Infinium I* basées sur une extension spécifique d'allèle (ASPE pour *allele-specific primer extension*)
- Les sondes *Infinium II* basées sur une incorporation simple base (SBE pour *single base extension*)

Dans les deux cas, les émissions de fluorescence sont obtenues après (1) une hybridation spécifique de l'ADN d'intérêt aux sondes et (2) une extension des sondes permettant de discriminer le génotype au SNP. Les allèles A ou T fluorescent en rouge, et les allèles C ou G fluorescent en vert (Figure 50).

La technologie *Infinium I* est basée sur l'utilisation de deux types de billes, chaque bille portant des oligonucléotides dont la séquence correspond à un seul allèle du SNP (le SNP étant localisé en 3' de la sonde). Au cours du génotypage, l'extension n'est alors possible que si l'hybridation de l'ADN cible est parfaite en 3' de l'oligonucléotide. L'interprétation est basée sur la présence ou non d'un signal à partir des sondes d'un seul ou des deux types d'allèles. La technologie *Infinium II* a permis de réduire le nombre de billes nécessaires à l'analyse de chaque SNP. Ce système de marquage à deux couleurs permet de détecter à partir d'une même bille les deux allèles d'un SNP. Dans ce cas, la séquence de la sonde correspond à la séquence précédant le SNP analysé, et l'incorporation d'un didesoxynucléotide rouge (ddA ou ddT) ou vert (ddC ou ddG) permet l'interprétation du génotype. Cependant, ce système ne permet pas de génotyper des SNP dont les allèles seraient A/T ou C/G. La puce porcine combine des marqueurs analysés à l'aide de l'une ou de l'autre de ces technologies selon les allèles aux marqueurs.

Afin d'améliorer la précision du génotypage, chaque SNP est représenté 20 fois sur la lame, et les billes sont distribuées de façon aléatoire sur les lames, selon une technologie mise au point

par Illumina, dite « BeadChip ». Pour chaque SNP, les 20 positions des billes sont identifiées préalablement et rassemblées dans un fichier, appelé *decode*, indispensable à l'analyse.

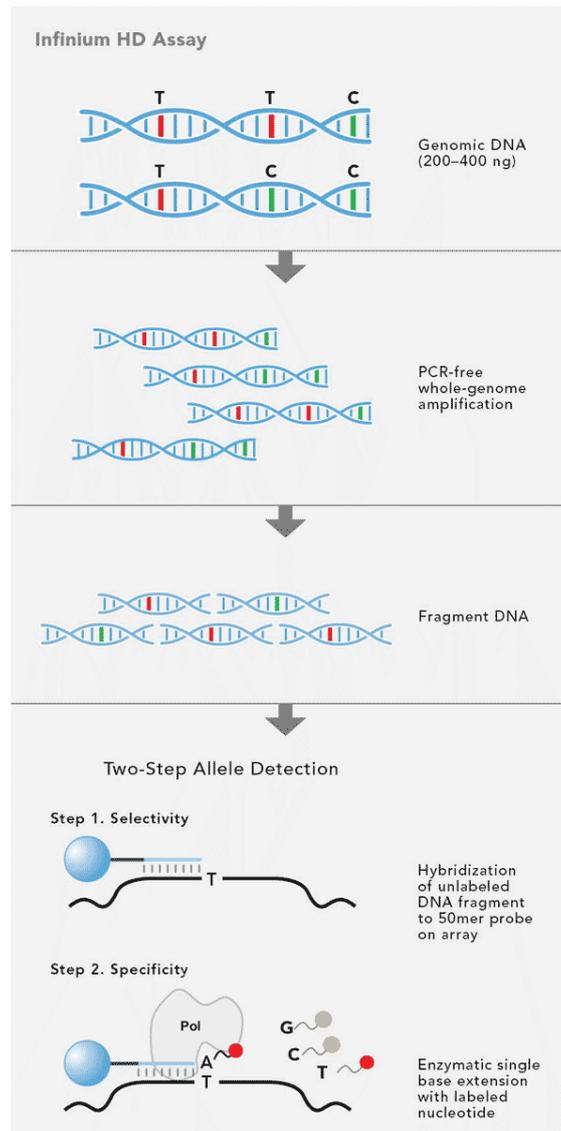


Figure 51 : Etapes du génotypage à l'aide de la puce porcine d'Illumina

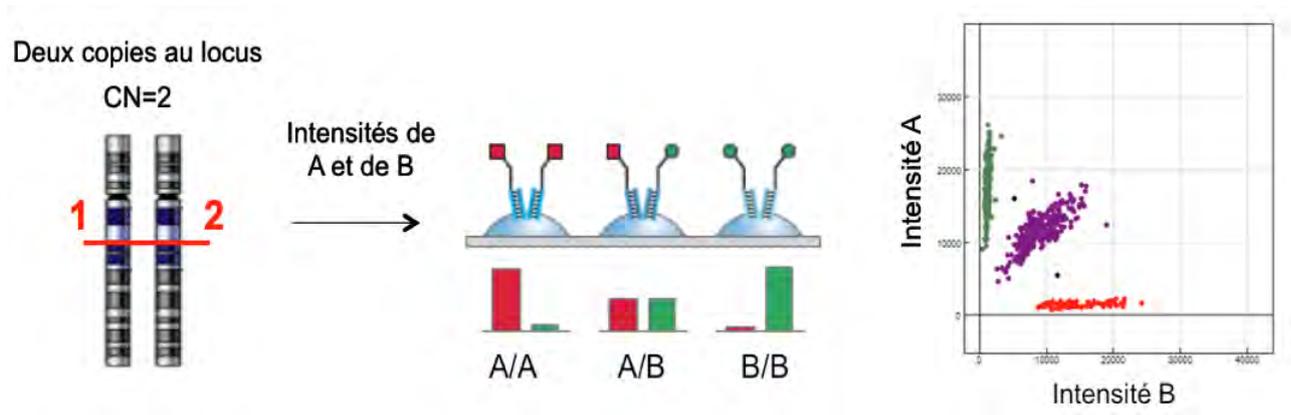


Figure 52 : Identification des génotypes à partir des émissions de fluorescence des billes de la puce porcine d'Illumina

**Protocole :** Le géotypage à l'aide de la puce Illumina se déroule sur trois jours (Figure 51):

Le premier jour, l'ADN est dénaturé à la soude, puis amplifié. Des bases Uracile sont incorporées de façon aléatoire à la place de certaines thymidines au cours de l'amplification.

Le deuxième jour, l'ADN amplifié est fragmenté grâce à une Uracyl-DNA-Glycosylase qui cible les Uraciles intégrés en jour 1. L'ADN fragmenté est alors dénaturé (formamide + 95°C), et déposé sur lame en chambre humide afin de s'hybrider sur les sondes.

Le troisième jour, les lames sont lavées afin d'éliminer l'ADN non hybridé, puis conservées en chambres capillaires humides afin d'éviter le dessèchement des fragments hybridés. Les bases marquées sont alors ajoutées pour incorporation. Les dernières étapes de lavage permettent d'éliminer les nucléotides marqués non-incorporés au cours de la réaction d'extension.

**Résultats :** Le géotype des individus est obtenu par traitement des valeurs d'intensités de fluorescence de chaque microbille, appelée également spot. Pour obtenir ces valeurs d'intensité, la lame est scannée (*iScan* d'Illumina) : pour chaque spot, deux valeurs d'intensité de fluorescence (rouge et vert) sont enregistrées. Selon la position de chaque spot sur la lame, et les informations de localisation du fichier *decode* spécifique de chaque puce, les valeurs d'intensités enregistrées sont associées au SNP correspondant. L'attribution d'un géotype à chaque individu est ensuite réalisée à l'aide du logiciel *GenomeStudio* (Figure 52). Pour chaque SNP, trois clusters correspondant aux trois types de géotypes A/A, A/B, B/B sont utilisés pour l'interprétation. Ces clusters peuvent être construits par l'utilisateur à partir d'un ensemble d'individus analysés, ou correspondre aux clusters proposés par Illumina pour l'interprétation des résultats et préalablement établis à partir d'une population de référence. Pour l'analyse réalisée en tout génome, les clusters utilisés sont ceux définis par le consortium international (Illumina). Pour la région candidate à l'intersexualité sur le chromosome 12, une vérification manuelle des attributions des clusters a été réalisée.

### **I.D. Géotypage à l'aide de la technologie Fluidigm**

**Contexte :** La première étape de GWAS a permis la primo-localisation d'un locus impliqué dans l'intersexualité. Une deuxième étape de cartographie fine a été envisagée pour réduire la taille de la région candidate. La technologie Fluidigm a permis le géotypage par PCR multiples de 96 SNP choisis spécifiquement dans la région.

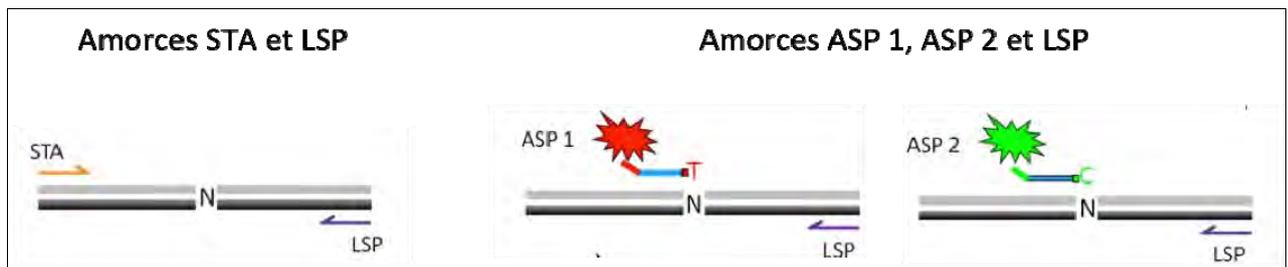


Figure 53 : Principe du génotypage Fluidigm. STA = Specific Target Amplification Primer, LSP = Locus Specific Primer et ASP = Allele Specific Primers

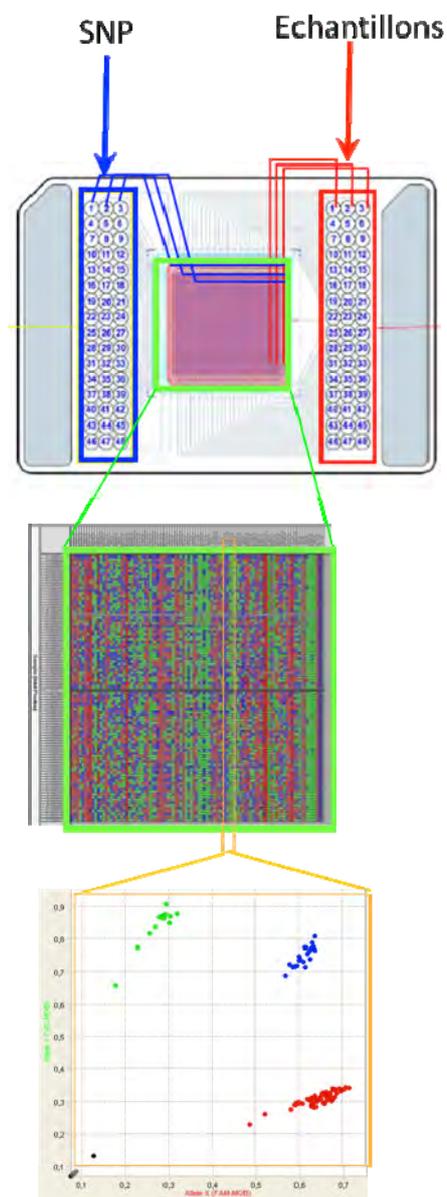


Figure 54 : Plaque BioMark, heatmap et attribution des génotypes

**Principe :** Le principe du génotypage Fluidigm est une amplification allèle spécifique, à partir d'amorces spécifiques de chaque allèle du SNP. Cette technique de génotypage a été réalisée à l'aide du Biomark (Fluidigm) qui permet d'analyser simultanément 48 SNP sur 48 individus ou 96 SNP sur 96 individus. Une première étape de pré-amplification est réalisée avec les amorces STA (*Specific Target Amplification*) et LSP (*Locus Specific Primer*), afin d'amplifier le fragment d'ADN contenant le SNP, quel que soit l'allèle présent (Figure 53). L'ADN de chaque individu analysé est ainsi pré-amplifié simultanément pour les 48 (96) loci. Une deuxième étape d'amplification est alors réalisée individuellement pour chaque marqueur, avec les amorces allèles spécifiques ASP (*Allele Specific Primers*) et LSP afin de mettre en évidence l'allèle présent sur le brin amplifié. L'amplification à partir d'une ou des amorces allèles spécifiques entraîne une émission de fluorescence au cours de l'amplification (Figure 53).

**Protocole :** Le génotypage à l'aide de la technologie Fluidigm a été réalisé selon la procédure du fabricant qui se déroule sur une journée : La première demi journée est consacrée à la pré-amplification avec les amorces STA et LSP (Figure 53) des fragments d'ADN contenant les SNP d'intérêt. Chaque réaction de pré-amplification est réalisée pour chaque échantillon d'ADN (1 individu), à partir des 96 couples d'amorces (STA et LSP). Le mélange est composé de 3,7 µl de mix comme défini dans le protocole du fabricant, et 1,3 µl d'ADN à 10ng/µl. Le mix est chauffé 15 minutes à 95°C puis subit 14 cycles de 15 secondes à 95°C suivi par 4 minutes à 60°C. Cette première réaction est réalisée dans une plaque 96 puits à l'aide d'un appareil PCR standard. La deuxième demi-journée correspond à la phase d'amplification avec les amorces ASP1/2 et LSP (Figure 53). Cette seconde amplification est réalisée à l'aide d'une puce Fluidigm (*Genotyping Dynamic Array*). Cette plaque est constituée à gauche de 48 (96) puits permettant de répartir les différents mix (un mix par SNP) et à droite de 48 (96) puits permettant de répartir les échantillons d'ADN pré-amplifiés lors de la première étape. Le mélange de chaque ADN avec chacun des 48 (96) mix est réalisé à l'aide de l'IFC (*Integrated Fluidic Circuit*) dans 2304 (9216) micropuits de 7 nanolitres (Figure 54). La PCR allèle spécifique est alors réalisée à l'aide du Biomark selon les conditions rapportées dans le Tableau 14.

<b>Conditions amplification</b>
<b>Dénaturation initiale :</b> 30 minutes à 70°C 10 minutes à 25°C 5 minutes à 95°C
Cycle 1: dénaturation : 15 secondes à 95°C hybridation : 45 secondes à 64°C élongation : 15 secondes à 70°C Cycle 2: dénaturation : 15 secondes à 95°C hybridation : 45 secondes à 63°C élongation : 15 secondes à 70°C Cycle 3: dénaturation : 15 secondes à 95°C hybridation : 45 secondes à 62°C élongation : 15 secondes à 70°C Cycle 4: dénaturation : 15 secondes à 95°C hybridation : 45 secondes à 61°C élongation : 15 secondes à 70°C 34 cycles : dénaturation : 15 secondes à 95°C hybridation : 45 secondes à 60°C élongation : 15 secondes à 72°C
<b>Arrêt de la réaction :</b> 10 secondes à 25°C

Tableau 14 : Conditions d'amplification pour le génotypage Fluidigm

**Résultats :** L'intensité de fluorescence émise au cours de l'amplification est enregistrée lors de quatre acquisitions en point final à l'issue des 26, 30, 34 et 36èmes cycles. Comme pour l'analyse des marqueurs de la puce 60K, les valeurs brutes d'intensité des deux fluorescences sont enregistrées et utilisées pour déduire le génotype de chaque individu.

### I.E. Génotypage du panel porcin d'hybrides d'irradiation

**Contexte :** Dans le cadre de cette thèse une étude de cartographie du chromosome X porcin a été réalisée afin, entre autres, (1) de définir la région pseudo-autosomale de ce chromosome et de la comparer aux régions PAR des autres espèces et (2) de valider la qualité de l'assemblage porcin de ce chromosome.

**Principe :** La cartographie physique du chromosome X a été réalisée à l'aide de deux types de marqueurs : des EST porcins, homologues des gènes connus chez les autres espèces de mammifères comme localisés sur la région pseudo-autosomale, ont été analysés par PCR à l'aide du panel IMNpRH2<sub>12,000</sub>. Une nouvelle carte de l'ensemble du chromosome a également été réalisée à l'aide des marqueurs de la puce porcine 60KSNP. Ces données ont été obtenues par le génotypage de l'ADN de chaque clone du panel à l'aide de la puce 60K.

**Protocole :** Les 90 clones du panel IMNpRH2<sub>12,000</sub> (Yerle et al., 2002) ont été utilisés pour la construction des cartes. Des couples d'amorces ont été dessinés dans *Primer3* (<http://frodo.wi.mit.edu/>) pour chacun des EST sélectionnés. Ils ont ensuite été testés par amplification PCR sur des échantillons d'ADN porcin (contrôle positif), d'ADN de hamster (contrôle parallèle) et d'eau (contrôle négatif).

Chacun des EST sélectionnés a ensuite été amplifié sur les 90 clones d'hybrides du panel par PCR. Le mix PCR d'un volume total de 15 µl, comprenait 1X de tampon PCR (GoTaq, Promega) 0.25 µM de chaque amorce, 0.25 unités de Taq Polymerase (GoTaq, Promega), 200 µM de chaque dNTP et 25 ng d'ADN. Les paramètres du cycle PCR étaient : 94°C à 5 min pour la dénaturation initiale, suivi par 35 cycles de 30 secondes de dénaturation à 94°C, 30 secondes d'hybridation à une température spécifique des amorces (Tableau 15) et 30 secondes d'élongation à 72°C. Une étape d'extension finale est réalisée à la fin des 35 cycles pendant 20 minutes à 72°C. Les produits de PCR ont été analysés par électrophorèse sur gel d'agarose à 2,5%, et visualisés par marquage au bromure d'éthidium. Une amplification PCR positive est notée H (*here*) et une amplification négative est notée A (*absent*). Lorsque le marquage est faible, le signe "-" est utilisé afin d'indiquer que le résultat peut être douteux.

Nom du gène	Amorces (5'→3')	Taille du fragment (bp)	Température (°C)	Nom de l'EST
PLCXD1	left : GGTCCCTGGAGAGGATGAC	235	60	CK466735
	right : ATCCCACAGGATGGAGGAG			
GTPBP6	left : CTCCTGTACCCCTTACCT	110	52	CV867557
	right : CGCTCTCAGATGCCCTAAC			
PPP2R3B	left : GAGGCAGCGACTTCTGTGT	109	52	DN103585
	right : TCCGCGTGAATAAATAACAAC			
SHOX	left : CTGGATAGCACCTTAGCAGGA	95	56	EW428396,2
	right : TCCCACCCCTTTTATTTT			
CSF2RA	left : GATGTCACCTCCAAAACATCTC	100	52	FD631245,1
	right : AAAACCGAGTCAAAGAACTGG			
SLC25A6	left : CCACGAAGGCAGACCTAGAC	117	60	AK239241
	right : TATTCGGCTGTGCAAAAGAG			
ASMTL	left : GCATGCTGGTGGAGCAC	100	62	FD632834,1
	right : CGGGGCGGGTAGTACAG			
P2RY8	left : AAGAGGAGACACAGACAGAGG	102	60	BF188931,1
	right : CAGCTTCTGGGGTTCCA			
AKAP17A	left : GCCTTCTCCCTTGTGTGTC	222	60	CF180526
	right : ACCTTCCTTTTCCACGGACT			
DHRSX	left : GGGACACCTTGTGGGGATAG	179	60	AK233613
	right : CGCCTCTGCTTCTCACTTC			
ZBED1	left : CGAGCGCCTCTTCTGCTC	155	60	FD641139,1
	right : CCTCGTCTCTGCTCCTCA			
XG	left : ACACGTTGCCTGGAGTAAGG	243	60	AK232227
	right : ACCCAACAGCATTCTCCATC			
GYG2	left : CGTGCTGCTGTATCCATCCT	181	60	DT332656
	right : CGTTCTGCAAAGCCTAAACC			
ARSD	left : CACTTCCCTTCTGCTCCT	103	60	AK234744
	right : CGGCATCTCTGACTTTCTGG			
ARSE	left : TGCATTTTGC GACTAAGAACA	115	56	CJ031183
	right : CCACCTCAAAGCAAGACGA			
MXRA5	left : GTATGTGGTGGAGGGGTGTT	136	60	CK455361
	right : CTTCAGGTGTCAGGGTGGTT			
PRKX	left : GGTGGTGTITTTGGGACTTTT	239	52	BI346943
	right : TTTTGTITTTTCGGGGTTTTT			
NLGN4X	left : CACACGCAAGAGAGAAAGGA	127	60	CN161317
	right : ATGGTCGTGGGAAAAGTCTG			
HDHD1A	left : CTTGGCAGAAAGCAGAGGAG	204	60	BG382475
	right : AAGGAGAGGGAGGGGATACA			
STS	left : CCGGGAAACATCATCTGG	152	58	DN128923
	right : CAGGGCTGCTTCCAGTGT			
PNPLA4	left : ATGTGCTGGTTGGCGATG	100	60	CO952450
	right : GCCGGACAGTGACCATCT			
KAL1	left : GAAGAGTCTACCGCAATGG	156	60	DN125663
	right : CACATTAACACCGACGCAAG			
TBL1XY	left : ACACAGGCAGGGTCGAAA	153	60	BP166851
	right : GTCCACACACCCCAAGAGAC			
GPR143	left : GGAGACCTGTGAGAGGACGA	167	60	CN158227
	right : ACGGATCTCTCTGCATCAC			
SHROOM2	left : GCCGAGCAAATCTCAAAGAA	292	60	BF199365
	right : TGGCGACTTCAGAGTGAGTG			
WWC3	left : CAGACCAAACCTTGACTTCCATC	117	58	EW148791
	right : ATGGGCTCTTTGAGGTTCTG			
MID1	left : CGGGTAATGGGATTCAGTTT	96	58	AK349446,1
	right : TGGCTCGTGTCTTTTATC			

Tableau 15 : Amorces utilisées pour l'amplification sur le panel d'hybrides d'irradiation

L'ADN de chacun des clones du panel RH2 a été génotypé avec la puce porcine 60SNP d'Illumina. Des échantillons d'ADN porcine et de hamster ont également été génotypés dans la même expérience. L'attribution du génotype a été réalisée selon la méthode décrite par Servin et al. (Servin et al., 2012).

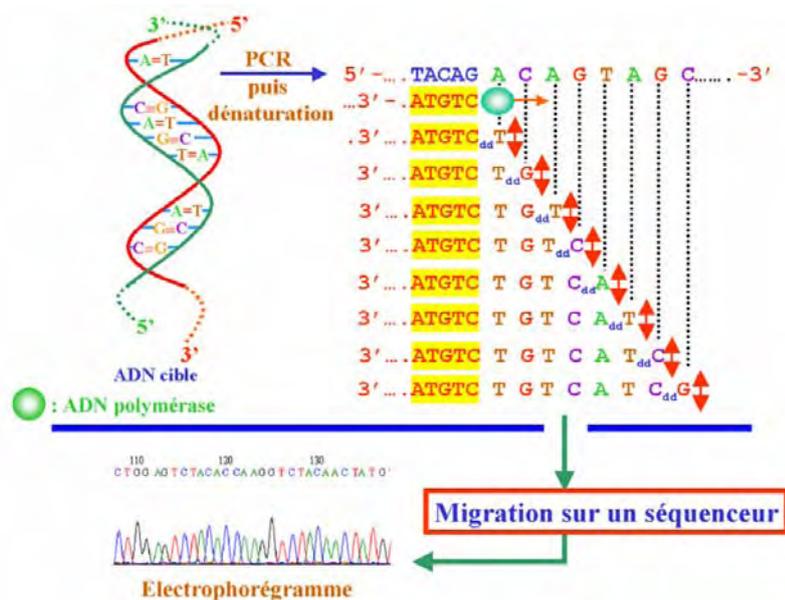


Figure 55 : Schéma des étapes de séquençage automatisé selon la méthode de Sanger.

Primer name	Stand	Primers (5'-> 3')	Product size (bp)	Annealing temperature (°C)	Elongation time (secondes)	Cycles number
SOX9_1	Forward	CAAGAAGAGCTGGGCACAC	830	64	30	38
	Reverse	CCTACAGCCCCTTCAACCTC				
SOX9_2	Forward	TGGTGGTCTGTGTAGTCGT	705	60	30	35
	Reverse	GCACAGCCCTGGTTGATT				
SOX9_3	Forward	GCCTGCTGCTTGGACATC	1214	65	60	45
	Reverse	CACAGAGCCCCTCTCTCTTTT				
SOX9_4	Forward	CTCCGCCTCCTCCACGA	832	60	30	36
	Reverse	GCTCCTTGCAGATTAAGTGGTT				
SOX9_5	Forward	CCCTCTCCCCTTCTCTTCT	540	60	30	36
	Reverse	TGACAAAGTTCTTGGATTGCTC				
SOX9_6	Forward	TTCTCCTCTTCTCCGGCTTC	924	60	30	35
	Reverse	CTCCTCTCCAATTCGCCTTT				
SOX9_1	Forward	TGTCCTCTTCGCTCTCCTTC	848	60	30	38
	Reverse	ACACACACACACACGCACAC				
TESCO_1	Forward	GGCAATCATCCCTGCATAAC	706	63	30	36
	Reverse	CCAGGAGTCTGAGGTTTAGGG				
TESCO_2	Forward	GCGCGTTGTGTATTGGAAA	690	63	30	36
	Reverse	TTAACCCTGAGCCACCA				
TESCO_3	Forward	GGCTATTGCACTGGGCTCT	642	63	45	35
	Reverse	CCAGGAGTCTGAGGTTTAGGG				
TESCO_4	Forward	ACCATCTCAAGGGCAACT	862	58	45	35
	Reverse	AAAAGGGAAAACCCAACAGG				

Tableau 16 : Amorces pour le séquençage de SOX9 et TESCO, et conditions PCR correspondantes

## II. Séquençage

### II.A. Séquençage automatisé selon la méthode de Sanger

**Contexte :** La première étape de GWAS a permis l'identification d'une région candidate à l'intersexualité. Dans le cadre de la cartographie fine de cette région, le séquençage du gène *SOX9* et de son enhancer testicule-spécifique (TESCO) a été réalisé.

**Principe :** Le principe du séquençage Sanger est basé sur l'incorporation aléatoire d'un didéoxynucleotide triphosphate (ddNTP : ddA-, ddC-, ddG- et ddT-TP), marqué par un fluorochrome différent selon la base, à la place d'un déoxynucleotide triphosphate (dNTP : dA-, dC-, dG- et dT-TP). L'utilisation de ddNTP comme terminateurs d'élongation stoppe la polymérisation par l'ADN polymérase lors de la réaction. La réaction de séquençage est une amplification linéaire réalisée à partir d'une seule amorce complémentaire de l'ADN matrice que l'on souhaite séquencer. A la fin de la réaction, les fragments obtenus sont de taille différente et émettent une fluorescence base-spécifique (Figure 55). Dans un second temps, les fragments d'ADN sont séparés selon leur taille (résolution à une base) et le signal de fluorescence émis, associé à chaque fragment, est enregistré. Les résultats sont présentés sous forme d'électrophorégramme. Le séquençage d'un fragment est le plus souvent réalisé sur les brins sens et antisens afin de confirmer les séquences obtenues.

**Protocole :** Des amorces de *SOX9* et son enhancer testicule-spécifique TESCO (*TEstis-Specific enhancer of SOX9 COre element*) ont été définies à l'aide du logiciel Primer3. La séquence porcine de *SOX9* a été obtenue dans Ensembl (Sscrofa10.2:12:9027039:9033810:-1, Accession : Acc:NM\_213843). La séquence de TESCO a été obtenue à partir de la séquence du draft porcin par homologie avec la séquence humaine de TESCO (Georg et al., 2010). Les amorces sont présentées dans le Tableau 16.

Les séquences d'ADN d'intérêt ont été amplifiées par PCR. Les produits d'amplification (1 à 12 µl) ont ensuite été incubés en présence de 0.5 unité de Shrimp Alkaline Phosphatase (Promega) et de 0.8 unités d'Exonuclease I (New England Biolabs) pendant 45 minutes à 37°C. Ces enzymes sont ensuite inactivées pendant 30 minutes à 80°C. L'exonucléase 1 permet d'éliminer les amorces en excès par digestion de l'ADN simple brin, et la SAP (Shrimp Alkaline Phosphatase) dé-phosphoryle le phosphate en 5' des nucléotides libres, empêchant ainsi leur utilisation lors de la réaction de séquence.

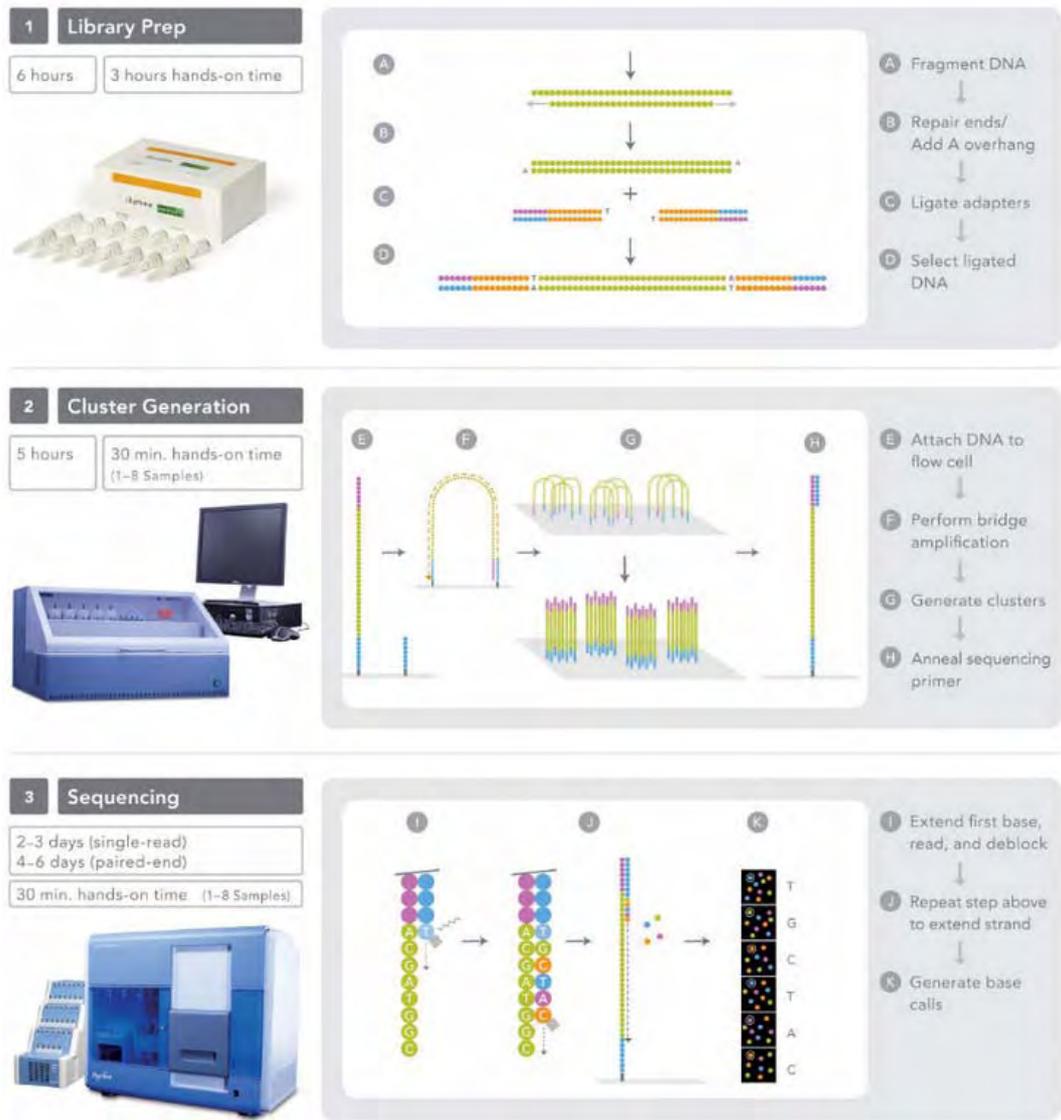


Figure 56 : Etapes du séquençage HiSeq 2000

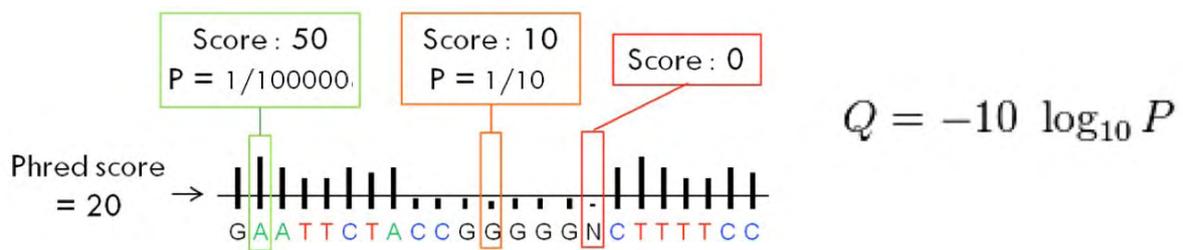


Figure 57 : Exemple de "Phred score" estimé pour les bases d'une lecture

La réaction de séquençage a ensuite été réalisée à l'aide du kit diChloroRhodamine Prism AmpliTaq FS Big Dye Terminator V3.1 kit (Applied Biosystems) selon les recommandations du fournisseur, et les produits de la réaction ont été analysés sur le séquenceur 3730 ABI capillary DNA (Applied Biosystems).

## II.B. Séquençage haut débit à l'aide du séquenceur HiSeq 2000

**Contexte :** En parallèle des études d'association en tout génome, deux individus intersexués et deux individus témoins ont été séquencés pour l'ensemble de leur génome. Le séquenceur nouvelle génération d'Illumina, HiSeq 2000, permet le séquençage de 600 Gbp (30 X) en lectures de 100 bp par ligne. Dans le cadre de cette thèse, chaque individu a été séquencé sur une ligne correspondant à une couverture moyenne de 4X.

**Protocole :** Le séquençage dure 10 jours environ (Figure 56). L'ADN est d'abord extrait et fragmenté de façon aléatoire par nébulisation ou ultra-sons. Les fragments de 400 bases +/- 100bases sont sélectionnées par migration sur gel d'agarose et des adaptateurs (A et B) sont ajoutés aux extrémités de chaque fragment. La préparation des librairies dure 1 jour et demi. Les molécules d'ADN sont fixées à l'aide des adaptateurs sur une plaque de silice comportant des oligonucléotides complémentaires des séquences de A et B. Les fragments sont ensuite amplifiés par PCR localement sur la plaque de silice afin de former des « clusters », puis dénaturés afin de ne conserver qu'un seul brin d'ADN matrice pour le séquençage. La réaction de séquençage est réalisée en parallèle pour des millions de fragments par l'incorporation de nucléotides fluorescents (bloqués en 3'), élimination des nucléotides non utilisés, mesure de la fluorescence émise par le nucléotide incorporé, puis élimination de la molécule bloquante afin de permettre l'incorporation d'un nouveau nucléotide au cycle suivant. Ce cycle est répété une centaine de fois et correspond aux lectures 1. A l'issue du séquençage du brin sens, une nouvelle série de cycles de séquençage est réalisée afin d'obtenir le séquençage du brin complémentaire (lectures 2). Le séquençage des brins sens et antisens dure 8 jours au total.

**Résultats :** À chaque cycle d'incorporation l'émission de fluorescence est mesurée en chaque position de la lame de silice. Une valeur de qualité est attribuée à chaque base selon l'intensité émise, c'est le score de qualité Phred Q, ou *Phred score* (Figure 57). Plus le score est élevé plus la probabilité que la base soit correctement attribuée est élevée. Un score de 50 signifie qu'il y a 1 chance sur 100 000 pour que la base soit mal identifiée, et un score de 10 signifie qu'il y a une chance sur 10 d'erreurs d'attribution de la base.

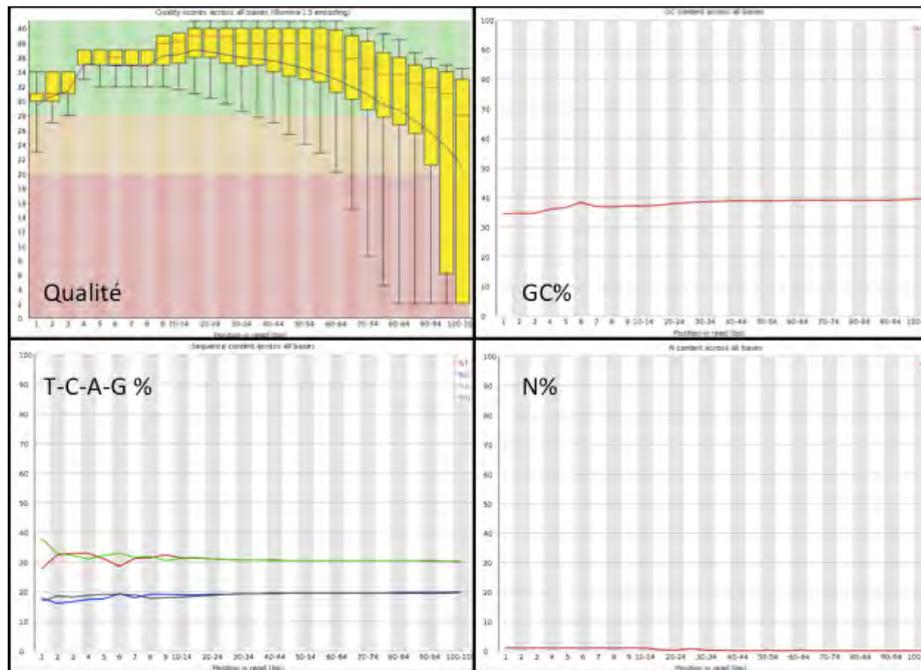


Figure 58 : contrôles qualité des séquences par base

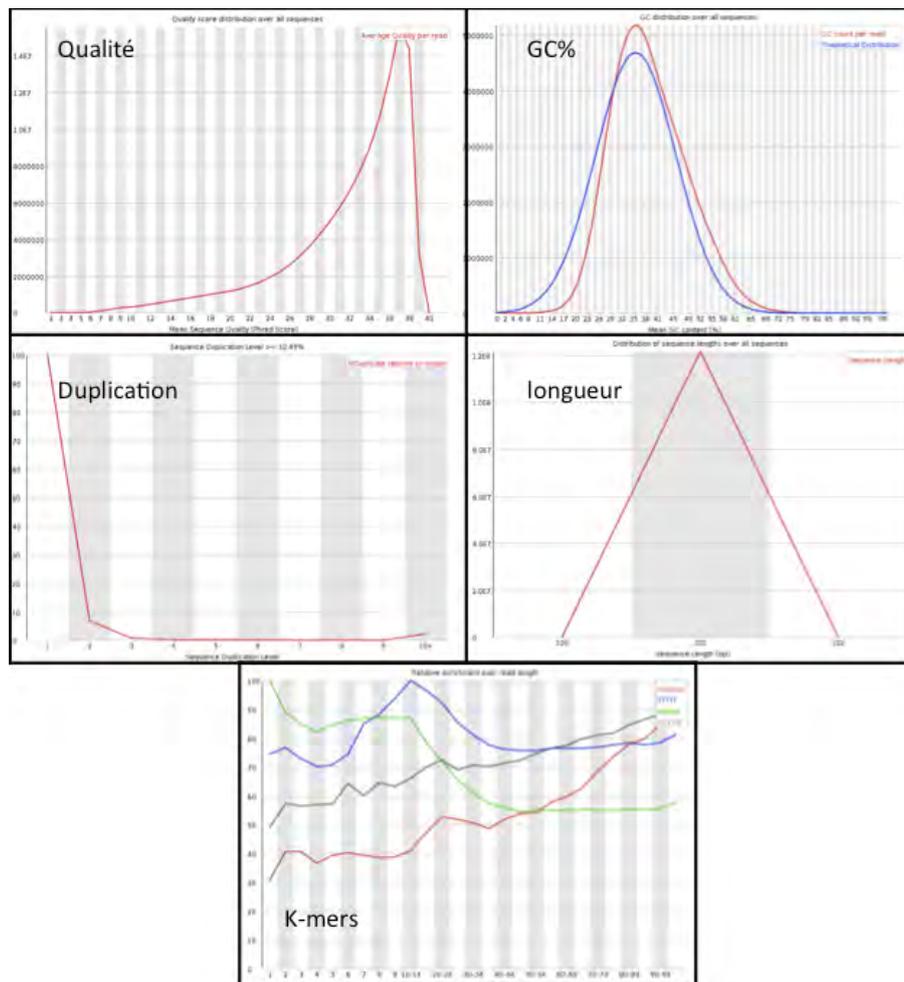


Figure 59 : Contrôles qualité des séquences par lecture

Le séquençage de 2 x 100 bases (séquençage « paired end ») à l'aide de cette technologie permet d'obtenir pour plus de 85% des bases lues avec une valeur de qualité de 30 (ce pourcentage atteint 90% pour des lectures de 2x35 bases). La valeur seuil sélectionnée dans le cadre de notre étude a été 20 (1 chance sur 100).

### **II.B.1. Contrôle de la qualité globale du séquençage**

Afin d'estimer la qualité globale d'une ligne de séquençage avant de l'exploiter, plusieurs contrôles sont réalisés. Ces contrôles n'ont pas pour but de supprimer des lectures, mais de savoir si les données peuvent être utilisées pour les analyses. Des contrôles sont réalisés selon la position de la base pour l'ensemble des lectures (pour toutes les bases en position 1, puis pour toutes les bases en position 2, etc. jusqu'à la dernière base en position 100), et lecture par lecture.

A chaque contrôle, les observations « Warn » et « Fail » sont attribuées selon les résultats (« Warn » lorsque les résultats sont près de la valeur seuil, et « fail » lorsqu'ils sont en dessous du seuil).

Les contrôles qualité selon la position de la base (Figure 58) permettent de vérifier :

- La qualité globale (warn : plus petit quartile d'une base inférieur à 10, ou si la médiane d'une base est inférieure à 25, Fail : plus petit quartile d'une base inférieur à 5, ou si la médiane d'une base est inférieure à 20)
- Le pourcentage de GC (warn : +/- 5% de la moyenne, fail : +/- 10%)
- Le pourcentage de Ns (warn : > 5%, fail : > 10%)
- La différence entre A et T ou G et C (warn : > 10%, fail : > 20%).

Les contrôles qualité selon les lectures (Figure 59) portent sur :

- La qualité globale par lecture (warn : qualité < 27 (taux d'erreur de 0.2%), fail : qualité < 20 (1%))
- Le pourcentage de GC (warn : déviation de la distribution normale du pourcentage de GC pour plus de 15% des lectures, fail : déviation pour plus de 30%)
- Les duplications de séquences (warn : les lectures non uniques représentent plus de 20% du total, fail : les lectures non uniques représentent plus de 50% du total)
- La longueur (warn : toutes les lectures ne sont pas de la même taille), fail : au moins une lecture de 0pb).

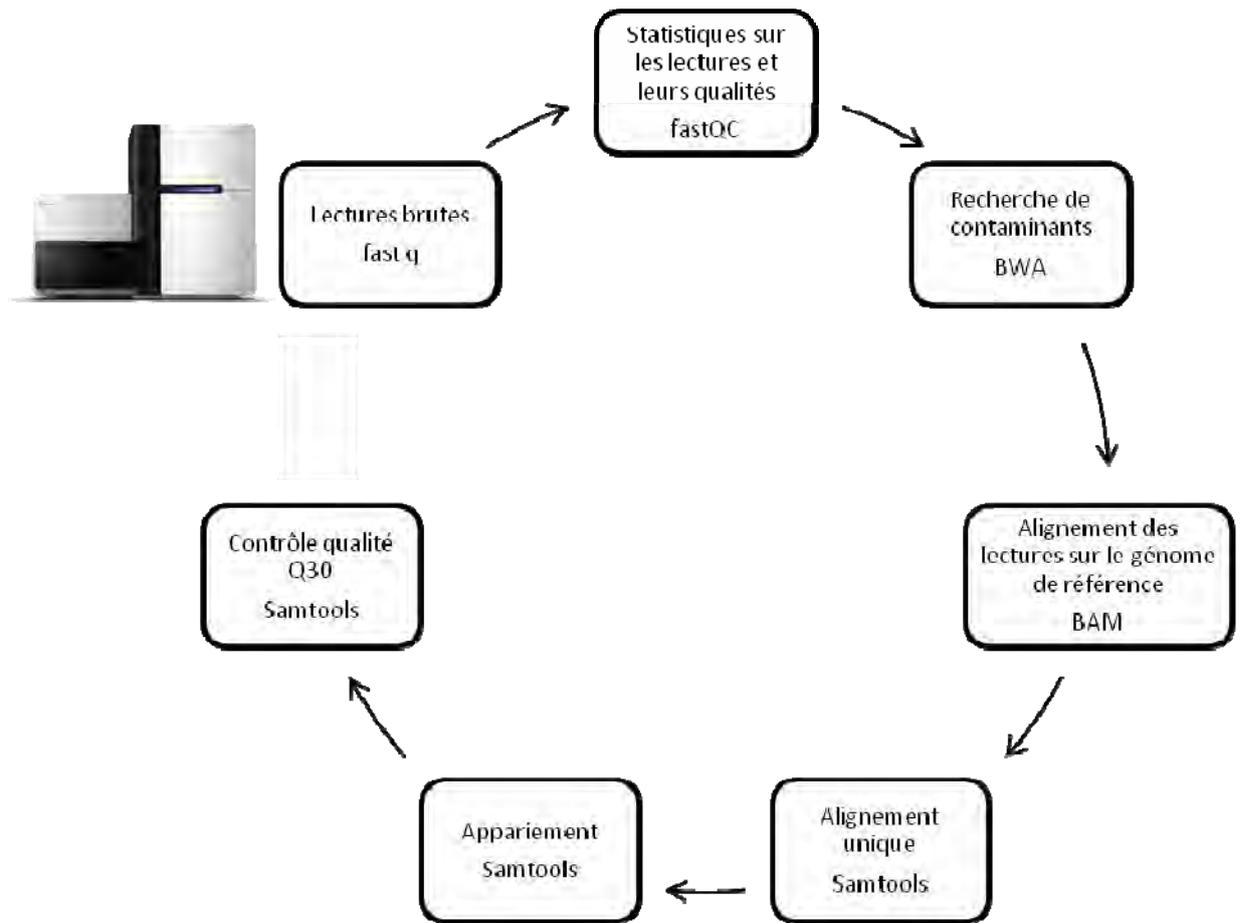


Figure 60 : Etapes d'analyses des séquences HiSeq

- Les K-mers (warn : un K-mer est enrichi plus de trois fois sur l'ensemble des lectures, ou plus de cinq fois à une position, fail : un K-mer est enrichi plus de dix fois à une position)
- Les séquences surreprésentées (warn : une séquence ne peut pas représenter plus de 0.1% du total, fail : une séquence ne peut pas représenter plus de 1% du total).

Si la qualité globale est validée, les séquences sont ensuite alignées sur des génomes bactérien, viral et levure afin de vérifier l'absence de contamination.

### **II.B.2. Alignement des séquences et critères de sélection lecture-spécifique :**

Seules les séquences pour lesquelles des bases sont attribuées sont conservées. Les séquences s'alignant sur N, c'est à dire pour lesquelles les bases ne sont pas attribuées, sont donc retirées.

Les séquences conservées sont ensuite alignées sur le génome de référence. Les séquences se positionnant à plusieurs endroits du génome, ou qui ne s'apparient pas correctement (lectures 1 sens et lectures 2 anti-sens, séparées par 300 pb environ) sont retirées (d'où l'intérêt d'avoir une séquence de référence de qualité correcte). Ces critères peuvent varier d'une étude à l'autre, la partie résultats de la thèse reprend l'analyse des séquences spécifique à cette étude.

Enfin, pour chaque séquence alignée répondant aux critères précédents, un score d'alignement est attribué (*quality score*). Seules les séquences pour lesquelles 2 erreurs sur 32 bases maximum sont observées sont sélectionnées. Ces critères permettent de conserver uniquement les séquences dont la qualité est correcte, minimisant ainsi les biais dans la recherche de polymorphismes.

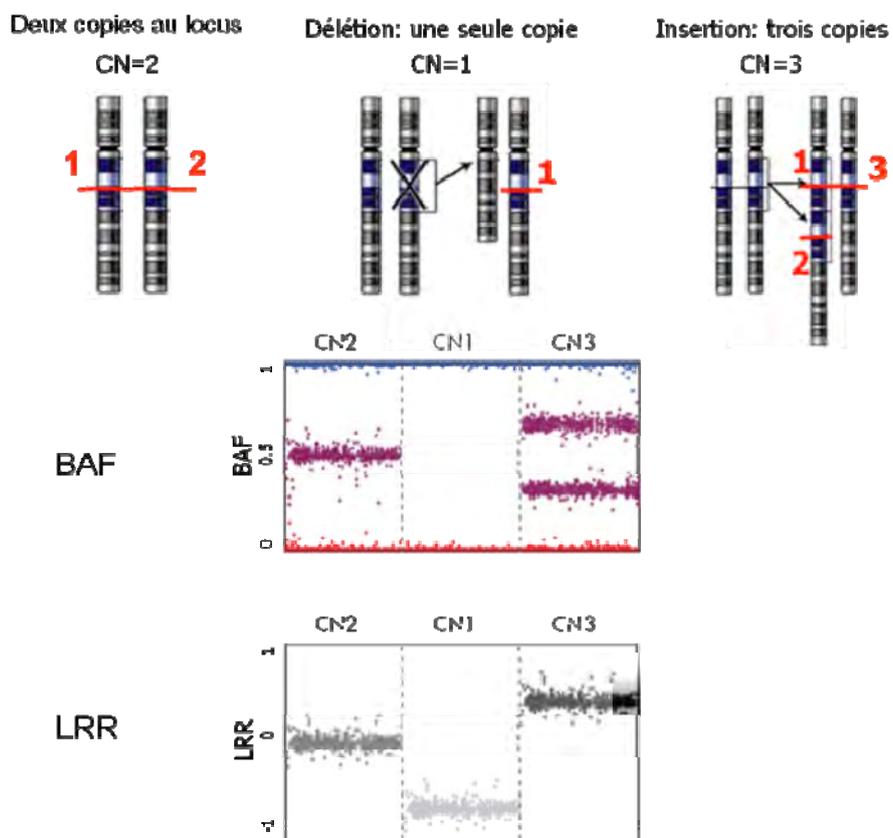


Figure 61 : BAF et LRR en absence de variation du nombre de copies (2 copies par locus)

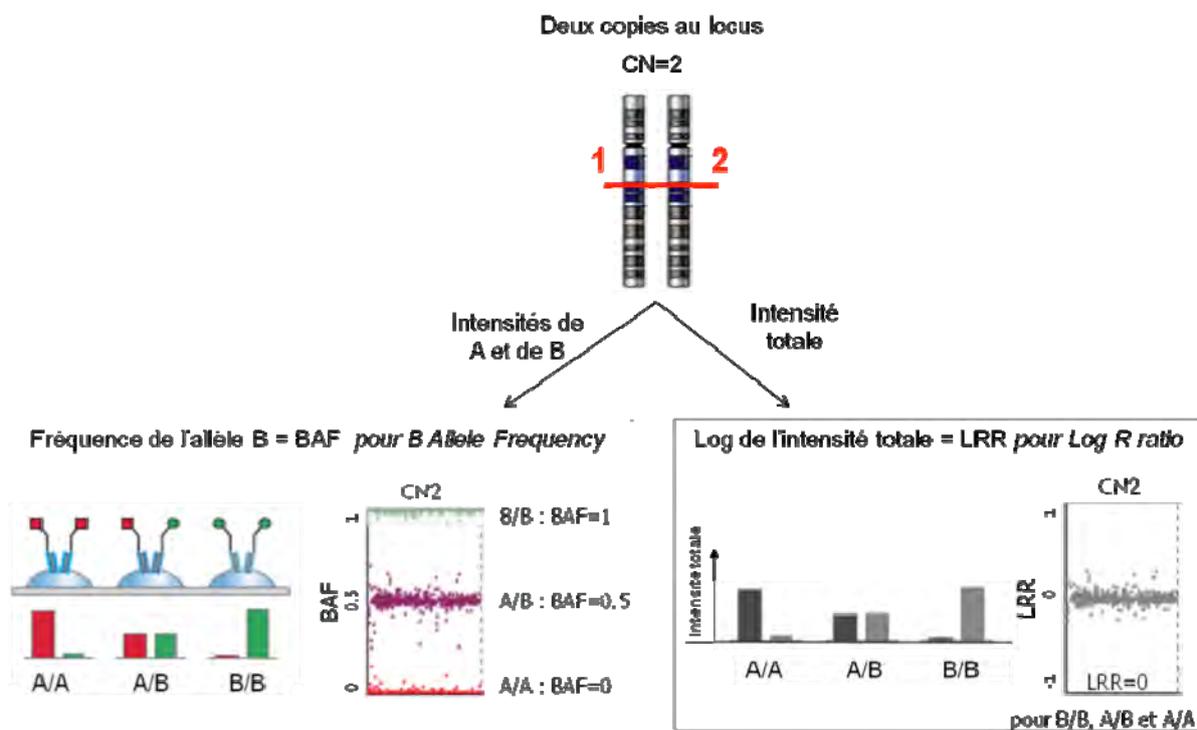


Figure 62 : BAF et LRR lors de CNV (1, 2 et 3 copies par locus)

### III. Analyses des données de génotypage

#### III.A. Recherche de motifs répétés

**Contexte :** Des duplications et des délétions responsables d'intersexualité chez l'homme dans la région de *SOX9* ont été rapportées dans la littérature. Afin de vérifier si un remaniement pouvait être la cause de la maladie chez le porc, des variations du nombre de copies (CNV pour *Copy Number Variation*) ont été recherchées chez les porcelets intersexués dans la région candidate du chromosome 12 contenant *SOX9*.

**Principe :** Les intensités obtenues à l'aide de la puce porcine de génotypage ont été utilisées pour la recherche de CNV. *PennCNV* (Wang et al., 2007) est un logiciel gratuit permettant ce type de détection à partir des données brutes de génotypage des puces Illumina et Affymetrics. *PennCNV* implémente un modèle de chaîne de Markov Caché (HMM) qui intègre plusieurs sources d'informations. Il prend en compte la distribution allélique globale et l'intensité des signaux de fluorescence de chaque échantillon. Il est également possible d'utiliser l'information familiale pour l'identification des CNV.

**Protocole :** Les données utilisées pour la détection sont la fréquence de l'allèle B (BAF ou *B allele frequency*) et le logarithme du ratio de l'intensité totale observée sur l'intensité totale attendue (LRR ou *Log R ratio*). Le LRR et le BAF de chaque échantillon analysé peuvent être exportés de GenomeStudio (Illumina San Diego, USA). La recherche des CNV a été réalisée avec les paramètres par défaut (facteur UF du modèle de la chaîne de Markov caché = 0.01).

Un fichier PFB (*population frequency of allele B*) est également nécessaire : il est généré avant les analyses et contient, pour chacun des SNP, la fréquence de l'allèle B dans l'ensemble de la population étudiée. Les génotypes provenant du CNG et ceux provenant du Labogena ont été traités séparément.

**Résultats :** Dans le cas où le nombre de copies par locus est égal à 2, le LRR est toujours égal à 0, quelque soit le génotype, et le BAF prend les valeurs de 0, 0.5 ou 1, selon le génotype, A/A, A/B ou B/B respectivement (Figure 61). Dans le cas d'une délétion, le BAF n'est jamais de 0.5, et le LRR est dévié négativement (Figure 62). Dans le cas d'une insertion, quatre valeurs de BAF sont observées, et le LRR est dévié positivement.

La recherche de CNV a été réalisée à l'échelle du génome.



### III.B. Haplotypes et construction de carte génétique

**Contexte :** Les premières analyses génétiques sont généralement réalisées marqueur par marqueur. Ensuite, l'analyse combinée de plusieurs marqueurs adjacents permet de gagner en informativité. La reconstruction d'haplotypes peut être réalisée au sein de famille (reconstruction de phases) ou au sein d'une population (reconstruction d'haplotypes). Dans le cadre de cette thèse, la variabilité haplotypique a essentiellement été caractérisée à l'aide des familles nucléaires utilisées pour le GWAS. Ces travaux ont été réalisés dans le cadre de deux études : (1) La reconstruction d'haplotype(s) (ou des phases au sein de familles) afin de rechercher et de caractériser un chromosome ancestral commun porteur d'une mutation présent chez tous les atteints, (2) La reconstruction des phases afin de réaliser une carte génétique (carte de recombinaison) du chromosome X.

**Principe :** La reconstruction d'haplotypes comprend deux étapes (1) l'identification de l'origine parentale de chaque allèle reçu par le descendant et (2) l'identification des événements de recombinaison. Pour un marqueur SNP ayant deux allèles A et B, l'origine parentale de l'allèle reçu par un descendant est rarement non-interprétable, excepté lorsqu'un descendant hétérozygote à un SNP est issu de deux parents également hétérozygote, quelques descendants par famille permettent ainsi facilement de déduire les allèles des marqueurs adjacents, portés ensemble sur chaque chromosome parental. Compte-tenu de la densité en marqueurs, la probabilité que deux descendants aient reçus un chromosome recombinant entre les mêmes marqueurs est faible voire nulle ; par paire de marqueurs, les combinaisons alléliques majoritaires sont donc considérées comme phases parentales, les minoritaires résultant de recombinaisons méiotiques.

**Protocole :** La reconstruction haplotypique au sein des familles a été réalisée à l'aide de YAPP, outil développé par B. Servin au laboratoire. Le principe de YAPP est similaire à l'outil développé et publié par Coop G (2008). Les résultats obtenus sont les phases paternelles et maternelles, les haplotypes des descendants et une carte de recombinaisons méiotiques (mâle, femelle et moyenne).

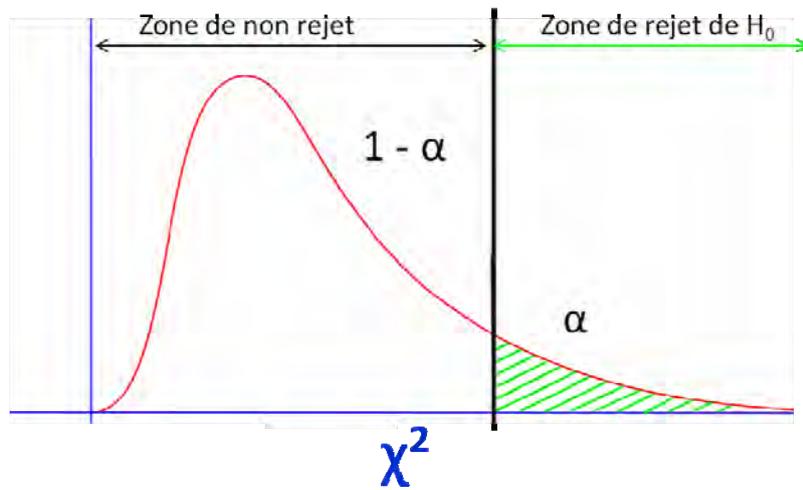


Figure 63 : Distribution du  $\chi^2$

		Allèle non transmis		Somme
		A1	A2	
Allèle transmis	A1	$a$	$b$	$a + b$
	A2	$c$	$d$	$c + d$
Somme		$a + c$	$b + d$	$2n$

Tableau 17 : exemple des valeurs observées utilisées pour le test du  $\chi^2$  dans une analyse TDT

### III.C. Analyses TDT

**Contexte :** Les analyses TDT ont permis de tester la liaison génétique et l'association entre les principales anomalies congénitales porcines et chacun des 60.000 marqueurs de la puce porcine.

**Principe :** Le test statistique d'indépendance Khi-carré  $\chi^2$  est utilisé pour rechercher une distorsion de ségrégation des allèles parentaux vers la descendance. Deux hypothèses sont testées. Sous hypothèse nulle ( $H_0$ ), il y a indépendance entre les variables. Sous hypothèse alternative ( $H_1$ ),  $H_0$  est rejeté et il y a dépendance des variables (Figure 63).

Estimer la valeur du  $\chi^2$  revient à calculer la différence entre la valeur observée  $O_{ij}$  (ou valeur empirique) et la valeur attendue s'il y avait indépendance  $E_{ij}$ ; (ou valeur théorique) :

$$\chi^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Plus la différence est grande, plus il y a dépendance entre les variables. Dans le test TDT, seuls les parents hétérozygotes sont utilisés (Tableau 17), on a donc l'équation simplifiée suivante :

$$\chi^2 = \frac{[b - (b+c)/2]^2}{(b+c)/2} + \frac{[c - (b+c)/2]^2}{(b+c)/2} = \frac{(b-c)^2}{b+c}$$

**Protocole :** L'outil *Plink* (PLINK version 1.07 <http://pngu.mgh.harvard.edu/purcell/plink/>) (Purcell et al., 2007) a été utilisé pour réaliser les analyses statistiques.

*Plink* permet de réaliser quelques analyses descriptives, comme le calcul de fréquence allélique (--freq) ou la structure du DL dans un échantillonnage (--r2) et de contrôles qualité des données (contrôle de filiation (--mendel), test d'équilibre de Hardy-Weinberg : --hardy).

Pour le TDT, le fichier d'entrée contient le génotype pour des familles complètes, c'est-à-dire avec les deux parents et les descendants atteints. L'allèle dont la fréquence est minoritaire est recherché pour chaque marqueur, et le nombre de fois où il est transmis par les parents hétérozygotes vers les descendants atteints, versus le nombre de fois où il n'est pas transmis au détriment de l'allèle majoritaire sont mesurés. Ces valeurs permettent d'estimer le CHI deux et la p value pour chacun des marqueurs.



La commande `--tdt` permet une analyse de l'ensemble du pedigree, la fonction `--tdt --poo` permet de ne prendre en compte que les méioses paternelles ou maternelles.

Le problème du test utilisé est que le risque  $\alpha$  de rejeter à tort  $H_0$  n'est valable que pour un seul test. Or les études réalisées testent l'association pour plusieurs marqueurs, voire pour un chromosome ou le génome entier (GWAS). Cette utilisation augmente le nombre de faux positifs. Il est possible de corriger ces tests multiples à l'aide de différentes méthodes. La correction de Sidak-Bonferroni est très robuste et donc fréquemment utilisée pour la correction des tests. Pour  $N$  événements indépendants, la probabilité de ne pas rejeter  $H_0$  à tort est  $(1-\alpha)^N$ . La correction pour chacun des événements individuels est  $\alpha_{1 \text{ test}} = 1 - (1-\alpha_{N \text{ tests}})^{1/N}$ . Cette formule a par la suite été simplifiée par Bonferroni en  $\alpha_{1 \text{ test}} \approx \alpha_{N \text{ tests}}/N$ . Ces corrections supposent l'indépendance des tests, or certains marqueurs sont en déséquilibre de liaison. Cette correction est donc très sévère et peut entraîner le rejet de certains SNP à tort.

### III.D. Construction de carte physique

**Contexte :** Dans le cadre de la cartographie du chromosome X, une carte complète de l'ensemble du chromosome a été construite à partir de marqueurs génétiques (marqueurs SNP de la puce 60K) et d'EST dans la région pseudo autosomale (PAR) de ce chromosome.

**Principe :** Les résultats de génotypage de marqueurs d'intérêt sur chacun des 90 hybrides du panel IMNpRH2 porcin permettent d'estimer leur distance les uns par rapport aux autres. En effet, deux marqueurs initialement proches ont plus de chance d'être présents sur les mêmes hybrides que deux marqueurs éloignés.

La probabilité de cassure  $\theta$  entre deux marqueurs A et B est calculée selon la formule suivante :

$$\theta = \frac{\left[ 1 - \left( \frac{2A^-B^-}{A^- + B^-} \right)^{\frac{1}{2}} \right]}{\left[ 1 - \left( \frac{A^- + B^-}{2N} \right)^{\frac{1}{2}} \right]}$$

Avec  $A^-$  le nombre d'hybrides où A est absent,

$B^-$  le nombre d'hybrides où B est absent,

$A^-B^-$  le nombre de d'hybrides où A et B sont absents ensemble,

Et  $N$  le nombre d'hybrides pour lesquels la présence de A et de B a été analysée.



La probabilité de cassure (Haldane, 1919) permet d'estimer la distance entre les marqueurs A et B selon la formule :

$$D_{AB} = -\ln(1 - \theta)$$

La distance des marqueurs deux à deux est utilisée pour estimer leur position les uns par rapport aux autres, et les localiser par rapport à des marqueurs dont la position sur les chromosomes est déjà connue.

**Protocole :** La carte physique du chromosome X a été établie selon trois étapes : (1) la construction d'une carte dite robuste, dont la probabilité est très significative, (2) l'intégration à la carte robuste de marqueurs annotés sur le chromosome X, mais non inclus initialement dans la carte robuste, ainsi que l'intégration des marqueurs sans position sur l'assemblage porcine (annotés sur des *unplaced scaffolds* sur Sscrofa 10.2) et (3) l'intégration des gènes de la région pseudo-autosomale (PAR) ancestrale des mammifères.

*Construction de la carte robuste :* A partir des 1586 marqueurs annotés sur le chromosome X et dont les données RH2 sont disponibles, des groupes de liaison ont été établis avec l'outil CarthaGene (<http://www.inra.fr/mia/T/CarthaGene/>) (de Givry et al., 2005) et l'option **group**. Seuls les groupes de 6 marqueurs liés (*LOD score* = 6) ou plus ont été conservés pour la carte physique. Les données RH2 pour ces marqueurs et les positions sur l'assemblage porcine (Sscrofa 10.2) ont été mergées dans CarthaGene. (Faraut et al., 2007) et une carte a été générée avec la commande LKH (pour *Lin-Kernighan heuristic*). Cette carte a été améliorée à l'aide de la commande MCMC (pour *Markov Chain Monte Carlo*). 5000 itérations ont été réalisées et les 1000 dernières ont été utilisées. Les cartes dont la probabilité est la plus forte ont été sélectionnées pour établir la carte robuste, grâce à l'outil *Metamap* (<http://snp.toulouse.inra.fr/~servin/index.cgi/Metamap>) (Servin, de Givry et al. 2010).

*Intégration des autres marqueurs annotés sur le chromosome X et des marqueurs sans position :* Des groupes de liaison ont été établis avec les données RH2 des marqueurs de la carte robuste (979 marqueurs) et les marqueurs non insérés et annotés sur le chromosome X (607 marqueurs), ainsi que les marqueurs sans position (2050 marqueurs). L'option **group** est similaire à celle citée précédemment. Pour chaque groupe de 6 marqueurs ou plus, une carte complète a été construite dans *CarthaGene* avec la fonction **buildfw**.



L'option `MrkTest` égale à 1, et la liste de marqueurs (option `MrkList`) correspond à l'ordre des marqueurs de la carte robuste. La carte la plus probable a été sélectionnée. Lorsque, pour un même marqueur, plusieurs positions sont équiprobables, les différentes positions sont conservées.

*Intégration des gènes de la PAR :* Enfin, des groupes de liaison ont été établis avec les données RH2 de la carte compréhensive précédemment obtenue et celles des 27 gènes de la PAR. Parallèlement des groupes de liaisons ont été recherchés entre ces 27 gènes et des marqueurs sans position sur l'assemblage. Les options pour les groupes de liaison et pour la construction de la carte complète avec les gènes restent similaires à celles utilisées précédemment.

**Résultats :** L'outil *MapChart version 2.0* (Voorrips, 2002) a été utilisé pour la représentation graphique de la carte compréhensive. L'ordre des gènes de la PAR porcine a été comparé à celui du chien, du bovin, du cheval et de l'homme (Das et al., 2009; Helena Mangs and Morris, 2007) et représenté graphiquement avec *MapChart*.



## IV. Analyses des données de séquençage

### IV.A. Recherche d'éléments répétés *in silico*

**Contexte :** Une analyse de séquence *in silico* a été réalisée dans la région de 50 Mb présentant un faible taux de recombinaison sur le chromosome X. L'objectif était de rechercher la présence de motifs particuliers (motifs répétés en tandem, motifs répétés-dispersés) dont la fréquence serait corrélée avec cette absence de recombinaison.

#### IV.A.1. Recherche de motifs répétés en tandem.

**Principe :** L'outil *Tandem Repeats Finder* ou TRF (Benson, 1999) permet de mettre en évidence des motifs répétés en tandem dans une séquence d'ADN donnée. Deux étapes sont nécessaires pour l'identification des motifs : une étape de détection des potentielles répétitions en tandem, et une étape d'analyse permettant de valider les répétitions et d'estimer les valeurs statistiques d'alignement. La séquence est scannée par fenêtre de taille variable, définie dans les paramètres d'entrée, et les occurrences de chaque motif sont enregistrées. Les motifs de répétitions potentiels sont ensuite recherchés à partir de ces résultats. L'analyse de chacun des alignements deux à deux permet d'estimer les scores d'alignement, de non alignement et d'insertions-délétions (indels), ainsi que la séquence de la répétition totale.

**Protocole :** La recherche de motifs répétés en tandem a été réalisée à partir du chromosome X porcine (Sscrofa 10.2). Les paramètres d'analyse utilisés avec TRF sont 500 (taille des motifs de 1 à 500pb), 2 (score d'alignement défini par défaut), 7 (score d'indels, valeur la moins permissive), 50 (score d'alignement minimum), 500 (taille maximum de la séquence en tandem), 80 et 10 (paramètres de détection définis par défaut, 80 pour la probabilité d'alignement, 10 pour la probabilité d'indel).

Les motifs identifiés sont ensuite recherchés sur le génome porcine (Sscrofa10.2) à l'aide de l'outil *Glnt* (Courcelle et al., 2008), permettant les alignements de séquences nucléotidiques basés sur une approche *seed-and-extend*. L'option `-align` a été utilisée et les fichiers d'entrée sont sous forme fasta. Des corrélations ont ensuite été établies entre le nombre de motifs par fenêtre non glissante d'1Mb sur tout le génome, et le taux de recombinaison des autosomes (Tortereau et al., 2012) et du chromosome X estimé préalablement dans ces travaux.



## IV.A.2. Recherche d'éléments dispersés connus et de séquences de faible complexité

**Principe :** *RepeatMasker* (Smit et al., 2010) est un programme permettant de rechercher des éléments dispersés connus et des séquences de faible complexité. Une banque de données est spécifique à chaque espèce de mammifères, et répertorie les différentes répétitions identifiées.

**Protocole :** Les motifs répétés ont été extraits de la base de données UCSC *Table Browser* (<http://genome.ucsc.edu>) (Karolchik et al., 2004) en sélectionnant les paramètres suivants :

- Version de l'assemblage porcin : Sscrofa 10.2,
- Groupe : *Variation and Repeats*
- Track : *RepeatMasker*
- Format de sortie : *BED*.

Des corrélations ont ensuite été établies entre le nombre de motifs pour chaque catégorie par fenêtre non glissante d'1Mb sur tout le génome, et le taux de recombinaison des autosomes (Tortereau et al., 2012) et du chromosome X estimé préalablement dans ces travaux.

## IV.B. Recherche de polymorphismes par comparaison de séquence

**Contexte :** Pour la cartographie fine de la région du chromosome 12, une recherche de mutations causales candidates et de polymorphismes nouveaux a été réalisée. Les données de séquence ont été obtenues à l'aide de la technique de séquençage Sanger pour les séquences du gène *SOX9* et de *TESCO*, et via le reséquençage complet de quatre individus à l'aide d'un séquenceur HiSeq2000 pour la caractérisation de l'ensemble de l'intervalle candidat du chromosome 12.

**Principe :** Dans le cadre de ces deux analyses, les comparaisons de séquence ont été réalisées en conditions homologues (comparaison avec la séquence de référence du draft 9 ou 10.2 du génome porcin). Lorsqu'une différence est identifiée entre les séquences obtenues et la séquence de référence, une valeur de "fiabilité" du polymorphisme identifié est estimée afin de rendre compte de la qualité de la séquence à cette position et aux positions adjacentes. Plusieurs outils sont disponibles pour la visualisation des alignements et la recherche de polymorphismes, et leur utilisation dépend essentiellement du type de données générées (quelques Kb ou plusieurs Mb).



**Protocole :** La recherche de polymorphismes dans les séquences de *SOX9* et de *TESCO* a été réalisée à l'aide du logiciel *CodonCode Aligner* v4.0.4 (<http://www.codoncode.com/aligner/>). Les données issues du séquenceur 3730 ont été importées directement dans *CodonCode Aligner*. Une recherche automatique de polymorphismes type SNP et Indels est réalisée par le logiciel au moment de l'alignement. L'avantage de *CodonCode Aligner* est qu'il prend en compte les indels dans l'alignement, notamment les cas d'hétérozygotie. Une vérification manuelle, base par base, est ensuite réalisée, chaque séquence pouvant être visualisée sous forme de chromatogramme.

Les séquences des deux animaux intersexués et des deux témoins, séquencés en tout génome à l'aide de la technologie HiSeq, ont été analysées avec *GATK* (*Genome Analysis Tool Kit*) (McKenna et al., 2010). *GATK* est un outil permettant la recherche de polymorphismes (SNP et indels) entre deux séquences. L'avantage est qu'il recalibre les lectures en prenant en compte les indels identifiés. De nombreux types de fichiers d'entrée sont acceptés (.fasta, .bam, .vcf ou .bed). La commande pour la recherche de polymorphismes est utilisée par défaut. Le fichier de sortie présente les différents polymorphismes identifiés, selon leur position sur les chromosomes, l'allèle de référence et l'allèle alternatif, ainsi que des valeurs de qualité d'attribution du génotype. La valeur « QUAL » indique la probabilité que le polymorphisme existe, et dépend de la qualité de la base et de la profondeur d'alignement.

L'ensemble des résultats (lectures issues des données de séquençage HiSeq, séquence de référence, polymorphismes identifiés, gènes annotés) a été visualisé grâce à l'outil *IGV* (*Integrative Genomics Viewer*) (Robinson et al., 2012; Thorvaldsdóttir et al., 2013). Les formats des fichiers d'entrée sont similaires à ceux de *CodonCode Aligner*.



# Chapitre I : GWAS pour la détection de QTL associés aux principales anomalies congénitales chez le porc

---

Les progrès récents en génomique permettent aujourd'hui de rechercher, à l'échelle du génome, des variants associés à des caractères phénotypiques. Le développement des puces de génotypage haute densité a notamment facilité ces études, et la densité en marqueurs est telle qu'elle permet de combiner les étapes de primo-localisation et de cartographie fine des régions mises en évidence. La puce porcine est disponible depuis 2009 et est composée de plus de 60.000 marqueurs SNP balisant le génome porcin. Parmi les différentes méthodes statistiques utilisées pour mettre en évidence une association entre un marqueur et un caractère, le test TDT (*Transmission Disequilibrium Test*) est très robuste et permet de pallier aux biais de stratification de la population, rencontrés avec les tests d'association simples. En effet, une association significative est uniquement identifiée si le marqueur est lié génétiquement et associé. Ce test est réalisé dans un dispositif familial de deux générations dont les descendants sont atteints. Afin de rechercher des régions du génome prédisposant les animaux aux anomalies congénitales les plus fréquentes, la première étape a consisté à mettre en place une collection (et une banque de données) permettant d'obtenir rapidement un grand nombre d'échantillons de porcelets atteints ainsi que de leurs parents. La collection a débuté en 2006, et au vu des résultats obtenus deux ans après, une première étude génétique a été entreprise pour les trois anomalies les plus fréquentes (la hernie inguino-scrotale, la cryptorchidie et l'intersexualité) dans le cadre du programme ANR SwAn.

Le chapitre I présente les étapes de la mise en place des analyses, et les résultats des tests TDT pour les trois maladies concernées. Une estimation de la qualité et de l'informativité de la puce porcine dans notre dispositif a également été réalisée.





## **I. Présentation du dispositif familial mis en place pour les études TDT**

### **I.A. Mise en place d'une collection d'échantillons d'animaux phénotypés**

L'objectif du programme SwAn est d'identifier les variants génétiques prédisposant les animaux à des anomalies congénitales. Afin d'obtenir des échantillons de porcelets atteints, deux stratégies étaient envisageables : (1) la création de familles sur une unité expérimentale de l'INRA, ou (2) la collecte d'échantillons au sein des populations commerciales. Malgré de grandes tailles de portées (douze porcelets en moyenne), une durée de gestation de 115 jours et un intervalle de génération d'un an induit une vitesse de production d'animaux relativement lente. A titre d'exemple, un premier dispositif familial permettant l'analyse de ségrégation de l'intersexualité et la production d'animaux atteints pour des études fines de physiologie a été mise en place dans les années 90 par Eric Pailhoux et Christian Legault sur le domaine expérimental de Galle. Par le croisement dirigé de parents, ayant produit des animaux intersexués au sein de premières portées, cinquante porcelets intersexués ont été obtenus en cinq ans. Le second élément pris en considération est que la création de familles rendait nécessaire le choix de la maladie à analyser. Il n'était pas envisageable de constituer des dispositifs familiaux différents pour chaque anomalie. Notre choix s'est donc porté sur la mise en place d'un système de collecte d'échantillons au sein des populations commerciales, en partenariat avec les organismes de sélection porcine (OSP).

#### **I.A.1. La collection ANOPORC**

Dans le cadre des actions innovantes soutenues par FranceAgriMer un système de collection, ANOPORC, a été mis en place à partir de 2006 entre les OSP et l'INRA. L'objectif d'ANOPORC est de collecter le maximum d'échantillons d'animaux atteints, quel que soit l'anomalie congénitale présente et la race des animaux. Lorsqu'au sein d'une portée, un ou plusieurs atteints sont identifiés, un échantillon biologique de ces animaux est prélevé (sang, morceau de queue ou d'oreille). Au moment du prélèvement, une description de l'anomalie est enregistrée par le technicien ou le vétérinaire. Les échantillons des deux parents, ainsi que de deux individus sains de la même portée et de même sexe que les atteints, sont également prélevés. Les études d'association envisagées sont basées sur un dispositif familial, il est donc important que la généalogie des individus malades soit confirmée.

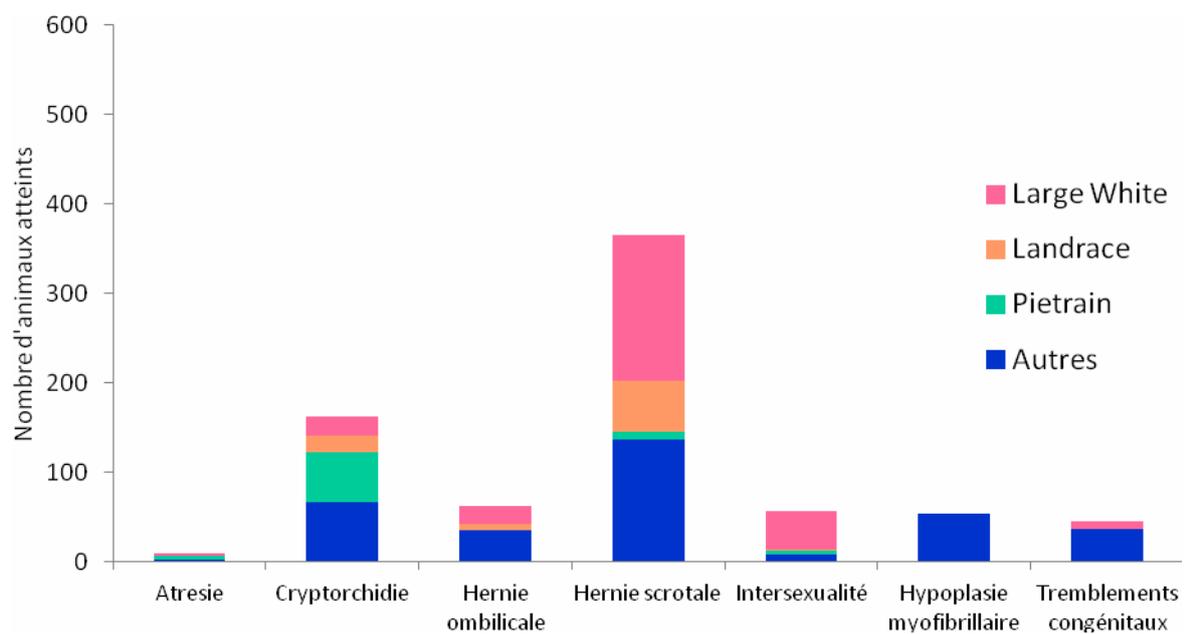


Figure 64 : bilan du nombre d'échantillons atteints collectés selon les différentes races et les principales anomalies en 2008

Le but de l'étude étant l'identification **des causes génétiques** déterminant les maladies ; tout facteur environnemental connu et maîtrisable qui pourrait augmenter la prévalence de ces anomalies est donc évité. Notamment, la castration peut augmenter la prévalence de certaines anomalies (Mattsson, 2011; Roy, 2011; Straw et al., 2009b) comme les hernies inguinales et scrotales. Le choix a donc été de ne collecter que des mâles entiers et d'écartier de la collection les mâles atteints castrés. Enfin ne sachant pas si les causes génétiques d'une anomalie sont identiques pour toutes les races, nous avons également considéré comme prioritaire de collecter des animaux de race pure afin de limiter les risques d'hétérogénéité génétique au sein des familles analysées.

Compte tenu de ces trois éléments (généalogie connue, mâles entiers, et animaux de race pure), la procédure de collecte n'a été mise en place que dans les élevages de sélection. Aucun animal issu des étages de multiplication et de production n'a été collecté et intégré à ANOPORC.

#### **I.A.2. Premier bilan de la collection ANOPORC :**

Un premier bilan a été réalisé deux ans après le début de la collection. Jusqu'en 2008, aucun *a priori* n'a été établi sur le type et la quantité d'échantillons collectés par anomalie.

Le nombre d'échantillons collectés en deux ans est présenté en Figure 64. Le système de collection s'est avéré très satisfaisant, puisque plus de 800 échantillons d'animaux atteints étaient disponibles, pour 3000 échantillons au total. Quatorze anomalies différentes avaient été enregistrées (anus imperforé, atrophie testiculaire, atrophie de la vulve, cryptorchidie, fente palatine, hernie inguinale, hernie ombilicale, hernie scrotale, hypertrophie de la vulve, intersexualité, prognathisme facial, rétention urinaire, splay leg (=hyperplasie myofibrillaire), trembleur). Cependant, la répartition des échantillons des animaux atteints par anomalie et par race était très hétérogène (Figure 64). Parmi les différentes races et anomalies, la majorité des échantillons concernait :

- la hernie inguino-scrotale en races Large White (164 atteints) et Landrace (57 atteints)
- la cryptorchidie en races Piétrain (56 atteints) et Large White (21 atteints)
- et l'intersexualité en race Large White (43 atteints).

L'objectif était d'obtenir rapidement des familles complètes de races pures afin de réaliser le plus tôt possible des études d'associations pan-génomiques.

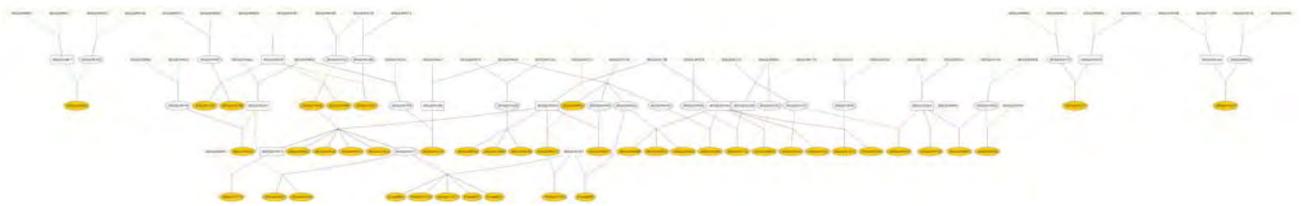


Figure 65 : pedigree INRA

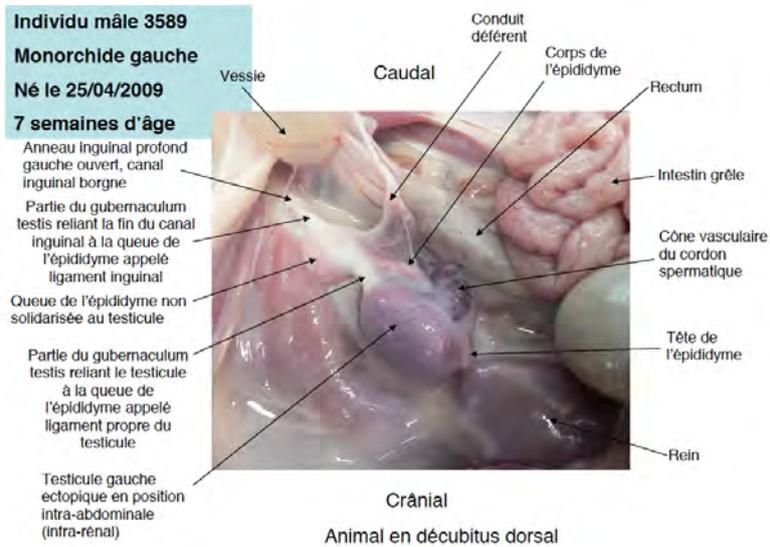


Figure 66 : Autopsie d'un animal cryptorchide

**Individu mâle 3486**  
**Hernie inguinale droite**  
**Né le 24/04/2009**  
**7 semaines d'âge**

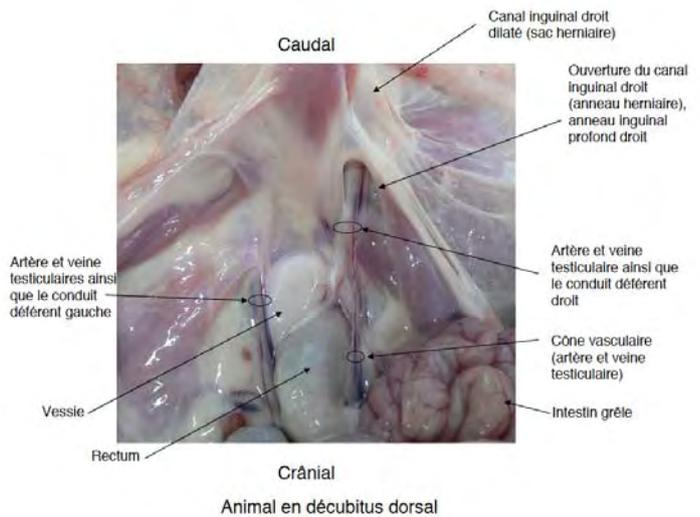


Figure 67 : Avant et après autopsie d'un animal atteint de hernie inguinale gauche

C'est pourquoi, au vu de ces premiers résultats, il a été conjointement décidé avec les OSP de privilégier la collection d'animaux atteints de hernie inguino-scrotale en race Large White et Landrace, d'animaux atteints de cryptorchidie en races Large White et Piétrain et d'animaux Large-White intersexués. Afin d'augmenter encore le nombre d'échantillons atteints, les familles d'animaux intersexués produites à l'INRA entre 1992 et 1997 (Pailhoux et al., 1997) ont été également ajoutées à la collection. Le pedigree complet de ces familles est représenté à la Figure 65. Pour constituer ce dispositif les parents de cas de mâles XX ou de vrais hermaphrodites nés en première portée, ont été croisés pour obtenir de nouveaux malades. Des croisements co-sanguins sur trois à quatre générations ont par la suite été réalisés afin d'augmenter la prévalence de la maladie dans ces familles. Au total, douze familles complémentaires (dont 51 atteints) ont ainsi été ajoutées à la collection ANOPORC.

### **I.A.3. Le phénotypage des animaux collectés**

L'élément critique dans la mise en place de ce type de collection est la multiplicité des interlocuteurs et, en conséquence, le nombre de personnes différentes en charge de la description phénotypique de l'animal atteint. Afin de s'assurer que tous les animaux étaient décrits de la manière la plus homogène possible, une fiche descriptive des caractéristiques morphologiques externes a été distribuée aux éleveurs pour ces trois anomalies majeures (fiches en Annexe 2). L'objectif de ces renseignements est essentiellement d'homogénéiser les informations transmises et non d'obtenir une description anatomique très fine de l'animal inenvisageable in vivo pour ce type de défauts.

Dans le cadre du projet SwAn une analyse phénotypique fine (euthanasie et autopsie) de cinq animaux atteints de hernie inguino-scrotale, et de cinq cas de monorchidie a été réalisé. L'autopsie a révélé que les animaux cryptorchides présentaient tous un testicule en position intra-abdominale à proximité du rein, de taille réduite, avec l'anneau inguinal profond ouvert et la queue de l'épididyme non relié au testicule (Figure 66). Chez les animaux atteints de hernie inguino-scrotale, l'anneau inguinal profond était ouvert et le conduit déférent souvent dilaté, que la hernie soit inguinale ou scrotale (Figure 67). Dans la suite de cette étude, les hernies inguinales et scrotales ont été regroupées dans la même catégorie « hernies inguino-scrotales ». Contrairement à ces deux anomalies, la caractérisation fine du phénotype intersexué est bien connue. Une analyse anatomique et histologique avait été réalisée par Eric Pailhoux en 1997 sur les animaux atteints du pedigree INRA (Figure 68).

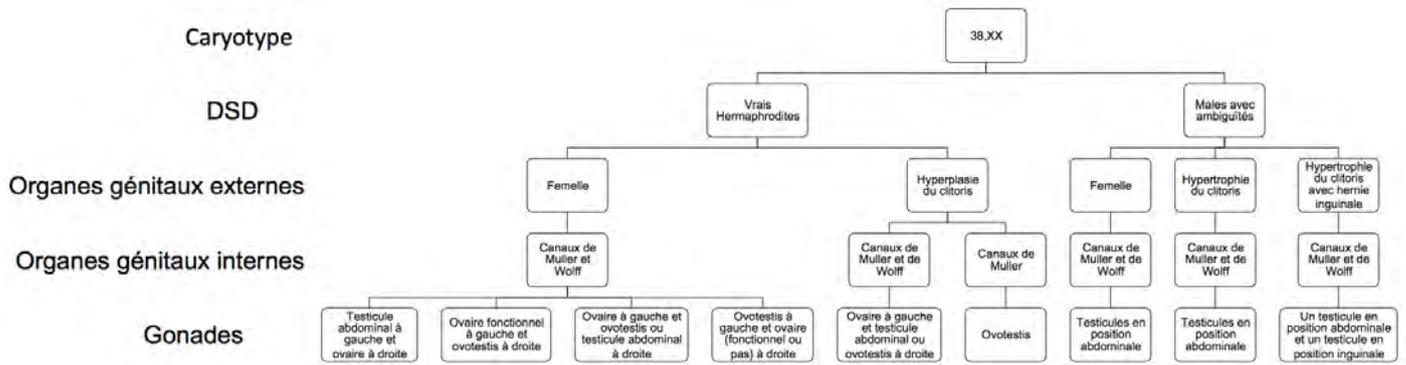


Figure 68 : Description des animaux intersexués issus du projet de Pailhoux et al. (Pailhoux et al., 1997). Le nombre d’animaux par catégorie est également réparti.

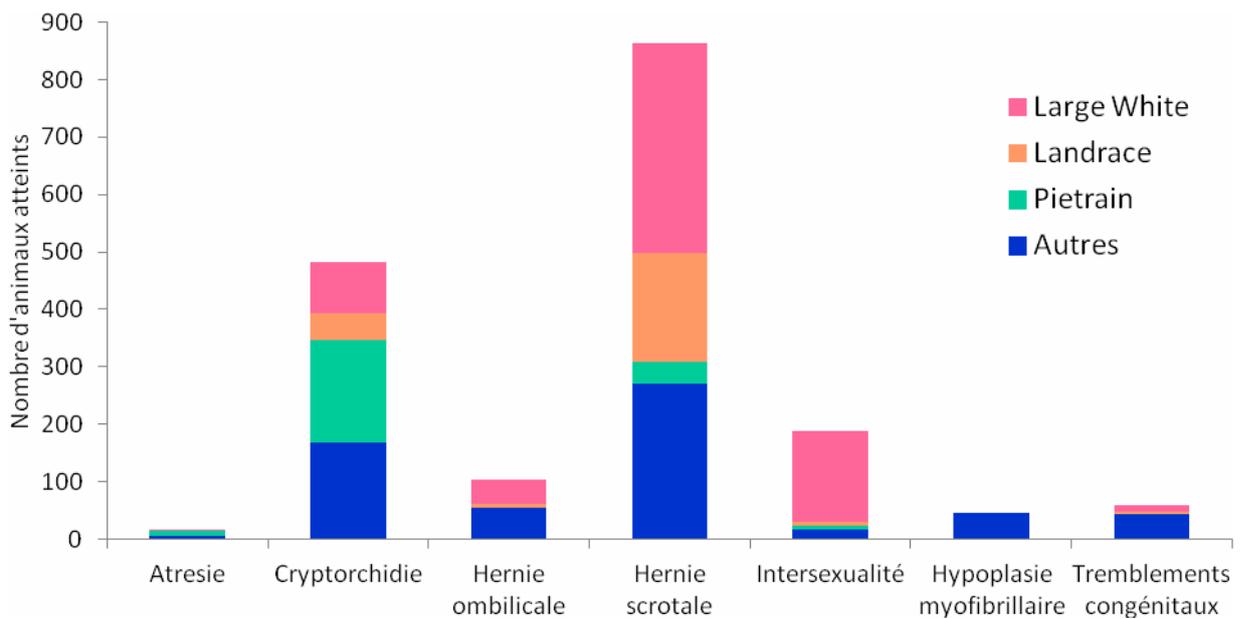


Figure 69 : bilan du nombre d’échantillons atteints collectés selon les différentes races et les principales anomalies en 2011

Ces analyses ont montré que les cas d'IS chez le porc étaient répartis en deux catégories : les vrais hermaphrodites et les mâles XX avec ambiguïtés. Au sein de ces catégories, le panel des caractéristiques sexuelles est large, mais les différentes formes sont retrouvées dans une même famille, ce qui semble indiquer que les différents types d'intersexualité sont causés par le ou les même(s) mécanisme(s) moléculaire(s).

Ces analyses fines de phénotypage, bien que succinctes, nous permettent de penser que la majorité des cas pour une maladie présente une physiopathologie identique. Au sein d'une même race et pour une anomalie, nous pouvons ainsi faire l'hypothèse que le(s) gène(s) prédisposant à ce défaut seront le(s) même(s) pour la grande majorité des atteints du dispositif.

## **I.B. Animaux génotypés dans le cadre du projet SwAn**

### **I.B.1. Critères de sélection des animaux génotypés**

A partir de 2010, nous avons estimé que le nombre d'atteints était suffisant pour débiter les analyses TDT. En effet, le nombre d'échantillons pour les races et les anomalies d'intérêt avait triplé depuis 2008, variant de 89 atteints pour les Large White cryptorchides, à 357 pour les Large White atteints de hernie inguino-scrotale (Figure 69). Parmi l'ensemble de la collection, un crible additionnel a été réalisé : le choix des familles à génotyper a été fait en fonction d'informations généalogiques et des résultats de contrôles moléculaires préalables.

**La taille des familles :** La priorité pour les analyses a été donnée aux familles complètes, c'est-à-dire pour lesquelles les échantillons des descendants atteints et **des deux parents sains** étaient disponibles. De plus, les familles comprenant au moins deux descendants atteints ont été choisies en priorité, sauf pour l'intersexualité où les familles présentant un seul atteint par famille ont été également analysées. Enfin, afin de réduire le nombre de puces de génotypage à utiliser, les familles pour lesquelles le père ou la mère étaient communs à plusieurs familles ont été sélectionnées. L'intérêt de ce type de famille est double ; cela permet (1) de réduire le nombre de puces de génotypage utilisées pour les parents sains (le génotype obtenu pour le parent commun sera utilisé pour l'analyse des deux familles) (Figure 70) et (2) de réduire la probabilité d'une cause environnementale.

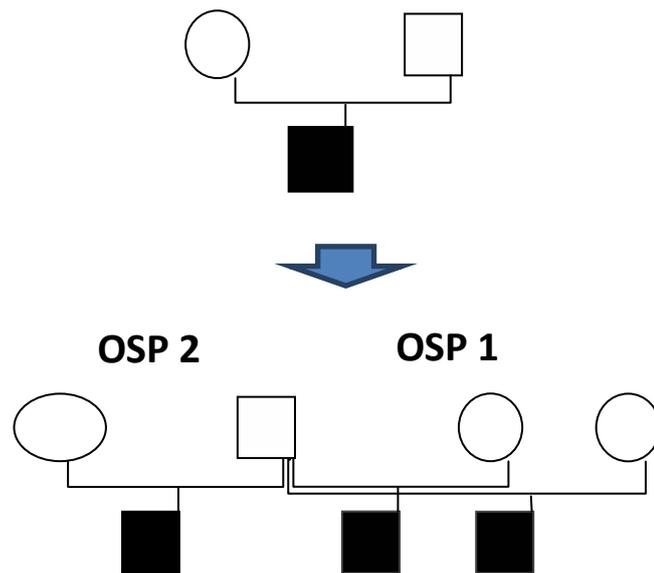


Figure 70 : Choix des familles pour lesquelles le même père a été utilisé dans plusieurs familles et dans différents OSP

Il est en effet établi que des facteurs environnementaux peuvent augmenter la prévalence des anomalies congénitales ; les familles issues d'un même verrat mais provenant de différents élevages (donc d'environnement différents) ont été ainsi privilégiées (Figure 70).

**La fiabilité des généalogies** : Des contrôles de filiation ont été réalisés préalablement, pour s'assurer des liens de parenté indiqués par les éleveurs lors du prélèvement. Le test TDT étant basé sur un déséquilibre de transmission familial, les structures des familles sont fondamentales.

Une analyse à l'aide de onze microsattellites a été réalisée pour chaque échantillon des différentes familles afin de rechercher de possibles erreurs mendéliennes. Lorsqu'un des deux parents présentait au moins deux erreurs mendéliennes sur les onze marqueurs testés, la famille était retirée des analyses. Lorsqu'au moins deux erreurs mendéliennes ont été détectées chez un des descendants atteints, cet individu n'a pas été sélectionné. Ce contrôle a permis de retirer 70 échantillons des analyses.

**Un génotype SRY-négatif** : Une cause possible d'intersexualité chez l'animal XX est la translocation du fragment du chromosome Y portant SRY, gène majeur de la différenciation mâle, sur un autosome. Parmi les intersexués XX, 6% des cas environ (Pailhoux et al., 2001b) seraient dus à la présence de SRY. L'objectif de cette thèse est d'identifier les mécanismes moléculaires impliqués dans le déterminisme de l'intersexualité chez le cochon pour les 94% de cas inexplicables. Ainsi, la présence de SRY chez les animaux atteints IS a été testée préalablement aux GWAS, et seules les familles dont les descendants IS ne sont pas porteurs de SRY ont été choisies.

D'après les analyses, 34% des animaux intersexués ont été typés « SRY positif » ainsi que 4% des mères de ces descendants atteints. Ce chiffre est très supérieur à celui obtenu par Pailhoux et al. (6%). Dans les deux études le génotypage de SRY a été réalisé de la même façon, c'est-à-dire par amplification d'un fragment de la séquence du gène. Dans les travaux de Pailhoux et al., 51 animaux IS avaient été testés, 30 venant de cas sporadiques, et 21 de grandes familles (dont 10 cas dans une des familles sur trois générations). La structure familiale très consanguine d'une partie du dispositif INRA pourrait expliquer cette différence, ce type d'échantillonnage n'étant pas adapté à une estimation de fréquence représentative d'une

population. En dehors des familles intersexuées, le même test diagnostique a été réalisé pour tous les animaux des familles HS et CR.

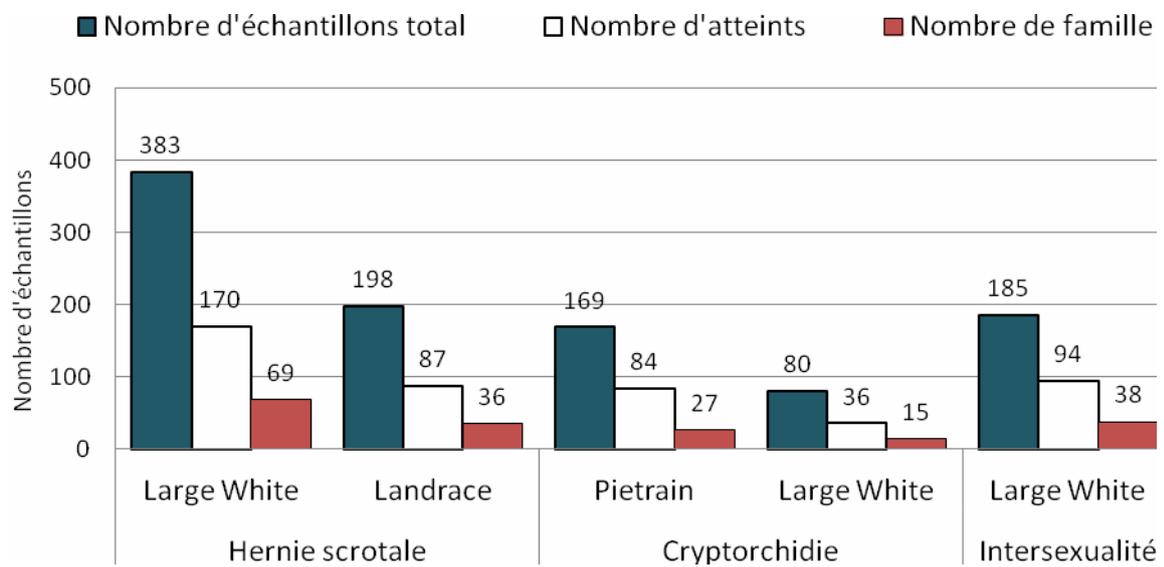


Figure 71 : Nombre total d'échantillons génotypés, nombre d'atteints et nombre de familles par anomalie et par race

Le typage s'est avéré positif pour 3% des mères chez les familles HS et CR. Il est établi que les cellules fœtales mâles peuvent circuler dans le sang maternel lors de la gestation (Wright et al., 2012). Ces données pourraient expliquer les résultats obtenus dans le cas où le prélèvement biologique des mères correspond à une prise de sang réalisée peu de temps après la mise bas ou au cours de la gestation.

Pour vérifier cette hypothèse, la date de la dernière portée des mères, ainsi que la date de l'insémination artificielle a été comparée à la date de prélèvement. Les dernières mises-bas varient de 3 à 63 jours avant le prélèvement. Pour trois femelles, les dates d'inséminations artificielles dataient de 2, 13 et 30 jours au moment du prélèvement. Ces données montrent qu'il n'y a pas de corrélation entre gestation/mise bas des truies, et la présence de Sry dans le sang maternel. Si l'on fait l'hypothèse que les 3-4% de résultats positifs obtenus sur les mères correspondent au taux d'erreur de notre méthode de génotypage, plus de 30% des cas d'intersexués de la collection peuvent néanmoins être expliqué par la présence de SRY. Sachant que 90% des cas de DSD XX chez l'homme sont dus à la présence Sry, cette valeur chez le porc semble plus plausible que les 6% observés préalablement.

#### **I.B.2. Bilan des génotypages réalisés dans le cadre de SwAn**

Au final, 1015 échantillons, dont 471 atteints, et 185 familles ont été choisis pour être génotypés (Figure 71). Les familles comprennent un à onze descendants atteints, avec une moyenne de 2-3 descendants par famille. Le nombre de familles par race et par anomalie est de:

- Hernie inguino-scrotale - Large White : 69 familles, 170 atteints
- Hernie inguino-scrotale – Landrace : 36 familles, 87 atteints
- Cryptorchidie – Piétrain : 27 familles, 84 atteints
- Cryptorchidie – Large White : 15 familles, 36 atteints
- Intersexualité – Large White : 38 familles, 94 atteints.

Tous les échantillons ont été génotypés à l'aide de la puce 60KSNP d'Illumina. Trois séries successives de génotypage, au fur et à mesure que la collection s'enrichissait, ont été réalisées entre 2010 et 2011 sur deux centres différents (CNG puis Labogena).



A l'issu du génotypage, une sélection des marqueurs et individus pour lesquels l'information génétique était suffisante a été effectuée : (1) chaque échantillon devait avoir un génotype pour 99% des SNP (valeur de call rate) et (2) pour chaque SNP, un génotype devait avoir été obtenu pour 95% des échantillons (valeur de call freq). En moyenne, 2% des échantillons par série de génotypage ont présenté un call rate < 0.99 et ont été éliminés. Comme contrôle additionnel, trois individus témoins ont également été génotypés dans chaque série d'analyses afin de vérifier la cohérence, entre les séries, des résultats de génotypage. Ces contrôles ont été réalisés avec l'aide de l'équipe SIGENAE et les génotypes ont été stockés dans une base de Biomart.

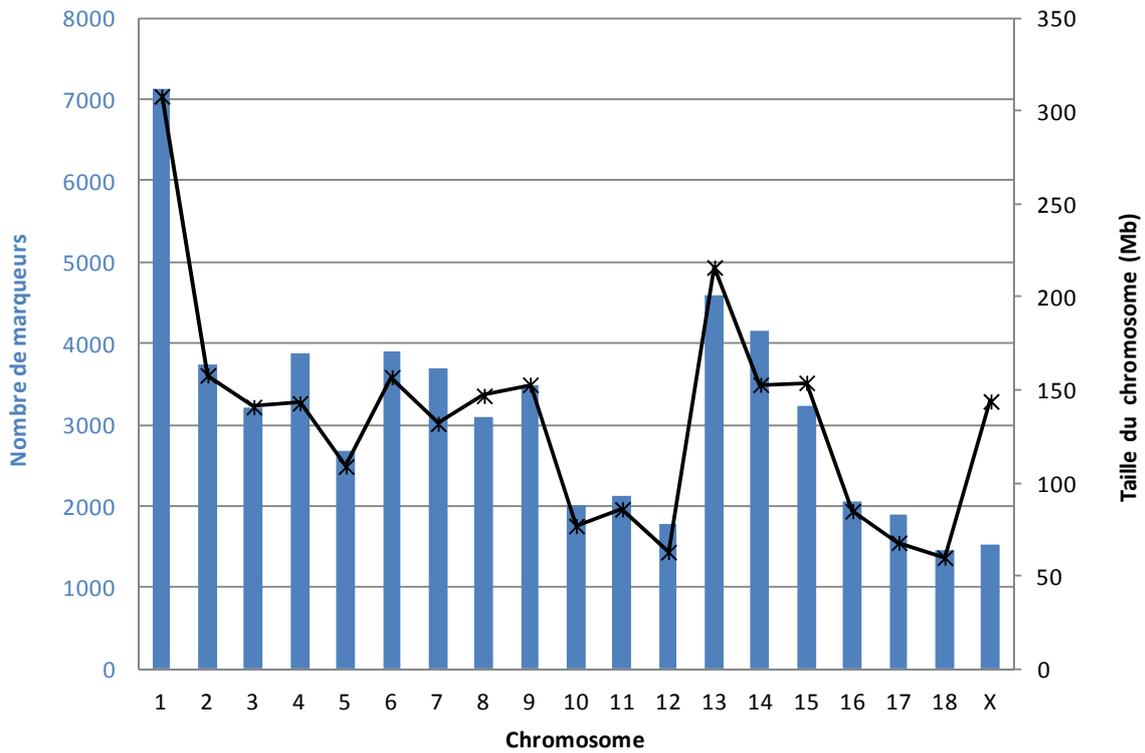


Figure 72 : Nombre de marqueurs par chromosome et selon la taille du chromosome

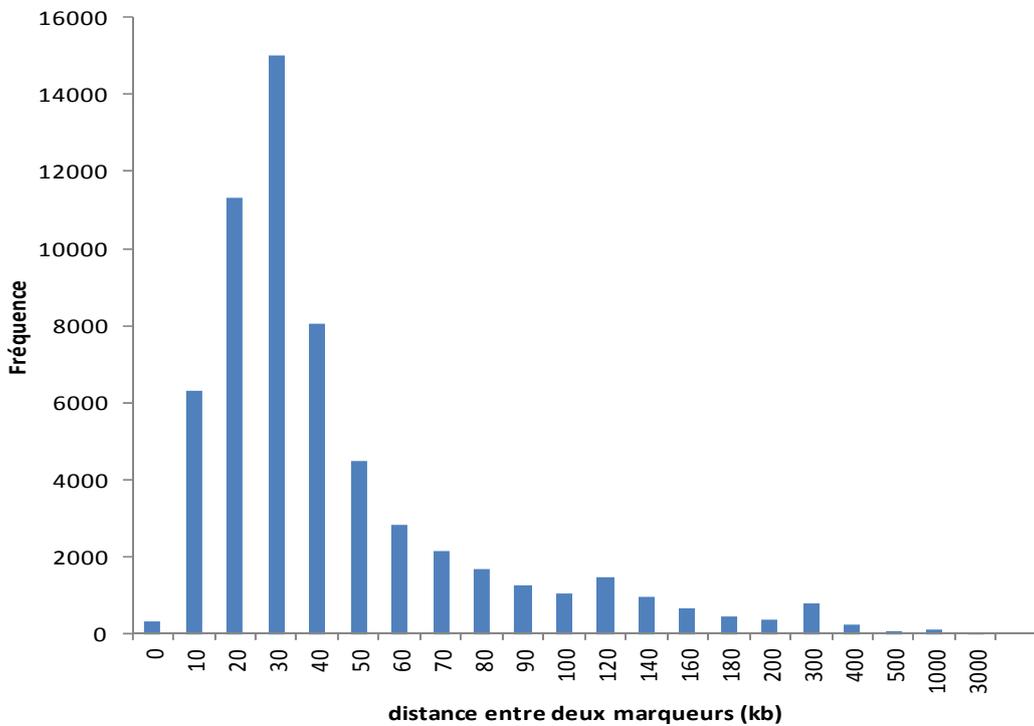


Figure 73 : Distribution des distances entre paires de marqueurs adjacents

## **II. Estimation de la qualité et de l'informativité de la puce porcine**

Depuis sa commercialisation, la puce de génotypage 60K porcine a été utilisée dans différents programmes de recherche. Des QTL impliqués dans la variabilité génétique de caractères de reproduction (Grindflek et al., 2011a; Onteru et al., 2011; Sironen et al., 2010a), d'odeurs sexuelles mâles dans la viande (Duijvesteijn et al., 2010; Grindflek et al., 2011b; Ramos et al., 2011), de production (Fan et al., 2011; Ponsuksili et al., 2011), de composition sanguine (Luo et al., 2012) et de coloration de la robe (Ren et al., 2011) ont été publiés. Dans le cadre de ces premières études, une analyse de la qualité et de l'informativité des marqueurs de la puce permet d'acquérir une bonne estimation de la fiabilité de cet outil générique de génotypage. J'ai donc réalisé une étude préliminaire aux GWAS afin d'estimer le niveau d'informativité des marqueurs dans le cadre du dispositif SwAn. Comme des analyses TDT intra race seront à terme réalisées, cette estimation de qualité des marqueurs de la puce a été faite pour chacune des races étudiées.

### **II.A. Estimation de la couverture du génome**

L'objectif d'une puce de génotypage est d'avoir une couverture homogène du génome par les marqueurs SNP de manière à pouvoir prendre en compte chacune des régions chromosomiques dans les analyses. La distribution des marqueurs par chromosome a donc été estimée à partir des SNP conservés à l'issue du tri basé sur la valeur de call freq de 95%. Pour plus de 92% des marqueurs de la puce (59651 SNP sur les 64232), le génotype est disponible dans au moins l'une des trois races (MAF de 0 à 0.5). Le nombre de marqueurs par autosome varie de 1468 SNP (chromosome 18, de taille 60Mb) à 7119 (chromosome 1, 308 Mb), soit un marqueur tous les 41,2 kb en moyenne.

Les chromosomes dont la densité est la plus forte sont SSC4, SSC7 et SSC14 (Figure 72). Les autosomes pour lesquels le nombre de marqueurs est relativement plus faible sont SSC8 et SSC16. La densité en marqueurs sur le chromosome X est la plus faible, puisqu'il n'y a que 1536 marqueurs pour 144 Mb, correspondant à une densité de 1 SNP/93.8kb.

La distance marqueur par marqueur a également été estimée, car des disparités peuvent exister entre région chromosomique. La distribution de la distance est présentée en Figure 73. Près de 90% des marqueurs sont distants de moins de 100kb. Cependant, la distance séparant certains marqueurs est de l'ordre du Mb (dix marqueurs distants de 1 à 2,7 Mb, localisés sur le chromosome X).

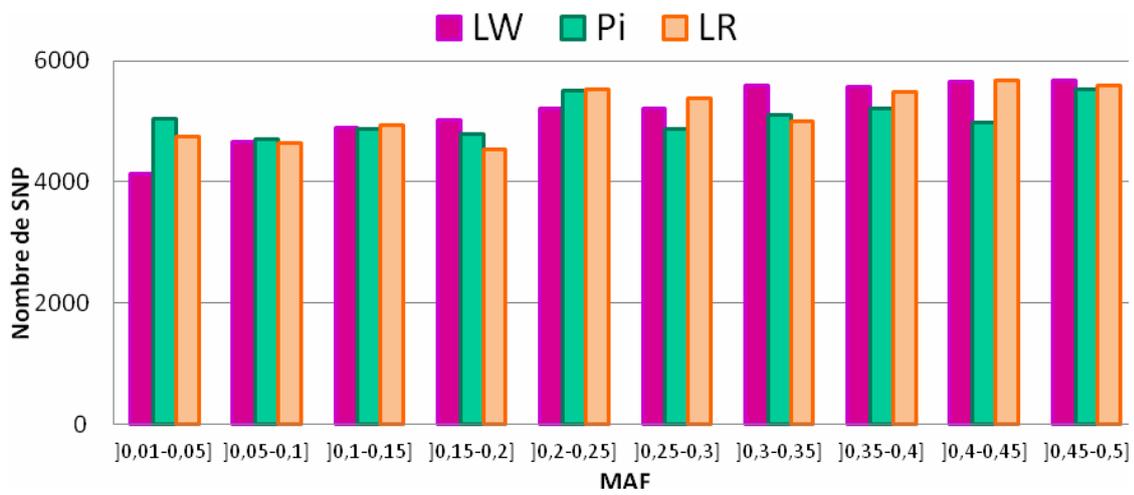


Figure 74 : Nombre de marqueurs par intervalles de MAF pour les races Large White (LW), Piétrain (Pi) et Landrace (LR)

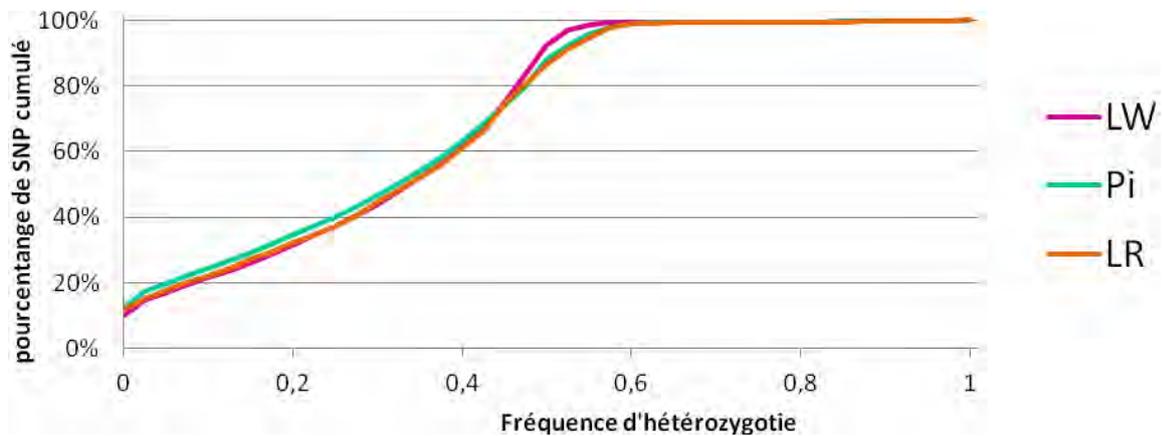


Figure 75 : Fréquence d'hétérozygotie chez les parents des descendants atteints par race

## **II.B. Informativité de la puce pour les analyses TDT**

Au delà de la répartition des marqueurs sur le génome, les marqueurs doivent être informatifs dans le dispositif analysé pour permettre la mise en évidence d'un signal d'association. L'analyse TDT prend en compte le nombre d'allèles transmis des parents hétérozygotes vers les descendants atteints. Ainsi, la statistique du test TDT est proportionnelle au nombre de parents hétérozygotes dans la population étudiée, elle est donc fonction des fréquences alléliques.

### **II.B.1. Informativité des marqueurs (MAF)**

Les marqueurs de la puce porcine ont été choisis pour être suffisamment polymorphes chez différentes races de porcs. Afin de vérifier la MAF des SNP dans les trois races étudiées, le nombre de marqueurs par intervalles de MAF a été estimé (Figure 74). Le nombre de marqueurs est équivalent pour les différentes fenêtres de valeurs de MAF en race Piétrain et Landrace. Pour la race Large white, le nombre de SNP croît légèrement lorsque la MAF augmente (4136 marqueurs dans l'intervalle] 0.01-0.05] contre 5671 dans l'intervalle] 0.45-0.5]). Ces différences entre populations sont néanmoins relativement faibles. Cependant, l'allèle minoritaire est parfois différent d'une race à l'autre. Pour la moitié des marqueurs (27322 SNP), l'allèle minoritaire est identique entre les trois races. Pour les autres, un tiers est commun à deux races et différent de la troisième (9168 SNP communs aux Landrace et aux Piétrain, 9112 communs aux Large White et aux Piétrain, et 9662 communs aux Large White et aux Landrace).

Les moyennes des MAF sont de 0,245, 0,242 et 0,233 pour les races LW, LR et Pi, respectivement. Lors de l'élaboration de la puce porcine par Illumina, les MAF moyennes, estimées à partir d'un échantillonnage indépendant des populations du dispositif SwAn, étaient de 0,26 pour les LW, 0,24 pour les LR et 0,23 pour les Piétrain. Les valeurs observées ici sont cohérentes avec les résultats d'Illumina.

### **II.B.2 Informativité du dispositif TDT - proportion de parents hétérozygotes**

La proportion de parents hétérozygotes sur le nombre de parents total a été estimée pour chacun des marqueurs, et pour chacune des races (Figure 75). Quelque soit la race, 11% des marqueurs sont monomorphes et un tiers des parents ne sont pas informatifs

(hétérozygotes) pour près de la moitié des marqueurs de la puce. Les trois courbes obtenues sont très comparables.

La valeur seuil utilisée lors des analyses TDT pour conclure à la significativité d'un signal sera une p-value inférieure à  $10^{-6}$ . D'après la table du CHI2, cette p-value (pour un degré de liberté) correspond à une valeur de CHI2 de 24. Dans le test TDT, seuls les parents hétérozygotes sont utilisés, et la valeur du CHI2 est obtenue à partir de l'équation simplifiée suivante :

$$\chi^2 = \frac{[b - (b+c)/2]^2}{(b+c)/2} + \frac{[c - (b+c)/2]^2}{(b+c)/2} = \frac{(b-c)^2}{b+c}$$

Avec  $b$  le nombre de fois où l'allèle 1 a été transmis, et  $c$  le nombre de fois où l'allèle 1 n'a pas été transmis.

Dans le cas où un des marqueurs de la puce présenterait une distorsion totale, au minimum 24 méioses informatives sont donc nécessaires pour obtenir un signal TDT significatif. Compte tenu de la proportion faible de parents hétérozygotes pour certains marqueurs, le nombre de descendants nécessaires pour "espérer" obtenir un signal significatif peut augmenter rapidement. Si pour les marqueurs les plus informatifs ( $\text{freq}_{\text{hétéro}} = 0.5$ ), 48 méioses en moyenne sont suffisantes, ce nombre passe à 80, 120 voire 240 pour les SNP dont les fréquences de parents hétérozygotes sont respectivement de 30%, 20% et 10%. L'analyse seule des familles Large-White atteintes de cryptorchidie, comprenant 36 descendants uniquement, ne nous permettra donc pas "par construction" d'obtenir un signal d'association significatif pour plus d'un tiers des marqueurs de la puce.

Enfin pour vérifier si l'informativité des marqueurs était répartie de façon homogène tout au long du génome, la proportion de parents hétérozygotes a été représentée en fonction de la position des marqueurs sur le génome.

Pour les trois races, si le pourcentage moyen de parents hétérozygotes est de 29%, des disparités existent cependant au sein du génome et entre les races. Au total, pour une centaine de régions chromosomiques de sept marqueurs ou plus ( $\pm 300\text{kb}$ ), au mieux 2% des parents sont hétérozygotes. La majorité de ces régions est observée chez les Piétrain (43 régions, contre 18 uniquement chez les Landrace et quatorze seulement chez les Large White). Enfin, six régions sont communes aux Landrace et aux Piétrain, et quatorze régions sont communes aux trois races.

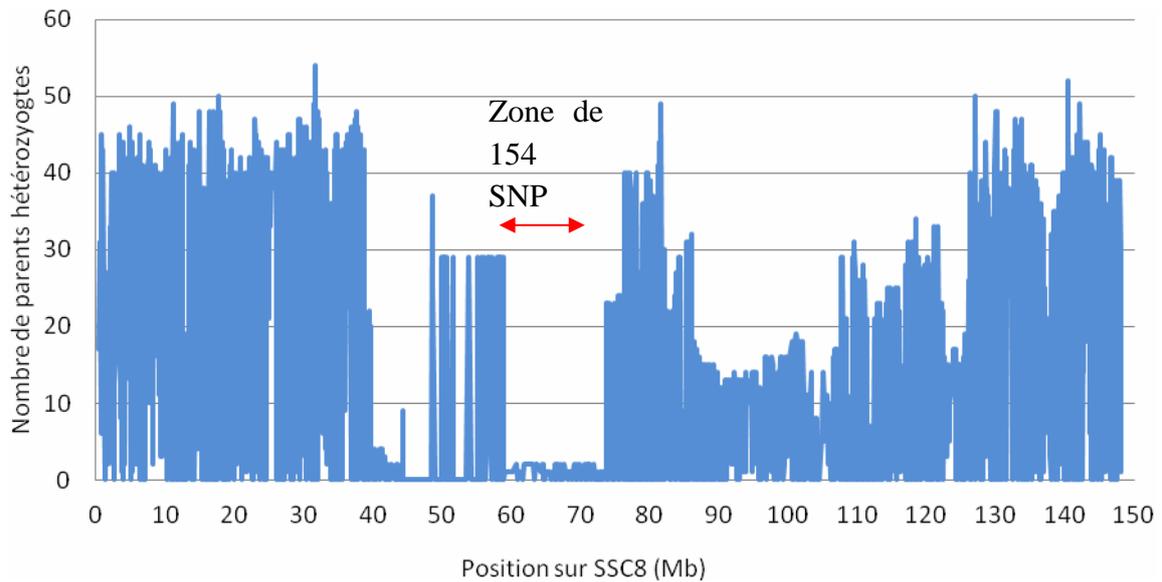


Figure 76 : Nombre de parents hétérozygotes dans les familles piétrain sur le chromosome 8

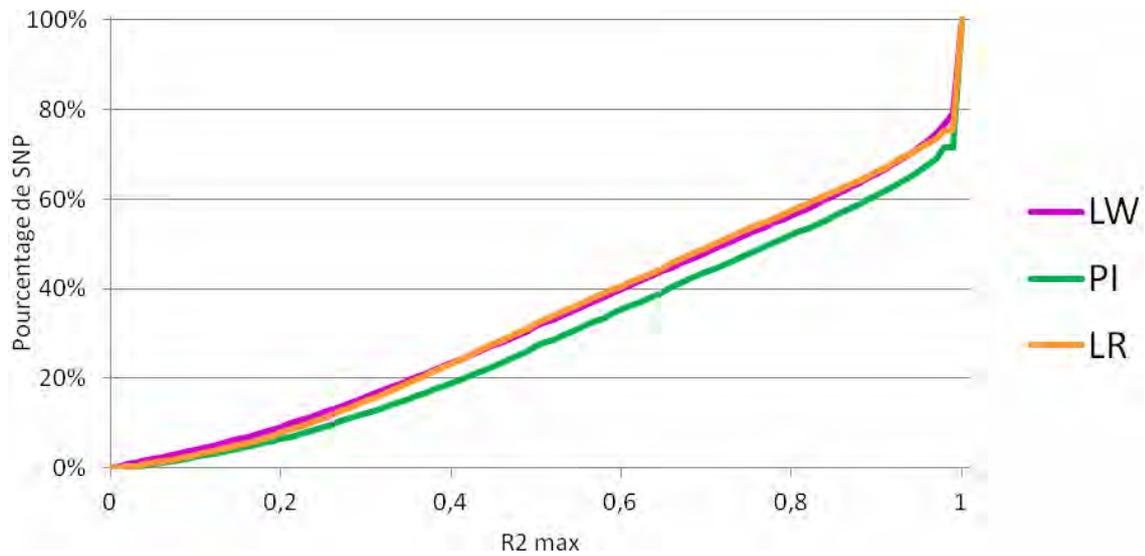


Figure 77 : Fréquence cumulée de SNP selon le  $r^2$  maximum obtenu entre chaque couple de SNP dans des intervalles de 2 Mb.

Les régions non informatives pour les trois races sont localisées sur les chromosomes :

- SSC13, entre INRA0040909 (12.788 Mb) et ALGA0071829 (12.913 Mb)
- SSC15, entre CASI0010485 (44.812 Mb) et MARC0046889 (45.224 Mb)
- SSC15, entre ALGA0085241 (53.667 Mb) et MARC0050100 (54.416 Mb)
- SSC15, entre DRGA0015150 (61.020 Mb) et MARC0022177 (61.968 Mb)
- Et 10 régions sur SSCX, entre MARC0027899 (49.617Mb) et ALGA0099838 (101.490 Mb)

Sur le chromosome X, près d'un tiers du chromosome est non informatif, alors que seules les méioses femelles sont prises en compte dans l'analyse.

Une région non informative très large est observée sur le SSC8 spécifiquement chez les Piétrain (Figure 76). Pour 154 marqueurs localisés entre les positions 58.9 à 73.8 Mb, deux parents au plus sont hétérozygotes. Près de 60 gènes sont annotés dans cette région. En race Large White et Landrace, cette région est plus informative car le nombre de parents hétérozygotes pour ces 154 marqueurs est de 32% et 22%, respectivement.

### **II.C. Estimation du déséquilibre de liaison entre marqueurs:**

Le dernier contrôle que nous avons réalisé est destiné à estimer la taille du déséquilibre de liaison (LD) dans notre échantillonnage à l'aide des marqueurs de la puce. En effet, afin de pouvoir identifier un ou plusieurs SNP associés à un caractère, il est nécessaire qu'un DL entre les marqueurs soit relativement important.

Pour réaliser ces estimations, chaque marqueur de la puce a été utilisé comme mutation potentielle afin de rechercher dans une fenêtre de deux Mb le marqueur présentant la plus forte DL. Pour chaque marqueur, la valeur du DL maximum ( $r^2_{\max}$ ) a été relevée ainsi que la distance entre le SNP de référence et le SNP présentant le  $r^2_{\max}$ . Les résultats sont présentés en Figure 77. La valeur du  $r^2_{\max}$  est inférieure à 0,5 pour 31, 32 et 27% des marqueurs (LW, LR et Pi respectivement) et 20 à 30% des SNP sont en DL total avec un marqueur à proximité selon la race (30% pour les Large White, 25% pour les Landrace et 20% pour les Piétrain).

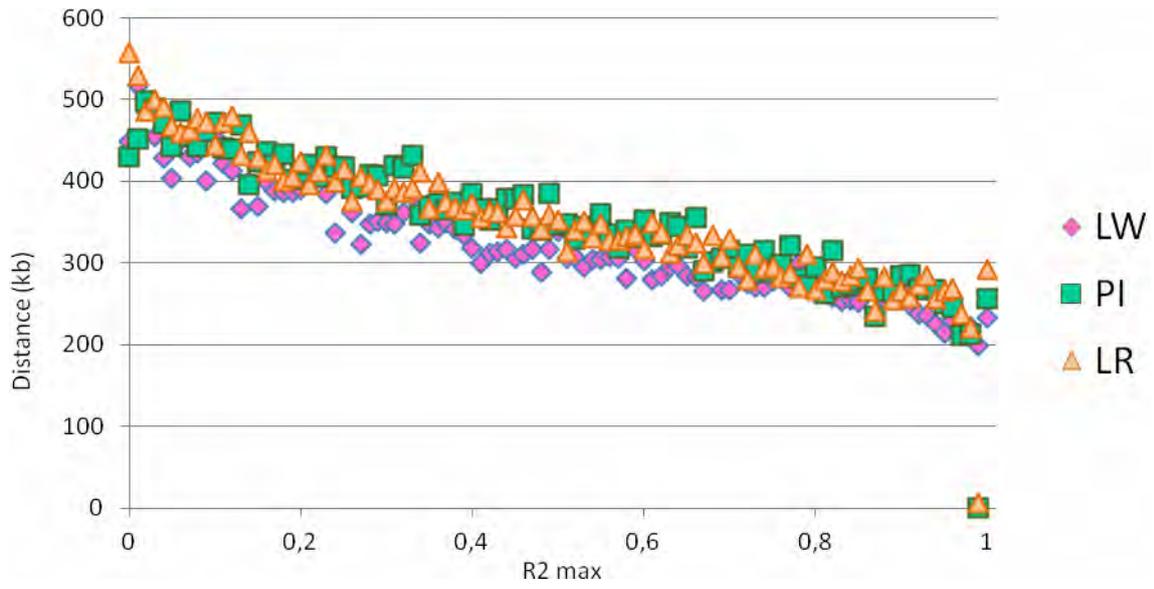


Figure 78 : Distance entre deux marqueurs selon le  $r^2$  max

Dans un second temps, les distances relevées ont permis d'estimer la taille moyenne du déséquilibre de liaison intra fenêtre de deux Mb. Les distances moyennes estimées par valeur de  $r^2_{\max}$  sont présentées en Figure 78. La distance moyenne pour un  $r^2_{\max}$  de 0,5 est de 337 kb pour les Large White, 352 kb pour les Landrace, et 346 kb pour les Piétrain. Au delà de 500 kb, il n'y a plus de déséquilibre de liaison entre les marqueurs de la puce.

Ces données signifient que 30% des marqueurs environ couvrent des régions de fort DL. Pour les autres marqueurs, il est nécessaire que la mutation soit à moins de 300kb environ pour tirer parti du DL existant et obtenir un signal d'association.

Au vu de la répartition des SNP sur le génome (un SNP tous les 41,2 kb environ sur les autosomes), les SNP sélectionnés pour constituer la puce porcine semblent adaptés à la mise en évidence de variants génétiques dans nos populations via des études GWAS.



### III. Résultats des analyses TDT pour les trois anomalies congénitales

#### III.A. Stratégie d'analyse adoptée

##### III.A.1. Hypothèses d'étude et modèles utilisés

Afin d'identifier des marqueurs SNP associés aux statuts des maladies et de mettre en évidence les allèles ségrégeant préférentiellement chez les individus atteints, des analyses TDT ont été réalisées pour les trois anomalies étudiées.

La première hypothèse que nous avons faite est qu'à chacune des trois anomalies étudiées était associé à un mécanisme biologique différent. Les analyses indépendantes ont donc été réalisées par maladie. Il n'est cependant pas exclu que des mêmes mécanismes moléculaires, intervenant dans la migration testiculaire, puissent être impliqués dans le déterminisme de la hernie inguino-scrotale (HS) et de la cryptorchidie (CR). C'est pourquoi les analyses ont également été réalisées en poolant les familles d'une même race atteintes de HS ou de CR.

Le second postulat que nous avons posé concerne la ségrégation de la ou des mutation(s) causale(s) dans les populations. Dans un premier temps, nous avons testé le scénario selon lequel les mutations causales seraient apparues **avant** la mise en place des races actuelles. L'intérêt, si ce scénario correspond à la réalité, est (1) de pouvoir pooler des familles de races différentes et augmenter ainsi la taille du dispositif et (2) de tirer parti du DL plus petit existant entre les races (comparativement au DL intra-race) pour les travaux de cartographie fine. Dans un second temps nous avons considéré que les mutations recherchées étaient apparues **récemment** (après la différenciation des trois races étudiées) et qu'elles différaient d'une race à l'autre. Pour chaque maladie, une première analyse a d'abord été réalisée en poolant les différentes familles, puis race par race.

Le test statistique utilisé est le TDT (*Transmission Disequilibrium Test*) permettant de mettre en évidence une distorsion de ségrégation au sein de familles, en tenant compte de la transmission des allèles paternels et maternels.

Cependant, il est possible que les gènes impliqués dans la maladie soient des gènes soumis à empreinte, dont l'expression dépend de l'origine parentale. Chez l'homme, plus de 200 gènes soumis à empreinte ont été identifiés et la majorité d'entre eux est impliquée dans le développement fœtal. Plusieurs maladies ont déjà été identifiées pour être associées à des gènes soumis à empreinte (Hirasawa and Feil, 2010).

Dans le cadre de SwAn, ce modèle a également été testé dans les dispositifs les plus grands.

### III.A.2. Procédure suivie et mise en forme des résultats

Les analyses ont été réalisées à l'aide de *Plink* en utilisant la fonction `-TDT`, et `-TDT-poo` pour tester des modèles de transmission parentale. L'option `--HW` a été choisie pour ne sélectionner que les SNP pour lesquels l'équilibre de Hardy Weinberg était respectée (p-value : 0.001). Sur la base de ce critère, 1000 à 1200 marqueurs ont été retirés. Dans la suite de ce chapitre, les résultats sont présentés sous forme de Manhattan plot, chaque point correspondant à un SNP et chaque couleur à un chromosome. L'axe des abscisses correspond au chromosome et à l'ordre des SNP sur chacun des chromosomes. L'axe des ordonnées correspond au logarithme de la probabilité d'obtenir la valeur de CHI2 par hasard.

Les valeurs seuils de significativité des résultats sont de 6 (p-value <  $10^{-6}$ ) pour le seuil *genome-wide* et de 4,72 (p-value <  $19 \cdot 10^{-6}$ ) pour le seuil *chromosome-wide*. Les régions significatives sont indiquées avec une flèche rouge, et les marqueurs isolés significatifs avec une flèche noire. Chaque SNP est traité indépendamment des 60.000 autres marqueurs. Ainsi, lorsque les SNP pour lesquels le signal TDT est significatif sont adjacents, les résultats semblent plus cohérents car ils permettent de mettre en évidence une région chromosomique. Lorsqu'un SNP isolé est significatif, les résultats sont à prendre avec précaution ; soit le SNP est le seul de la région à être en DL avec la mutation (région peu dense en marqueurs), soit le SNP peut être mal positionné, ou être mal génotypé (des erreurs de génotypage pouvant entraîner un faux positif).

Au vu des études précédentes, le déterminisme de l'intersexualité est supposé monogénique (Pailhoux et al., 2001a), alors que la hernie et la cryptorchidie semblent être des maladies polygéniques complexes (Ding et al., 2009; Grindflek et al., 2006). Une région significative serait donc attendue pour l'intersexualité. Concernant les deux autres maladies, il est possible qu'un ou plusieurs signaux soient observés si le nombre de méioses informatives (= le nombre de parents hétérozygotes par marqueur) est suffisant. Dans le cas où les effectifs sont faibles, comme pour les Large White cryptorchides où 15 familles comprenant 36 atteints ont été génotypés, il est possible que le nombre de parents hétérozygotes ne soit pas suffisant, et donc qu'aucun signal significatif ne soit observé.

Dans la suite de cette partie, les résultats TDT seront présentés par anomalie, pour toutes les familles, puis par race. Enfin les résultats TDT sous hypothèse d'empreinte parentale seront rapportés.

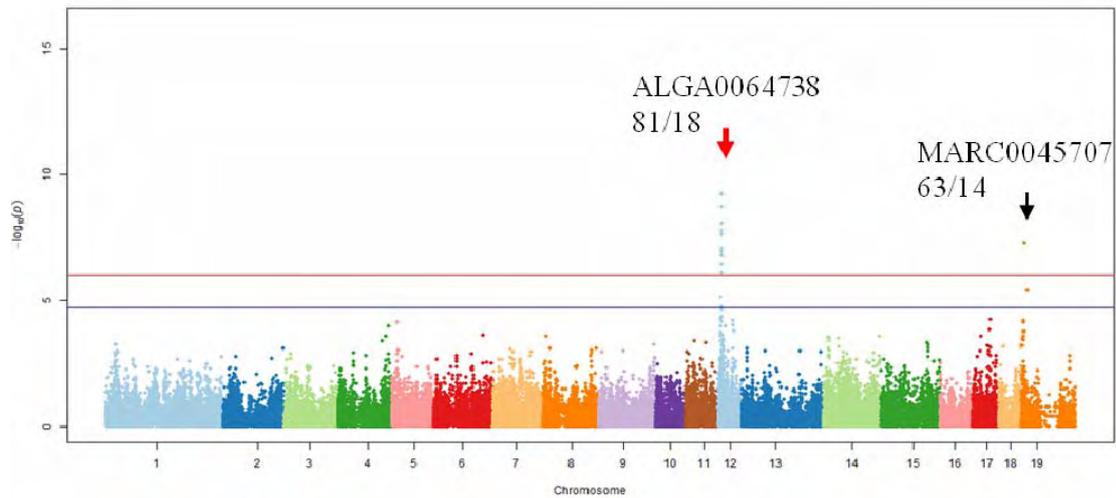


Figure 79

re 79 : Résultats TDT en tout génome pour l'intersexualité - familles Large White

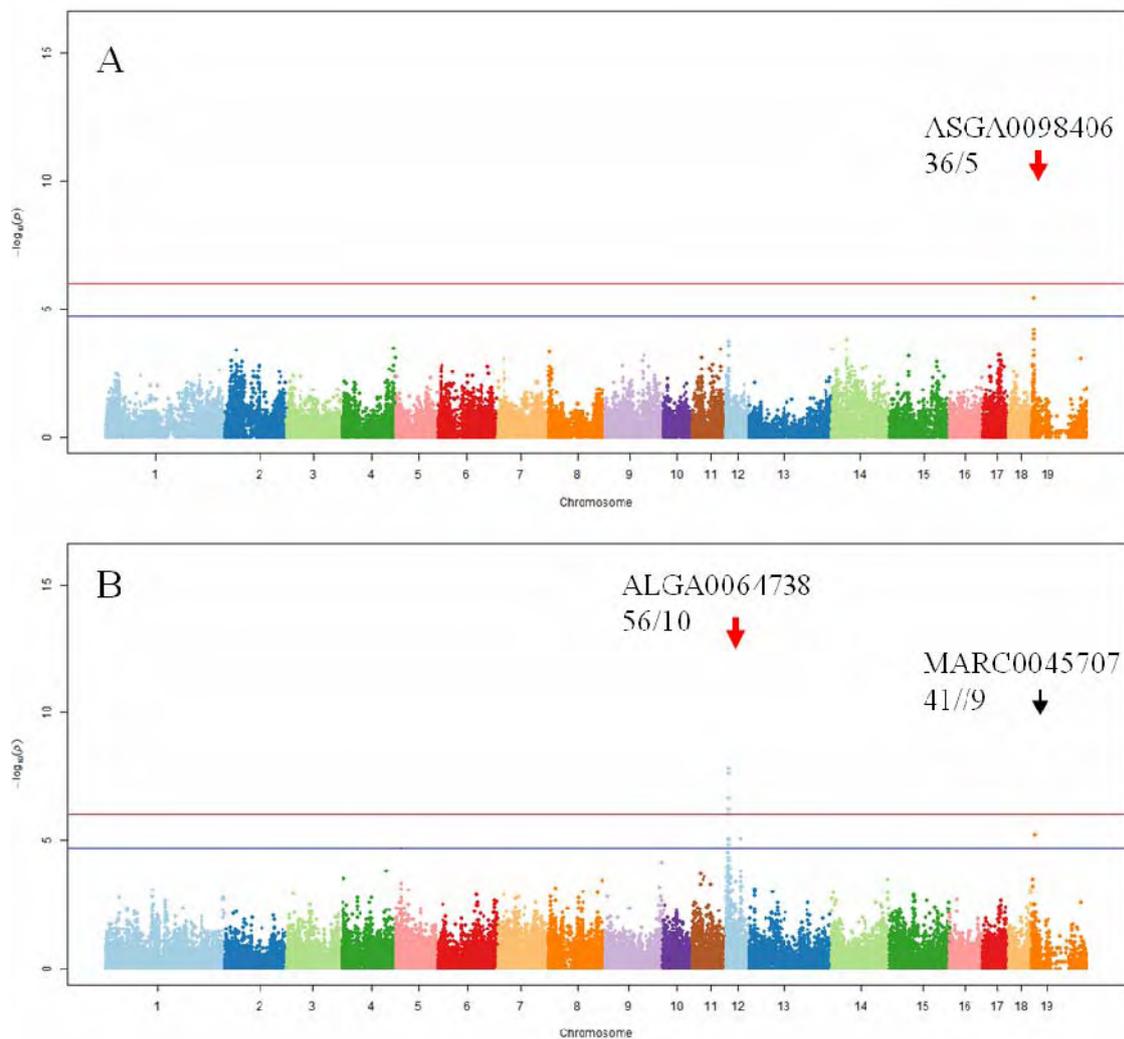


Figure 80 : Résultats TDT en tout génome pour A. les familles INRA, et B. les familles des OSP

## III.B. Résultats obtenus pour chaque anomalie congénitale

### III.B.1. L'intersexualité

Les résultats TDT pour les 38 familles Large White dont les descendants sont intersexués sont présentés en Figure 79. En moyenne par marqueur SNP, 50 méioses sont informatives.

Deux chromosomes présentent des résultats significatifs : le chromosome 12 et le chromosome X. Un marqueur isolé sur le chromosome X en position 6,79 Mb (MARC0045707) a une p-value de  $5,48.10^{-8}$ . Sur SSC12, les résultats sont significatifs pour quinze marqueurs localisés entre les positions 7,10 et 9,59 Mb (de MARC0046612 à MARC0041079). Le marqueur le plus significatif ( $p = 5,65.10^{-10}$ ) est localisé à la position 9,053 Mb. Pour ce marqueur, les animaux atteints ont reçu 81 fois l'allèle G, et 18 fois l'allèle A.

Parmi toutes les familles prises en compte dans l'analyse, deux types d'échantillons sont présents : les familles issues de la collection ANOPORC dans le cadre de SwAn (familles OSP), et les familles issues du dispositif expérimental INRA (familles INRA). Les analyses TDT ont été réalisées séparément pour ces deux groupes (Figure 80) afin de s'assurer que les signaux observés étaient observés dans ces groupes, et donc que les mêmes mécanismes moléculaires semblaient impliqués.

Pour les familles INRA, seuls deux marqueurs sur SSCX sont significatifs : ASGA0098406 (5,916 Mb) et MARC0085130 (5,972 Mb). La région du chromosome 12 n'est pas significative mais les p-values observées pour les marqueurs de la région sont parmi les plus faibles. (MARC0041079, en position 9,597 Mb, a une p-value de 0,0001828). Pour le marqueur ALGA0064738, la p-value est de 0,00607, l'allèle G est transmis 25 fois et l'allèle A 8 fois.

Pour les analyses réalisées à partir des familles OSP seules, les résultats les plus significatifs ont été obtenus sur SSC12 (entre ASGA0098350, 8,826 Mb et ASGA0097657, 9,449 Mb). Le signal le plus fort ( $p = 1,49.10^{-8}$ ) est observé pour le marqueur ALGA0064738, 9,053 Mb. Pour ce marqueur, l'allèle G est transmis 56 fois, et l'allèle A 10 fois. Un signal significatif est également présent sur SSCX, pour MARC0045707 en position 6,798 Mb.



Dans le cadre de la recherche de gènes prédisposant les animaux à l'intersexualité un signal très significatif a été obtenu sur le chromosome 12. Nous verrons par la suite qu'un seul gène est localisé dans l'intervalle défini par les marqueurs les plus significatifs, et que ce gène est un acteur majeur de la différenciation sexuelle chez le mâle. Les travaux de cartographie fine de la région sont traités dans le **chapitre II** de la thèse.

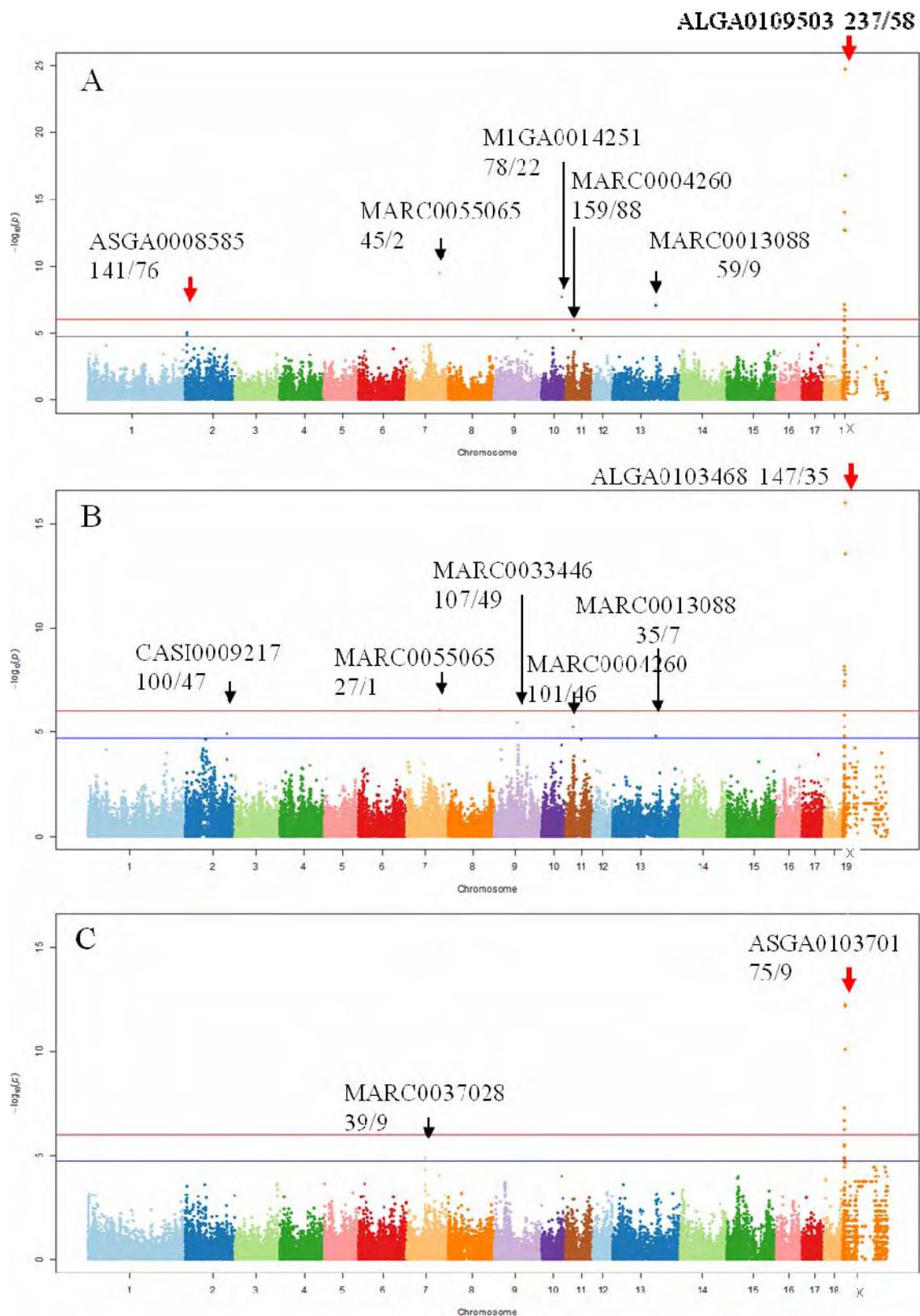


Figure 81 : Résultats TDT en tout génome pour la hernie scrotale - familles Large White+Landrace (A), familles Large White seules (B) et familles Landrace seules (C).

### **III.B.2. La hernie inguino-scrotale**

Toutes les familles (Large White et Landrace) ont d'abord été traitées ensemble, dans l'hypothèse d'un mécanisme moléculaire commun aux deux races. Puis les analyses ont été réalisées race par race.

#### ***Résultats TDT obtenus à partir de l'ensemble des familles***

Les analyses ont été réalisées avec les 105 familles Large White et Landrace, soit à partir de 180 méioses informatives en moyenne par marqueur. Les résultats sont présentés en Figure 81. A.

Ces analyses ont permis de mettre deux régions et quatre marqueurs isolés en évidence. La plus grande région (1,7 Mb) est localisée sur le chromosome X. 18 marqueurs (entre ASGA0080573 et MARC0045707) sont significatifs entre les positions 5,09 et 6,79 Mb. Une deuxième région est localisée sur le chromosome 2 et concerne deux marqueurs (ASGA0008585 et DIAS0002167, respectivement en positions 4,77 et 4,83 Mb). Enfin, quatre marqueurs isolés semblent également associés à la maladie sur les chromosomes 7, 10, 11 et 13 (MARC0055065 sur SSC7 à 104,61 Mb, M1GA0014251 sur SSC10 à 65,09 Mb, MARC0004260 sur SSC11 à 23,73 Mb et MARC0013088 sur SSC13 à 140,49 Mb).

En comparant les résultats obtenus pour l'intersexualité et pour la hernie, un marqueur est significatif dans les deux analyses ; il s'agit de MARC0045707 (Intersexualité :  $p = 5,48.10^{-8}$ , hernie :  $p = 1.70.10^{-17}$ ). Cependant, l'allèle préférentiellement transmis n'est pas le même entre les deux maladies : les intersexués ont reçu 63 fois l'allèle T, et 14 fois l'allèle C, et les animaux atteints de hernie ont reçu 61 fois l'allèle T, et 198 fois l'allèle C).

#### ***Résultats obtenus à partir des familles Large White seules***

En ne sélectionnant que les familles Large White HS, 69 familles sont disponibles pour les analyses, correspondant en moyenne à 110 méioses. Les tests TDT sont significatifs pour six chromosomes (Figure 81, B). Un seul groupe de marqueurs adjacents est présent, localisé sur le chromosome X, entre les positions 4,83 et 6,79 Mb (12 marqueurs, de ALGA0099025 à MARC0045707). Le signal maximum est observé pour le marqueur ALGA0103468 en position 6,77 Mb ( $p$ -value =  $1.024.10^{-16}$ ). Cette région correspond à celle identifiée précédemment en poolant les familles LW et LR.

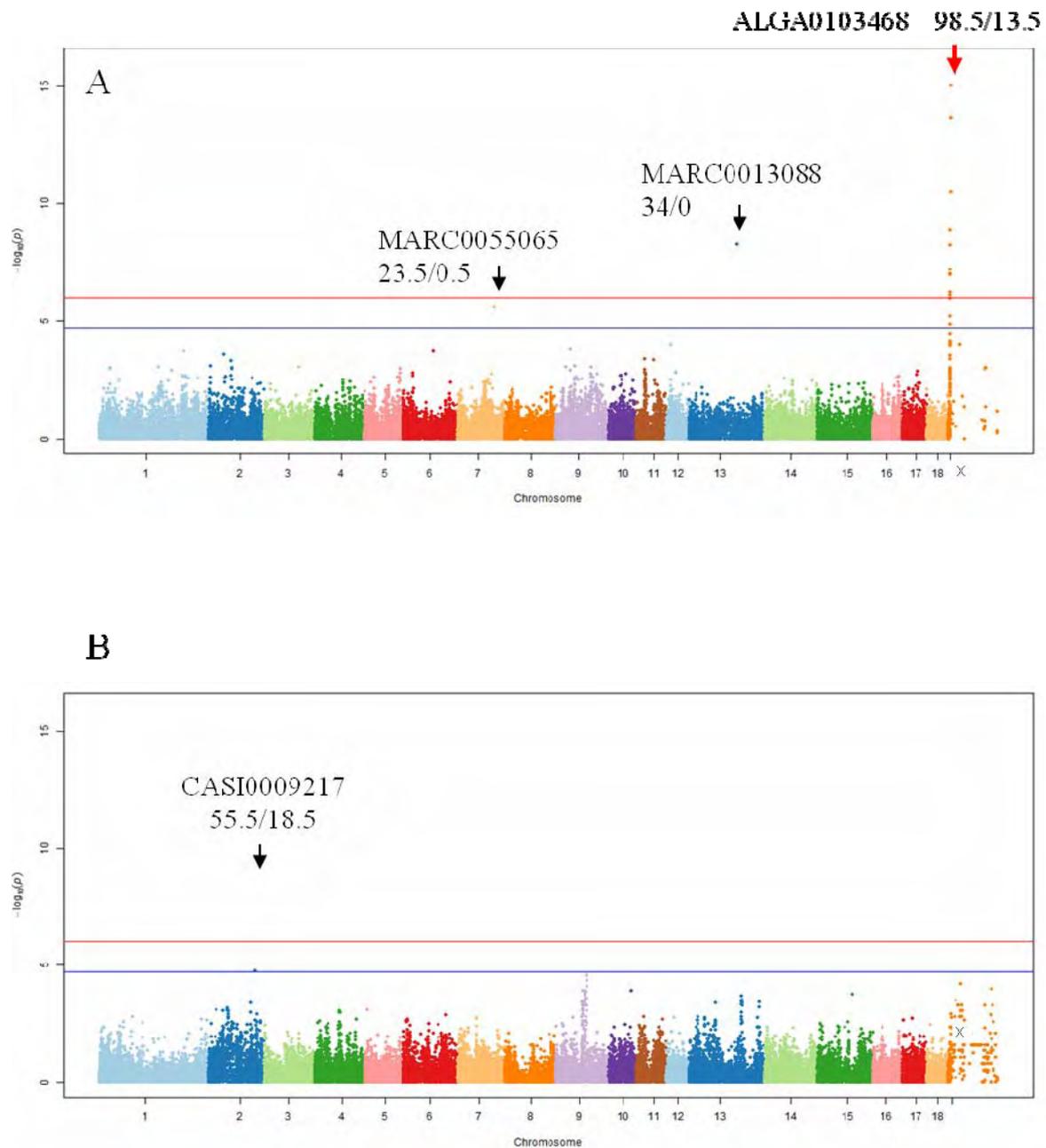


Figure 82 : Résultats TDT en tout génome pour la hernie scrotale selon l'origine parentale. A. brin paternel, B. brin maternel.

Cinq marqueurs isolés sont également mis en évidence (CASI0009217 sur SSC2 à 138.02 Mb, MARC0055065 sur SSC7 à 104,61 Mb, MARC0033446 sur SSC9 à 74,62 Mb, MARC0004260 sur SSC11 à 23,73 Mb et MARC0013088 sur SSC13 à 140,49 Mb).

Parmi les marqueurs isolés, les SNP sur les chromosomes 7, 11 et 13 sont les mêmes marqueurs que ceux identifiés en poolant les familles LW et LR.

### ***Résultats TDT obtenus à partir des familles Landrace***

Des analyses TDT ont également été réalisées en ne considérant que les familles Landrace (36 familles, 70 méioses informatives) (Figure 81. C). Deux chromosomes présentent des signaux significatifs : SSC7 et SSCX. Sur le chromosome 7, un seul marqueur a une p-value significative ; il s'agit de MARC0037028, à 59,26 Mb. Ce marqueur n'est pas significatif lorsque toutes les familles HS sont poolées.

Sur le chromosome X, la même région que celle identifiée précédemment est significative. Elle est localisée entre 5,80 et 6,79 Mb (onze marqueurs, de ASGA0080690 à MARC0045707). Au-delà de l'hypothèse de mécanismes race-spécifiques, la significativité des résultats entre les Large White et les Landrace semble néanmoins proportionnelle à la taille des effectifs des deux dispositifs familiaux.

Les résultats de l'analyse TDT selon l'origine parentale sont présentés en Figure 82. Au vu du manque de puissance avec les familles Landrace, seules les familles Large White ont été étudiées.

### ***Résultats obtenus sous une hypothèse d'empreinte parentale***

*Origine paternelle* : Une région est significative sur le chromosome X de 4,83 à 6,79 Mb (13 marqueurs, de ALGA0099025 à MARC0045707). Le signal maximum est observé pour le marqueur ALGA0103468 en position 6,77 Mb (p-value =  $9.61.10^{-16}$ ).

Pour ce marqueur, les pères ont transmis 98,5 fois l'allèle C, contre 13,5 l'allèle A, et les mères ont transmis 48,5 fois l'allèle C contre 21,5 fois l'allèle A (p = 0,00125). On peut donc supposer que la distorsion de ségrégation observée dans la région est essentiellement due à une distorsion de transmission des allèles paternels.

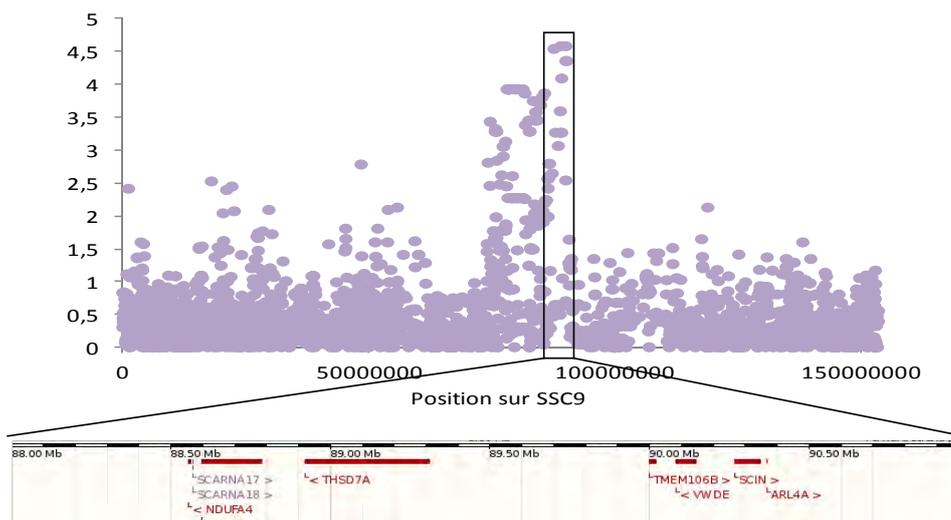


Figure 83 : Résultats TDT selon l'origine maternelle sur le chromosome 9

CHR	SSC2		SSC2	SSC7	SSC7	SSC9
Mrk	ASGA0008585	DIAS0002167	CASI0009217	MARC0037028	MARC0055065	MARC0033446
pos	4771317	4839440	138026728	59269003	104612789	74629827
HS	<b>4,991</b> 141/76	<b>4,827</b> 126/66	3,361 140/87	0,977 94/71	<b>9,449</b> 45/2	4,542 149/85
HS LW	3,003 74/39	2,251 61/34	<b>4,908</b> 100/47	0,286 55/62	<b>6,084</b> 27/1	<b>5,466</b> 107/49
HS LR	2,486 67/37	3,093 65/32	<b>0</b> 40/40	<b>4,827</b> 39/09	4,017 18/1	0,304 42/36

CHR	SSC10	SSC11	SSC13	SSCX		
Mrk	M1GA0014251	MARC0004260	MARC0013088	ALGA0109503	ALGA0103468	ASGA0103701
pos	65098291	23735278	140497748	6722207	6769971	6778001
HS	<b>7,669</b> 78/22	<b>5,204</b> 159/88	<b>7,026</b> 50/9	<b>24,705</b> 237/48	<b>6,677</b> 184/97	<b>12,224</b> 77/9
HS LW	4,371 49/16	<b>5,242</b> 101/46	<b>4,808</b> 35/7	<b>13,5</b> 139/38	<b>15,99</b> 147/35	0,803 2/0
HS LR	3,995 29/38	0,96 58/72	2,791 15/2	<b>12,158</b> 98/20	1,921 62/37	<b>12,224</b> 75/9

Tableau 18 : Bilan des marqueurs significatifs pour la hernie inguino-scrotale. Le fond vert indique que les marqueurs sont significatifs pour les analyses combinant les familles des deux races, le fond rose ne concerne que les familles LW, et le fond orange que les familles LR. HS : hernie inguino-scrotale, LW : Large White, LR : Landrace, CHR : chromosome, SSC : chromosome porcine, Mrk : marqueur, pos : position en pb sur le génome de référence porcine.

Deux marqueurs isolés sont également significatifs selon l'origine paternelle : MARC0055065 sur SSC7 à 104,61 Mb et MARC0013088 sur SSC13 à 140,49 Mb. Le nombre de méioses informatives pour les femelles pour ces deux marqueurs est très faible (quatre pour MARC0055065 et huit pour MARC0013088), il est donc difficile de conclure à une distorsion de ségrégation d'origine paternelle uniquement.

*Origine maternelle* : Un seul résultat est significatif sur le chromosome 2 en position 138,02 Mb (CASI0009217). Bien que non significative, une région se dégage sur le chromosome 9 (Figure 83). Six marqueurs ont une p-value supérieure à  $10^{-4}$  entre les positions 87,59 et 90,18 Mb. Dans cette région, plusieurs gènes sont annotés chez le porc, notamment NDUFA4, identifié chez l'homme pour être soumis à empreinte paternelle. Bien que les analyses manquent de puissance avec ce dispositif, il serait intéressant de refaire les analyses avec un plus grand nombre d'échantillons.

Le Tableau 18 est un bilan des signaux significatifs observés race par race, et en poolant les races pour les analyses réalisées à partir des familles atteintes de hernie inguino-scrotale. D'après ces résultats, certains des signaux significatifs résultent de la contribution des familles de deux races (en vert), et d'autres semblent spécifiques de la race LW ou LR (rouge et orange, respectivement). Il est ici très clair que le faible nombre de famille est limitant, et qu'augmenter la taille du dispositif en combinant les différentes familles augmente la puissance du test. A titre d'exemple, la région de deux marqueurs sur le chromosome 2 (ASGA0008585 et DIAS0002167) n'est identifiée que lorsque l'on pool toutes les familles. Si l'on regarde le nombre d'allèles 1 transmis et non transmis par les parents Large White, et les parents Landrace seuls, les valeurs sont 74/39 pour les LW et 67/37 pour LR, les valeurs du  $-\log(p\text{-value})$  sont donc de 3 et 2,49, respectivement. En poolant les deux races, l'allèle 1 est alors transmis 141 fois contre 76 fois ou il n'a pas été transmis. Ce ratio permet d'obtenir un  $-\log(p\text{-value})$  de 4,99.

En bilan, les signaux retenus pour toutes les familles (Large White + Landrace) sont :

- **SSC2** (ASGA0008585 et DIAS0002167),
- **SSC7** (MARC0055065),
- **SSC13** (MARC0013088),
- **SSCX** (signal maximum : ALGA0109503) d'origine paternelle.

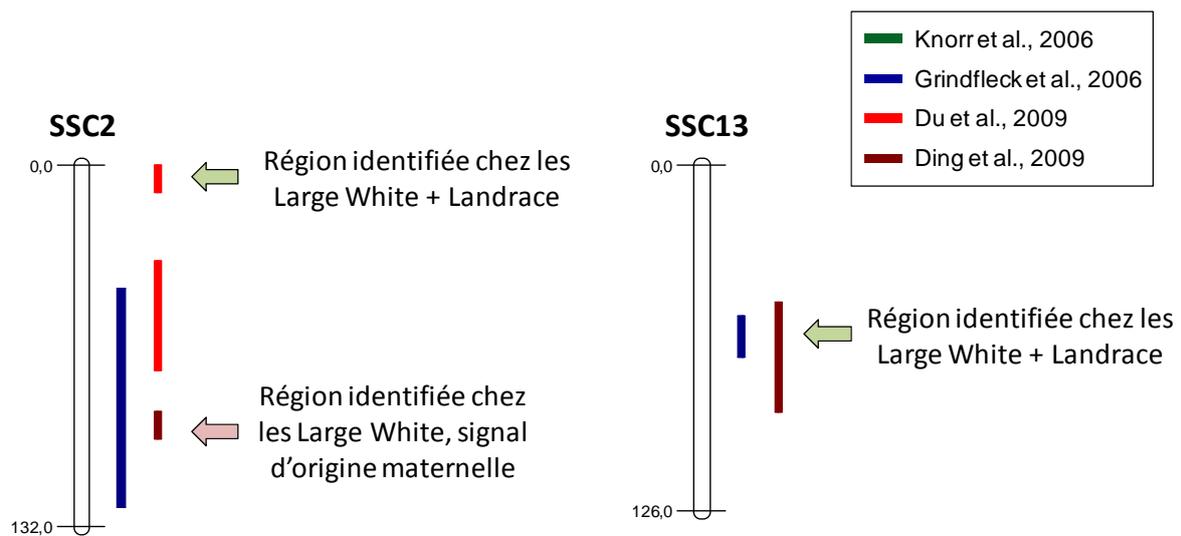


Figure 84 : comparaison entre les signaux obtenus dans cette étude et les résultats des études précédentes

Pour les Large White seulement, quatre chromosomes sont impliqués :

- **SSC2** (CASI0009217) d'origine maternelle,
- **SSC9** (MARC0033446),
- **SSC10** (M1GA0014251)
- et **SSC11** (MARC0004260).

Pour les Landrace, un signal est retenu sur :

- **SSC7** (MARC0037028)

Antérieurement à ce projet, plusieurs études internationales avaient été réalisées afin de rechercher des régions QTL déterminant la prédisposition des animaux à développer une hernie inguino-scrotale (Ding et al., 2009; Du et al., 2009; Grindflek et al., 2006; Knorr et al., 2006). Ces études avaient été réalisées à l'échelle du génome, à l'aide de microsatellites (moins de 200 marqueurs pour l'ensemble des chromosomes) et avaient permis de mettre en évidence certaines régions chromosomiques. Trois de ces régions publiées sont communes aux régions candidates que nous avons identifiées.

La région localisée en extrémité du bras court du chromosome 2 avait également été identifiée par Du et al. (Du et al., 2004) (Figure 84) entre les positions 0,95 et 4,74 Mb. Huit gènes candidats (FADD, SPTBN2, ACTN3, T175C13, CTSF, SP426E24, EFEMP2, CFL1) avaient été sélectionnés dans cette région pour la recherche de mutation causale, mais aucun variant n'avait été identifié comme associé à la maladie. Dans la région identifiée avec la puce, sept gènes sont annotés sur le clone contenant les deux marqueurs significatifs. Le gène CTSF, précédemment analysé par Du et al, fait partie de ces sept gènes. La résolution obtenue à l'aide de la puce étant supérieure à la résolution obtenue via l'utilisation de microsatellites, il pourrait être intéressant si ce résultat était confirmé, d'étudier de manière plus exhaustive les six autres gènes du BAC et les gènes à proximité de ce BAC.

Le second signal sur SSC2, mis en évidence uniquement à partir des familles LW (en position 138.024 Mb) sous une hypothèse d'expression maternelle, est proche du marqueur significatif identifié par Ding ( $p < 0,05$ ) : SWR2157, en position 133,376 Mb (accession AF253920). Cette région est également significative d'après les travaux de Grindfleck.



Dans ces deux études, aucun modèle d'empreinte n'avait été testé. A proximité du SNP le plus significatif, en position 137,455 Mb, le gène ADAMTS19, (*ADAM metallopeptidase with thrombospondin type 1 motif, 19*) pourrait être un bon candidat pour la hernie inguino-scrotale. La famille d'enzymes ADAM est connue pour être impliquée dans la dégradation de la matrice extracellulaire, et notamment du collagène. De plus, le gène codant pour ADAMTS16, dont la séquence est très similaire à celle de ADAMTS19, est un gène soumis à empreinte parentale chez l'homme.

Une région du chromosome 13 avait également été mise en évidence par Grindfleck et Ding. Cette région co-localise avec la position du signal obtenu avec le marqueur MARC0013088. Dans cette région, un seul gène est annoté dans un intervalle de 500 kb : HES1 (*Hairy and enhancer of split 1*). Il est connu pour être impliqué dans la prolifération cellulaire et la différenciation au cours de l'embryogénèse. HES1 est soumis à empreinte parentale chez l'homme. Les analyses TDT selon l'origine parentale ont montré que le signal provenait d'une distorsion de ségrégation paternelle. Mais nous ne pouvons pas exclure que ce résultat résulte dans notre dispositif d'un manque d'informativité des mères des familles analysées.

Tous ces résultats confirment la complexité du déterminisme de la hernie inguino-scrotale, ainsi que son caractère polygénique. Certains mécanismes pourraient être race-spécifiques, tandis que d'autres déterminés par des mécanismes d'empreinte parentale. Quelques résultats intéressants (données corroborant des données publiées, présence de gènes candidats...) ont été néanmoins obtenus. Il est cependant clair que le dispositif est encore trop petit et il est donc essentiel de poursuivre la collecte d'échantillon pour cette anomalie. Une étape de confirmation ou d'invalidation des résultats obtenus est essentielle avant d'entreprendre une étude de gènes candidats dans les régions identifiées.

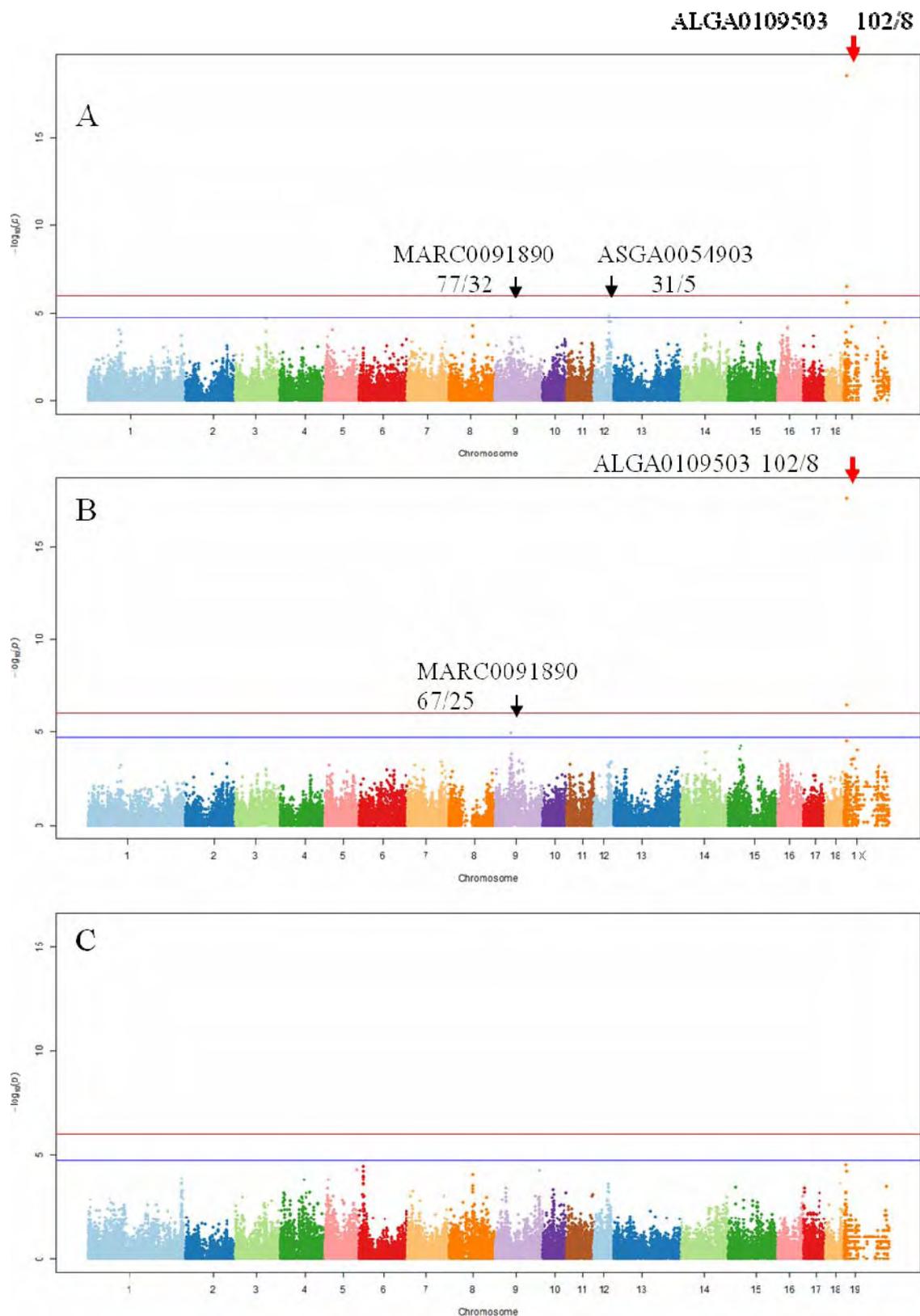


Figure 85 : Résultats TDT en tout génome pour la cryptorchidie – familles Piétrain+Large White (A), familles Piétrain (B) et familles Large White (C)

### **II.B.3. La Cryptorchidie**

Les résultats TDT pour des familles dont les descendants sont atteints de cryptorchidie sont présentés en Figure 85.

#### ***Résultats TDT obtenues à partir de l'ensemble des familles***

Les résultats TDT pour les 53 familles Piétrain et Large White (73 méioses informatives en moyenne) sont présentés en Figure 85(A). Une région de trois marqueurs adjacents est identifiée sur le chromosome X, entre les positions 6,72 et 6,79 Mb (marqueurs ALGA0109503, ASGA0103701 et MARC0045707). Le signal maximum est observé pour ALGA0109503, en position 6.722 Mb ( $p = 3,17.10^{-19}$ ).

Deux marqueurs isolés sont également significatifs (MARC0091890 sur SSC9 à 51,24 Mb, ASGA0054903 sur SSC12 à 50,66 Mb).

Le signal sur le chromosome X est similaire à celui précédemment observé. Des corrélations entre la hernie inguino-scrotale et la cryptorchidie ont été établies dans la littérature (Zhao, 2009), il est donc possible que des mécanismes moléculaires soient communs aux deux anomalies.

#### ***Résultats TDT issus de l'analyse des familles Piétrain seules***

Les résultats TDT pour les 27 familles Piétrain (51 méioses informatives) sont présentés dans la Figure 85(B). Les signaux significatifs mis en évidence ici ont été précédemment identifiés lorsque les familles Piétrain et Large White ont été poolées. Ils correspondent à la région sur le chromosome X (ALGA0109503 et ASGA0103701), et au marqueur MARC0091890 sur le chromosome 9.

#### ***Résultats TDT obtenus à partir des familles Large White***

Les résultats TDT pour les 15 familles Large White (22 méioses informatives en moyenne) sont présentés dans la Figure 85(C). Aucun résultat significatif n'est observé. Le marqueur dont le signal est le plus fort est localisé sur SSCX (ASGA0098910, en position 6,119 Mb) mais le signal n'est pas significatif ( $p = 3,22.10^{-5}$ ).

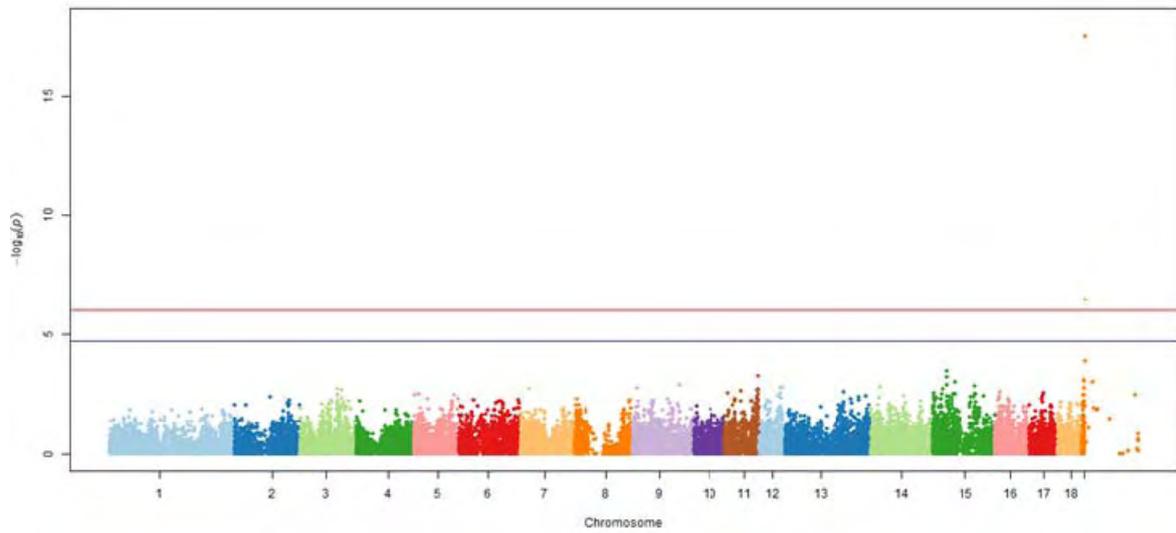


Figure 86 : Résultats TDT en tout génome pour la cryptorchidie selon le brin paternel chez les Piétrain

En testant l'hypothèse d'un gène candidat soumis à empreinte parentale, seul deux marqueurs présentent un signal significatif d'expression paternelle en race Piétrain (Figure 86). Ces deux marqueurs sont localisés sur le chromosome X, il s'agit de ALGA0109503 ( $p = 2,818.10^{-18}$ ) et ASGA0103701 ( $p = 3,414.10^{-7}$ ).

En bilan, les signaux retenus pour toutes les familles (Large White + Piétrain) sont :

- **SSC9** (MARC0091890),
- **SSC12** (ASGA0054903),
- **SSCX** (signal maximum : ALGA0109503) d'origine paternelle.

Aucune étude rapportant la primo localisation de QTL déterminant cette anomalie n'a été jusqu'à présent publiée chez le porc. Les seuls travaux publiés portent sur une estimation de la concentration calcique qui est augmentée chez les porcelets cryptorchides, en comparaison avec les animaux sains. Dans l'état actuel de la collection, le nombre de méioses informatives est trop faible pour obtenir plus de résultats. Pour les Large White, seules 15 familles sont disponibles, ce qui ne permet pas la mise en évidence de régions candidates. Bien que deux signaux soient significatifs pour l'ensemble des familles, ces résultats ne permettent pas d'envisager une recherche de gènes candidats. Au vu de ces résultats, il est donc impératif de poursuivre la collection pour cette anomalie.



### III.C. Discussion

Les anomalies congénitales sont un problème majeur pour l'industrie porcine. Bien qu'il soit établi qu'une part du déterminisme des anomalies est génétique, les études réalisées jusqu'à présent n'ont pas permis d'identifier les mécanismes moléculaires sous-jacents.

Pour l'intersexualité et la cryptorchidie, aucune recherche de QTL à l'échelle du génome n'avait été réalisée chez le porc. Seules les expressions de quelques gènes associés au développement sexuel pour l'intersexualité et les concentrations calciques pour la cryptorchidie ont fait l'objet d'études chez le porc. Ces analyses n'ont pas permis l'identification de mécanismes moléculaires responsables des anomalies. Les études chez l'homme et chez la souris montrent que de nombreux gènes pourraient être responsables de ces maladies, mais la liste des gènes candidats n'est pas exhaustive.

Pour la hernie inguino-scrotale, plusieurs travaux portent sur la recherche de QTL à l'échelle du génome. Ces études ont été réalisées avec une faible densité de marqueurs (moins de deux cents marqueurs microsatellites au maximum pour couvrir l'ensemble des autosomes porcins). Plusieurs loci ont été mis en évidence, mais les résultats varient d'une étude à l'autre, et selon les races étudiées.

Le développement de la puce porcine 60KSNP offre de nouvelles possibilités pour les études d'association en tout génome chez le porc, et permet une recherche haute densité de régions associées à un caractère. De plus, la mise en place d'un dispositif, en collaboration avec les organismes de sélection porcine, a permis de récolter plus de 1500 échantillons d'animaux atteints d'anomalies de naissance, ainsi que les échantillons de leurs deux parents. Ces données ont permis d'envisager des études d'association familiales, en utilisant le test statistique TDT, qui teste à la fois la liaison et l'association d'un marqueur à un caractère. L'avantage de ce test est qu'il est particulièrement robuste, dû au fait que les problèmes de stratification de population, rencontrés dans les études classiques d'associations, sont corrigés par l'étude de la liaison.

Pour la première fois, des analyses TDT ont été réalisées pour les trois anomalies les plus fréquentes dans les élevages (hernie inguino-scrotale, cryptorchidie et intersexualité) en tout génome et à l'aide de la puce porcine haute densité. Ces analyses ont permis d'identifier quelques régions candidates associées aux défauts de naissance.

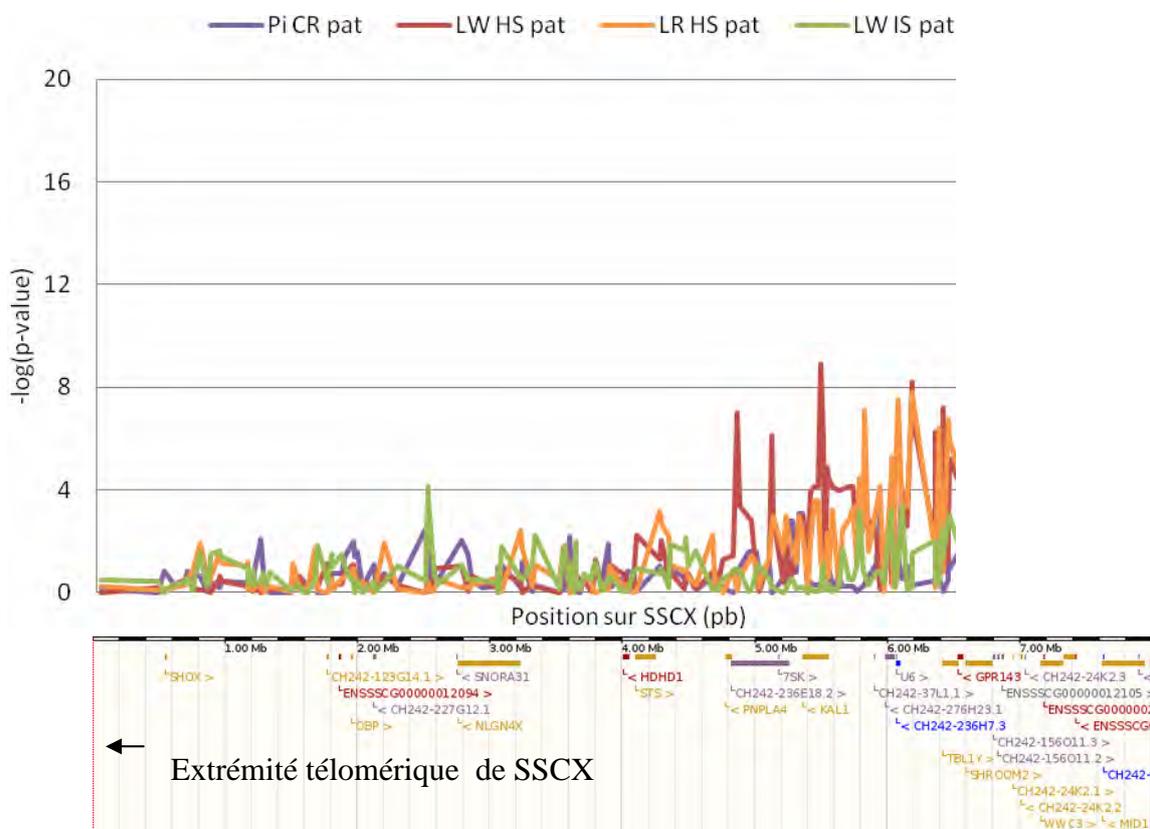


Figure 87 : Résultats TDT significatifs sur l'extrémité du chromosome X pour les différentes populations, et les gènes annotés dans la région

Au vu de la significativité des marqueurs adjacents et des gènes localisés au niveau des signaux significatifs, le signal le plus prometteur a été obtenu pour l'intersexualité sur le chromosome 12.

Un seul gène, *SOX9*, est annoté dans cette région, connu pour être un acteur majeur de la différenciation sexuelle. De plus, des variations dans la séquence encadrant ce gène ont été identifiées chez la souris et chez l'homme pour être associées à des défauts de différenciation sexuelle. *SOX9* est donc un excellent candidat, et la suite des travaux de la thèse a été orientée sur la cartographie fine de la région et sur la recherche de variants causaux. Le **chapitre II** traite de ces travaux de cartographie.

En considérant l'ensemble des résultats, un signal très significatif est également observé sur le chromosome X, et est associé aux trois maladies étudiées. Les analyses TDT, sous hypothèse d'empreinte parentale, laissent supposer que seule la distorsion de ségrégation paternelle est responsable du signal observé. En effet, si l'on compare dans cette région tous les signaux observés pour les différents groupes étudiés (Figure 87), la similarité de la zone candidate est confirmée. Seuls les Large White cryptorchides ne présentent pas de résultats significatifs. Cependant, les analyses n'ayant été réalisées qu'avec 15 familles, le nombre de méioses informatives est relativement faible.

Bien que des corrélations entre la hernie scrotale et la cryptorchidie aient été établies dans la littérature, il est très surprenant que le même locus puisse être impliqué dans trois pathologies différentes.

Chez les autres mammifères, l'extrémité du bras court du chromosome X correspond à la région pseudoautosomale (PAR), région dont la séquence est homologue entre le chromosome X et le chromosome Y (Helena Mangs and Morris, 2007). Un appariement est nécessaire entre les régions PAR des chromosomes X et Y (PARx et la PARy) lors de la méiose pour assurer la gamétogénèse (Kauppi et al., 2011).

Chez le porc, très peu de données sont disponibles concernant la PAR (Quilter et al., 2002; Skinner et al., 2013), et le chromosome X en général. Une des priorités des travaux de thèse a donc été de réaliser une analyse approfondie du chromosome X, et de la PAR porcine en particulier, afin de comprendre notamment les résultats obtenus par analyses TDT dans cette région. Ces travaux correspondent au **chapitre III**.



D'autres signaux significatifs ont été mis en évidence à l'issue des analyses TDT. Ils concernent, dans la plupart des cas, un marqueur isolé (hernie : 7 marqueurs isolés et 2 marqueurs adjacents ; cryptorchidie : 2 marqueurs isolés). Chaque SNP étant traité indépendamment des 60.000 autres marqueurs, plusieurs marqueurs adjacents présentant des résultats significatifs permettent plus certainement d'être considérés comme candidats d'une région associée à la maladie, que des régions où seul un marqueur isolé est significatif. Nous avons donc considéré que ces résultats ne permettaient pas d'envisager une recherche de gènes candidats, bien que certains d'entre-eux puissent être prometteurs.

Malgré l'utilisation d'une puce de génotypage haute densité et d'un dispositif de près de 150 familles de descendants atteints (300 cryptorchides Large White et Piétrain, et 550 animaux Large White et Landrace atteints de hernies), cette étude n'a pas permis d'identifier de régions associées et très significatives pour la hernie scrotale et la cryptorchidie. Plusieurs raisons pourraient expliquer qu'aucun intervalle n'ait été identifié : (1) plusieurs variants à effet faible sont responsables des anomalies, mais la taille du dispositif et/ou la méthode ne permettent pas de les identifier, (2) les variants causaux sont rares, dans ce cas l'outil utilisé (la puce porcine) n'est pas adapté, (3) des variants structuraux sont responsables des anomalies et l'étude des SNP ne permet pas facilement leur identification, et (4) au sein d'une même pathologie, plusieurs maladies différentes et donc plusieurs causes sont en fait rassemblées. Chacun des points présentés ici sera abordé et des perspectives seront proposées.

#### *Plusieurs variants à effet faible :*

D'après les études portant sur le mode de ségrégation de la hernie scrotale et de la cryptorchidie, l'hypothèse la plus probable est que ces maladies sont multifactorielles (Ivell, 2003; Matthews and Neumayer, 2008). Cette hypothèse pourrait expliquer que plusieurs régions peu significatives aient été identifiées. Afin de réduire les risques de causes environnementales, des critères comme la sélection de familles produites dans différents élevages mais issues d'un même verrat, ou l'absence d'animaux castrés dans le dispositif ont été pris en compte pour cette étude. Les chances sont donc faibles pour que l'absence de signal significatif soit expliquée par les seules causes environnementales responsables de la maladie. Cependant, un nombre d'échantillons trop faible et/ou un test peu puissant peuvent expliquer l'absence de résultats.



### *Taille du dispositif*

La taille optimale du dispositif pour l'identification de variants génétiques est difficile à estimer car elle dépend de nombreux paramètres, tels que la MAF des SNP sélectionnés, le nombre de variants causaux et leur effet respectif, le nombre de phénotypes réellement étudiés et la structure de la population (Marian, 2012).

Toutes les études s'accordent sur le fait que le nombre d'échantillons doit être très important. Chez l'homme de nombreuses études GWA ont été réalisées en utilisant des échantillonnages de 2.000 à 10.000 individus permettant d'avoir une puissance statistique suffisante pour identifier des variants communs avec confiance (avec des valeurs de *odds ratio* de 1,5 ou plus) (Newton-cheh et al., 2009; Sotoodehnia et al., 2010; Teslovich et al., 2011). Il est cependant probable que de nombreux loci restent à découvrir et Frazer et al. rapportent qu'une taille d'échantillon de 60.000 individus serait nécessaire pour avoir une puissance suffisante permettant d'identifier la majorité des variants avec des *odds ratios* de 1,1 (Frazer et al., 2009). Afin d'augmenter la taille des dispositifs, des méta-analyses ont été réalisées par le *Wellcome Trust Case Control Consortium* en regroupant les cohortes de 2.000 patients (une cohorte par maladie) issus d'études indépendantes pour les sept maladies humaines majeures et un même groupe de 3.000 individus contrôles (The Wellcome Trust Case Control Consortium, 2007). Cette étude a permis de mettre en évidence 24 signaux d'association très significatifs ( $p < 5 \cdot 10^{-7}$ ) et 58 signaux additionnels ( $10^{-5} < p < 5 \cdot 10^{-7}$ ) permettant ainsi de mettre en évidence des facteurs de risque génétiques associés à la plupart des maladies ciblées (un pour les troubles bipolaires, un pour la maladie coronarienne, neuf pour la maladie de Crohn, trois pour la polyarthrite rhumatoïde, sept pour le diabète de type I et trois pour le diabète de type II).

Chez le porc, plusieurs GWAS ont été réalisées depuis 2010 avec la puce porcine. L'essentiel de ces études porte sur des caractères quantitatifs (production d'hormones sexuelles (Duijvesteijn et al., 2010; Grindflek et al., 2011b; Ramos et al., 2011), caractères de production (Fan et al., 2011; Ponsuksili et al., 2011) et de reproduction (Grindflek et al., 2011a; Onteru et al., 2011)). Le nombre d'échantillons utilisés varie de 800 à 2.500 animaux par GWAS. Ces études ont permis de confirmer l'implication de régions précédemment identifiées par analyse de liaison à l'aide d'une faible densité de marqueurs, et de mettre en évidence de nouvelles régions candidates.



A titre d'exemple, Fan et al. ont identifié dix-huit régions associées aux caractères de carcasse, dont neuf avaient déjà été mises en évidence (régions contenant MC4R (associé au lard dorsal) et IGF2 (associé au muscle de la longe)). D'autres gènes candidats ont également été mis en évidence (CHCHD3, BMP2 et la famille de HOXA). Jusqu'à présent, seules deux études portent sur des phénotypes qualitatifs. L'étude de Sironen et al. a permis d'identifier une région (deux gènes candidats : STK17b et HECW2) associée à un défaut de structure des acrosomes, responsable d'infertilité chez le porc (Sironen et al., 2010b). Seuls quatorze atteints et vingt et un témoins ont été suffisants dans cette étude. Ren et al. ont identifié un gène, TYRP1, responsable de la coloration brune des porcs chinois en utilisant 121 porcs bruns et 745 porcs non bruns comme contrôle.

Le phénotype binaire de ces deux études correspond au type de caractère de notre analyse (atteint = 1 versus non atteint = 0), mais les deux maladies ayant fait l'objet de ces études sont monogéniques, alors que l'hypothèse la plus probable est que la hernie inguino-scrotale et la cryptorchidie sont polygéniques. La taille du dispositif nécessaire correspondrait donc aux effectifs utilisés pour les GWAS des caractères quantitatifs (800 à 2.500). Sachant que notre dispositif, collectés en quatre ans, est actuellement composé de 300 cryptorchides et 550 animaux atteints de hernie inguino-scrotale, il est impératif de poursuivre la collecte d'échantillons quelques années encore afin d'augmenter significativement la taille du dispositif et augmenter la puissance des analyses.

Comme alternative, des méta-analyses permettraient d'augmenter la taille du dispositif plus rapidement. Le principe des méta-analyses est de combiner différents dispositifs d'études portant sur le même caractère. Ces analyses ont déjà fait leurs preuves, notamment chez l'homme (Heid et al., 2011; Zeggini et al., 2009) chez le bovin (Minozzi et al., 2012), et chez le cochon (Silva et al., 2011). Dans notre étude, cette méthode a été réalisée à moindre échelle pour l'intersexualité, en intégrant des familles issues d'un autre dispositif (Pailhoux et al., 1997) aux familles collectées par les OSP. Le signal observé sur le chromosome 12 avec les familles des OSP uniquement a ainsi été confirmé lors de l'ajout des familles OSP. Pour la hernie scrotale, les races sont différentes d'une étude à l'autre, mais trois signaux observés dans nos travaux ont également été mis en évidence dans les autres études (Ding et al., 2009; Du et al., 2004; Grindflek et al., 2006). Cela signifie que certains mécanismes moléculaires associés à la maladie pourraient être communs aux différentes races.



Nous avons ainsi envisagé de mettre en place une collaboration avec une équipe de recherche belge, rencontrée lors de l'ISAG (*International Society for Animal Genetics*). Leur dispositif est composé de 188 trios dont les descendants sont atteints de hernie inguino-scrotale, et de 300 trios dont les descendants sont atteints de cryptorchidie (Stinckens et al., 2012), génotypés à l'aide de la puce 60K porcine.

Une collaboration internationale pourrait également être envisagée avec les équipes travaillant sur le déterminisme génétique de la hernie inguino-scrotale, dont les études sont publiées dans la littérature (Ding et al., 2009; Du et al., 2004; Grindflek et al., 2006; Knorr et al., 2006; Zhao et al., 2009). Il s'agirait de pooler les échantillons de différentes études portant sur le même phénotype. Néanmoins ces dispositifs ont été jusqu'à présent analysés à l'aide de microsatellites et il serait donc nécessaire de génotyper ces animaux à l'aide de la puce 60K porcine.

#### *Test statistique*

Nous avons choisis pour cette étude d'utiliser le test TDT, réputé pour sa robustesse, mais dont la puissance est faible. Au moment du choix du test, le problème majeur des tests d'association était les biais dus à la stratification des populations, et le TDT basé sur des études familiales d'association permettait de s'astreindre de ces biais. Depuis, il est considéré comme acquis que les données de génotypages permettent de disposer d'un contrôle interne : en cas de stratification, le nombre de loci présentant des différences de fréquences alléliques entre cas et contrôles est beaucoup plus important, et réparti de façon plus homogène dans le génome, que lorsque les différences de fréquences (estimés dans un dispositif homogène) résultent uniquement des signaux d'association. En génétique humaine et animale, des modèles tirant parti des données de génotypage, comme par exemple les matrices d'apparentement génomique, permettent désormais de corriger les biais de structure au sein d'un échantillonnage.

Parallèlement, d'autres modèles ont été développés afin de chercher à estimer non plus l'effet potentiel de chaque SNP un par un sur le caractère, mais simultanément pour l'ensemble des marqueurs. Ces modèles permettent de tester des hypothèses de déterminisme oligo-génique (peu de QTL à effets forts) ou polygénique (nombreux QTL à effets faibles) de la maladie. Ces modèles, utilisés pour les études d'association cas-contrôles, pourraient être utilisés dans notre dispositif. Dans notre cas, une population d'animaux contrôles serait nécessaire pour utiliser ce modèle.

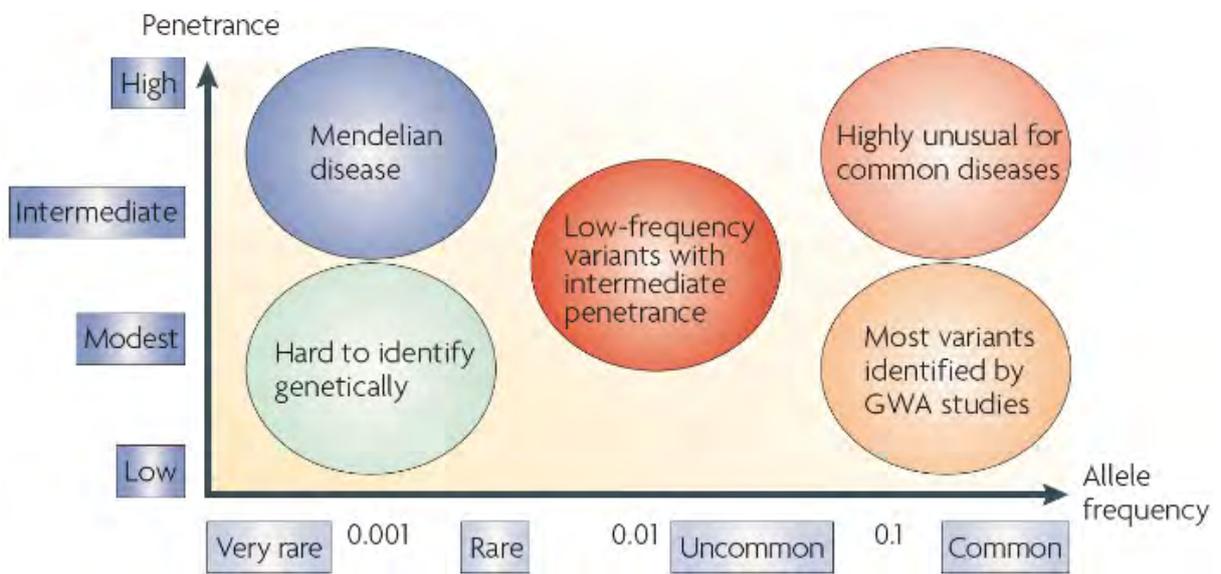


Figure 88 : distribution des variants responsables de maladies (McCarthy et al., 2008).

Au laboratoire, un grand nombre d'animaux de différentes races ont été génotypés dans le cadre d'autres dispositifs, et pourraient donc servir de contrôles pour notre étude.

Les modèles récemment développés comme celui de Zhou et al. (Zhou et al., 2013) n'ont, à notre connaissance, jamais été utilisés dans une étude GWAS réalisée avec la puce porcine. Dans notre cas, le fait qu'une région ait été identifiée pour l'intersexualité à l'aide du TDT, test très robuste, permettra de valider le modèle statistique pour notre dispositif, qui pourra alors être utilisé pour la hernie et la cryptorchidie. En effet, les résultats obtenus avec deux tests différents virent parfois, comme c'est le cas dans les analyses chez le porc réalisées pour la hernie scrotale (Ding et al., 2009; Du et al., 2009; Grindflek et al., 2006; Knorr et al., 2006).

### *Variants rares*

Malgré des tailles de dispositifs importantes, et des méthodes statistiques puissantes, les loci identifiés n'expliquent souvent qu'une part très faible de la variation phénotypique. Selon la pénétrance de la maladie et la fréquence du ou des variants causaux, les maladies peuvent être réparties en cinq catégories (Figure 88). Les maladies mendéliennes (un seul variant est un facteur majeur de prédisposition et la pénétrance de la maladie est forte) firent l'objet des premières études de liaison dès les années 70-80 (Altshuler et al., 2008). Néanmoins pour la plupart des maladies, plusieurs variants dont la fréquence dans la population est importante (fréquence de 0,1 environ) semblent impliqués dans des caractères d'expressivité variable (Bodmer and Bonilla, 2008; Pritchard, 2001; Reich and Lander, 2001). Ces maladies sont désignées par le terme « common disease, common variant » (les maladies communes sont polygéniques, c'est-à-dire dues à plusieurs variants). Le risque qu'un individu soit malade est déterminé par la présence de chaque variant contribuant à un effet faible additif sur le phénotype de la maladie. Les GWAS, réalisées avec une haute densité en marqueurs recouvrant tout le génome (puces de génotypage), et un grand nombre d'échantillons ont permis d'identifier de nombreux variants impliqués dans ces maladies communes, pour lesquelles les analyses de liaison n'avaient pas été concluantes. Cependant, les GWAS n'ont pas été à la hauteur des espérances, car les variants identifiés n'expliquaient souvent qu'une part très faible de la maladie. Il fut par la suite montré que pour certaines maladies, des variants rares (fréquence inférieure à 0,05) pouvaient être impliqués.



Les marqueurs des puces de génotypage haute densité étant sélectionnés sur critères de MAF 0,05 (« common variant »), les allèles rares sont souvent indétectables lorsque les GWAS sont réalisées avec des puces de génotypage (Cirulli and Goldstein, 2010), à l'exception des régions où des variants causaux rares sont en DL avec des SNP communs (Dickson et al., 2010).

Dans les cas des anomalies congénitales porcines, la prévalence est de 1 à 3% toutes anomalies confondues. Les trois anomalies étudiées ici concernent donc chacune 1% au maximum de la population. Pour une fréquence génotypique de l'état "homozygote muté" de 0.01, et dans le cas d'un déterminisme monogénique récessif, la fréquence de l'allèle muté serait donc de 0.10. Les SNP de la puce porcine ont été sélectionnés pour leur MAF supérieure ou égale à 5%. Dans l'hypothèse de variants causaux communs, la puce porcine est adaptée. Dans l'hypothèse où il existe plusieurs causes différentes, dont des variants rares, la puce porcine n'est plus nécessairement adaptée car le DL est peut-être trop faible entre les marqueurs et la mutation causale pour qu'un signal soit détecté.

Afin de pallier ces limites, les puces de génotypage sont peu à peu remplacées par le séquençage NGS (*New Generation Sequencing*) (Koeleman et al., 2013). L'idéal serait de disposer de l'ensemble des polymorphismes, obtenus par séquençage, différenciant les génomes d'individus sains et atteints. Cette approche est pour l'heure très chère et donc limitée, le génotypage par puce restant la stratégie la plus courante (Gilissen et al., 2011a).

Des approches ont donc été développées pour identifier les variants rares présents dans les populations. Chez l'homme, le projet 1000 génomes ([www.1000genomes.org](http://www.1000genomes.org)) a été mis au point pour disposer d'une base de données exhaustive des variants rares dans les populations humaines. Ce projet consiste au séquençage haut débit d'un millier d'individus afin de disposer d'une base de données génériques utilisable pour tous les projets. Ces données de génotypages sont utilisées pour les différentes études d'association, en intégrant par inférence les génotypes probables des individus étudiés à ces variants rares, au vu des génotypes réels obtenus aux marqueurs des puces de génotypage (Huyghe et al., 2013). Chez le bovin, un projet 1000 génomes est également en cours (<http://www.1000bullgenomes.com>). En dehors des études d'association, l'avantage de ce type de projet chez les animaux domestiques est l'exploitation des données pour la sélection génomique sans que les variants génétiques responsables des caractères d'intérêt n'aient été identifiés.



En attendant un projet 1000 génomes chez le porc, une méthode alternative serait nécessaire dans notre étude pour identifier des marqueurs plus informatifs. Deux types d'approches pourraient être utilisées : 1) une analyse tout génome sans *a priori* qui pourrait être envisagée dès à présent, et 2) une étude exhaustive des régions mises en évidence, dans le cas où les méthodes proposées pour augmenter la puissance des analyses auraient permis d'identifier des régions candidates.

Dans le cas d'une analyse tout génome sans *a priori*, mais à coût réduit, deux approches sont envisageables (Gilissen et al., 2011b): Le RADseq (*Restriction site Associated DNA (RAD) sequencing*) et le séquençage d'exome (séquençage de l'ensemble des exons, *Exome Sequencing*). Le RADseq consiste à isoler par digestion enzymatique puis à séquencer une petite portion du génome de façon systématique chez l'ensemble des individus d'un dispositif. Cette stratégie permet de découvrir et simultanément de génotyper des milliers de variants (SNP et indels) pour un grand nombre d'échantillons (jusqu'à 384) (Elshire et al., 2011) sans biais particulier basé sur la fréquence d'un allèle mineur. Bien que non exhaustive, cette approche permet d'analyser des variants rares à l'échelle du génome.

Le séquençage d'exome est une méthode de plus en plus utilisée car elle permet une étude de régions ciblées potentiellement impliquées dans de nombreuses variations phénotypiques. La capture de séquence permet un enrichissement spécifique d'une librairie d'ADN génomique dans des régions ciblées avant séquençage. En effet, l'exome ne représente que 1% du génome (30Mb approximativement) mais 85% des mutations identifiées dans les maladies mendéliennes sont localisées dans ces régions (Choi et al., 2009a). Depuis ces trois dernières années, plus de 100 gènes ont été identifiés dans ces régions (Wang et al., 2013) comme responsables de nombreuses maladies (Marian, 2012). Cette stratégie est cependant plus adaptée à la recherche de mutations à effet important qu'à la recherche de variants rares, car les polymorphismes identifiés sont localisés au sein des gènes. Le ciblage des régions codantes du génome est réalisé à l'aide de puce générique de capture d'exons. Cette approche est donc facile à mettre en œuvre lorsqu'une puce d'exome est disponible et commercialisée pour une espèce étudiée.

Dans l'hypothèse de la mise en évidence de régions candidates suite aux méta-analyses et/ou à l'utilisation d'une autre méthode statistique, il est possible de procéder au séquençage de ces régions par capture, technologie couramment utilisée chez l'homme (Choi et al., 2009b).



Contrairement aux puces d'exomes, il est nécessaire pour cette approche de constituer spécifiquement une puce correspondant aux régions candidates. Réduire la taille du génome séquencé permet d'augmenter la profondeur en séquences, et ainsi d'obtenir une profondeur optimale pour déterminer le génotype des variants séquencés.

### *Recherche de variants structuraux*

D'autres variants de type CNV (*Copy Number Variation*) pourraient être impliqués dans la maladie (Hirasawa and Feil, 2010; Perry et al., 2007). Cependant, les méthodes citées ici ne permettent pas toujours d'identifier ces variants dits structuraux. Des méthodes spécifiques ont été développées les concernant et pourraient être utilisées dans le cadre de notre étude.

### *CNV*

Les CNV ont été rapportés comme étant associés à de nombreuses maladies communes chez l'homme telles que la susceptibilité au HIV, le lupus érythémateux systémique, la polyartérite microscopique, la maladie de Crohn,...(McCarroll et al., 2008a). Ils sont également reconnus comme pouvant être associés à des maladies rares (Stankiewicz and Lupski, 2010).

Bien que les GWAS puissent permettre d'identifier des CNV, soit par la mise en évidence de SNP en DL avec un CNV (McCarroll et al., 2008b), soit par l'étude de l'intensité de fluorescence enregistrée pour chaque SNP génotypé (Wang et al., 2007), des méthodes d'investigation spécifiques existent. Les deux méthodes utilisées pour la détection de CNV sont l'hybridation génomique comparative (CGH) et le séquençage. La détection de CNV par la méthode CGH permet de mettre en évidence des duplications ou des délétions de petite taille. Mais elle nécessite, comme pour l'approche précédemment décrite, de disposer ou de développer une puce d'hybridation représentative de la région étudiée. L'approche alternative est d'utiliser des approches de séquençage. Deux types de données sont utilisées pour la mise en évidence de CNV : les données de séquençage *paired-ends* (identification de régions présentant une profondeur de lecture plus importante que dans le reste du génome) et les données de séquençage *mate-pairs* (distance entre deux lectures plus importante qu'attendu). Par rapport aux approches CGH, les données de séquençage permettent en outre d'identifier les remaniements de structure (sans modification du nombre de copie). Dans le cadre d'un programme d'étude de la variabilité de structure du génome chez des porcs de races différentes, des séquençages *mate-pairs* sont en cours. Il est prévu qu'un des pools d'individus Large White étudié dans ce projet soit constitué d'animaux intersexués. Selon les résultats obtenus, nous envisagerons d'étendre cette approche aux autres maladies.



### *Empreinte parentale*

La majorité des mécanismes moléculaires identifiés comme responsables de maladies est de type récessif ou dominant. Cependant, certains mécanismes sont plus complexes, comme par exemple les gènes soumis à empreinte parentale, pour lesquels l'expression d'un gène dépend de son origine parentale. Ces gènes pourraient expliquer certains cas de variations phénotypiques complexes. A titre d'exemple de complexité chez l'homme, un allèle associé au diabète de type 2 est un avantage lorsqu'il est transmis par la mère, alors qu'il est défavorable lorsqu'il est transmis par le père (Kong et al., 2013). Des phénomènes de dominance polaire sont également observés (un seul type d'hétérozygote est différent des homozygotes, comme c'est le cas pour le phénotype Callipyge chez le mouton (Davis et al., 2004)). Chez le porc, le gène soumis à empreinte le plus étudié est IGF2, associé à des caractères de carcasse (Van Laere et al., 2003).

Chez l'homme, plus de deux cent gènes soumis à empreinte ont été identifiés et sont impliqués pour la plupart d'entre eux dans la croissance fœtale et le développement. Ils pourraient donc être de bons candidats à l'intersexualité et/ou la cryptorchidie. Cependant, la taille du dispositif nécessaire est très importante, puisque seule la moitié des parents est considérée. Chez le porc, trente gènes ont été identifiés comme soumis à empreinte (<http://www.geneimprint.com/>) mais la liste n'est pas exhaustive. En effet, les homologies sont fortes entre l'homme et le porc, il est donc probable que les gènes soumis à empreinte chez l'homme le soient également chez le porc.

Dans le cadre de cette étude, une hypothèse d'empreinte a été testée par analyse TDT et des résultats prometteurs ont été obtenus. L'étude de gènes soumis à empreinte parentale n'avait jamais été envisagée auparavant, et pourrait expliquer qu'aucun mécanisme n'ait été identifié. Des analyses TDT prenant en compte l'origine parentale sont donc à considérer pour la recherche de mécanismes sous-jacents à certaines maladies. Mais ces études nécessitent de disposer des génotypes des parents (afin d'identifier chez chaque descendant l'origine parentale des allèles reçus). De plus, seules les méioses paternelles ou maternelles sont considérées, ce qui réduit encore le nombre de méioses informatives, et nécessite une taille de dispositif très importante.



Au delà de la variabilité des mécanismes moléculaires pouvant être impliqués dans la maladie, l'absence de résultats significatifs dans notre étude pourrait être expliquée par une variabilité phénotypique des caractères, c'est-à-dire que plusieurs types de hernie ou de cryptorchidie, dont les causes diffèrent, pourraient être regroupés dans les mêmes phénotypes.

### *Plusieurs maladies décrites sous le même caractère*

Avec les avancées des technologies de génotypage, une des principales étapes limitant l'identification des variants responsables des maladies est le phénotypage. En effet, certaines études ont montré que le diagnostic des maladies pouvait biaiser les recherches du déterminisme moléculaire (Alzheimer, Parkinson, cancer du sein, syndrome de Gilles de la Tourette) (Burd et al., 2001).

Le phénotypage systématique (clinique, biochimique, biologique) de tous les individus (Bennett et al., 2012) est aujourd'hui envisagé dans les études génétiques. Chez le porc, les quelques GWAS réalisées ont été très concluantes. Malgré l'étude de caractères polygéniques, les phénotypes ont été finement caractérisés (Duijvesteijn et al., 2010; Fan et al., 2011; Grindflek et al., 2011a, 2011b; Onteru et al., 2011; Ponsuksili et al., 2011; Ramos et al., 2011). Pour chacun des caractères, de nombreux paramètres ont été mesurés ou dosés sur chacun des animaux.

Dans le cas des trois anomalies étudiées, l'autopsie d'animaux atteints permet de réaliser un phénotypage fin, tel que l'étude d'expression de gènes dans des tissus et organes potentiellement impliqués dans la maladie, ou une caractérisation morphologique fine. Ce type de travaux a été réalisé pour certains cas d'intersexués chez le porc (Pailhoux et al., 2001a). Les fœtus issus de croisements d'animaux dont les descendants étaient intersexués ont été autopsiés et les tissus gonadiques ont été utilisés pour le dosage d'ARN de certains gènes candidats. Mais ce type de caractérisation phénotypique très fin ne peut être envisagé que sur un petit nombre d'individus. Les analyses GWAS pour des caractères complexes nécessitent inévitablement de disposer d'une mesure phénotypique fiable, facile à réaliser et peu coûteuse. Dans le cas des anomalies congénitales porcines, il n'existe malheureusement aucun marqueur biologique sanguin, ou tissu dont le prélèvement n'est pas invasif, qui n'ait été identifié comme caractéristique de ces maladies. Le seul phénotypage non invasif est celui réalisé dans notre étude.



Pour la première fois, des GWAS avec la puce porcine ont été réalisées pour rechercher des variants associés à la hernie inguino-scrotale, la cryptorchidie et l'intersexualité. Bien qu'elles n'aient pas permis d'identifier de gènes candidats, les résultats sont très prometteurs. D'une part, ils ont permis de confirmer et réduire l'intervalle de régions candidates potentielles identifiées dans des analyses antérieures. En effet, trois régions (deux sur le chromosome 2 et une sur le chromosome 13) associées à la hernie scrotale avaient également été mises en évidence dans une autre étude (Ding et al., 2009; Du et al., 2004; Grindflek et al., 2006). Dans nos travaux ces régions ont été réduites à un voire deux marqueurs SNP. Une hypothèse d'empreinte parentale a également été testée pour une des trois régions. D'autre part, de nouvelles régions potentielles ont été identifiées (cinq régions pour la hernie et une pour la cryptorchidie). Pour ces différentes régions, le manque de puissance n'a pas permis d'envisager des études de cartographie fine.

En conclusion, la piste prioritaire de cette étude, afin d'améliorer ces travaux, est l'augmentation de la taille du dispositif. La mise en place de la banque ANOPORC permet à la collection de s'enrichir progressivement d'échantillons supplémentaires. Dans quelques mois, une série complémentaire d'animaux atteints pourra donc être génotypée. L'apport de ces données et l'utilisation de méthodes d'analyses autres que le TDT devraient nous permettre de gagner en puissance et de confirmer ou d'infirmer les associations trouvées dans le cadre de cette thèse.



## Chapitre II : Cartographie fine de la région candidate à l'intersexualité chez le porc

---

Les travaux de GWAS, présentés dans le chapitre I, ont permis de mettre en évidence un signal d'association significatif entre l'intersexualité et des marqueurs du chromosome 12.

La région de 2,49 Mb est localisée entre les positions 7,10 (MARC0046612) et 9,59 Mb (MARC0041079). Cette région comprend quinze marqueurs significatifs ( $p < 19 \times 10^{-5}$ ) Le marqueur le plus significatif ( $p=5,65 \times 10^{-10}$ ) est en position 9,05 Mb (ALGA0064738).

Un seul gène est annoté dans la région, Il s'agit de *SOX9* (*Sex-Determining Region Y-Box 9*) (Sscrofa10.2:12:9028879:9033246:1), un des acteurs majeurs de la différenciation sexuelle chez le mâle. Des défauts de régulation de *SOX9* chez la souris (Bishop et al., 2000a; Lavery et al., 2011) et l'homme (Benko et al., 2009; Cox et al., 2011; Vetro et al., 2011) sont impliqués dans des désordres du développement sexuel. Ce gène est donc un excellent candidat à l'intersexualité chez le porc.

Au vu de ces résultats, une première étape de cartographie fine de la région a été entreprise en 2012 dans le but de réduire l'intervalle de localisation, voire d'identifier la mutation causale.

Bien que dans l'état actuel des données disponibles, tout l'intervalle soit candidat, une première étude des éléments connus de la région a été réalisée. Ces éléments correspondent au gène *SOX9*, et son enhancer testicule-spécifique, appelé TESCO (*testis-specific enhancer of SOX9 core*). TESCO n'est pas annoté sur la séquence de référence du porc, mais est caractérisé chez d'autres espèces de mammifères. L'alignement de la séquence humaine de TESCO sur le génome porcine a permis d'obtenir la séquence porcine (Sscrofa10.2:12:9045199:9047423:1).

L'analyse TDT ayant été réalisée marqueur par marqueur, une analyse multi-point par analyse haplotypique peut permettre l'identification d'un haplotype candidat et ainsi réduire la zone de recherche. Ce type d'analyse a été réalisé à partir des données de génotypage des marqueurs de la puce.

L'ensemble de ces travaux a fait l'objet d'un article accepté dans PlosOne en Septembre 2013.



**I. ARTICLE 1 : A genome-wide association study points out the causal implication of *SOX9* in the sex-reversal phenotype in XX pigs**

**PlosOne**

Sarah Rousseau<sup>1</sup>, Nathalie Iannuccelli<sup>1</sup>, Marie-José Mercat<sup>2</sup>, Claire Naylies<sup>1</sup>, Jean-Claude Thouly<sup>3</sup>, Bertrand Servin<sup>1</sup>, Denis Milan<sup>1</sup>, Eric Pailhoux<sup>4</sup>, Juliette Riquet<sup>1</sup>

**1.** INRA, UMR444 LGC F-31326 Castanet-Tolosan, France, **2.** IFIP-Pôle génétique, la Motte au Vicomte, F-35651 Le Rheu, France, **3.** INRA, UE0332 Domaine de Galles, F-18520 Avord, France, **4.** INRA, UMR1198 BDR, F-78352 Jouy-en-Josas, France.

# A Genome-Wide Association Study Points out the Causal Implication of *SOX9* in the Sex-Reversal Phenotype in XX Pigs

Sarah Rousseau<sup>1</sup>, Nathalie Iannuccelli<sup>1</sup>, Marie-José Mercat<sup>2</sup>, Claire Naylies<sup>1</sup>, Jean-Claude Thouly<sup>3</sup>, Bertrand Servin<sup>1</sup>, Denis Milan<sup>1</sup>, Eric Pailhoux<sup>4</sup>, Juliette Riquet<sup>\*</sup>

1 INRA, UMR444 LGC, Castanet-Tolosan, France, 2 IFIP, Pôle génétique, la Motte au Vicomte, Le Rheu, France, 3 INRA, UE0332 Domaine de Galles, 4 INRA, UMR1198 BDR, Jouy-en-Josas, France

## Abstract

Among farm animals, pigs are known to show XX sex-reversal. In such cases the individuals are genetically female but exhibit a hermaphroditism, or a male phenotype. While the frequency of this congenital disease is quite low (less than 1%), the economic losses are significant for pig breeders. These losses result from sterility, urogenital infections and the carcasses being downgraded because of the risk of boar taint. It has been clearly demonstrated that the *SRY* gene is not involved in most cases of sex-reversal in pigs, and that autosomal recessive mutations remain to be discovered. A whole-genome scan analysis was performed in the French Large-White population to identify candidate genes: 38 families comprising the two non-affected parents and 1 to 11 sex-reversed full-sib piglets were genotyped with the PorcineSNP60 BeadChip. A Transmission Disequilibrium Test revealed a highly significant candidate region on SSC12 (most significant  $p$ -value  $< 4.65 \cdot 10^{-10}$ ) containing the *SOX9* gene. *SOX9*, one of the master genes involved in testis differentiation, was sequenced together with one of its main regulatory region Tesco. However, no causal mutations could be identified in either of the two sequenced regions. Further haplotype analyses did not identify a shared homozygous segment between the affected pigs, suggesting either a lack of power due to the SNP properties of the chip, or a second causative locus. Together with information from humans and mice, this study in pigs adds to the field of knowledge, which will lead to characterization of novel molecular mechanisms regulating sexual differentiation and dysregulation in cases of sex reversal.

**Citation:** Rousseau S, Iannuccelli N, Mercat M-J, Naylies C, Thouly J-C, et al. (2013) A Genome-Wide Association Study Points out the Causal Implication of *SOX9* in the Sex-Reversal Phenotype in XX Pigs. PLoS ONE 8(11): e79882. doi:10.1371/journal.pone.0079882

**Editor:** Stephen Moore, University of Queensland, Australia

**Received:** June 3, 2013; **Accepted:** September 26, 2013; **Published:** November 6, 2013

**Copyright:** © 2013 Rousseau et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The project was supported by the French National Research Agency (Agence Nationale de la Recherche, ANR-08-GENM-040 SwAn project) and by BIOPORC. SR PhD is funded by LABOGENA and the Research Ministry (Cliffre grant). The ANR and LABOGENA had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. BIOPORC has contributed to the construction of the collection used for the design; BIOPORC had no role in the analysis and the preparation of the manuscript.

**Competing interests:** This study was partly funded by BIOPORC and LABOGENA. There are no patents, products in development or marketed products to declare. This does not alter our adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

\* E-mail: juliette.riquet@toulouse.inra.fr

## Introduction

In mammals, sex differentiation is a genetically and hormonally controlled process. Sex differentiation first relies on the establishment of genetic sex at fertilization by the combination of X or Y chromosome containing gametes to produce XX or XY zygotes. This initial step is then followed by the activation of the appropriate male or female genetic cascades involved in gonad differentiation. In a normal situation, the switch between the male or female pathway is driven by the action of the Y-located testis-determining gene *SRY* (Sex-determining Region of the Y chromosome) [1]. Once gonad differentiation is engaged, the secretion of testicular anti-Müllerian hormone (AMH) and androgens ensure that

sexual characteristics develop correctly into their final form [2]. At any of these steps, flaws in the process can lead to disorders of sexual development. When chromosomal, gonadic and somatic sexes are not consistent, sexual attribution is ambiguous, and both male and female characteristics may be found in a single individual. In mammals, sex reversal has been described in numerous species including humans [3], pigs [4,5], goats [6,7], horses [8], dogs [9,10], mice [11], marsupials [12] and moles [13]. Naturally occurring models of *SRY*-negative XX sex reversal in XX individuals could provide important clues as to the etiology of some cases of the human disorder and increase our knowledge about sex differentiation in vertebrates. These pathologies can result either from loss-of-function of crucial pro-ovarian genes (for example, *FOXL2* in

goats [6] or *RSPO1* in humans [14]) or from gain-of-function of a testis-promoting factor (for example, *SRY* or *SOX9* in humans [15,16] or in mice [17,18]). Among farm animals, pigs and goats present a XX intersex phenotype at a higher frequency than other livestock species. These XX sex-reversed individuals have a normal female karyotype (XX) but some degree of testicular differentiation of their gonads [19]. In goats, *SRY*-negative XX sex reversal is due to a recessive mutation, referred to as the "PIS" (Polled Intersex Syndrome) mutation, which is associated with the dominant autosomal trait for hornlessness [6,20]. The causative mutation, identified in 2001, is a large 11.7 kb deletion that affects the transcription of *PISRT1* (PIS-regulated transcript 1) and *FOXL2* (Forkhead box L2). In the homozygous state, this mutation causes the masculinization of all polled genetic females [6]. Despite this successful result in goats, the causal genes in *SRY*-negative XX sex reversal in pigs remain unidentified.

Pig intersex phenotypes range from true hermaphrodites (48%) to male phenotypes with (50%) or without (2%) ambiguities [4]. As in humans some of these cases result either from the presence of XX and XY cells within the same animal (4%) or from the presence of a small fragment of the Y chromosome (containing the *SRY* gene) in XX individuals (2%). However, most XX sex-reversed individuals (94%) do not have *SRY* or any of the other genes normally located on the Y chromosome (*SRY*-negative XX sex reversal). In these cases, testis induction is therefore due, to alterations of autosomal genes in the absence of *SRY* [4]. In pigs, the frequency of XX sex reversal varies from 0.08 to 0.75% depending on the herd, and available heritability estimates for this disorder are high, ranging from 0.72 to 0.81 [21]. Previous segregation analyses with pedigrees of limited size are consistent with autosomal recessive transmission of a single gene [4]. A more recent study however, suggests the possible involvement of multiple autosomal regions [22].

Until recently, the identification of genes causing inherited disorders was performed using genome-wide pedigree-based linkage analyses, or by focusing on candidate genes using population-based association studies. With the recent development of commercial high throughput SNP genotyping BeadChips, these two approaches have merged into a single approach referred to as genome-wide association studies (GWAS). GWAS methods have been shown to be more powerful than linkage based methods [23]. The power of GWAS has been demonstrated by the identification of several genes underlying a range of diseases in humans, dogs and cattle which had been difficult to study previously [24,25].

In this contribution we report the results of a GWAS study of XX intersex in French Large White pigs. We genotyped nuclear families identified by one or several sex reversed offspring using the Illumina PorcineSNP60 BeadChip tool to identify chromosome regions and markers associated with sex reversal. A unique association signal was identified and suggests a probable role for the *SOX9* gene in sex reversal in pigs. The *SOX9* region has been described as a gene desert, presenting extremely conserved elements among mammals. It is now clear that *SOX9* is expressed by means of a complex regulatory region thought to cover at least 1 Mb. Among these

sequences, Sekido and Lovell-Badge first located a 3.2-kb testis-specific enhancer of *Sox9* (TES) that was [26] further refined to 1.4 kb and is referred to as the Testis Enhancer Sequence core element (TESCO). Even if *SOX9* was retained as the main candidate gene, TESCO was also candidate for the causative mutation in pigs.

## Materials and Methods

### Animals and Phenotypes

DNA samples were collected from experimental families in the 1990s, and more recently from nuclear families during the national selection program of the French Large-White breed.

The collection of experimental XX sex-reversed cases was established over a 6 year period starting in 1992 in an experimental herd (Domaine de Galles, Avord, France) of the Institut National de la Recherche Agronomique (INRA). This collection has been described previously [4] and some cases were used for in-depth descriptions of XX sex-reversal, including fetal development [5]. Briefly, all cases (n=30) were tested for the presence of the Y chromosome, especially *SRY*. Then, the parents of true XX sex-reversed cases (94%) were retained in the experimental herd and mated to determine the mode of inheritance of the intersex condition in the pig. The reproductive systems of all intersex cases were examined post-mortem in adults (sporadic cases) or at 5 weeks postnatally (familial cases). As a result, 12 nuclear families comprising 1 to 9 affected piglets and their two non-affected parents were selected.

In addition, samples from 26 half/full-sib families were collected between 2006 and 2010 from the Large-White population of the French breeding companies belonging to the BIOPORC association. In addition to blood sampling, the cases were described by the farmer, the technician or the veterinary surgeon using standard nomenclature. The phenotype was defined using only external anatomical criteria: enlarged vulva, presence of a penile clitoris, number of testicles (0, 1 or 2) located in a scrotum-like structure, and presence of a mid-ventral penile sheath. Animals presenting one or more of these criteria were considered as affected and included in the sample. Samples were also collected from their non-affected parents. These two phenotypic categories, affected and non-affected, were used for association analysis. The animals in the experimental families produced between 1992 and 1998 were described by Pailhoux & al. in the frame of another study [4]. All procedures were approved by the animal welfare commission of the Institute at the time of this study. In our study, only the DNA collected from these animals was used. The second set of samples was collected by the BIOPORC association. The collection of blood samples and external phenotyping of the animals were performed in compliance with the guidelines of the French Ministry of Agriculture and Fisheries.

### Sample genotyping

Genomic DNA was extracted from blood samples using standard protocols. For each sample, the presence/absence of the *SRY* gene was determined by PCR amplification with

primer pairs (forward -CTGTAGCCTCTGTGCCTCCT and reverse - TTTTCATTGTGTGGTCTCGTG). Amplicons were identified following electrophoresis on agarose gels and samples with amplified products were not retained for SNP genotyping. Additional paternity tests were performed using 10 microsatellites to exclude any families presenting Mendelian inconsistencies from the design. In total, 185 samples were retained, corresponding to 38 Large-White families with in all 89 affected piglets. The samples were genotyped at the Animal Genotyping Platform LABOGENA, (Jouy-en-Josas, France) using the Illumina PorcineSNP60 BeadChip platform (Illumina, San Diego, USA), according to the manufacturer's protocol.

### Quality control and genome-wide association analyses

The analysis was conducted for autosomal SNPs with the PLINK whole genome association analysis toolset (PLINK version 1.07 <http://pngu.mgh.harvard.edu/purcell/plink/>) [27]. All individuals with a call rate lower than 95% were discarded. Genotype data were obtained for 173 animals (including 89 affected piglets). Markers deviating from the Hardy-Weinberg equilibrium ( $p < 10^{-3}$ ) (741 SNPs), markers with a call rate lower than 95% (4944 SNPs), and markers with a minimum allele frequency (MAF) of less than 1% (8804 SNPs) were excluded from the analysis. Using these quality control criteria, 47,155 SNPs out of the initial set of 62,163 SNPs remained available for the Transmission Disequilibrium Tests. Nominal  $p$  values corresponding to genome-wide thresholds for significance or suggestive association were determined. We applied a conservative Bonferroni correction [28] considering 47,155 independent tests (i.e. the number of SNPs used), which yielded thresholds of  $1.06 \times 10^{-6}$  (significant,  $\log(1/p) = 5.97$ ) and  $1.91 \times 10^{-5}$  (suggestive,  $\log(1/p) = 4.72$ ).

### Identification of CNV's

PennCNV software [29] was used to detect pig CNVs surrounding the SOX9 gene on SSC12. Analysis was performed with all the SNP of the chip (62,163) but only the results obtained on SSC12 were considered. PennCNV software integrates, in a joint-calling algorithm, a Hidden Markov Model (HMM) including the pedigree information when available, the total signal intensity (Log R ratio - LRR) and allelic intensity ratio (B allele frequency - BAF) at each SNP marker, the distance between neighboring SNPs and the population frequency of allele B (PFB). Both LRR and BAF were exported from GenomeStudio (Illumina San Diego, USA) and the PFB file was calculated based on the BAF of each marker. CNV calling was performed using the default parameters of the HMM model with 0.01 as UF factor. The marker positions on SSC12 were derived from the swine genome sequence assembly (10.2) ([http://www.ensembl.org/Sus\\_scrofa/Info/Index](http://www.ensembl.org/Sus_scrofa/Info/Index)).

### TESCO and SOX9 sequencing

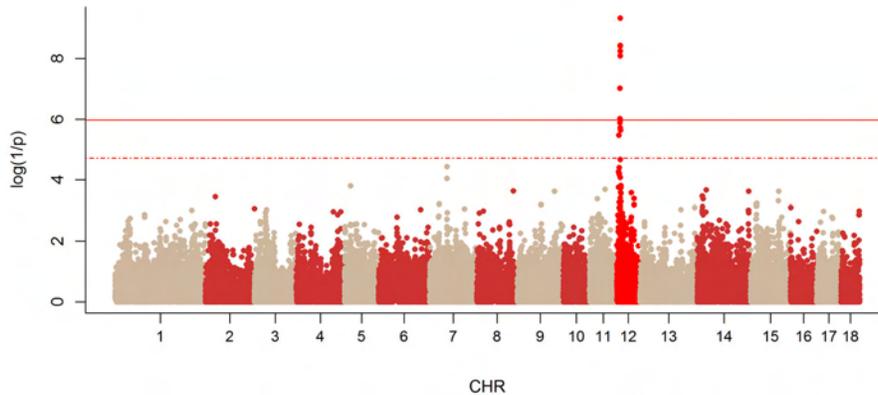
Five affected pigs, carrying different mutated haplotypes (including the 3 most frequent) were selected for sequencing. These five unrelated individuals were selected from five different nuclear families of the design. In addition, five individuals were selected as controls in families from another

project in which no cases of intersexuality were reported. Pairs of primers, covering the entire SOX9 gene and TESCO regulatory region, were chosen to amplify and sequence these 2 candidate regions (Table S1). The SOX9 pig sequence is available and annotated in Ensembl (Sscrofa10.2:12:9028879:9033246:1), whereas the TESCO pig sequence was obtained by aligning TESCO human sequence against the pig draft genome (Sscrofa10.2:12:9045199:9047423:1). To sequence PCR products, an aliquot (1–12  $\mu$ l) was purified in a single step (45 min at 37°C followed by 30 min at 80°C) using 0.5U of Shrimp Alkaline Phosphatase (Promega) and 0.8 U of Exonuclease I (New England Biolabs). Sequencing was performed on a 3730 ABI capillary DNA sequencer using a Big Dye terminator V3.1 cycle sequencing kit (Applied Biosystems, Inc.). Sequences were analyzed using the CodonCode Aligner software v4.0.4 (<http://www.codoncode.com/aligner/>), and any polymorphisms were detected by comparison with the reference sequence. Multiple alignments of the core region of TESCO was performed with multalin software [30]. In total, five distinct sequences (alleles) were obtained for SOX9 (Genbank Accession numbers KF422597, KF422598, KF422599, KF422600, KF422601), and 3 for the TESCO region (Genbank Accession numbers KF422602, KF422603, KF422604).

## Results

### Informativeness of the SNP panel

Thirty-eight half/full-sib families comprising 89 sex-reversed piglets were genotyped with the Illumina PorcineSNP60 BeadChip. After quality controls, 173 individuals and 47,155 autosomal SNP were retained for transmission disequilibrium (TDT) analysis. The number of SNPs per chromosome ranged from 1,051 on SSC18 to 5,216 on SSC1. MAFs of polymorphic markers were uniformly distributed between 0.01 and 0.5 (Figure S1). The power of the TDT was directly proportionate to the number of heterozygous parents for markers associated with causative variants, and the level of linkage disequilibrium (LD) between causative variants and SNP markers. The proportion of heterozygous parents for each of the 47,155 SNPs is represented in Figure S2. At least half of the parents of the design were not informative for more than 85% of the SNPs. The LD estimation was computed as reported by Dupuis & al [31] to assess the level of genome coverage provided by the 47,155 remaining SNPs. Each SNP was considered as a pseudocausative polymorphism in order to identify the SNP (and its distance) with the highest  $r^2$  value ( $r^2_{\max}$ ) among the remaining markers on the chip located within a 2 Mb surrounding window. Calculations were performed using only the parents of the design. The cumulative frequency distribution of  $r^2_{\max}$  is reported in the Figure S3(A). In our pedigree,  $r^2_{\max}$  values were lower than 0.5 for 28% of the SNPs and only 25% of the SNPs were in total linkage disequilibrium ( $r^2_{\max} = 1$ ) with another SNP on the chip. The average distance for SNP with a  $r^2_{\max}$  of 0.5 was 300 kb (Figure S3(B)). The number of remaining markers and the informativeness of the design indicate that the coverage of the SNP panel may not be optimal for all chromosomal regions.



**Figure 1. Manhattan plot of genome-wide TDT analysis for sex-reversal phenotype.** For each marker the nominal significance ( $\log(1/p)$ ) is indicated. The dotted and the solid lines correspond to the genome-wide suggestive and significant thresholds respectively.

doi: 10.1371/journal.pone.0079882.g001

### Transmission Disequilibrium Test

Due to the two periods of birth of the founders, the population structure was quantified using genome-wide pairwise identity-by-state distances estimated by PLINK. All animals were grouped in a same cluster; these results indicate that despite the differences in the collection periods the two populations have not diverged. The SNPs showing significant genome-wide ( $p < 1.06 \times 10^{-6}$ ) and chromosome-wide ( $1.06 \times 10^{-6} < p < 1.33 \times 10^{-5}$ ) association with the intersexuality (IS) phenotype as determined by TDT are presented in Table 1 and the Manhattan plot in Figure 1. All these SNPs were located on SSC12 between position 6,459,472 and 9,597,291 bp, the most significant marker being located at position 8,881,355 bp (M1GA0024789) (Table 1). Because all these SNPs were localized within the same short interval, it seemed likely that they were linked to a single causal locus. No other significant or suggestive regions were detected by the genome scan. No signals were identified close to PISRT 1 and FOXL2 on chromosome 13, in spite of very informative markers in that region (ALGA0119502 : 39 informative meioses ( $p=0.53$ ) and ASGA0089628: 43 informative meioses ( $p=1$ )). The 3.14 Mb segment encompassing the IS locus on swine chromosome 12 includes several transcriptional units, but only one gene (*SOX9*) was localized in the 882,642 kb segment defined by the position of the significant genome-wide markers (Figure 2). The *SOX9* gene is one of the most important genes downstream from *SRY* in the pathway of testicular differentiation. In the significant interval, *SOX9* was therefore considered as a strong candidate gene.

### Detection of CNV

To investigate whether a chromosomal rearrangement (deletion or duplication) surrounding the *SOX9* gene could be causal, genotyping data obtained with the 60K SNP chip were analyzed with PennCNV. The analyses of SNP arrays provide normalized total signal intensity and allelic intensity ratios which represent overall copy numbers and allelic contrasts. To detect CNV, we jointly analyzed the 60KSNP genotypes of the

**Table 1.** Significant SNP association hits in the Transmission Disequilibrium Test for sex-reversal.

SNP name	Chr	Position (bp)	$\chi^2$	p-value	$\log(1/p)$
ASGA0052652	12	6 459 472	21.6	3.36E-06	5.47
ASGA0052953	12	8 586 694	28.47	9.51E-08	7.02
ASGA0098350	12	8 826 005	24	9.63E-07	6.02
ALGA0112008	12	8 835 743	23.4	1.32E-06	5.88
ALGA0108362	12	8 836 960	23.4	1.32E-06	5.88
MARC0011974	12	8 851 013	22.7	1.89E-06	5.72
<b>M1GA0024789</b>	<b>12</b>	<b>8 881 355</b>	<b>38.82</b>	<b>4.65E-10</b>	<b>9.33</b>
MARC0009109	12	8 894 982	34.68	3.89E-09	8.41
ALGA0106073	12	8 897 127	23.82	1.06E-06	5.98
ALGA0064738	12	9 053 320	34.77	3.71E-09	8.43
ASGA0053002	12	9 128 147	33.91	5.78E-09	8.24
H3GA0033370	12	9 257 006	33.91	5.78E-09	8.24
ASGA0097657	12	9 469 336	33.24	8.14E-09	8.09
MARC0041079	12	9 597 291	22.35	2.27E-06	5.64

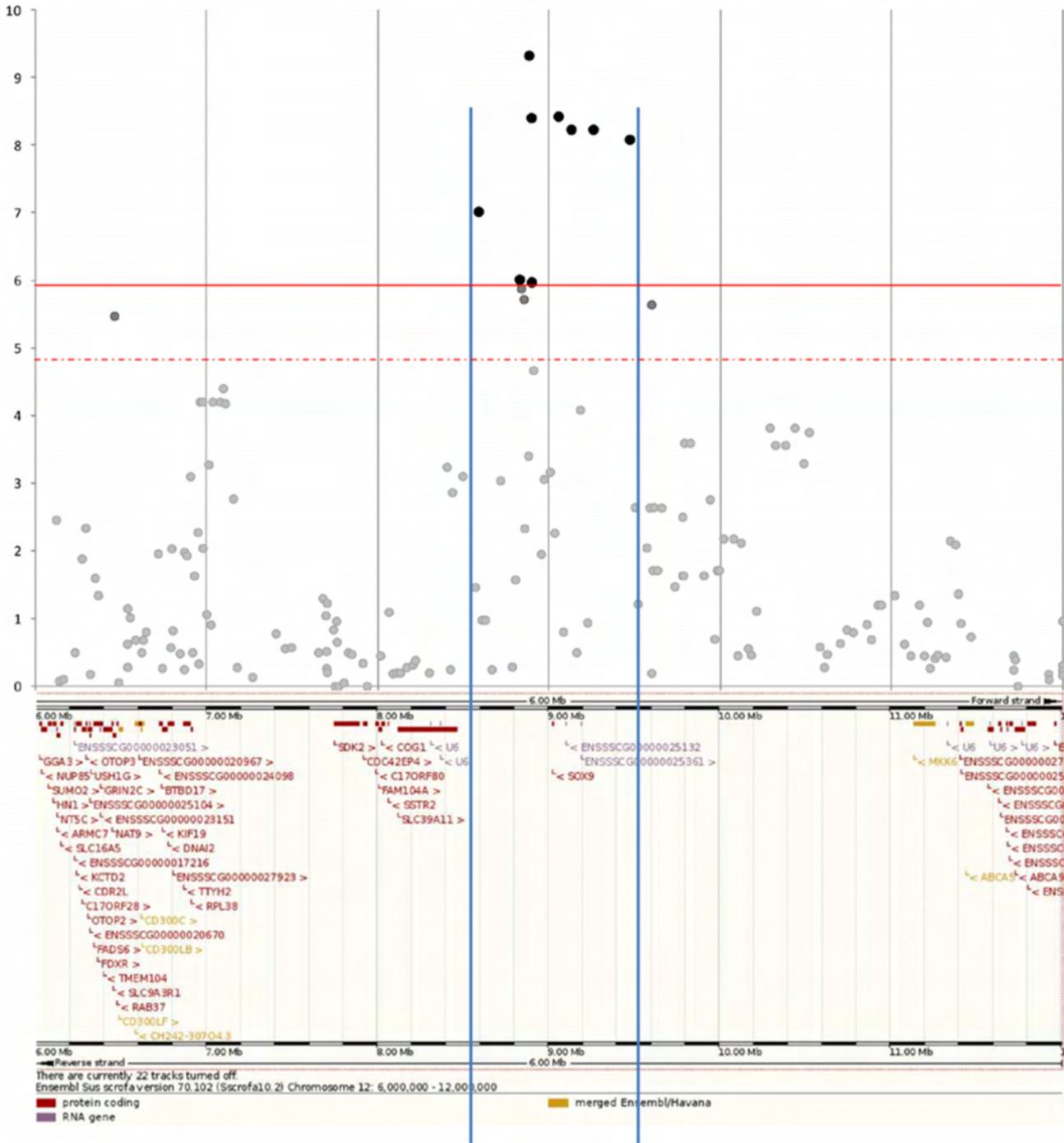
Chromosome-wide significant SNP ( $1.06 \times 10^{-6} < p < 1.33 \times 10^{-5}$ ) are in italic, genome-wide significant SNP ( $p < 1.06 \times 10^{-6}$ ) in normal font, and the SNP corresponding to the lowest p-value is highlighted in bold. Positions were defined on the Sscrofa10.2 draft sequence.

doi: 10.1371/journal.pone.0079882.t001

89 affected piglets. Successive windows of 3 markers, corresponding to 150kb windows on average (given the density of the chip) were analyzed. Among all the affected individuals of our design, no structural variations were identified in the 3.14 Mb segment encompassing *SOX9*.

### Haplotypic characterization of *SOX9* surrounding interval

To further refine the map position of the causative mutation, 35 SNPs of the chip located between positions 8,566,755 and 9,604,096 bp (interval containing "the significant markers" bounded by an additional marker on each side) were selected

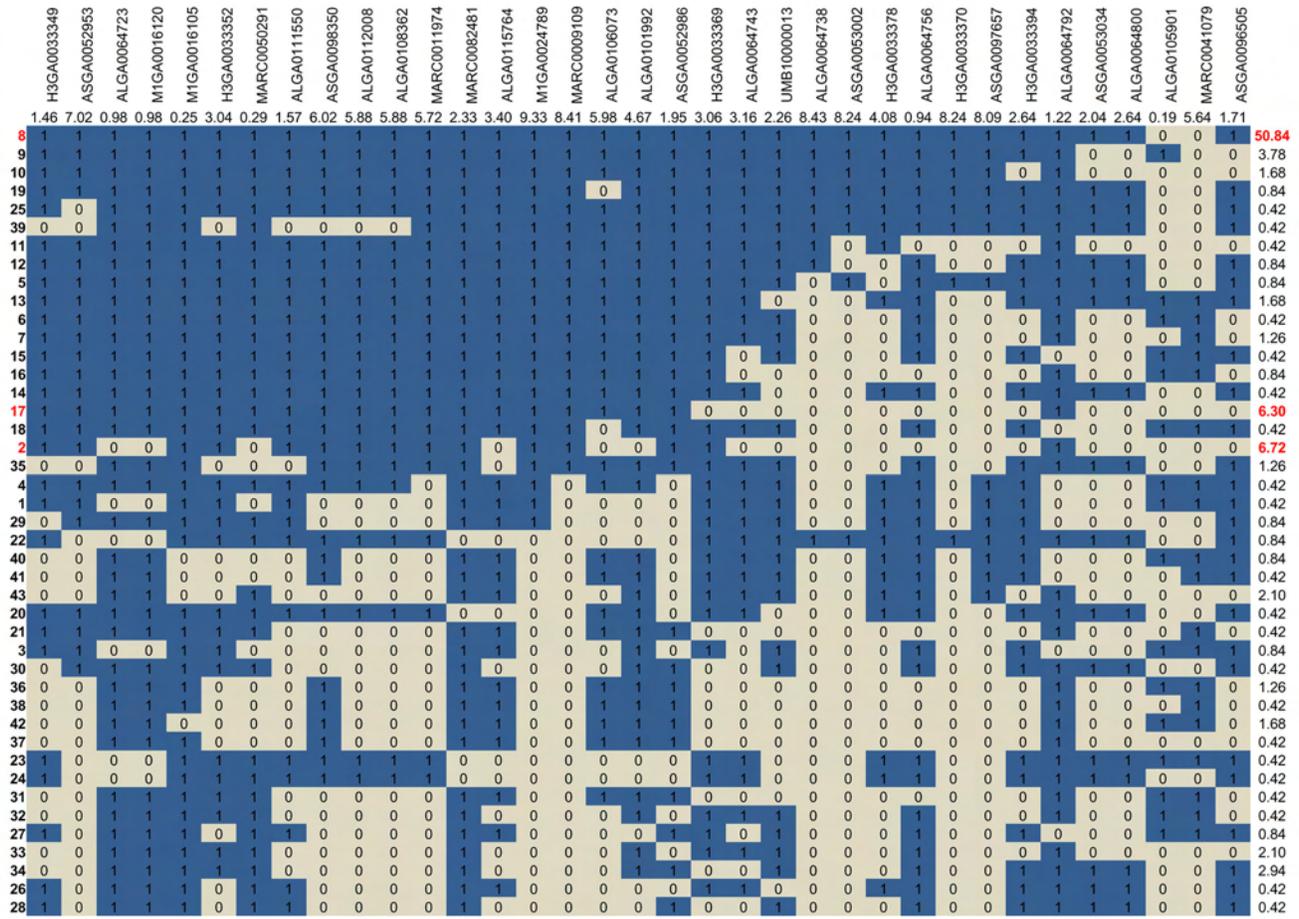


**Figure 2. Log(1/p) values of the SNP localized in the 6Mb-12Mb region on SSC12.** The non-significant markers are presented in grey, suggestive ones in medium grey, the significant SNP in black. The suggestive (dotted line) and significant (solid line) thresholds are indicated in red. The two vertical lines delimit the minimum interval defined by the first and least significant SNP in the region. At the bottom, genes and transcripts annotated in Ensembl (Sscrofa10.2) are reported.

doi: 10.1371/journal.pone.0079882.g002

to characterize the haplotypes segregating among affected offspring. The phases of each descendant were reconstructed using familial allelic transmission from parents to offspring.

Comparing the haplotype-based genotypes, no single homozygous chromosome segment could be identified among the affected piglets of the design. In total 43 different



**Figure 3. List of the 43 haplotypes defined with the 35 SNP of the chip surrounding SOX9.** The haplotypes were classified according to the p-value obtained for each marker (sharing of the transmitted allele from the most significant SNP to the less significant one). For each marker the preferentially transmitted allele was noted "1", the alternative allele "0" and the log(1/p) obtained with the TDT analysis was reported. On the right, the frequency of each haplotype among the affected animals is reported (the 3 most frequent ones are indicated in red). The marker shown in red box is the marker closest to the gene SOX9.  
doi: 10.1371/journal.pone.0079882.g003

haplotypes segregate within the 38 families. However, among the affected offspring, the frequency of haplotype "8" (50.8%) was much higher than the frequencies of other haplotypes (Figure 3). In addition the frequencies of two other haplotypes, haplotypes 2 (6.7%) and 17 (6.3%) were slightly higher than the others.

No IBS segment was found in common in the 43 haplotypes, but most of them shared segments of variable size (2 to 30 adjacent markers) identical to the most frequent haplotype. These results suggest that a mutation may be shared by several haplotypes, including the three more frequent ones.

**Candidate region sequencing**

The minimum interval mapped by genetic analysis covered 882,642 kb and the causative mutation could be located anywhere within it. Before undertaking an exhaustive search for all polymorphisms in this region, we chose to target the regions annotated in the interval: the SOX9 gene and its

regulatory region TESCO. We could not exclude that a substitution of an amino-acid residue might increase the lifespan or the DNA affinity of the protein, or block the nuclear exportation signal and lead to abnormal nuclear retention of SOX9 in the XX gonads [32]. To test whether the causal mutation might be located in the SOX9 gene or in its regulatory TESCO region, these elements were sequenced in five individuals. Four were homozygous or heterozygous carriers for haplotypes "8", "2" and "17", and the fifth was heterozygous for two rare haplotypes. The aim was (1) to search for candidate mutations that might affect the functional structure of the gene or (2) to identify a small segment shared between the "carrier" haplotypes. Even if a candidate mutation was not identified in the sequences of SOX9 and TESCO, the presence of a single haplotype shared among all affected individuals would suggest that the mutation is close to these two elements.

Additionally, five healthy animals were also sequenced. The genomic sequence of SOX9, including the 5'UTR, the 3 exons,

		TESCO												SOX9																
														5'UTR			Intron 1				Intron 2			3'UTR						
		361	527	555	571	636-637	735	788	1001	1229	1256	1268	1278	1286	1297	1334	134	217	247	942	1075	1421	1563	1661	2223	2409	2483	2591	3569	3570
Case 8	8	A	G	C	C	A	G	G	T	T	A	C	C	T	A	A	T	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	C	G	C	C	-
	8	A	G	C	C	A	G	G	T	T	A	C	C	T	A	A	T	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	C	G	C	C	-
Case 8	2	A	G	C	C	A	G	G	T	T	A	C	C	T	A	A	T	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	C	G	C	C	-
	2	T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	G	T	T	C	-
Case 34	43	T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 6	C	G	C	C	T
	43	T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 6	C	G	C	C	T
Case 2	36	T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	G	T	T	C	-
	36	T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	G	T	T	C	-
Case 17	17	T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	G	T	T	C	-
	17	T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	G	T	T	C	-
Control		T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	G	T	T	C	-
		T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	G	T	T	C	-
Control		T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	-----	CGGGG	G	G	C	T	(CG) 7	G	T	T	C	-
		T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	-----	CGGGG	G	G	C	T	(CG) 7	G	T	T	C	-
Control		A	G	C	C	A	G	A	T	C	G	A	T	G	G	C	T	T	-----	-----	A	A	T	T	(CG) 4	C	T	T	T	T
		T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 6	C	G	C	C	T
Control		A	G	C	C	A	G	A	T	C	G	A	T	G	G	C	T	T	-----	-----	A	A	T	T	(CG) 4	C	T	T	T	T
		T	A	T	A	--	A	A	C	C	G	A	C	G	G	C	C	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 6	C	G	C	C	T
Control		A	G	C	C	A	G	G	T	T	A	C	C	T	A	A	T	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	C	G	C	C	-
		A	G	C	C	A	G	G	T	T	A	C	C	T	A	A	T	G	TGCAGCCGAGTGACGCGC	-----	G	G	C	A	(CG) 7	C	G	C	C	-

**Figure 4. Polymorphisms identified for the 5 cases and 5 controls individuals used for sequencing.** For each case, the number in red corresponds to the haplotype number defined with the 35 SNP of the chip surrounding SOX9. For each individual the phases defined in *TESCO* and *SOX9* with the detected polymorphisms is reported. Positions were defined from position 9033579 (=1) on the pig draft sequence for *SOX9* and 9047123 (=1) for *TESCO*. Five different alleles were detected for *SOX9* (Genbank Accession numbers: KF422597, KF422598, KF422599, KF422600, KF422601), and 3 for the *TESCO* (KF422602, KF422603, KF422604).

doi: 10.1371/journal.pone.0079882.g004

the 2 introns and the 3' UTR regions, was obtained using 7 pairs of primers and the 1.4kb core region of *TESCO* was obtained with 4 pairs of primers. In the genomic sequence of *SOX9*, 14 different polymorphisms were identified (10 substitutions and 4 indels from 1 to 18 nucleotides). All these polymorphisms were located outside of the gene exons and outside of the intron / exon junctions. One Indel and 14 substitutions were characterized in the 1.4 kb *TESCO* region. Consequently, no functional candidate mutation was highlighted (Figure S4). The distribution of the polymorphisms among the sequenced individuals (Figure 4) showed five different alleles for *SOX9* and 3 for *TESCO*. No shared allele was present between the different "carrier" chromosomes, and identical alleles were identified on "carrier" and "control" chromosomes. Among the control animals, two were

homozygous for the same alleles as those characterized in case pigs (Figure S4).

### Discussion

This study focused on the identification of the gene(s) underlying XX sex reversal in the pig. A TDT was conducted using the Illumina PorcineSNP60 array to detect risk loci. In this family-based study, we found genome-wide association with intersexuality for 9 SNPs on SSC12. All these markers are localized in a same small interval that also contains the strong candidate gene, *SOX9*. Some of the experimental families included in the present analysis had been analyzed in 2001 for the expression level of *SOX9* in gonads at 4 developmental stages (50 and 70 dpc; 5 weeks after birth and adult) [5]. This included gonads of XX males with genital ambiguities (mainly

testicular tissue) and of true hermaphrodites for which testis-like gonads and ovarian-like gonads had been studied separately. Results had shown that the level of SOX9 increases in testis-like gonads of XX males and of true hermaphrodites, whereas it remains at a normal female level in the ovarian-like tissues of the latter. However, whereas these results alone were not proof for genetic causality, a SOX9 genetic determinism has now stronger support with the association signal obtained here. No additional suggestive regions were identified in accordance with a monogenic autosomal recessive mode of inheritance as previously proposed [4].

In the mouse, the transcription factor SOX9 has been shown to be the direct target of the protein encoded by SRY [26]. It has several relevant functions during development and is essential for testis differentiation. In the mouse, Sox9 is initially expressed in the bi-potential gonad of both sexes at 10.5 day post conception (dpc). It then becomes highly up-regulated in XY gonads and down-regulated in XX gonads at 11.5 dpc [33]. In males, it plays an essential role during Sertoli cell differentiation from supporting cell precursors, which is the preliminary step in testis differentiation [34]. In females, the down-regulation of SOX9 is maintained in the granulosa cells throughout fetal development. In pigs, SOX9 is expressed in a similar pattern to that described in mice [35]. Porcine SOX9 expression during embryonic and fetal development begins at 21 dpc, in both sexes and then increases in the testis from 28dpc on. As in mice, expression in the female is intense as early as 21 dpc, and then decreases from 28 dpc, the key stage of gonadal switch in pigs.

Both loss- and gain-of-function studies in the mouse support the implication of the Sox9 gene in gonadal differentiation. A conditional deletion of Sox9 in the gonads results in XY male-to-female sex reversal [36,37] whereas misexpression of a Sox9 transgene in XX mice induces female-to-male sex reversal [17,38]. Similar phenotypes have been found in humans. Mutations within and outside of SOX9 leading to haploinsufficiency are known to cause a bone disorder, campomelic dysplasia, in both sexes and gonadal dysgenesis in 75% of XY individuals presenting as females [39,40]. Even if female-to-male sex reversal in humans is rare, Huang & al. reported in 1999 [16] evidence supporting that SOX9 duplication, through a *de novo* mosaic 46,XX,dup(17)(q23.1q24.3), can cause XX sex reversal. In 2011, Cox & al described a familial 46, XX Developmental Testicular Disorder due to an approximately 178-kb duplication located 600 kb upstream of SOX9 [41] and Vetro & al., a 96 kb triplication 500 kb upstream of SOX9 [42]. Several copy number variants (CNV) and translocations within a large gene desert region upstream of SOX9 have been characterized as responsible for some of these developmental disorders affecting the skeleton and the genitalia [43,44]. It is now clear that Sox9 is expressed in various tissues by means of a complex regulatory region thought to cover at least 1 Mb [45]. Based on the large amount of information obtained in other species, we looked for candidate CNVs within the interval detected by TDT that might disrupt regulatory regions. We thus mined our 60K genotyping data for CNVs, but were not able to identify any structural

variations. No CNVs have been reported in this interval in previously published studies in different pig breeds [46,47]. Although CNV detection is feasible with this technology, it is impaired by low marker density, irregular distribution of SNPs along chromosomes and the lack of non-polymorphic probes specifically designed for CNV identification. Hence, only the largest CNVs are expected to be detected using the Porcine 60K SNP Chip. We cannot therefore exclude the possibility that CNVs of smaller sizes might exist in the candidate interval. Results obtained in humans with a dense SNP chip reported much higher numbers of CNVs with smaller average sizes than with a less dense SNP chip. Based on our analysis, we cannot exclude that the causal mutation is a small deletion, but it is unlikely that a major structural variation is responsible for the sex-reversed cases studied.

To further refine the localization of the mutation, haplotype analysis was performed on the entire interval. Forty-three different haplotypes were identified among the affected individuals. Although the hypothesis tested by the TDT analysis was that intersexuality is determined by a single autosomal recessive locus, no shared segment was identified among the different "carrier" haplotypes. However one of the haplotypes (haplotype "8") was shown to be more common than any other, and more than half of the 43 haplotypes had a shared segment of 2 to 30 adjacent markers, with haplotype "8". Although these similarities are consistent with the strong association found in this study, some of the haplotypes identified in sex-reversed individuals were very different from the major haplotype "8".

The second targeted analysis that we performed within the candidate interval was the re-sequencing of the SOX9 gene and TESCO region on different "carrier" and "control" chromosomes. The different elements regulating site- and stage-specific transcription of the SOX9 gene and all the core regulation elements are still not characterized in detail; however in 2008, Sekido and Lovell-Badge [26] identified a 3.2 kb element responsible for the testis-specific enhancer (TES) of Sox9, located 13 kb upstream of the mouse transcription start of the gene. A 1.4 kb core region within TES (TESCO) has been found to be highly conserved across mouse, rat, dog and human genomes, and a 180-bp sequence within TESCO is conserved even in amphibians [48]. No functional candidate mutations were detected within these sequences. The most informative results of this comprehensive analysis of the SOX9 gene and TESCO variability were (1) the absence of a shared segment between the different "carrier" chromosomes, and to the contrary (2) the presence of identical alleles among "carrier" and "control" chromosomes. Altogether, these results suggest that the causative mutation is not in the sequenced regions, meaning that it is not "in" or "near" the SOX9 gene or TESCO.

Up to now haplotype analysis and sequencing have not revealed a single haplotype associated with the causal mutation. Some affected individuals carry rare haplotypes very different from haplotype "8". Different hypotheses can be proposed to explain these results: (1) First, these results could be explained if some individuals were incorrectly phenotyped as could be the case when managing collections from numerous breeding farms. Although an information sheet

describing the disease was drawn up to ensure that phenotyping was as consistent as possible, we cannot exclude that a few of the "affected animals" were in fact, unaffected. Pig intersex phenotypes range from true hermaphrodites to male phenotypes with or without ambiguities [4]. Most of the sex-reversed animals used in this study were included based on an external morphological description only, and for the TDT all the phenotypic variants were pooled in a single category of affected animals. The wide range of possible phenotypes indicates that the "expressivity" of the trait is variable, and a small number of animals might be poorly phenotyped and misclassified (2). Secondly the absence of a common segment shared by different "carrier" haplotypes could also be due to the informativeness of the markers used for the analysis. The data obtained so far indicate that none of the SNPs on the chip is in total linkage disequilibrium with the causal mutation. SNP markers on the Illumina PorcineSNP60 BeadChip were selected in order to achieve pan-genetic analyzes, regardless of the breed and whatever the nature of the traits of interest [49]. The LW breed was one of the seven economically important pig breeds used to validate the SNPs selected for the chip, and the highest number of polymorphic loci was obtained for the LW samples. All the families of the XX sex reversal design were from LW breed, which is the best situation (from the breed point of view) to take fully advantage of the informativeness of the markers on the chip. An additional quality criterion used to select the SNPs used on the chip was the MAF value obtained in a validation population panel. The average MAF for the SNPs on the chip is 0.274, and 90% of the SNPs have a MAF >0.15 (criteria of the three first waves of selection). The frequency of XX sex reversed animals in pig populations is less than 1%, and consequently the expected frequency of the mutated allele must be low. It was therefore unlikely that the SNPs on the chip would show strong linkage disequilibrium with the causal mutation. If the chip is of good quality for the analysis of purebred LW animals *a priori*, its use for mapping loci influencing rare trait variations is not necessarily optimal. This sub-optimal situation could explain why no common homozygous haplotype for all the XX sex reversed animals was identified (3). Thirdly, it is necessary to clarify whether allelic (or a genetic) heterogeneity is possible within the analyzed pedigree. The signal obtained on chromosome 12 in the region of SOX9 targets one mutation affecting gene regulation and certainly explains most cases of XX sex-reversed animals in this design. However, if a second independent gene mutation, also inducing a deregulation of SOX9, segregates in the pedigree, some animals could have two mutated copies but be double-heterozygous "affected/+". "affected/+" with two different haplotypes. In the SOX9 region, these individuals would be heterozygous for several polymorphisms. Finally, we also cannot exclude that another gene in another chromosomal region might induce an intersex phenotype that is therefore independent of SOX9. This hypothesis could not however be confirmed in the present design, as most individuals shared part of haplotype "8". Further analyses should be conducted by excluding these pigs, leading to a highly reduced number of affected pigs in the residual test population.

So far, the genetic determinism of only a small percentage of sex-reversed pigs is known. The presence and expression of the *SRY* gene (due to chromosomal rearrangement or trisomy) does not however, explain the vast majority of the cases. Here we report the identification of a strong association with genetic markers near the *SOX9* gene. To date, this result has not been reported in pigs and the identification of the causative mutation will add to the knowledge and understanding of sexual differentiation in mammals. The data published here support the continued exploration of the regions that regulate *SOX9* and further mapping of the interval to find the causal mutation. At first, a special effort should be made to increase the power of the design by adding (1) additional individuals and families and (2) SNP markers. Having a larger number of cases would allow help to refine the phenotypes into different sub-categories and understand the variability of the phenotype expression by identifying other potential regulatory genes. In parallel the full re-sequencing of cases and controls in the candidate interval will provide more informative SNP markers presenting a higher LD with the causal mutation. However, the identification of the candidate "8" haplotype means that its frequency can now be estimated in different populations and therefore the number of potentially sex-reversed unexplained cases predicted. From a functional point of view, this haplotype can be used as of now to produce controlled crosses for a detailed exploration of the underlying physiological mechanisms.

## Supporting Information

**Figure S1. Frequency distributions of minor allele frequencies (MAF).**

(TIF)

**Figure S2. Cumulative frequency of SNP as a function of the heterozygous parents frequency.**

(TIF)

**Figure S3. Estimation of the linkage disequilibrium among the founders of the design.**

(TIF)

**Figure S4. Localization of the detected polymorphisms in *TESCO* and *SOX9*.**

(TIF)

**Table S1. List of primer pairs used for the amplifications of the *SOX9* and *TESCO* regions.**

(TIF)

## Acknowledgements

We gratefully acknowledge BIOPORC gathering the French pig breeding companies (ADN, NUCLEUS, Gene+ and Pen Ar Lan) for their contribution in the constitution of the collection. We also thank LABOGENA for its technical assistance with the microarray genotyping. We thank J. Demars, H. Gilbert and *P. mulsant* for helpful discussion and revision of the manuscript.

The authors also thank Byron C Jones for its precious help for the English revision of the manuscript.

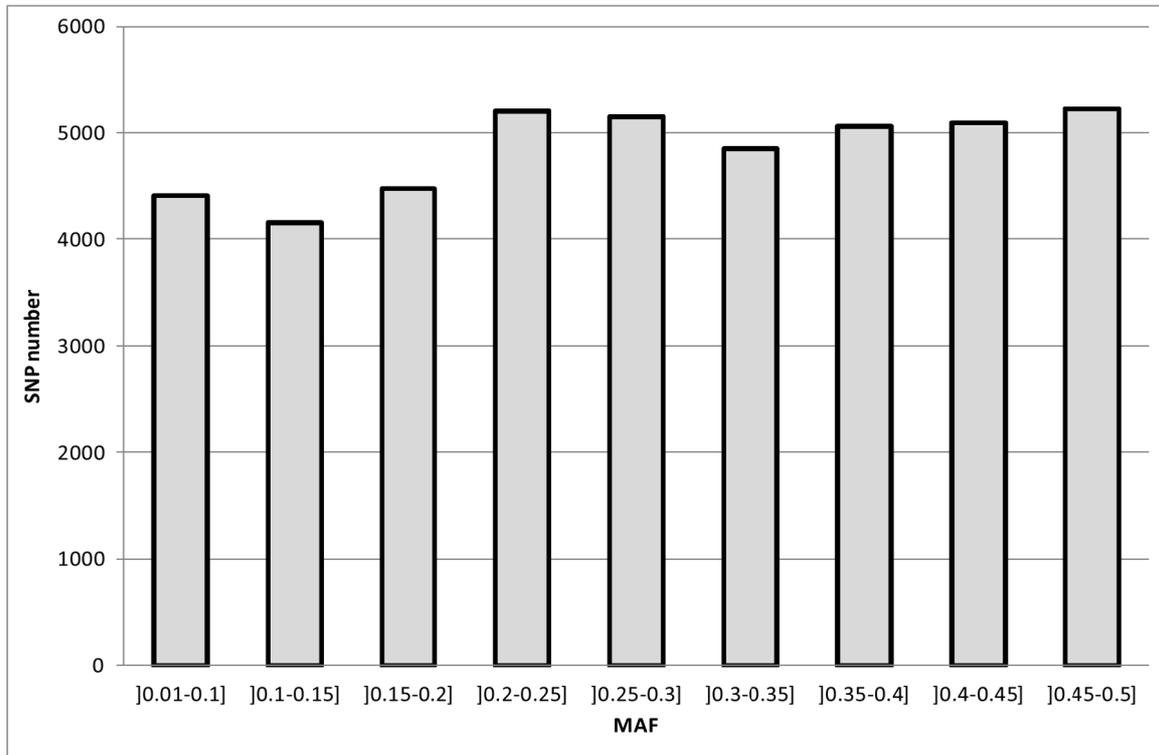
## Author Contributions

Conceived and designed the experiments: DM EP JR.  
Performed the experiments: SR NI. Analyzed the data: SR BS

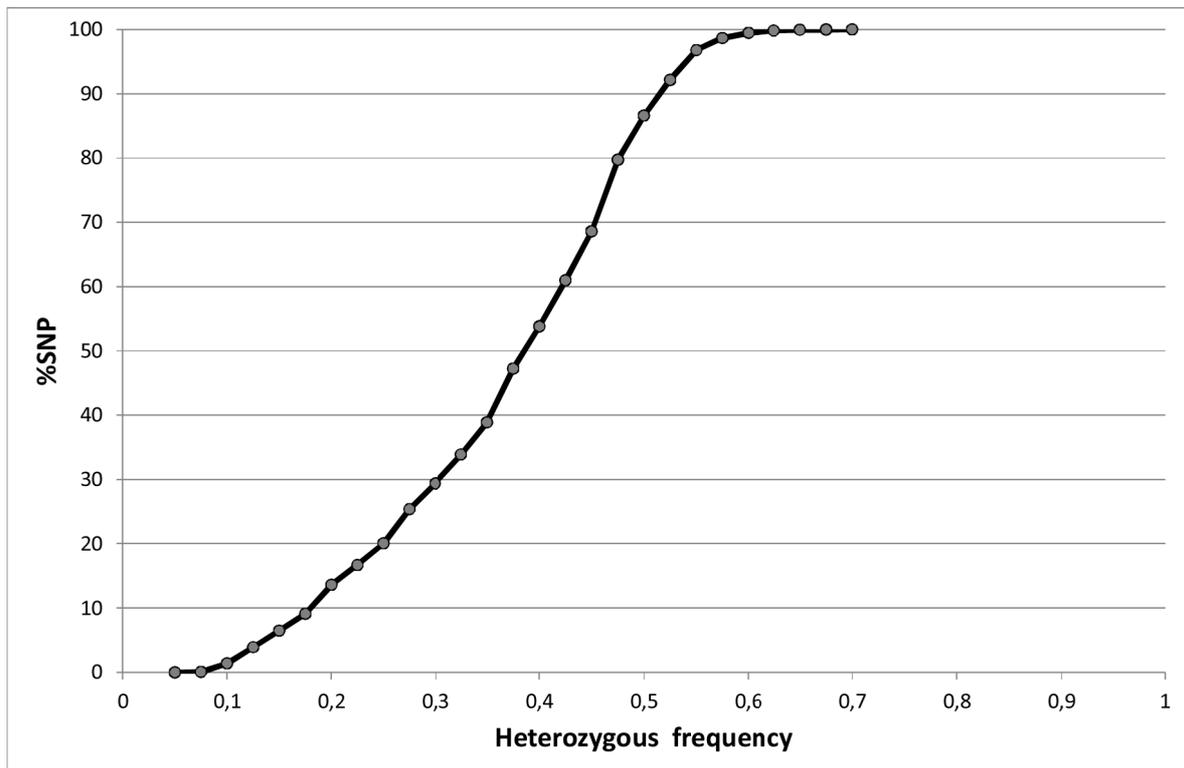
## References

- Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL et al. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* 346: 240–244. doi: 10.1038/346240a0. PubMed: 1695712.
- Eggers S, Sinclair A (2012) Mammalian sex determination—insights from humans and mice. *Chromosome\_Res* 20: 215–238. PubMed: 22290220.
- Hughes IA, Houk C, Ahmed SF, Lee PA (2006) Consensus statement on management of intersex disorders. *J Pediatr Urol* 2: 148–162. doi: 10.1016/j.jpuro.2006.03.004. PubMed: 18947601.
- Pailhoux E, Pelliniemi L, Barbosa A, Parma P, Kuopio T et al. (1997) Relevance of intersexuality to breeding and reproductive biotechnology programs; XX sex reversal in pigs. *Theriogenology* 47: 93–102. doi: 10.1016/S0093-691X(96)00343-3.
- Pailhoux E, Parma P, Sundström J, Vigier B, Servel N et al. (2001) Time course of female-to-male sex reversal in 38,XX fetal and postnatal pigs. *Dev Dyn* 222: 328–340. doi:10.1002/dvdy.1194. PubMed: 11747069.
- Pailhoux E, Vigier B, Chaffaux S, Servel N, Taourit S et al. (2001) A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nat Genet* 29: 453–458. doi:10.1038/ng769. PubMed: 11726932.
- Pailhoux E, Vigier B, Vaiman D, Schibler L, Vaiman A et al. (2001) Contribution of domestic animals to the identification of new genes involved in sex determination. *J Exp Zool* 290: 700–708. doi:10.1002/jez.1120. PubMed: 11748618.
- Lear TL, McGee RB (2012) Disorders of sexual development in the domestic horse, *Equus caballus*. *Sex Dev* 6: 61–71. doi: 10.1159/000334048. PubMed: 22095202.
- Meyers-Wallen VN, Schlafer D, Barr I, Lovell-Badge R, Keyzner A (1999) Sry-negative XX sex reversal in purebred dogs. *Mol Reprod Dev* 53: 266–273. doi:10.1002/(SICI)1098-2795(199907)53:3. PubMed: 10369387.
- Meyers-Wallen VN (2012) Gonadal and sex differentiation abnormalities of dogs and cats. *Sex Dev* 6: 46–60. doi: 10.1159/000332740. PubMed: 22005097.
- Nagamine CM, Morohashi K, Carlisle C, Chang DK (1999) Sex reversal caused by *Mus musculus domesticus* Y chromosomes linked to variant expression of the testis-determining gene *Sry*. *Dev Biol* 216: 182–194. doi:10.1006/dbio.1999.9436. PubMed: 10588871.
- Whitworth DJ, Shaw G, Renfree MB (1996) Gonadal sex reversal of the developing marsupial ovary in vivo and in vitro. *Development* 122: 4057–4063. PubMed: 9012525.
- Jiménez R, Burgos M, Sánchez A, Sinclair AH, Alarcón FJ et al. (1993) Fertile females of the mole *Talpa occidentalis* are phenotypic intersexes with ovotestes. *Development* 118: 1303–1311. PubMed: 8269856.
- Parma P, Radi O, Vidal V, Chaboissier MC, Dellambra E et al. (2006) R-spondin1 is essential in sex determination, skin differentiation and malignancy. *Nat Genet* 38: 1304–1309. doi:10.1038/ng1907. PubMed: 17041600.
- Berta P, Hawkins JR, Sinclair AH, Taylor A, Griffiths BL et al. (1990) Genetic evidence equating *SRY* and the testis-determining factor. *Nature* 348: 448–450. doi:10.1038/348448A0. PubMed: 2247149.
- Huang B, Wang S, Ning Y, Lamb AN, Bartley J (1999) Autosomal XX sex reversal caused by duplication of *SOX9*. *Am J Med Genet* 87: 349–353. doi:10.1002/(SICI)1096-8628(19991203)87:4. PubMed: 10588843.
- Bishop CE, Whitworth DJ, Qin Y, Agoulnik AI, Agoulnik IU et al. (2000) A transgenic insertion upstream of *sox9* is associated with dominant XX sex reversal in the mouse. *Nat Genet* 26: 490–494. doi:10.1038/82652. PubMed: 11101852.
- Koopman P, Gubbay J, Vivian N, Goodfellow P, Lovell-Badge R (1991) Male development of chromosomally female mice transgenic for *Sry*. *Nature* 351: 117–121. doi:10.1038/351117a0. PubMed: 2030730.
- Hunter RH (1996) Aetiology of intersexuality in female (XX) pigs, with novel molecular interpretations. *Mol Reprod Dev* 45: 392–402. doi: 10.1002/(SICI)1098-2795(199611)45:3. PubMed: 8916051.
- Pannetier M, Elzaïat M, Thépot D, Pailhoux E (2012) Telling the story of XX sex reversal in the goat: highlighting the sex-crossroad in domestic mammals. *Sex Dev* 6: 33–45. doi:10.1159/000334056. PubMed: 22094227.
- Larzul C, Delaunay I, Schwob S, Mercat MJ (2008) Paramètres génétiques des principales anomalies congénitales porcines. 40es Journées de la Recherche Porcine, Paris FRANCE 40: 141–142
- Pailhoux E, Mandon-Pepin B, Cotinot C (2001) Mammalian gonadal differentiation: the pig model. *Reprod Suppl* 58: 65–80. PubMed: 11980203.
- Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* 273: 1516–1517. doi:10.1126/science.273.5281.1516. PubMed: 8801636.
- Charlier C, Coppieters W, Rollin F, Desmecht D, Agerholm JS et al. (2008) Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nat Genet* 40: 449–454. doi:10.1038/ng.96. PubMed: 18344998.
- Lequarré AS, Andersson L, André C, Fredholm M, Hitte C et al. (2011) LUPA: a European initiative taking advantage of the canine genome architecture for unravelling complex disorders in both human and dogs. *Vet J* 189: 155–159. doi:10.1016/j.tvjl.2011.06.013. PubMed: 2175267526.
- Sekido R, Lovell-Badge R (2008) Sex determination involves synergistic action of *SRY* and *SF1* on a specific *Sox9* enhancer. *Nature* 453: 930–934. doi:10.1038/nature06944. PubMed: 18454134.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575. doi:10.1086/519795. PubMed: 17701901.
- Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11: 241–247. doi:10.1038/ng1195-241. PubMed: 7581446.
- Wang K, Li M, Hadley D, Liu R, Glessner J et al. (2007) PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res*, 17: 1665–1674. PubMed: 17921354.
- Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 16: 10881–10890. doi:10.1093/nar/16.22.10881. PubMed: 2849754.
- Dupuis MC, Zhang Z, Druet T, Denoix JM, Charlier C et al. (2011) Results of a haplotype-based GWAS for recurrent laryngeal neuropathy in the horse. *Mamm Genome* 22: 613–620. doi:10.1007/s00335-011-9337-3. PubMed: 2169847232.
- Gasca S, Canizares J, De Santa Barbara P, Mejean C, Poulat F et al. (2002) A nuclear export signal within the high mobility group domain regulates the nucleocytoplasmic translocation of *SOX9* during sexual determination. *Proc Natl Acad Sci U S A* 99: 11199–11204. doi: 10.1073/pnas.172383099. PubMed: 12169669.
- Morais da Silva S, Hacker A, Harley V, Goodfellow P, Swain A et al. (1996) *Sox9* expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. *Nat Genet* 14: 62–68. doi:10.1038/ng0996-62. PubMed: 8782821.
- Sekido R, Lovell-Badge R (2009) Sex determination and *SRY*: down to a wink and a nudge? *Trends Genet* 25: 19–29. doi:10.1016/j.tig.2008.10.008. PubMed: 19027189.
- Parma P, Pailhoux E, Cotinot C (1999) Reverse transcription-polymerase chain reaction analysis of genes involved in gonadal differentiation in pigs. *Biol Reprod* 61: 741–748. doi:10.1095/biolreprod61.3.741. PubMed: 10456852.
- Chaboissier MC, Kobayashi A, Vidal VI, Lützkendorf S, van de Kant HJ et al. (2004) Functional analysis of *Sox8* and *Sox9* during sex

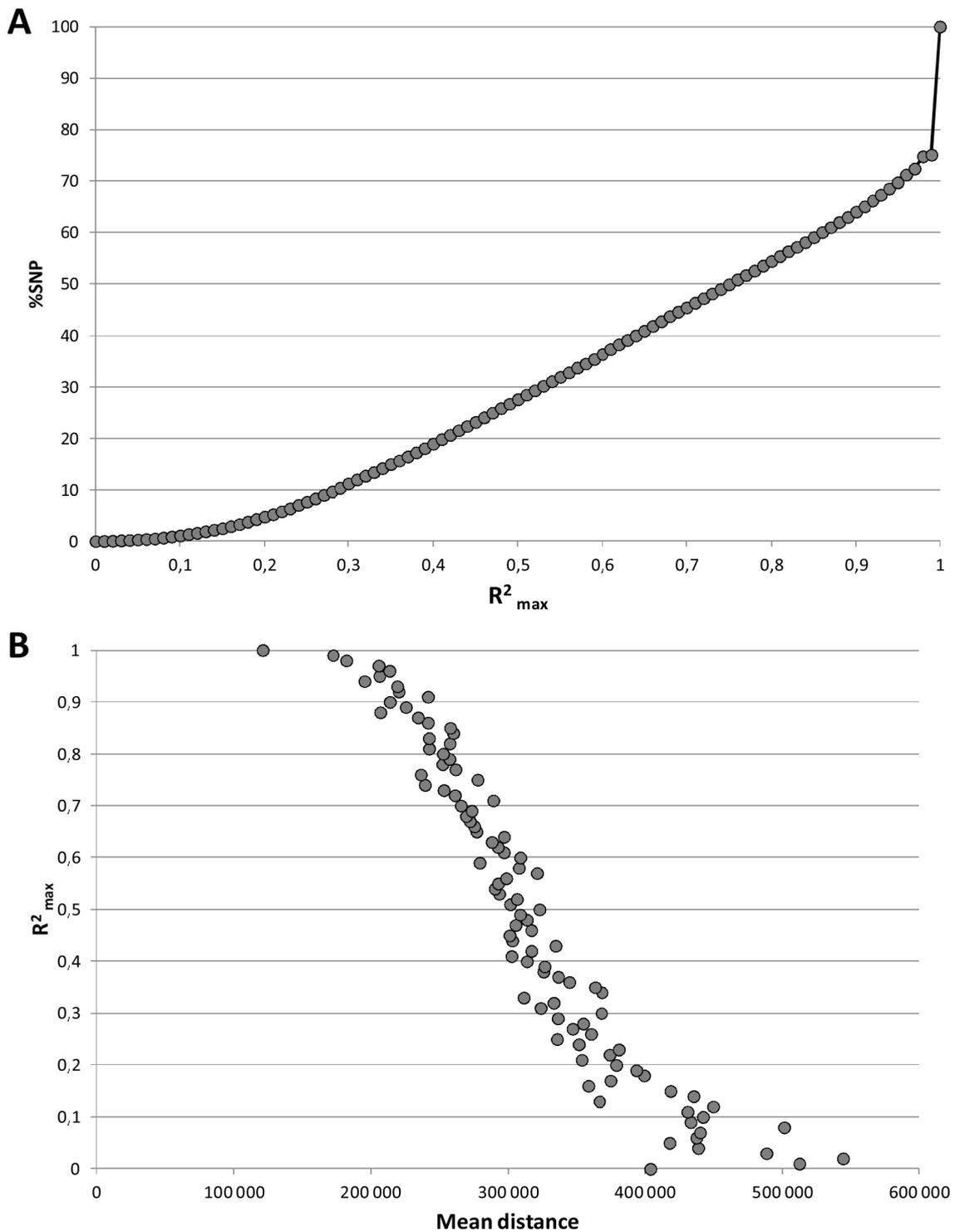
- determination in the mouse. *Development* 131: 1891–1901. doi: 10.1242/dev.01087. PubMed: 1505661537.
37. Lavery R, Lardenois A, Ranc-Jianmotamedi F, Pauper E, Gregoire EP et al. (2011) XY Sox9 embryonic loss-of-function mouse mutants show complete sex reversal and produce partially fertile XY oocytes. *Dev Biol* 354: 111–122. doi:10.1016/j.ydbio.2011.03.029. PubMed: 21466799.
  38. Vidal VP, Chaboissier MC, De Rooij DG, Schedl A (2001) Sox9 induces testis development in XX transgenic mice. *Nat Genet* 28: 216–217. doi: 10.1038/90046. PubMed: 11431689.
  39. Foster JW, Dominguez-Steglich MA, Guioli S, Kwok C, Weller PA et al. (1994) Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature* 371: 525–530. PubMed: 7990924.
  40. Wagner T, Wirth J, Meyer J, Zabel B, Held M et al. (1994) Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell* 79: 1111–1120. doi: 10.1016/0092-8674(94)90041-8. PubMed: 8001137.
  41. Cox JJ, Willatt L, Homfray T, Woods CG (2011) A SOX9 duplication and familial 46,XX developmental testicular disorder. *N Engl J Med* 364: 91–93. doi:10.1056/NEJMc1010311. PubMed: 21208124.
  42. Vetro A, Ciccone R, Giorda R, Patricelli MG, Della Mina E et al. (2011) males SRY negative: a confirmed cause of infertility XX. *J Med Genet* 48: 710–712. doi:10.1136/jmedgenet-2011-100036. PubMed: 21653197.
  43. White S, Ohnesorg T, Notini A, Roeszler K, Hewitt J et al. (2011) Copy number variation in patients with disorders of sex development due to 46,XY gonadal dysgenesis. *PLOS ONE* 6: e17793. doi:10.1371/journal.pone.0017793. PubMed: 21408189.
  44. Jakubiczka S, Schröder C, Ullmann R, Volleth M, Ledig S et al. (2010) Translocation and deletion around SOX9 in a patient with acampomelic campomelic dysplasia and sex reversal. *Sex Dev* 4: 143–149. doi: 10.1159/000302403. PubMed: 20453475.
  45. Bagheri-Fam S, Barrionuevo F, Dohrmann U, Günther T, Schüle R et al. (2006) Long-range upstream and downstream enhancers control distinct subsets of the complex spatiotemporal Sox9 expression pattern. *Dev Biol* 291: 382–397. doi:10.1016/j.ydbio.2005.11.013. PubMed: 16458883.
  46. Wang J, Jiang J, Fu W, Jiang L, Ding X et al. (2012) A genome-wide detection of copy number variations using SNP genotyping arrays in swine. *BMC Genomics* 13: 273. doi:10.1186/1471-2164-13-273. PubMed: 22726314.
  47. Ramayo-Caldas Y, Castelló A, Pena RN, Alves E, Mercadé A et al. (2010) Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. *BMC Genomics* 11: 593. doi: 10.1186/1471-2164-11-593. PubMed: 20969757.
  48. Bagheri-Fam S, Sreenivasan R, Bernard P, Knowler KC, Sekido R et al. (2012) Sox9 gene regulation and the loss of the XY/XX sex-determining mechanism in the mole vole *Ellobius lutescens*. *Chromosome Res* 20: 191–199. doi:10.1007/s10577-011-9269-5. PubMed: 22215485.
  49. Ramos AM, Crooijmans RP, Affara NA, Amaral AJ, Archibald AL et al. (2009) Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLOS ONE* 4: e6524. doi:10.1371/journal.pone.0006524. PubMed: 19654876.



**Figure S1. Frequency distributions of minor allele frequencies (MAF).** MAF were calculated for the 47,155 SNPs of the porcine 60K chip conserved for the TDT analysis.

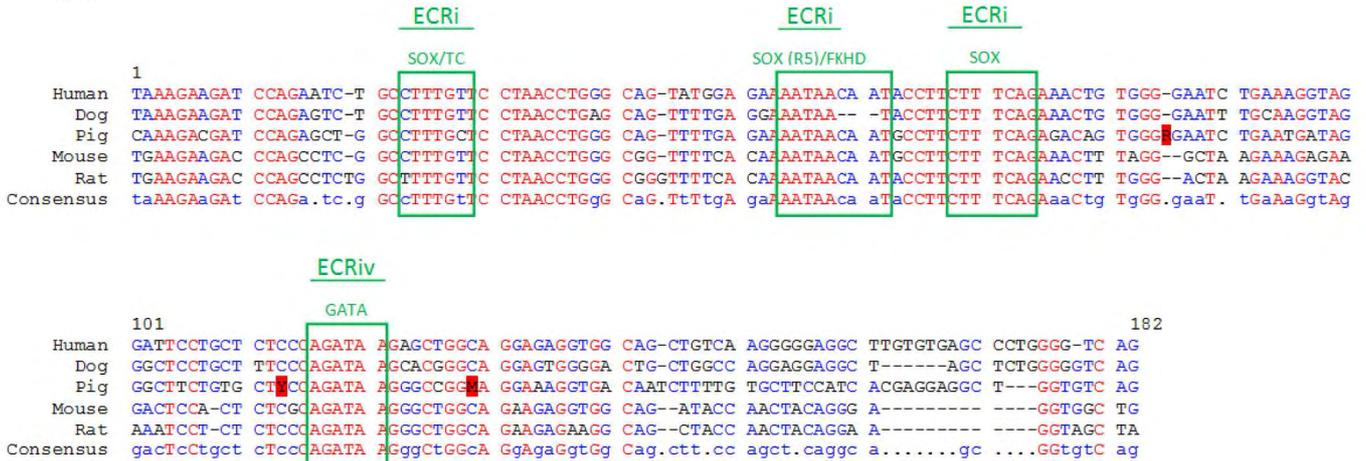


**Figure S2. Cumulative frequency of SNP as a function of the heterozygous parents frequency.** Heterozygous frequency was estimated for each SNP used for TDT analysis.

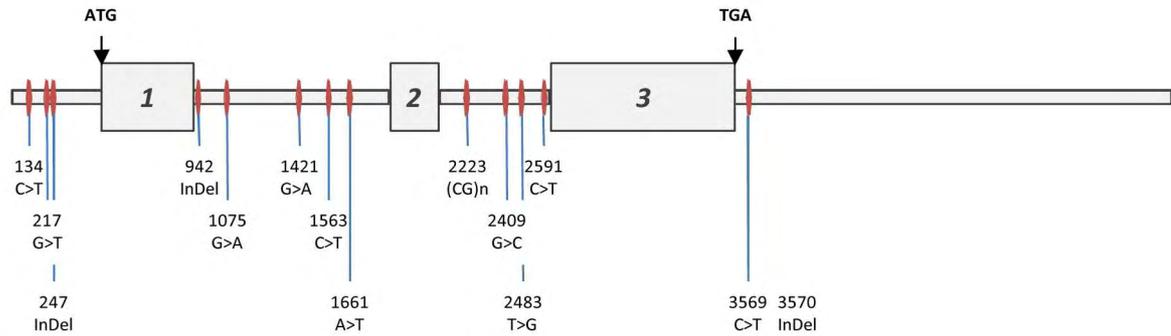


**Figure S3. Estimation of the linkage disequilibrium among the founders of the design.** (A) Cumulative frequency of SNP as a function of the highest  $r^2$  values (divided in classes of  $r^2_{max}$  values of length 0.01) for each core SNP used for TDT analysis and the SNP presenting the highest LD ( $r^2_{max}$ ) with this SNP. (B)  $r^2_{max}$  value (divided in classes of  $r^2_{max}$  values of length 0.01) expressed as a function of the corresponding mean distance (bases) between each core SNP and the SNP presenting the highest LD ( $r^2_{max}$ ) with this SNP.

**A**



**B**



**Figure S4. Localization of the detected polymorphisms in *TESCO* and *SOX9*.** (A) Localization of the 3 polymorphisms identified in the core region of *TESCO* element. Sequence comparison of the 180bp ECR shared by mammals. Polymorphisms identified by sequencing are highlighted in red. The boxed sequences indicate conserved putative transcription factor binding sites and the *SOX* site R5 (Bagheri-Fam et al. (2012)). (B) Position in *SOX9* of the different polymorphisms identified by sequencing. Representation corresponds to the Sscrofa10.2:12:9029179:9033579 interval (in reverse-complement). Positions were defined from position 9033579 (=1) on the pig draft sequence.

Primer name	Stand	Primers (5'-> 3')	Product size (bp)	Annealing temperature (°C)	Elongation time (sec)	number of cycles
Sox9_1	Forward	CAAGAAGAGCTGGGCACAC	830	64	30	38
	Reverse	CCTACAGCCCCTTCAACCTC				
Sox9_2	Forward	TGGTGGTCTGTGTAGTCGT	705	60	30	35
	Reverse	GCACAGCCCTGGTTGATT				
Sox9_3	Forward	GCCTGCTGCTTGGACATC	1214	65	60	45
	Reverse	CACAGAGCCCCTCTCTCTTTT				
Sox9_4	Forward	CTCCGCCTCCTCCACGA	832	60	30	36
	Reverse	GCTCCTTGACAGATTAAGTGTTT				
Sox9_5	Forward	CCCTCTCCCCTTCTCTTCT	540	60	30	36
	Reverse	TGACAAAGTTCTTGATTGCTC				
Sox9_6	Forward	TTCTCCTTCTCCGGCTTC	924	60	30	35
	Reverse	CTCCTCTCCAATTCGCCTTT				
Sox9_7	Forward	TGCCTCTTCGCTCTCCTTC	848	60	30	38
	Reverse	ACACACACACACGCACAC				
TESCO_1	Forward	GGCAATCATCCCTGCATAAC	706	63	30	36
	Reverse	CCAGGAGTCTGAGGTTTAGGG				
TESCO_2	Forward	GCGCGTTGTGTATTGGAAA	690	63	30	36
	Reverse	TTAACCCCTGAGCCACCA				
TESCO_3	Forward	GGCTATTGCACTGGGCTCT	642	63	45	35
	Forward	CCAGGAGTCTGAGGTTTAGGG				
TESCO_4	Forward	ACCCATCTCAAGGGCAACT	862	58	45	35
	Reverse	AAAAGGGAAAACCCAACAGG				

**Table S1. List of primer pairs used for the amplifications of the *SOX9* and *TESCO* regions. Product sizes and amplification conditions are reported.**



## I.A. Résumé de l'article

### *Primo-localisation*

Une analyse TDT a été réalisée à partir d'un dispositif familial (38 familles) et a permis d'identifier une région candidate sur le chromosome 12. Ces analyses TDT ont été réalisées avec les marqueurs de la puce porcine remplissant les critères de qualité (MAF > 0,05, *Call Rate* > 0,05 et équilibre de Hardy-Weinberg < 0,001). 47155 SNP ont ainsi été conservés et neuf SNP ont permis d'obtenir un signal significatif ( $p < 1,33 \cdot 10^{-5}$ ) sur le bras court du chromosome 12. En tenant compte des marqueurs les plus significatifs ( $p < 10^{-6}$ ), la région est localisée entre les positions 8,59 Mb (ASGA0052953) et 9,47 Mb (ASGA0097657), et le marqueur le plus significatif est en position 8,881 Mb (M1GA0024789). Dans cet intervalle un seul gène est annoté, le gène *SOX9*.

### *Recherche de CNV*

Des duplications dans la région chez l'homme ont été mises en évidence chez des individus XX atteints de DSD testiculaire ou d'inversion sexuelle (Benko et al., 2009; Cox et al., 2011; Vetro et al., 2011). De plus, des délétions dans cet intervalle ont également été identifiées chez des patients XY atteints de DSD ovarien (White et al., 2011). La recherche de CNV dans la région a donc été réalisée avec les données de génotypage de la puce porcine et de l'outil PennCNV. Aucune variation de 150kb ou plus n'a été observée en utilisant des fenêtres de trois SNP consécutifs, ce qui n'exclut pas la possibilité de CNV de taille inférieure à 150 kb.

### *Analyse haplotypique*

Afin d'affiner la localisation de la mutation causale, 35 marqueurs localisés entre les positions 8,566 et 9,604 Mb ont été sélectionnés pour caractériser les haplotypes ségrégeant chez les descendants atteints. Ces marqueurs ont été choisis pour contenir et encadrer les marqueurs significatifs des analyses TDT. Cette étude n'a pas permis d'identifier de segment chromosomique commun à tous les haplotypes portés par les animaux atteints. Cependant, un des 43 haplotypes est majoritairement représenté (« haplotype « 8 »), avec une fréquence de 50,7%. Deux autres haplotypes sont également surreprésentés, avec une fréquence de plus de 6% chacun (haplotypes « 2 » et « 17 »).



### *Analyse de la séquence codante et de l'élément fonctionnel*

Le gène *SOX9* (trois exons, deux introns et les régions 5' et 3' UTR) et son enhancer testicule-spécifique TESCO (1,4 kb de séquence) ont été séquencés chez des animaux atteints et non atteints pour rechercher des mutations associées à la maladie. Cinq animaux atteints ont été choisis sur base de leur haplotype dans la région : quatre animaux atteints sont porteurs d'une ou deux des haplotypiques les plus fréquentes (« 8 », « 2 » et/ou « 17 »), et un animal atteint est porteur d'un haplotype rare. Cinq animaux LW non malades, collectés dans le cadre d'un autre programme, mais issus de la même population ont également été séquencés. Quatorze mutations (dix substitutions et quatre indels de un à dix-huit nucléotides) ont été identifiées dans la séquence de *SOX9*. Cependant, aucun de ces polymorphismes n'est localisé dans un exon (9 polymorphismes concernent les introns\_1 et \_2, trois polymorphismes sont localisés en 5'UTR et deux en 3'UTR). Un indel et quatorze substitutions ont été caractérisés dans la séquence de TESCO, mais ne sont pas localisés dans les sites de fixation conservés entre les espèces (Bagheri-Fam et al., 2010). Les vingt-et-un polymorphismes identifiés dans la séquence de *SOX9* et TESCO sont également répartis entre les animaux atteints et les animaux sains, aucune de ces variations n'est donc associée au statut à la maladie : les polymorphismes sont répartis en cinq haplotypes pour *SOX9*, et trois pour TESCO, mais aucun segment haplotypique n'est spécifiquement partagé que par les atteints.

### **I.B. Conclusion**

Bien que plusieurs marqueurs SNP de la région SSC12 soient significativement associés à la maladie, et que le gène localisé dans cette région soit un très bon gène candidat fonctionnel, aucune mutation ou variation associée à la maladie n'a pu être identifiée dans la région. L'étude haplotypique a montré qu'un haplotype en particulier était surreprésenté, mais aucun segment chromosomique n'est commun à tous les individus atteints. Deux hypothèses sont proposées au vu de ces résultats :

1) L'intersexualité chez le porc est une maladie polygénique, une variation dans la région de *SOX9* est responsable du déterminisme de la maladie pour une partie des atteints (ceux porteurs de l'haplotype surreprésenté). Pour les autres animaux atteints, une ou plusieurs autres régions dans le génome sont impliquées. Dans ce cas, il est nécessaire d'augmenter la taille du dispositif pour gagner de la puissance et pour pouvoir identifier d'autres régions chromosomiques en distorsion de ségrégation.



Le test TDT est peu puissant car il ne prend en compte que les parents hétérozygotes. C'est pourquoi, dans le cas de maladie polygénique, le nombre d'animaux doit être élevé.

2) La maladie est monogénique, le gène *SOX9* est impliqué dans le déterminisme de la maladie chez tous les atteints. Dans ce cas, il est possible qu'une ou plusieurs mutations associées à *SOX9* soient responsables de la maladie. Chez l'homme et chez la souris, des variants situés jusqu'à 1Mb de *SOX9* sont responsables de modifications dans la régulation du gène, induisant des désordres du développement sexuel. La ou les mutations chez le porc peuvent donc être localisées dans tout l'intervalle mis en évidence, mais les marqueurs de la puce pourraient ne pas être suffisamment en DL avec la mutation causale. Pour confirmer ou infirmer cette seconde hypothèse, il est nécessaire de densifier en marqueurs la région candidate à l'aide de marqueurs plus discriminants.

A l'issue des premiers travaux rapportés dans cet article nous avons décidé de réaliser une analyse exhaustive de la région afin de rechercher une ou plusieurs mutations dans l'intervalle, et de caractériser la ou les haplotypes présents chez les animaux malades. Nous avons donc choisi de reséquencer entièrement l'intervalle. D'une part, nous estimions qu'un séquençage exhaustif et de qualité pouvait permettre l'identification de mutation(s) causale(s) ou de CNV, par comparaison des séquences d'individus malades et d'individus sains. D'autre part, les données de séquences obtenues par cette approche permettaient à minima d'identifier des polymorphismes complémentaires en fort DL avec la mutation causale.

La deuxième partie de chapitre porte sur les travaux de reséquençage haut débit.



## II. Reséquençage complet

Les analyses de données de reséquençage sont grandement facilitées par la disponibilité d'une séquence de référence de l'espèce étudiée. Le premier assemblage du génome porcin a été publié en septembre 2009. Ce draft (Sscrofa9) avait été obtenu par assemblage de clones de BAC (couverture de 4 X environ). Au moment des travaux de reséquençage des animaux atteints IS que nous avons réalisés, cette première version était en cours de révision. Depuis 2011, une nouvelle version 10.2 (Sscrofa10.2) est disponible. Elle correspond à une version issue de l'intégration de données de WGS (*Whole Genome Sequence*) pour obtenir une couverture 30 X. Dans le cadre de cette thèse les travaux de reséquençage ont donc été débutés avec la version V9 puis repris avec la nouvelle version V10.2.

### II.A. Stratégie de séquençage utilisée

L'objectif du séquençage est 1) de fournir une couverture optimale de la région candidate (3 Mb environ) et 2) de fournir une profondeur suffisante pour pouvoir identifier des variants de type SNP ou CNV. Deux alternatives se présentaient pour le séquençage : le reséquençage ciblé de la région après capture de séquence ou le reséquençage complet du génome (et analyse de la région candidate). La capture de séquence permet un enrichissement spécifique d'une librairie d'ADN génomique en fragments d'ADN issus de régions ciblées. Cette étape de capture, au moment de ces travaux, devait être réalisée par hybridation d'ADN génomique totale sur une puce comprenant plusieurs milliers d'oligonucléotides choisis spécifiquement dans la région cible. L'intérêt de cette approche est qu'elle permet d'obtenir une profondeur en séquences importante et optimale pour l'identification de SNP et de CNV et le reséquençage de plusieurs individus différents simultanément. A contrario, les oligonucléotides de la puce de capture sont définis à partir de la séquence de référence, dans notre cas la séquence de référence V9. Afin d'être fiable, cette stratégie nécessite par conséquent une séquence de référence de bonne qualité. Le reséquençage haut débit permet quant à lui de séquencer tout le génome, indépendamment de l'assemblage de référence, mais la profondeur peut être faible dans certaines régions (régions riches en GC par exemple). Au vu de la qualité du draft porcin au moment des analyses et de l'annonce de la publication d'une nouvelle version V10, le reséquençage haut débit a été choisi pour l'étude de la région. Bien que la profondeur attendue soit significativement plus faible avec la technique de reséquençage tout génome que celle attendue avec un enrichissement local, nous avons estimé plus sûr, afin d'être exhaustif, de ne pas se fier à la version V9.

IS_35	G	G	A	G	G	C	C	T	G	G	T	T	T	G	G	G	G	A	C	T	T	C	A	A	C	T	T	T	G	G	A	G	G	A	A	G	C	C	C	A	C	A	A	A	G	T	A	A	T	A	T	G	T
	G	G	A	G	G	C	C	T	G	G	T	T	T	G	G	G	G	A	C	T	T	C	A	A	C	T	T	T	G	G	A	G	G	A	A	G	C	C	C	G	T	A	G	A	T	C	G	G	C	G	A	G	T
IS_49	A	G	A	G	G	C	C	T	G	G	T	T	T	G	G	G	G	A	C	T	T	C	A	A	C	T	T	T	G	G	A	G	G	A	A	G	C	C	C	G	T	A	G	A	T	C	G	G	C	G	A	G	T
	G	G	A	G	G	C	C	T	G	G	T	T	T	G	G	G	G	A	C	T	T	C	A	A	C	T	T	T	G	G	A	G	G	A	A	G	C	C	C	G	T	A	G	A	T	C	G	G	C	G	A	G	T
Tem_29	A	G	G	G	A	C	C	T	G	T	C	C	G	T	G	T	A	G	T	C	C	T	A	A	C	G	C	C	A	A	A	G	A	A	G	T	T	T	C	A	C	G	A	G	G	T	A	A	T	A	T	A	C
	A	G	G	G	A	C	C	T	G	T	C	C	T	C	G	T	A	G	C	C	T	A	A	C	T	T	T	G	A	A	G	A	A	G	T	T	C	A	C	G	A	G	G	T	A	A	T	A	T	A	C		
Tem_44	A	G	G	G	A	C	C	T	G	T	C	C	G	T	G	T	A	G	?	C	C	T	A	G	C	G	C	C	A	A	A	G	A	A	G	T	T	C	A	C	G	A	G	G	T	A	A	T	A	T	A	C	
	A	G	G	G	A	C	C	T	G	T	C	T	T	C	G	T	A	G	?	C	C	T	A	A	C	G	T	C	G	A	A	G	A	A	G	T	C	C	C	G	T	A	G	A	T	C	A	G	C	G	A	G	T

Figure 89 : Haplotypes dans la région IS pour les deux animaux atteints et les deux témoins sélectionnés pour le reséquençage

De plus, dans l'hypothèse où une autre région serait impliquée dans la maladie, les données en tout génome pouvaient être utilisées pour d'autres régions candidates.

L'objectif est de comparer des animaux atteints d'intersexualité et des animaux sains. Quatre animaux ont été sélectionnés (deux sains et deux atteints) sur base de leur haplotype dans la région candidate à l'intersexualité ou IS (Figure 89). Les deux animaux intersexués (n°35 et 49) ont été choisis car ils présentaient une forte homozygotie pour l'haplotype le plus représenté chez tous les animaux intersexués (« haplotype 8 »).

Les deux animaux témoins (n°29 et 44) ont été choisis comme présentant des haplotypes les plus distincts de l'haplotype 8 dans la zone IS.

Ce reséquençage a été réalisé à l'aide du système HiSeq d'Illumina, disponible sur la plateforme génomique GeT-PlaGe de Toulouse. Ce séquenceur nouvelle-génération permet d'obtenir 100 Gb de séquence par flow-cell, chaque flow-cell étant composé de huit lignes. Une ligne offre donc une capacité de séquençage de 12,5 Gb en moyenne. Le génome porcine étant composé de près de 3 Gb, l'analyse d'un animal sur une ligne permet d'obtenir une couverture de 4X environ. Cette couverture a donc été choisie dans un premier temps pour nos analyses, nous permettant ainsi de disposer en moyenne de 8X de séquences issues d'animaux contrôles et 8X de séquences issues d'animaux malades. Les études chez les autres espèces comme l'homme et le bovin suggèrent qu'une couverture minimum de 15X voire 30X est nécessaire pour déterminer de façon fiable les génotypes des individus séquencés (Mertes et al., 2011). A terme nous envisageons de compléter ces premières données par des travaux de re-séquençage complémentaires : le choix du nombre de lignes de reséquençage dédié à chaque animal sera décidé au vu des premiers résultats.



Figure 90 : Séquençage HiSeq des fragments d'ADN

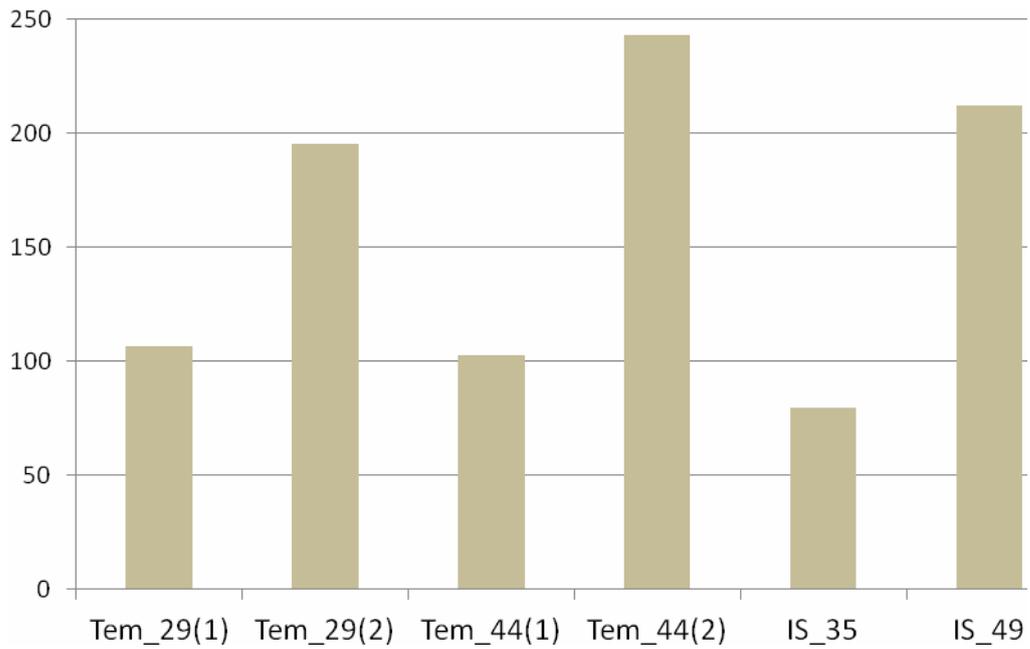


Figure 91 : Nombre total de séquences obtenues par animal

## II.B. Résultats

### II.B.1. Nombre de séquences obtenues

Le principe du séquençage HiSeq est de séquencer les deux extrémités, sur une longueur de 100 bases, de fragments de 500 bases en moyenne (Figure 90). Au cours d'une première réaction, les premières 100 bases sont séquencées sur le brin sens (lectures 1) ; dans un second temps, les extrémités comprises entre les positions 400 et 500 sont obtenues par le séquençage de l'extrémité du brin anti-sens (lectures 2). L'appariement lecture1/lecture2 (séparées de 300 bases environ) est pris en compte au moment de l'analyse.

Lors du reséquençage des animaux choisis, deux lignes n'ont pas permis d'obtenir suffisamment de lectures (24 millions de séquences environ, contre 120 millions attendues), et deux lignes ont été séquencées dans un sens seulement, dû à l'arrêt du séquençage en raison d'un incident technique. Afin de compléter les données de séquençage, huit lignes au total ont été séquencées : les deux lignes dont le nombre de lectures était trop faible n'ont pas été conservées pour la suite des analyses, et les deux lignes pour lesquelles le séquençage n'a permis d'obtenir que les lectures-1 ont été conservées.

Le nombre total de lectures brutes obtenues par animal et ligne retenus est présenté en Figure 91. Ainsi, 80 à 350 millions de lectures ont été obtenues par animal. Les deux lignes séquencées dans un seul sens concernent les deux témoins (Tem\_44(1) et Tem\_29(1)), qui ont donc été séquencés deux fois. Les différences entre les quatre animaux sont importantes, d'une part car les animaux intersexués n'ont été séquencés qu'une seule fois, alors que les animaux atteints de hernie inguino-scrotale l'ont été au cours d'une ligne et demi, d'autre part car le nombre de lectures obtenues entre deux lignes présente de grandes disparités.

Pour l'animal IS\_35, 80 millions de lectures sont disponibles alors que l'animal IS\_49 présente plus de 200 millions de lectures. Le nombre de lectures pour les deux témoins est comparable (300 millions pour Tem\_29, et 350 millions pour Tem\_44).

Une ligne offre une capacité de séquençage de 12,5Gb (4X), soit 125 millions de lectures de 100pb. Les données obtenues ici varient de 80 à 350 millions de lectures, correspondant à une couverture de 2,6X (IS\_35) jusqu'à 11,6X (Tem\_44). Les résultats obtenus sont au dessus des estimations (couverture IS\_49 : 7 X et Tem\_29 : 10 X). Pour l'animal IS\_35, le nombre de lecture (et couverture) attendue n'a pas été obtenue, et ce résultat devra être pris en compte pour la suite des analyses.



L'objectif à terme est de réaliser un alignement des séquences obtenues sur le génome de référence porcine, puis de rechercher des polymorphismes discriminant les individus atteints et non atteints. Cependant une étape préalable d'estimation de la qualité globale du séquençage, et de celle de chacune des lectures doit être réalisée afin de ne conserver que les lectures les plus fiables.

### **II.B.2. Contrôle qualité**

Un premier contrôle a été réalisé afin d'estimer la qualité globale de chaque ligne de séquençage, et de savoir si les données pouvaient être utilisées pour la suite des analyses. Ces contrôles se sont avérés satisfaisants pour chacune des six lignes de séquençage, et ont donc été conservés.

Les séquences ont ensuite été alignées sur des génomes bactérien (E-coli), viral et de levure pour vérifier que l'ADN séquencé n'a pas été contaminé par un microorganisme. Une séquence sur 100.000 en moyenne s'aligne sur un des trois génomes cités (dont 99% environ sur la séquence virale, 0,003% sur la séquence bactérienne et 0,001% sur la levure). Ces valeurs ne sont pas aberrantes et signifient que les produits séquencés ne sont pas contaminés.

Ces premiers contrôles sont réalisés de façon systématique, quelque soit la matrice séquencée et quelque soit le domaine d'exploitation de ces données de séquence. Dans le cadre de l'étude que nous réalisons, l'objectif est de rechercher toutes les différences nucléotidiques différenciant les individus IS des individus contrôles. Les contrôles additionnels que nous avons fait sont donc destinés à ne conserver que les paires de lectures de bonne qualité (sans bases indéterminées), appariée en un seul endroit du génome (les deux lectures étant alignées sur un brin (lecture 1) et le brin complémentaire (lecture 2) à une distance de 300 pb environ). Les lectures ne répondant pas à ces critères seront éliminées.

La qualité de l'assemblage de référence est donc importante, car des fragments de séquences absents ou redondants dans l'assemblage de référence biaisent l'alignement, et les lectures sont éliminées à tort.

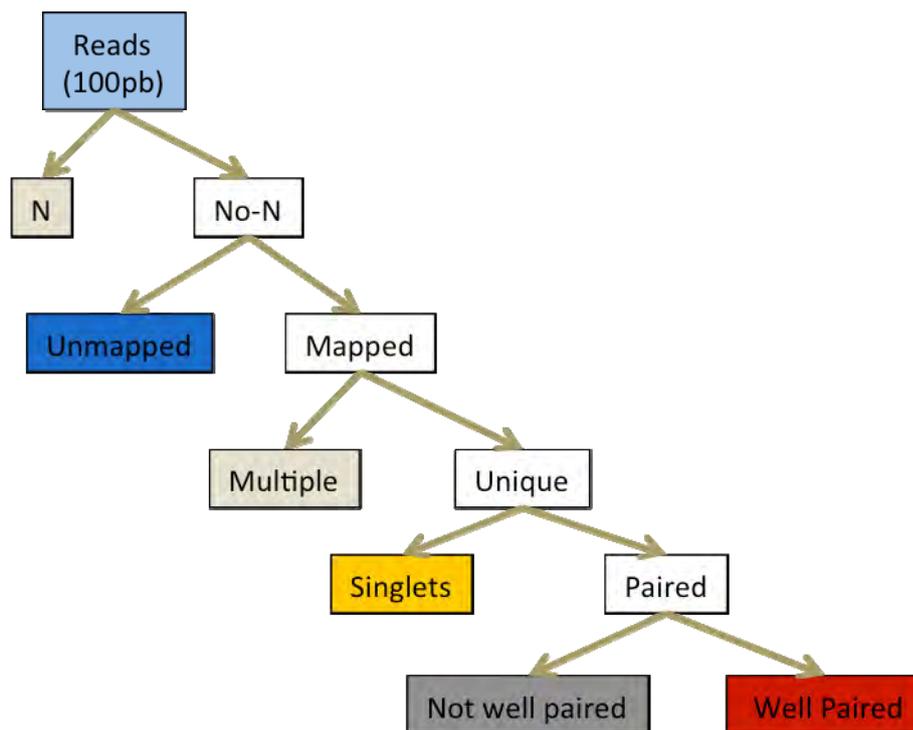


Figure 92 : Démarche générale de l'alignement des séquences issues du reséquençage HiSeq

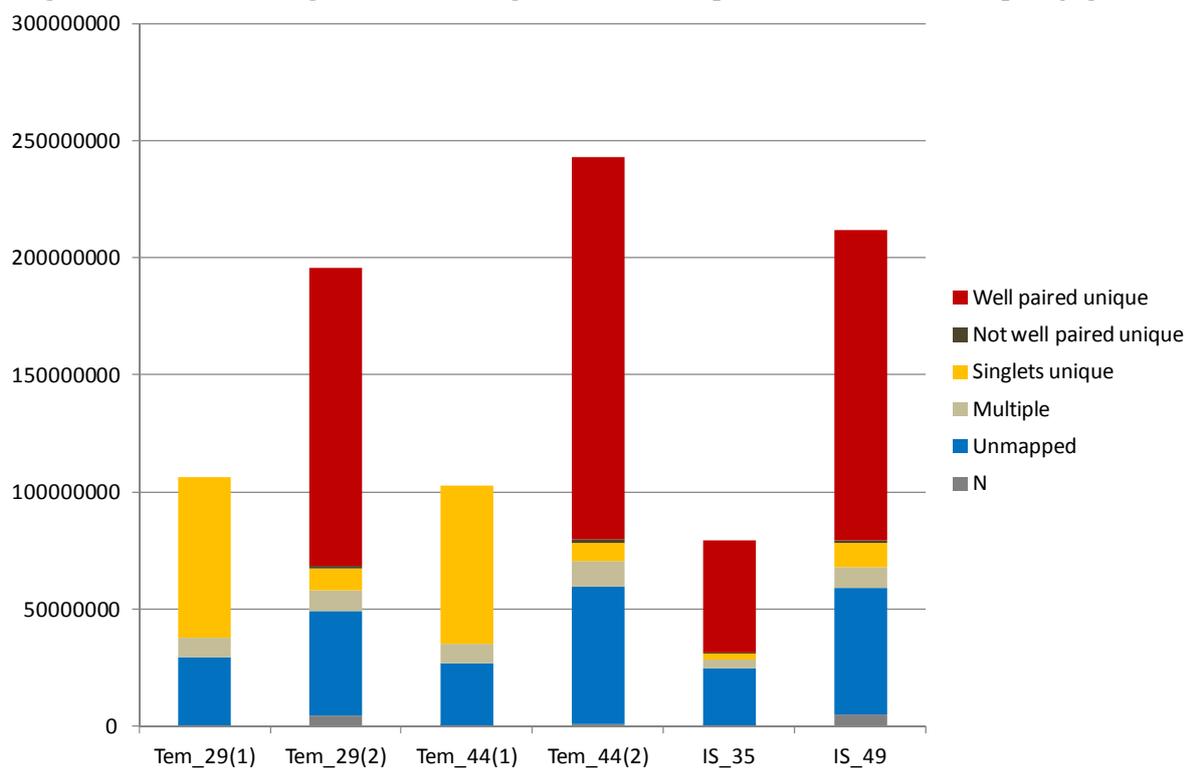


Figure 93 : Résultats d'alignement des lectures sur le draft 9

### II.B.3. Alignement des lectures

Les étapes successives de la sélection des lectures « sûres » qui permettront la recherche de polymorphismes sont présentées en Figure 92. Au moment du reséquençage, le draft porcine disponible était la version 9. Les premières analyses ont donc été réalisées sur ce draft. La Figure 93 présente les résultats d'alignement obtenus pour les six lignes de séquençage réalisées, selon les différentes étapes de la Figure 92.

**Biais d'attribution des bases :** Chaque lecture est composée de 100 bases, mais il est possible que l'intensité pour certaines bases soit trop faible et qu'elle n'ait pas permis l'attribution d'un nucléotide. Les lectures possédant au moins un « N » (attribué lorsque le nucléotide n'est pas déterminé) sont éliminées, la qualité de la lecture étant jugée trop faible pour le reste des analyses. Le nombre de séquences de N représente de 0,3% à 2% du nombre de séquences totales que nous avons obtenues.

**Alignement sur le génome porcine :** Les séquences ont ensuite été alignées sur le génome de référence (draft 9 au moment du reséquençage). Le pourcentage d'alignement varie de 69% (IS\_35) à 75% (Tem\_29), pour une moyenne de 73% d'alignement. Ces données ne sont pas très satisfaisantes mais elles sont cohérentes avec les données d'alignements obtenues pour d'autres projets de reséquençage réalisés sur la plateforme. Ces résultats reflètent que près d'un quart du génome porcine n'est pas représenté dans la séquence de référence publiée dans *Ensembl*. Cet état est un fait non négligeable pour une recherche exhaustive de polymorphismes.

**Alignement unique :** Si une lecture de 100 bases est une séquence correspondant à un gène membre d'une famille de gènes, il est quelquefois difficile de déterminer à quelle copie de gène correspond cette lecture. Le risque est donc d'aligner au sein du même contig (en une ou plusieurs positions dans le génome) des lectures, présentant quelques bases de différences les unes par rapport aux autres, qui correspondent en réalité à différentes localisations. Afin de limiter les risques d'identification de polymorphismes erronés, les séquences présentant plusieurs sites d'appariement ont donc été éliminées. Le pourcentage d'alignement unique varie de 64% (IS\_35) à 71% (Tem\_44(2)). La moyenne est de 67% d'alignement unique.

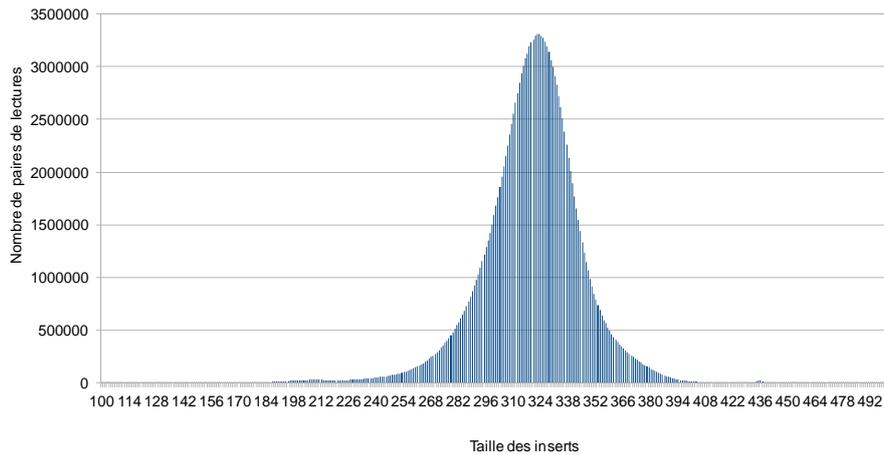


Figure 94 : taille des inserts entre les lectures appariées

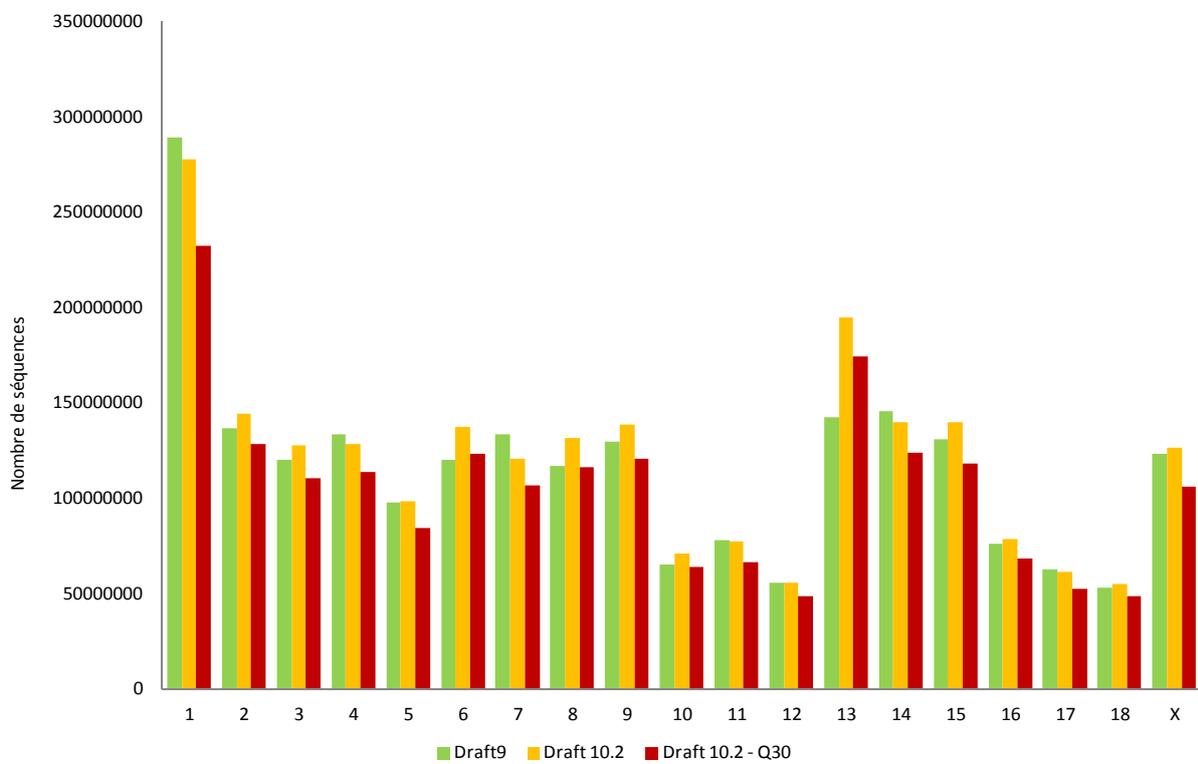


Figure 95 : couverture des lectures par chromosomes sur le draft 9, le draft 10.2 et après contrôle qualité Q30 sur le draft 10.2

**Appariement des lectures :** À chaque lecture-1 correspond une lecture-2 en sens opposé séparée de 300 pb environ. Les lectures ont été appariées, et la Figure 94 présente la taille des inserts pour l'ensemble des lectures appariées. La taille moyenne de 322, correspond à la taille attendue moyenne entre les lectures. La distribution de la taille des inserts suit une loi normale. Au final, le pourcentage de lectures correctement appariées est de 60% (IS\_35) à 67% (Tem\_44(2)) des lectures initialement obtenues grâce au séquenceur. La moyenne est de 64% pour les quatre échantillons, sachant que les lignes Tem\_29(1) et Tem\_544(1) n'ont pas été appariés.

**Contrôle qualité de chaque lecture :** Enfin, pour chaque lecture, un contrôle de la qualité d'attribution des bases a été réalisé. Il s'agit de ne conserver que les lectures pour lesquelles l'attribution des bases est sûre. Le seuil q30 a été choisi, et signifie que l'attribution des bases est sûre à 99,9%.

Quelques temps après l'obtention des données de séquençage, une nouvelle version corrigée de l'assemblage porcin (v10.2) a été disponible. L'alignement a donc été réalisé sur cette nouvelle version. La couverture par chromosome a été comparée entre le draft 9 et le draft 10.2 (Figure 95), avant et après sélection des lectures de bonne qualité (Q30). Le contrôle qualité Q30 permet de garder 85% des lectures environ. La couverture selon le draft 10.2 par rapport au draft 9 est supérieure pour la quasi-totalité des chromosomes, exceptés pour SSC1, 7, 14 et 17.

Concernant le chromosome 12, sur lequel se trouve la région candidate à l'intersexualité, la taille des séquences des draft 9 et 10.2 est très comparable. La sélection des lectures q30 permet de conserver 87% des lectures.

Le reséquençage complet pour les quatre animaux choisis (deux intersexués et deux contrôles) a permis d'obtenir la profondeur attendue, voire deux fois supérieure pour trois animaux sur quatre. Cependant, un quart des lectures sont perdues. Parmi les différents critères de filtres utilisés, le critère ayant le plus fort impact est l'appariement des lectures sur le génome de référence. Nous avons cherché à estimer la qualité du draft porcin dans la région candidate étudiée afin de voir si en améliorant la qualité de la séquence de référence dans cet intervalle nous pourrions réduire ce taux de 25% de « déchets ».

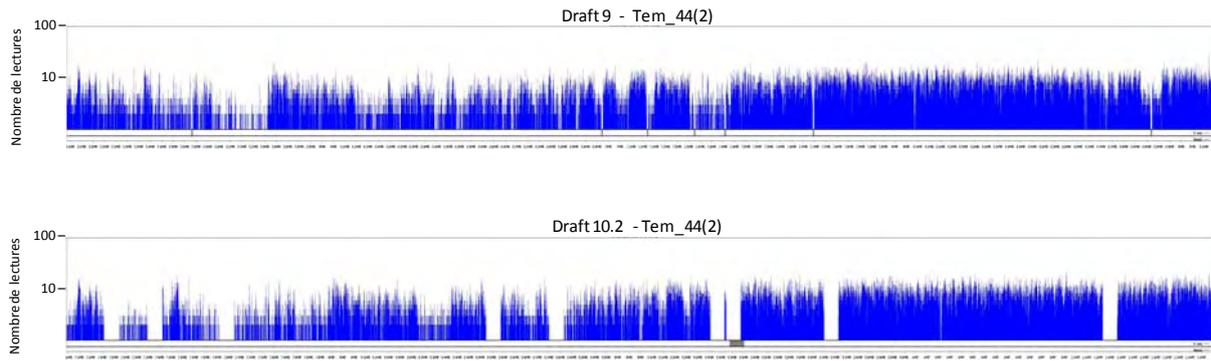


Figure 96 : Alignement des lectures de l'individu Tem\_44(2) sur les drafts 9 et 10.2. Chaque trait bleu correspond à un nucléotide, selon sa position sur l'assemblage en abscisse, et sa couverture en ordonnée.

## II.C. Etude de la couverture dans la région candidate

Notre objectif étant d'identifier des polymorphismes sur le chromosome 12, entre les positions 6,53 Mb et 7,19 Mb (draft9) (et les positions 8,59 Mb et 9,47 Mb (draft 10.2)) nous avons entrepris une analyse fine de la qualité des données de reséquençage dans cet intervalle. Les deux critères majeurs pour cette étude sont de disposer d'une couverture optimale dans la région candidate afin de réaliser une recherche la plus exhaustive possible des variants existants et de disposer d'une séquence de référence de très bonne qualité.

Afin d'étudier la couverture et la profondeur dans la zone, une région plus large a été sélectionnée pour bien couvrir l'ensemble de l'intervalle. La région étudiée est localisée entre les positions 5 Mb à 9 Mb sur le draft 9, puis 7 Mb et 12 Mb sur le draft 10.2. Les lectures obtenues ont donc été alignées sur les deux versions du draft et un exemple des résultats est présenté en Figure 96 (l'échantillon présenté ici est Tem\_44(2)).

Bien que le draft 10.2 soit une version améliorée du draft 9, de nombreux « trous » (séquences manquantes ou N) sont encore présents dans la séquence de la région d'intérêt du draft 10.2. En effet, la région candidate du draft 10.2 comporte 5 séquences de 50.000 N et 28 séquences de 100 N, correspondant à plus de 250 kb de séquences manquantes. Une reconstruction manuelle d'une séquence de référence à partir des données disponibles dans les bases de données publiques a donc été envisagée.

Parmi les filtres utilisés pour le tri des lectures, les filtres "alignement multiple en plusieurs endroits du génome" ou à contrario "absence d'alignement" entraînent l'élimination de la lecture. La reconstruction d'un draft de l'intervalle 7 à 12 Mb était donc destinée (1) à éliminer d'éventuelles séquences redondantes, et (2) à rechercher des séquences, dans les bases de données, permettant de combler des zones de "N". Une troisième source d'élimination des lectures est un appariement des paires de lectures en désaccord avec la théorie. Par exemple si les lectures-1 et -2 sont distantes de plusieurs kb l'une de l'autre (et non de 300-500 pb) ou sont alignées sur le même brin (et non sur un brin et le brin complémentaire respectivement), les paires de lecture sont éliminées. Pour la reconstruction du draft, une réorganisation des contigs intra-BAC a donc été réalisée en utilisant l'ordre du génome humain en référence. Bien que des inversions puissent exister entre les génomes de l'homme et du porc, aucun remaniement n'a jusqu'à présent été mis en évidence dans cette zone et la probabilité qu'une rupture de synténie survienne à la limite d'un contig intra-BAC est faible.

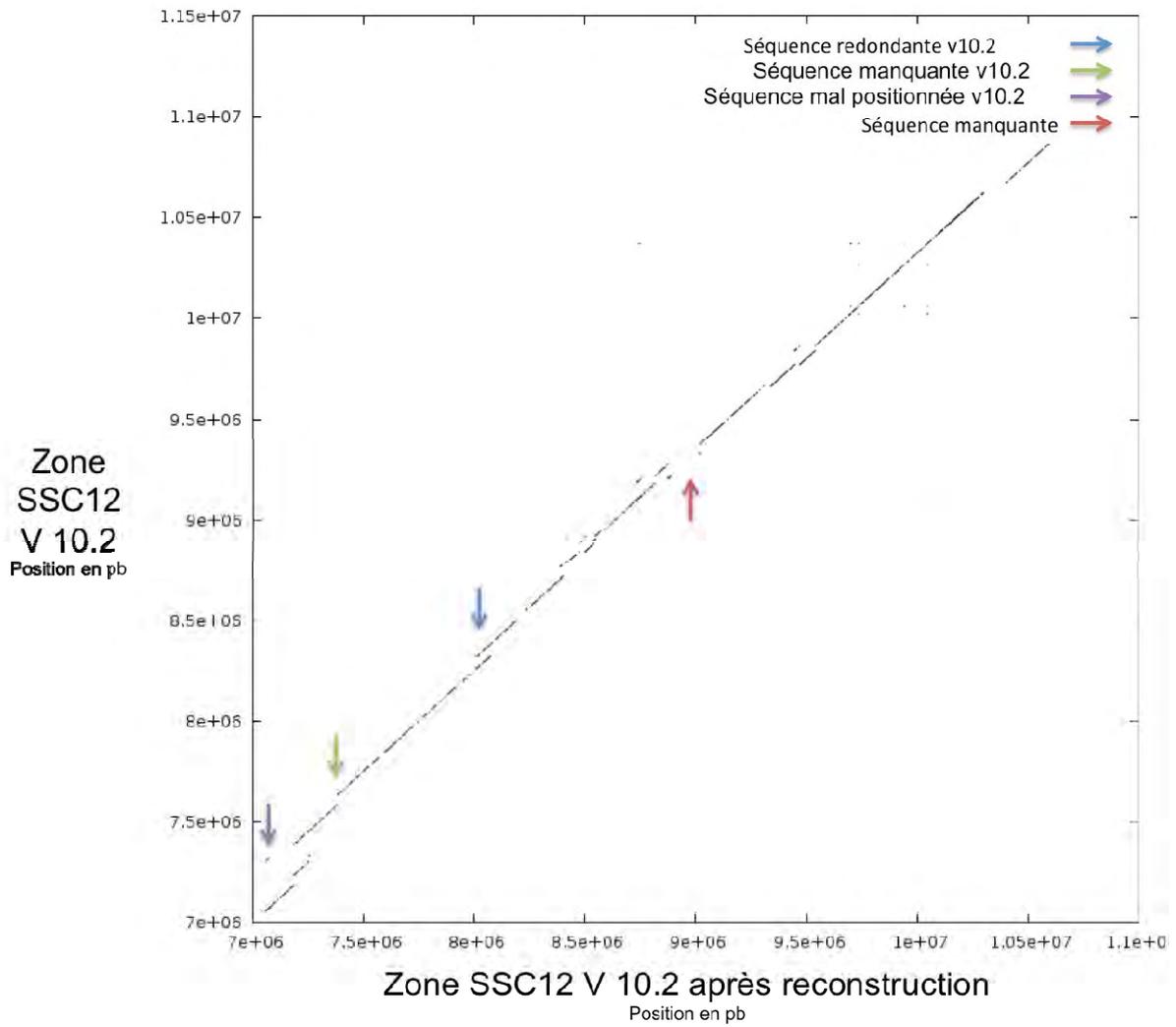


Figure 97 : Alignement de la région candidate de v10.2 après reconstruction sur la v10.2

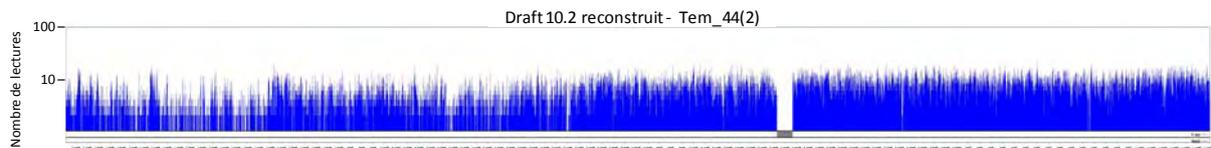


Figure 98 : Alignement des lectures de l'individu Tem\_44(2) sur le draft 10.2 après reconstruction

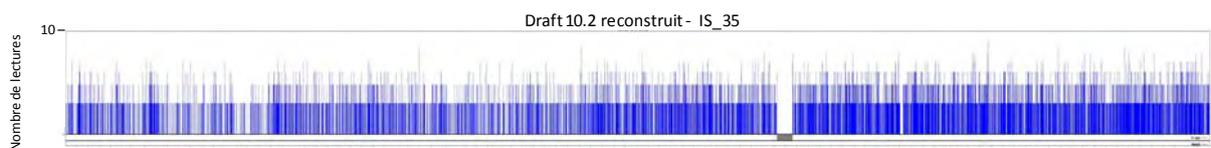


Figure 99 : Alignement des lectures de l'individu IS\_35 sur le draft 10.2 après reconstruction

Après reconstruction de la région, huit blocs de 50.000N et neuf blocs de 100N ont été éliminés, car les séquences de part et d'autre des zones de N étaient en fait chevauchantes sur quelques kb, ou parce que des séquences disponibles dans les bibliothèques porcines ont été ajoutées pour combler les régions manquantes. Trois contigs de BAC ont été identifiés comme redondants, et cinq contigs étaient manquants dans la séquence de la V10.2. Cela a permis de retirer plus de 400 kb de séquences de N et 190 kb de séquences redondantes, et de compléter la région avec 330 kb de séquences initialement manquantes. La Figure 97 présente une comparaison *DotPlot* de la nouvelle séquence de référence avec la séquence de la version 10.2. Des exemples de redondances et de séquences manquantes sont indiqués par les flèches. Ces résultats indiquent que la qualité du draft dans la région a été fortement améliorée, ce qui a permis de réduire le nombre de lectures éliminées à l'issue du reséquençage.

L'alignement des lectures a été renouvelé sur la nouvelle séquence de référence (Figure 98). Le nombre de séquences obtenues par échantillon a conduit à l'obtention d'une profondeur variable d'un animal à l'autre. Pour l'animal Tem\_44(2), la profondeur varie de une à dix lectures environ (Figure 98) alors que pour l'animal IS\_35, présentant une faible couverture, la profondeur atteint rarement 10 lectures (Figure 99). La majorité de la région est couverte par une seule lecture. Ces résultats sont à prendre en compte dans la recherche de polymorphismes, qui pourraient être biaisée lorsque la profondeur est trop faible.

Le reséquençage complet pour quatre animaux a été concluant bien que la profondeur soit faible dans certaines zones, et l'échec principal est dû à la qualité du draft porcin, qui entraîne la perte de 25% des lectures. Afin de caractériser finement la région d'intérêt, une reconstruction de la séquence dans la zone a été entreprise et a permis d'améliorer la qualité du draft dans la région candidate. A l'issue de cette analyse, une recherche de variants a pu être entreprise dans l'intervalle.

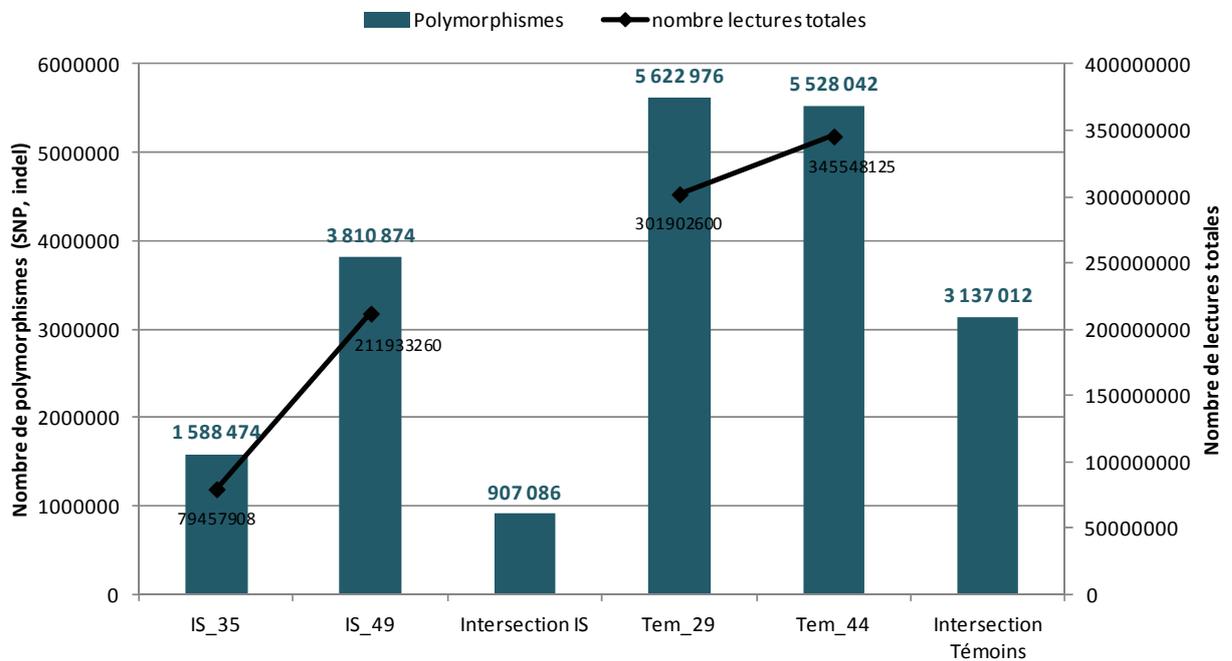


Figure 100 : Nombre de polymorphismes (SNP et indels) et nombre total de lectures chez les quatre animaux séquencés en tout génome. L'intersection correspond aux polymorphismes identifiés chez les deux animaux d'un même groupe

### **III. Recherche de polymorphismes dans la région candidate à l'intersexualité**

Les variants ont été recherchés entre les séquences des individus atteints IS et les témoins, afin d'identifier celui ou ceux qui pourraient être responsables de l'intersexualité. Les polymorphismes peuvent être de type SNP ou indels.

#### **III.A. Recherche exhaustive de polymorphismes**

Les variants entre les séquences obtenues pour chaque animal et la séquence porcine de référence ont été recherchés dans un premier temps animal par animal. Les variants communs aux animaux IS d'une part, et aux témoins d'autre part, ont été ensuite identifiés (intersections). Les animaux séquencés sont de race Large White, alors que les séquences utilisées pour la construction du génome de référence sont essentiellement issues du séquençage d'un animal Duroc. Pour cette recherche aucun filtre sur le nombre de lectures couvrant la séquence de référence n'a été appliqué : si une seule lecture est disponible et diffère de la séquence de l'animal Duroc, le polymorphisme identifié est conservé dans la liste. Afin de couvrir la région candidate, les polymorphismes ont été recherchés entre 8,0 Mb et 9,2 Mb (positions sur la nouvelle séquence de référence). Les résultats obtenus sont présentés dans la Figure 100. Plus de cinq millions de polymorphismes sont identifiés chez les témoins, et 1,5 à 3,8 millions de variations sont présents chez les animaux IS (IS\_35 et IS\_49, respectivement). Le nombre de polymorphismes tient compte du nombre de SNP et du nombre d'insertions et délétions (indels). Cependant le nombre d'indel est bien plus faible que le nombre de SNP. Le pourcentage d'indels par rapport au nombre total de polymorphismes est de 2,7% et 6,47% respectivement pour les IS\_35 et 49. Pour les animaux témoins, le pourcentage est de 7,9% pour Tem\_29 et 8,57% pour Tem\_44. Parmi l'ensemble des polymorphismes identifiés, certains sont communs aux quatre individus séquencés et diffèrent uniquement avec la séquence de référence. Ils sont donc race-spécifiques, et correspondent à 15% de tous les polymorphismes identifiés.

Le nombre de polymorphismes identifiés chez les deux animaux IS est proportionnel au nombre de séquences obtenues initialement. Chez les individus témoins, le nombre de polymorphismes est similaire alors que le nombre de lectures disponible pour la recherche était différent. Ces données montrent que le nombre de lectures obtenu pour les animaux IS

n'est pas suffisant pour identifier tous les polymorphismes, alors que la couverture 10X (Tem\_29) semble convenable.

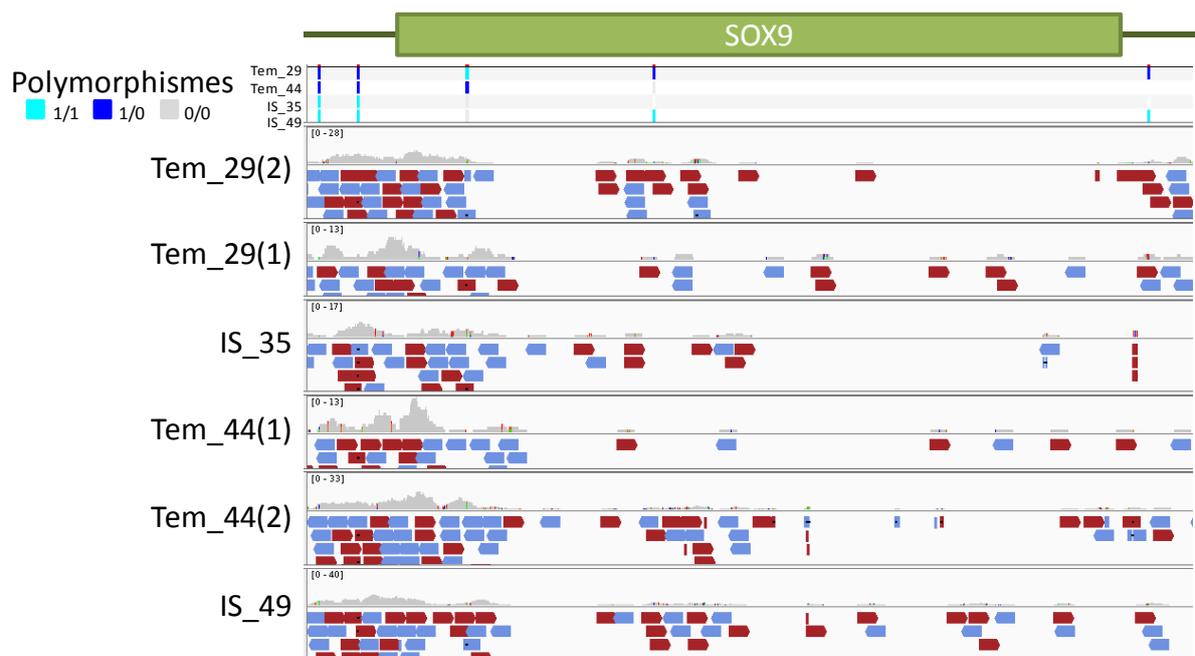


Figure 101: Couverture et polymorphismes dans la séquence de *SOX9*. *SOX9* est indiqué par un rectangle vert

Le nombre de polymorphismes communs à chacun des groupes (groupe des deux individus IS, et groupe des deux animaux témoins) est plus faible que le nombre obtenu pour chacun des animaux, ce qui signifie que les animaux au sein d'un même groupe présentent des différences. Les animaux IS ont été sélectionnés pour être très similaires dans la région IS au vu de leur haplotype, et la faible proportion de polymorphismes communs aux deux animaux peut être expliquée par la profondeur peu importante. Pour le groupe des témoins, les animaux ont été choisis parce qu'ils présentaient des haplotypes différents, et ces différences de variabilité sont retrouvées ici.

### **III.B. Validation des polymorphismes identifiés**

Afin de vérifier les génotypes identifiés, les résultats ont été comparés aux données de séquençage du gène *SOX9* et de *TESCO*, qui avaient été obtenues par séquençage Sanger. Le design d'amorces pour le séquençage Sanger de *SOX9* s'était avéré difficile car la séquence est très riche en GC (jusqu'à 64%). La Figure 101 présente la couverture et les polymorphismes identifiés dans la région *SOX9*. La couverture est très faible, voire absente à certains endroits, et ce pour les quatre animaux. Le biais semble plutôt concerner le contenu en séquence, difficile à séquencer avec la technologie HiSeq. Cependant, les deux polymorphismes identifiés par l'analyse des données de reséquençage avaient bien été identifiés par séquençage Sanger automatisé. Pour *TESCO*, 3 des 15 polymorphismes identifiés ont été retrouvés.

Ces résultats montrent que pour une séquence peu riche en GC, et une couverture et profondeur suffisante, l'identification des polymorphismes par analyse de données HiSeq est fiable.

IS_35	IS_49	Tem_29	Tem_44	
1	1	0	0	
1	0.5	0	0	Dominant
0.5	1	0	0	Dominant
0.5	0.5	0	0	Dominant
1	1	0.5	0.5	Récessif
1	1	0.5	0	Récessif
1	1	0	0.5	Récessif
0	0	1	1	
0	0	1	0.5	Récessif
0	0	0.5	1	Récessif
0	0	0.5	0.5	Récessif
0.5	0.5	1	1	Dominant
0.5	0	1	1	Dominant
0	0.5	1	1	Dominant

Tableau 19 : Règle de décision pour le choix des polymorphismes. Les valeurs correspondent aux fréquences alléliques de l'allèle 1. Les valeurs 1 et 0 signifient que l'animal est homozygote, la valeur 0.5 est attribuée à un animal hétérozygote

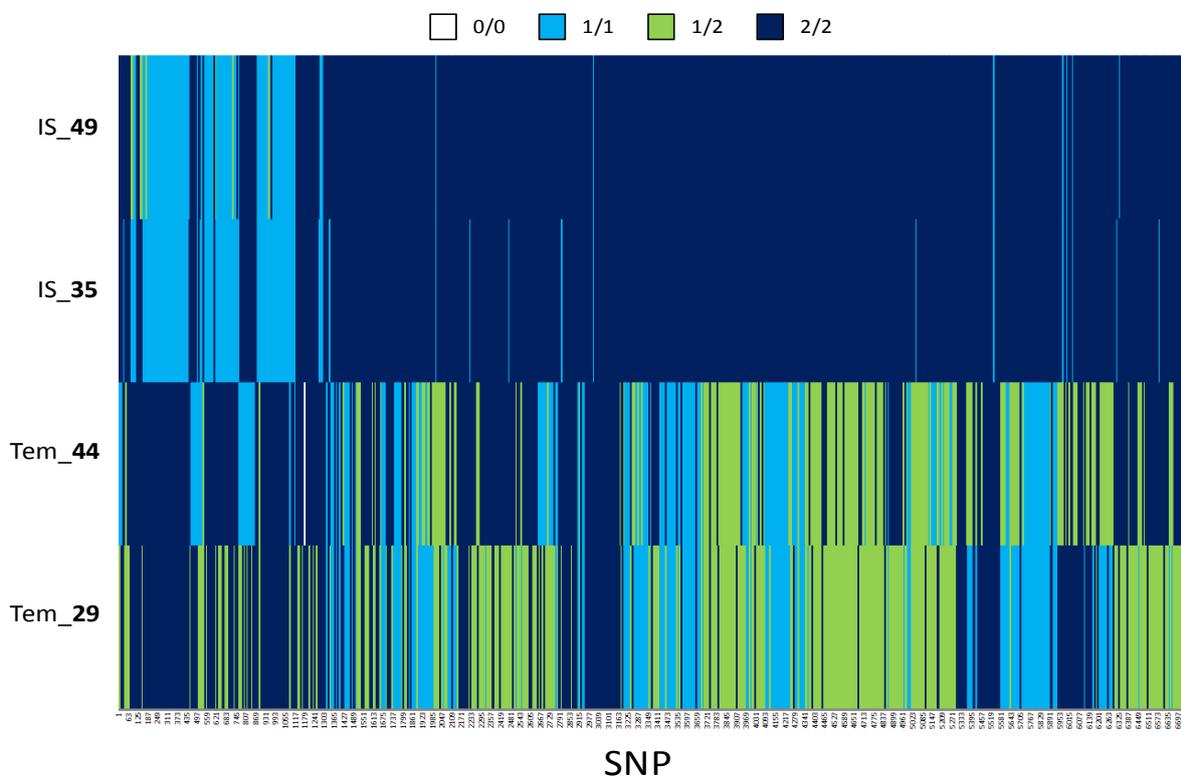


Figure 102 : Génotypes pour chacun des 6801 polymorphismes observés dans la région IS, et pour chacun des quatre animaux. Blanc 0/0 : pas de génotype attribué, Bleu : homozygote, vert : hétérozygote

### III.C. Identification de polymorphismes associés à l'intersexualité

A ce stade de l'analyse l'ensemble des variants identifiés n'est pas candidat à l'intersexualité. Afin de ne sélectionner que les SNP d'intérêt, qui pourraient être associés à l'intersexualité, une règle de décision a été établie. Selon des modèles de dominance et de récessivité : (1) un seul allèle muté suffit pour que l'individu soit atteint ; dans ce cas aucun allèle muté ne doit être présent chez les animaux sains (modèle dominant), (2) il faut deux allèles mutés pour que l'animal soit atteint, dans ce cas les témoins peuvent présenter au plus 1 allèle muté (modèle récessif) (Tableau 19).

Sur la base de ces deux hypothèses, un filtre a été appliqué à la liste des variants identifiés dans l'intervalle candidat et 6801 polymorphismes ont été conservés. La Figure 102 est une représentation des polymorphismes sélectionnés. Sur l'ensemble de la région, les deux atteints sont majoritairement homozygotes. Les témoins sont homozygotes dans certaines zones (exemple : entre 8,90 et 8,95 Mb), mais hétérozygotes dans d'autres régions (exemple dans la région 8,95 et 9 Mb). Initialement, les deux animaux atteints ont été sélectionnés sur la base de leurs génotypes aux marqueurs de la puce car ils étaient identiques et homozygotes dans la région. Les deux témoins ont été choisis comme très différents des atteints pour ces mêmes marqueurs, et présentaient moins d'homologie entre eux. Ces données sont cohérentes avec les résultats de la Figure 102: les atteints sont homozygotes et très comparables, et les témoins très différents l'un de l'autre ainsi que des deux animaux IS.

Néanmoins, pour 98 polymorphismes, les résultats attendus pour les deux atteints diffèrent (pour un même marqueur, un atteint homozygote d'un type et le deuxième atteint homozygote d'un autre type). De plus, l'animal IS\_49 est hétérozygote pour 66 polymorphismes en début de région, alors que le deuxième animal atteint est homozygote. La profondeur de lecture pour les marqueurs dont les génotypes sont différents entre les deux animaux atteints peut expliquer ces différences. En effet, il est possible que l'attribution soit faite à partir d'un nombre très faible de lectures, elle sera donc moins fiable. Pour les 98 polymorphismes pour lesquels un atteint est homozygote d'un type et le deuxième atteint est homozygote d'un autre type (Figure 103), trois quart des génotypes ne sont attribués qu'à partir d'une voire deux lectures pour un des deux animaux. La même réserve peut être apportée pour les 66 variants présentant un génotype déduit "hétérozygote" chez l'animal IS\_49 (Figure 104).

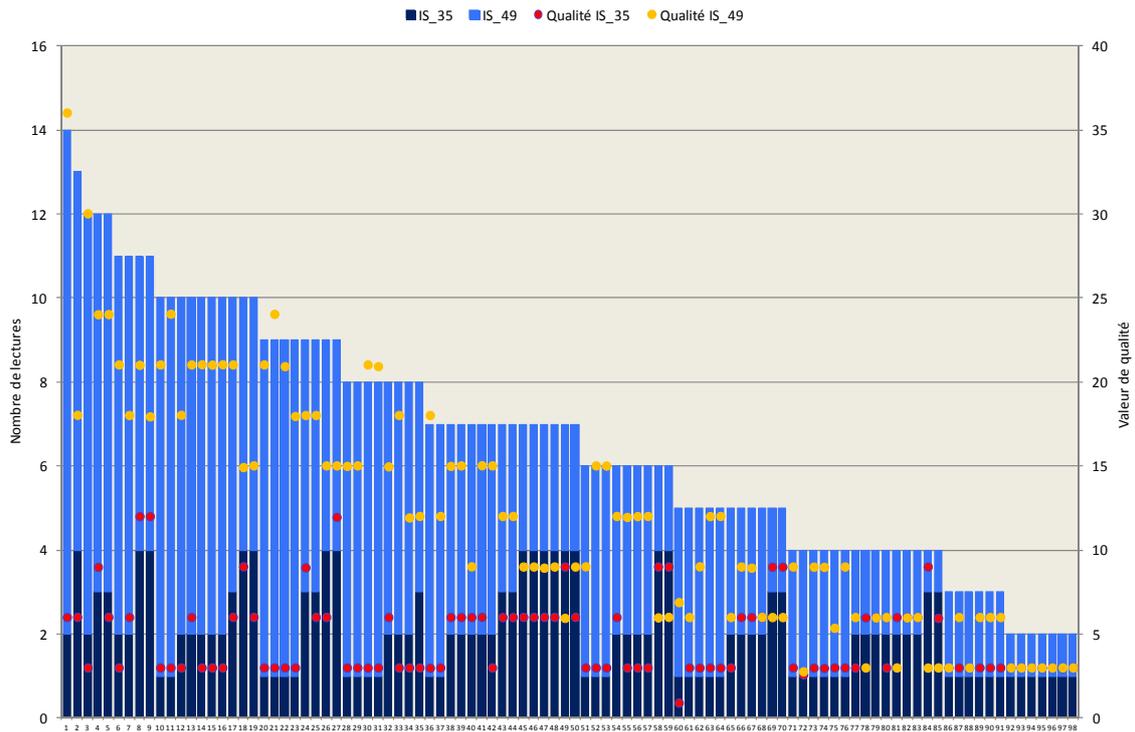


Figure 103 : Nombre de lectures observées et valeur de qualité pour les 98 SNP homozygotes pour chaque allèle respectivement pour les deux individus IS (IS\_35 : 1/1 et IS\_49 : 2/2 ou IS\_35 : 2/2 et IS\_49 : 1/1)

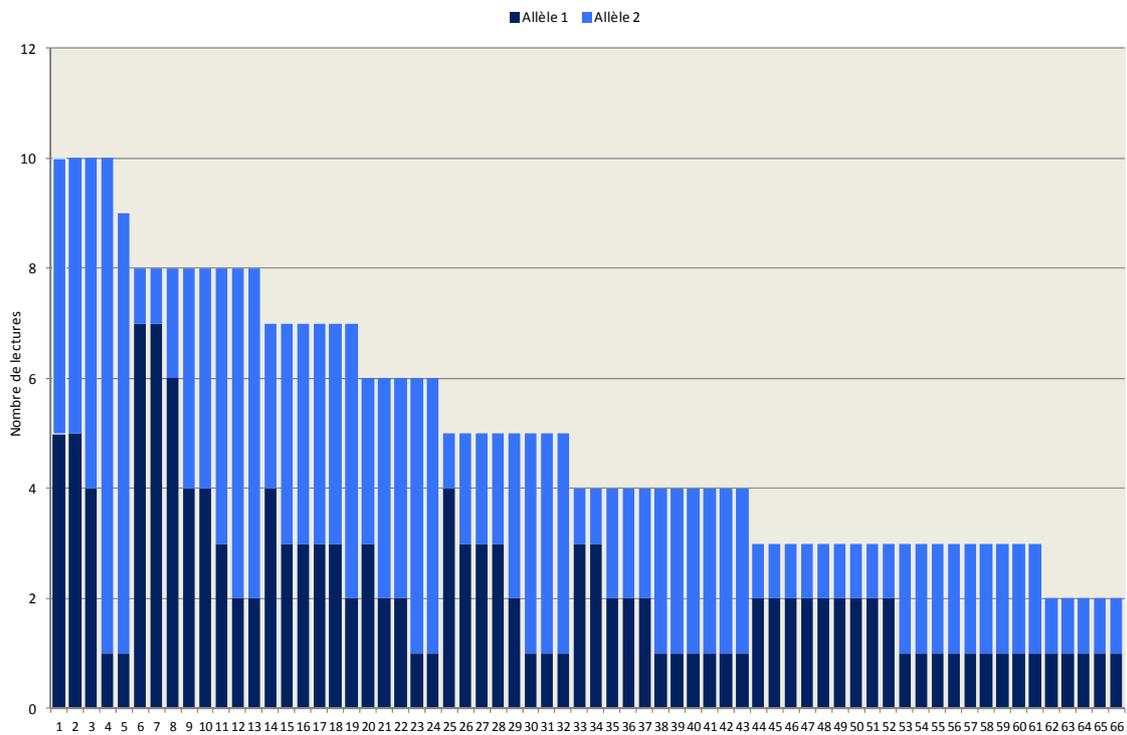


Figure 104 : Nombre de lectures observées pour chaque allèle des 66 SNP hétérozygotes pour l'individu IS\_49

Pour certains polymorphismes, le nombre de lectures correspondant à l'allèle 1 est très différent du nombre de lectures portant l'allèle 2, et pour la moitié des marqueurs une seule lecture est disponible pour l'un des deux allèles (Figure 104).

A l'issue de cette analyse, 6801 SNP peuvent être candidats au déterminisme génétique de l'intersexualité. Cette liste importante ne permet pas d'envisager une analyse de ces variants un à un, nous avons donc choisi de poursuivre cette étude par une analyse *in silico* de ces polymorphismes. La différenciation sexuelle fait intervenir de nombreuses voies de signalisation, il est donc possible qu'un des 6801 polymorphismes modifie un site de fixation nécessaire, par exemple, à la répression de *SOX9* chez la femelle par un facteur de transcription. Préalablement à cette étude des variants répertoriés, une recherche de CNV à partir des données de séquence a également été réalisée. Chez l'homme, des délétions ou des duplications à proximité de *SOX9* sont impliquées dans des cas de DSD. Les données de reséquençage permettent une recherche plus exhaustive de CNV dans la région que celle réalisée avec les données de génotypage de la puce.

Dans la suite de ce travail les données de reséquençage ont donc été utilisées pour la recherche de CNV dans la région, et la recherche de motifs de fixation dont la séquence pourrait être modifiée entre atteints et sains.

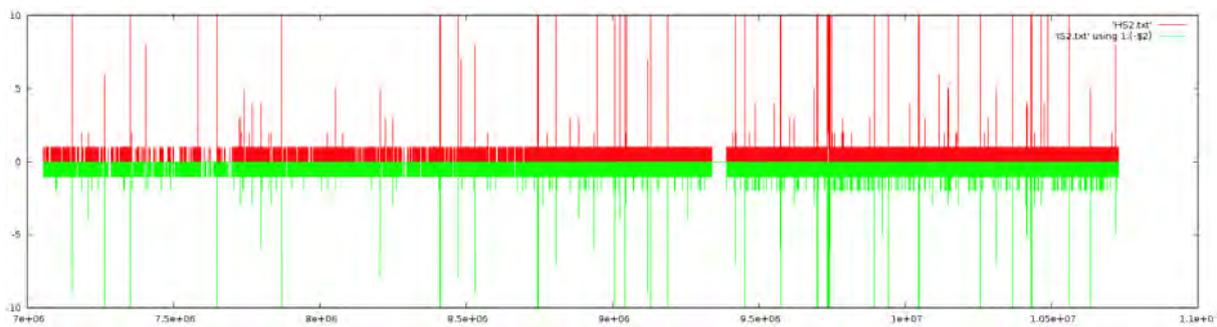


Figure 105 : Profondeurs des lectures par base entre les Témoins (rouge) et les IS (vert) dans la région du chromosome 12. L'axe des ordonnées correspond à la variation de la profondeur de séquençage par rapport à la moyenne de chacun des lots témoins / IS. Le seuil a été fixé à dix afin de mieux voir les petites fluctuations

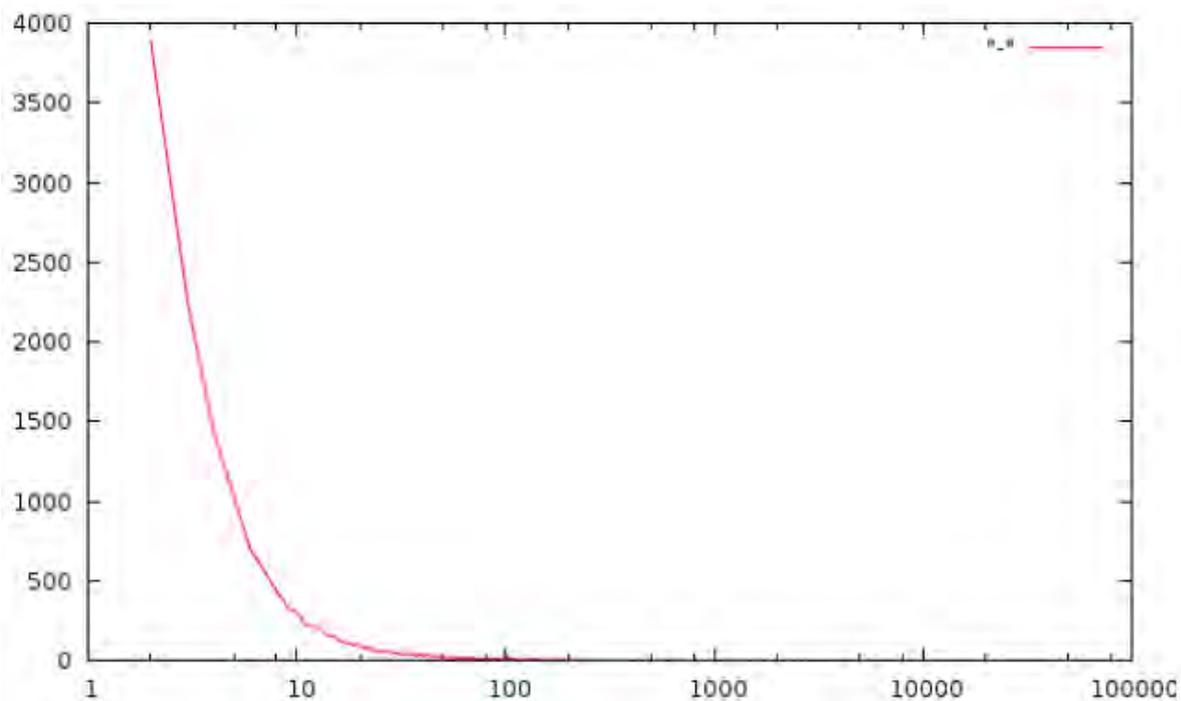


Figure 106 : Nombre de variations (axe des ordonnées) en fonction de la taille de la variation

## **IV. Recherche de sources de variabilité autres que les polymorphismes**

### **IV.A. Recherche de CNV**

Les CNV correspondent à une variation du nombre de copies d'une séquence d'ADN. La présence de CNV dans la région séquencée se traduirait par une variation de la profondeur en lectures à un endroit donné chez les atteints ou chez les sains. Cette variation peut être une absence de couverture (délétion), ou un nombre de lecture plus élevé dans une zone entre les cas et les contrôles.

La méthode utilisée ici pour la recherche de CNV est la comparaison du nombre de lectures chez les animaux IS et chez les animaux témoins, pour chacune des positions dans la région candidate. Afin de couvrir très largement la zone, les CNV ont été recherchés entre les positions 7 Mb et 11 Mb. Les résultats sont présentés en Figure 105. De nombreuses variations du nombre de lectures sont observées pour les deux catégories d'individus, mais ces variations sont très comparables entre les deux groupes. De plus ces variations de couverture ne s'étendent que sur une dizaine de bases en moyenne. La Figure 106 présente le nombre de variations selon la taille. Beaucoup de variations de profondeur ne portent que sur dix bases successives, voire cent bases, mais aucune variation de plus grande taille n'est observée.

D'après ces résultats, aucun CNV dans la région ne serait impliqué dans le déterminisme de la maladie. Il est néanmoins important de souligner que la couverture de la région n'est pas exhaustive et que la profondeur obtenue à partir de ces lectures est très faible dans certaines zones. De plus nous ne pouvons écarter l'hypothèse que des cas sporadiques d'intersexualité pourraient être dus, comme chez l'homme, à des CNV. Les deux cas d'intersexués séquencés sont des individus homozygotes pour l'haplotype 8. La conclusion tirée de cette étude est que la mutation causale associée à l'haplotype 8 n'est certainement pas un CNV.

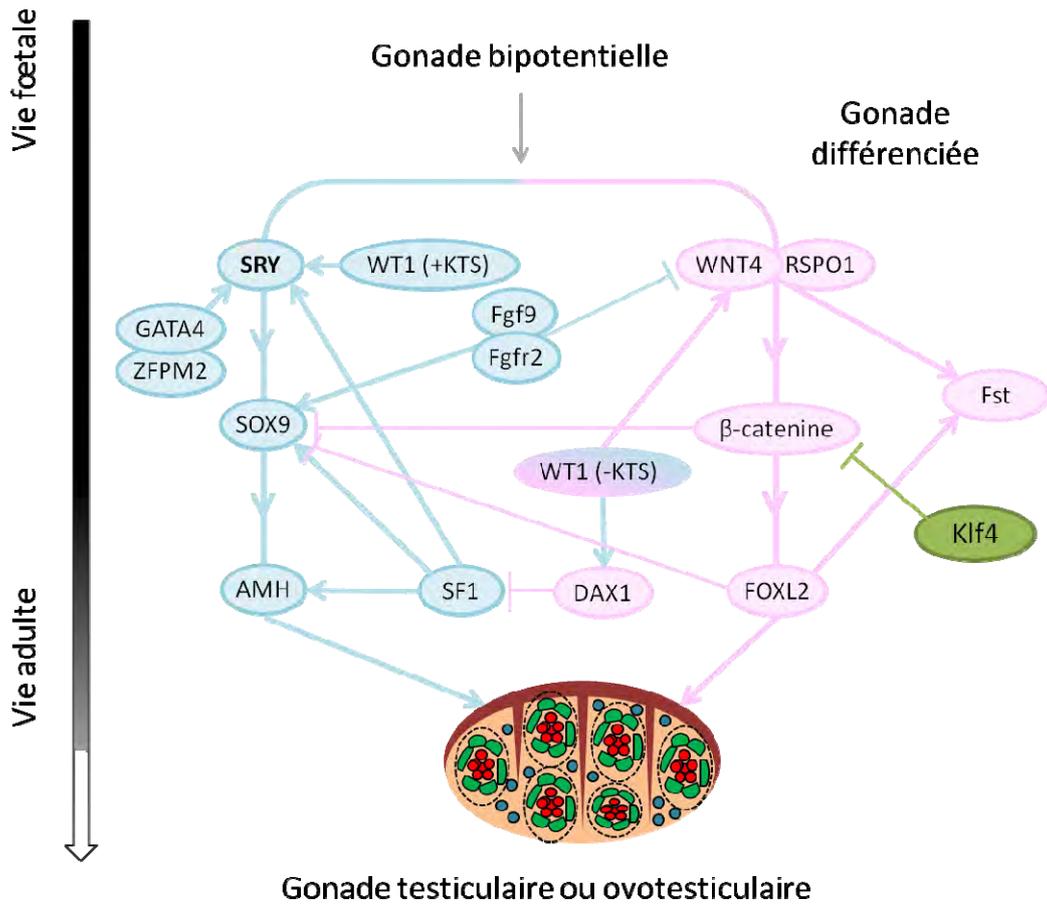


Figure 107 : Gènes impliqués dans le différenciation gonadique et rôle de Klf4

## IV.B. Etude des sites de fixation

Les cas d'intersexués étudiés dans cette thèse sont des animaux XX masculinisés. De nombreuses voies de signalisation sont impliquées dans la différenciation sexuelle. Chez le mâle, la voie *SOX9* est activée par la présence de *SRY*, puis l'expression de *SOX9* est maintenue par d'autres mécanismes. Chez la femelle, la voie *SOX9* est réprimée par Foxl2 et par les voies Wnt4/Rspo1/b-caténine et LEF/TCF (Figure 107). Tous ces facteurs de transcription agissent via leur liaison à des motifs nucléotidiques présents dans la séquence à proximité de *SOX9*.

Il est possible qu'un ou plusieurs variants identifiés différenciant les atteints et les individus témoins modifie(nt) un site de fixation d'un facteur de transcription de la différenciation sexuelle. Deux hypothèses sont envisagées : (1) les sites de fixation des facteurs de transcription de la voie « femelle » n'existent plus chez les atteints, *SOX9* n'est donc plus réprimé, (2) les sites de fixation des facteurs « mâles » sont anormalement présents, et la voie *SOX9* est surexprimée chez les intersexués XX.

Une recherche de motifs présents dans les séquences des individus atteints d'une part, et témoins d'autre part, a été réalisée dans la région candidate pour laquelle les polymorphismes avaient été identifiés (8 à 9,2 Mb). Les données dans l'un et l'autre des groupes ont ensuite été comparées. Au total, 1.688 régions d'un ou plusieurs motifs ont été identifiées chez les témoins, et 1.775 chez les intersexués. Après comparaison de ces régions entre les deux groupes d'individus, 427 correspondent à un ou plusieurs sites créés chez les animaux malades, et 408 sites disparaissent en présence de l'allèle muté.

Afin de filtrer le nombre de régions (et de mutations) candidates, nous avons fait l'hypothèse que les sites importants pour le développement gonadique devaient être conservés dans le génome de différentes espèces présentant les mêmes mécanismes de différenciation sexuelle. Nous avons donc répertorié les sites candidats mutés, présents dans les régions dont les séquences étaient conservées chez quatre autres espèces de mammifères (homme, souris, bovin, chien). Selon ce critère, seules quatorze régions candidates sont conservées, quatre régions correspondant à une perte de motif chez les atteints et dix à un gain.



Malheureusement, aucun de ces quatorze motifs affectés par l'allèle muté du variant ne correspond au site de fixation d'un facteur connu pour réprimer ou induire l'expression de *SOX9* lors de la différenciation gonadique femelle. Parmi les sites créés dans la séquence des individus atteints, un motif pourrait être intéressant à considérer. Il s'agit de Klf4, qui code pour un facteur de transcription régulant de nombreuses voies durant le développement embryonnaire. Klf4 est impliqué dans la différenciation des cellules de Sertoli chez la souris (Godmann et al., 2009). *KLF4* inhibe la voie  $\beta$ -caténine/Wnt et interagit avec TCF4 indépendamment de la  $\beta$ -caténine (Evans et al., 2010) (Figure 107).

Par manque de temps, cette étude n'a pas pu être poursuivie au-delà de cette analyse *in silico*. Avant d'entreprendre une analyse fonctionnelle fine d'un (ou de plusieurs) de ces sites de fixation candidats, il serait nécessaire de vérifier par le génotypage d'un plus grand nombre de cas et de contrôles que les polymorphismes affectant ces sites sont bien associés au statut de la maladie.

Cette étude de reséquençage complet a permis l'identification de plus de 6.000 polymorphismes dans la zone candidate, chacun d'entre eux pouvant potentiellement correspondre à la mutation causale. L'analyse de la couverture et de la profondeur du séquençage obtenu dans la région a montré que la recherche de polymorphismes n'était pas exhaustive. Néanmoins, à partir des données disponibles, une première recherche de mutation candidate (variant de structure ou mutation ponctuelle) a été réalisée. Bien qu'une première piste d'étude soit envisageable à l'issue de l'analyse *in silico* des sites de fixation, les travaux réalisés jusqu'à présent ne permettent pas de confirmer la causalité fonctionnelle du polymorphisme modifiant la séquence du site de fixation du facteur Klf4.

Parallèlement à ce travail, nous avons donc choisi d'exploiter ces polymorphismes pour une étude génétique, afin de réduire l'intervalle de la région candidate en recherchant des marqueurs en fort DL avec la mutation causale. Pour cela, nous avons décidé de sélectionner un jeu de marqueurs SNP complémentaires afin de génotyper un grand nombre d'individus et de réaliser de nouvelles analyses TDT. Une étape de densification en marqueurs a ainsi été réalisée dans la région.



## V. Densification de la région candidate en marqueurs SNP

Afin d'identifier des marqueurs en fort DL avec la mutation causale, nous avons choisi de génotyper un grand nombre d'individus avec certains des SNP identifiés par reséquençage dans la région. Pour maximiser les chances qu'un des SNP soit en DL total avec la mutation causale, nous avons choisi d'ajouter un nombre de marqueurs dans la zone permettant d'obtenir une densité comparable aux puces de génotypage HD. La région candidate étant localisée dans un intervalle de 1,3 Mb environ, il convenait d'ajouter au minimum une cinquantaine de marqueurs. Nous avons fait de choix pour ce travail de densification d'utiliser la technologie Fluidigm qui permet le génotypage simultanément de 96 SNP.

### V.A. Choix des marqueurs et des animaux

La première étape a consisté au choix des marqueurs complémentaires. La liste des 6.801 polymorphismes identifiés en amont est adéquate pour cette étude car la densité en marqueurs identifiés est très importante, et ces SNP sont discriminants des atteints et des sains. Afin de sélectionner les 96 SNP qui seront utilisés pour la densification, des critères de sélection ont d'abord été définis en fonction de la technologie utilisée. La technologie Fluidigm permet de génotyper des polymorphismes bialléliques type SNP, or les 6.801 polymorphismes regroupent des indels et des SNP bilalléliques et trialléliques. Une première sélection a permis de ne conserver que 3.567 SNP bialléliques.

Le deuxième critère de sélection a été défini en fonction de l'haplotype porteur. D'après les génotypes obtenus avec les marqueurs de la puce porcine, ainsi que pour les marqueurs identifiés par reséquençage, les animaux atteints sont très comparables, homozygotes pour la majorité des marqueurs, et différents des témoins. Quelque soit le mode de ségrégation de la maladie, la mutation ne sera jamais présente sur un « haplotype sain ». Ainsi, seuls les marqueurs discriminants (individus IS homozygotes pour un allèle, et individus témoins homozygotes pour l'autre allèle) ont été sélectionnés. Ce critère de choix a permis de réduire la liste de polymorphismes à 2.314 marqueurs.

Les animaux atteints ont été choisis initialement pour être homozygotes dans la région candidate. *A contrario* les deux individus contrôles sont plus variables et leurs génotypes moins certains. C'est pourquoi seuls les polymorphismes dont les génotypes ont été définis à partir d'au moins quatre lectures par individu sain ont été retenus, et 1.868 marqueurs ont répondu à ces différents critères.

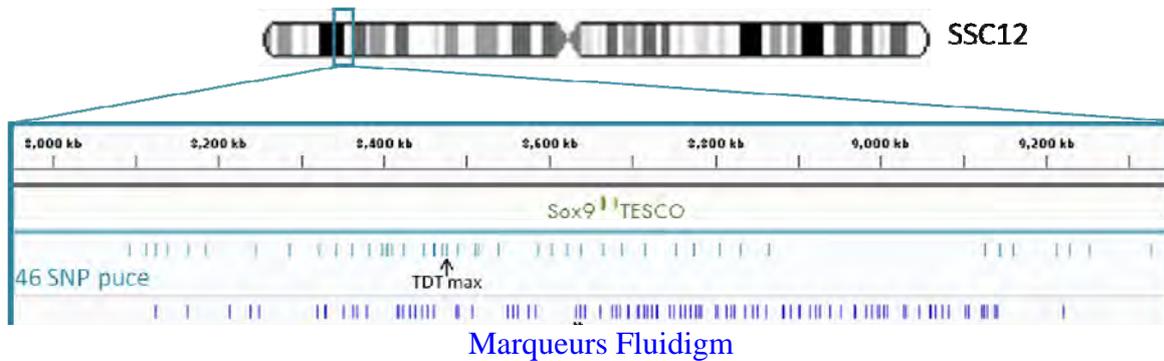


Figure 108 : Résultats TDT à partir des données de génotypage Fluidigm. La courbe bleue correspond au  $-\log$  de la p-value du CHI2, et les points rouges indiquent le nombre de parents hétérozygotes par marqueur (axe de droite), et donc leur informativité.

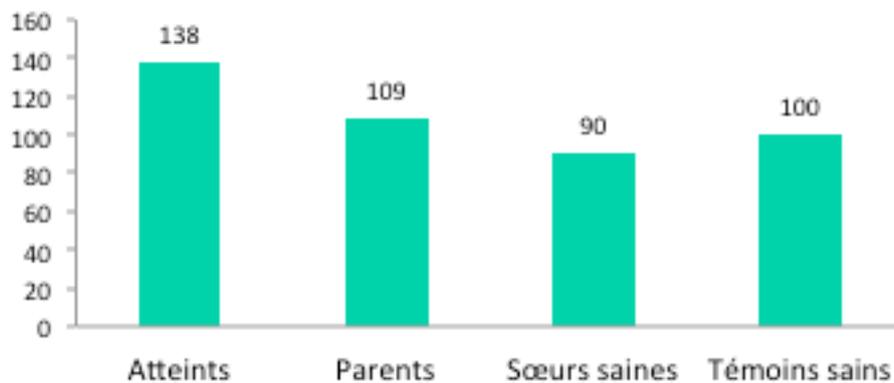


Figure 109 : Nombre d'échantillons selon le statut à la maladie

D'un point de vue fonctionnel, il est connu que les régions ayant une fonction importante pour les organismes présentent un niveau de conservation de séquence plus important. Nous avons donc veillé à ce que les marqueurs choisis soient à proximité des différentes régions conservées (+/- 1 kb). Cette analyse a été réalisée par comparaison des séquences porcine et murine. La liste de marqueurs a ainsi été réduite à 664.

A partir de cette liste, un dernier critère de sélection conditionné par la méthode de génotypage a été défini. Les amorces SNP spécifiques s'hybrident à une distance maximum de 60 pb du SNP d'intérêt, ainsi des polymorphismes de part et d'autre du SNP considéré peuvent empêcher l'hybridation des amorces, et donc biaiser le génotypage. Les marqueurs définitifs ont donc été choisis pour ne pas être à 60 pb minimum d'un autre polymorphisme.

Au final, 149 SNP ont ainsi été choisis et soumis au fabricant pour vérifier que le choix d'amorces était possible pour chaque marqueur. Enfin, une dernière sélection de 96 SNP a été réalisée pour le génotypage. Les 96 SNP sélectionnés sont présentés dans la Figure 108. Ils apportent une couverture homogène de la région et couvrent la région qui était initialement manquante (et non couverte par les marqueurs de la puce 60K) dans la version V10.2 du génome de référence, entre 8,94 et 9,12 Mb.

Au moment du génotypage, une quarantaine de nouveaux échantillons d'individus atteints ainsi que leurs parents avaient été réceptionnés au laboratoire depuis les précédentes analyses génétiques. Au total, 138 échantillons IS et leurs 109 parents (Figure 109) ont donc été génotypés pour pouvoir réaliser de nouvelles analyses TDT. En parallèle, des échantillons d'animaux sains sont venus compléter le dispositif. Deux types de populations contrôles ont été ajoutées : un échantillonnage de mâles LW issus d'un dispositif indépendant, et un échantillonnage de sœurs saines, ayant des parents communs aux animaux atteints, mais phénotypées non atteintes par les éleveurs. Il est important de souligner que ces deux groupes ne sont pas des populations contrôles idéales : les sœurs saines peuvent être porteuses (voire présenter le même génotype qu'un individu malade dans le cas où la maladie serait à pénétrance incomplète), et les mâles sains peuvent être porteurs ou présenter un génotype malade car l'intersexualité ne concerne que les femelles (masculinisation d'animaux XX). Cet échantillonnage nous permettra néanmoins d'estimer la fréquence d'haplotypes mutés candidats dans la population LW.

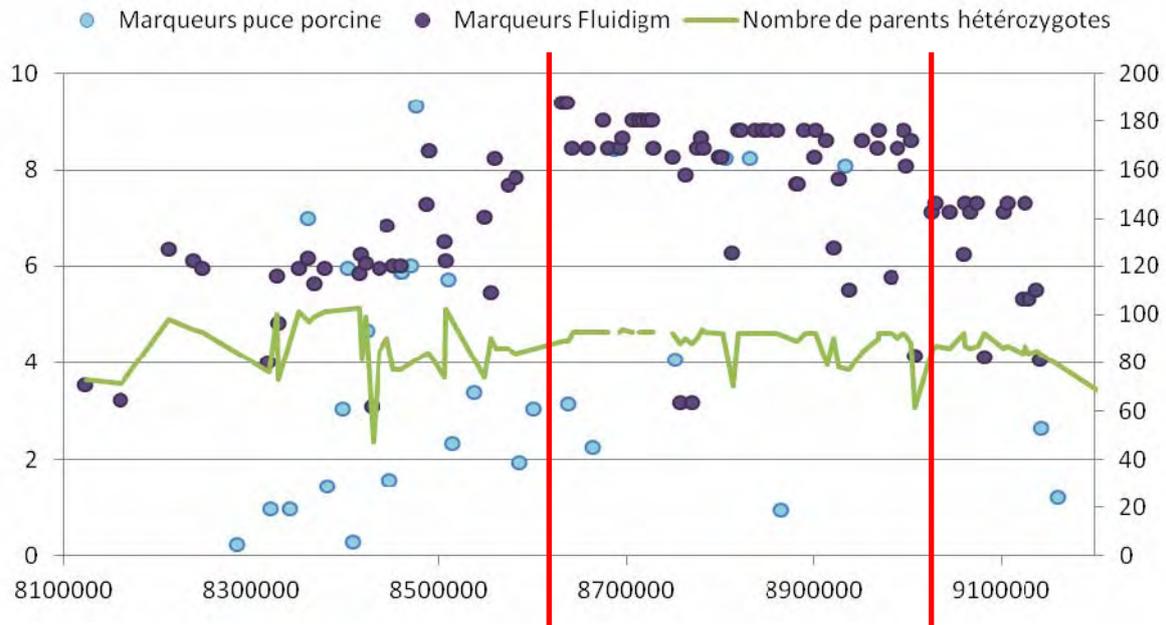


Figure 110 : Résultats TDT avec les marqueurs complémentaires Fluidigm, et les marqueurs de la puce porcine, ainsi que le nombre de parents hétérozygotes par marqueurs Fluidigm



Figure 111 : DL dans la région candidate à l'intersexualité avec les marqueurs de la puce porcine. Les carrés rouges indiquent un DL proche de 1



Figure 112 : DL dans la région candidate à l'intersexualité avec les marqueurs complémentaires Fluidigm

## V.B. Analyses TDT à l'aide des données de génotypage Fluidigm

Des analyses TDT ont été réalisées avec les nouvelles données de génotypage, afin de mettre en évidence un ou plusieurs marqueurs en distorsion de ségrégation. Les résultats des analyses sont présentés dans la Figure 110. Les résultats TDT précédemment obtenus avec les marqueurs de la puce porcine sont également indiqués sur la figure, afin de pouvoir estimer le gain de résolution obtenu. A l'issue de ces nouvelles analyses, les marqueurs les plus significatifs sont les marqueurs « 30 » et « 31 », adjacents et localisés en position 8,632 et 8,638 Mb, respectivement. Pour ces deux marqueurs, les animaux atteints ont reçu l'allèle n°1 78 fois, et l'allèle n°2 15 fois. Cependant, les valeurs pour les signaux les plus forts sont identiques à ceux qui avaient été obtenus avec un des marqueurs de la puce (M1GA0024789). A droite du signal le plus fort, entre les positions 8,6 et 9Mb, plus d'une trentaine de marqueurs ont également une p-value inférieure à  $10^{-8}$ . Bien que les nouvelles analyses n'aient pas permis de gagner en puissance, le nombre de marqueurs significatifs dans la région est très supérieur au nombre de marqueurs significatifs de la puce. Ce « plateau » (entre les positions 8,6 et 9Mb) signifie qu'un fort DL a été capturé avec les marqueurs que nous avons ajoutés. Cette différence de profil de déséquilibre de liaison, obtenue par l'ajout des 96 SNP complémentaires, a été estimée à l'aide d'*Haploview* et est représenté Figure 111 (DL estimé à l'aide des marqueurs SNP de la puce 60K) et Figure 112 (DL estimé à l'aide des 96 SNP complémentaires).

Bien qu'aucun marqueur ne soit en DL total avec la mutation causale, ces données ont permis d'identifier une région centrale dans l'intervalle primo-détectée, présentant une plus forte association avec le statut « intersexué ». La région candidate pour la recherche d'une mutation causale est désormais réduite à l'intervalle 8,6 – 9 Mb. La suite de ce travail a donc consisté à caractériser plus finement cette région par une analyse haplotypique.

n°_1	Freq_1	HAPLOTYPE_1	n°_2	n°_2	HAPLOTYPE_2		
8	0,508	[Haplotype 1 pattern]	1	0,6019	[Haplotype 2 pattern]		
39	0,004						
10	0,017						
19	0,008						
22	0,008						
20	0,004						
24	0,004						
5	0,008					2	0,0049
11	0,004					3	0,0049
12	0,008					4	0,0097
41	0,004					5	0,034
40	0,008						
1	0,004						
43	0,021						
4	0,004						
29	0,008						
26	0,004						
23	0,004						
14	0,004	10	0,2524				
13	0,017						
15	0,004						
18	0,004						
16	0,008						
2	0,067						
17	0,063						
34	0,029						
33	0,021						
42	0,017						
36	0,013						
3	0,008						
27	0,008						
28	0,004						
30	0,004						
31	0,004						
32	0,004						
37	0,004						
38	0,004						
6	0,004						
21	0,004						
		11	0,0097				
35	0,013	12	0,0146				

Figure 113 : Comparaison des haplotypes obtenus avec les marqueurs de la puce porcine (haplotype\_1), et après ajout de nouveaux marqueurs dans la région 8,6 à 9 Mb (haplotype\_2)

## V.C. Etude haplotypique

Les haplotypes des animaux atteints ont été reconstruits à l'aide des cinquante SNP complémentaires localisés dans la zone de plus forte association (377 kb). Au total douze haplotypes ont été identifiés, alors que la première analyse réalisée, dans un intervalle de 1,3 Mb avec les marqueurs de la puce avait aboutie à la définition de 43 haplotypes différents. La correspondance entre les haplotypes définis lors de ces deux études est rapportée dans la Figure 113.

Au sein de l'échantillonnage d'animaux atteints les fréquences des douze haplotypes identifiés sont variables. La fréquence de l'haplotype n°1 est de 60,2%. Le second haplotype le plus fréquent est le n°10, porté par 25,2% des atteints, suivi par le n°5, présent chez 3,4% des atteints. La fréquence des autres haplotypes varie de 2,9% (n°7) à 0,5% (n°2, 3, et 8). Une fréquence de 0,5% signifie que l'haplotype n'a été observé qu'une seule fois (soit un animal hétérozygote), une fréquence de 1%, deux fois. Bien que l'intervalle candidat et le nombre d'haplotype aient été réduits, aucune région identique et partagée ne peut être identifiée parmi tous ces haplotypes. *A contrario*, dans cette région de 377kb, les haplotypes n°1 et n°10, correspondent aux phases complémentaires (un allèle versus l'allèle alternatif pour chacun des cinquante SNP). Ces haplotypes sont les plus fréquents dans la population d'intersexués. Ce type de résultats laisse supposer qu'il pourrait y avoir au moins deux déterminismes différents à l'intersexualité.

Afin de savoir si un ou plusieurs de ces haplotypes pourraient être porteurs d'une mutation responsable de la maladie, les fréquences de ces haplotypes ont été estimées chez les populations contrôles, génotypées puis comparées aux fréquences haplotypiques des animaux atteints.

Les résultats des fréquences sont présentés en Figure 114. Seuls les haplotypes dont la fréquence est supérieure à 2% chez les atteints sont représentés. Les haplotypes n°1 et n°10 sont les plus représentés chez les trois groupes pris en compte. L'haplotype 1 est présent chez 60% des atteints et 40% des sœurs saines, ce qui en fait l'haplotype le plus représenté chez ces deux groupes. Chez les mâles, sa fréquence est de 26%. L'haplotype 10 a une fréquence comprise entre 25 à 30% pour les trois groupes.

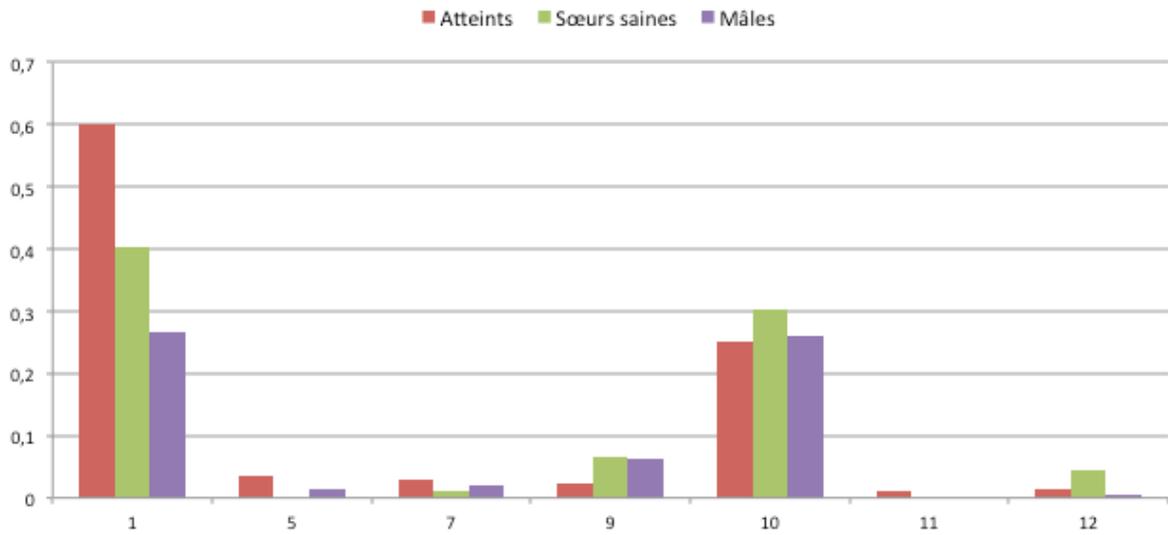


Figure 114 : Fréquence haplotypique chez des animaux intersexués, leurs sœurs saines et des mâles d'un autre projet

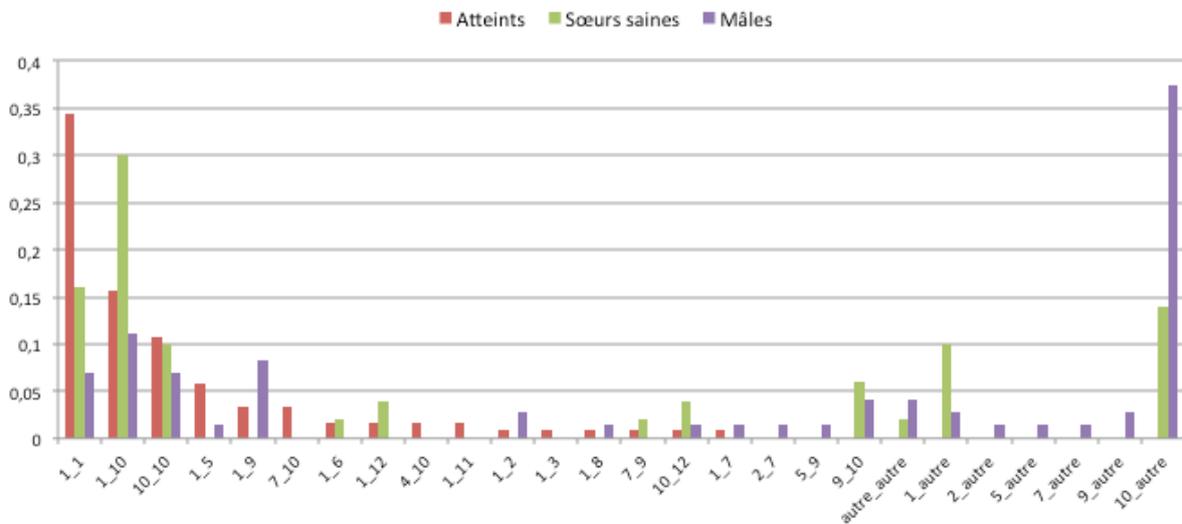


Figure 115 : Fréquence génotypique des haplotypes

Enfin, deux haplotypes, bien que de faible fréquence dans l'échantillonnage d'atteints, sont encore moins représentés dans les deux autres groupes : l'haplotype 5 (fréquence de 3,4% chez les atteints contre 1,5% chez les mâles et absent chez les sœurs saines) et l'haplotype 7 (3% chez les atteints contre 1% chez les sœurs et 2% chez les mâles).

Sous une hypothèse de récessivité, les individus sains peuvent être porteurs à l'état hétérozygote de la mutation. Il n'est donc pas aberrant que les haplotypes identifiés chez des individus IS présentent une fréquence non nulle dans les deux échantillonnages de contrôle. Les fréquences génotypiques des haplotypes ont alors été estimées pour les trois groupes, notre objectif étant d'identifier un génotype surreprésenté chez les animaux atteints par rapport aux sains.

Les fréquences haplotypiques chez des animaux sains ont été estimées, et comparées aux fréquences chez les atteints. Les résultats sont présentés dans la Figure 114. Près de 35% des atteints sont homozygotes pour l'haplotype 1, alors que seuls 16% des sœurs saines et 7% des mâles sont homozygotes pour cet haplotype (8 sœurs et 5 mâles contre 42 atteints). Le second résultat intéressant que nous avons relevé est que l'haplotype 5 (haplotype absent chez les sœurs saines) est toujours associé chez les individus IS à l'haplotype 1.

Au vu de ces résultats, nous faisons l'hypothèse que dans cette région deux haplotypes différents pourraient être porteurs d'une mutation induisant une dérégulation de *SOX9*. Le premier haplotype candidat est l'haplotype n°1. Cet haplotype est surreprésenté à l'état homozygote chez les atteints par rapport aux populations contrôles, ce qui est cohérent avec une hypothèse de récessivité. Dans la population LW, la fréquence de cet haplotype serait de 25% (estimation obtenue à partir de l'échantillonnage de verrats LW), et l'estimation d'une fréquence de sœurs saines « 1 / 1 » supérieure à la fréquence des verrats « 1 / 1 » indiquerait que la pénétrance du caractère est incomplète.

Le second haplotype candidat est l'haplotype n°5. Cet haplotype n'a jamais été identifié dans le groupe de sœurs saines et est systématiquement trouvé associé à un premier haplotype « muté » (haplotype n°1). Ces animaux malades pourraient donc également être homozygotes (au sens « porteurs de deux copies induisant une dérégulation de *SOX9* ») tout en étant hétérozygotes à chacune des deux mutations et aux polymorphismes environnants.

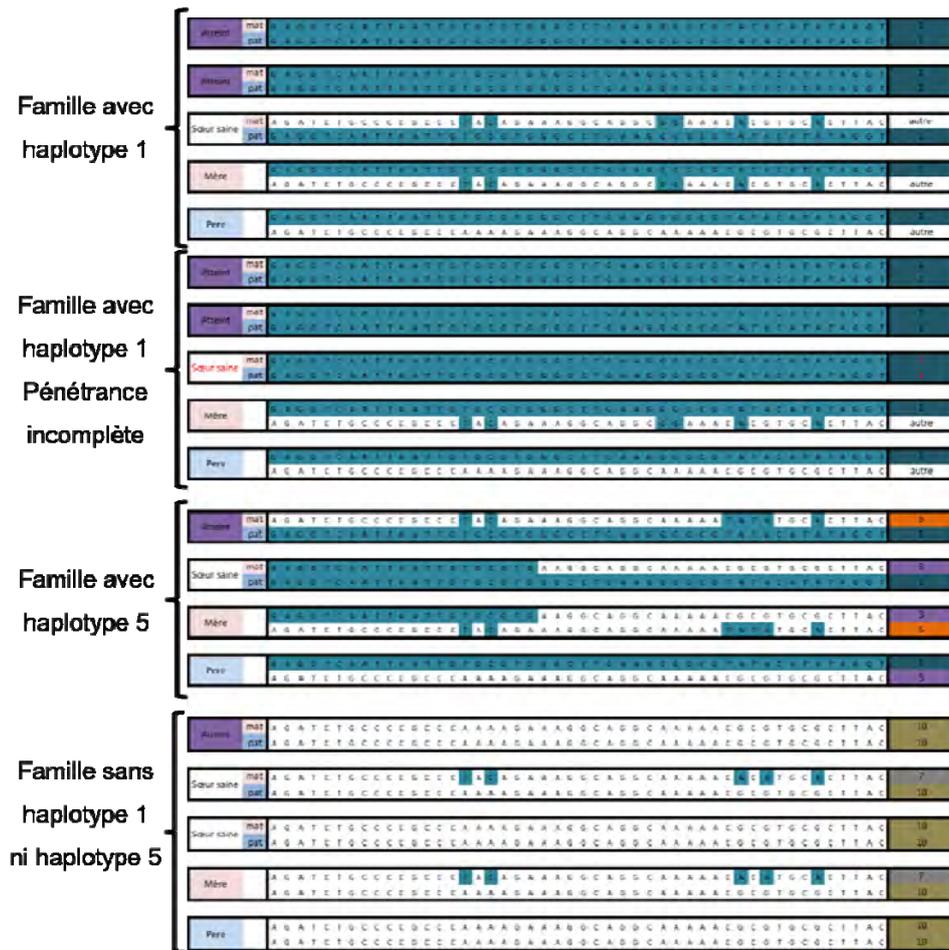


Figure 116 : Exemples de familles homozygotes pour l'haplotype 1, hétérozygote pour l'haplotype 5 ou sans haplotype 1 ou 5. L'origine du brin des descendants est indiquée en rose pour l'origine maternelle et bleu pour l'origine paternelle. Les numéros des haplotypes sont indiqués à droite.

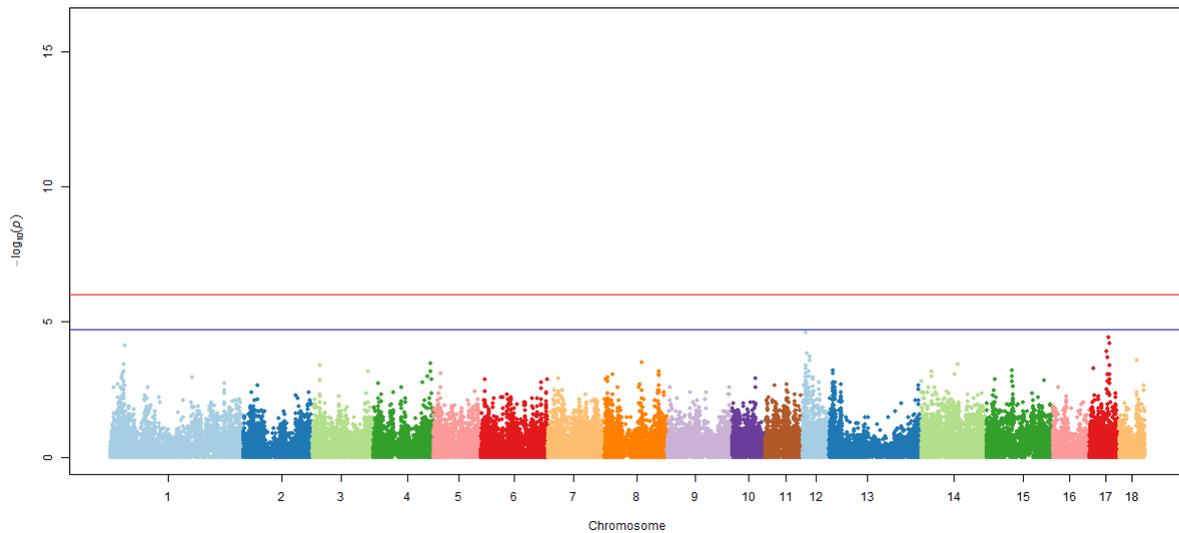


Figure 117 : Résultats TDT en retirant les descendants atteints homozygotes pour l'haplotype 1 et hétérozygotes pour l'haplotype 5

Au sein du dispositif, l'haplotype n°5 est en ségrégation dans les familles expérimentales INRA et n'a jamais été identifié dans les familles collectées par les OSP.

Afin d'estimer le nombre de cas familiaux ainsi expliqués, les différentes familles du dispositif ont été classées en fonction des haplotypes en ségrégation. Les 76 familles (père+mère+atteint(s)+sœur(s) saine(s)) sont réparties ainsi :

- Pour 25 familles (issues du dispositif INRA et des familles OSP), les descendants atteints sont homozygotes pour l'**haplotype n°1** (Figure 116). Parmi ces familles, treize d'entre elles comportent une ou plusieurs sœurs saines homozygotes pour cet haplotype. Pour les autres familles, aucune sœur saine n'est homozygote pour l'haplotype n°1. Le fait que certaines sœurs saines soient homozygotes résulterait d'une pénétrance incomplète.
- Pour cinq familles, les descendants atteints sont hétérozygotes pour les **haplotypes « 1 / 5 »** (Figure 116). Toutes ces familles sont des familles INRA, ajoutées à la collection. Ce dispositif a été mis en place en collectant comme fondateurs plusieurs mâles et femelles ayant eu, au sein de premières portées, des porcelets intersexués. Ces familles ne sont pas issues d'un seul verrat fondateur, nous ne pouvons donc pas exclure que différentes mutations causales soient en ségrégation dans ce dispositif.
- Néanmoins, pour les 46 familles restantes, le déterminisme génétique des descendants intersexués reste difficilement « compréhensible ». Aucun génotype particulièrement discriminant ne semble se différencier dans cette région (Figure 116). Afin de tester si une autre région que cet intervalle du chromosome 12 pouvait être impliquée nous avons réalisé une analyse TDT en retirant du dispositif les animaux homozygotes pour l'haplotype n°1, et les animaux porteurs de l'haplotype n°5. Les résultats des analyses TDT sont présentés dans la Figure 117. Aucun signal n'est significatif, et le marqueur pour lequel le signal est le plus fort est aussi localisé dans la région de *SOX9* sur le chromosome 12. L'analyse des allèles préférentiellement transmis indique que cette distorsion est également due à la transmission de l'haplotype n°1, mais cette fois chez des porcelets atteints, hétérozygotes. Une interaction avec une autre région du génome peut être suspectée mais l'absence de résultats significatifs peut également être due à un manque de puissance du test, car trente familles ont été retirées de ces analyses.



## VI. Discussion

Les analyses TDT réalisées à l'échelle du génome avec la puce porcine ont permis d'identifier une région de 1,3 Mb sur le chromosome 12 associée à l'intersexualité chez le porc. Le seul gène (*SOX9*) annoté dans la région est un excellent candidat à l'intersexualité, car ce gène est connu pour être impliqué dans des défauts de différenciation sexuelle chez l'homme et chez la souris.

Le reséquençage haut débit de la région que nous avons entrepris dans le cadre de cette thèse n'a pas permis d'identifier la mutation causale, mais plus de 6.000 polymorphismes ont été mis en évidence entre deux individus intersexués et deux individus contrôles. Une sous-sélection d'un jeu de marqueurs complémentaires nous a néanmoins permis de réduire l'intervalle de la région candidate. A l'heure actuelle, une zone candidate de 377 kb a été définie et deux haplotypes (les haplotypes "1" et "5") sont surreprésentés chez les animaux atteints par rapport aux animaux sains. Cependant, tous les cas d'intersexués ne correspondent pas à des individus dont le génotype serait "1"/"1", "5"/"5" ou "1"/"5", et les cas d'intersexualités d'une grande partie des familles du pedigree restent inexpliqués. Au vu des données disponibles actuellement, l'hypothèse la plus probable serait qu'un gène (*SOX9*) ou plusieurs gènes pourraient être impliqués, et que plusieurs mutations causales du gène *SOX9* pourraient également être en ségrégation dans la population Large White.

Ce type de déterminisme oligo-génique dû à différentes mutations à effets modérés (voire importants) a déjà été décrit. Chez les ovins, cette situation a été décrite chez les animaux des populations synthétiques Belclare et Cambridge présentant un phénotype hyperprolifique et des cas, à faible fréquence, de stérilité. Dans ces populations deux mutations indépendantes dans la séquence codante du gène *BMP15* et une mutation du gène *GDF9* ont été identifiées. Tous les animaux hétérozygotes pour une de ces trois mutations présentent un phénotype hyperprolifique (environ +1 ovulation pour les allèles mutés du gène *BMP15*, +2 pour l'allèle muté de *GDF9*) ; par effet d'additivité les animaux porteurs de la mutation *GDF9* et d'une des deux mutations *BMP15* présentent une prolificité d'environ +3 ovulation. Les cas de stérilité correspondent aux animaux homozygotes mutés à l'un des 3 loci **et** aux animaux porteurs des deux allèles différents mutés *BMP15* (Hanrahan et al., 2004).



Si le déterminisme génétique de l'intersexualité chez le porc résulte d'une situation complexe de ce type, impliquant différents gènes, il nous semble important de comprendre dans un premier temps l'implication de la région candidate du chromosome 12 dans la maladie par l'identification des mutations qui pourraient être associées aux haplotypes "1" et "5".

Les situations de ces haplotypes ne sont pas équivalentes : l'haplotype "1" est très fréquent dans la population Large White (60% chez les atteints contre 40% chez les sœurs saines et 25% chez les mâles) et majoritairement représenté à l'état homozygote chez les atteints (42 atteints, 8 sœurs saines et 5 mâles homozygotes). A l'inverse, l'haplotype "5" est rare et spécifique des atteints (3,4% chez les atteints, aucun chez les sœurs saines et 1,5% chez les mâles). Chez les atteints cet haplotype est toujours associé à l'haplotype "1". L'absence d'individus "5"/"5" (homozygote pour l'haplotype "5") n'est pas nécessairement surprenant compte tenu de la fréquence de cet haplotype. Afin de tenter d'identifier les mutations causales de *SOX9* associées à ces haplotypes, la stratégie la plus adaptée sera de reséquencer spécifiquement et de manière exhaustive ces deux haplotypes afin de comparer leurs séquences à des séquences d'individus sains. Les premières données de séquençage que nous avons obtenues dans le cadre de cette thèse indiquent qu'intra-Large White la variabilité différenciant deux chromosomes peut être importante et le choix des individus de référence sera donc primordial.

### **Obtenir une séquence de référence de qualité**

Afin de faciliter l'analyse des données de reséquençage le premier objectif reste l'obtention d'une séquence de référence de la région candidate. Malgré la reconstruction manuelle de l'assemblage de référence que nous avons réalisée, certaines séquences sont encore manquantes dans la région (près de 54.000 "N"). En parallèle près de 25% des lectures obtenues pour les quatre individus séquencés n'ont pu être alignées sur le draft. Il est possible que certaines de ces séquences correspondent à des portions de l'intervalle du chromosome 12. Il est donc possible d'envisager l'utilisation des lectures issues du séquençage tout génome réalisé dans le cadre de cette thèse pour réaliser un assemblage *de novo* de la région. Nous avons constaté que les génomes humain et porcine dans ce petit intervalle de quelques Mb ne présentaient pas de remaniements de synténie et pour faciliter l'assemblage *de novo* des régions manquantes nous pourrions envisager d'utiliser la séquence humaine en référence. A l'issue de ce travail, si quelques régions restent non couvertes, la stratégie la plus efficace sera de reséquencer spécifiquement les zones manquantes.

Pour les régions de jonction entre contig intra-BAC (correspondant aux séquences de 100 "N"), des amorces pourront être choisies aux extrémités des séquences encadrant des régions manquantes. Les produits d'amplification obtenus seront séquencés. Pour la région manquante dont la taille actuellement est représentée par 50.000"N", le reséquençage à faible profondeur d'un clone de BAC couvrant cet intervalle pourra faciliter la construction de l'assemblage de cette région. En dehors des lectures obtenues dans la cadre de cette thèse, ce travail d'assemblage pourra être réalisé en utilisant les données de séquençage *paired-ends* et *mate-pairs* issues d'autres projets.

### Quels individus séquencer ?

Lorsque la séquence de référence dans la région sera de qualité satisfaisante, le reséquençage de nouveaux haplotypes porteurs et non porteurs pourra être envisagé. Dans le cadre de cette thèse, les individus séquencés ont été choisis afin, entre autre, de répertorier des polymorphismes additionnels pour pouvoir génotyper le dispositif familial à l'aide d'un jeu de marqueurs complémentaires. Force est de constater que la différence haplotypique entre les chromosomes "sains" et "porteurs" sélectionnés sur la base des génotypes des SNP de la puce, s'est très largement confirmée ; constituer un jeu de marqueurs complémentaires à partir de ces données a été simple mais *a contrario* il n'est pas envisageable de rechercher parmi 6.000 variants une mutation candidate. Afin de réduire le nombre de polymorphismes candidats, il est nécessaire de pouvoir comparer des haplotypes sains et malades ayant divergés depuis peu. Si le temps de divergence est court le nombre de nouvelles mutations apparues dans l'un ou l'autre chromosome sera faible. L'idéal serait de disposer, comme chromosome "sain" de référence, du chromosome non-muté d'origine, celui dans lequel la mutation est apparue. Entre les séquences de ce chromosome de référence et du chromosome muté, une seule base serait différente : la mutation. Cette situation a été rapportée dans deux cas de clonage positionnel, l'identification de la mutation callipyge chez les ovins et la mutation du gène IGF2 chez le porc. Dans l'exemple ovin, un individu, hétérozygote au locus CLPG (estimé par testage sur descendance) présentait une forte homozygotie dans l'intervalle de localisation de la mutation. L'analyse généalogique de cet individu a permis de constater que cette identité était due à une boucle de consanguinité et que dans cette région ses deux copies étaient issues d'un unique chromosome hérité par les voies mâle et femelle d'un ancêtre à la troisième génération. Mais l'élément déterminant de cette analyse réside dans le fait que, par la voie mâle, cet ancêtre est le père du bélier "Solid Gold" chez lequel la mutation CLPG est apparue.



Le descendant séquencé était donc homozygote par descendance dans l'intervalle candidat à l'exception de la mutation CLPG (Freking et al., 2002). Ce "cas d'école" familial est une situation rare.

Néanmoins dans le cas du clonage positionnel du gène IGF2 chez le porc, la même stratégie a été adoptée et a prouvé son efficacité à l'échelle des populations. Une analyse haplotypique fine de chromosomes testés sur descendance, issus de races différentes, a permis d'identifier que l'haplotype muté Large White et Piétrain dans la région du QTL était identique à un haplotype sauvage d'origine Meishan, à l'exception de la mutation. Historiquement, la race Large White est une race synthétique constituée au cours du 18<sup>ème</sup>-19<sup>ème</sup> siècle à partir d'animaux de différentes populations dont des animaux originaires d'Asie. La mutation du gène IGF2 serait apparue dans un chromosome d'origine Meishan introgressé (puis sélectionné) dans la lignée Large White (Van Laere et al., 2003). La difficulté de ce type d'approche réside dans le fait que l'intervalle de localisation candidat doit être le plus précis et le plus sûr possible. La combinaison, au sein d'une même analyse, des séquences issues de races différentes nécessite que la taille de l'intervalle candidat soit comparable à la taille du DL entre ces races. La comparaison des séquences permettant l'identification de la mutation du gène IGF2 a été réalisée dans un intervalle dont la taille était inférieure à 30kb. Le second élément déterminant est la qualité de l'estimation de l'allèle, à la mutation recherchée, porté pour chaque chromosome pris en compte dans l'analyse. La façon la plus sûre d'estimer l'allèle à la mutation de chaque chromosome séquencé est de le tester sur descendance. En fonction du niveau d'effet de la mutation, le nombre de descendants produits devra être plus ou moins important.

Dans le cas de nos travaux, notre objectif serait donc de séquencer un haplotype "1" et un haplotype "5" et de comparer leurs séquences à des haplotypes sains identiques, *a minima* très semblables sur marqueurs, porteurs de l'allèle sauvage. Pour l'heure, l'haplotype "5" n'a pas été identifié chez des témoins sœurs saines et les haplotypes les plus comparables sont les haplotypes "6", "7" et "8", à la fois représentés dans les pools d'individus sains et intersexués. Concernant l'haplotype "1", il est également représenté dans les deux types de pools. Cette situation semble donc idéale pour une recherche des mutations causales par séquençage. Néanmoins l'analyse de ségrégation familiale de ces haplotypes au sein du dispositif ne semble pas en faveur de l'hypothèse de chromosomes identiques, porteur ou non-porteurs de la mutation, en co-ségrégation dans la population Large White :

au sein de certaines familles du dispositif, des descendants "1"/"1" issus de deux parents hétérozygotes "1"/"-" sont classés sur base de leur statut phénotypique dans le pool de malades, d'autres dans le pool de sœurs saines. Une même combinaison haplotypique, héritée par descendance n'aboutit donc pas au même phénotype. L'hypothèse la plus probable est donc qu'un seul haplotype "1" soit en ségrégation dans le dispositif et que la pénétrance observée résulte d'une expressivité variable de l'allèle muté ou d'une interaction avec un autre loci.

De façon comparable aux travaux réalisés pour l'identification de la mutation d'IGF2, l'ajout à nos analyses d'animaux issus d'autres races permettrait peut-être de réduire la taille de l'intervalle de localisation de la mutation en tirant parti du DL existant inter-race (Shearin and Ostrander, 2010). Cette méthode a également fait ses preuves chez le chien, où elle a permis d'identifier plusieurs mutations responsables de maladies monogéniques (Lequarré et al., 2011; Merveille et al., 2012). Jusqu'à présent, et afin de pouvoir rapidement disposer d'un échantillonnage de taille suffisante pour une première étude d'association, la collecte de famille Large White a été favorisée. Dans la collection, les familles d'autres races, comme le Landrace et le Piétrain ainsi que des familles de lignées composites sino-européennes sont présentes mais en petit nombre. Il serait intéressant de pouvoir tester si la même région chromosomique est impliquée dans les cas d'intersexualité de ces populations. Il sera nécessaire dans un premier temps d'augmenter le nombre de malades collectés, et de vérifier que le caractère intersexué de ces animaux n'est pas dû au locus SRY. Si un signal d'association est identifié dans la région de *SOX9*, une analyse comparative des haplotypes de la région pourra être réalisée afin de rechercher si un même segment chromosomique de petite taille pourrait être partagé entre les animaux Large White et les animaux d'une ou plusieurs autres races. A contrario, si aucune association n'est identifiée dans la région de *SOX9* une analyse par haplotype pourra également être réalisée. En effet, si un haplotype très proche de l'haplotype "1" était en ségrégation, l'absence d'association permettrait de faire l'hypothèse que ce chromosome est un chromosome non muté apparenté à l'haplotype muté. Le séquençage de ce chromosome et la comparaison de sa séquence à la séquence de l'haplotype "1" permettront de réduire fortement le nombre de mutations candidates.

### **Estimation des effets phénotypiques de l'haplotype "1"**

Dans l'état actuel, un des manques majeurs de nos travaux est l'estimation précise de la corrélation entre génotype et phénotype.

Jusqu'à présent les données phénotypiques fines sont issues des autopsies réalisées par Eric Pailhoux sur les animaux du dispositif INRA. L'analyse de ces familles à l'aide de la puce 60K révèle, comme pour les familles des OSP, une grande variabilité haplotypique dans la région candidate. Aucune "cohérence" entre variabilité haplotypique et phénotypique n'a pu être déduite.

Afin de poursuivre les travaux portant sur cette anomalie, il nous semblerait important de renouveler une étude phénotypique fine à la lumière des résultats génétiques obtenus. Actuellement, l'haplotype "1" est l'haplotype le plus "prometteur" (c'est l'haplotype qui contribue au signal statistique significatif des analyses TDT). Nous aimerions mettre en place un dispositif de testage de cet haplotype afin de pouvoir précisément décrire les effets phénotypiques qui lui sont associés. A partir des marqueurs génotypés, un jeu de SNP-tag permettant d'identifier les individus porteurs de cet haplotype pourrait être défini. L'objectif serait de sélectionner dans la population Large White des verrats et des cochettes homozygotes et/ou hétérozygote pour cet haplotype, afin de réaliser en unité expérimentale des croisements dirigés, permettant d'obtenir un nombre important de porcelets homozygotes "1"/"1" et hétérozygotes "1"/"-". L'objectif de ce dispositif serait multiple. En premier lieu, ces animaux permettraient potentiellement d'identifier des familles ne présentant jamais de descendants intersexués contrairement à d'autres familles et ainsi d'identifier, si cet haplotype existe, un haplotype "1" sauvage (non-muté). Le second intérêt de ces familles serait de pouvoir réellement estimer la pénétrance du caractère en estimant par famille présentant au moins un descendant intersexué, le nombre de porcelets "1"/"1" classés "sains" sur base d'une analyse morphologique externe. Le troisième intérêt de cette étude serait de réaliser une autopsie de ces porcelets sains (bien que théoriquement double-mutés) afin de vérifier si une faible expressivité du caractère ne pourrait passer inaperçue par un simple phénotypage morphologique. Enfin, ces animaux permettraient de comparer les différents phénotypes d'individus porteurs versus double-mutés.

Certaines estimations de l'effet de cet haplotype "1" pourraient également être réalisées à partir des données collectées par les sélectionneurs. Il serait intéressant, si la pénétrance est incomplète, d'étudier la carrière des truies porteuses de l'haplotype "1" à l'état hétérozygote ou homozygote. Pour certaines mutations, alors que les animaux homozygotes mutés présentent un défaut de performance, l'état hétérozygote peut conférer un « avantage » sur l'état homozygote non muté.

A titre d'exemple, dans le cas de l'intersexualité chez la chèvre, la délétion récessive responsable des cas d'intersexués (affectant l'expression des gènes PISRT1 et FOXL2) est associée à une hyperprolificité chez les animaux hétérozygotes (Pailhoux et al., 2001c). Chez le porc, une insertion dans le gène SPEF2 est responsable d'infertilité à l'état homozygote, alors que la taille de portée est augmentée lorsque les animaux sont hétérozygotes (Sironen et al., 2012). Bien que les animaux homozygotes soient contre-sélectionnés dans les élevages, la fréquence anormalement élevée de l'allèle muté, présent chez les individus hétérozygotes, est alors expliquée par l'avantage sélectif que ce génotype présente. Dans le cas de l'intersexualité chez le porc, les données de génotypage disponibles ne permettent pas de suspecter ce type de situation. Compte tenu de la fréquence allélique de l'haplotype "1" dans la population Large White, les fréquences des trois classes de génotypes possibles ne dévient pas des fréquences théoriques ( $p^2$ ,  $2pq$ ,  $q^2$ ). *A contrario*, nous ne pouvons pas écarter l'hypothèse que l'haplotype "1" affecterait défavorablement les performances de certains caractères à l'état hétérozygote. Les données que nous pourrions prendre en compte, pour cette estimation, serait la taille de portée, les difficultés de retour en chaleur et de réussite à l'IA, et les causes et âge de réforme. Sous l'hypothèse que l'intersexualité correspondrait au phénotype extrême observé pour le génotype "1"/"1", il devient alors nécessaire de mieux décrire les formes phénotypiques "intermédiaires" associées à ce génotype et au génotype hétérozygote. Pour l'heure, le taux d'animaux intersexués est relativement faible dans les élevages et l'intersexualité est donc considérée comme une anomalie moins problématique que les hernies. Si la même mutation peut également permettre de comprendre les causes de certains problèmes de reproduction des mères, le domaine d'application des résultats de ces travaux sera élargi.

### Quelle application en sélection...

L'objectif de ce travail de recherche collaboratif entre l'INRA, l'IFIP et les organisations de sélection porcine est à terme de pouvoir mettre en place une contre-sélection des allèles défavorables des gènes de prédisposition aux anomalies les plus fréquentes. La situation idéale serait de disposer de mutations causales. *A minima*, des marqueurs dont un des allèles serait en très fort DL avec l'allèle muté peuvent également permettre, intra race, une sélection tout aussi efficace. Si l'hypothèse d'un seul haplotype 1 muté se confirme, l'identification de la mutation causale n'est pas nécessaire pour entreprendre une contre-sélection de cet haplotype : les allèles des SNP localisés dans l'intervalle de 377kb présentent un très fort DL

et l'haplotype 1 est peu semblable aux autres haplotypes identifiés au sein de la population Large White dans cet intervalle.

Contre-sélectionner cet haplotype devrait donc permettre en théorie de réduire la fréquence de l'intersexualité. Néanmoins, la fréquence de cet haplotype dans la population Large White est de 25% ; il n'est donc pas envisageable d'éliminer 43,5% des individus, homozygotes (6%) ou porteur de cet haplotype (37,5%). Différents travaux destinés à proposer des schémas de sélections adaptés aux situations rencontrées ont été publiés. L'objectif de ces études est de proposer des stratégies permettant de réduire progressivement la fréquence des allèles à contre-sélectionner, tout en maintenant la diversité existante des régions adjacentes (et du reste du génome) dans la population sélectionnée. Dans le cas présent, une proposition pourrait être par exemple d'éliminer les verrats homozygotes pour l'haplotype 1 (à l'exception des animaux présentant un fort potentiel génétique), la population de femelle LW étant maintenue en l'état. A l'issue de treize générations de sélection, la fréquence de l'haplotype 1 serait inférieure à 10%, et la fréquence des individus homozygotes mutés, inférieure à 1%. Néanmoins, dans l'état actuel de nos connaissances, la pénétrance semblant incomplète chez les femelles (16% de sœurs saines homozygotes 1/1) et cet haplotype ne permettant pas à lui seul d'expliquer l'ensemble des cas d'intersexualités, il ne serait pas pertinent aujourd'hui d'entreprendre une contre-sélection assistée par marqueur de cet haplotype. Si l'étude phénotypique des individus homozygotes 1/1 (non-intersexués) et des individus hétérozygotes révèle, qu'en dehors d'un phénotype extrême intersexué, cet haplotype est également impliqué dans des cas de baisse de fertilité, cette question pourrait être alors reconsidérée. Il est également important de souligner que, quelque soit le caractère, l'utilisation d'une mutation (ou de marqueurs en DL) dans un schéma de sélection nécessite dans un premier temps de vérifier que cette mutation ne présente pas d'effet pléiotrope ou n'est pas liée à d'autres mutations qui risqueraient de diminuer la valeur génétique des animaux pour différents caractères d'intérêt agronomique.

Bien que l'identification de la mutation causale ne soit pas nécessaire pour la mise en place d'un schéma de contre-sélection, elle permettrait d'apporter des connaissances nouvelles sur cette maladie pouvant affecter l'homme ou d'autres espèces. En effet, le gène *SOX9* est impliqué dans des défauts du développement sexuel chez l'homme (Cox et al., 2011; Sekido and Lovell-Badge, 2008; Vetro et al., 2011; White et al., 2011), chez la souris (Bishop et al., 2000b; Lavery et al., 2011) mais également chez le chien (Nowacka et al., 2005) et le cerf (Kropatsch et al., 2013). Cependant, les associations identifiées chez ces différentes espèces n'ont pas nécessairement abouti à la compréhension des mécanismes sous-jacents.



L'identification des mutations causales chez ces différentes espèces, dont le porc, apporterait donc des données précieuses sur le fonctionnement et les moyens de dérégulation du gène *SOX9*. A l'issu des travaux complémentaires réalisés dans la région de *SOX9*, il serait également nécessaire de rechercher si une autre région ne serait pas impliquée dans le déterminisme de certains des cas d'intersexués. Pour l'heure, les effectifs analysés n'ont pas permis d'identifier un autre signal d'association dans une autre région du génome. Poursuivre l'effort de collecte de familles complémentaires reste donc une priorité dans la suite de ces travaux.



# Chapitre III : Cartographies physique et génétique du chromosome X porcin

---

Des GWAS ont été réalisées afin d'identifier des régions chromosomiques associées aux principales anomalies porcines. Une région en extrémité du chromosome X est très significativement associée aux trois maladies, quelque soit la race étudiée. Les analyses TDT sous hypothèse d'empreinte parentale ont montré que la distorsion de ségrégation était uniquement paternelle. Bien que des corrélations aient été établies entre la hernie inguino-scrotale et la cryptorchidie (anomalies de la migration testiculaire), l'implication d'un même locus dans les trois maladies a motivé des travaux de cartographie plus approfondis de la région.

Il est établi chez les autres mammifères que les extrémités des chromosomes X et Y sont homologues, et correspondent à la région pseudo-autosomale (PAR) (Helena Mangs and Morris, 2007). En utilisant la souris comme espèce modèle, il a été montré qu'un appariement est nécessaire entre les régions PAR des chromosomes X et Y (PARx et la PARy) lors de la méiose pour assurer la gamétogénèse (Kauppi et al., 2011). Peu de données sont disponibles concernant le chromosome X porcin, et la PAR porcine n'était pas caractérisée au moment des résultats des analyses TDT (Quilter et al., 2002; Skinner et al., 2013). Afin de comprendre pourquoi un signal d'association a systématiquement été observé dans cette région pour les trois maladies, il nous a paru nécessaire de caractériser finement la région.

La première partie de ce chapitre porte sur les travaux de cartographie fine de la région (cartographie génétique et études de fréquences alléliques) réalisés afin de confirmer ou non l'implication de la région dans le déterminisme des trois anomalies.

La deuxième partie du chapitre est la valorisation des travaux réalisés pour l'ensemble du chromosome X porcin, sous forme d'un article qui sera soumis à *BMC Genomics*. Très peu de données étant disponibles concernant le chromosome X, des études de cartographies physique et génétique ont été réalisées pour l'ensemble du chromosome.

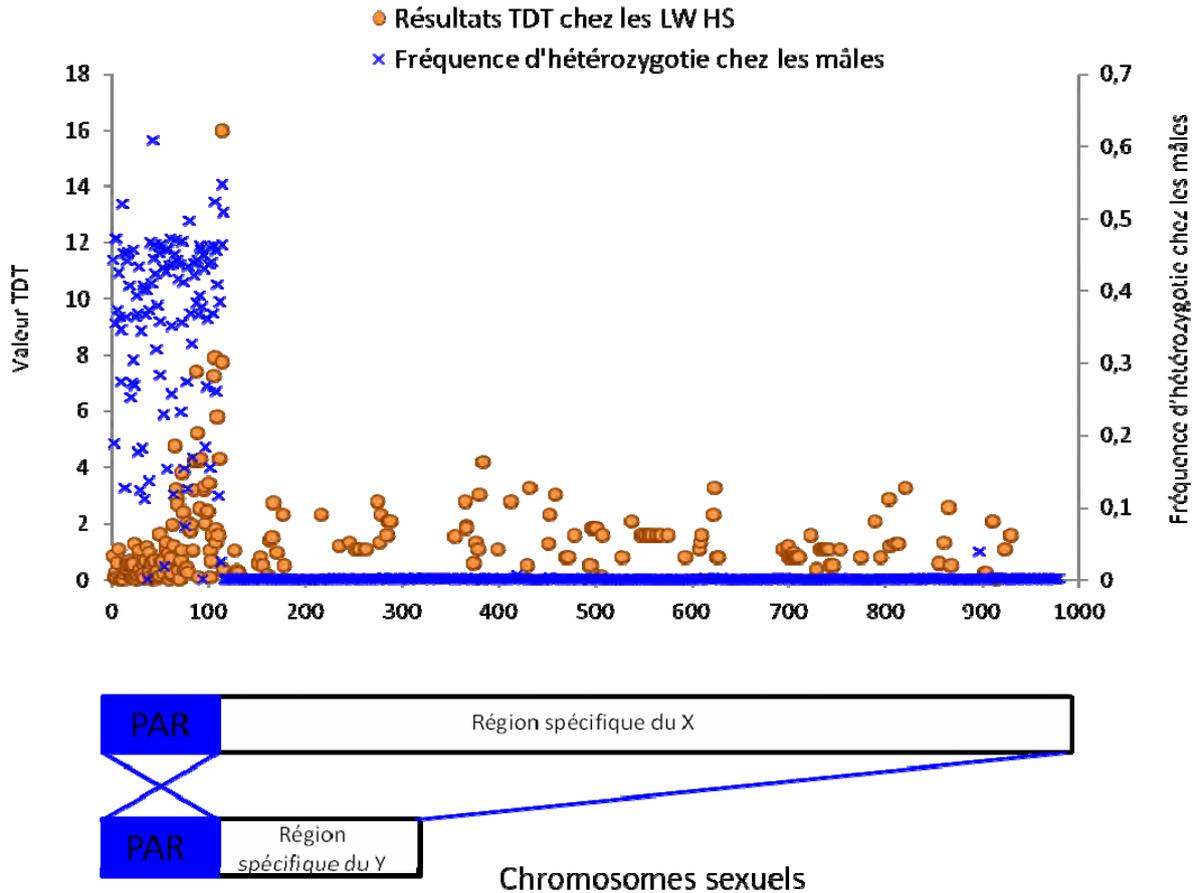


Figure 118 : Localisation de la région pseudo-autosomale porcine. Le signal TDT est indiqué par les points orange, et les croix bleues correspondent à la fréquence d'hétérozygotie. La PAR est localisée en bleu sur les chromosomes sexuels en bas de la figure. Les résultats TDT présentés ici proviennent des analyses chez les Large White atteints de hernie scrotale, mais la localisation du signal significatif est identique pour tous.

## I. Etude de l'implication de la région du chromosome X dans le déterminisme de la hernie inguino-scrotale, de la cryptorchidie et de l'intersexualité

L'objectif de la thèse étant d'identifier des gènes candidats responsables du déterminisme des principales anomalies chez le porc, une première étape a été de confirmer ou non l'association entre cette région et les trois anomalies. Les analyses TDT ont montré que les pères des atteints transmettaient, **dans plus de 90% des cas**, le même allèle à leurs descendants malades.

Le signal TDT est localisé en extrémité du bras court du chromosome X. Il augmente progressivement et atteint son maximum à la position 6,722 Mb (ALGA0109503), puis chute radicalement à partir de la position 6,811 Mb (Figure 118). Chez l'homme, les deux extrémités des chromosomes sexuels correspondent aux régions pseudo-autosomales PAR1 et PAR2 (Helena Mangs and Morris, 2007), alors que chez les autres espèces de mammifères, seule une région PAR1 a jusqu'à présent été détectée. Elle est localisée en extrémité du bras court chez la plupart des mammifères (Das et al., 2009; Raudsepp and Chowdhary, 2008). Le résultat TDT significatif étant positionné en extrémité du bras court du chromosome X porcin, une première étape de localisation de la région pseudo-autosomale porcine par rapport au signal a été entreprise.

### I.A. Localisation de la région pseudo-autosomale (PAR) porcine

Il est établi chez les mammifères que les mâles possèdent un chromosome X et un chromosome Y, alors que les femelles possèdent deux chromosomes X. Les mâles possèdent donc deux allèles au niveau de la PAR, un sur le chromosome X (PAR<sub>x</sub>) et un sur le chromosome Y (PAR<sub>y</sub>), et peuvent être hétérozygotes. En dehors de la PAR, ils sont hémizygotés pour les allèles des marqueurs spécifiques du chromosome X, et hémizygotés pour les allèles des marqueurs spécifiques du chromosome Y.

Afin d'identifier les marqueurs de la puce porcine potentiellement localisés sur la PAR, la fréquence d'hétérozygotie chez les mâles a été estimée pour l'ensemble des marqueurs annotés sur le chromosome X (Figure 118). Comme attendu, la fréquence d'hétérozygotie est supérieure à zéro pour les marqueurs en extrémité télomérique du bras court du chromosome X, puis devient nul pour l'ensemble des autres marqueurs annotés sur le chromosome X.

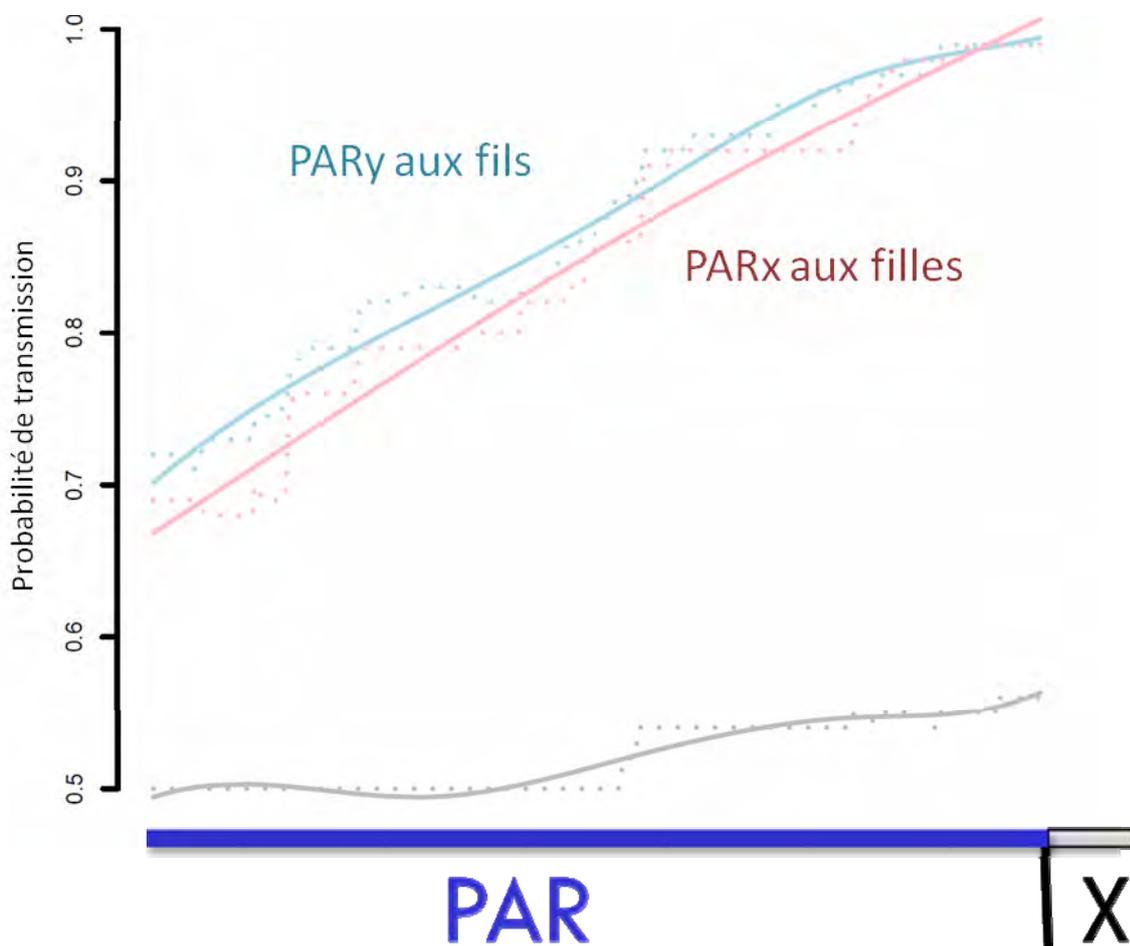


Figure 119 : Probabilité de transmission sur la PAR. Probabilité de transmission des allèles PARy des pères vers les descendants mâles (courbe bleue), des allèles PARx des pères vers les descendantes femelles (courbe rose), et des deux allèles des marqueurs de la PAR vers tous les descendants (courbe grise).

Cette région télomérique correspond donc à la région pseudo-autosomale porcine, et le signal significatif TDT maximum est localisé sur la PAR, à la limite de la région spécifique du chromosome X. Cette limite est appelée PAB (*PAr Boundary*).

Sachant que les pères transmettent leur chromosome Y à leurs descendants mâles, et leur chromosome X à leurs descendantes femelles, nous avons décidé d'étudier plus finement la ségrégation de la PAR. Bien que la transmission soit supposée équiprobable, il est possible que cette région soit affectée par le mode de ségrégation des chromosomes sexuels. La probabilité de transmission des allèles portés par la PAR du chromosome Y (allèles PARy) et la PAR du chromosome X (allèles PARx) selon le sexe des descendants a ainsi été estimée.

### **I.B. Transmission des allèles PARx et PARy**

Afin de réaliser cette étude, l'ensemble des familles génotypées à l'aide de la puce 60K dans le cadre de différents programmes, autres que le programme SwAn, a été utilisé. Les phases ont d'abord été estimées à l'aide de l'outil *Yapp*, puis la probabilité de la transmission des allèles PARy vers les descendants mâles, et de la transmission des allèles PARx aux descendantes femelles a été calculée. Les résultats sont présentés en Figure 119.

Près de la limite de la PAR, les probabilités de transmission sont égales à 1, ce qui signifie que tous les allèles PARy sont transmis aux fils, et que tous les allèles PARx sont transmis aux filles (comme observés pour les marqueurs localisés dans les régions spécifiques des chromosomes sexuels). Sur un autosome, cette valeur serait égale à 0,5 : la transmission de l'un ou l'autre allèle aux descendants mâles et femelles est équiprobable. Afin de vérifier qu'il n'y avait pas de biais de transmission des allèles au sein de familles présentant autant de descendants des deux sexes, la probabilité de transmission de chacun des deux allèles paternels à sa descendance a été estimée et est égale à 0,5. Les pères transmettent bien leurs deux allèles à leur descendance. Au vu de ces résultats, nous avons conclu que les allèles aux marqueurs situés près de la PAB ont le même mode de ségrégation que les allèles des marqueurs chromosome sexuel-spécifiques.

En extrémité télomérique de la PAR, la probabilité de transmission est de 0,7. Cette valeur indique que les pères ont 7 chances sur 10 de transmettre aux descendants mâles, les allèles portés par leur PARy plutôt que les allèles portés par leur PARx. Elle indique qu'une faible distorsion de ségrégation est également présente en extrémité télomérique de la PAR.

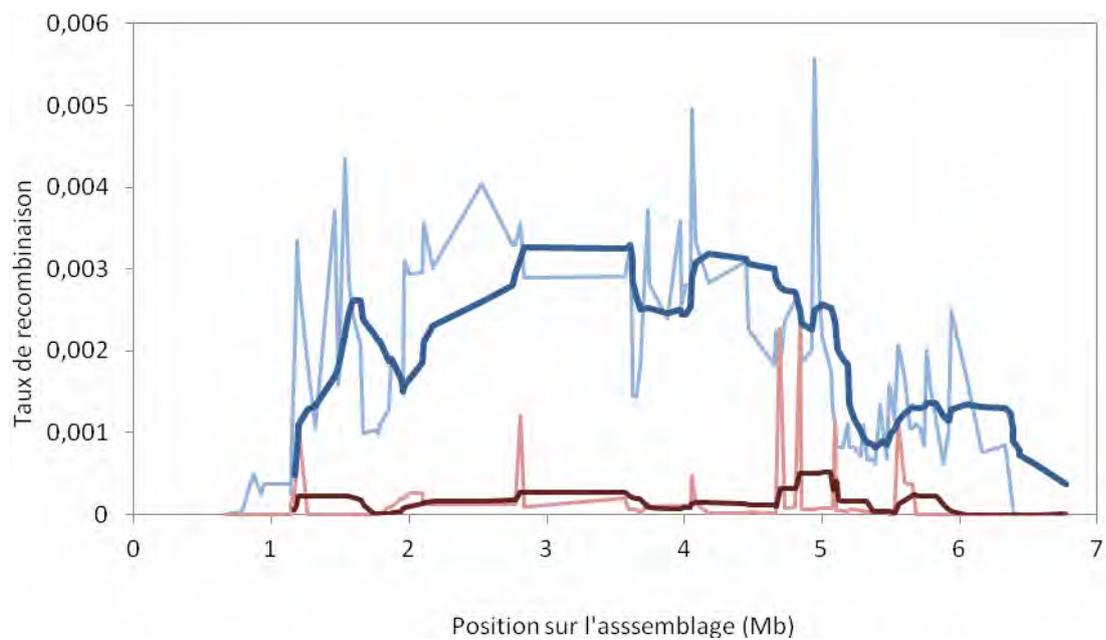


Figure 120 : Taux de recombinaison sur la région pseudo-autosomale porcine. La courbe bleu clair représente le taux de recombinaison chez les mâles (bleu foncé pour les valeurs lissées), et les courbes roses (données brutes et données lissées) le taux de recombinaison chez les femelles.

Cependant ce résultat nécessite d'être confirmé ; il est en effet possible que des évènements de recombinaison ne soient pas mis en évidence en extrémité télomérique par manque de marqueurs SNP et que la probabilité soit par conséquent sous-estimée.

Ces résultats montrent qu'une distorsion de ségrégation importante est observée sur la région pseudo-autosomale lorsque l'on se rapproche de la limite avec la région spécifique du chromosome X.

Il a pourtant été montré qu'un appariement est nécessaire au niveau de la PAR pour assurer la spermatogénèse chez le mâle (Kauppi et al., 2011). Des évènements de recombinaison sont donc attendus entre les régions PAR<sub>x</sub> et les PAR<sub>y</sub> des mâles. Des cartes de recombinaisons ont donc été calculées.

### **I.C. Taux de recombinaison sur PAR chez les mâles et les femelles**

A partir des familles utilisées précédemment, une carte génétique mâle et une carte génétique femelle ont été estimées à partir des méioses paternelles (recombinaisons entre les PAR<sub>y</sub> et le PAR<sub>x</sub>) et des méioses maternelles (points de recombinaisons, localisés dans la PAR, entre deux chromosomes X) (Figure 120). Comme attendu, le taux de recombinaison chez les mâles est supérieur dans cette région à celui des femelles. En effet au cours de la méiose, chez les femelles, les évènements de recombinaison peuvent avoir lieu tout le long du chromosome X.

Cependant, sur la carte mâle, les points des remaniements méiotiques ne sont pas répartis de façon homogène le long de la PAR. La recombinaison est faible près de la limite avec la région spécifique du chromosome X. Elle est proche de zéro entre les positions 6 et 6,8 Mb, et faible entre les positions 5,5 et 6,8Mb.

La diminution du taux de recombinaison vers la PAB pourrait avoir un impact sur l'histoire évolutive des fréquences alléliques entre les copies PAR<sub>x</sub> et PAR<sub>y</sub> dans cette région. En effet, (1) la diminution du taux de recombinaison empêche le brassage allélique entre PAR<sub>x</sub> et PAR<sub>y</sub> et (2) la distorsion de ségrégation induit une absence de transmission des allèles PAR<sub>y</sub> aux femelles.

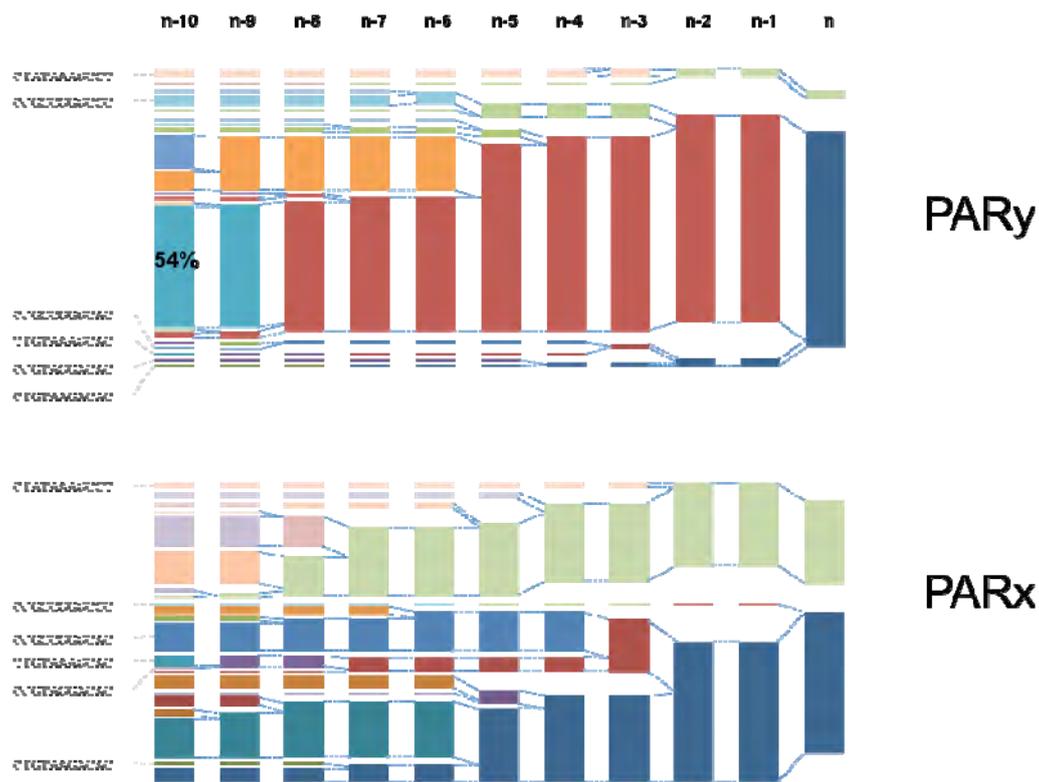


Figure 121 : Variabilité haplotypique estimée respectivement à partir de pools de régions chromosomiques PARy et PARx. La reconstruction haplotypique s'est fait de la droite (marqueur « n ») vers la gauche (marqueur « n-10 »), soit du marqueur le plus proche de la limite de la PAR vers le télomère (pour les 11 derniers marqueurs de la PAR). Les haplotypes, dont la séquence est indiquée, sont communs aux PARx et au PARy.

## I.D. Fréquences alléliques et haplotypiques des marqueurs PARx et PARy

A partir des phases préalablement reconstruites, des pools de régions chromosomiques PARx et PARy ont été constitués afin de comparer les fréquences alléliques et haplotypiques des copies PARx et PARy.

Dans un premier temps la reconstruction des haplotypes a été réalisée à partir du marqueur en limite de PAR (marqueur  $n$ ), puis par ajout successif du SNP en position  $n-1$ , puis  $n-2$ , jusqu'à  $n-10$  (Figure 121).

Dès le premier marqueur (marqueur  $n$ ) à la limite de la PAR, les fréquences alléliques entre PARy et PARx sont différentes. Pour ce marqueur, la fréquence de l'allèle C (Figure 121 et Figure 122) est de 92,5% pour PARy, et de 62,4% pour PARx, l'allèle C est donc très majoritaire sur PARy.

En ne tenant compte que des marqueurs  $n$  et  $n-1$ , trois haplotypes sont identifiés, et les différences de fréquences entre les pools PARx et les pools PARy s'amplifient : l'haplotype CT représente 37,1% pour PARx et 4,2% pour PARy (vert sur la Figure 121), l'haplotype CC 0,5% pour PARx contre 92% pour PARy (rouge), et l'haplotype AC 62,4% contre 3,3%.

Le nombre d'haplotypes identiques entre les deux pools diminue rapidement, et les haplotypes présents sur PARx sont très rares dans les pools de PARy.

En prenant en compte les 50 derniers marqueurs de la PAR, aucun haplotype n'est commun entre PARx et PARy. En effet la reconstruction des haplotypes pour ces 50 marqueurs et leur fréquence respective a montré que 113 haplotypes différents étaient présents. Un peu plus de la moitié des haplotypes est PARx spécifique, les autres étant PARy spécifiques (61 brins maternels contre 52 brins paternels), mais aucun haplotype n'est commun. Un haplotype paternel est surreprésenté chez les atteints LW (fréquence de 16%), alors que la fréquence maximum observée pour les brins maternels est de 5,7%. Les autres haplotypes ont une fréquence de 17 à 1%.

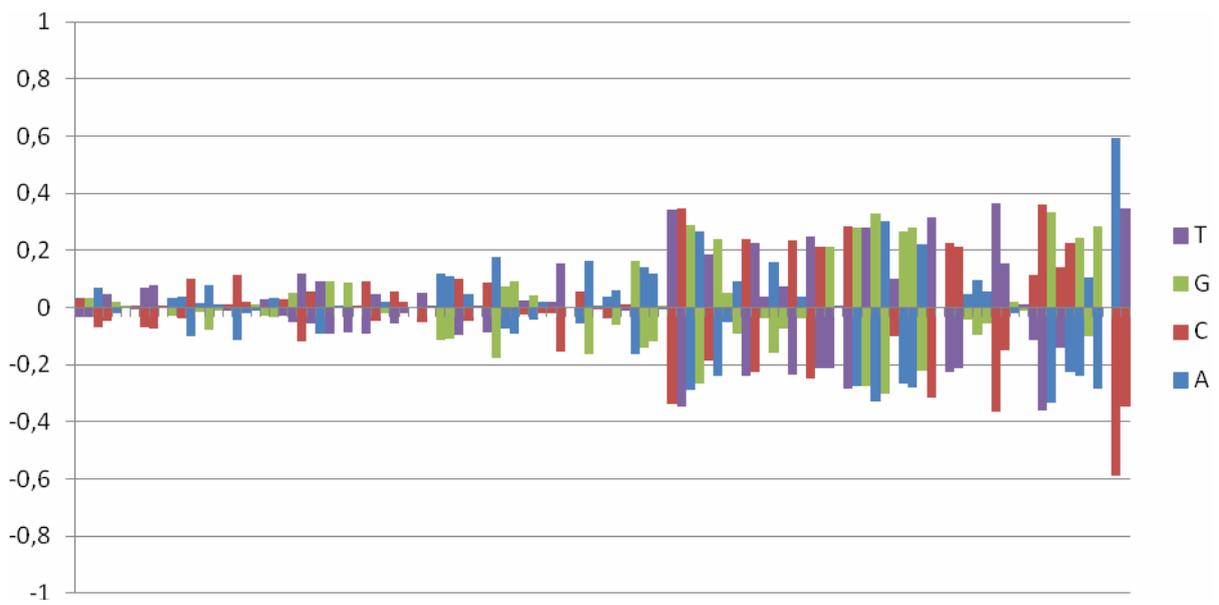


Figure 122 : Différence de fréquence entre les allèles PAR<sub>x</sub> et PAR<sub>y</sub>. L'axe des abscisses correspond à l'ordre des marqueurs sur la PAR, et l'axe des ordonnées à la différence de fréquences des allèles PAR<sub>x</sub>-PAR<sub>y</sub>.

Enfin, les différences de fréquences des allèles PAR<sub>x</sub> et PAR<sub>y</sub> ont été estimées pour l'ensemble des marqueurs de la PAR (Figure 122). La différence est très faible vers le télomère du chromosome, puis elle augmente progressivement jusqu'à atteindre son maximum à proximité de la région spécifique du chromosome X. Des différences entre PAR<sub>x</sub> et PAR<sub>y</sub> sont principalement observées pour les 50 marqueurs près de la limite, en accord avec les résultats précédemment obtenus.

L'ensemble de ces résultats souligne une spécificité génétique de la PAR. D'une part, une distorsion de ségrégation importante est observée plus on se rapproche de la limite de la PAR. D'autre part, l'histoire évolutive différente entre les allèles PAR<sub>x</sub> et PAR<sub>y</sub> a entraîné une différence de fréquence allélique.

		PARx	T	C
			37%	63%
PARy				
T	95,8%		35,45	60,35
C	4,2%		1,55	2,65

		PARx	T	C
			37	63
PARx				
T	37		13,69	23,31
C	63		23,31	39,69

**Pères hétérozygotes : 61,9%**

**T<sub>y</sub>C<sub>x</sub> : 60,35%**

**C<sub>y</sub>T<sub>x</sub> : 1,55%**

**Mères hétérozygotes : 46,62%**

Figure 123 : Fréquence des allèles sur PARx et PARy pour le marqueur le plus proche de la limite de la PAR

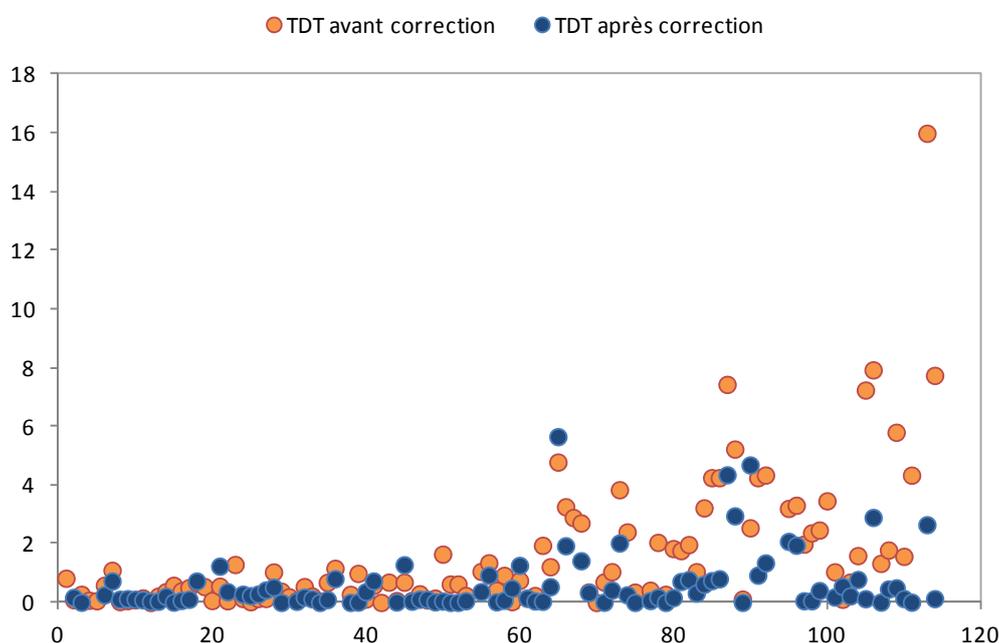


Figure 124 : Signal TDT observé sur la PAR pour les Large White atteints de hernie inguino-scrotale, avant et après correction du signal. La correction a été apportée en définissant une nouvelle hypothèse nulle, spécifique de la PAR, qui tient compte de la distorsion de ségrégation présente sur la PAR.

## I.E. Conclusion

Sous hypothèse nulle du test TDT, il n'y a pas de distorsion de ségrégation lors de la transmission des deux allèles d'un marqueur, et il n'y a pas de différences de fréquences entre les allèles d'un brin chromosomique et ceux de l'autre brin d'une même paire de chromosomes.

Dans le cas de la PAR, un fort DL entre les allèles des marqueurs de la PAR et des allèles des marqueurs des régions spécifiques des chromosomes sexuels est présente, en raison (1) d'une distorsion de ségrégation qui augmente plus on se rapproche de la PAB et (2) d'une absence de recombinaison entre les PARx et PARy.

Comme les fréquences respectives des deux types d'hétérozygotes (allèles1(PARx)/allèles2(PARy) et allèles1(PARy)/allèles2(PARx)) ne sont pas équiprobables, ainsi que la transmission des deux allèles (1 et 2) des SNP, aux descendants du même sexe, les résultats significatifs obtenus dans les analyses TDT pourraient être artificiels. La Figure 123 présente, à titre d'exemple, les proportions des allèles T et C sur les PARx et PARy pour le marqueur le plus proche de la limite de la PAR. La fréquence des pères hétérozygotes, type TyCx, est de 60%, alors qu'elle est inférieure à 2% pour les cas d'hétérozygotie CyTx. Sachant qu'il y a une distorsion de ségrégation des allèles PARy vers les fils, ces résultats suffisent à expliquer un signal TDT significatif. En effet dans notre dispositif, des descendants du même sexe par maladie sont sélectionnés. Pour rappel, les analyses TDT montraient que les pères des atteints transmettaient, **dans plus de 90% des cas**, le même allèle à leurs descendants mâles de hernie inguino-scrotale par exemple.

La correction de l'un ou l'autre des biais de la PAR (différences de fréquences alléliques PARx/PARy et distorsion de ségrégation) devrait être suffisant pour corriger le signal TDT s'il est uniquement dû aux biais de la PAR. De plus, il ne devrait pas y avoir de différences de fréquences génotypiques entre des mâles atteints et des mâles non atteints dans la région.

Nous avons redéfini une nouvelle hypothèse nulle spécifique de la région pseudo-autosomale, selon laquelle une distorsion de ségrégation existe indépendamment du statut à la maladie. En utilisant cette hypothèse H0, le signal TDT diminue et n'est pas significatif (Figure 124).

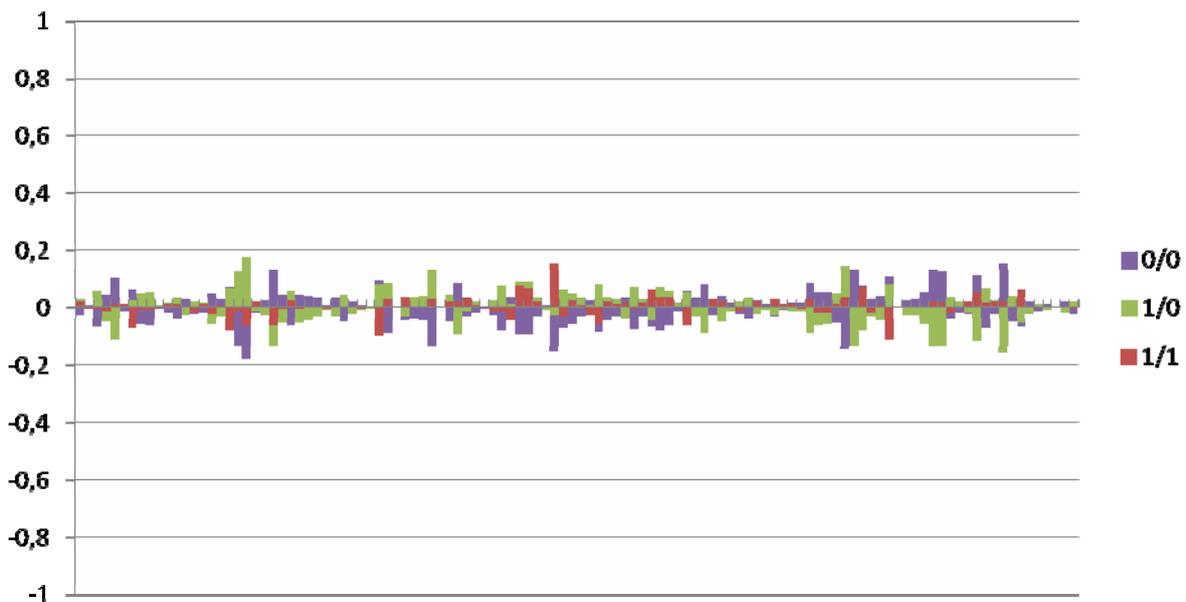


Figure 125 : Différence de fréquence génotypique entre les Large White atteints de hernie inguino-scrotale et des témoins Large White sur la PAR

Enfin, comme contrôle ultime, nous avons réalisé une étude d'association à l'aide des marqueurs de la PAR. Les fréquences génotypiques ont été estimées dans la population d'atteints étudiée dans SwAn et dans une population de mâles LW témoin issus d'un autre programme. Les différences entre les deux lots d'animaux pour les 3 classes de génotypes sont rapportées en Figure 125. Ces résultats montrent qu'un génotype particulier n'est pas surreprésenté chez les atteints par rapport aux témoins.

L'ensemble de ces résultats montre qu'il est nécessaire de corriger le signal sur la région pseudo-autosomale pour des analyses d'associations réalisées avec les marqueurs de cette région. Beaucoup de modèles (et de logiciels d'analyse) proposent pour le chromosome X d'étudier indépendamment les marqueurs de la PAR (selon un modèle autosomal) et les marqueurs localisés dans la région spécifique de l'X. Les résultats présentés ici laissent à penser qu'il serait plus juste (et moins source d'erreurs) de développer des modèles d'analyse spécifiquement pour cette petite région chromosomique afin de prendre en compte son mode de ségrégation très particulier. Des travaux de recherche de QTL impliqués dans les caractères de carcasse chez le porc (Duthie et al., 2009) avaient permis l'identification de QTL sur la PAR porcine, près de la limite de la région spécifique du chromosome X. D'après les auteurs, ce QTL serait soumis à empreinte maternelle, signifiant que seuls les allèles paternels étaient exprimés. Ce résultat est très comparable à nos observations initiales. Compte tenu des études complémentaires que nous avons menées, il serait intéressant de reconsidérer les données sur ce QTL

Les travaux que nous avons réalisés étaient essentiellement destinés à confirmer ou infirmer la présence d'une région chromosomique en extrémité du chromosome X porcine en forte association avec les différentes anomalies étudiées. Bien que notre attention portait sur la région PAR, le peu de données disponibles sur le chromosome X porcine nous a amené à réaliser une caractérisation fine de l'ensemble de ce chromosome : des études de cartographies physique et génétique ont été réalisées et devront faire l'objet d'un article à soumettre à *BMC genomics*.



## II.

### II. ARTICLE 2: A genetic and physical fine characterization of pig SSCX

*En vue d'une soumission à BMC genomics*

Sarah Rousseau<sup>1</sup>, Bertrand Servin<sup>1</sup>, Sylvain Foissac<sup>1</sup>, Thomas Faraut<sup>1</sup>,  
Denis Milan<sup>1</sup>, Juliette Riquet<sup>1</sup>

1. INRA, UMR444 LGC F-31326 Castanet-Tolosan, France



## INTRODUCTION

Le porc domestique (*Sus scrofa domesticus*) est une espèce d'intérêt majeur, d'une part car l'industrie porcine est une filière importante en agroalimentaire, d'autre part car il est un excellent modèle médical de par ses similitudes anatomiques, physiopathologiques et génétiques avec l'homme (Lunney, 2007). Depuis peu, une séquence de référence porcine (SusScrofa v10.2) et une puce de génotypage haute densité (*Porcine 60SNP BeadChip* d'Illumina) sont disponibles chez le cochon. Des travaux de cartographies physique (Servin et al., 2012) et génétique (Tortereau et al., 2012) haute densité ont également été réalisés pour les autosomes. Cependant, peu de données sont disponibles concernant les chromosomes sexuels. Ce manque de données s'explique par leur contenu en séquence très particulier, riche en motifs répétés, et un mode de ségrégation différent des autosomes. Pourtant, plusieurs QTL associés à des caractères de production ont été identifiés sur le chromosome X porcine (Duthie et al., 2009; Ma et al., 2013) mais le manque de données génétiques et les spécificités du chromosome X n'a pas permis d'identifier les mécanismes moléculaires sous-jacents.

Les chromosomes sexuels suscitent beaucoup d'intérêts. Chez les mammifères le chromosome X et le chromosome Y sont morphologiquement et génétiquement différents, contrairement aux paires d'autosomes. Les mâles possèdent un chromosome Y, porteur du gène majeur de la différenciation sexuelle mâle, SRY, et un chromosome X, alors que les femelles possèdent deux chromosomes X. Afin de compenser le niveau de transcription des gènes localisés sur le chromosome X, qui devrait être deux fois plus important chez les femelles que chez les mâles, un des deux chromosomes X est inactivé chez la femelle (Heard et al., 1997). Cette inactivation se fait grâce à la présence d'un ARN non codant, Xist (Chow et al., 2005), qui recouvre le chromosome X à inactiver. Une région du chromosome X échappe à l'inactivation, il s'agit de la région pseudo-autosomale (PAR) située à une extrémité télomérique des chromosomes sexuels (Helena Mangs and Morris, 2007). La PAR est une région dont la séquence est homologue entre le chromosome X et le chromosome Y, vestige de l'origine commune des chromosomes sexuels comme paire d'autosomes il y a 300 millions d'années (Livernois et al., 2012)

La recombinaison entre le chromosome X et le chromosome Y n'est limitée qu'à la région pseudo-autosomale (PAR), située en extrémité télomérique des chromosomes sexuels (Helena Mangs and Morris, 2007).

Chez l'homme, il existe deux régions PAR, aux deux extrémités du chromosome, alors qu'une seule région, PAR1, a été identifiée jusqu'à présent chez les autres espèces de mammifères (Das et al., 2009; Helena Mangs and Morris, 2007; Raudsepp and Chowdhary, 2008). Un appariement entre la PAR du chromosome X (PAR<sub>x</sub>) et la PAR du chromosome Y (PAR<sub>y</sub>) est nécessaire lors de la méiose chez la souris pour assurer la spermatogénèse (Kauppi et al., 2011).

Afin de caractériser finement le chromosome X porcine et la région pseudo-autosomale porcine, nous avons réalisé des travaux de cartographies physique et génétique à l'aide de la puce porcine haute densité pour l'ensemble du chromosome. Les travaux de cartographie physique ont permis de confirmer l'ordre de l'assemblage porcine sur ce chromosome, et de corriger certains défauts d'assemblage. De plus, une carte physique exhaustive de la PAR a

été réalisée, permettant ainsi de définir le contenu et l'ordre des gènes dans la région, et de confirmer la limite avec la région spécifique du chromosome X.

Nom du gène	Amorces (5'→3')	Taille du fragment (bp)	Température (°C)	Nom de l'EST
PLCXD1	left : GGTCCCTGGAGAGGATGAC	235	60	CK466735
	right : ATCCACAGGATGGAGGAG			
GTPBP6	left : CTCCTGTACCCCTTCACCT	110	52	CV867557
	right : CGCTCTCAGATGCCCTAAC			
PPP2R3B	left : GAGGCAGCGACTTCTGTGT	109	52	DN103585
	right : TCCGCGTGAATAAATAACAAC			
SHOX	left : CTGGATAGCACCTTAGCAGGA	95	56	EW428396,2
	right : TCCCACCCCTTTTATTTT			
CSF2RA	left : GATGTCACCTCCAAAACATCTC	100	52	FD631245,1
	right : AAAACCGAGTCAAAGAATCGG			
SLC25A6	left : CCACGAAGGCAGACCTAGAC	117	60	AK239241
	right : TATTGCGCTGTGCAAAAGAG			
ASMTL	left : GCATGCTGGTGGAGCAC	100	62	FD632834,1
	right : CGGGGCGGGTAGTACAG			
P2RY8	left : AAGAGGAGACACAGACAGAGG	102	60	BF188931,1
	right : CAGCTTCTGGGGTTCCA			
AKAP17A	left : GCCTTTCTCCCTTGTGTGTC	222	60	CF180526
	right : ACCTTCCTTTCCACGGACT			
DHRSX	left : GGGACACCTTGTGGGGATAG	179	60	AK233613
	right : CGCCTCTGCTTCTCACTTC			
ZBED1	left : CGAGCGCCTTCTGCTC	155	60	FD641139,1
	right : CCTCGTCTCTGCTCTCCA			
XG	left : ACACGTTGCCTGGAGTAAGG	243	60	AK232227
	right : ACCCAACAGCATTCTCCATC			
GYG2	left : CGTGCTGCTGTATCCATCCT	181	60	DT332656
	right : CGTTCTGCAAAGCCTAAACC			
ARSD	left : CACTTCCCTTCTGCTCCT	103	60	AK234744
	right : CGGCATCTCTGACTTCTGG			
ARSE	left : TGCATTTGCGACTAAGAACA	115	56	CJ031183
	right : CCACCTCAAAGCAAGACGA			
MXRA5	left : GTATGTGGTGGAGGGTGT	136	60	CK455361
	right : CTTAGGTGTGAGGGTGT			
PRKX	left : GGTGGTGTGTTGGGACTTTT	239	52	BI346943
	right : TTTTGTGTTTTCGGGGTTTT			
NLGN4X	left : CACACGCAAGAGAGAAAGGA	127	60	CN161317
	right : ATGGTCGTGGGAAAAGTCTG			
HDHD1A	left : CTTGGCAGAAAGCAGAGGAG	204	60	BG382475
	right : AAGGAGAGGGAGGGGATACA			
STS	left : CCGGGAAACATCATCTGG	152	58	DN128923
	right : CAGGGCTGCTTCCAGTGT			
PNPLA4	left : ATGTGCTGTTGGCGATG	100	60	CO952450
	right : GCCGGACAGTGACCATCT			
KAL1	left : GAAGAGTCTACCGCAATGG	156	60	DN125663
	right : CACATTAACACCGACGCAAG			
TBL1XY	left : ACACAGGCAGGGTCGAAA	153	60	BP166851
	right : GTCCACACACCCCAAGAGAC			
GPR143	left : GGAGACCTGTGAGAGGACGA	167	60	CN158227
	right : ACGGATCTCTCTGCATCAC			
SHROOM2	left : GCCGAGCAAATCTCAAAGAA	292	60	BF199365
	right : TGGCGACTTCAGAGTGAGTG			
WWC3	left : CAGACCAAACCTTGACTTCCATC	117	58	EW148791
	right : ATGGGCTCTTTGAGGTTCTG			
MID1	left : CGGGTAATGGGATTCAGTTC	96	58	AK349446,1
	right : TGGCTGCGTTGCTTTATC			

Tableau A : Couples d'amorces et conditions PCR pour chacun des 27 EST porcins sélectionnés pour la carte physique de la région pseudo-autosomale porcine

Par les travaux de cartographie génétique du chromosome X, nous soulignons le caractère très particulier de la région pseudo-autosomale quant à son mode de ségrégation, et la nécessité de prendre en compte cette spécificité dans les études d'association et de liaison. De plus, en dehors de la PAR, une région correspondant à un tiers du chromosome X présente un taux de recombinaison très faible. Enfin, nous avons réalisé des travaux approfondis du contenu en séquence de cette région qui ont permis d'identifier un motif particulier associé aux points froids de recombinaison.

## **MATERIEL ET METHODES**

### **Sélection des marqueurs**

*Sélection des STS:* Parmi les 64232 marqueurs de la puce porcine, 1586 marqueurs annotés sur le chromosome X (SSCX pour *Sus scrofa X chromosome*) selon la version 10.2 de l'assemblage porcin, et 3477 marqueurs SNP sans position sur l'assemblage ont été utilisés pour la construction de la carte physique.

*Sélection des gènes de la région pseudo-autosomale ancestrale et choix des amorces :* 27 gènes de la région pseudo-autosomale ancestrale des mammifères (Das et al., 2009; Helena Mangs and Morris, 2007) ont été sélectionnés. Les transcrits humains des 27 gènes ont permis de sélectionner des EST (*Expressed Sequence Tags*) porcins par homologie à l'aide d'ICCARE (*Interspecific Comparative Clustering and Annotation foR Ests*) (<http://bioinfo.genopole-toulouse.prd.fr/iccare>) (Muller et al., 2004) ou NCBI *Basic Local Alignment Search Tool* (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Des amorces ont été choisies à partir de la séquence des EST porcins, à l'aide de Primer3 (<http://frodo.wi.mit.edu/>) (Tableau A). Chaque couple d'amorce a été testé par amplification PCR sur de l'ADN porcin (contrôle positif), de l'ADN de hamster et de l'eau (contrôles négatifs).

### **Génotypage des clones d'hybrides**

*Génotypage du panel d'hybrides irradiés (RH) à l'aide de la puce porcine 60KSNP :* l'ADN des 90 clones du panel IMNpRH2<sub>12,000</sub> (Yerle et al., 2002) ainsi que de l'ADN génomique de porc et de hamster ont été génotypés à l'aide de la puce porcine d'Illumina. L'attribution des génotypes a été réalisée selon la méthode décrite dans les travaux de Servin et al. (Servin et al., 2012). Parmi les 50220 SNP obtenus, un vecteur de données a été obtenu pour 3138 marqueurs (1088 marqueurs annotés sur le chromosome X, et 2050 marqueurs sans position sur l'assemblage).

*Amplification PCR des séquences d'EST :* Les 27 EST porcins sélectionnés ont été amplifiés sur les 90 hybrides du panel IMNpRH2<sub>12,000</sub> (Yerle et al., 2002) par PCR. Le mix PCR est composé d'un volume réactionnel de 15 µl, avec 0,25 µM de chaque amorce, 0,25 Unité de Taq Polymerase, 200 µM de dNTP et 25 ng d'ADN. Les conditions PCR sont : 5 min à 94°C pour la dénaturation, suivi par 35 cycles de trois phases, chaque cycle étant composé d'une phase de 30 secondes à 94°C pour la dénaturation, d'une phase de 30 secondes à température gène-spécifique (Tableau A) pour l'hybridation des amorces, et d'une phase de 30 secondes à 72°C pour l'extension. En fin de réaction une extension finale de 20 min à 72°C est réalisée. Les produits PCR sont analysés par électrophorèse sur gel d'agarose à 3%.



Selon l'amplification positive ou négative, les résultats sont codés respectivement par "H" (pour *here*), and "A" (pour *absent*). En cas de marquage très faible, le résultat est annoté "-" (pour *missing data*).

### **Construction de la carte physique du chromosome X porcin**

Deux outils libres d'accès ont été utilisés pour la construction de la carte physique :

- L'outil CarthaGene (<http://www.inra.fr/mia/T/CarthaGene/>) (de Givry et al., 2005)
- L'outil Metamap (<http://snp.toulouse.inra.fr/~servin/index.cgi/Metamap>) (Servin et al., 2010a).

La carte physique a été construite en trois phases : 1) La construction, à partir des marqueurs annotés sur le chromosome X, d'une carte robuste, dont la probabilité de l'ordre des marqueurs est très forte, 2) l'intégration des marqueurs annotés sur SSCX mais exclus de la carte robuste, ainsi que les marqueurs sans position sur l'assemblage de référence porcin, puis 3) l'intégration des gènes de la région pseudo-autosomale ancestrale aux mammifères.

Les marqueurs physiquement liés sur le chromosome X ont été sélectionnés par une analyse de liaison entre les 1088 marqueurs SNP annotés sur SSCX, les 2050 marqueurs sans position et les 27 gènes de la PAR ancestrale. Les groupes de liaison ont été générés avec la commande `group` de CarthaGene et un LOD minimum de 6. Seuls les groupes de liaison de 6 marqueurs annotés sur SSCX ou plus ont été sélectionnés pour la construction de la carte robuste. Les données RH2 et la position *a priori* des marqueurs sélectionnés sur la séquence de référence porcine (version 10.2) ont été mergées à l'aide de CarthaGene (Faraut et al., 2007). Une première carte a été construite avec la commande LKH (pour *Lin-Kernighan heuristic*), puis améliorée avec la commande MCMC (for *Markov Chain Monte Carlo*). 5000 itérations ont été réalisées et les 1000 dernières ont été conservées. Les 25 meilleures cartes, pour lesquelles la probabilité 2-points est la plus forte, ont été utilisées dans Metamap pour construire la carte finale, dite robuste, dont la probabilité est proche de 1. La carte robuste est composée de 979 marqueurs compris dans 25 groupes de liaison. A partir de l'ordre des 979 marqueurs de la carte robuste, une carte physique a été construite pour chacun des 25 groupes de liaison avec la fonction `buildfw` dans CarthaGene. L'option `MrkTest` est égale à 1, et la `MrkList` correspond à l'ordre des marqueurs de la carte robuste. L'ensemble des 979 marqueurs de la carte robuste, 99 marqueurs additionnels non placés sur l'assemblage, 41 SNP non intégrés à la carte robuste mais localisés sur le chromosome X, ainsi que les 27 gènes de la PAR ancestrale ont été rassemblés dans une carte complète grâce à la commande `sem`. La carte complète correspondant à la région pseudo-autosomale porcine a été représentée graphiquement grâce à MapChart version 2.0 (Voorrips, 2002). L'ordre des gènes de la PAR porcine a été comparé au chien, au bovin, au cheval et à l'homme (Das et al., 2009; Helena Mangs and Morris, 2007) et représenté graphiquement avec MapChart.

### **Construction de la carte génétique**

Tous les animaux génotypés à l'aide de la puce issus de différents programmes ont été utilisés pour la construction de la carte génétique. Au total, 882 méioses femelles et 2264 méioses mâles d'animaux issus de races différentes étaient disponibles.



Les 979 marqueurs SNP de la carte robuste ont été utilisés pour estimer le taux de recombinaison entre les paires de marqueurs selon la méthode de Coop et al. (Coop et al., 2008).

### **Etude du contenu en séquence**

L'assemblage porcine actuel a été utilisé pour l'étude du contenu en séquence. Le pourcentage en GC a été estimé avec un script Perl dans des fenêtres non chevauchantes d'une Mb (Tortereau et al., 2012). La présence du motif consensus CCNCCNTNNCCNC, site de fixation du Prdm9 (Myers et al., 2008a) a été estimée grâce à fuzznuc (<http://150.185.138.86/cgi-bin/emboss/fuzznuc>). Une recherche d'éléments répétés *in silico* en tandem a été réalisée pour l'ensemble du chromosome X avec l'outil TandemRepeatFinder ou TRF (Benson, 1999). Les paramètres d'analyse utilisés sont : 500 (taille des motifs de 1 à 500pb); 2 (score d'alignement défini par défaut); 7 (score d'indels, valeur la moins permissive) - 50 (score d'alignement minimum); 500 (taille maximum de la séquence en tandem); 80 et 10 (paramètres de détection définis par défaut, 80 pour la probabilité d'alignement, 10 pour la probabilité d'indel). Les motifs identifiés sont alignés sur le génome porcine (Sscrofa10.2) à l'aide de l'outil Glint, (Courcelle et al., 2008) et l'option `-align`. Les éléments dispersés connus et les séquences de faible complexité ont été extraits de la base de données UCSC Table Browser (<http://genome.ucsc.edu>) (Karolchik et al., 2004) en sélectionnant les paramètres suivants : version de l'assemblage porcine : Sscrofa 10.2, groupe : *Variation and Repeats*, track : *RepeatMasker*, format de sortie : *BED*. Le nombre de motifs et le pourcentage de GC sont ensuite estimés sur tout le génome par fenêtres non glissantes d'1Mb.

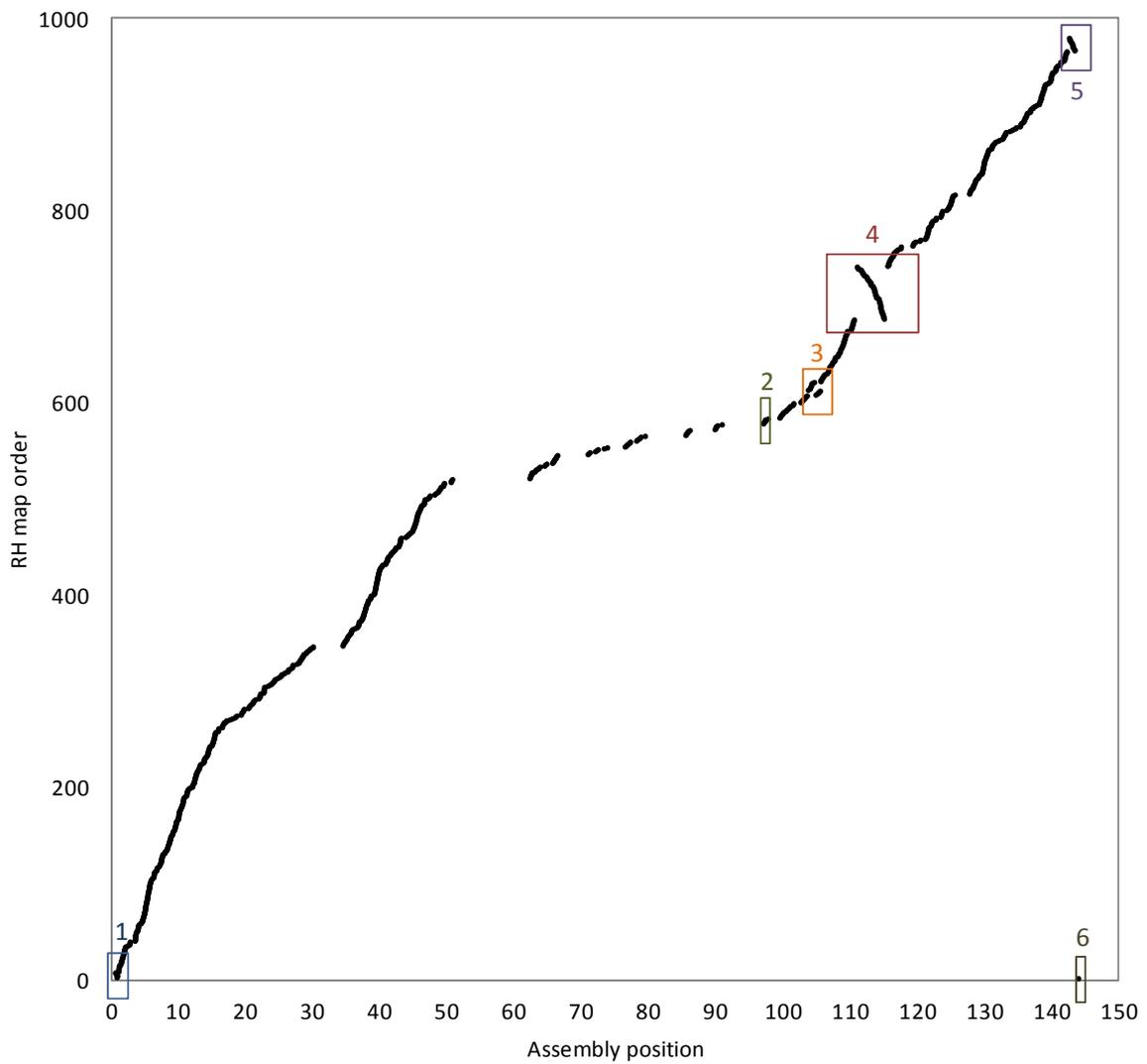


Figure A : comparaison entre la position des marqueurs SNP sur l'assemblage porcine et leur ordre d'après la carte RH robuste. Chaque point correspond à un marqueur SNP. L'axe des abscisses est la position sur l'assemblage porcine (v10.2) et l'axe des ordonnées l'ordre sur la carte RH. Les réarrangements sont indiqués par des groupes de couleur et des numéros.

## RESULTATS

### Carte physique du chromosome X porcin

*Carte physique complète* : Une carte physique du chromosome X porcin a été réalisée à l'aide des données de génotypage haute densité de la puce porcine d'Illumina (*porcine60SNP Beadchip*), obtenues pour les 90 clones du panel IMNpRH2<sub>12,000</sub> (Yerle et al., 2002). Une première carte dont la probabilité est très forte, dite robuste, a été construite à l'aide de CarthaGene (de Givry et al., 2005) selon une méthode proposée par Faraut et al. (Faraut et al., 2007), utilisant les vecteurs RH de ces 1088 marqueurs et leur position *a priori* sur l'assemblage porcin (version 10.2). La carte robuste (Servin et al., 2010b) a permis d'ordonner 979 des marqueurs (Annexe A). La comparaison entre la position sur l'assemblage porcin et l'ordre des marqueurs sur la carte robuste est présenté en figure A, et les détails de la carte sont regroupés dans l'annexe A.

L'alignement est globalement bien conservé, excepté pour six groupes de marqueurs. Le plus grand réarrangement est une inversion (n°4), concernant 56 marqueurs, localisés entre les positions 110,666 Mb et 115,647 Mb. La comparaison avec le génome humain dans cette région est en accord avec l'ordre de la carte robuste (résultats non présentés). Bien que des remaniements existent entre l'homme et le porc, la structure du (ou l'ordre des gènes) chromosome X est bien conservée sur l'ensemble de la séquence. De plus, aucune séquence porcine n'est chevauchante entre les extrémités de la potentielle inversion et les séquences encadrants cette inversion (cf contigs dans l'annexe A). Ces données apportent des éléments de validation de l'ordre proposé sur la carte RH. Concernant le réarrangement n°3, entre les marqueurs ASGA0081223 (position: 103,657 Mb) et MARC0067876 (position: 105,694 Mb), l'ordre de la carte RH est également en accord avec l'assemblage humain. Ce réarrangement peut être divisé en deux groupes de marqueurs, le premier concerne les marqueurs MARC0031940 (103,804 Mb) à ASGA0104371 (104,719 Mb), et le deuxième les marqueurs ASGA0081230 (104,939 Mb) à MARC0075831 (105,603 Mb). D'après les données de la carte RH, la séquence constituée du second groupe de marqueurs est localisée avant la séquence composée du premier groupe. Le réarrangement n°6 est composé d'un seul marqueur, ASGA0101795 en position 144,033 Mb d'après l'assemblage, et est localisé à l'autre extrémité du chromosome sur la carte RH. Les blocs n°1 et 5 correspondent aux marqueurs en extrémité des bras courts et longs, respectivement. Le bloc n°1 est composé de trois marqueurs (ALGA0101715, ASGA0090472, ALGA0116319, positions 0,670 à 0,783 Mb) et le bloc n°5 de 14 marqueurs, entre DRGA0017329 en position 142,269Mb et CAHM0000013 en position 143,502 Mb. Chacun des blocs est inversé d'après la carte RH. L'ordre RH est comparable à celui de l'homme pour ces deux blocs, et aucune séquence chevauchante n'est présente entre les blocs d'intérêts et le reste du chromosome pour justifier l'ordre de l'assemblage. Cependant, les extrémités des cartes RH sont souvent peu robustes car elles ne sont bornées par d'autres marqueurs que d'un seul côté. Les différents points de discordance entre l'ordre RH et l'ordre de l'assemblage tendent vers des erreurs de l'assemblage, au vu de l'homologie avec l'homme et de la robustesse des résultats.

Au-delà des erreurs potentielles détectées, la construction d'une carte RH complète a permis l'ordonnement de 1061 marqueurs SNP et 27 gènes au total, répartis en 25 groupes de liaison.

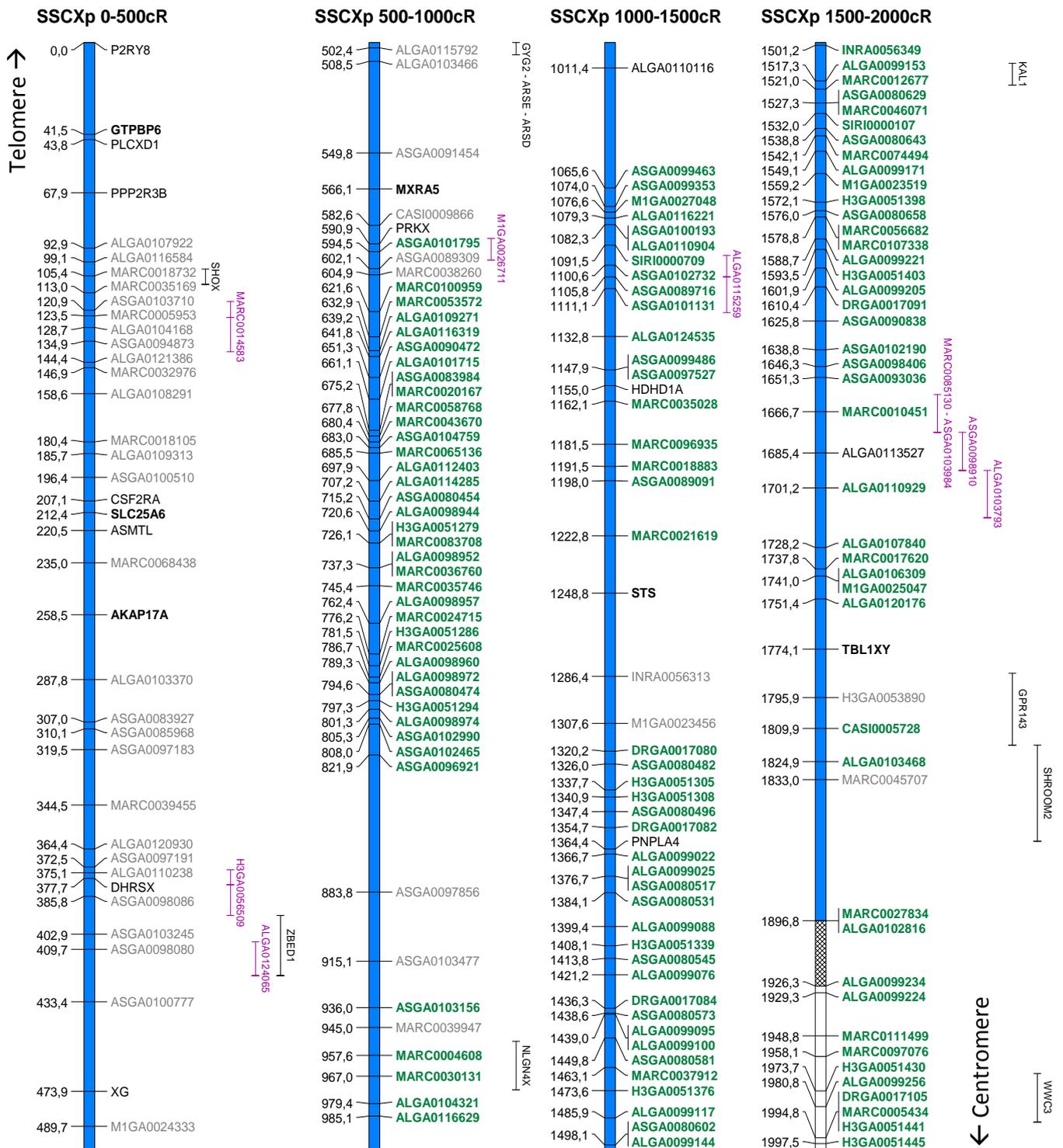


Figure B : Carte physique complète de la région pseudo-autosomale porcine. La région bleue correspond à la PAR et la région hachurée à la limite de la PAR. Les gènes en noir et les marqueurs en vert proviennent de la carte robuste. Les gènes et marqueurs verticaux possèdent plusieurs positions équiprobables. SSCXp : bras court du chromosome X porcine.

Près de cent marqueurs sans position liés aux SNP annotés sur le chromosome X ont pu être localisés sur le chromosome X. Une des applications de la cartographie RH est en effet la prédiction de position de ces marqueurs sans position sur l'assemblage (Servin et al., 2012).

*Région pseudo-autosomale* : Peu d'information étant disponible concernant la région pseudo-autosomale (PAR) porcine (Das et al., 2013; Quilter et al., 2002; Skinner et al., 2013), un des objectifs de ces travaux a été de définir la limite de la région pseudo-autosomale et la région spécifique du chromosome X, ainsi que le contenu en gènes de la PAR. Deux types de données ont permis d'estimer la position de la limite de la PAR porcine : la fréquence d'hétérozygotie mâle, et les groupes de liaison des marqueurs SNP obtenus avec les données RH. Les mâles sont porteurs d'un chromosome X et d'un chromosome Y, alors que les femelles possèdent deux chromosomes X. Cependant, les extrémités des chromosomes sexuels sont homologues (Helena Mangs and Morris, 2007) et s'apparient durant la méiose. Cette région est appelée région pseudo-autosomale (PAR). Les mâles possèdent également deux copies de la PAR, une sur le chromosome X (PARx) et une sur le chromosome Y (PARy). Ils peuvent donc être hétérozygotes dans cette région, alors qu'ils sont hémizygotés pour les marqueurs localisés sur la région spécifique du chromosome X.

La fréquence d'hétérozygotie mâle a été estimée à l'aide de 1950 mâles dont les données de génotypage pour les marqueurs de la puce porcine étaient disponibles au laboratoire, issus d'autres projets de recherche. La fréquence d'hétérozygotie est supérieure à zéro pour les SNP des deux premiers groupes de liaison de la carte RH (155 marqueurs) (Annexe A), en extrémité du bras court du chromosome X porcine. Pour les marqueurs des groupes 3 à 25, la fréquence d'hétérozygotie est nulle. D'après ces estimations, la limite de la PAR porcine est localisée entre les positions 6,811 Mb (ALGA0102816) et 7,075 Mb (ALGA0099234) (Figure B).

Les 27 EST porcins, sélectionnés par homologie avec les gènes de la PAR ancestrale aux mammifères (Das et al., 2009), sont répartis dans trois groupes de liaison : le n° 1, le n° 2 (définis précédemment comme PAR) et le n°3 (Annexe A). 17 gènes sont liés aux marqueurs du groupe n°1, 8 gènes sont liés aux marqueurs du groupe n°2 et deux gènes sont liés aux marqueurs du groupe n°3. La limite entre la PAR et la région spécifique du chromosome X est donc définie entre les gènes SHROOM2 et WWC3 (Figure C). Il est établi que l'ordre des gènes en extrémité du bras court du chromosome X est conservé chez les mammifères. Cependant, la limite entre la PAR et la XSR varie d'une espèce à l'autre : la PAR humaine ne mesure que 2,6 Mb, alors que la PAR canine est de 6,7 Mb.

D'après nos résultats, la limite de la PAR porcine est similaire à celle de la PAR canine (Figure C). L'ordre des gènes de la PAR porcine est identique à celui observé chez les autres espèces de mammifères (Das et al., 2009; Helena Mangs and Morris, 2007; Raudsepp and Chowdhary, 2008), excepté pour trois gènes : P2RY8, PLCXD1 et GTPBP6 (Figure C). L'ordonnement du gène P2RY8 en extrémité de carte peut être dû à un biais de la méthode. En effet, quand la probabilité de l'ordonnement est faible, les marqueurs sont généralement positionnés à tort en extrémité de carte RH. Concernant P2RY8, son taux de cassure avec le marqueur suivant est fort (34%%) et il est important de souligner que la

probabilité qu'il soit positionné en extrémité de carte est faible (LOD=8.2). Sa position est donc incertaine.

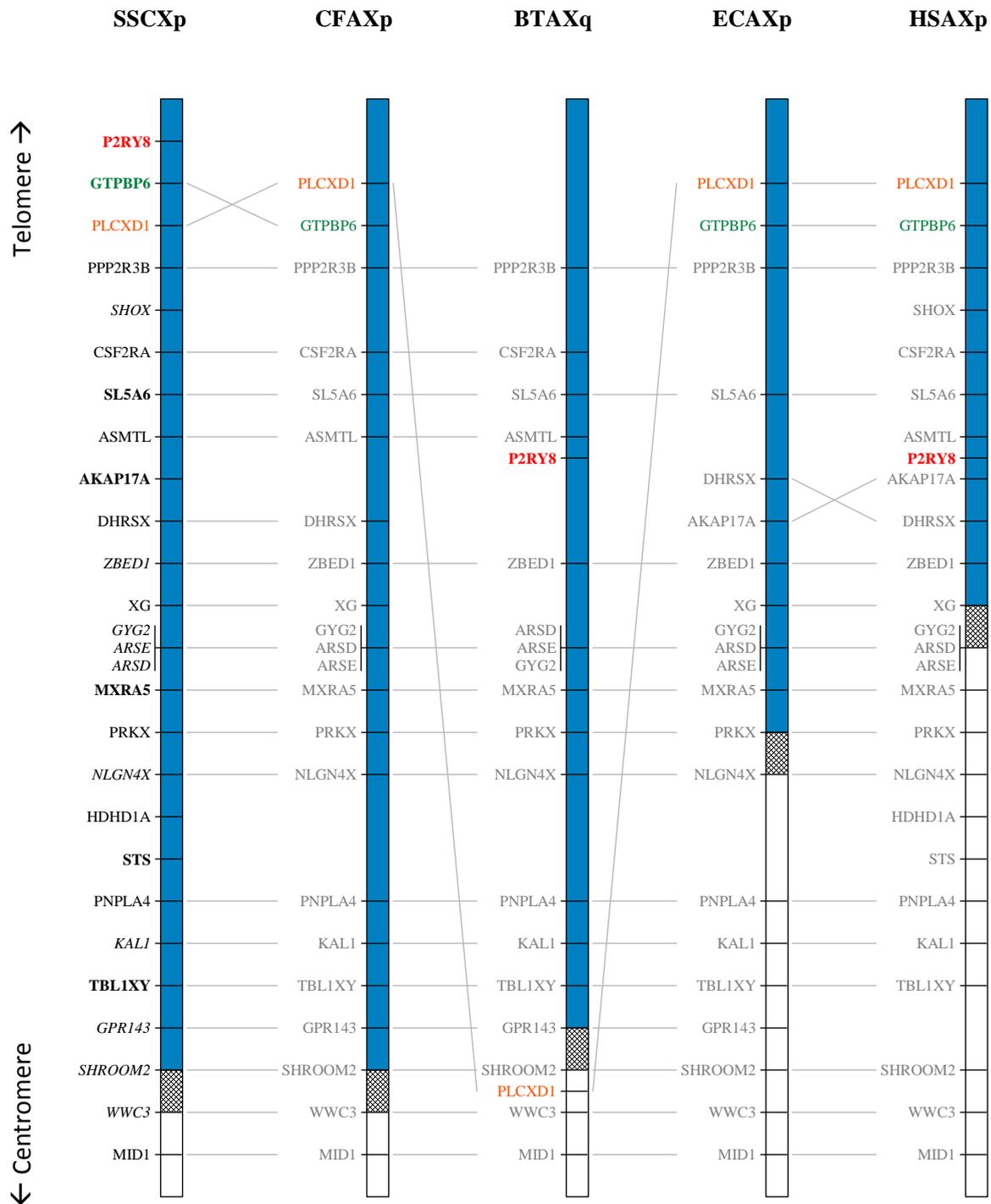


Figure C : Ordre des gènes sur la PAR chez le porc, le chien, le bovin, le cheval et l'homme. La région en bleu correspond à la PAR, et la région hachurée à la limite de la PAR. Les trois gènes dont l'ordre est différent entre le porc et les autres mammifères sont indiqués de couleur (rouge, vert et orange). SSCXp : bras court du chromosome X porc, CFAxp : bras court du chromosome X canin, BTAXq : bras long du chromosome X bovin, HSAXp : bras court du chromosome X humain.

Les vecteurs des deux gènes PLCXD1 et GTPBP6 ne diffèrent que pour un seul clone, les probabilités d'ordonnement entre ces deux gènes l'un par rapport à l'autre sont donc très similaires. Les ordres PLCXD1/GTPBP6 et GTPBP6/PLCXD1 sont quasi équiprobables.

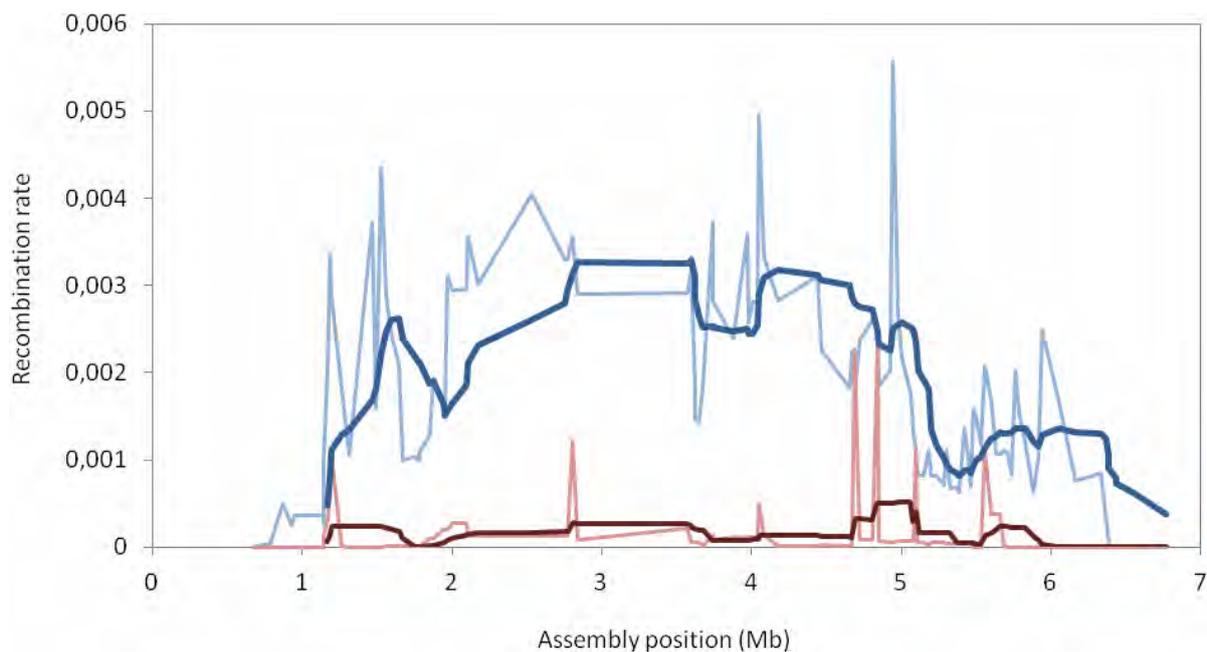


Figure D : Taux de recombinaison sur la PAR. Les courbes claires correspondent aux valeurs absolues du taux de recombinaison chez les mâles (bleu) et les femelles (rose), les courbes foncées sont les valeurs lissées.

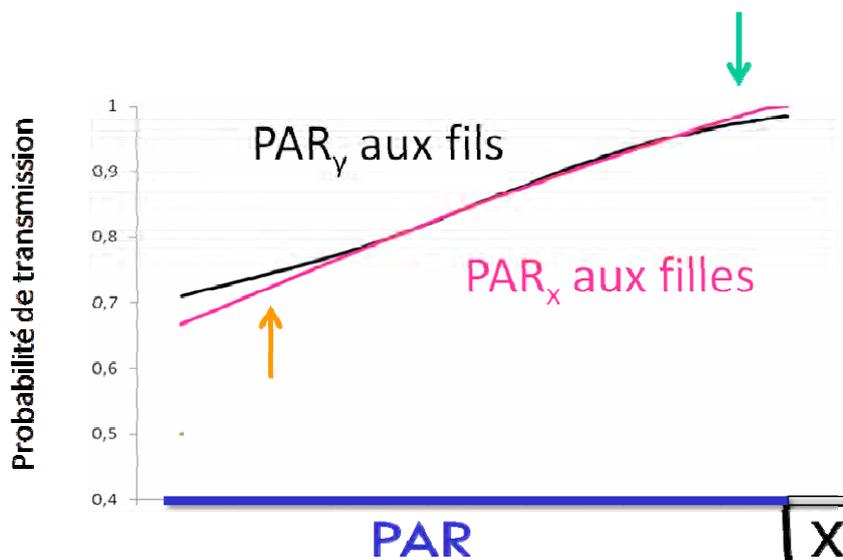


Figure E : Probabilité de transmission sur la PAR

## **Cartographie génétique du chromosome X**

*Région pseudo-autosomale porcine* : L'appariement des chromosomes sexuels durant la méiose est crucial pour la spermatogénèse, et est limité à la région pseudo-autosomale PAR1 (Kauppi et al., 2011). Un des objectifs de ces travaux est d'estimer le taux recombinaison de la PAR porcine. Les travaux de cartographie physique réalisés précédemment ont permis d'identifier les marqueurs localisés sur la PAR porcine à utiliser pour la construction de la carte génétique de la PAR. Le taux de recombinaison a été estimé à l'aide des 114 marqueurs des groupes de liaison 1 et 2 de la carte physique robuste, et à partir de 3456 animaux, correspondant à 882 méioses femelles et 2264 méioses mâles. La probabilité de recombinaison sur la PAR porcine est présentée en Figure D. Chez les mâles, le taux de recombinaison est cinq fois plus élevé au milieu de la PAR qu'aux extrémités (télomère et limite avec la région spécifique du chromosome X). La distance génétique de la PAR mâle est de 19,9 cM, et celle de la PAR femelle de 1,7 cM. Chaque mâle transmet son chromosome Y à ses descendants mâles et son chromosome X à ses descendantes femelles.

Sachant qu'un appariement est nécessaire au niveau de la PAR lors de la méiose, la transmission aux fils et aux filles des allèles initialement localisés sur les PAR des chromosomes X (PAR<sub>x</sub>) et des chromosomes Y (PAR<sub>y</sub>) chez les pères a été estimée. Pour chaque descendant mâle, la proportion d'allèles initialement localisés sur la PAR<sub>y</sub> a été estimée. Les résultats sont présentés en Figure E. Une distorsion de ségrégation importante est observée sur la PAR, plus on se rapproche de la limite entre la PAR et la XSR. Lorsque la probabilité de transmission est égale à 1 (flèche verte), cela signifie que tous les descendants mâles ont reçu l'allèle porté par la PAR<sub>y</sub> du père, et que toutes les descendantes femelles ont reçu l'allèle porté par la PAR<sub>x</sub> du père. Bien que cette région soit pseudo-autosomale, le mode de transmission n'est pas comparable à un autosome, où la probabilité de transmission d'un allèle vers un descendant est égale à 0,5. Même en extrémité télomérique, la probabilité est égale à 0,7 (flèche orange), et non 0,5.

*Région spécifique du chromosome X* : Les mâles ne possédant qu'un chromosome X, le taux de recombinaison sur la région spécifique du chromosome X ne peut être estimé que chez les femelles. Une carte génétique a donc été construite en utilisant les 882 méioses femelles et les 865 marqueurs SNP localisés sur la région spécifique du chromosome X et sélectionnés d'après la carte physique robuste (Figure F). La distance génétique de cette région XSR est de 148,8 cM. Le taux de recombinaison est très variable d'une région du chromosome à l'autre : il peut atteindre  $1,14 \cdot 10^{-2}$  (région entre 125 et 128 Mb), alors que d'autres régions ne recombinent pas. Une large région notamment présente un taux de recombinaison très faible, entre les positions 50 et 100 Mb approximativement. Une région aussi large avec un taux de recombinaison si faible n'a jamais été observée chez les autres espèces. Chez le porc, cette région de faible recombinaison avait été préalablement mise en évidence, mais à partir d'une carte génétique de plus faible résolution (Ma et al., 2010). Au vu de la position du centromère, autour duquel un faible taux recombinaison est observé (Jensen-Seaman et al., 2004), la région peu recombinante peut être divisée en deux régions : un « plateau » au niveau du centromère, dont le faible taux recombinaison peut être expliqué par la présence du

centromère, et une région de 40 Mb environ, pour laquelle le taux de recombinaison est proche de zéro, qui sera annotée NR (pour *Non-recombinant Region*).

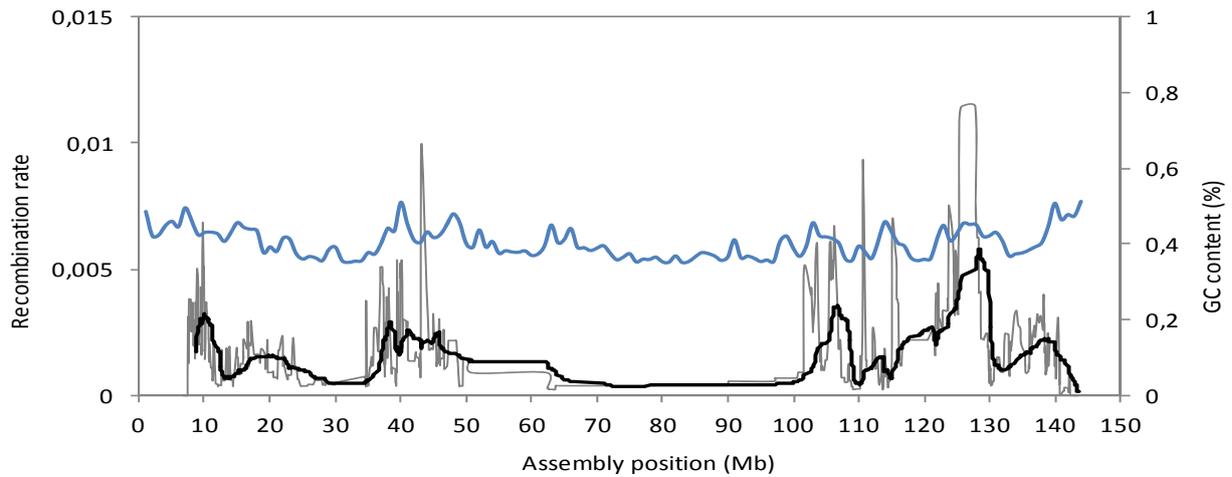


Figure F : Taux de recombinaison sur le chromosome X porcine (ligne gris clair pour les valeurs brutes, et foncée pour les valeurs lissées) et le pourcentage de GC (ligne bleue).

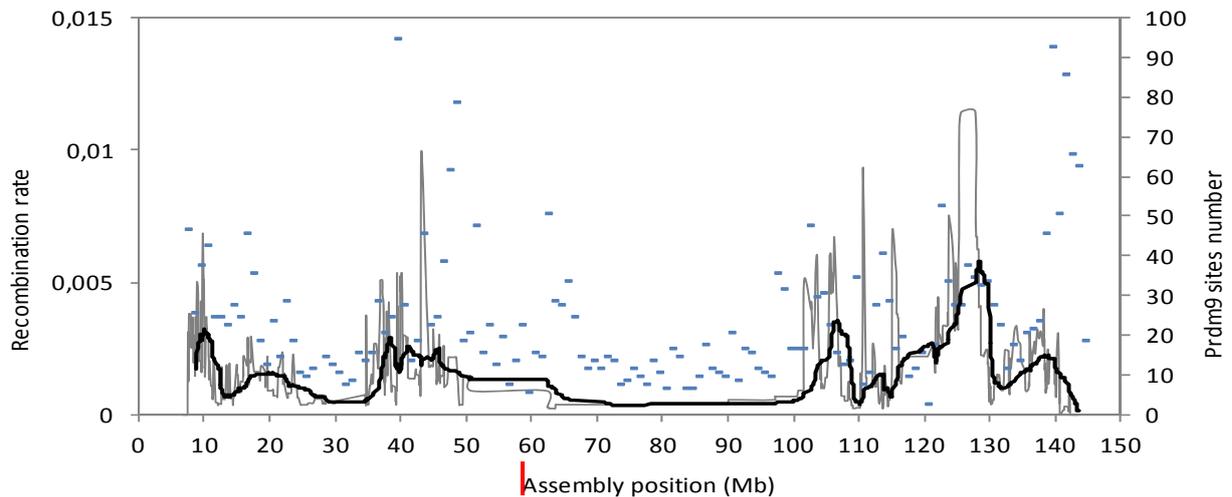


Figure G : Taux de recombinaison sur le chromosome X porcine (ligne gris clair pour les valeurs brutes, et foncée pour les valeurs lissées) et nombre de sites de fixation Prdm9 (traits bleus).

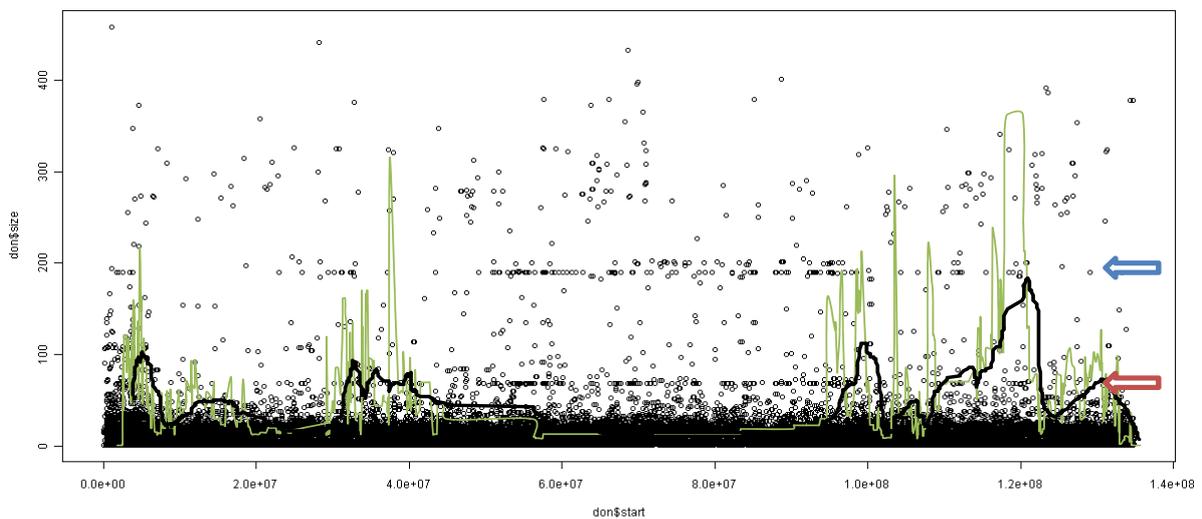


Figure H : Taille des motifs répétés en tandem selon leur position sur le chromosome X (un point par motif) et taux de recombinaison (courbe verte). La flèche bleue indique les motifs de 190 pb et la flèche rouge les motifs de 86 pb.

Afin d'identifier des mécanismes pouvant être impliqués ou associés à la NR, le contenu en séquence du chromosome X a fait l'objet d'études complémentaires. Il est établi que le taux de recombinaison est corrélé au pourcentage en GC et certains éléments répétés chez l'homme (Kong et al., 2002), la souris (Jensen-Seaman et al., 2004), le poulet (Groenen et al., 2009) et le chien (Wong et al., 2010). De plus, une protéine a été identifiée pour être impliquée dans le mécanisme de recombinaison, il s'agit de PRDM9 (Baudat et al., 2010). Un site de fixation de 13pb, CCNCCNTNNCCNC, de cette protéine a également été rapportée (Myers et al., 2008a).

#### *Contenu en séquence du chromosome X*

Le pourcentage de GC par fenêtres non chevauchantes d'une Mb, ainsi que le nombre de motifs CCNCCNTNNCCNC (Myers et al., 2008a) par fenêtre d'une Mb ont été estimés. Les résultats sont présentés en Figure F et Figure G. Au vu de ces résultats, le taux de GC reste compris entre 40 et 50% dans la région peu recombinante, et suit la tendance observée de part et d'autre de la région. Concernant le nombre de site de fixation de Prdm9 dans la région peu recombinante, il varie entre 10 et 35 motifs par Mb, alors que près de 95 motifs par Mb sont observés dans deux autres régions du chromosome.

Afin d'approfondir la caractérisation de la région, une recherche de motifs en tandem a été réalisée à l'échelle du chromosome avec TandemRepeatFinder (Benson, 1999). Ces analyses ont permis de mettre en évidence deux tailles de motifs particulièrement présents dans la NR : des motifs de 68 pb, et des motifs de 190 pb (Figure H). L'ensemble des motifs de 68 pb correspond à la même séquence, à quelques bases près. Les motifs de 190 pb représentent également la même séquence, à quelques différences nucléotidiques. L'alignement de ces deux types de motifs sur le génome de référence porcine a montré qu'ils avaient été identifiés dans la séquence 5'UTR d'un LINE-1. Dans ce LINE, ils sont également répétés en Tandem (le motif de 68 pb est répété en moyenne 3,8 fois, et le motif de 190 pb 4,5 fois (Sironen et al., 2007). Ces deux motifs sont séparés par un motif de 540 pb, qui n'est pas répété en tandem. Ce motif de 540 pb (appelé B) est également présent dans la séquence NR du chromosome X entre les motifs de 68 et 190 pb (appelés A et C, respectivement), mais il n'a pas été identifié par TandemRepeatFinder car il n'est pas répété en tandem.

Afin de savoir si ce motif surreprésenté dans la région peu recombinante était spécifique à cette région, ou associé à un faible taux de recombinaison à l'échelle du génome, le nombre de motifs ABC a été estimé par fenêtres non chevauchantes d'une Mb sur tous les chromosomes. A titre de comparaison, le nombre d'éléments dispersés connus chez le porc (type LINE, SINE, LTR, ...) présents dans la région peu recombinante et hors de la région a également été estimé. En parallèle, des corrélations ont été estimées entre le taux de recombinaison et l'occurrence des motifs (ABC et éléments dispersés). L'objectif de cette analyse était 1) de rechercher les motifs les plus corrélés, chez le porc, avec de faible taux de recombinaison, quel que soit le chromosome, 2) de rechercher les motifs sur représentés dans la région NR du chromosome X et 3) de combiner ces résultats afin d'identifier si un motif

généralement associé à de faible taux de recombinaison était particulièrement sur représenté dans la région NR du chromosome X.

## LINE-1 identifié chez le porc



Figure I : Comparaison des motifs identifiés en tandem dans la région peu recombinante du chromosome X, et ceux en 5'UTR d'un LINE-1 porcine (Sironen et al., 2007).

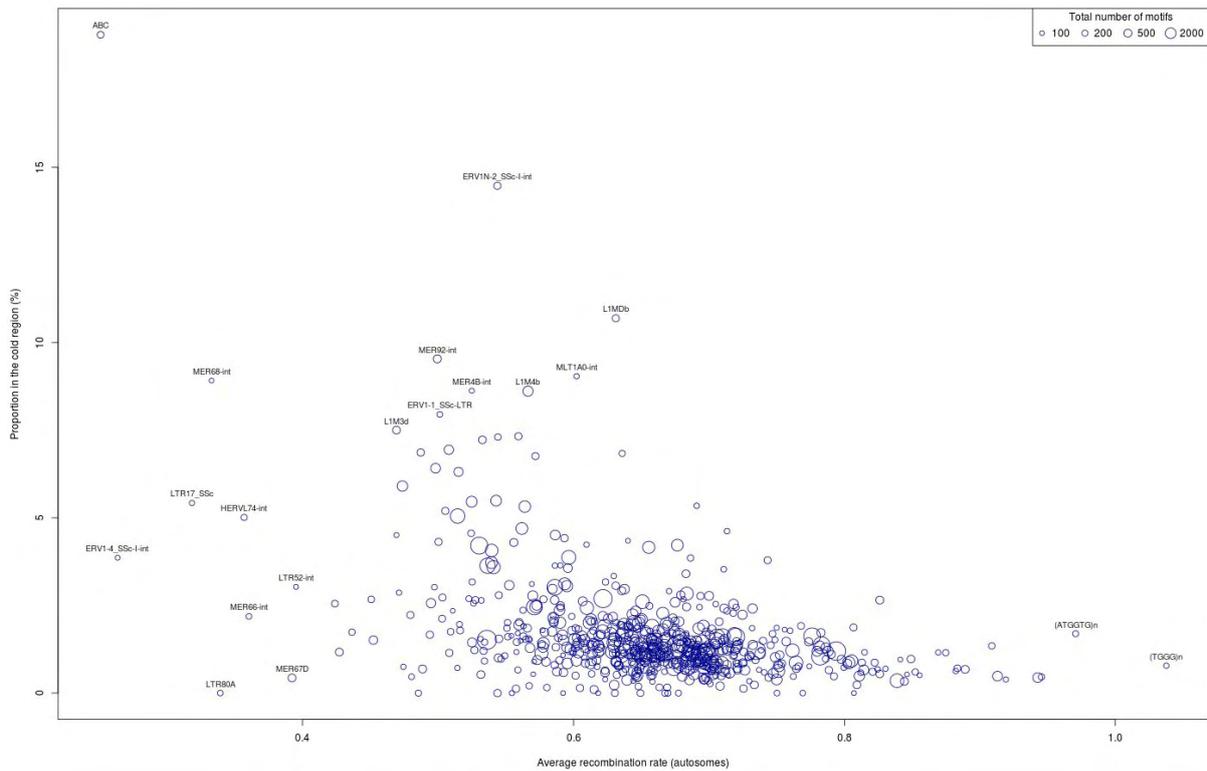


Figure J : Motifs répétés et taux de recombinaison à l'échelle du chromosome, et dans la région peu recombinante. Les motifs issus de l'annotation de RepeatMasker sont indiqués par des cercles bleus, dont le diamètre varie selon le nombre de motifs à l'échelle du génome. Seuls les motifs d'occurrence supérieure à 100 sont représentés. L'axe des X correspond au taux de recombinaison moyen pour chaque motif. L'axe des Y est la proportion de motifs dans la région 50-100Mb du chromosome X par rapport à son occurrence sur le reste du génome.

Les résultats de cette analyse sont présentés en Figure J. Seuls les motifs présents au moins 100 fois à l'échelle du génome ont été conservés pour les analyses, car ils doivent être suffisamment représentés pour qu'une association puisse être établie entre leur présence et le taux de recombinaison.

Au total, 656 motifs ont été identifiés selon ces critères, et la plupart des motifs (90%) ont été identifiés dans des régions du génome dont le taux de recombinaison est compris entre 0,5 et 0,8 (axe des abscisses) et sont répartis de façon uniforme entre la région NR du chromosome X et le reste du génome.

Seuls 4,4% (29 motifs) des motifs sont localisés dans des régions où le taux de recombinaison est inférieur à 0,5. Ces éléments correspondent aux catégories des LTR (ERV1, ERVL, ERVK), des LINE/L1 et le motif ABC. Le motif ABC est le motif le plus souvent associé à un faible taux de recombinaison. Chez l'homme les éléments les plus corrélés avec le faible taux de recombinaison sont les LTR/ERK (-0,29) et les LINE/L1 (-0,16) (Lee et al., 2011). Chez le porc, les éléments identifiés jusqu'à présent dans les points froids de recombinaison, dits *coldspots* sont de catégories LINE (-0,05) et *Low complexity* (duplications de motifs de 1 à 5 pb) (-0,26) (Tortereau et al., 2012). Cependant dans cette étude l'ensemble des LINE ont été analysés comme une seule famille, alors que dans notre étude le motif L1 de la famille des LINE est particulièrement surreprésenté.

Concernant l'occurrence des motifs, 96% des motifs identifiés sont présents moins de 5 fois dans la NR du chromosome X par rapport au reste du génome. Sachant que la région d'intérêt mesure 50Mb, et que le génome porcin compte près de 2,8 Gb de séquences, un motif, dont 2% des occurrences seraient identifiées dans la région NR, ne peut pas être considéré comme un motif surreprésenté. Seuls quelques motifs sont plus fréquents dans la région NR. Ils correspondent aux mêmes motifs identifiés comme associés à de faibles taux de recombinaison. Cependant, un seul motif est très largement surreprésenté : Le motif ABC est dix fois plus représenté dans la région peu recombinante du chromosome X que dans le reste du génome.



## DISCUSSION

Pour la première fois, une carte physique et génétique haute densité du chromosome X porcine a été réalisée. Les travaux de cartographie physique ont permis de vérifier l'ordre de l'assemblage de référence (version 10.2) et d'identifier quelques remaniements, notamment entre les positions 110 à 115 Mb, et 103 à 105 Mb. Trois marqueurs (ASGA0091454, CASI0009866 et ASGA0101795) sont annotés en extrémité télomérique du bras long sur l'assemblage, alors qu'ils sont localisés en extrémité télomérique du bras court sur la carte RH. Il est établi chez l'homme que les deux extrémités des chromosomes sexuels sont homologues et qu'elles correspondent aux PAR1 et PAR2. Cependant chez les autres mammifères, aucune PAR2 n'est annotée (bovin (Das et al., 2009), cheval (Raudsepp and Chowdhary, 2008), chien (Young et al., 2008)). D'après notre étude, les marqueurs annotés en extrémité du bras long du chromosome X chez le porc, qui auraient pu correspondre à la PAR2, sont également localisés dans la PAR1. Aucune séquence connue ne semble correspondre à une région PAR2 dans le génome porcine. Des travaux récents chez le porc ont pourtant montré par FISH qu'il existait une région homologue entre SSCX et SSCY en extrémité du bras long du chromosome X, potentielle PAR2 (Skinner et al., 2013). La séquence de référence porcine ne permet pas à ce jour d'identifier cette région.

Près de cent marqueurs sans position sur l'assemblage porcine de référence ont été positionnés sur le chromosome X porcine. L'utilisation de la cartographie physique pour intégrer des marqueurs sans position sur le génome de référence avait fait ses preuves dans les travaux de Servin et al. (Servin et al., 2012), où plus de 2000 marqueurs sans position sur l'assemblage avaient été localisés sur les autosomes porcins. Dans le cas du chromosome X, beaucoup moins riche en marqueurs SNP que les autosomes (1 SNP tous les 132 kb sur le chromosome X, contre 1 SNP tous les 2 kb sur les autosomes (Tortereau et al., 2012)), l'intégration de nouveaux marqueurs est particulièrement utile pour compléter la séquence manquante.

D'après la carte génétique de la PAR porcine, la distance génétique chez les mâles est dix fois plus importante que chez les femelles (19,9 cM chez les mâles contre 1,7 cM chez les femelles). Sur les autosomes porcins, les femelles ont un taux de recombinaison plus important que chez les mâles (Tortereau et al., 2012). Cette différence entre le sexe hétérogamétique et le sexe homogamétique est également observée chez d'autres espèces (Haldane, 1922). Le fait que le taux de recombinaison sur la PAR soit plus fort chez les mâles que chez les femelles peut être expliqué par la nécessité chez les mâles d'un appariement entre les chromosomes sexuels pour assurer la gamétogénèse. Or l'appariement ne peut survenir qu'au niveau de la PAR. Chez les femelles, la recombinaison peut survenir tout au long du chromosome X, ce qui peut expliquer que le taux de recombinaison sur la PAR soit si faible. A l'échelle du chromosome, le taux de recombinaison est de 0,95 cM/Mb. Cette valeur est légèrement supérieure à celle estimée sur les autosomes porcins (environ 0,8 cM/Mb). Pour les autosomes de taille comparable, tel que le chromosome 4 (143 Mb contre 144 Mb pour le chromosome X), le taux de recombinaison est de 0,67 cM/Mb.

Un tiers du chromosome X a un taux de recombinaison très faible. Cette région peu recombinante avait déjà été observée dans des études précédentes (Ma et al., 2010; Nagaraja et al., 1997), avec une résolution bien plus faible que celle utilisée dans cette étude. Il est

établi que le taux de recombinaison est réduit au niveau du centromères (Jensen-Seaman et al., 2004).

D'après les travaux de Ma et al., le centromère est localisé aux alentours de 57 Mb (Ma et al., 2010). La présence du centromère dans cette région pourrait expliquer un faible taux de recombinaison de part et d'autre du centromère, entre les positions 51 et 62 Mb environ. Cependant, la région entre les positions 62 et 101 Mb ne peut être expliquée par un effet centromérique.

Chez les mammifères, le taux de recombinaison est corrélé avec le pourcentage en GC et certains éléments répétés (Groenen et al., 2009; Jensen-Seaman et al., 2004; Kong et al., 2002; Wong et al., 2010). Ces corrélations ont également été montrées sur les autosomes porcins (Tortereau et al., 2012). La région peu recombinante n'est cependant pas expliquée par un taux très faible en GC ou une absence de sites de fixation de Prdm9. Cependant, l'étude du contenu en séquence du chromosome X a montré qu'un motif était surreprésenté dans les régions où le taux de recombinaison est faible, et plus particulièrement dans la région non recombinante du chromosome X. Ce motif est composé d'un élément de 540 pb, encadré par un ou plusieurs éléments de 68 pb en 5', et un ou plusieurs éléments de 190 pb en 3'. Ce motif de 800 pb ou plus a été identifié dans la région 5'UTR d'un élément transposable type LINE-1 identifié chez le porc (Sironen et al., 2007). La surreprésentation de ce motif dans la région non recombinante du chromosome X peut être une conséquence du faible taux de recombinaison. En effet, il est établi que les éléments dispersés s'accumulent dans les régions où le taux de recombinaison est très faible (Dolgin and Charlesworth, 2008). Ce mécanisme d'accumulation de séquences lorsque la recombinaison est faible est désigné par le Cliquet de Muller (Muller, 1964) et est l'un des processus les plus connus influençant l'évolution du chromosome Y (Gordo and Charlesworth, 2001). Le faible taux de recombinaison dans cette région pourrait expliquer l'accumulation de motifs répétés type ABC. Cependant, la proportion d'autres motifs dispersés accumulés dans cette région n'est pas aussi importante que la proportion du motif ABC.

L'hypothèse alternative est que la surreprésentation de ce motif serait la cause du faible taux de recombinaison de cette région. Le motif ABC correspond à la région 5'UTR d'une classe de retrotransposons, les LINE-1, qui sont associés à l'hétérochromatisation (Chow et al., 2010). Chez l'homme la densité en L1 sur le chromosome X est deux fois plus importante que sur les autosomes (Bailey et al., 2000), notamment au niveau de Xist, essentiel pour l'inactivation du chromosome X (Chow et al., 2005). Xist chez le porc est annoté en position 67,181 Mb, dans la région du chromosome X dont le taux de recombinaison est faible. Les L1 joueraient un rôle de "*booster stations*" de Xist, c'est à dire qu'ils faciliteraient la propagation du signal Xist le long du chromosome X. Chez l'homme, la région de 17 Mb encadrant la position de Xist a un taux de recombinaison très faible (Jensen-Seaman et al., 2004) et est très riche en motifs L1.

Néanmoins, la présence du motif ABC, région 5'UTR d'un LINE-1, est plus corrélée avec le taux de recombinaison que la présence d'autres motifs L1. Cela suppose que le motif a un rôle privilégié dans le génome pour être conservé. Chez l'homme, les études ont montré que la région 5'UTR des LINE L1 comportait un promoteur sens, et un promoteur anti-sens (ASP) (Speek, 2001). Le promoteur anti-sens permet d'activer la transcription de nombreux gènes dans le génome (Alexandrova et al., 2012). Chez le porc, un ASP a été prédit dans la région

5'UTR du LINE-1 entre les positions 691 et 941 pb (Sironen et al., 2007) (motif B dans nos travaux).

D'après les travaux de Sironen et al., de nombreux sites de fixation de facteurs de transcription ont été identifiés dans cette région par méthode prédictive. Sur modèle du mécanisme d'initiation de recombinaison induit par PRDM9 se fixant sur un motif particulier (Myers et al., 2008b), il est possible qu'un mécanisme similaire intervienne dans l'inhibition de la recombinaison, par reconnaissance d'un motif de fixation spécifique qui pourrait être localisé dans ou à proximité du motif ABC.

Nos résultats apportent des données nouvelles concernant le chromosome X. D'une part, une cartographie physique haute densité du chromosome X a été réalisée pour la première fois chez le porc, apportant des informations précieuses quant à la qualité du draft porcin sur ce chromosome, et la caractérisation génétique de la région pseudo-autosomale porcine. D'autre part, les études génétiques ont montré une ségrégation très particulière de la région pseudo-autosomale, dont la prise en compte dans les études d'association et de liaison est nécessaire, et la présence d'une région représentant un tiers du chromosome X qui recombine très peu chez le porc. Des études complémentaires ont permis d'identifier un motif de 800 pb environ, très corrélé avec le faible taux de recombinaison. L'approfondissement de ces résultats pourrait permettre de comprendre les mécanismes impliqués dans la présence des *coldspots* dans le génome, jusqu'alors peu connus.



## ANNEXE A

Marker	position (cR)	break (%)	LOD	Linkage group	Chr.	position (pb)	Robust map order	rearrangement group	Clone	Male heterozygoty
P2RY8	0	34	8,2	1						
GTPBP6	41,5	2,3	24,6	1						
PLCXD1	43,8	21,4	13,5	1						
PPP2R3B	67,9	22,1	12,9	1						
ALGA0107922	92,9	6,1	21,4	1	U	334				0,36
ALGA0116584	99,1	6,1	21,6	1	U	475				0,40
MARC0018732	105,4	7,3	20,1	1	X	462585				0,18
MARC0035169	113	7,6	19,5	1	X	590511				0,45
ASGA0103710	120,9	2,6	22,7	1	U	1023				0,27
MARC0005953	123,5	5,1	18,2	1	U	6067703				0,24
ALGA0104168	128,7	6,1	18	1	U	6065622				0,35
ASGA0094873	134,9	9	19,5	1	U	1601058				0,40
ALGA0121386	144,4	2,5	23,2	1	U	6061873				0,32
MARC0032976	146,9	11	16,9	1	U	6099682				0,18
ALGA0108291	158,6	19,6	12,7	1	U	1597336				0,38
MARC0018105	180,4	5,1	20,1	1	U	1663311				0,23
ALGA0109313	185,7	10,2	17,6	1	U	10035671				0,22
ASGA0100510	196,4	10,2	17,6	1	U	20526336				0,14
CSF2RA	207,1	5,1	20,8	1						
SLC25A6	212,4	7,8	18,5	1						
ASMTL	220,5	13,5	14,9	1						
MARC0068438	235	20,9	12,8	1	U	1784				0,00
AKAP17A	258,5	25,4	10,6	1						
ALGA0103370	287,8	17,5	13,7	1	U	274				
ASGA0083927	307	3	23,4	1	U	792				0,89
ASGA0085968	310,1	8,9	20,5	1	U	829				0,66
ASGA0097183	319,5	22,1	13,1	1	U	27226784				0,13
MARC0039455	344,5	18,1	13,7	1	U	19951269				0,31
ALGA0120930	364,4	7,8	18,4	1	U	20001353				0,28
ASGA0097191	372,5	2,6	20,7	1	U	12755266				0,29
ALGA0110238	375,1	2,6	20,9	1	U	15935581				0,24
DHRX	377,7	7,8	17,8	1						
ASGA0098086	385,8	15,7	13,8	1	U	12737566				0,31
ASGA0103245	402,9	6,6	18,1	1	U	10229205				0,18
ASGA0098080	409,7	21,1	13	1	U	942				0,29
ASGA0100777	433,4	33,3	8,8	1	U	27160306				0,44
XG	473,9	14,6	15,6	1						
M1GA0024333	489,7	12	14,5	1	U	22092340				0,45
ALGA0115792	502,4	5,8	16,7	1	U	15653456				0,44
ALGA0103466	508,5	33,9	8,8	1	U	276				0,18
ASGA0091454	549,8	15	15,8	1	X	143833542		6bis	CU928462.11	0,34
MXRA5	566,1	15,2	15,2	1						
CASI0009866	582,6	7,9	17,8	1	X	144039023		6bis	CU928353.3	0,40
PRKX	590,9	3,5	20,7	1						
ASGA0101795	594,5	7,3	18,4	1	X	144033871	1	6	CU928353.3	0,44
ASGA0089309	602,1	2,8	18,9	1	U	18528979				0,37
MARC0038260	604,9	15,4	13,4	1	U	1664				0,42
MARC0100959	621,6	10,7	17,4	1	X	865659	2			0,19
MARC0053572	632,9	6,1	19,3	1	X	927637	3			0,47
ALGA0109271	639,2	2,6	21,7	1	X	947241	4		FP015888.26	0,35
ALGA0116319	641,8	9,1	18,6	1	X	783067	5	1	FP102150.5	0,47
ASGA0090472	651,3	9,3	17,3	1	X	692756	6	1	FP102150.5	0,37
ALGA0101715	661,1	13,2	15,2	1	X	670786	7	1	FP102150.5	0,43
ASGA0083984	675,2	0	24,6	1	X	1141455	8		FP015839.7	0,36
MARC0020167	675,2	2,6	21,9	1	X	1142159	9			0,35
MARC0058768	677,8	2,6	22,2	1	X	1144870	10			0,27
MARC0043670	680,4	2,5	22,9	1	X	1163540	11			0,52
ASGA0104759	683	2,5	22,9	1	X	1184375	12			0,45
MARC0065136	685,5	11,7	17,1	1	X	1194955	13			0,13
ALGA0112403	697,9	8,8	18,3	1	X	1259632	14			0,36
ALGA0114285	707,2	7,7	18,9	1	X	1313392	15			0,45

ASGA0080454	715,2	5,2	19,7	1	X	1462758	16		0,44
ALGA0098944	720,6	5,3	19,2	1	X	1487899	17		0,45
H3GA0051279	726,1	0	21,7	1	X	1558589	19		0,25
MARC0083708	726,1	10,6	16,8	1	X	1528885	18		0,41
ALGA0098952	737,3	0	24,1	1	X	1645943	21		0,46
MARC0036760	737,3	7,8	18,5	1	X	1591152	20		0,27
MARC0035746	745,4	15,6	14,7	1	X	1666702	22		0,30
ALGA0098957	762,4	12,9	16,2	1	X	1747673	23		0,27
MARC0024715	776,2	5,1	20,6	1	X	1778709	24		0,36
H3GA0051286	781,5	5,1	19,7	1	X	1781663	25		0,39
MARC0025608	786,7	2,6	21,6	1	X	1852374	26		0,37
ALGA0098960	789,3	5,1	20,1	1	X	1872325	27		0,18
ALGA0098972	794,6	0	24	1	X	1954776	29		0,12
ASGA0080474	794,6	2,6	22	1	X	1943169	28		0,43
H3GA0051294	797,3	3,9	21,6	1	X	1968265	30		0,35
ALGA0098974	801,3	3,9	21,6	1	X	2004788	31		0,18
ASGA0102990	805,3	2,6	21	1	X	2095311	32		0,40
ASGA0102465	808	13	15,4	1	X	2108186	33		0,41
ASGA0096921	821,9	46,1	5	1	X	2174683	34		0,11
ASGA0097856	883,8	26,9	10,1	1	X	2486601			0,10
ASGA0103477	915,1	18,8	13,4	1	U	1017			0,06
ASGA0103156	936	8,7	19	1	X	2530828	35		0,37
MARC0039947	945	11,8	17,4	1	U	1675			0,41
MARC0004608	957,6	9	17,3	1	X	2754871	36		0,40
MARC0030131	967	11,7	16,8	1	X	2770561	37		0,00
ALGA0104321	979,4	5,5	19,9	1	X	2807565	38		0,37
ALGA0116629	985,1	23,2	12,1	1	X	2838767	39		0,14
ALGA0110116	1011,4	41,8	7,4	1	U	366			0,47
ASGA0099463	1065,6	8,1	17,5	2	X	3567624	40		0,47
ASGA0099353	1074	2,6	20,4	2	X	3592827	41		0,41
M1GA0027048	1076,6	2,6	20,1	2	X	3600043	42		0,61
ALGA0116221	1079,3	3	20,4	2	X	3608787	43		0,44
ASGA0100193	1082,3	0	24,2	2	X	3619744	44		0,42
ALGA0110904	1082,3	8,8	18,8	2	X	3623792	45		0,46
SIRI0000709	1091,5	8,7	19,2	2	X	3650232	46		0,32
ASGA0102732	1100,6	5,1	20,7	2	X	3677084	47		0,38
ASGA0089716	1105,8	5,2	19	2	X	3739006	48		0,45
ASGA0101131	1111,1	19,5	12,7	2	X	3744916	49		0,28
ALGA0124535	1132,8	14	16,7	2	X	3876373	50		0,36
ASGA0099486	1147,9	0	23	2	X	3970981	51		0,46
ASGA0097527	1147,9	13,2	15,9	2	X	3979293	52		0,43
MARC0035028	1162,1	17,7	14,1	2	X	4006006	53		0,23
MARC0096935	1181,5	9,5	18,7	2	X	4043888	54		0,02
MARC0018883	1191,5	6,3	20,1	2	X	4051867	55		0,46
ASGA0089091	1198	22	13,8	2	X	4078589	56		0,43
MARC0021619	1222,8	22,8	12,7	2	X	4173588	57		0,15
STS	1248,8	31,3	8,6	2					
INRA0056313	1286,4	19,1	13,4	2	U	1423			0,00
M1GA0023456	1307,6	11,9	17,4	2	U	1455			0,25
DRGA0017080	1320,2	5,6	21,3	2	X	4444258	58		0,46
ASGA0080482	1326	11,1	18,6	2	X	4468500	59		0,44
H3GA0051305	1337,7	3,2	23,7	2	X	4651387	60		0,47
H3GA0051308	1340,9	6,3	22,2	2	X	4666913	61		0,35
ASGA0080496	1347,4	7	21,1	2	X	4691395	62		0,26
DRGA0017082	1354,7	9,3	19	2	X	4725085	63		0,44
PNPLA4	1364,4	2,3	23,4	2					
ALGA0099022	1366,7	9,4	18,7	2	X	4806364	64		0,12
ALGA0099025	1376,7	0	25,9	2	X	4838407	65		0,45
ASGA0080517	1376,7	7,2	20,1	2	X	4853651	66		0,47
ASGA0080531	1384,1	14,1	16,5	2	X	4927032	67		0,44
ALGA0099088	1399,4	8,3	20,6	2	X	4942454	68		0,44
H3GA0051339	1408,1	5,6	21,3	2	X	4977485	69		0,42
ASGA0080545	1413,8	7,1	19,7	2	X	5001421	70		0,44
ALGA0099076	1421,2	14	17,5	2	X	5061093	71		0,23
DRGA0017084	1436,3	2,3	25,2	2	X	5079013	72		0,47
ASGA0080573	1438,6	0,4	26	2	X	5095480	73		0,36
ALGA0099095	1439	0	25,2	2	X	5103986	74		0,41
ALGA0099100	1439	10,3	17,7	2	X	5116100	75		0,15

ASGA0080581	1449,8	12,5	16,6	2	X	5147277	76		0,07
MARC0037912	1463,1	9,9	19	2	X	5183703	77		0,27
H3GA0051376	1473,6	11,6	17,9	2	X	5200314	78		0,43
ALGA0099117	1485,9	11,5	19	2	X	5230043	79		0,13
ASGA0080602	1498,1	0	25,5	2	X	5283450	80		0,50
ALGA0099144	1498,1	3,1	22,9	2	X	5304422	81		0,37
INRA0056349	1501,2	14,9	19	2	X	5326424	82		0,33
ALGA0099153	1517,3	3,6	23,6	2	X	5371409	83		0,17
MARC0012677	1521	6,1	21,8	2	X	5385943	84		0,42
ASGA0080629	1527,3	0	25,5	2	X	5419281	85		0,44
MARC0046071	1527,3	4,6	22,9	2	X	5444729	86		0,44
SIRI0000107	1532	6,5	21,4	2	X	5464915	87		0,38
ASGA0080643	1538,8	3,3	23,3	2	X	5484206	88		0,44
MARC0074494	1542,1	6,7	19,8	2	X	5496330	89		0,37
ALGA0099171	1549,1	9,6	17,3	2	X	5528883	90		0,46
M1GA0023519	1559,2	12,1	19,2	2	X	5557301	91		0,39
H3GA0051398	1572,1	3,9	22,9	2	X	5600087	92		0,46
ASGA0080658	1576	2,7	22,7	2	X	5636073	93		0,38
MARC0056682	1578,8	0	24,6	2	X	5657660	94		0,00
MARC0107338	1578,8	9,4	18,8	2	X	5682684	95		0,43
ALGA0099221	1588,7	4,7	22	2	X	5708949	96		0,45
H3GA0051403	1593,5	8,1	20,3	2	X	5737743	97		0,18
ALGA0099205	1601,9	8,1	20,9	2	X	5759751	98		0,27
DRGA0017091	1610,4	14,3	16,9	2	X	5782952	99		0,36
ASGA0090838	1625,8	12,2	17,5	2	X	5828771	100		0,44
ASGA0102190	1638,8	7,3	20,3	2	X	5881619	101		0,16
ASGA0098406	1646,3	4,9	21,6	2	X	5916542	102		0,46
ASGA0093036	1651,3	14,3	16,9	2	X	5937610	103		0,44
MARC0010451	1666,7	17	16	2	X	6049852	104		0,37
ALGA0113527	1685,4	14,6	15,9	2	X	6079959			0,24
ALGA0110929	1701,2	23,7	12,4	2	X	6160267	105		0,46
ALGA0107840	1728,2	9,1	20,4	2	X	6328557	106		0,52
MARC0017620	1737,8	3,1	23,4	2	X	6364521	107		0,26
ALGA0106309	1741	0	25,5	2	X	6382577	108		0,46
M1GA0025047	1741	10	21,3	2	X	6387129	109		0,41
ALGA0120176	1751,4	20,3	14,1	2	X	6432615	110		0,12
TBL1XY	1774,1	19,6	13,6	2					
H3GA0053890	1795,9	13	16,4	2	X	6433914	111		0,39
CASIO005728	1809,9	14	16,6	2	X	6545438	112		0,02
ALGA0103468	1824,9	7,7	19,1	2	X	6769971	113		0,55
MARC0045707	1833	47,2	5,8	2	X	6798065	114		0,46
MARC0027834	1896,8	0	25,2	2	X	6878442	116		
ALGA0102816	1896,8	25,6	11,7	2	X	6811166	115		0,51
ALGA0099234	1926,3	2,9	20,5	3	X	7075307	117		0,00
ALGA0099224	1929,3	17,7	13,3	3	X	7185187	118		0,00
MARC0111499	1948,8	8,9	17,1	3	X	7269183	119		0,00
MARC0097076	1958,1	14,4	15,9	3	X	7285934	120		0,00
H3GA0051430	1973,7	6,9	19,5	3	X	7407981	121		0,00
ALGA0099256	1980,8	13	15,7	3	X	7438484	122		0,00
DRGA0017105	1994,8	0	23,3	3	X	7505070	125		0,00
MARC0005434	1994,8	0	23	3	X	7497664	124		0,00
H3GA0051441	1994,8	2,7	20,5	3	X	7461488	123		0,00
H3GA0051445	1997,5	11,8	15,9	3	X	7515894	126		0,00
ASGA0080716	2010,1	12	16,1	3	X	7608878	127		0,00
H3GA0051448	2022,9	3	20,4	3	X	7638707	128		0,00
H3GA0051451	2025,9	0	22,8	3	X	7706280	130		0,00
ASGA0080719	2025,9	5,3	21,2	3	X	7674339	129		0,00
MID1	2031,4	23,8	11,6	3					
ASGA0080725	2058,5	6	17,8	3	X	7979887	131		0,00
MARC0032319	2064,7	0	22,5	3	X	8055841	133		0,00
M1GA0023548	2064,7	10,6	16,3	3	X	8008220	132		0,00
ASGA0080726	2075,9	15,9	14,4	3	X	8264173	134		0,00
ALGA0099275	2093,3	8	17,8	3	X	8329500	135		0,00
ASGA0080734	2101,6	0	23,6	3	X	8407407	137		0,00
ALGA0099279	2101,6	7,8	18	3	X	8364141	136		0,00
M1GA0023551	2109,8	5,2	18,8	3	X	8447070	138		0,00
ALGA0099283	2115,2	2,6	20,5	3	X	8475276	139		0,00
H3GA0051467	2117,8	13,5	15	3	X	8547577	140		0,00

H3GA0051468	2132,3	7,3	18,5	3	X	8636250	141		0,00
ASGA0080745	2139,9	3,5	20,8	3	X	8655152	142		0,00
ALGA0099288	2143,5	4	20,8	3	X	8692219	143		0,00
ASGA0080749	2147,6	4,1	20,4	3	X	8718698	144		0,00
INRA0056407	2151,8	8,3	15,8	3	X	8818096	145		0,00
INRA0056408	2160,4	2,8	18,2	3	X	8831742	146		0,00
ASGA0080751	2163,2	11,4	14	3	X	8850003	147		0,00
ALGA0099290	2175,3	4,3	17,8	3	X	8887538	148		0,00
MARC0089538	2179,7	4,3	17,4	3	X	8938817	149		0,00
DRGA0017113	2184,1	8,6	14,9	3	X	9095567	150		0,00
ALGA0099297	2193,1	2,9	18,1	3	X	9142750	151		0,00
ASGA0092320	2196	2,9	18,1	3	X	9164433	152		0,00
ASGA0104911	2198,9	9,5	14,3	3	X	9226461	153		0,00
ASGA0104040	2208,9	25,2	9,9	3	X	9261433	154		0,00
ASGA0095424	2237,9	0	21,4	3	X	9474764	156		0,00
ALGA0119789	2237,9	2,7	20,3	3	X	9410480	155		0,00
ASGA0095102	2240,7	3,6	20,4	3	X	9505261	157		0,00
ASGA0097714	2244,3	7,3	18,5	3	X	9513017	158		0,00
ASGA0099808	2251,9	2,7	20,5	3	X	9566999	159		0,00
ALGA0099301	2254,7	7,3	18,5	3	X	9609979	160		0,00
M1GA0023561	2262,2	3,6	20,4	3	X	9650001	161		0,00
ALGA0099305	2265,9	2,7	20,4	3	X	9671130	162		0,00
MARC0063719	2268,6	2,7	19,7	3	X	9684933	163		0,00
H3GA0051482	2271,4	5,5	18,4	3	X	9724808	164		0,00
ASGA0080764	2277	5,5	17,5	3	X	9823731	165		0,00
ALGA0099314	2282,7	8,3	16,2	3	X	9970257	166		0,00
MARC0040504	2291,4	4,1	20,4	3	X	10005053	167		0,00
ALGA0099322	2295,5	4,1	20,4	3	X	10032407	168		0,00
MARC0062773	2299,7	2,7	20,3	3	X	10059758	169		0,00
M1GA0023567	2302,4	8,1	16,4	3	X	10075688	170		0,00
H3GA0051494	2310,9	8	16,9	3	X	10093556	171		0,00
ALGA0099333	2319,2	0	23,7	3	X	10134168	173		0,00
M1GA0023569	2319,2	5,3	19,6	3	X	10120828	172		0,00
H3GA0051499	2324,7	5,4	19,3	3	X	10145695	174		0,00
ASGA0080788	2330,2	8,3	16,8	3	X	10215451	175		0,00
H3GA0051503	2338,8	11	15,7	3	X	10282739	176		0,00
H3GA0053490	2350,5	0	22,8	3	X	10390751	178		0,00
ASGA0080789	2350,5	5,4	18,4	3	X	10346377	177		0,00
MARC0102536	2356	2,8	19,5	3	X	10504345	179		0,00
ASGA0103293	2358,8	3,4	20	3	X	10506715	180		0,00
ASGA0092384	2362,3	10,3	16,7	3	X	10561004	181		0,00
H3GA0052195	2373,2	6,4	17,4	3	X	10653783	182		0,00
H3GA0051511	2379,8	0	21,5	3	X	10712229	184		0,00
H3GA0051510	2379,8	13,1	14,2	3	X	10683401	183		0,00
MARC0090683	2393,8	8,4	15,8	3	X	10735656	185		0,00
ASGA0080794	2402,6	4,1	19,4	3	X	10756156	186		0,00
DRGA0017114	2406,8	4,2	19,1	3	X	10776985	187		0,00
ASGA0080795	2411,1	6,5	17,5	3	X	10810539	188		0,00
INRA0056441	2417,8	28,8	10,4	3	X	10875335	189		0,00
ASGA0080798	2451,8	16,3	12,7	3	X	11060922	190		0,00
ASGA0080800	2469,7	13,8	13,2	3	X	11234223	191		0,00
H3GA0051520	2484,6	8,3	16,2	3	X	11260891	192		0,00
CASI0001318	2493,2	5,8	17,5	3	X	11273819	193		0,00
INRA0056448	2499,2	40	7,7	3	X	11292127	194		0,00
ASGA0080809	2550,2	2,8	23	3	X	11345780	195		
DRGA0017115	2553,1	37,6	8,5	3	X	11384380	196		
MARC0090477	2600,2	5,5	17,2	3	X	11443525	197		0,00
MARC0033471	2605,9	14,7	14	3	X	11547719	198		0,00
H3GA0051533	2621,8	23,8	10,7	3	X	11783693	199		0,00
CAHM0000159	2648,9	0	20,8	3	X	12118563	201		0,00
INRA0056454	2648,9	3	18,3	3	X	12108725	200		0,00
ALGA0099375	2651,9	8,3	16,3	3	X	12129907	202		0,00
ALGA0099376	2660,6	8,2	14,5	3	X	12250158	203		0,00
ALGA0099377	2669,1	8	16,3	3	X	12368988	204		0,00
H3GA0051536	2677,5	0	22,9	3	X	12411027	206		0,00
ASGA0080825	2677,5	7,7	19,2	3	X	12396227	205		0,00
MARC0031765	2685,5	2,6	22,5	3	X	12454020	207		0,00
MARC0016228	2688,1	14	16,5	3	X	12467952	208		0,00

MARC0060456	2703,2	7	20,5	3	X	12508028	209			0,00
ALGA0099380	2710,5	0	22,7	3	X	12551969	210			0,00
H3GA0051539	2710,5	7,1	19	3	X	12624276	211			0,00
CASIO007528	2717,9	3,4	21,7	3	X	12634044	212			0,00
ASGA0080826	2721,4	3,1	21,7	3	X	12659659	213			0,00
ALGA0099383	2724,5	12,7	17,2	3	X	12702717	214			0,00
INRA0056460	2738,1	0	22,7	3	X	12850075	216			0,00
M1GA0023581	2738,1	11,7	15,4	3	X	12824275	215			0,00
CAILO000240	2750,5	23,4	11,7	3	X	12899541	217			0,00
MARC0047046	2777,2	13,6	14,8	3	X	13061465	218			0,00
ASGA0080840	2791,8	19,3	12,2	3	X	13085711	219			0,00
ASGA0080850	2813,3	8	16,8	3	X	13152222	220			0,00
H3GA0051544	2821,7	0	22,7	3	X	13245808	222			0,00
ALGA0099399	2821,7	16,3	13,7	3	X	13190474	221			0,00
DIAS0003540	2839,4	22,3	10,5	3	X	13284646	223			0,00
ALGA0104407	2864,7	21,5	11,7	3	X	13448866	224			0,00
H3GA0051546	2888,9	20,9	12,6	3	X	13692150	225			0,00
ALGA0099407	2912,4	0	22,7	3	X	13872249	228			0,00
ASGA0080854	2912,4	0	23,7	3	X	13860513	227			0,00
ALGA0099404	2912,4	2,6	21,9	3	X	13843857	226			0,00
ALGA0120288	2915	0	24,4	3	X	13957508	230			0,00
ALGA0099408	2915	15,7	14,9	3	X	13898655	229			0,00
ALGA0099412	2932,1	3,7	19,7	3	X	14154082	231			0,00
INRA0056499	2935,9	23,6	12,7	3	X	14182919	232			0,00
ASGA0080857	2962,8	0	20,3	3	X	14387866	234			0,00
MARC0072193	2962,8	2,8	18,2	3	X	14359593	233			0,00
M1GA0023599	2965,6	0	22,5	3	X	14444686	236			0,00
ASGA0080858	2965,6	5,4	19	3	X	14429098	235			0,00
M1GA0023603	2971,2	2,7	20,8	3	X	14488730	237			0,00
ALGA0099416	2973,9	0	23,3	3	X	14587254	239			0,00
CASIO003795	2973,9	0	23,3	3	X	14513143	238			0,00
H3GA0051567	2973,9	0	23,3	3	X	14656171	241			0,00
ASGA0080864	2973,9	2,7	21,3	3	X	14613767	240			0,00
ALGA0099419	2976,6	13,5	14,6	3	X	14694394	242			0,00
ASGA0093713	2991,1	13,5	14,6	3	X	15010224	243			0,00
ALGA0099420	3005,6	13	15,5	3	X	15038886	244			0,00
M1GA0023606	3019,5	3,2	19,9	3	X	15081305	245			0,00
M1GA0023609	3022,8	5,3	19,8	3	X	15179599	246			0,00
H3GA0051576	3028,2	3,1	21,2	3	X	15221016	247			0,00
H3GA0051577	3031,4	13,1	15,7	3	X	15246857	248			0,00
M1GA0023611	3045,5	10,4	15,2	3	X	15281878	249			0,00
INRA0056524	3056,5	3,4	18,6	3	X	15360592	250			0,00
INRA0056526	3060	8,1	17,5	3	X	15376436	251			0,00
H3GA0051581	3068,4	0	23,3	3	X	15446569	254			0,00
ALGA0099429	3068,4	0	22,8	3	X	15418248	253			0,00
ALGA0099427	3068,5	8	17,5	3	X	15414393	252			0,00
H3GA0051580	3076,8	3,2	20,7	3	X	15487505	255			0,00
INRA0056528	3080	13	15,5	3	X	15506170	256			0,00
ALGA0099433	3093,9	13,5	14,5	3	X	15592452	257			0,00
ASGA0080877	3108,5	8,2	16,7	3	X	15939865	258			0,00
ASGA0080878	3117	13,5	14,4	3	X	15982273	259			0,00
ASGA0080879	3131,5	5,7	18,8	3	X	16023625	260			0,00
ASGA0085569	3137,4	25,8	11,1	3	X	16036762	261		FP085429.5	0,00
MARC0110056	3167,3	14	16,7	3	X	16543737	262	7	CU856551.5	0,00
H3GA0051587	3182,4	8,1	20,3	3	X	16517599		7	CU856551.5	0,00
H3GA0051588	3190,8	0	26,5	3	X	16608916	264		CU856551.5	0,00
MARC0075377	3190,8	7,3	20,5	3	X	16571375	263			0,00
ALGA0099436	3198,3	10,8	17,1	3	X	16696492	265			0,00
ASGA0089140	3209,8	0	23,4	3	X	16831825	267			0,00
DRGA0017139	3209,8	28	10	3	X	16821520	266			0,00
ALGA0107322	3242,6	5,5	19,5	3	X	17147091	268			0,00
ALGA0113568	3248,3	35,5	8,8	3	X	17175776	269			0,00
MARC0007534	3292,1	18,8	14	3	X	17648713	270		FP085383.2	0,00
H3GA0051589	3312,9	5,7	18,5	3	X	18027206	271	8	CU207383.6	0,00
INRA0056532	3318,8	17,7	12,6	3	X	17921534		8	CU207383.6	0,00
INRA0056541	3338,2	21,8	11,6	3	X	18397206	272		CU550673.3	0,00
ALGA0099445	3362,8	14,2	14,6	3	X	18578960	273			0,00
ALGA0110447	3378,1	41,8	6,9	3	X	18720914	274			0,00

DRGA0017142	3432,2	13,5	14,6	3	X	19357189	275			0,00
ASGA0080894	3446,7	15,6	13,3	3	X	19391263	276			0,00
MARC0084083	3463,7	18,7	12,4	3	X	19484109	277			0,00
ALGA0099449	3484,4	9,4	15,4	3	X	19571422	278			0,00
ASGA0080895	3494,3	6,1	16,4	3	X	19701277	279			0,00
H3GA0051592	3500,6	6	16,9	3	X	19745935	280			0,00
ALGA0099453	3506,8	23,2	11	3	X	19819712	281			0,00
MARC0098373	3533,1	9,8	15,7	3	X	20449406	282			0,00
ALGA0099455	3543,4	16,9	12,5	3	X	20556792	283			0,00
MARC0011285	3561,9	2,9	18,1	3	X	20627071	284			0,00
ALGA0099456	3564,9	28,3	8,8	3	X	20730332	285			0,00
ALGA0099457	3598,1	11,1	14,7	3	X	20932608	286			0,00
INRA0056550	3609,9	0	20,4	3	X	21123895	288			0,00
MARC0074079	3609,9	3,2	18,3	3	X	21103213	287			0,00
ALGA0099460	3613,2	19,7	11,4	3	X	21214195	289			0,00
ALGA0099461	3635,1	19,7	11,6	3	X	21355809	290			0,00
MARC0103569	3657	29,6	9,6	3	X	21506559	291			0,00
ASGA0080907	3692,1	2,9	20,3	3	X	22009202	292			0,00
MARC0096091	3695,1	13,8	14,2	3	X	22058835	293			0,00
INRA0056566	3709,9	0	22,1	3	X	22243187	296			0,00
INRA0056562	3709,9	0	21,6	3	X	22196317	295			0,00
INRA0056561	3709,9	2,8	18,9	3	X	22182736	294			0,00
ASGA0080914	3712,8	6,3	16,9	3	X	22260185	297			0,00
M1GA0023618	3719,3	18,7	12,7	3	U	20473505				0,00
ALGA0099476	3740	5,4	18,4	3	X	22730354	298			0,00
ASGA0080917	3745,6	10,7	15,9	3	X	22756962	299			0,00
H3GA0051605	3756,9	3,5	20,5	3	X	22775852	300			0,00
ALGA0099479	3760,5	7,1	19	3	X	22800247	301			0,00
ALGA0099482	3767,8	2,9	20,5	3	X	22820590	302			0,00
ASGA0100414	3770,8	5,8	18,8	3	X	22836936	303			0,00
INRA0059349	3776,7	23,4	11,6	3	X	22855768	304			0,00
ALGA0099486	3803,3	13,1	15,9	3	X	23323011	305			0,00
MARC0056932	3817,3	11,4	16,5	3	X	23614448	306			0,00
ASGA0080927	3829,5	13,6	14,6	3	X	23884239	307			0,00
ASGA0080929	3844,2	16,7	12,8	3	X	23980272	308			0,00
DRGA0017172	3862,4	3,2	19,7	3	X	24126684	309			0,00
ALGA0099496	3865,6	0	21,2	3	X	24345955	311			0,00
ALGA0099497	3865,6	16,8	12,8	3	X	24303646	310			0,00
MARC0114348	3884	14,6	11,7	3	X	24440411	312			0,00
ALGA0099498	3899,9	30,7	8,1	3	X	24769857	313			0,00
MARC0041073	3936,5	24,1	10,8	3	X	25015600	314		FP245421.2	0,00
ALGA0099508	3964,1	5,3	19,6	3	X	25407106		9	CU468992.2	0,00
DRGA0017183	3969,6	8,1	16,5	3	X	25264447	315		CU855538.2	0,00
ALGA0099509	3978	13,8	13,1	3	X	25351691	316			0,00
MARC0095098	3992,9	11,1	15	3	X	25598073				0,00
INRA0056584	4004,7	0	20,8	3	X	25750686	318		CU914767.3	0,00
ASGA0080944	4004,7	39,7	6	3	X	25391786	317	10	CU468992.2	0,00
INRA0056578	4055,4	47,9	4,4	3	U	1424				0,00
MARC0088489	4120,6	24,4	10,8	3	X	25930877	319		FQ014219.3	0,00
MARC0046092	4148,6	14,5	13,8	3	X	26333452	320			0,00
MARC0094939	4164,3	5,9	17,4	3	X	26393606	321			0,00
MARC0058538	4170,4	17,9	12,9	3	X	26404011	322			0,00
SIRI0001474	4190,1	16,3	13,7	3	X	26693958	323			0,00
MARC0038149	4207,9	8	17,5	3	X	26951074	324			0,00
DRGA0017189	4216,2	5,4	18,6	3	X	27034249	325			0,00
ASGA0080948	4221,7	5,5	17,1	3	X	27063320	326			0,00
ASGA0091695	4227,3	27,6	8,2	3	X	27092597	327			0,00
DRGA0017191	4259,6	8,2	17,2	3	U	22389669				0,00
DRGA0017199	4268,1	13,6	14,6	3	X	27678225	328			0,00
ASGA0080952	4282,8	5,4	18,9	3	X	27962279	329			0,00
ALGA0099522	4288,3	8,1	16	3	X	28014791	330			0,00
ALGA0099524	4296,8	2,7	18,2	3	X	28116041	331			0,00
DRGA0017205	4299,6	18,9	12,5	3	X	28149721	332			0,00
DBNP0000059	4320,5	0	23,9	3	X	28389686	334			0,00
ALGA0099529	4320,5	10,7	16,6	3	X	28337468	333			0,00
ALGA0099530	4331,8	21,3	9,9	3	X	28418793	335			0,00
INRA0056602	4355,8	9,1	14,2	3	X	28617226	336			0,00
H3GA0051621	4365,3	10,8	16	3	X	28632007	337			0,00

ASGA0080960	4376,7	23,5	11,8	3	X	28688598	338		0,00
MARC0072612	4403,5	5,8	18,7	3	X	28988256	339		0,00
CASIO006594	4409,5	5,8	18,5	3	X	29072100	340		0,00
ALGA0099539	4415,5	29,7	9,7	3	X	29216623	341		0,00
ALGA0099540	4450,7	6,1	17,1	3	X	29454951	342		0,00
MARC0011063	4456,9	8,4	16,3	3	X	29604046	343		0,00
H3GA0051628	4465,7	22	10,6	3	X	29704103	344		0,00
DRGA0017225	4490,5	6,2	16,5	3	X	30064026	345		0,00
ALGA0099545	4497	66,3	2	3	X	30087545	346		0,00
ALGA0099648	4605,9	2,8	20	4	X	34511133	347		0,00
ASGA0081084	4608,7	13,9	14,5	4	X	34551801	348		0,00
ASGA0081086	4623,7	12,5	16,2	4	X	34603527	349		0,00
INRA0056693	4637	6,3	17,1	4	X	34748484	350		0,00
ALGA0099653	4643,5	13,8	12,7	4	X	34765863	351		0,00
MARC0002024	4658,4	16,7	12,5	4	X	34925038	352		0,00
INRA0061605	4676,7	6,7	17	4	X	35034444	353		0,00
ALGA0108064	4683,6	10	15,8	4	X	35120284	354		0,00
H3GA0051726	4694,1	0	23,1	4	X	35153341	355		
ASGA0081089	4694,2	16,7	13	4	X	35194326	356		0,00
ALGA0099661	4712,4	11,4	14,4	4	X	35343870	357		0,00
H3GA0051735	4724,5	11,4	13,8	4	X	35429400	358		0,00
ALGA0099666	4736,6	10,3	14,2	4	X	35602686	359		0,00
ALGA0099668	4747,4	6,9	15,3	4	X	35680718	360		0,00
CASIO008758	4754,5	8,6	15,4	4	X	35737712	361		0,00
ALGA0110700	4763,5	2,9	18,6	4	X	35815889	362		0,00
MARC0017264	4766,3	5,7	17,2	4	X	35838914	363		0,00
MARC0009475	4772,2	5,6	17,7	4	U	1530			0,00
ASGA0104326	4778	22	11,4	4	X	35973541	364		0,00
INRA0060785	4802,8	24,3	11	4	X	36313690	365		0,00
ALGA0103600	4830,6	2,7	21	4	X	36632018	366		0,00
ALGA0110939	4833,3	11	15,8	4	X	36750664	367		0,00
DRGA0017467	4845	2,8	20,1	4	X	36828555	368		0,00
ASGA0103680	4847,8	2,7	20,5	4	X	36885147	369		0,00
ASGA0100973	4850,5	2,7	18,8	4	X	36914927	370		0,00
H3GA0056568	4853,3	29,1	8,5	4	X	36919815	371		0,00
ASGA0092599	4887,8	45,5	4,7	4	X	36956021	372		
INRA0056612	4948,5	0	20,4	4	X	37248345	373		0,00
H3GA0051635	4948,5	4,2	18,4	4	X	37274618	374		0,00
ASGA0080976	4952,8	0	20,5	4	X	37430671	375		0,00
ALGA0099559	4952,8	4,2	18,4	4	X	37453542	376		0,00
H3GA0051638	4957,1	21,4	12,1	4	X	37474047	377		0,00
ASGA0080982	4981,2	6	17,4	4	X	37583277	378		0,00
DRGA0017231	4987,4	0	19,7	4	X	37624574	379		0,00
ALGA0099569	4987,4	0	19,1	4	X	37647128	380		0,00
ALGA0099572	4987,4	2,8	17,1	4	X	37678424	381		0,00
ASGA0080989	4990,2	9,4	17,6	4	X	37710596	382		0,00
ALGA0099574	5000,1	9,4	17,6	4	X	37803629	383		0,00
ALGA0099575	5009,9	0	23,1	4	X	37833279	384		0,00
ALGA0099576	5009,9	2,7	21	4	X	37863545	385		0,00
ALGA0099578	5012,7	2,7	21	4	X	37909815	386		0,00
MARC0060376	5015,4	2,7	20,4	4	X	37947367	387		0,00
ASGA0081002	5018,1	5,4	18,9	4	X	37990102	388		0,00
ASGA0081003	5023,7	8	17,6	4	X	38023389	389		0,00
ASGA0081004	5032	2,6	21,4	4	X	38086949	390		0,00
INRA0056638	5034,7	0	23,9	4	X	38214530	391		0,00
ALGA0099584	5034,7	10,5	17	4	X	38235904	392		0,00
ALGA0099585	5045,7	0	22,2	4	X	38268046	393		0,00
ALGA0099588	5045,8	6,7	18,6	4	X	38289148	394		0,00
SIRIO000669	5052,7	6,7	17,4	4	X	38473276	395		0,00
ASGA0081008	5059,6	2,6	21	4	X	38607037	396		0,00
ASGA0081018	5062,2	5,3	19,1	4	X	38670937	397		0,00
ALGA0099595	5067,7	2,7	20,3	4	X	38680048	398		0,00
ALGA0099603	5070,4	9,7	16,5	4	X	38708793	399		0,00
INRA0056654	5080,6	6,4	19,3	4	X	39062455	400		0,00
ASGA0081023	5087,2	15,4	14,9	4	X	39204865	401		0,00
ALGA0099608	5103,8	3	21,7	4	X	39230661	402		0,00
ASGA0081025	5106,8	9,1	18,5	4	X	39254782	403		0,00
ALGA0099612	5116,4	4,9	20,7	4	X	39305636	404		0,00

ALGA0099614	5121,3	11,8	17,3	4	X	39325975	405		0,00
H3GA0051674	5133,9	16,1	14,2	4	X	39360955	406		0,00
ASGA0081029	5151,5	3,5	21,2	4	X	39392195	407		0,00
H3GA0051677	5155	7,2	18,9	4	X	39405671	408		0,00
ALGA0099619	5162,5	13,5	15,2	4	X	39434792	409		0,00
ASGA0081033	5177	11,5	16,7	4	X	39470159	410		0,00
MARC0112877	5189,2	7,8	17,5	4	X	39496837	411		0,00
ASGA0081036	5197,2	3,3	18,9	4	X	39513512	412		0,00
M1GA0023654	5200,6	13,4	14,1	4	X	39535835	413		0,00
H3GA0051688	5215,1	0	21,4	4	U	1322			0,00
ASGA0081046	5215,1	0	22,4	4	X	39591719	414		0,00
ALGA0099626	5215,1	15,9	15,2	4	X	39616763	415		0,00
ASGA0081055	5232,4	0	24,7	4	X	39658724	416		0,00
ASGA0081057	5232,4	5,6	20,4	4	X	39671416	417		0,00
H3GA0051697	5238,2	20,5	12,5	4	X	39695484	418		0,00
H3GA0051701	5261,2	13,5	14,9	4	X	39731548	419		0,01
ASGA0081060	5275,7	0	22,7	4	X	39762262	420		0,00
H3GA0051704	5275,7	12,9	16,4	4	X	39782830	421		0,00
ASGA0081063	5289,5	27,5	10,4	4	X	39805849	422		0,00
ASGA0090264	5321,7	33,5	8,5	4	U	6281791			0,00
M1GA0023676	5362,5	5,3	19,8	4	X	39870323	423		0,00
H3GA0051711	5367,9	8,1	16,9	4	X	39893402	424		0,00
H3GA0051713	5376,4	16,1	13,7	4	X	39959010	425		0,00
M1GA0023684	5393,9	10,6	16,9	4	X	39994318	426		0,00
H3GA0051716	5405,1	19,4	11,1	4	X	40047092	427		0,00
M1GA0023689	5426,7	13,9	13,2	4	X	40134610	428		0,00
MARC0046838	5441,6	24,6	10,8	4	X	40237740	429		0,00
DIAS0002807	5469,9	3,6	17,9	4	X	40383724	430		0,00
ASGA0105181	5473,5	20,5	10	4	X	40435781	431		0,00
ALGA0099669	5496,4	6	14,4	4	X	40851499	432		0,00
H3GA0051739	5502,7	3	16,3	4	X	40929401	433		0,00
SIRI0000553	5505,7	11,9	11,8	4	X	40954819	434		0,00
ASGA0081096	5518,4	7,9	12,9	4	U	37447406			0,00
ASGA0081097	5526,6	20,3	8,2	4	X	41097603	435		0,00
H3GA0051743	5549,3	6,1	12,9	4	X	41157602	436		0,00
CAHM0000185	5555,6	0	16,8	4	X	41172411	437		0,00
ASGA0081099	5555,6	3	15,6	4	X	41189836	438		0,00
ASGA0081102	5558,7	6	14,3	4	X	41270474	439		0,00
ALGA0099682	5564,9	15,1	10,8	4	X	41381724	440		0,00
ALGA0099684	5581,3	3	15,6	4	X	41572198	441		0,00
ALGA0099686	5584,3	6	14,3	4	X	41654469	442		0,00
ALGA0099688	5590,4	5,9	15,6	4	X	41695550	443		0,00
ALGA0099692	5596,5	0	19,6	4	X	41879564	444		0,00
H3GA0051748	5596,5	8,8	14,3	4	X	41984701	445		0,00
ASGA0081112	5605,7	17,1	12	4	X	42168505	446		0,00
ASGA0081113	5624,5	11	15,7	4	X	42307123	447		0,00
ASGA0104984	5636,1	0	22,7	4	X	42381862	448		0,00
MARC0053932	5636,1	22,3	10,7	4	X	42401261	449		0,00
H3GA0051753	5661,3	0	20,4	4	X	42831924	450		0,00
ASGA0081114	5661,4	8,7	14,7	4	X	42843771	451		0,00
ASGA0081117	5670,4	9,7	14,6	4	X	42942863	452		0,00
ALGA0099700	5680,6	4,6	18,4	4	X	42984220	453		0,00
ASGA0081121	5685,4	0	21,5	4	X	43046630	454		0,00
INRA0056710	5685,4	11,4	13,7	4	X	43059926	455		0,00
INRA0056714	5697,5	5,8	16,1	4	X	43073855	456		0,00
MARC0112186	5703,4	14	13,3	4	X	43111314	457		0,00
ASGA0095463	5718,5	2,8	19,7	4	X	43119073	458		0,00
H3GA0051758	5721,3	25,6	9,9	4	X	43222426	459		0,00
ALGA0048882	5750,8	60,5	2,6	4	U	140			0,00
ASGA0084081	5843,6	0	21	5	X	43875783	460		0,00
H3GA0051764	5843,6	14,1	13,5	5	X	44092699	461		0,00
ASGA0081128	5858,9	11,5	12	5	X	44255664	462		0,00
ALGA0099709	5871,1	2,9	16,1	5	X	44343366	463		0,00
ALGA0099714	5874,1	5,8	15,5	5	X	44567520	464		0,00
MARC0113187	5880,1	20,1	10,1	5	X	44669558	465		0,00
ASGA0089444	5902,5	5,6	16,9	5	X	44901083	466		0,00
MARC0094196	5908,3	17,1	11,6	5	X	44925074	467		0,00
ASGA0083522	5927,1	2,8	18,3	5	X	45004996	468		0,00

ASGA0094952	5930	16,3	12,2	5	X	45005429	469			0,00
ASGA0081139	5947,7	3,3	17,4	5	X	45147562	470			0,00
DRGA0017252	5951	6,5	16	5	X	45171592	471			0,00
INRA0056718	5957,7	20,7	9,7	5	X	45193927	472			0,00
ASGA0081141	5980,9	3	16,7	5	X	45290257	473			0,00
ASGA0081143	5983,9	3	16,7	5	X	45299243	474			0,00
CASI0006838	5986,9	15,3	8,8	5	X	45375726	475			0,00
ASGA0081148	6003,5	12,3	9,5	5	X	45395917	476			0,00
H3GA0051772	6016,7	3	15,5	5	X	45432843	477			0,00
ALGA0099721	6019,7	3	15,4	5	X	45506116	478			0,00
ASGA0081151	6022,8	8,8	14,2	5	X	45518635	479			0,00
ASGA0081152	6032	5,8	15,3	5	X	45539525	480			0,00
DRGA0017254	6038	5,9	14,8	5	X	45561094	481			0,00
MARC0030812	6044	3	16,2	5	X	45611014	482			0,00
MARC0000172	6047	0	18	5	X	45653306	483			0,00
ALGA0099725	6047	0	18,8	5	X	45696218	484			0,00
ASGA0081155	6047,1	3	16,3	5	X	45762955	485			0,00
H3GA0051778	6050,1	6	14,9	5	X	45779996	486			0,00
ALGA0099728	6056,2	3	17	5	X	45914845	487			0,00
ASGA0081159	6059,2	17,6	11,1	5	X	45948905	488			0,00
ALGA0099732	6078,5	8,7	16	5	X	46048134	489			0,00
H3GA0051786	6087,6	0	19,5	5	X	46085263	490			0,00
MARC0090296	6087,6	11,7	13,3	5	X	46124768	491			0,00
M1GA0023718	6100	2,9	18	5	X	46229089	492			0,00
H3GA0051790	6103	11,8	12,9	5	X	46258200	493			0,00
CAIL0000254	6115,5	5,9	15,4	5	X	46583822	494			0,00
H3GA0052592	6121,6	44,5	3,8	5	X	46614790	495			0,00
M1GA0023719	6180,5	27,8	8,5	5	U	1457				0,00
MARC0020505	6213	7,8	13,4	5	U	1585				0,00
ASGA0098428	6221,2	0	18,8	5	X	46683760	496			0,00
CASI0005642	6221,2	0	18,1	5	X	46709850	497			0,00
CASI0005731	6221,2	22,4	8,8	5	X	46709947	498			0,00
ALGA0103643	6246,5	41,2	4,7	5	X	46745547	499			0,00
MARC0036635	6299,7	22,9	10	5	X	47168747	500			0,00
MARC0095030	6325,7	20,2	10,1	5	X	47338233	501			0,00
ALGA0115848	6348,2	6,2	15,9	5	X	47464201	502			0,00
ALGA0123708	6354,6	47,6	5	5	X	47500409	503			0,00
MARC0064721	6419,2	0	21,7	6	X	48174948	504			0,00
CASI0012104	6419,2	11,2	16,4	6	X	48308192	505			0,00
ASGA0089807	6431,1	20,1	10,1	6	X	48537427	506			0,00
H3GA0051793	6453,6	5,7	17,2	6	X	48751658	507			0,00
INRA0056729	6459,4	19,7	10,7	6	X	48773949	508			0,00
INRA0056730	6481,3	26,4	10,3	6	X	48929906	509			0,00
ASGA0081177	6512	9,6	17	6	X	49013199	510			0,00
INRA0056732	6522,1	5,6	17,8	6	X	49072513	511			0,00
ALGA0099737	6527,9	33,7	7,4	6	X	49144733	512			0,00
H3GA0051797	6569	6,5	19,5	7	X	49466141	513			0,00
ALGA0099741	6575,8	7	19,3	7	X	49515649	514			0,00
ASGA0081179	6583	5,6	19,6	7	X	49582630	515			0,00
MARC0027899	6588,8	29,2	8,4	7	X	49617748	516			0,00
ALGA0099744	6623,3	18,7	12,3	7	X	50625708	517			0,00
INRA0056733	6644,1	4,4	18,5	7	X	50662741	518			0,00
INRA0056737	6648,6	14,7	15,7	7	X	50719962	519			0,00
ASGA0081182	6664,5	5,1	20,8	7	U	759				0,00
MARC0038794	6669,7	46,6	5,7	7	X	50839151	520			0,00
MARC0009397	6732,4	14	14,5	8	X	51700353				0,00
BGIS0007835	6747,5	22,3	12	8	X	51804104				0,00
MARC0018661	6772,8	29,5	8,9	8	U	1577				0,00
INRA0056739	6807,7	23	9,6	8	X	51851733				0,00
MARC0056564	6833,8	8,6	15,5	8	X	52288809				0,00
MARC0051269	6842,8	77,6	0,6	8	X	52526690		FP340254.2		0,00
H3GA0051809	6992,4	16,5	13,6	9	X	56808512		11	CU856223.2	0,00
INRA0056742	7010,4	27,5	9,8	9	X	56659659		11	CU856223.2	0,00
H3GA0055482	7042,7	29,1	9,6	9	U	19473000				0,00
MARC0086069	7077	22,8	10,3	9	U	1862				0,00
ASGA0090927	7102,9	73,3	0,6	9	U	861				0,00
CAIL0000258	7234,8	14,9	10,8	10	X	62342620	521		CU694617.2	0,00
ALGA0099771	7250,9	8,8	13,8	10	X	62429427	522			0,00

INRA0056751	7260,1	3,7	16,8	10	X	62454725	523			0,00
INRA0056753	7263,9	22	10	10	X	62473840	524			0,00
ALGA0099767	7288,8	16,2	10,7	10	X	62629920	525			0,00
ALGA0099769	7306,5	0	19,9	10	X	62650052	526			0,00
ALGA0099770	7306,5	11,5	13,3	10	X	62675407	527			0,00
H3GA0051816	7318,7	11,5	13,3	10	X	63109515	528			0,00
H3GA0051815	7331	8,6	14,9	10	X	63195121	529			0,00
MARC0014028	7340	31	8,1	10	X	63297208	530			0,00
DBNP0001718	7377	17,4	12,7	10	X	63646243	531			0,00
ALGA0099761	7396,2	20,4	12,5	10	X	63708027	532			0,00
ALGA0099760	7419,1	25	9,7	10	X	63891837	533			0,00
MARC0060021	7447,8	5,7	15,2	10	X	64571859	534			0,00
INRA0056764	7453,7	11,3	12,9	10	X	64665979	535			0,00
ASGA0081194	7465,7	54,8	3	10	X	64882657	536			0,00
INRA0056765	7545,2	14,3	13	11	X	65733803	537			0,00
ASGA0081195	7560,7	6,8	16,5	11	X	65857445	538			0,00
H3GA0051821	7567,7	10,4	14,2	11	X	65945170	539			0,00
H3GA0051820	7578,7	20,4	9,7	11	X	65997012	540			0,00
INRA0056771	7601,5	7,7	14,6	11	X	66093757	541			0,00
DBNP0002253	7609,6	3,9	16,1	11	X	66247075	542			0,00
INRA0056772	7613,5	6,7	14,7	11	X	66297930	543			0,00
INRA0056775	7620,5	13,5	13,7	11	X	66356168	544			0,00
INRA0056776	7635	85,8	0,3	11	X	66488605	545			0,00
ALGA0099787	7829,9	27,1	11	12	X	71022228	546			0,00
INRA0056800	7861,4	17,3	14,5	12	X	71175600	547			0,00
INRA0056799	7880,4	18,6	14,1	12	X	71320128	548			0,00
ALGA0099790	7901	13,6	15,9	12	X	72098978	549			0,00
ALGA0111007	7915,7	18,5	13,4	12	X	72313380	550			0,00
INRA0056802	7936,1	22,4	9,9	12	X	72523273	551			0,00
ALGA0099791	7961,5	22,4	9,9	12	X	73417360	552			0,00
MARC0062362	7986,8	40,4	6,2	12	X	73903037	553			0,00
INRA0056816	8038,6	23,6	11,2	13	X	76542585	554			0,00
H3GA0051830	8065,5	16,9	13,9	13	X	76751110	555			0,00
H3GA0051829	8084	11,4	15,8	13	X	76830183	556			0,00
ALGA0099803	8096,1	14,2	15	13	X	77046979	557			0,00
ALGA0099805	8111,4	20	11,8	13	X	77257605	558			0,00
INRA0056821	8133,8	49,7	3,6	13	X	77333713	559			0,00
ASGA0081208	8202,4	3	16,9	14	X	78247995	560			0,00
INRA0056822	8205,4	8,8	14,3	14	X	78366088	561			0,00
CASI0001379	8214,6	11,7	13,4	14	X	78582885	562			0,00
CAILO000002	8227	20	10,8	14	X	78757423	563			0,00
ALGA0123953	8249,4	14,1	13,4	14	X	78893355	564			0,00
INRA0056833	8264,6	67,3	0,6	14	X	79536055	565			0,00
H3GA0051832	8376,4	8,9	14	15	X	85563023	566			0,00
ALGA0099817	8385,7	0	21,5	15	X	85670297	567			0,00
INRA0056853	8385,7	8,4	15,3	15	X	85729920	568			0,00
DRGA0017261	8394,5	8,3	15,7	15	X	85891025	569			0,00
INRA0056857	8403,1	34	7,4	15	X	85914898	570			0,00
INRA0056863	8444,6	89,9	0,2	15	X	86201492	571			0,00
MARC0075556	8674,2	15	16,3	16	X	89918983	572			0,00
MARC0012275	8690,4	2,9	22,4	16	X	90038587	573			0,00
ASGA0081209	8693,3	7,5	19,7	16	X	90058905	574			0,00
MARC0018104	8701,1	12,7	16	16	X	90112349	575			0,00
MARC0024249	8714,7	24,6	11,5	16	X	90315328	576			0,00
ALGA0120071	8742,9	10,7	17	16	X	90967583	577			0,00
ASGA0104296	8754,3	10,1	17,4	16	U	1035				0,00
INRA0056883	8764,9	2,5	21,9	16	X	97152327	578			0,00
ASGA0081212	8767,4	12,6	16,3	16	X	97188651	579			0,00
INRA0056888	8780,9	5,1	20	16	X	97346910	580		CU480760.10	0,00
MARC0006336	8786,2	12,7	16,3	16	X	97471020	581	2	CU480760.10	0,00
ALGA0099835	8799,9	25,5	12,1	16	U	19611306				0,00
H3GA0051838	8829,4	15,6	15,9	16	X	97419673	582	2	CU480760.10	0,00
DBUN0000917	8846,3	81	0,9	16	X	97832104	583		FP102403.6	0,00
INRA0056907	9012,2	8,2	17,6	17	X	99565570	584			0,00
INRA0056908	9020,8	10,5	17,4	17	X	99671322	585			0,00
H3GA0051842	9031,8	15,4	15,6	17	X	99725484	586			0,00
INRA0056913	9048,5	15,4	15,5	17	X	99836761	587			0,00
INRA0056914	9065,2	5,2	20,5	17	X	100022640	588			0,00

INRA0056917	9070,5	6,8	20,6	17	X	100056217	589			0,00
CASI0010225	9077,6	3,4	22,5	17	X	100225284	590			0,00
INRA0056920	9081	5,2	20,3	17	X	100478965	591			0,00
INRA0056929	9086,4	20,2	13,1	17	X	100580571	592			0,00
ALGA0099836	9108,9	12,9	16	17	X	100779395	593			0,00
ALGA0099837	9122,7	14,5	15,4	17	X	100911159	594			0,00
DBNP0000897	9138,4	2,9	21,4	17	X	101068864	595			0,00
H3GA0051844	9141,4	20,4	13,7	17	X	101133867	596			0,00
ALGA0099838	9164,2	9	18,5	17	X	101490507	597			0,00
ASGA0081216	9173,6	9,3	18,3	17	X	101602201	598			0,00
ALGA0099840	9183,4	46,7	5,4	17	X	101658281	599			0,00
INRA0056930	9246,4	0	23,5	18	X	102723523	600			0,00
INRA0056937	9246,4	19,1	11,6	18	X	102773567	601			0,00
H3GA0051849	9267,7	5,6	16,5	18	X	103011287	602			0,00
ALGA0099851	9273,5	17,1	12,3	18	X	103037035	603			0,00
INRA0056946	9292,2	0	19,9	18	X	103285641	604			0,00
H3GA0051850	9292,2	5,7	16,7	18	X	103335800	605			0,00
H3GA0051852	9298,1	20,4	9,8	18	X	103413196	606			0,00
ASGA0081223	9320,9	22	10,9	18	X	103657659	607		CU914429.2	0,00
ASGA0081230	9345,8	21	12,2	18	X	104939296	608	3	FP016024.5	0,00
ALGA0099866	9369,4	29,4	9,1	18	X	105147853	609	3	CU633273.3	0,00
MARC0093863	9404,3	2,7	19,9	18	X	105360782	610	3	FP102602.4	0,00
MARC0054761	9407,1	11	15,2	18	X	105453804	611	3	FP102602.4	0,00
MARC0075831	9418,7	16,5	12,9	18	X	105603846	612	3	FP102252.3	0,00
ASGA0093307	9436,7	13,7	14,2	18	U	14401848				0,00
ALGA0116559	9451,4	13,5	15	18	U	473				0,00
MARC0031940	9465,9	16,1	14,1	18	X	103804320	613	3	FP015865.7	0,00
ASGA0105262	9483,5	8,2	16,7	18	X	104086577	614	3	FP565759.2	0,00
ASGA0102433	9492	3,1	19,4	18	X	104115747	615	3	CU929914.2	0,00
ALGA0122955	9495,1	19	12,9	18	X	104170699	616	3	CU929914.2	0,00
ALGA0099859	9516,2	0	21,5	18	X	104275689	617	3	CU618307.10	0,00
MARC0071952	9516,2	2,8	19,5	18	X	104300020	618	3	CU618307.10	0,00
MARC0114274	9519	14,7	14,6	18	X	104313388	619	3	CU618307.10	0,00
H3GA0051862	9534,9	20,4	12,8	18	X	104399270	620	3	CU618307.10	0,00
ASGA0104371	9557,7	2,7	21	18	X	104719181	621	3	FP102359.6	0,00
MARC0067876	9560,4	0	23,5	18	X	105694258	622		CU928844.7	0,00
ASGA0104139	9560,4	7,9	18,5	18	X	105695392	623			0,00
ALGA0099868	9568,6	7,7	18,5	18	X	105793106	624			0,00
ALGA0099869	9576,6	2,5	22	18	X	105918634	625			0,00
ASGA0081231	9579,2	10,3	17,9	18	X	105978406	626			0,00
ALGA0099872	9590	15,1	15,8	18	X	106043823	627			0,00
ALGA0099874	9606,3	25	11,3	18	X	106189920	628			0,00
H3GA0051868	9635,1	17,4	14	18	X	106272459	629			0,00
MARC0013382	9654,3	8,8	17,3	18	X	106609109	630			0,00
INRA0056957	9663,5	13,5	14	18	X	106695438	631			0,00
ASGA0092129	9678	7,9	17,8	18	U	875				0,00
CASI0005735	9686,3	7,8	18,8	18	U	1087				0,00
ASGA0081235	9694,4	2,6	22	18	X	106812075	632			0,00
ASGA0081236	9697,1	2,6	21	18	X	106842550	633			0,00
ALGA0099877	9699,7	7,8	17,7	18	X	106871777	634			0,00
ASGA0081240	9707,9	7,7	18,8	18	X	106970823	635			0,00
INRA0056964	9715,9	2,6	22,2	18	X	107000024	636			0,00
MARC0026850	9718,5	15,4	15,2	18	X	107023141	637			0,00
ALGA0099884	9735,3	9,2	17,8	18	X	107190983	638			0,00
ALGA0099882	9745	6,1	19,5	18	X	107286936	639			0,00
ALGA0104344	9751,3	6,2	19,1	18	X	107360497	640			0,00
MARC0030222	9757,6	0	22,8	18	X	107455627	641			0,00
ALGA0111404	9757,6	23,2	13,5	18	X	107478542	642			0,00
DRGA0017268	9784	11,2	17,1	18	X	107724077	643			0,00
INRA0056972	9795,9	10,4	16,6	18	X	107774669	644			0,00
ALGA0099887	9806,9	0	23	18	X	107803722	645			0,00
MARC0029939	9806,9	8,4	18,2	18	X	107832820	646			0,00
MARC0017716	9815,6	25	12,2	18	X	107876735	647			0,00
H3GA0053709	9844,3	12,7	18,8	18	X	108206131	648			0,00
ASGA0081242	9858	0	24,5	18	X	108258250	649			0,00
ALGA0099889	9858	0	23,9	18	X	108341482	650			0,00
ASGA0081243	9858	15,7	13,8	18	X	108379145	651			0,00
CASI0006622	9875,1	2,7	20,7	18	X	108524253	652			0,00

MARC0034652	9877,8	0	22,9	18	X	108576436	653			0,00
ALGA0099895	9877,8	0	22,9	18	X	108605952	654			0,00
ALGA0099896	9877,8	8	17,5	18	X	108647614	655			0,00
DRGA0017276	9886,2	5,4	18,5	18	X	108716572	656			0,00
DRGA0017277	9891,7	13,8	15,2	18	X	108816272	657			0,00
H3GA0051884	9906,6	0	21,5	18	X	108881292	658			0,00
ALGA0099901	9906,6	11,1	15,2	18	X	108915748	659			0,00
MARC0071577	9918,4	5,5	18,2	18	X	109003374	660			0,00
DRGA0017278	9924,1	0	22,2	18	X	109056537	661			0,00
DRGA0017281	9924,1	2,8	19,7	18	X	109091645	662			0,00
DIAS0000787	9926,9	0	21,6	18	X	109121920	663			0,00
ASGA0081247	9926,9	8,6	15,4	18	X	109154253	664			0,00
INRA0056981	9935,9	5,8	16	18	X	109215742	665			0,00
ALGA0099904	9941,9	11,7	13,2	18	X	109274518	666			0,00
ASGA0081249	9954,3	5,8	16,2	18	X	109314982	667			0,00
ASGA0081252	9960,3	5,8	16,2	18	X	109335040	668			0,00
ALGA0099906	9966,3	8,8	14,4	18	X	109357428	669			0,00
MARC0015493	9975,5	11,8	12,8	18	X	109461719	670			0,00
H3GA0051890	9988	2,9	17,5	18	X	109515494	671			0,00
H3GA0055497	9991	5,8	16,1	18	X	109562447	672			0,00
MARC0003320	9997	3	17,9	18	X	109581718	673			0,00
ASGA0093715	10000,1	39,6	6,3	18	X	109626292	674			0,00
MARC0110449	10050,5	5,5	19,5	19	X	110165834	675			0,00
MARC0048855	10056,2	0	21,5	19	X	110208365	676			0,00
ALGA0099918	10056,2	2,8	19,1	19	X	110250748	677			0,00
INRA0056990	10059	0	21,2	19	X	110344256	679			0,00
ASGA0081259	10059	5,7	17,1	19	X	110313959	678			0,00
ALGA0099921	10064,9	2,8	18,9	19	X	110376363	680			0,00
ALGA0099922	10067,7	8,5	15,7	19	X	110455710	681			0,00
ASGA0081262	10076,6	2,8	19,1	19	X	110485560	682			0,00
ASGA0081263	10079,5	5,7	16,6	19	X	110503443	683			0,00
ASGA0081264	10085,3	5,8	15,5	19	X	110528573	684			0,00
H3GA0051893	10091,3	2,9	17,6	19	X	110622688	685			0,00
ALGA0099925	10094,3	21,4	11,2	19	X	110666451	686		CU914323.23	0,00
MARC0063079	10118,4	14,3	15,4	19	X	115135321	687	4	FP015843.2	0,00
ASGA0103380	10133,8	0	24,4	19	X	115075713	688	4	FP015843.2	0,00
H3GA0052615	10133,8	18,1	14,4	19	X	115010565	689	4	FP015843.2	0,00
DIAS0000708	10153,8	3,1	19,9	19	X	114972696	690	4	CU856642.2	0,00
H3GA0051929	10156,9	8,1	17,2	19	X	114951380	691	4	CU856642.2	0,00
ASGA0081339	10165,3	10	18	19	X	114935340	692	4	CU856642.2	0,00
INRA0057021	10175,8	3,3	19,6	19	X	114810073	693	4	CU207320.4	0,00
ASGA0081338	10179,2	15	14,3	19	X	114781818	694	4	CU207320.4	0,00
ALGA0100008	10195,4	5,9	20,1	19	X	114728235	695	4	CU207320.4	0,00
ALGA0100003	10201,5	10,5	17	19	X	114692426	696	4	CU207320.4	0,00
ASGA0081335	10212,6	5,4	17,9	19	X	114661312	697	4	CU207320.4	0,00
ALGA0100001	10218,1	15,7	13,7	19	X	114622837	698	4	CU406985.3	0,00
CASI0009790	10235,2	2,6	22,2	19	X	114595666	699	4	CU914673.2	0,00
M1GA0023756	10237,8	2,6	22,2	19	X	114577246	700	4	CU914673.2	0,00
H3GA0051924	10240,4	2,6	22,5	19	X	114564150	701	4	CU914673.2	0,00
ASGA0081324	10243	21	12,2	19	X	114535864	702	4	CU914673.2	0,00
MARC0040146	10266,5	21	12,1	19	U	1677				0,00
ALGA0099994	10290,1	6	20	19	X	114499816	703	4	CU914673.2	0,00
ALGA0099993	10296,3	12,4	16,2	19	X	114472252	704	4	CU856501.10	0,00
ALGA0099992	10309,5	15,7	14,4	19	X	114422078	705	4	CU856501.10	0,00
ALGA0099989	10326,6	2,9	22	19	X	114375835	706	4	CU856501.10	0,00
ASGA0081313	10329,6	15	15,6	19	X	114342231	707	4	CU856501.10	0,00
ASGA0081312	10345,9	18	14	19	X	114300217	708	4	CU856501.10	0,00
ALGA0099999	10365,8	37,5	8,1	19	U	253				0,00
ASGA0102812	10412,8	2,5	23,2	19	U	19048930				0,00
ASGA0098634	10415,3	17,8	14,5	19	U	19047503				0,00
ASGA0105482	10434,9	0	23,6	19	X	113962590	709	4	FP245395.5	0,00
ASGA0103238	10434,9	5,2	21,4	19	X	113960640	710	4	FP245395.5	0,00
CAHM0000007	10440,3	11,3	17,7	19	X	113927719	711	4	FP245395.5	0,00
ALGA0099984	10452,3	3,5	22	19	X	113821548	712	4	CU467108.12	0,00
ASGA0081310	10455,8	10,8	18,5	19	X	113806399	713	4	CU467108.12	0,00
ASGA0081299	10467,3	0	23,4	19	X	113728388	715	4	CU467108.12	0,00
ALGA0099981	10467,3	13,9	16,4	19	X	113780071	714	4	CU467108.12	0,00
DIAS0000476	10482,3	3,4	21,9	19	X	113688502	716	4	CU467108.12	0,00

H3GA0051907	10485,7	3,4	18,9	19	X	113654570	717	4	CU467108.12	0,00
H3GA0051904	10489,1	3,4	19,2	19	X	113618068	718	4	CU915797.5	0,00
ASGA0081287	10492,6	6,7	20,9	19	X	113577253	719	4	CU915797.5	0,00
ASGA0081284	10499,5	5,4	20,7	19	X	113468186	720	4	CU929805.4	0,00
ASGA0081283	10505,1	25,2	11,7	19	X	113443344	721	4	CU929805.4	0,00
MARCO070291	10534,1	10,5	16,8	19	X	113174617	722	4	CU633285.4	0,00
ASGA0081281	10545,2	7,7	18,2	19	X	113142886	723	4	CU633285.4	0,00
ALGA0099958	10553,3	5,1	19,7	19	X	113114941	724	4	CU633285.4	0,00
ALGA0099955	10558,5	15,7	14,3	19	X	113078996	725	4	CU633285.4	0,00
ALGA0099948	10575,6	5,3	19,3	19	X	112884986	726	4	CU468368.18	0,00
H3GA0051897	10581,1	10,7	16,6	19	X	112807067	727	4	CU468368.18	0,00
INRA0057006	10592,4	11	15,8	19	X	112668183	728	4	CU693400.11	0,00
ALGA0099944	10604,1	14	13,4	19	X	112615161	729	4	CU693400.11	0,00
ALGA0099945	10619,1	16,7	12,5	19	X	112534471	730	4	CU693400.11	0,00
ASGA0081271	10637,4	19,1	12,6	19	X	112459164	731	4	CU927922.3	0,00
ASGA0103791	10658,6	0	22,7	19	X	112108617	733	4	FP101971.13	0,00
ASGA0089897	10658,6	8,1	17,7	19	X	112147706	732	4	FP101971.13	0,00
H3GA0051894	10667	0	22,7	19	X	111903249	735	4	CU861980.6	0,00
ALGA0099936	10667,1	2,7	20,1	19	X	111944157	734	4	CU861980.6	0,00
ALGA0099937	10669,8	10,6	16,8	19	X	111836660	736	4	CU861980.6	0,00
DRGA0017292	10681	2,6	21,4	19	X	111697699	737	4	CU207357.8	0,00
SIRI0000380	10683,6	22,6	12,9	19	X	111676396	738	4	CU207357.8	0,00
DRGA0017289	10709,3	3,2	20,5	19	X	111240833	739	4	CU856263.19	0,00
ALGA0099930	10712,5	2,8	21,9	19	U	252				0,00
ALGA0099931	10715,4	32,7	8	19	X	111209478	740	4	CU856263.19	0,00
ALGA0099927	10755	16,8	12,3	19	X	110912743		4bis	CU207262.4	0,00
DRGA0017287	10773,3	6,2	18,2	19	X	110841417		4bis	CU207262.4	0,00
MARCO064134	10779,7	16,1	13,2	19	X	111015335		4bis	CU856402.16	0,00
ASGA0081268	10797,2	62,6	2,6	19	X	111107255	741	4	CU469179.5	0,00
ALGA0100027	10895,5	5,1	21,2	20	X	115647616	742		CU856192.6	0,00
CAHM0000011	10900,7	3,3	22,8	20	X	115788286	743			0,00
INRA0057030	10904,1	6,6	20,9	20	X	115800499	744			0,00
ALGA0100025	10910,9	7,5	19,5	20	X	115829136	745			0,00
MARCO028428	10918,7	0	24,2	20	X	115856986	746			0,00
ASGA0081353	10918,7	10,2	16,6	20	X	115895028	747			0,00
ALGA0100024	10929,4	10,4	15,6	20	X	116033785	748			0,00
H3GA0051932	10940,4	18,5	12,1	20	X	116066402	749			0,00
ASGA0081347	10960,9	10,6	16,3	20	X	116163275	750			0,00
ALGA0100014	10972,1	2,7	20,7	20	X	116291308	751			0,00
ALGA0100018	10974,8	4,4	20	20	X	116418847	752			0,00
H3GA0051930	10979,3	6,2	19	20	X	116493459	753			0,00
ASGA0081341	10985,7	9,4	18,1	20	X	116527064	754			0,00
CASI0006640	10995,6	12	16,3	20	X	116547127	755			0,00
MARCO004517	11008,4	2,6	22,4	20	X	116686255	756			0,00
CASI0004714	11011	10,3	17,9	20	X	116843785	757			0,00
INRA0057024	11021,8	9,4	17,8	20	X	117010968	758			0,00
ALGA0100019	11031,7	22,2	13,1	20	X	117148065	759			0,00
ALGA0100029	11056,8	0	24,1	20	X	117574951	760			0,00
INRA0057031	11056,8	16,6	12,2	20	X	117611622	761			0,00
ALGA0100031	11074,9	56,1	1,7	20	X	117643900	762			0,00
H3GA0051937	11157,2	3,6	17,5	21	X	119371829	763			0,00
INRA0057039	11160,9	11	14,3	21	X	119428342	764			0,00
INRA0057041	11172,5	0	20	21	X	119507239	765			0,00
ALGA0100033	11172,5	22,9	9,6	21	X	119565063	766			0,00
ASGA0081357	11198,4	25,2	9,9	21	X	119809654	767			0,00
ALGA0100036	11227,5	11,2	14,9	21	X	120411731	768			0,00
ASGA0081359	11239,4	34	7,3	21	X	120540026	769			0,00
ASGA0090123	11281	16,7	13	22	X	121196405	770			0,00
ALGA0105613	11299,2	7,5	17	22	X	121347665	771			0,00
ASGA0100231	11307	3,7	18,9	22	X	121399063	772			0,00
MARCO019077	11310,7	0	21,6	22	X	121464408	773			0,00
MARCO109042	11310,7	2,8	19,5	22	X	121496988	774			0,00
ASGA0081364	11313,5	8,4	15,3	22	X	121532845	775			0,00
ALGA0100039	11322,2	4,2	17,3	22	X	121574868	776			0,00
ALGA0100043	11326,6	4,3	17,9	22	X	121626473	777			0,00
ALGA0100046	11331	2,9	18,6	22	X	121650383	778			0,00
ALGA0100052	11333,9	9,5	15,3	22	X	121677792	779			0,00
MARCO045418	11343,8	0	21,1	22	X	121711665	780			0,00

ASGA0091448	11343,8	6,3	16,6	22	X	121712991	781			0,00
DIAS0001174	11350,4	3,2	17,5	22	X	121796977	782			0,00
ALGA0101975	11353,6	10,1	8,6	22	X	121945530	783			0,00
ASGA0081378	11364,3	15	7,2	22	X	122081895	784			0,00
H3GA0051942	11380,5	0	20,6	22	X	122125166	785			0,00
ALGA0100059	11380,5	2,9	18,6	22	X	122135967	786			0,00
H3GA0051948	11383,4	5,7	16,1	22	X	122198848	787			0,00
ALGA0100073	11389,3	14,4	12,3	22	X	122284202	788			0,00
M1GA0023767	11404,8	32,7	8,3	22	X	122407577	789			0,00
ASGA0100035	11444,3	5,3	19,8	22	X	122792622	790			0,00
ASGA0090034	11449,7	0	23,9	22	X	122843107	791			0,00
ASGA0095043	11449,7	22,6	12,5	22	X	122863117	792			0,00
MARC0056699	11475,4	32,6	9,3	22	X	123518440	793			0,00
ASGA0098084	11514,8	15,5	15	22	U	15454217				0,00
ASGA0087691	11531,6	38,1	8,2	22	U	833				0,00
ASGA0081399	11579,6	7,4	20	22	X	123578482	794			0,00
INRA0057055	11587,3	9,4	18,5	22	X	123628691	795			0,00
ASGA0081398	11597,2	0	24,9	22	X	123659477	796			0,00
ASGA0081396	11597,2	14,4	19,7	22	X	123698101	797			0,00
ASGA0081391	11612,8	0	25,2	22	X	123752552	798			0,00
INRA0057054	11612,8	24,5	12,4	22	X	123800527	799			0,00
ASGA0083267	11640,9	34,5	8,6	22	U	14302099				0,00
MARC0022743	11683,2	12	17,7	23	X	124405844	800			0,00
ALGA0100083	11695,9	9,5	19,2	23	X	124573141	801			0,00
M1GA0023768	11705,9	4,7	22,2	23	X	124719392	802			0,00
ALGA0100084	11710,8	7,1	20,5	23	X	124730576	803			0,00
H3GA0051959	11718,1	11,5	17,7	23	X	124790927	804			0,00
H3GA0051963	11730,3	4,5	22,6	23	X	124937136	805			0,00
ASGA0081413	11734,9	0	27,1	23	X	124994817	806			0,00
ASGA0081409	11735	0	27,1	23	X	125006304	807			0,00
H3GA0051967	11735	6,7	21,4	23	X	125043360	808			0,00
H3GA0051971	11741,9	4,5	22,9	23	X	125118254	809			0,00
ALGA0100095	11746,5	2,3	24,4	23	X	125186910	810			0,00
M1GA0027014	11748,8	0	26,2	23	X	125193471	811			0,00
MARC0056935	11748,8	13,6	15,8	23	X	125199118	812			0,00
H3GA0051975	11763,5	20	13,1	23	X	125268690	813			0,00
H3GA0051981	11785,8	8,8	18,9	23	X	125334164	814			0,00
H3GA0051983	11795	8,7	19,4	23	X	125364187	815		CU606897.13	0,00
ALGA0100119	11804	15,4	15,9	23	X	125674247	816	12	CU571211.2	0,00
ASGA0081434	11820,8	89,1	0,1	23	X	125422463		12	CU207404.4	0,00
ALGA0100147	12042,9	0	25,9	24	X	127826215	817		FP102148.7	0,00
ALGA0100146	12042,9	9,5	20,3	24	X	127863725	818			0,00
ASGA0081470	12052,9	4,7	22	24	X	127912501	819			0,00
ASGA0081471	12057,7	22,2	13,4	24	X	127945580	820			0,00
ALGA0100148	12082,8	0	25,7	24	X	128085069	821			0,00
INRA0057102	12082,8	22,7	13,9	24	X	128108120	822			0,00
M1GA0023799	12108,6	4,9	21,6	24	X	128323441	823			0,00
H3GA0052032	12113,6	2,5	23,4	24	X	128352443	824			0,00
ALGA0100151	12116,1	0	25,5	24	X	128402834	825			0,00
INRA0057104	12116,1	11,7	17,5	24	X	128430326	826			0,00
MARC0069777	12128,5	6	18,4	24	X	128572285	827			0,00
INRA0057107	12134,7	2,6	20,4	24	X	128607665	828			0,00
H3GA0052033	12137,3	2,6	22,2	24	X	128639837	829			0,00
ASGA0081480	12139,9	10,5	17,1	24	X	128669464	830			0,00
INRA0057111	12151	2,6	21,6	24	X	128789248	831			0,00
ALGA0100157	12153,6	18,7	13,3	24	X	128842928	832			0,00
ALGA0100159	12174,3	10,7	16,1	24	X	129032125	833			0,00
ALGA0100160	12185,6	6,3	19,1	24	X	129128714	834			0,00
H3GA0052035	12192,2	9,5	17,7	24	X	129224101	835			0,00
H3GA0052036	12202,1	14,8	14,6	24	X	129291562	836			0,00
ASGA0081483	12218,1	8,8	17,3	24	X	129436153	837			0,00
ALGA0100164	12227,4	5,2	20	24	X	129678182	838			0,00
H3GA0052038	12232,7	2,6	21	24	X	129708478	839			0,00
M1GA0023810	12235,4	2,7	20,1	24	X	129734045	840			0,00
MARC0001039	12238,1	0	21,7	24	X	129773291	841			0,00
ALGA0100173	12238,1	0	22,3	24	X	129829247	842			0,00
MARC0059701	12238,1	5,3	19	24	X	129858672	843			0,00
M1GA0023815	12243,5	7	19,2	24	X	129878841	844			0,00

ASGA0081500	12250,8	3,4	21,7	24	X	129909517	845			0,00
ALGA0100182	12254,2	10,1	18	24	X	129930801	846			0,00
ALGA0100185	12264,9	6,5	20,7	24	X	129975535	847			0,00
H3GA0052060	12271,6	5,1	22,6	24	X	130002728	848			0,00
ALGA0100193	12276,8	3,5	23	24	X	130018690	849			0,00
ALGA0100195	12280,4	0	25	24	X	130037489	850			0,00
ALGA0100201	12280,4	11,1	19,6	24	X	130061019	851			0,00
ASGA0081517	12292,1	15,2	15,6	24	X	130139882	852			0,00
ASGA0081522	12308,7	18,8	14,1	24	X	130187597	853			0,00
INRA0057201	12329,5	2,5	23,6	24	X	130253058	854			0,00
ALGA0100206	12331,9	9,5	19,4	24	X	130295423	855			0,00
ASGA0081528	12341,9	7,1	20,9	24	X	130343379	856			0,00
ALGA0100210	12349,3	7,3	20,3	24	X	130417352	857			0,00
ALGA0100211	12357	13,6	17,3	24	X	130476837	858			0,00
M1GA0023827	12371,6	10,9	17,9	24	X	130497670	859			0,00
ASGA0081529	12383,1	7,5	18,9	24	X	130551842	860			0,00
MARCO051008	12390,9	18,2	12,9	24	X	130591934	861			0,00
DRGA0017306	12411	0	21,7	24	X	130643601	862			0,00
MARCO026568	12411,1	33,8	9	24	X	130717706	863			0,00
M1GA0023830	12452,3	11,9	16,2	24	X	131104850	864			0,00
ASGA0081532	12465	6	18,7	24	X	131164954	865			0,00
ALGA0100214	12471,2	0	25	24	X	131194931	866			0,00
H3GA0052072	12471,3	2,6	22,5	24	X	131221918	867			0,00
ASGA0081536	12473,8	5,1	20,6	24	X	131326695	868			0,00
ASGA0101549	12479,1	5,1	20,8	24	X	131470694	869			0,00
ASGA0104113	12484,4	26,9	9,9	24	X	131601376	870			0,00
ALGA0111431	12515,7	15,2	12,8	24	X	131708696	871			0,00
MARCO036201	12532,2	21,4	10,7	24	X	132088343	872			0,00
DRGA0017308	12556,3	19,6	11,2	24	X	132326739	873			0,00
H3GA0052073	12578,2	19,9	9,7	24	X	132730463	874			0,00
ALGA0100218	12600,4	14	12,2	24	X	132851730	875			0,00
ASGA0081540	12615,5	2,8	20,1	24	X	132904121	876			0,00
ALGA0100223	12618,3	8,3	16,9	24	X	132945039	877			0,00
ALGA0100228	12626,9	16,9	12,6	24	X	133074194	878			0,00
H3GA0052075	12645,4	7,7	16,1	24	X	133166090	879			0,00
INRA0057216	12653,5	3,8	17,9	24	X	133186862	880			0,00
CASIO001438	12657,3	14	13,8	24	X	133273953	881			0,00
ALGA0117672	12672,4	22,3	9,4	24	X	133792858	882		CU928352.4	0,00
ASGA0081541	12697,7	8,4	14,6	24	X	133995643		13	CU618313.2	0,00
ASGA0081542	12706,5	2,7	20,4	24	X	133974413		13	CU618313.2	0,00
ALGA0100236	12709,3	16,1	14,1	24	X	134261711		13	CU407313.4	0,00
INRA0057223	12726,8	5,3	19,3	24	X	134108864	883	13	CU407313.4	0,00
INRA0057225	12732,3	5,3	19,3	24	X	134501175		13	CU207403.3	0,00
INRA0057227	12737,7	0	23,9	24	X	134364813	884	13	CU207403.3	0,00
ALGA0100237	12737,7	13,3	15,5	24	X	134715561	885		CU856496.2	0,00
H3GA0052082	12752	34	7,3	24	X	134779719	886			0,00
ALGA0100238	12793,5	2,9	17,6	25	X	135350565	887			0,00
ALGA0100239	12796,5	5,7	16,1	25	X	135397934	888			0,00
ALGA0100240	12802,4	2,8	18,6	25	X	135463309	889			0,00
M1GA0023834	12805,3	0	21,1	25	X	135519159	890			0,00
SIRIO001106	12805,3	7,5	17,6	25	X	135678628	891			0,00
ALGA0106052	12813	3,7	19,5	25	X	135914828	892			0,00
ASGA0100836	12816,8	13,7	14,5	25	X	135937088	893			0,00
MARCO027616	12831,5	8,1	17,5	25	X	136018066	894			0,00
CASIO009917	12839,9	2,7	20,5	25	X	136071186	895			0,00
ALGA0111114	12842,7	5,4	19	25	X	136095959	896			0,00
ALGA0100242	12848,2	2,7	20,5	25	X	136234003	897			0,04
H3GA0052087	12851	0	22,7	25	X	136287816	898			0,00
MARCO008763	12851	5,4	19	25	X	136338191	899			0,00
ASGA0081547	12856,6	2,7	20,7	25	X	136433302	900			0,00
ALGA0100247	12859,3	28,6	9,6	25	X	136445696	901			0,00
H3GA0052088	12893	5	20,9	25	X	136846599	902			0,00
CASIO001440	12898,2	2,5	23,2	25	X	136863474	903			0,00
DRGA0017316	12900,7	5	21,6	25	X	136889801	904			0,00
ALGA0100249	12905,8	9,8	19,9	25	X	136986704	905			0,00
ALGA0111800	12916,2	0	25,7	25	X	137271066	906			0,00
MARCO045098	12916,2	21,2	13,6	25	X	137323125	907			0,00
MARCO046821	12940	14,3	16,3	25	X	137562054	908			0,00

INRA0057238	12955,4	26,5	11,6	25	X	137814306	909			0,00
H3GA0052090	12986,2	2,5	23	25	X	138175988	910			0,00
H3GA0052091	12988,7	13,8	16,3	25	X	138204177	911			0,00
CASIO006266	13003,6	14,1	15,9	25	X	138268315	912			0,00
MARC0049678	13018,8	5,2	19,7	25	X	138335440	913			0,00
MARC0115244	13024,1	24,1	10,6	25	X	138394250	914			0,00
ALGA0100254	13051,7	34,8	7,9	25	X	138410450	915			0,00
ASGA0081560	13094,5	5,6	19,8	25	X	138484005	916			0,00
H3GA0052097	13100,2	5,5	19,8	25	X	138510438	917			0,00
ALGA0100257	13105,9	0	23,9	25	X	138533262	918			0,00
ASGA0081561	13105,9	11,2	17,1	25	X	138563721	919			0,00
ALGA0100258	13117,8	14,3	14,9	25	X	138620285	920			0,00
ASGA0081564	13133,3	12,8	15,9	25	X	138684852	921			0,00
ASGA0081565	13147	3,2	21,2	25	X	138716109	922			0,00
H3GA0052101	13150,2	24	10,9	25	X	138748129	923			0,00
DRGA0017321	13177,7	14,8	14,5	25	X	138830938	924			0,00
H3GA0052103	13193,7	14,2	15,7	25	X	138884889	925			0,00
ASGA0081569	13209	2,5	22,6	25	X	138932299	926			0,00
M1GA0023842	13211,6	7,7	19,2	25	X	138944474	927			0,00
ALGA0100262	13219,6	7,8	17,8	25	X	138988497	928			0,00
M1GA0023845	13227,7	6,1	19	25	X	139042393	929			0,00
M1GA0023858	13234	12	17,4	25	X	139062103	930			0,00
MARC0095079	13246,8	31,9	8,5	25	X	139204105	931			0,00
MARC0110934	13285,3	13,9	13,5	25	X	139568819	932			0,00
MARC0013075	13300,3	19,6	12,5	25	X	139723313	933			0,00
ASGA0088407	13322,1	16,1	14	25	X	139856936	934			0,00
H3GA0052117	13339,7	14	13,9	25	X	139913062	935			0,00
MARC0031172	13354,8	9,8	15,3	25	X	139935048	936			0,00
ALGA0100267	13365,1	9,7	15,5	25	X	139965455	937			0,00
ASGA0081580	13375,3	9,5	15,5	25	X	139986269	938			0,00
ALGA0100266	13385,4	12,7	14,5	25	X	140000914	939			0,00
ASGA0081577	13398,9	16,7	13,6	25	X	140043606	940			0,00
MARC0001310	13417,2	0	21,5	25	X	140124291	941			0,00
MARC0071986	13417,2	5,6	17,7	25	X	140132987	942			0,00
ALGA0124350	13422,9	25,3	9,6	25	X	140180976	943			0,00
ALGA0112324	13452,1	13,7	14,4	25	X	140471268	944			0,00
SIRI0000378	13466,9	10,7	16,6	25	X	140634325	945			0,00
H3GA0052122	13478,2	8	18,1	25	X	140674130	946			0,00
ASGA0081587	13486,5	4	21,2	25	X	140694550	947			0,00
H3GA0052120	13490,6	3,9	21,4	25	U	1324				0,00
ALGA0100279	13494,6	7,8	18,6	25	X	140716292	948			0,00
H3GA0052119	13502,7	2,6	21,4	25	X	140793755	949			0,00
ALGA0105938	13505,4	37,1	7,2	25	X	140920145	950			0,00
MARC0050328	13551,7	2,7	21	25	X	141270018	951			0,00
ASGA0081595	13554,5	8,1	17,7	25	X	141319763	952			0,00
M1GA0023874	13562,9	5,4	18,5	25	X	141391933	953			0,00
H3GA0052129	13568,5	13,6	14,1	25	X	141412550	954			0,00
M1GA0023883	13583,2	5,5	18,5	25	U	1459				0,00
ALGA0108973	13588,8	11,2	14,7	25	X	141793169	955			0,00
MARC0047250	13600,7	0	21	25	X	141950120	956			0,00
ALGA0108022	13600,7	0	21,1	25	X	141996499	957			0,00
ASGA0093173	13600,7	5,6	19,1	25	X	142000976	958			0,00
MARC0055594	13606,5	5,6	17,1	25	X	142047331	959			0,00
M1GA0023899	13612,3	0	20,6	25	X	142081042	960			0,00
ASGA0081614	13612,3	0	21,2	25	X	142118497	961			0,00
ASGA0081608	13612,3	2,8	19,2	25	X	142126925	962			0,00
ASGA0081616	13615,2	19,8	11,3	25	X	142189065	963			0,00
ASGA0081621	13637,3	2,9	18,7	25	X	142314708	964			0,00
ALGA0100304	13640,2	11,5	13,9	25	X	142327782	965		CU914490.13	0,00
CAHM0000013	13652,4	0	20,7	25	X	143502838	966	5	CU862037.17	0,00
H3GA0052166	13652,4	11,5	13,9	25	X	143488665	967	5	CU862037.17	0,00
ASGA0081645	13664,6	2,9	18,7	25	X	143372799	968	5	CU862037.17	0,00
M1GA0023920	13667,5	2,9	18,7	25	X	143256700	969	5	CU914693.14	0,00
M1GA0023927	13670,4	2,9	17,5	25	X	143223350	970	5	CU914693.14	0,00
H3GA0052163	13673,4	2,9	17,5	25	X	143191017	971	5	CU914693.14	0,00
H3GA0052162	13676,3	0	20,7	25	X	143168666	972	5	CU914693.14	0,00
ASGA0081629	13676,3	11,4	14,4	25	X	143146945	973	5	CU914693.14	0,00
M1GA0023917	13688,4	2,8	19,2	25	X	143013614	974	5	CU861979.13	0,00

ALGA0100307	13691,2	20,4	10,4	25	X	142984588	975	5	CU861979.13	0,00
ALGA0100305	13714	3	16,9	25	X	142814456	976	5	CU914687.13	0,00
H3GA0052153	13717	0	18,8	25	X	142780076	977	5	CU914687.13	0,00
ALGA0100306	13717	20,7	9,7	25	X	142738807	978	5	CU914687.13	0,00
MARC0051265	13740,2	24,3	8,9	25	U	27225824				0,00
ASGA0102172	13768	14	11,3	25	U	1006				0,00
DRGA0017329	13783,1			25	X	142697541	979	5	CU914687.13	0,00



- Adham, I.M., and AgoulNIK, A.I. (2004). Insulin-like 3 signalling in testicular descent. *International Journal of Andrology* 27, 257–265.
- Affara, N.A., Magenis, R.E., Tolmie, J.L., Boyd, E., Cooke, A., Jamieson, D., Kwok, K., Mitchell, M., and Snadden, L. (1987). Mapping the testis determinants by an analysis of Y-specific sequences in males with apparent. *Nucleic Acids Research* 15, 7325–7342.
- Ai, H., Huang, L., and Ren, J. (2013). Genetic diversity, linkage disequilibrium and selection signatures in chinese and Western pigs revealed by genome-wide SNP markers. *PLoS One* 8, e56001.
- Akhtar, N., and Karim, S. (2012). Varieties of foetal congenital abnormality. *Mymensingh Medical Journal : MMJ* 21, 662–667.
- Al-Agha, A.E., Ocheltree, A.H., and Al-Tamimi, M.D. (2012). Association between genotype, clinical presentation, and severity of congenital adrenal hyperplasia: a review. *The Turkish Journal of Pediatrics* 54, 323–332.
- Alexandrova, E. a, Olovnikov, I. a, Malakhova, G. V, Zabolotneva, A. a, Suntsova, M. V, Dmitriev, S.E., and Buzdin, A. a (2012). Sense transcripts originated from an internal part of the human retrotransposon LINE-1 5' UTR. *Gene* 511, 46–53.
- Altshuler, D., Daly, M.J., and Lander, E.S. (2008). Genetic Mapping in Human Disease. *Science* 322, 881–888.
- Amann, R.P., and Veeramachaneni, D.N.R. (2006). Cryptorchidism and associated problems in animals 1. *Animal Reproduction Science* 1, 108–120.
- Amaral, A.J., Megens, H.-J., Crooijmans, R.P.M. a, Heuven, H.C.M., and Groenen, M. a M. (2008). Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics* 179, 569–579.
- Andersson, L., Haley, C.S., Ellegren, H., Knott, S. a, Johansson, M., Andersson, K., Andersson-Eklund, L., Edfors-Lilja, I., Fredholm, M., and Hansson, I. (1994). Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science (New York, N.Y.)* 263, 1771–1774.
- Andresen, E. (1964). The Inheritance of the Blood Factors I a and I b in Pigs of the Duroc and Hampshire Breeds. *Vox Sanguinis* 9, 617–621.
- Andrieux, J. (2008). [Array-CGH for routine diagnosis of cryptic chromosomal imbalances]. *Pathologie-biologie* 56, 368–374.
- Antoniou, G.A., Tentes, I.K., Antoniou, S.A., Simopoulos, C., and Lazarides, M.K. (2011). Matrix Metalloproteinase Imbalance in Inguinal Hernia Formation. *Journal of Investigative Surgery* 24, 145–150.
- Arango, N.A., Lovell-badge, R., and Behringer, R.R. (1999). Targeted Mutagenesis of the Endogenous Mouse *Mis* Gene Promoter : In Vivo Definition of Genetic Pathways of Vertebrate Sexual Development. *Cell* 99, 409–419.
- Archibald, A.L., Haley, C.S., Brown, J.F., Couperwhite, S., McQueen, H. a, Nicholson, D., Coppieters, W., Van de Weghe, A., Stratil, A., and Winterø, a K. (1995). The PiGMaP consortium linkage map of the pig (*Sus scrofa*). *Mammalian Genome : Official Journal of the International Mammalian Genome Society* 6, 157–175.
- Arias, J. a, Keehan, M., Fisher, P., Coppieters, W., and Spelman, R. (2009). A high density linkage map of the bovine genome. *BMC Genetics* 10, 18.

- Armstrong, J.F., Pritchard-Jones, K., Bickmore, W. a, Hastie, N.D., and Bard, J.B. (1993). The expression of the Wilms' tumour gene, WT1, in the developing mammalian embryo. *Mechanisms of Development* 40, 85–97.
- Asakawa, S., Abe, I., Kudoh, Y., Kishi, N., Wang, Y., Kubota, R., Kudoh, J., Kawasaki, K., Minoshima, S., and Shimizu, N. (1997). Human BAC library: construction and rapid screening. *Gene* 191, 69–79.
- Auger, J., D, L.D., Berges, R., Doridot, L., Salmon, B., and Mc, C. (2013). Environmental levels of oestrogenic and antiandrogenic compounds feminize digit ratios in male rats and their unexposed male progeny. *Proceedings of the Royal Society B: Biological Sciences* 280, 20131532.
- Badke, Y.M., Bates, R.O., Ernst, C.W., Schwab, C., and Steibel, J.P. (2012). Estimation of linkage disequilibrium in four US pig breeds. *BMC Genomics* 13, 24.
- Bagheri-Fam, S., Barrionuevo, F., Dohrmann, U., Günther, T., Schüle, R., Kemler, R., Mallo, M., Kanzler, B., and Scherer, G. (2006). Long-range upstream and downstream enhancers control distinct subsets of the complex spatiotemporal Sox9 expression pattern. *Developmental Biology* 291, 382–397.
- Bagheri-Fam, S., Sinclair, A.H., Koopman, P., and Harley, V.R. (2010). Conserved regulatory modules in the Sox9 testis-specific enhancer predict roles for SOX, TCF/LEF, Forkhead, DMRT, and GATA proteins in vertebrate sex determination. *The International Journal of Biochemistry & Cell Biology* 42, 472–477.
- Bagheri-Fam, S., Sreenivasan, R., Bernard, P., Knowler, K.C., Sekido, R., Lovell-Badge, R., Just, W., and Harley, V.R. (2012). Sox9 gene regulation and the loss of the XY/XX sex-determining mechanism in the mole vole *Ellobius lutescens*. *Chromosome Research : an International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 20, 191–199.
- Bailey, J. a, Carrel, L., Chakravarti, a, and Eichler, E.E. (2000). Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 97, 6634–6639.
- Bampton, P.R. (1994). The investigation of genetic and hereditary disorders of pigs. *Pig Journal*.
- Barreiro, L.B., Laval, G., Quach, H., Patin, E., and Quintana-Murci, L. (2008). Natural selection has driven population differentiation in modern humans. *Nature Genetics* 40, 340–345.
- Bashamboo, a, Ledig, S., Wieacker, P., Achermann, J.C., Achermann, J., and McElreavey, K. (2010). New technologies for the identification of novel genetic markers of disorders of sex development (DSD). *Sexual Development : Genetics, Molecular Biology, Evolution, Endocrinology, Embryology, and Pathology of Sex Determination and Differentiation* 4, 213–224.
- Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G., and de Massy, B. (2010). PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* 327, 836–840.
- Beck, J., Lanoi, W., Taeubert, H., Knorr, C., and Brenig, B. (2006a). Porcine Hernia – functional/positional candidate genes. In *1st European Conference on Pig Genomics, (Lodi)*, p. 512.
- Beck, J., Bornemann-Kolatzki, K., Knorr, C., Taeubert, H., and Brenig, B. (2006b). Molecular characterization and exclusion of porcine GUSB as a candidate gene for congenital hernia inguinalis/scrotalis. *BMC Veterinary Research* 2, 14.
- Behl, R., Sheoran, N., Behl, J., Tania, M.S., and Vijn, R.K. (1999). Microsatellite Sequences of Mammals and Their Applications in Genome Analysis in Pigs - A Review. *Animal Genetics* 001, 1822–1830.
- Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., and Haussler, D. (2004). Ultraconserved elements in the human genome. *Science* 304, 1321–1325.

Benko, S., Fantes, J. a, Amiel, J., Kleinjan, D.-J., Thomas, S., Ramsay, J., Jamshidi, N., Essafi, A., Heaney, S., Gordon, C.T., et al. (2009). Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. *Nature Genetics* 41, 359–364.

Bennett, S.N., Caporaso, N., Fitzpatrick, A.L., Agrawal, A., Boyd, H.A., Cornelis, M.C., Hansel, N.N., Heiss, G., Heit, A., Kang, J.H., et al. (2012). Phenotype harmonization and cross-study collaboration in GWAS consortia: the GENEVA experience. *Genetic Epidemiology* 35, 159–173.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* 27, 573–580.

Bernard, P., Ryan, J., Sim, H., Czech, D.P., Sinclair, A.H., Koopman, P., and Harley, V.R. (2012). Wnt signaling in ovarian development inhibits Sf1 activation of Sox9 via the Tesco enhancer. *Endocrinology* 153, 901–912.

Bersaglieri, T., Sabeti, P.C., Patterson, N., Vanderploeg, T., Schaffner, S.F., Drake, J. a, Rhodes, M., Reich, D.E., and Hirschhorn, J.N. (2004). Genetic signatures of strong recent positive selection at the lactase gene. *American Journal of Human Genetics* 74, 1111–1120.

Berta, P., Hawkins, J.B., Sinclair, A.H., Taylor, A., Griffiths, B.L., Goodfellow, N., and Fellous, M. (1990a). Genetic evidence equating SRY and the testis-determining factor. *450*, 348448.

Berta, P., Hawkins, J., Sinclair, A., Taylor, A., Griffiths, B., Goodfellow, P., and Fellous, M. (1990b). Genetic evidence equating SRY and the testis-determining factor. *Nature* 348, 448–450.

Beuermann, C., Beck, J., Schmelz, U., Dunkelberg, H., Schütz, E., Brenig, B., and Knorr, C. (2009). Tissue calcium content in piglets with inguinal or scrotal hernias or cryptorchidism. *Journal of Comparative Pathology* 140, 182–186.

Beverdam, A., and Koopman, P. (2006). Expression profiling of purified mouse gonadal somatic cells during the critical time window of sex determination reveals novel candidate genes for human sexual dysgenesis syndromes. *Human Molecular Genetics* 15, 417–431.

Biason-Lauber, A. (2010). Control of sex development. *Best Practice & Research. Clinical Endocrinology & Metabolism* 24, 163–186.

Biason-Lauber, A., Konrad, D., Navratil, F., and Schoenle, E.J. (2004). A WNT4 mutation associated with Müllerian-duct regression and virilization in a 46,XX woman. *The New England Journal of Medicine* 351, 792–798.

Biason-Lauber, A., Konrad, D., Meyer, M., DeBeaufort, C., and Schoenle, E.J. (2009). Ovaries and female phenotype in a girl with 46,XY karyotype and mutations in the CBX2 gene. *American Journal of Human Genetics* 84, 658–663.

Bidanel, J.P., Milan, D., Iannuccelli, N., Amigues, Y., Boscher, M.Y., Bourgeois, F., Caritez, J.C., Gruand, J., Le Roy, P., Lagant, H., et al. (2001). Detection of quantitative trait loci for growth and fatness in pigs. *Genetics, Selection, Evolution : GSE* 33, 289–309.

Bidanel, J.-P., Ducrocq, V., and Ollivier, L. (1990). Les nouvelles méthodes d'évaluation des reproducteurs - Intérêt et applications du modèle individuel chez le porc. *Journées Recherche Porcine*.

Birk, O.S., Casiano, D.E., Wassif, C. a, Cogliati, T., Zhao, L., Zhao, Y., Grinberg, a, Huang, S., Kreidberg, J. a, Parker, K.L., et al. (2000). The LIM homeobox gene Lhx9 is essential for mouse gonad formation. *Nature* 403, 909–913.

Bishop, C.E., Whitworth, D.J., Qin, Y., Agoulnik, a I., Agoulnik, I.U., Harrison, W.R., Behringer, R.R., and Overbeek, P. a (2000a). A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. *Nature Genetics* 26, 490–494.

Bishop, C.E., Whitworth, D.J., Qin, Y., Agoulnik, a I., Agoulnik, I.U., Harrison, W.R., Behringer, R.R., and Overbeek, P. a (2000b). A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. *Nature Genetics* 26, 490–494.

Bladon, B. (2002). Surgical management of cryptorchidism in the horse. In *Practice* 24, 126–134.

Le Blond, J. (2013). Germany becomes first European country to recognise “undetermined” sex. *The Guardian*.

Bodmer, W., and Bonilla, C. (2008). Europe PMC Funders Group Common and rare variants in multifactorial susceptibility to common diseases. *40*, 695–701.

Bogatcheva, N. V, Truong, A., Feng, S., Engel, W., Adham, I.M., and Agoulnik, A.I. (2003). GREAT/LGR8 is the only receptor for insulin-like 3 peptide. *Molecular Endocrinology (Baltimore, Md.)* 17, 2639–2646.

Bogatcheva, N. V, Ferlin, A., Feng, S., Truong, A., Giancesello, L., Foresta, C., and Agoulnik, A.I. (2007). T222P mutation of the insulin-like 3 hormone receptor LGR8 is associated with testicular maldescent and hinders receptor expression on the cell surface membrane. *American Journal of Physiology. Endocrinology and Metabolism* 292, E138–44.

Botstein, D., and Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature Genetics* 33 *Suppl*, 228–237.

BPEX (2010). 2010 PIG COST OF PRODUCTION IN SELECTED INTRODUCTION METHODOLOGY.

Brandão, M.P., Costa, E.M.F., Fukami, M., Gerdulo, M., Natalia P. Pereira, S., Domenice, O., Ogata, T., and Mendonca, B.B. (2011). MAMLD1 (Mastermind-Like Domain Containing 1) Homozygous Gain-of-Function Missense Mutation Causing 46,XX Disorder of Sex Development in a Virilized Female. In *Hormonal and Genetic Basis of Sexual Differentiation Disorders and Hot Topics in Endocrinology: Proceedings of the 2nd World Conference*,.

Breeuwsma, J. (1967). A case of XXY sex chromosome constitution in an intersex pig. *Journal of Reproduction and Fertility* 16, 119–120.

Bullejos, M., and Koopman, P. (2001). Spatially Dynamic Expression of Sry in Mouse. *Developmental Dynamics* 205, 201–205.

Burd, L., Kerbeshian, J., and Klug, M.G. (2001). Neuropsychiatric genetics: misclassification in linkage studies of phenotype-genotype research. *Journal of Child Neurology* 16, 499–504.

Cai, L., Taylor, J.F., Wing, R. a, Gallagher, D.S., Woo, S.S., and Davis, S.K. (1995). Construction and characterization of a bovine bacterial artificial chromosome library. *Genomics* 29, 413–425.

Camerino, G., Parma, P., Radi, O., and Valentini, S. (2006). Sex determination and sex reversal. *Current Opinion in Genetics & Development* 16, 289–292.

Cameron, F.J., and Sinclair, a H. (1997). Mutations in SRY and SOX9: testis-determining genes. *Human Mutation* 9, 388–395.

Campbell, C.D., Ogburn, E.L., Lunetta, K.L., Lyon, H.N., Freedman, M.L., Groop, L.C., Altshuler, D., Ardlie, K.G., and Hirschhorn, J.N. (2005). Demonstrating stratification in a European American population. *Nature Genetics* 37, 868–872.

Canada Pork International (2013). [www.canadapork.com](http://www.canadapork.com).

Canfield, M. a, Honein, M. a, Yuskiv, N., Xing, J., Mai, C.T., Collins, J.S., Devine, O., Petrini, J., Ramadhani, T. a, Hobbs, C. a, et al. (2006). National estimates and race/ethnic-specific variation of selected birth defects in the United States, 1999-2001. *Birth Defects Research. Part A, Clinical and Molecular Teratology* 76, 747–756.

Care Effective health program (2012a). Evaluation and treatment of cryptorchidism.

Care Effective health program (2012b). Evaluation and treatment of cryptorchidism.

Cesalli, F., and Farion, F. (2009). *Entre deux sexes*.

Chaboissier, M.-C., Kobayashi, A., Vidal, V.I.P., Lützkendorf, S., van de Kant, H.J.G., Wegner, M., de Rooij, D.G., Behringer, R.R., and Schedl, A. (2004). Functional analysis of Sox8 and Sox9 during sex determination in the mouse. *Development (Cambridge, England)* 131, 1891–1901.

Chang, C., So, L., Rs, W., Yeh, S., and Tm, C. (2013). Androgen Receptor ( AR ) Physiological Roles in Male and Female Reproductive Systems : Lessons Learned from AR- Knockout Mice Lacking AR in Selective Cells . 89, 109132.

Charlier, C., Coppieters, W., Rollin, F., Desmecht, D., Agerholm, J.S., Cambisano, N., Carta, E., Dardano, S., Dive, M., Fasquelle, C., et al. (2008). Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nature Genetics* 40, 449–454.

Chassot, A., Bradford, S., Auguste, A., Gregoire, E., Pailhoux, E., De Rooij, D., Schedl, A., and Chaboissier, M. (2012). WNT4 and RSPO1 together are required for cell proliferation in the early mouse gonad . *Development* 139, 23095882.

Chassot, A.-A., Ranc, F., Gregoire, E.P., Roepers-Gajadien, H.L., Taketo, M.M., Camerino, G., de Rooij, D.G., Schedl, A., and Chaboissier, M.-C. (2008). Activation of beta-catenin signaling by Rspo1 controls differentiation of the mammalian ovary. *Human Molecular Genetics* 17, 1264–1277.

Chen, C., Qiao, R., Wei, R., Guo, Y., Ai, H., Ma, J., Ren, J., and Huang, L. (2012). A comprehensive survey of copy number variation in 18 diverse pig populations and identification of candidate copy number variable genes associated with complex traits. *BMC Genomics* 13, 733.

Chen, S.-R., Zheng, Q.-S., Zhang, Y., Gao, F., and Liu, Y.-X. (2013). Disruption of genital ridge development causes aberrant primordial germ cell proliferation but does not affect their directional migration. *BMC Biology* 11, 22.

Cheung, C. (2005). *Investigating Birth Defects*.

Chevillon, P., Mircovich, C., Fleho, S., and Dubroca, J. (2004). Euthanasie en élevage de porc. *Techniporc* 27, 21–27.

Choi, M., Scholl, U.I., Ji, W., Liu, T., Tikhonova, I.R., Zumbo, P., Nayir, A., Bakkaloğlu, A., Ozen, S., Sanjad, S., et al. (2009a). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* 106, 19096–19101.

Choi, M., Scholl, U.I., Ji, W., Liu, T., Tikhonova, I.R., Zumbo, P., Nayir, A., Bakkaloğlu, A., Ozen, S., Sanjad, S., et al. (2009b). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* 106, 19096–19101.

Chow, A., Purkayastha, S., Athanasiou, T., Tekkis, P., and Darzi, A. (2008). Inguinal hernia Digestive system disorders Inguinal hernia. *Clinical Evidence* 1–63.

- Chow, J., Ciaudo, C., Fazzari, M., Mise, N., Servant, N., Ji, G., Attreed, M., Avner, P., Barillot, E., Jm, G., et al. (2010). LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation . *Cell* *141*, 2010.
- Chow, J.C., Yen, Z., Ziesche, S.M., and Brown, C.J. (2005). Silencing of the mammalian X chromosome. *Annual Review of Genomics and Human Genetics* *6*, 69–92.
- Chung, W.C.J., and Auger, A.P. (2013). Gender differences in neurodevelopment and epigenetics. *Pflügers Archiv : European Journal of Physiology* *465*, 573–584.
- Ciobanu, D.C., Bastiaansen, J.W.M., Lonergan, S.M., Thomsen, H., Dekkers, J.C.M., Rothschild, M.F., and Plastow, G.S. (2004). New alleles in calpastatin gene are associated with meat quality traits in pigs. *Journal of Animal Science* 2829–2839.
- Cirulli, E.T., and Goldstein, D.B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Reviews. Genetics* *11*, 415–425.
- Clancy, B.S., Ph, D., Shaw, K.M., and Education, P.D.N. (2008). DNA Deletion and Duplication and the Associated Genetic Disorders Clustering of Breakpoints : Recombination Hotspots. *Nature Education* *1*.
- Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., et al. (2012). Origins and functional impact of copy number variation in the human genome. *464*, 704–712.
- Coop, G., Wen, X., Ober, C., Pritchard, J.K., and Przeworski, M. (2008). High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science* *319*, 1395–1398.
- Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* *16*, 10881–10890.
- Courcelle, E., Beausse, Y., Letort, S., Stahl, O., Fremez, R., Ngom-Bru, C., Gouzy, J., and Faraut, T. (2008). Narcisse: a mirror view of conserved synteny. *Nucleic Acids Research* *36*, D485–90.
- Cox, J.E., Edwards, G.B., and Neal, P. a (1979). An analysis of 500 cases of equine cryptorchidism. *Equine Veterinary Journal* *11*, 113–116.
- Cox, J.J., Willatt, L., Homfray, T., and Woods, C.G. (2011). A SOX9 duplication and familial 46,XX developmental testicular disorder. *The New England Journal of Medicine* *364*, 91–93.
- Damgaard, I.N., Jensen, T.K., Petersen, J.H., Skakkebaek, N.E., Toppari, J., and Main, K.M. (2007). Cryptorchidism and maternal alcohol consumption during pregnancy. *Environmental Health Perspectives* *115*, 272–277.
- Daneau, I., Ethier, J.F., Lussier, J.G., and Silversides, D.W. (1996). Porcine SRY gene locus and genital ridge expression. *Biology of Reproduction* *55*, 47–53.
- Das, P., Mishra, D., Ghosh, S., Avila, F., Johnson, G., Chowdhary, B., and Raudsepp, T. (2013). Comparative organization and gene expression profiles of the porcine pseudoautosomal region. *Cytogenetic and Genome Research* 2013.
- Das, P.J., Chowdhary, B.P., and Raudsepp, T. (2009). Characterization of the bovine pseudoautosomal region and comparison with sheep, goat, and other mammalian pseudoautosomal regions. *Cytogenetic and Genome Research* *126*, 139–147.
- Davidoff, A.M. (2012). Wilms Tumor. *Advances in Pediatrics* *59*, 247–267.

- Davis, E., Jensen, C.H., Schroder, H.D., Shay-hadfield, T., Kliem, A., Cockett, N., Georges, M., and Charlier, C. (2004). Ectopic Expression of DLK1 Protein in Skeletal Muscle of Padumnal Heterozygotes Causes the Callipyge Phenotype. *14*, 1858–1862.
- Dekkers, J.C.M. (2004). Commercial application of marker- and gene-assisted selection in livestock : Strategies and lessons. *Journal of Animal Science*.
- Desarda, M.P. (2003). Surgical physiology of inguinal hernia repair--a study of 200 cases. *BMC Surgery* *3*, 2.
- Dickson, S.P., Wang, K., Krantz, I., Hakonarson, H., and Goldstein, D.B. (2010). Rare variants create synthetic genome-wide associations. *PLoS Biology* *8*, e1000294.
- Ding, N.S., Mao, H.R., Guo, Y.M., Ren, J., Xiao, S.J., Wu, G.Z., Shen, H.Q., Wu, L.H., Ruan, G.F., Brenig, B., et al. (2009). A genome-wide scan reveals candidate susceptibility loci for pig hernias in an intercross between White Duroc and Erhualian. *Journal of Animal Science* *87*, 2469–2474.
- Dolgin, E.S., and Charlesworth, B. (2008). The effects of recombination rate on the distribution and abundance of transposable elements. *Genetics* *178*, 2169–2177.
- Du, F.-X., Mathialagan, N., Dyer, C.J., Grosz, M.D., Messer, L.A., Clutter, A.C., Lohuis, M.M., and Byatt, J.C. (2004). Discovery and mapping of a QTL affecting scrotal hernia incidence in chromosome 2 in domestic pigs. In 29th International Conference on Animal Genetics, (Tokyo), p. 143.
- Du, F.-X., Clutter, A.C., and Lohuis, M.M. (2007a). Characterizing linkage disequilibrium in pig populations. *International Journal of Biological Sciences* *3*, 166–178.
- Du, Z.-Q., Vincent-Naulleau, S., Gilbert, H., Vignoles, F., Créchet, F., Shimogiri, T., Yasue, H., Leplat, J.-J., Bouet, S., Gruand, J., et al. (2007b). Detection of novel quantitative trait loci for cutaneous melanoma by genome-wide scan in the MeLiM swine model. *International Journal of Cancer. Journal International Du Cancer* *120*, 303–320.
- Du, Z.-Q., Zhao, X., Vukasinovic, N., Rodriguez, F., Clutter, A.C., and Rothschild, M.F. (2009). Association and haplotype analyses of positional candidate genes in five genomic regions linked to scrotal hernia in commercial pig lines. *PloS One* *4*, e4837.
- Du, Z.-Q., Gorbach, D.M., Hu, Z., and Rothschild, M.F. (2010). Segmental Duplications In Pig Genome. In *Plant & Animal Genomes XVIII Conference*.
- Ducos, a, Revay, T., Kovacs, a, Hidas, a, Pinton, a, Bonnet-Garnier, a, Molteni, L., Slota, E., Switonski, M., Arruga, M. V, et al. (2008). Cytogenetic screening of livestock populations in Europe: an overview. *Cytogenetic and Genome Research* *120*, 26–41.
- Ducos, A., Berland, H., Pinton, A., Séguéla, A., Blanc, M.F., Darré, A., Sans, P., Darré, R., Nationale, É., and Toulouse, V. De (1997). Les translocations réciproques chez le porc : état des lieux et perspectives. *Journées Recherche Porcine* *375–382*.
- Ducos, A., Pinton, A., Berland, H., Séguéla, A., Brun-baronnat, C., Bonnet, N., Darré, R., and Cytogénétique, U.M.R.I. De (2002). Contrôle chromosomique des populations porcines en France : bilan de cinq années d'activité. *269–275*.
- Duijvesteijn, N., Knol, E.F., Merks, J.W.M., Crooijmans, R.P.M. a, Groenen, M. a M., Bovenhuis, H., and Harlizius, B. (2010). A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6. *BMC Genetics* *11*, 42.

- Dupuis, M.-C., Zhang, Z., Druet, T., Denoix, J.-M., Charlier, C., Lekeux, P., and Georges, M. (2011). Results of a haplotype-based GWAS for recurrent laryngeal neuropathy in the horse. *Mammalian Genome : Official Journal of the International Mammalian Genome Society* 22, 613–620.
- Duthie, C.-A., Simm, G., Pérez-Enciso, M., Doeschl-Wilson, A., Kalm, E., Knap, P.W., and Roehe, R. (2009). Genomic scan for quantitative trait loci of chemical and physical body composition and deposition on pig chromosome X including the pseudoautosomal region of males. *Genetics, Selection, Evolution : GSE* 41, 27.
- Eggers, S., and Sinclair, A. (2012). Mammalian sex determination—insights from humans and mice. *Chromosome Research : an International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 20, 215–238.
- Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J. a, Kawamoto, K., Buckler, E.S., and Mitchell, S.E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6, e19379.
- Emile, J.-F., Leteurtre, E., and Guyétant, S. (2012). *Pathologie générale*.
- Erichsen, H.C., and Chanock, S.J. (2004). SNPs in cancer research and treatment. *British Journal of Cancer* 90, 747–751.
- Evans, P.D., Gilbert, S.L., Mekel-Bobrov, N., Vallender, E.J., Anderson, J.R., Vaez-Azizi, L.M., Tishkoff, S. a, Hudson, R.R., and Lahn, B.T. (2005). Microcephalin, a gene regulating brain size, continues to evolve adaptively in humans. *Science (New York, N.Y.)* 309, 1717–1720.
- Evans, P.M., Chen, X., Zhang, W., and Liu, C. (2010). KLF4 interacts with beta-catenin/TCF4 and blocks p300/CBP recruitment by beta-catenin. *Molecular and Cellular Biology* 30, 372–381.
- Fadista, J., Nygaard, M., Holm, L.-E., Thomsen, B., and Bendixen, C. (2008). A snapshot of CNVs in the pig genome. *PLoS One* 3, e3916.
- Fan, B., Du, Z., Gorbach, D.M., and Rothschild, M.F. (2010). Development and Application of High-density SNP Arrays in Genomic Studies of Domestic Animals methodologies and technologies have been developed. *The Asian-Aust. J. Anim. Sci.* 23, 833–847.
- Fan, B., Onteru, S.K., Du, Z.-Q., Garrick, D.J., Stalder, K.J., and Rothschild, M.F. (2011). Genome-wide association study identifies Loci for body composition and structural soundness traits in pigs. *PLoS One* 6, e14726.
- Faraut, T., de Givry, S., Chabrier, P., Derrien, T., Galibert, F., Hitte, C., and Schiex, T. (2007). A comparative genome approach to marker ordering. *Bioinformatics (Oxford, England)* 23, e50–6.
- Farce, M.-H. (2000). *Génétique moléculaire*.
- Ferguson, L., and Agoulnik, A.I. (2013). Testicular cancer and cryptorchidism. *Frontiers in Endocrinology* 4, 32.
- Ferlin, a, Bogatcheva, N. V, Giancesello, L., Pepe, a, Vinanzi, C., Agoulnik, a I., and Foresta, C. (2006). Insulin-like factor 3 gene mutations in testicular dysgenesis syndrome: clinical and functional characterization. *Molecular Human Reproduction* 12, 401–406.
- Van Ferneij, J.P., Raynaud, F., and Rieu, M. (2012). Commerce international du porc : les principaux courants et leurs évolutions. *Journées Recherche Porcine* 229–234.
- Fiegel, H.C., Rolle, U., Metzger, R., Gfroerer, S., and Kluth, D. (2011). Embryology of the testicular descent. *Seminars in Pediatric Surgery* 20, 170–175.

Fonseca, A.C.S., Bonaldi, A., Bertola, D.R., Kim, C. a, Otto, P. a, and Vianna-Morgante, A.M. (2013). The clinical impact of chromosomal rearrangements with breakpoints upstream of the SOX9 gene: two novel de novo balanced translocations associated with acampomelic campomelic dysplasia. *BMC Medical Genetics* 14, 50.

Food and Agriculture Organization of the United Nations (2008). L'état des ressources zoogénétiques pour l'alimentation et l'agriculture dans le monde.

Food and agriculture organization of the United Nations (2012). <http://faostat.fao.org/>.

Foresta, C., Zuccarello, D., Garolla, A., and Ferlin, A. (2008). Role of hormones, genes, and environment in human cryptorchidism. *Endocrine Reviews* 29, 560–580.

Foster, J., Dominguez-Steglich, M., Guioli, S., Kwok, C., Weller, P., Stevanović, M., Weissenbach, J., Mansour, S., Young, I., Goodfellow, P., et al. (1994). Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature* 8, 525–530.

Fraietta, R., Zylberstejn, D., and Esteves, S. (2013). Hypogonadotropic Hypogonadism Revisited. *Clinics* 68, 81–88.

Frazer, K. a, Murray, S.S., Schork, N.J., and Topol, E.J. (2009). Human genetic variation and its contribution to complex traits. *Nature Reviews. Genetics* 10, 241–251.

Freking, B.A., Murphy, S.K., Wylie, A.A., Rhodes, S.J., Keele, J.W., Leymaster, K.A., Jirtle, R.L., and Smith, T.P.L. (2002). Identification of the Single Base Change Causing the Callipyge Muscle Hypertrophy Phenotype , the Only Known Example of Polar Overdominance in Mammals. *Genome Research* 1496–1506.

Fujii, J., Otsu, K., Zorzato, F., de Leon, S., Khanna, V.K., Weiler, J.E., O'Brien, P.J., and MacLennan, D.H. (1991). Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. *Science (New York, N.Y.)* 253, 448–451.

Gaboreanu, a-M., Grapes, L., Ramos, a M., Kim, J.-J., and Rothschild, M.F. (2004). Characterization of an X-chromosome PCR-RFLP marker associated with fat deposition and growth in the pig. *Animal Genetics* 35, 401–403.

Galani, A., Kitsiou-Tzeli, S., Sofokleous, C., Kanavakis, E., and Kalpini-Mavrou, A. (2008). Androgen insensitivity syndrome: clinical features and molecular defects. *Hormones (Athens, Greece)* 7, 217–229.

Gall, J.G., and Pardue, M. Lou (1969). Formation and detection of rna-dna hybrid molecules in cytological preparations. *Proceedings of the National Academy of Sciences of the United States of America* 378–383.

Gasca, S., Canizares, J., De Santa Barbara, P., Mejean, C., Poulat, F., Berta, P., and Boizet-Bonhoure, B. (2002). A nuclear export signal within the high mobility group domain regulates the nucleocytoplasmic translocation of SOX9 during sexual determination. *Proceedings of the National Academy of Sciences of the United States of America* 99, 11199–11204.

Georg, I., Bagheri-Fam, S., Knower, K.C., Wieacker, P., Scherer, G., and Harley, V.R. (2010). Mutations of the SRY-responsive enhancer of SOX9 are uncommon in XY gonadal dysgenesis. *Sexual Development : Genetics, Molecular Biology, Evolution, Endocrinology, Embryology, and Pathology of Sex Determination and Differentiation* 4, 321–325.

George, J.W., Dille, E. a, and Heckert, L.L. (2011). Current concepts of follicle-stimulating hormone receptor gene regulation. *Biology of Reproduction* 84, 7–17.

- Georges, M., Nielsen, D., Mackinnon, M., Mishra, a, Okimoto, R., Pasquino, a T., Sargeant, L.S., Sorensen, a, Steele, M.R., and Zhao, X. (1995). Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics* *139*, 907–920.
- Ghazzal, a M.H. (2006). Inguinal hernias and genital abnormalities in young Jordanian males. *Eastern Mediterranean Health Journal = La Revue de Santé de La Méditerranée Orientale = al-Majallah Al-šihhīyah Li-sharq Al-mutawassiṭ* *12*, 483–488.
- Ghirri, P., Ciulli, C., Vuerich, M., Cuttano, a, Faraoni, M., Guerrini, L., Spinelli, C., Tognetti, S., and Boldrini, a (2002). Incidence at birth and natural history of cryptorchidism: a study of 10,730 consecutive male infants. *Journal of Endocrinological Investigation* *25*, 709–715.
- Giannopoulos, M.F., Vlachakis, I.G., and Charissis, G.C. (2001). 13 Years' Experience with the Combined Hormonal Therapy of Cryptorchidism. *Hormone Research* *55*, 33–37.
- Gierl, M.S., Gruhn, W.H., von Seggern, A., Maltry, N., and Niehrs, C. (2012). GADD45G functions in male sex determination by promoting p38 signaling and Sry expression. *Developmental Cell* *23*, 1032–1042.
- Gilissen, C., Hoischen, A., Brunner, H.G., and Veltman, J. a (2011a). Unlocking Mendelian disease using exome sequencing. *Genome Biology* *12*, 228.
- Gilissen, C., Hoischen, A., Brunner, H.G., and Veltman, J. a (2011b). Unlocking Mendelian disease using exome sequencing. *Genome Biology* *12*, 228.
- Gillard, E.F., Otsu, K., Fujii, J., Khanna, V.K., de Leon, S., Derdemezi, J., Britt, B. a., Duff, C.L., Worton, R.G., and MacLennan, D.H. (1991). A substitution of cysteine for arginine 614 in the ryanodine receptor is potentially causative of human malignant hyperthermia. *Genomics* *11*, 751–755.
- De Givry, S., Bouchez, M., Chabrier, P., Milan, D., and Schiex, T. (2005). CARHTA GENE: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics (Oxford, England)* *21*, 1703–1704.
- Goddard, M.E., and Hayes, B.J. (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature Reviews. Genetics* *10*, 381–391.
- Godmann, M., Katz, J.P., Guillou, F., Simoni, M., Klaus, H., Behr, R., National, I., Recherche, D., Centre, A., and Universite, S. (2009). Krüppel-like factor 4 is involved in functional differentiation of testicular Sertoli cells. *Development Biology* *315*, 552–566.
- Goizet, C., and Lacombe, D. (2009). Le dépistage néonatal.
- Gordo, I., and Charlesworth, B. (2001). The speed of Muller's ratchet with background selection, and the degeneration of Y chromosomes. *Genetics Research* *78*, 2001.
- Gorlov, I.P., Kamat, A., Bogatcheva, N. V, Jones, E., Lamb, D.J., Truong, A., Bishop, C.E., McElreavey, K., and Agoulnik, A.I. (2002). Mutations of the GREAT gene cause cryptorchidism. *Human Molecular Genetics* *11*, 2309–2318.
- Griffiths, A.J.F., Gelbart, W.M., Miller, J.H., and Lewontin, R.C. (2001). *Analyse génétique moderne*.
- Grindflek, E., Moe, M., Taubert, H., Simianer, H., Lien, S., and Moen, T. (2006). Genome-wide linkage analysis of inguinal hernia in pigs using affected sib pairs. *BMC Genetics* *7*, 25.
- Grindflek, E., Meuwissen, T.H.E., Aasmundstad, T., Hamland, H., Hansen, M.H.S., Nome, T., Kent, M., Torjesen, P., and Lien, S. (2011a). Revealing genetic relationships between compounds affecting boar taint and reproduction in pigs. *Journal of Animal Science* *89*, 680–692.

- Grindflek, E., Lien, S., Hamland, H., Hansen, M.H.S., Kent, M., van Son, M., and Meuwissen, T.H.E. (2011b). Large scale genome-wide association and LDLA mapping study identifies QTLs for boar taint and related sex steroids. *BMC Genomics* *12*, 362.
- Grocock, C. a, Charlton, H.M., and Pike, M.C. (1988). Role of the fetal pituitary in cryptorchidism induced by exogenous maternal oestrogen during pregnancy in mice. *Journal of Reproduction and Fertility* *83*, 295–300.
- Groenen, M. a M., Wahlberg, P., Foglio, M., Cheng, H.H., Megens, H.-J., Crooijmans, R.P.M. a, Besnier, F., Lathrop, M., Muir, W.M., Wong, G.K.-S., et al. (2009). A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Research* *19*, 510–519.
- Groenen, M. a M., Archibald, A.L., Uenishi, H., Tuggle, C.K., Takeuchi, Y., Rothschild, M.F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H.-J., et al. (2012). Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* *491*, 393–398.
- Groth, K. a, Skakkebaek, A., Høst, C., Gravholt, C.H., and Bojesen, A. (2013). Clinical review: Klinefelter syndrome--a clinical update. *The Journal of Clinical Endocrinology and Metabolism* *98*, 20–30.
- Gyapay, G., Schmitt, K., Fizames, C., Jones, H., Vega-czarny, N., Spillett, D., Muselet, D., Homme, P., Dib, C., Auffray, C., et al. (1996). A radiation hybrid map of the human genome. *Human Molecular Genetics* *5*, 339–346.
- Hacker, a, Capel, B., Goodfellow, P., and Lovell-Badge, R. (1995). Expression of Sry, the mouse sex determining gene. *Development (Cambridge, England)* *121*, 1603–1614.
- Hadziselimovic, F., Hocht, B., Herzog, B., and Buser, M.W. (2007). Infertility in cryptorchidism is linked to the stage of germ cell development at orchidopexy. *Hormone Research* *68*, 46–52.
- Haldane, J.B.S. (1919). The combination of linkage values, and the calculation of distances between the loci of linked factors. *Journal of Genetics* *8*, 299–309.
- Haldane, J.B.S. (1922). Sex-ratio and unisexual sterility in hybrid animal. *Journal of Genetics* *101*–109.
- Haldane, J.B.S., Sprunt, B.A., and Haldane, N.M. (1915). Reduplication in mice.
- Haley, C.S., and Lee, G.J. (1993). Genetic basis of prolificacy in Meishan pigs. *Journal of Reproduction and Fertility. Supplement* *48*, 247–259.
- Hammes, a, Guo, J.K., Lutsch, G., Leheste, J.R., Landrock, D., Ziegler, U., Gubler, M.C., and Schedl, a (2001). Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. *Cell* *106*, 319–329.
- Han, J.I., and Na, K.J. (2011). Wnt /  $\beta$  Catenin Signaling Pathway in Canine Skin Melanoma and a Possibility as a Cancer Model for Human Skin Melanoma. In *Melanoma in the Clinic - Diagnosis, Management and Complications of Malignancy*,.
- Hancock, J.L., and Daker, M.G. (1981). Testicular hypoplasia in a boar with abnormal sex chromosome constitution ( 39 XXY ). *Journal of Reproduction and Fertility* *61*, 395–397.
- Hanley, N.A., Hagan, D.M., Ball, S.G., Strachan, T., Mcelreavey, K., Lindsay, S., Robson, S., Bullen, P., Ostrer, H., and Wilson, D.I. (2000). SRY , SOX9 , and DAX1 expression patterns during human sex determination and gonadal development. *91*, 403–407.
- Hanrahan, J.P., Gregan, S.M., Mulsant, P., Mullen, M., Davis, G.H., Powell, R., and Galloway, S.M. (2004). Mutations in the genes for oocyte-derived growth factors GDF9 and BMP15 are associated with both increased ovulation rate and sterility in Cambridge and Belclare sheep (*Ovis aries*). *Biology of Reproduction* *70*, 900–909.

- Harmegnies, N., Farnir, F., Davin, F., Buys, N., Georges, M., and Coppieters, W. (2006). Measuring the extent of linkage disequilibrium in commercial pig populations. *Animal Genetics* 37, 225–231.
- Hawken, R.J., Murtaugh, J., Flickinger, G.H., Yerle, M., Robic, a, Milan, D., Gellin, J., Beattie, C.W., Schook, L.B., and Alexander, L.J. (1999). A first-generation porcine whole-genome radiation hybrid map. *Mammalian Genome : Official Journal of the International Mammalian Genome Society* 10, 824–830.
- Hayes, H. m. (1986). Epidemiological features of 5009 cases of equine. *Equine Veterinary Journal* 18, 5009.
- Heard, E., Clerc, P., and Avner, P. (1997). X-chromosome inactivation in mammals. *Annual Review of Genetics* 31, 571–610.
- Heid, I.M., Jackson, A.U., Randall, J.C., Winkler, T.W., Steinthorsdottir, V., Thorleifsson, G., and Zillikens, M.C. (2011). Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nature Genetics* 42, 949–960.
- Helena Mangs, a, and Morris, B.J. (2007). The Human Pseudoautosomal Region (PAR): Origin, Function and Future. *Current Genomics* 8, 129–136.
- Henricson, B., and Backstrom, L. (1964). Translocation heterozygosity in a boar.
- Van Der Heyde, H., De Mets, J.P., Porreya, L., Henderickx, H., Calus, a., Bekaert, H., and Buysse, F. (1989). Influence of season, litter size, parity, gestation length, birth weight, sex and farrowing pen on frequency of congenital splayleg in piglets. *Livestock Production Science* 21, 143–155.
- Hirasawa, R., and Feil, R. (2010). Genomic imprinting and human disease . *Essays in Biochemistry* 48, 20822494.
- Hou, Y., Liu, G.E., Bickhart, D.M., Cardone, M.F., Wang, K., Kim, E.-S., Matukumalli, L.K., Ventura, M., Song, J., VanRaden, P.M., et al. (2011). Genomic characteristics of cattle copy number variations. *BMC Genomics* 12, 127.
- El Houate, B., Rouba, H., Sibai, H., Barakat, A., Chafik, A., Chadli, E.B., Imken, L., Bogatcheva, N. V, Feng, S., Agoulnik, A.I., et al. (2007). Novel mutations involving the INSL3 gene associated with cryptorchidism. *The Journal of Urology* 177, 1947–1951.
- Houmard, B., Small, C., Yang, L., Naluai-Cecchini, T., Cheng, E., Hassold, T., and Griswold, M. (2009). Global gene expression in the human fetal testis and ovary. *Biology of Reproduction* 81, 438–443.
- Hu, Z.-L., Park, C. a, Wu, X.-L., and Reecy, J.M. (2013). Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Research* 41, D871–9.
- Huang, B., Wang, S., Ning, Y., Lamb, a N., and Bartley, J. (1999). Autosomal XX sex reversal caused by duplication of SOX9. *American Journal of Medical Genetics* 87, 349–353.
- Hughes, I. a (2001). Minireview: sex differentiation. *Endocrinology* 142, 3281–3287.
- Hughes, I. a, Houk, C., Ahmed, S.F., and Lee, P. a (2006). Consensus statement on management of intersex disorders. *Journal of Pediatric Urology* 2, 148–162.
- Hunter, R.H. (1996). Aetiology of intersexuality in female (XX) pigs, with novel molecular interpretations. *Molecular Reproduction and Development* 45, 392–402.
- Husmann, D. a, and Levy, J.B. (1995). Current concepts in the pathophysiology of testicular undescend. *Urology* 46, 267–276.

Huyghe, J.R., Jackson, A.U., Fogarty, M.P., Buchkovich, M.L., Stringham, H.M., Sim, X., Yang, L., Cederberg, H., Chines, P.S., Teslovich, T.M., et al. (2013). Exome array analysis identifies novel loci and low-frequency variants for insulin processing and secretion. *Nature Genetics* 45, 197–201.

Irving, D.N. (2006). Chart of all 23 Stages , Detailed Descriptions of Stages 1 - 6 " Part One : Introduction. Lifeissues.net.

Ivell, R. (2003). The molecular basis of cryptorchidism. *Molecular Human Reproduction* 9, 175–181.

Jääskeläinen, M., Prunskaitė-Hyyryläinen, R., Naillat, F., Parviainen, H., Anttonen, M., Heikinheimo, M., Liakka, A., Ola, R., Vainio, S., Vaskivuo, T.E., et al. (2010). WNT4 is expressed in human fetal and adult ovaries and its signaling contributes to ovarian cell survival. *Molecular and Cellular Endocrinology* 317, 106–111.

Jackson, P. (2003). Performative Genders, Perverse Desires: A Bio-History of Thailand's Same-Sex and Transgender Cultures. *Intersections: Gender, History and Culture in the Asian Context*.

Jacquet, P. (2004). Sensitivity of germ cells and embryos to ionizing radiation. *18*, 15471212.

Jäger, R.J., Anvret, M., Hall, K., and Scherer, G. (1990). A human XY female with a frame shift mutation in the candidate testis-determining gene SRY. *Nature* 348, 452–454.

Jain, V., Srivastava, R., Jha, S., Misra, S., Rawat, N.S., and Amla, D.V. (2009). Study of matrix metalloproteinase-2 in inguinal hernia. *Journal of Clinical Medicine Research* 1, 285–289.

Jakob, S. (2010). FOXL2 represses the testis-specific enhancer of Sox9 to maintain ovary differentiation. University College London.

Jakob, S., and Lovell-Badge, R. (2011). Sex determination and the control of Sox9 expression in mammals. *The FEBS Journal* 278, 1002–1009.

Jakobsen, L.P., Ullmann, R., Christensen, S.B., Jensen, K.E., Mølsted, K., Henriksen, K.F., Hansen, C., Knudsen, M. a, Larsen, L. a, Tommerup, N., et al. (2007). Pierre Robin sequence may be caused by dysregulation of SOX9 and KCNJ2. *Journal of Medical Genetics* 44, 381–386.

Jakubiczka, S., Schröder, C., Ullmann, R., Volleth, M., Ledig, S., Gilberg, E., Kroisel, P., and Wieacker, P. (2010). Translocation and deletion around SOX9 in a patient with acampomelic campomelic dysplasia and sex reversal. *Sex Dev* 4, 143–149.

Jašarević, E., Geary, D.C., and Rosenfeld, C.S. (2012). Sexually selected traits: a fundamental framework for studies on behavioral epigenetics. *ILAR Journal / National Research Council, Institute of Laboratory Animal Resources* 53, 253–269.

Jenkins, J.T., and O'Dwyer, P.J. (2008). Inguinal hernias. *BMJ (Clinical Research Ed.)* 336, 269–272.

Jensen, M.S., Wilcox, A.J., Olsen, J., Bonde, J.P., Thulstrup, A.M., Ramlau-Hansen, C.H., and Henriksen, T.B. (2012). Cryptorchidism and hypospadias in a cohort of 934,538 Danish boys: the role of birth weight, gestational age, body dimensions, and fetal growth. *American Journal of Epidemiology* 175, 917–925.

Jensen-Seaman, M.I., Furey, T.S., Payseur, B.A., Lu, Y., Roskin, K.M., Chen, C.-F., Thomas, M.A., Haussler, D., and Jacob, H.J. (2004). Comparative recombination rates in the rat, mouse, and human genomes. *Genome Research* 14, 528–538.

Jiang, J., Ma, H., Lu, Y., Wang, Y., Wang, Y., Li, Q., and Ji, S. (2003). Transmission disequilibrium test for congenital dislocation of the hip and HOXB9 gene or COL1A1 gene. *Chinese Journal of Medical Genetics* 20, 193–195.

- Jiang, T., Hou, C.-C., She, Z.-Y., and Yang, W.-X. (2012). The SOX gene family: function and regulation in testis determination and male fertility maintenance. *Molecular Biology Reports* 2187–2194.
- Jiménez, R., Burgos, M., Sánchez, a, Sinclair, a H., Alarcón, F.J., Marín, J.J., Ortega, E., and Díaz de la Guardia, R. (1993). Fertile females of the mole *Talpa occidentalis* are phenotypic intersexes with ovotestes. *Development (Cambridge, England)* 118, 1303–1311.
- Jørgensen, C.B., Cirera, S., Anderson, S.I., Archibald, a L., Raudsepp, T., Chowdhary, B., Edfors-Lilja, I., Andersson, L., and Fredholm, M. (2003). Linkage and comparative mapping of the locus controlling susceptibility towards *E. COLI* F4ab/ac diarrhoea in pigs. *Cytogenetic and Genome Research* 102, 157–162.
- Jorgez, C.J., Klysik, M., Jamin, S.P., Behringer, R.R., and Matzuk, M.M. (2004). Granulosa cell-specific inactivation of follistatin causes female fertility defects. *Molecular Endocrinology (Baltimore, Md.)* 18, 953–967.
- Jost, A. (1952). Investigation of hormonal control of genesis of the sex organs in rabbit and notes on certain malformations of the genital apparatus in man. *Ginecología y Obstetricia de México* 7, 13010521.
- Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., and Kent, W.J. (2004). The UCSC Table Browser data retrieval tool. *Nucleic Acids Research* 32, D493–6.
- Kashimada, K., Pelosi, E., Chen, H., Schlessinger, D., Wilhelm, D., and Koopman, P. (2011). FOXL2 and BMP2 act cooperatively to regulate follistatin gene expression during ovarian development. *Endocrinology* 152, 272–280.
- Katoh-Fukui, Y., Owaki, A., Toyama, Y., Kusaka, M., Shinohara, Y., Maekawa, M., Toshimori, K., and Morohashi, K.-I. (2005). Mouse Polycomb M33 is required for splenic vascular and adrenal gland formation through regulating Ad4BP/SF1 expression. *Blood* 106, 1612–1620.
- Katoh-Fukui, Y., Miyabayashi, K., Komatsu, T., Owaki, A., Baba, T., Shima, Y., Kidokoro, T., Kanai, Y., Schedl, A., Wilhelm, D., et al. (2012). Cbx2, a polycomb group gene, is required for Sry gene expression in mice. *Endocrinology* 153, 913–924.
- Kauppi, L., Barchi, M., Baudat, F., Romanienko, P.J., Keeney, S., and Jasin, M. (2011). Distinct properties of the XY pseudoautosomal region crucial for male meiosis. *Science* 331, 916–920.
- Kikuchi, K., and Hamaguchi, S. (2013). Novel sex-determining genes in fish and sex chromosome evolution. *Developmental Dynamics : an Official Publication of the American Association of Anatomists* 242, 339–353.
- Kim, K.S., Larsen, N., Short, T., Plastow, G., and Rothschild, M.F. (2000). A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. *Mammalian Genome : Official Journal of the International Mammalian Genome Society* 11, 131–135.
- Kim, Y., Kobayashi, A., Sekido, R., DiNapoli, L., Brennan, J., Chaboissier, M.-C., Poulat, F., Behringer, R.R., Lovell-Badge, R., and Capel, B. (2006). Fgf9 and Wnt4 act as antagonistic signals to regulate mammalian sex determination. *PLoS Biology* 4, e187.
- Klein, R.J., Zeiss, C., Chew, E.Y., Tsai, J., Sackler, R.S., Haynes, C., Henning, A.K., Sangiovanni, J.P., Mane, S.M., Susan, T., et al. (2005). NIH Public Access. *Science* 308, 385–389.
- Knorr, C., Bornemann-Kolatzki, K., and Brenig, B. (2006). Genomic mismatch scanning (GMS) and a genome-wide microsatellite scan identify identical chromosome regions associated with porcine hernia inguinalis/scrotalis. In *Pig Genome I - An Overview of Cutting-edge Genomics with Emphasis on the Pig.*
- Knower, K.C., Kelly, S., Ludbrook, L.M., Bagheri-Fam, S., Sim, H., Bernard, P., Sekido, R., Lovell-Badge, R., and Harley, V.R. (2011). Failure of SOX9 regulation in 46XY disorders of sex development with SRY, SOX9 and SF1 mutations. *PLoS One* 6, e17751.

- Koeleman, B., Al-Ali, A., van der Laan, S.W., and Asselbergs, F.C. (2013). A concise history of genome-wide association studies. *Saudi Journal of Medicine and Medical Sciences* 1, 4.
- Kondrashov, A.S. (2003). Direct estimates of human per nucleotide mutation rates at 20 loci causing mendelian diseases. *Human Mutation* 21, 12–27.
- Kong, A., Gudbjartsson, D.F., Sainz, J., Jonsdottir, G.M., Gudjonsson, S. a, Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., et al. (2002). A high-resolution recombination map of the human genome. *Nature Genetics* 31, 241–247.
- Kong, A., Steinthorsdottir, V., Masson, G., Thorleifsson, G., Sulem, P., Besenbacher, S., Jonasdottir, A., Sigurdsson, A., Th, K., Jonasdottir, A., et al. (2013). Parental origin of sequence variants associated with complex diseases. *462*, 868–874.
- Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1991). Male development of chromosomally female mice transgenic for Sry. *Nature* 351, 117–121.
- Kopera, I., Durlej, M., Hejmej, a., Knapczyk-Stwora, K., Duda, M., Slomczynska, M., Koziowski, M., and Bilinska, B. (2010). Effects of pre- and postnatal exposure to flutamide on connexin 43 expression in testes and ovaries of prepubertal pigs. *European Journal of Histochemistry* 54, 67–73.
- Kosambi, D.D. (1944). The estimation of map distances from recombination values. *Annals of Eugenics* 12, 172–175.
- Kozakowski, J., and Zgliczyński, W. (2013). Body composition, glucose metabolism markers and serum androgens - association in women with polycystic ovary syndrome. *Endokrynologia Polska* 64, 94–100.
- Kramerov, D. a, and Vassetzky, N.S. (2005). Short retroposons in eukaryotic genomes. *International Review of Cytology* 247, 165–221.
- Kreidberg, J.A., Sariola, H., Loring, J.M., Maeda, M., Pelletier, J., Housman, D., and Jaenisch, R. (1993). WT-1 is required for early kidney development. *Cell* 74, 679–691.
- Kropatsch, R., Dekomien, G., Akkad, D. a, Gerding, W.M., Petrasch-Parwez, E., Young, N.D., Altmüller, J., Nürnberg, P., Gasser, R.B., and Epplen, J.T. (2013). SOX9 Duplication Linked to Intersex in Deer. *PloS One* 8, e73734.
- Kwok, C., Weller, P. a, Guioli, S., Foster, J.W., Mansour, S., Zuffardi, O., Punnett, H.H., Dominguez-Steglich, M. a, Brook, J.D., and Young, I.D. (1995). Mutations in SOX9, the gene responsible for Campomelic dysplasia and autosomal sex reversal. *American Journal of Human Genetics* 57, 1028–1036.
- Kyrölahti, A., Euler, R., Bielinska, M., Schoeller, E.L., Moley, K.H., Toppari, J., Heikinheimo, M., and Wilson, D.B. (2012). GATA4 regulates Sertoli cell function and fertility in adult male mice. *Molecular and Cellular Endocrinology* 333, 85–95.
- De la Chapelle, a (1987). The Y-chromosomal and autosomal testis-determining genes. *Development (Cambridge, England)* 101 Suppl, 33–38.
- Van Laere, A.-S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald, A.L., Haley, C.S., Buys, N., Tally, M., et al. (2003). A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* 425, 832–836.
- Laity, J.H., Dyson, H.J., and Wright, P.E. (2000). Molecular basis for modulation of biological function by alternate splicing of the Wilms' tumor suppressor protein. *Proceedings of the National Academy of Sciences of the United States of America* 97, 11932–11935.

- Lander, E., and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet.* *11*, 241–247.
- Larzul, C., Delaunay, I., Schwob, S., Ur, I., Motte, L., and Rheu, L. (2008). Paramètres génétiques des principales anomalies congénitales porcines. 141–142.
- Laus, E.J. (1984). inheritance of swine intersexuality (*Sus scrofa domestica*) : an experimental study. *Brazilian Journal of Genetics.*
- Lavery, R., Lardenois, A., Ranc-Jianmotamedi, F., Pauper, E., Gregoire, E.P., Vigier, C., Moreilhon, C., Primig, M., and Chaboissier, M.-C. (2011). XY Sox9 embryonic loss-of-function mouse mutants show complete sex reversal and produce partially fertile XY oocytes. *Developmental Biology* *354*, 111–122.
- Lear, T., and McGee, R. (2012). Disorders of sexual development in the domestic horse, *Equus caballus*. *Sex Dev* *6*, 61–71.
- Lee, J.H., and Jeon, J.T. (2008). Methods to detect and analyze copy number variations at the genome-wide and locus-specific levels. *Cytogenetic and Genome Research* *123*, 333–342.
- Lee, P. a, Houk, C.P., Ahmed, S.F., and Hughes, I. a (2006). Consensus statement on management of intersex disorders. *International Consensus Conference on Intersex. Pediatrics* *118*, e488–500.
- Lee, Y., Chao, A., Chen, C., Chou, T., Wang, S.M., and Wang, T. (2011). Analysis of human meiotic recombination events with a parent-sibling tracing approach. *BMC Genomics.*
- Lequarré, A., Andersson, L., André, C., Fredholm, M., Hitte, C., Leeb, T., Lohi, H., Lindblad-toh, K., and Georges, M. (2011). LUPA : A European initiative taking advantage of the canine genome architecture for unravelling complex disorders in both human and dogs. *The Veterinary Journal* *189*, 155–159.
- Li, Y., Mei, S., Zhang, X., Peng, X., Liu, G., Tao, H., Wu, H., Jiang, S., Xiong, Y., and Li, F. (2012). Identification of genome-wide copy number variations among diverse pig breeds by array CGH. *BMC Genomics* *13*, 725.
- Litten-Brown, J.C., Corson, a M., and Clarke, L. (2010). Porcine models for the metabolic syndrome, digestive and bone disorders: a general overview. *Animal : an International Journal of Animal Bioscience* *4*, 899–920.
- Liu, C.-F., Bingham, N., Parker, K., and Yao, H.H.-C. (2009). Sex-specific roles of beta-catenin in mouse gonadal development. *Human Molecular Genetics* *18*, 405–417.
- Livernois, a M., Graves, J. a M., and Waters, P.D. (2012). The origin and evolution of vertebrate sex chromosomes and dosage compensation. *Heredity* *108*, 50–58.
- Lourenço, D., Brauner, R., Rybczynska, M., Nihoul-Fékété, C., McElreavey, K., and Bashamboo, A. (2011). Loss-of-function mutation in GATA4 causes anomalies of human testicular development. *Proceedings of the National Academy of Sciences of the United States of America* *108*, 1597–1602.
- Lunney, J.K. (2007). Advances in swine biomedical model genomics. *International Journal of Biological Sciences* *3*, 179–184.
- Luo, W., Chen, S., Cheng, D., Wang, L., Li, Y., Ma, X., Song, X., Liu, X., Li, W., Liang, J., et al. (2012). Genome-wide association study of porcine hematological parameters in a Large White × Minzhu F2 resource population. *International Journal of Biological Sciences* *8*, 870–881.
- Luo, X., Ikeda, Y., and Parker, K.L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell* *77*, 481–490.

Ma, J., Iannuccelli, N., Duan, Y., Huang, W., Guo, B., Riquet, J., Huang, L., and Milan, D. (2010). Recombinational landscape of porcine X chromosome and individual variation in female meiotic recombination associated with haplotypes of Chinese pigs. *BMC Genomics* 11, 159.

Ma, J., Gilbert, H., Iannuccelli, N., Duan, Y., Guo, B., Huang, W., Ma, H., Riquet, J., Bidanel, J.-P., Huang, L., et al. (2013). Fine mapping of fatness QTL on porcine chromosome X and analyses of three positional candidate genes. *BMC Genetics* 14, 46.

Magenis, R.E., Webb, M.J., McKean, R.S., Tomar, D., Allen, L.J., Kammer, H., Van Dyke, D.L., and Lovrien, E. (1982). Translocation(X;Y)(p22.33;p11.2) in XX males: etiology of male phenotype. *Human Genetics* 62, 271–276.

Mäkinen, a, Andersson, M., and Nikunen, S. (1998). Detection of the X chromosomes in a Klinefelter boar using a whole human X chromosome painting probe. *Animal Reproduction Science* 52, 317–323.

Mandel, H., Shemer, R., Borochowitz, Z.U., Okopnik, M., Knopf, C., Indelman, M., Drugan, A., Tiosano, D., Gershoni-baruch, R., and Choder, M. (2008). SERKAL Syndrome : An Autosomal-Recessive Disorder Caused by a Loss-of-Function Mutation in WNT4. *The American Journal of Human Genetics* 39–47.

Manuylov, N.L., Zhou, B., Ma, Q., Fox, S.C., Pu, W.T., and Tevosian, S.G. (2011). Conditional Ablation of Gata4 and Fog2 Genes in Mice Reveals their Distinct Roles in Mammalian Sexual Differentiation. *Development Biology* 353, 229–241.

Marian, A.J. (2012). Molecular Genetic Studies of Complex Phenotypes. *Translational Research* 159, 64–79.

Marklund, L., Winterö, a K., Thomsen, P.D., Johansson, M., Fredholm, M., Gustafsson, U., and Andersson, L. (1993). A linkage group on pig chromosome 4 comprising the loci for blood group L, GBA, ATP1B1 and three microsatellites. *Animal Genetics* 24, 333–338.

Marklund, S., Kijas, J., Rodriguez-Martinez, H., Rönstrand, L., Funa, K., Moller, M., Lange, D., Edfors-Lilja, I., and Andersson, L. (1998). Molecular basis for the dominant white phenotype in the domestic pig. *Genome Research* 8, 826–833.

Marumudi, E., Khadgawat, R., Surana, V., Shabir, I., Joseph, A., and Ammini, A.C. (2013). Diagnosis and management of classical congenital adrenal hyperplasia. *Steroids* 78, 741–746.

Matson, C.K., Murphy, M.W., Sarver, A.L., Griswold, M.D., and Vivian, J. (2012). mammalian testis. *Nature* 476, 101–104.

Matthews, R.D., and Neumayer, L. (2008). Inguinal hernia in the 21st century: an evidence-based review. *Current Problems in Surgery* 45, 261–312.

Mattos, M.R.F., Somões-Mattos, L., and Domingues, S.F.S. (2000). Cryptorchidism in dog. *Ciência Animal* 2000 10, 61–70.

Mattsson, P. (2011). Prevalence of congenital defects in Swedish Hampshire , Landrace and Yorkshire pig breeds and opinions on their prevalence in Swedish commercial herds.

McCarroll, S. a, Kuruvilla, F.G., Korn, J.M., Cawley, S., Nemes, J., Wysoker, A., Shapero, M.H., de Bakker, P.I.W., Maller, J.B., Kirby, A., et al. (2008a). Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature Genetics* 40, 1166–1174.

McCarroll, S. a, Kuruvilla, F.G., Korn, J.M., Cawley, S., Nemes, J., Wysoker, A., Shapero, M.H., de Bakker, P.I.W., Maller, J.B., Kirby, A., et al. (2008b). Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature Genetics* 40, 1166–1174.

- McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P. a, and Hirschhorn, J.N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews. Genetics* 9, 356–369.
- McClintock, B. (1983). The origin and behavior of matable loci in maize. *Genetics*.
- Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit : A MapReduce framework for analyzing next-generation DNA sequencing data. 1297–1303.
- McLaren, A. (1998). Gonad development: Assembling the mammalian testis. *Current Biology*.
- McLaren, A. (2003). Primordial germ cells in the mouse. *Developmental Biology* 262, 1–15.
- McMahon, D.R., Kramer, S. a, and Husmann, D. a (1995). Antiandrogen induced cryptorchidism in the pig is associated with failed gubernacular regression and epididymal malformations. *The Journal of Urology* 154, 553–557.
- Mead, T.J., Wang, Q., Bhattaram, P., Dy, P., Afelik, S., Jensen, J., and Lefebvre, V. (2013). A far-upstream (-70 kb) enhancer mediates Sox9 auto-regulation in somatic tissues during development and adult regeneration. *Nucleic Acids Research* gkt140–.
- Meijerink, E., Neuenschwander, S., Fries, R., Dinter, a, Bertschinger, H.U., Stranzinger, G., and Vögeli, P. (2000). A DNA polymorphism influencing alpha(1,2)fucosyltransferase activity of the pig FUT1 enzyme determines susceptibility of small intestinal epithelium to Escherichia coli F18 adhesion. *Immunogenetics* 52, 129–136.
- Mertes, F., Elsharawy, A., Sauer, S., van Helvoort, J.M.L.M., van der Zaag, P.J., Franke, A., Nilsson, M., Lehrach, H., and Brookes, A.J. (2011). Targeted enrichment of genomic DNA regions for next-generation sequencing. *Briefings in Functional Genomics* 10, 374–386.
- Merveille, A.-C., Davis, E.E., Becker-Heck, A., Legendre, M., Amirav, I., Bataille, G., Belmont, J., Beydon, N., Billen, F., Clément, A., et al. (2012). CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs. 43, 72–78.
- Meuwissen, T.H., Hayes, B.J., and Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829.
- Meyer, A.L.M., Berger, E., Monteiro Jr, O., Alonso, P.A., Stavale, J.N., and Gonçalves, M.P.S. (2007). Quantitative and qualitative analysis of collagen types in the fascia transversalis of inguinal hernia patients. *Arq Gastroenterol* 230–234.
- Meyers-Wallen, V.N. (2012). Gonadal and sex differentiation abnormalities of dogs and cats. *Sexual Development : Genetics, Molecular Biology, Evolution, Endocrinology, Embryology, and Pathology of Sex Determination and Differentiation* 6, 46–60.
- Meyers-Wallen, V.N., Schlafer, D., Barr, I., Lovell-Badge, R., and Keyzner, a (1999). Sry-negative XX sex reversal in purebred dogs. *Molecular Reproduction and Development* 53, 266–273.
- Miang-Lon Ng, P., and Lufkin, T. (2012). Embryonic stem cells: protein interaction networks. *Biomol Concepts* 2, 13–25.
- Mikami, H., and Fredeen, H.T. (1979). A genetic study of cryptorchidism and scrotal hernia in pigs. *Canadian Journal of Genetics and Cytology. Journal Canadien de Génétique et de Cytologie* 21, 9–19.

- Milan, D., Hawken, R., Cabau, C., Leroux, S., Genet, C., Lahbib, Y., Tosser, G., Robic, a, Hatey, F., Alexander, L., et al. (2000a). IMpRH server: an RH mapping server available on the Web. *Bioinformatics (Oxford, England)* *16*, 558–559.
- Milan, D., Jeon, J.-T., Looft, C., Amarger, V., Robic, A., Thelander, M., Rogel-Gaillard, C., Paul, S., Iannuccelli, N., Rask, L., et al. (2000b). A Mutation in PRKAG3 Associated with Excess Glycogen Content in Pig Skeletal Muscle. *Science* *288*, 1248–1251.
- Miller, N.A., Lue, S.J. Van, and Rawlings, C.A. (2004). Use of laparoscopic-assisted cryptorchidectomy in dogs and cats. *J Am Vet Med Assoc.* 875–878.
- Mills, R.E., Bennett, E.A., Iskow, R.C., and Devine, S.E. (2007). Which transposable elements are active in the human genome? *Trends in Genetics : TIG* *23*, 183–191.
- Minozzi, G., Williams, J.L., Stella, A., Strozzi, F., Luini, M., Settles, M.L., Taylor, J.F., Whitlock, R.H., Zanella, R., and Neibergs, H.L. (2012). Meta-analysis of two genome-wide association studies of bovine paratuberculosis. *PLoS One* *7*, e32578.
- Miyamoto, N., Yoshida, M., Kuratani, S., Matsuo, I., and Aizawa, S. (1997). Defects of urogenital development in mice lacking *Emx2*. *Development (Cambridge, England)* *124*, 1653–1664.
- Moniot, B., Declosmenil, F., Barrionuevo, F., Scherer, G., Aritake, K., Malki, S., Marzi, L., Cohen-Solal, A., Georg, I., Klattig, J., et al. (2009). The PGD2 pathway, independently of FGF9, amplifies SOX9 activity in Sertoli cells during male sexual differentiation. *Development (Cambridge, England)* *136*, 1813–1821.
- Montazer-Torbati, F., Kocer, A., Auguste, A., Renault, L., Charpigny, G., Pailhoux, E., and Pannetier, M. (2010). A study of goat SRY protein expression suggests putative new roles for this gene in the developing testis of a species with long-lasting SRY expression. *Developmental Dynamics : an Official Publication of the American Association of Anatomists* *239*, 3324–3335.
- Morais da Silva, S., Hacker, A., Harley, V., Goodfellow, P., Swain, A., and Lovell-Badge, R. (1996a). Sox9 expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. *Nature Genetics*.
- Morais da Silva, S., Hacker, A., Harley, V., Goodfellow, P., Swain, A., and Lovell-Badge, R. (1996b). Sox9 expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. *Nature Genetics* *14*, 62–68.
- Mormède, P., and Mercat, M.-J. (2002). génétiques sur les processus d'adaptation et le comportement alimentaire chez le porc. *INRA Productions Animales* *15*, 349–356.
- Muller, H.J. (1964). the Relation of Recombination To Mutational Advance. *Mutation Research* *106*, 2–9.
- Muller, C., Denis, M., Gentzbittel, L., and Faraut, T. (2004). The Iccare web server: an attempt to merge sequence and mapping information for plant and animal species. *Nucleic Acids Research* *32*, W429–34.
- Mulley, R.C., and Edwards, M.J. (1984). Prevalence of congenital abnormalities in pigs. *Australian Veterinary Journal* *61*, 116–120.
- Murgiano, L., Tammen, I., Harlizius, B., and Drögemüller, C. (2012). A de novo germline mutation in MYH7 causes a progressive dominant myopathy in pigs. *BMC Genetics* *13*, 99.
- Myers, S., Freeman, C., Auton, A., Donnelly, P., and G, M. (2008a). A common sequence motif associated with recombination hot spots and genome instability in humans. *Nature Genetics* *40*, 2008.

- Myers, S., Freeman, C., Auton, A., Donnelly, P., and G, M. (2008b). A common sequence motif associated with recombination hot spots and genome instability in humans. *Nature Genetics* 40, 2008.
- Nagamine, C.M., Morohashi, K., Carlisle, C., and Chang, D.K. (1999). Sex reversal caused by *Mus musculus* domesticus Y chromosomes linked to variant expression of the testis-determining gene *Sry*. *Developmental Biology* 216, 182–194.
- Nagaraja, R., MacMillan, S., Kere, J., Jones, C., Griffin, S., Schmatz, M., Terrell, J., Shomaker, M., Jermak, C., Hott, C., et al. (1997). X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. *Genome Research* 7, 210–222.
- Natoli, J.L., Ackerman, D.L., McDermott, S., and Edwards, J.G. (2012). Prenatal diagnosis of Down syndrome: a systematic review of termination rates (1995-2011). *Prenatal Diagnosis* 32, 142–153.
- Nef, S., and Parada, L.F. (1999). Cryptorchidism in mice mutant for *Insl3*. *Nature Genetics* 22, 295–299.
- Nef, S., Verma-Kurvari, S., Merenmies, J., Vassalli, J.-D., Efstratiadis, A., Accili, D., and Parada, L.F. (2003). Testis determination requires insulin receptor family function in mice. *Nature* 426, 291–295.
- Newton-cheh, C., Johnson, T., Gateva, V., Tobin, M.D., Bochud, M., Coin, L., Najjar, S.S., Zhao, J.H., Heath, S.C., Eyheramendy, S., et al. (2009). Eight blood pressure loci identified by genome-wide association study of 34,433 people of European ancestry. *Nature Genetics* 41, 666–676.
- Ninomiya, S., Isomura, M., Narahara, K., Seino, Y., and Nakamura, Y. (1996). Isolation of a testis-specific cDNA on chromosome 17q from a region adjacent to the breakpoint of t(12;17) observed in a patient with acampomelic campomelic dysplasia and sex reversal. *Human Molecular Genetics* 5, 69–72.
- Nowacka, J., Nizanski, W., Klimowicz, M., Dzimira, S., and Switonski, M. (2005). Lack of the *SOX9* gene polymorphism in sex reversal dogs (78,XX; *SRY* negative). *The Journal of Heredity* 96, 797–802.
- Nsengimana, J., Baret, P., Haley, C.S., and Visscher, P.M. (2004). Linkage disequilibrium in the domesticated pig. *Genetics* 166, 1395–1404.
- Observatoire Bruxellois de la Consommation Durable (2007). Consommation de viande : un lourd tribut environnemental.
- Öçal, G. (2011). Current concepts in disorders of sexual development. *Journal of Clinical Research in Pediatric Endocrinology* 3, 105–114.
- Ollivier, L. (1979). Les anomalies héréditaires dans l'espèce porcine. *Journées Recherche Porcine* 371–382.
- Ollivier, L., and Sellier, P. (1982). Pig genetics : a review.
- OMS (2009). International statistical classification of diseases, 10th [tenth] revision ICD-10 (Genève).
- OMS (2012). Congenital anomalies.
- Ono, M., and Harley, V.R. (2013). Disorders of sex development: new genes, new concepts. *Nature Reviews Endocrinology*.
- Onteru, S.K., Fan, B., Nikkilä, M.T., Garrick, D.J., Stalder, K.J., and Rothschild, M.F. (2011). Whole-genome association analyses for lifetime reproductive traits in the pig. *Journal of Animal Science* 89, 988–995.

Orton, F., Rosivatz, E., Scholze, M., and Kortenkamp, A. (2012). Competitive Androgen Receptor Antagonism as a Factor Determining the Predictability of Cumulative Anti androgenic Effects of Widely Used Pesticides. 1578–1584.

Ott, J. (1999). Analysis of human genetic linkage.

Ottolenghi, C., Omari, S., Garcia-Ortiz, J.E., Uda, M., Crisponi, L., Forabosco, A., Pilia, G., and Schlessinger, D. (2005). Foxl2 is required for commitment to ovary differentiation. *Human Molecular Genetics* 14, 2053–2062.

Pailhoux, E., Pelliniemi, L., Barbosa, A., Parma, P., Kuopio, T., and Cotinot, C. (1997). Relevance of intersexuality to breeding and reproductive biotechnology programs XX sex reversal in pigs. *Theriogenology*.

Pailhoux, E., Vaiman, D., Barboux, S., Cribiu, E., Cotinot, C., and Fellous, M. (1998). Etude de l'intersexualité chez l'homme et l'animal domestique. Apport en matière de différenciation sexuelle chez les mammifères.

Pailhoux, E., Parma, P., Sundström, J., Vigier, B., Serval, N., Kuopio, T., Locatelli, a, Pelliniemi, L.J., and Cotinot, C. (2001a). Time course of female-to-male sex reversal in 38,XX fetal and postnatal pigs. *Developmental Dynamics : an Official Publication of the American Association of Anatomists* 222, 328–340.

Pailhoux, E., Vigier, B., Vaiman, D., Schibler, L., Vaiman, a, Cribiu, E., Nezer, C., Georges, M., Sundström, J., Pelliniemi, L.J., et al. (2001b). Contribution of domestic animals to the identification of new genes involved in sex determination. *The Journal of Experimental Zoology* 290, 700–708.

Pailhoux, E., Vigier, B., Chaffaux, S., Serval, N., Taourit, S., Furet, J.P., Fellous, M., Grosclaude, F., Cribiu, E.P., Cotinot, C., et al. (2001c). A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nature Genetics* 29, 453–458.

Pailhoux, E., Mandon-Pepin, B., and Cotinot, C. (2001d). Mammalian gonadal differentiation: the pig model. *Reprod Suppl* 58, 65–80.

Pannetier, M., Elzaiat, M., Thépot, D., and Pailhoux, E. (2012). Telling the story of XX sex reversal in the goat: highlighting the sex-crossroad in domestic mammals. *Sex Dev.* 6, 33–45.

Papatsiros, V.G. (2012). The Splay Leg Syndrome in Piglets : A Review. *American Journal of Animal and Veterinary Sciences* 7, 80–83.

Park, S.Y., and Jameson, J.L. (2005). Minireview: transcriptional regulation of gonadal development and differentiation. *Endocrinology* 146, 1035–1042.

Parma, P., Pailhoux, E., and Cotinot, C. (1999). Reverse transcription-polymerase chain reaction analysis of genes involved in gonadal differentiation in pigs. *Biology of Reproduction* 61, 741–748.

Parma, P., Radi, O., Vidal, V., Chaboissier, M.C., Dellambra, E., Valentini, S., Guerra, L., Schedl, A., and Camerino, G. (2006). R-spondin1 is essential in sex determination, skin differentiation and malignancy. *Nature Genetics* 38, 1304–1309.

Partlow, G.D., Fisher, K.R., Page, P.D., MacMillan, K., and Walker, a F. (1993). Prevalence and types of birth defects in Ontario swine determined by mail survey. *Canadian Journal of Veterinary Research = Revue Canadienne de Recherche Vétérinaire* 57, 67–73.

Pascal, L., Rambourg Schepens, M.-O., and Lasalle, J.-L. (2001). Les malformations congénitales du petit garçon en Languedoc-Roussillon - Étude des cryptorchidies et des hypospadias opérés à partir des données du PMSI de 1998 à 2001.

- Pascual, G., Corrales, C., Gómez-Gil, V., Buján, J., and Bellón, J.M. (2007). TGF-beta1 overexpression in the transversalis fascia of patients with direct inguinal hernia. *European Journal of Clinical Investigation* 37, 516–521.
- Pasternak, B. (2008). Towards surgical use of matrix metalloproteinase biology.
- Pathirana, I.N., Yamasaki, H., Kawate, N., Tsuji, M., Büllesbach, E.E., Takahashi, M., Hatoya, S., Inaba, T., and Tamada, H. (2012). Plasma insulin-like peptide 3 and testosterone concentrations in male dogs : changes with age and effects of cryptorchidism . *Theriogenology* 77, 2011–2012.
- Paty, a.-C., Gomes do Espirito Santo, E., and Suzan, F. (2012). Étude des cryptorchidies et hypospadias opérés en France de 1998 à 2008 chez le petit garçon âgé de moins de sept ans à partir des données du Programme de médicalisation des systèmes d'information (PMSI). *Revue d'Épidémiologie et de Santé Publique* 60, S30.
- Pearlman, A., Loke, J., Le Caignec, C., White, S., Chin, L., Friedman, A., Warr, N., Willan, J., Brauer, D., Farmer, C., et al. (2010). Mutations in MAP3K1 cause 46,XY disorders of sex development and implicate a common signal transduction pathway in human testis determination. *American Journal of Human Genetics* 87, 898–904.
- Pereira, P.M.D.C.C., and Vicente, A.F.D.R.B. (2013). Meat nutritional composition and nutritive role in the human diet. *Meat Science* 93, 586–592.
- Perry, G.H., Dominy, N.J., Claw, K.G., Lee, A.S., Fiegler, H., Redon, R., Werner, J., Villanea, F. a, Mountain, J.L., Misra, R., et al. (2007). Diet and the evolution of human amylase gene copy number variation. *Nature Genetics* 39, 1256–1260.
- Pienkowskic, C., and Cartlault, A. (2008). Anomalie du développement sexuel du nouveau né.
- Pilon, N., Daneau, I., Paradis, V., Hamel, F., Lussier, J.G., Viger, R.S., and Silversides, D.W. (2003). Porcine SRY promoter is a target for steroidogenic factor 1. *Biology of Reproduction* 68, 1098–1106.
- Pinton, A., Barasc, H., Raymond Letron, I., Bordedebat, M., Mary, N., Massip, K., Bonnet, N., Calgaro, A., Dudez, a M., Feve, K., et al. (2011). Meiotic studies of a 38,XY/39,XXY mosaic boar. *Cytogenetic and Genome Research* 133, 202–208.
- Pinton, A., Calgaro, A., Bonnet, N., Mary, N., Dudez, A.M., Barasc, H., Plard, C., Yerle, M., and Ducos, A. (2012). Contrôle chromosomique des populations porcines en France : 43–44.
- Pitetti, J.-L., Calvel, P., Romero, Y., Conne, B., Truong, V., Papaioannou, M.D., Schaad, O., Docquier, M., Herrera, P.L., Wilhelm, D., et al. (2013). Insulin and IGF1 receptors are essential for XX and XY gonadal differentiation and adrenal development in mice. *PLoS Genetics* 9, e1003160.
- Polanco, J.C., and Koopman, P. (2007). Sry and the hesitant beginnings of male development. *Developmental Biology* 302, 13–24.
- Polanco, J.C., Wilhelm, D., Davidson, T.-L., Knight, D., and Koopman, P. (2010). Sox10 gain-of-function causes XX sex reversal in mice: implications for human 22q-linked disorders of sex development. *Human Molecular Genetics* 19, 506–516.
- Ponsuksili, S., Murani, E., Brand, B., Schwerin, M., and Wimmers, K. (2011). Integrating expression profiling and whole-genome association for dissection of fat traits in a porcine model. *Journal of Lipid Research* 52, 668–678.
- Pritchard, J.K. (2001). Are rare variants responsible for susceptibility to complex diseases? *American Journal of Human Genetics* 69, 124–137.

- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. a R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* *81*, 559–575.
- Quigley, C., De Bellis, A., Marschke, K., El-Awady, M., Wilson, E., and French, F. (1995). Androgen receptor defects : historical , clinical , and molecular. *Endocr Rev.* *16*, 7671849.
- Quilter, C.R., Blott, S.C., Mileham, A.J., Nabeel, A.A., Sargent, C.A., and Gri, D.K. (2002). A mapping and evolutionary study of porcine sex chromosome genes. *Mammalian Genome* *594*, 588–594.
- Rainelli, P. (2001). L'image de la viande de Porc en France : attitudes des consommateurs.
- Ramayo-Caldas, Y., Castelló, A., Pena, R.N., Alves, E., Mercadé, A., Souza, C. a, Fernández, A.I., Perez-Enciso, M., and Folch, J.M. (2010). Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. *BMC Genomics* *11*, 593.
- Ramos, A.M., Crooijmans, R.P.M. a, Affara, N. a, Amaral, A.J., Archibald, A.L., Beever, J.E., Bendixen, C., Churcher, C., Clark, R., Dehais, P., et al. (2009). Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS One* *4*, e6524.
- Ramos, A.M., Duijvesteijn, N., Knol, E.F., Merks, J.W.M., Bovenhuis, H., Crooijmans, R.P.M. a, Groenen, M. a M., and Harlizius, B. (2011). The distal end of porcine chromosome 6p is involved in the regulation of skatole levels in boars. *BMC Genetics* *12*, 35.
- Raudsepp, T., and Chowdhary, B.P. (2008). The horse pseudoautosomal region (PAR): characterization and comparison with the human, chimp and mouse PARs. *Cytogenetic and Genome Research* *121*, 102–109.
- Read, A., and Donnai, D. (2008). Génétique médicale ; de la biologie à la pratique clinique.
- Reich, D.E., and Lander, E.S. (2001). On the allelic spectrum of human disease. *Trends in Genetics* *17*, 502–510.
- Reiner, G., Eckert, J., Peischl, T., Bochert, S., Jäkel, T., Mackenstedt, U., Joachim, a., Dausgchies, a., and Geldermann, H. (2002). Variation in clinical and parasitological traits in Pietrain and Meishan pigs infected with *Sarcocystis miescheriana*. *Veterinary Parasitology* *106*, 99–113.
- Ren, J., Mao, H., Zhang, Z., Xiao, S., Ding, N., and Huang, L. (2011). A 6-bp deletion in the TYRP1 gene causes the brown colouration phenotype in Chinese indigenous pigs. *Heredity* *106*, 862–868.
- Rettenberger, G., Bruch, J., Fries, R., Archibald, a L., and Hameister, H. (1996). Assignment of 19 porcine type I loci by somatic cell hybrid analysis detects new regions of conserved synteny between human and pig. *Mammalian Genome : Official Journal of the International Mammalian Genome Society* *7*, 275–279.
- Richter, A., Wissel, J., Harlizius, B., Simon, D., Schelosky, L., Scholz, U., Poewe, W., and Löscher, W. (1995). The “Campus Syndrome” in Pigs: Neurological, Neurophysiological, and Neuropharmacological Characterization of a New Genetic Animal Model of High-Frequency Tremor. *Experimental Neurology*.
- Risch, N., and Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science* *273*, 1516–1517.
- Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2012). Integrative genomics viewer. *Nature Biotechnology* *29*, 24–26.
- Rodrigues Junior, A.J., Rodrigues, C.J., da Cunha, A.C.P., and Jin, Y. (2002). Quantitative analysis of collagen and elastic fibers in the transversalis fascia in direct and indirect inguinal hernia. *Revista Do Hospital Das Clínicas* *57*, 265–270.

Rohrer, G. a, Alexander, L.J., Hu, Z., Smith, T.P., Keele, J.W., and Beattie, C.W. (1996). A comprehensive map of the porcine genome. *Genome Research* 6, 371–391.

Rosch, R., Junge, K., Lynen, P., Mertens, P.R., Klinge, U., and Schumpelick, V. (2003). Hernia - A Collagen Disease? *European Surgery-Acta Chirurgica Austriaca* 35, 11–15.

Ross, D.G.F., Bowles, J., Koopman, P., and Lehnert, S. (2008). New insights into SRY regulation through identification of 5' conserved sequences. *BMC Molecular Biology* 9, 85.

Rothschild, M.F. (2004). Porcine genomics delivers new tools and results: this little piggy did more than just go to market. *Genetical Research* 83, 1–6.

Rothschild, M.F., and Ruvinsky, A. (2011). *The Genetics of the Pig*, 2nd Edition.

Rothschild, M., Jacobson, C., Vaske, D., Tuggle, C., Wang, L., Short, T., Eckardt, G., Sasaki, S., Vincent, a, McLaren, D., et al. (1996). The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proceedings of the National Academy of Sciences of the United States of America* 93, 201–205.

Rothschild, M.F., Christian, L.L., and Blanchard, W. (1988). Evidence for multigene control of cryptorchidism in swine. *The Journal of Heredity* 79, 313–314.

Rothschild, M.F., Messer, L., Day, a, Wales, R., Short, T., Southwood, O., and Plastow, G. (2000). Investigation of the retinol-binding protein 4 (RBP4) gene as a candidate gene for increased litter size in pigs. *Mammalian Genome : Official Journal of the International Mammalian Genome Society* 11, 75–77.

Rouault, K., Scotet, V., Autret, S., Gaucher, F., Dubrana, F., Tanguy, D., Rassi, C.Y. El, Fenoll, B., and Férec, C. (2010). Evidence of association between GDF5 polymorphisms and congenital dislocation of the hip in a Caucasian population Objective. *Osteoarthritis and Cartilage* 18, 2010.

Roy, M. (2011). La problématique des hernies dans les élevages porcins.

Rudkin, G.T., and Stollar, B.D. (1977). High resolution detection of DNA-RNA hybrids in situ by indirect immunofluorescence. *Nature* 265, 5593.

Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z.P., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J. V, Patterson, N.J., Mcdonald, G.J., et al. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419.

Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., et al. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409, 928–933.

De Santa Barbara, P., Méjean, C., Moniot, B., Malclès, M.H., Berta, P., and Boizet-Bonhoure, B. (2001). Steroidogenic factor-1 contributes to the cyclic-adenosine monophosphate down-regulation of human SRY gene expression. *Biology of Reproduction* 64, 775–783.

Schaeffer, L.R. (2006). Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics = Zeitschrift Für Tierzüchtung Und Züchtungsbiologie* 123, 218–223.

Schibler, L., Vaiman, D., Oustry, a, Guinec, N., Dangy-Caye, a L., Billault, a, and Cribiu, E.P. (1998). Construction and extensive characterization of a goat bacterial artificial chromosome library with threefold genome coverage. *Mammalian Genome : Official Journal of the International Mammalian Genome Society* 9, 119–124.

Schlotz, W., and Phillips, D.I.W. (2009). Fetal origins of mental health: evidence and mechanisms. *Brain, Behavior, and Immunity* 23, 905–916.

- Schneider, J. a, Peto, T.E. a, Boone, R. a, Boyce, A.J., and Clegg, J.B. (2002). Direct measurement of the male recombination fraction in the human beta-globin hot spot. *Human Molecular Genetics* *11*, 207–215.
- Van der Schoot, P. (1992). Disturbed testicular descent in the rat after prenatal exposure to the antiandrogen flutamide. *Journal of Reproduction and Fertility* *96*, 483–496.
- Schull, W.J., and Otake, M. (1999). Cognitive function and prenatal exposure to ionizing radiation. *Teratology* *59*, 222–226.
- Schwob, S., Riquet, J., Bellec, T., Kernaleguen, L., Tribout, T., and Bidanel, J.-P. (2009). Mise en place d' un programme de sélection assistée par marqueurs dans la population sino-européenne Duochan. *Journées Recherche Porcine* *1*, 1–2.
- Seeherunvong, T., Perera, E.M., Bao, Y., Benke, P.J., Benigno, A., Donahue, R.P., and Berkovitz, G.D. (2004). 46,XX sex reversal with partial duplication of chromosome arm 22q. *American Journal of Medical Genetics. Part A* *127A*, 149–151.
- Sekido, R., and Lovell-Badge, R. (2008). Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature* *453*, 930–934.
- Sekido, R., and Lovell-Badge, R. (2009). Sex determination and SRY: down to a wink and a nudge? *Trends in Genetics : TIG* *25*, 19–29.
- Sela, N., Kim, E., and Ast, G. (2010). The role of transposable elements in the evolution of non-mammalian vertebrates and invertebrates. *Genome Biology* *11*, R59.
- Servin, B., de Givry, S., and Faraut, T. (2010a). Statistical confidence measures for genome maps: application to the validation of genome assemblies. *Bioinformatics (Oxford, England)* *26*, 3035–3042.
- Servin, B., de Givry, S., and Faraut, T. (2010b). Statistical confidence measures for genome maps: application to the validation of genome assemblies. *Bioinformatics (Oxford, England)* *26*, 3035–3042.
- Servin, B., Faraut, T., Iannuccelli, N., Zelenika, D., and Milan, D. (2012). High-resolution autosomal radiation hybrid maps of the pig genome and their contribution to the genome sequence assembly. *BMC Genomics* *13*, 585.
- She, X., Cheng, Z., Zöllner, S., Church, D.M., and Eichler, E.E. (2008). Mouse segmental duplication and copy number variation. *Nature Genetics* *40*, 909–914.
- Shearin, A.L., and Ostrander, E. a (2010). Leading the way: canine models of genomics and disease. *Disease Models & Mechanisms* *3*, 27–34.
- Sherwood, L. (2006). *Physiologie humaine: A Human Perspective*.
- Shifman, S., Bell, J.T., Copley, R.R., Taylor, M.S., Williams, R.W., Mott, R., and Flint, J. (2006). A high-resolution single nucleotide polymorphism genetic map of the mouse genome. *PLoS Biology* *4*, e395.
- Silva, K., Bastiaansen, J., Knol, E., Merks, J., Lopes, P., Guimarães, S., and van Arendonk, J. (2011). Meta-analysis of results from quantitative trait loci mapping studies on pig chromosome 4. *Animal Genetics* *42*, 2145.
- Sinclair, a H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, a M., Lovell-Badge, R., and Goodfellow, P.N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* *346*, 240–244.

- Sironen, a, Uimari, P., Iso-Touru, T., and Vilkki, J. (2012). L1 insertion within SPEF2 gene is associated with increased litter size in the Finnish Yorkshire population. *Journal of Animal Breeding and Genetics = Zeitschrift Für Tierzucht Und Züchtungsbiologie* 129, 92–97.
- Sironen, A., Thomsen, B., Andersson, M., Ahola, V., and Vilkki, J. (2006). An intronic insertion in KPL2 results in aberrant splicing and causes the immotile short-tail sperm defect in the pig. *Proceedings of the National Academy of Sciences of the United States of America* 103, 5006–5011.
- Sironen, A., Vilkki, J., Bendixen, C., and Thomsen, B. (2007). Infertile Finnish Yorkshire boars carry a full-length LINE-1 retrotransposon within the KPL2 gene. *Molecular Genetics and Genomics : MGG* 278, 385–391.
- Sironen, A., Uimari, P., Nagy, S., Paku, S., Andersson, M., and Vilkki, J. (2010a). Knobbed acrosome defect is associated with a region containing the genes STK17b and HECW2 on porcine chromosome 15. *BMC Genomics* 11, 699.
- Sironen, A., Uimari, P., Nagy, S., Paku, S., Andersson, M., and Vilkki, J. (2010b). Knobbed acrosome defect is associated with a region containing the genes STK17b and HECW2 on porcine chromosome 15. *BMC Genomics* 11, 699.
- Skinner, B.M., Lachani, K., Sargent, C. a, and Affara, N. a (2013). Regions of XY homology in the pig X chromosome and the boundary of the pseudoautosomal region. *BMC Genetics* 14, 3.
- Smit, A., Hubley, R., and Green, P. (2010). RepeatMasker Open-3.2.8.
- Smith, J., Sones, K., Grace, D., MacMillan, S., Tarawali, S., and Herrero, M. (2012). Beyond milk, meat, and eggs: Role of livestock in food and nutrition security. *Animal Frontiers* 3, 6–13.
- Sotoodehnia, N., Isaacs, A., Bakker, P.I.W. de, Dörr, M., Newton-Cheh, C., Nolte, I.M., Harst, P. van der, Müller, M., Eijgelsheim, M., and Alonso, A. (2010). Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. *Nature Genetics* 42, 1068–1076.
- Speedy, A.W. (2003). Global production and consumption of animal source foods. *The Journal of Nutrition* 4048–4053.
- Speek, M. (2001). Antisense Promoter of Human L1 Retrotransposon Drives Transcription of Adjacent Cellular Genes Antisense Promoter of Human L1 Retrotransposon Drives Transcription of Adjacent Cellular Genes. *Molecular and Cellular Biology* 21.
- Stankiewicz, P., and Lupski, J.R. (2010). Structural variation in the human genome and its role in disease. *Annual Review of Medicine* 61, 437–455.
- Steemers, F.J., and Gunderson, K.L. (2007). Whole genome genotyping technologies on the BeadArray platform. *Biotechnology Journal* 2, 41–49.
- Stewart, E.A., Mckusick, K.B., Aggarwal, A., Brady, S., Chu, A., Fang, N., Hadley, D., Hussain, S., Lee, R., Maratukulam, A., et al. (1997). An STS-Based Radiation Hybrid Map of the Human Genome An STS-Based Radiation Hybrid Map of the Human Genome. *Genome Research* 422–433.
- Stinckens, A., Janssens, S., Spincemaille, G., and Buys, N. (2012). Promising loci for the susceptibility to both scrotal hernia and cryptorchidism in pigs. In ISAG, (Caims, Australia),.
- Straw, B., Bates, R., and May, G. (2009a). prevalence and pig performance. 17, 28–31.
- Straw, B., Bates, R., and May, G. (2009b). Anatomical abnormalities in a group of finishing pigs: prevalence and pig performance. *Journal of Swine Health and Production* 17, 28–31.

Sturtevant, A.H. (1913). The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *Journal of Experimental Zoology* 14, 43–59.

Sutton, E. (1943). Bar Eye in *Drosophila Melanogaster*: A Cytological Analysis of Some Mutations and Reverse Mutations. *Genetics* 28, 97–107.

Sutton, E., Hughes, J., White, S., Sekido, R., Tan, J., Arboleda, V., Rogers, N., Knowler, K., Rowley, L., Eyre, H., et al. (2011). Identification of SOX3 as an XX male sex reversal gene in mice and humans. *The Journal of Clinical Investigation* 121.

Suzuki, K., Asakawa, S., Iida, M., Shimanuki, S., Fujishima, N., Hiraiwa, H., Murakami, Y., Shimizu, N., and Yasue, H. (2000). Construction and evaluation of a porcine bacterial artificial chromosome library. *Animal Genetics* 31, 8–12.

Suzuki, Y., Yeung, A.C., and Ikeno, F. (2011). The representative porcine model for human cardiovascular disease. *Journal of Biomedicine & Biotechnology* 2011, 195483.

Tambasco, A.J. (1990). Cytogenetic study of intersex swine. *Brazilian Journal of Genetics*.

Tanyel, F.C., Ulusu, N.N., Tezcan, E.F., and Büyükpamukçu, N. (2003). Total Calcium Content of Sacs Associated with Inguinal Hernia, Hydrocele or Undescended Testis Reflects Differences Dictated by Programmed Cell Death. *Urologia Internationalis* 70, 211–215.

Teffene, O., Rieu, M., Dagorn, J., Mainsant, P., Marouby, H., and Porin, F. (1998). Trente ans d' évolution du secteur porcin en France : de l' autarcie à la compétitivité internationale. *Journées Recherche Porcine* 133–152.

Tennant, P.W.G., Pearce, M.S., Bythell, M., and Rankin, J. (2010). 20-Year Survival of Children Born With Congenital Anomalies: a Population-Based Study. *Lancet* 375, 649–656.

Teslovich, T.M., Musunuru, K., Smith, A. V., Edmondson, A.C., Stylianou, I.M., Koseki, M., and Pirruccello, J.P. (2011). Biological, Clinical, and Population Relevance of 95 Loci for Blood Lipids. *Nature* 466, 707–713.

Tevosian, S.G., Albrecht, K.H., Crispino, J.D., Fujiwara, Y., Eicher, E.M., and Orkin, S.H. (2002). Gonadal differentiation, sex determination and normal Sry expression in mice require direct interaction between transcription partners GATA4 and FOG2. *Development (Cambridge, England)* 129, 4627–4634.

Thaller, G., Dempfle, L., and Hoeschele, I. (1996). Investigation of the inheritance of birth defects in swine by complex segregation analysis. *Journal of Animal Breeding and Genetics* 113, 77–92.

The Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661–678.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* 14, 178–192.

Tomaselli, S., Megiorni, F., De Bernardo, C., Felici, A., Marrocco, G., Maggiulli, G., Grammatico, B., Remotti, D., Saccucci, P., Valentini, F., et al. (2008). Syndromic true hermaphroditism due to an R-spondin1 (RSPO1) homozygous mutation. *Human Mutation* 29, 220–226.

Tomaselli, S., Megiorni, F., Lin, L., Mazzilli, M.C., Gerrelli, D., Majore, S., Grammatico, P., and Achermann, J.C. (2011). Human RSPO1/R-spondin1 is expressed during early ovary development and augments  $\beta$ -catenin signaling. *PLoS One* 6, e16366.

- Tomizuka, K., Horikoshi, K., Kitada, R., Sugawara, Y., Iba, Y., Kojima, A., Yoshitome, A., Yamawaki, K., Amagai, M., Inoue, A., et al. (2008). R-spondin1 plays an essential role in ovarian development through positively regulating Wnt-4 signaling. *Human Molecular Genetics* *17*, 1278–1291.
- Torres, M., Gómez-Pardo, E., Dressler, G.R., and Gruss, P. (1995). Pax-2 controls multiple steps of urogenital development. *Development (Cambridge, England)* *121*, 4057–4065.
- Tortereau, F., Servin, B., Frantz, L., Megens, H.-J., Milan, D., Rohrer, G., Wiedmann, R., Beever, J., Archibald, A.L., Schook, L.B., et al. (2012). A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content. *BMC Genomics* *13*, 586.
- Turleau, C., and Prieur, M. (2000). Types, fréquences et mécanismes de formation des anomalies chromosomiques.
- Uimari, P., and Tapio, M. (2011). Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. *Journal of Animal Science* *89*, 609–614.
- Vainio, S., Heikkilä, M., Kispert, a, Chin, N., and McMahon, a P. (1999). Female development in mammals is regulated by Wnt-4 signalling. *Nature* *397*, 405–409.
- Val, P., and Swain, A. (2010). Gene dosage effects and transcriptional regulation of early mammalian adrenal cortex development. *Molecular and Cellular Endocrinology* *323*, 105–114.
- VanRaden, P.M., Van Tassell, C.P., Wiggans, G.R., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F., and Schenkel, F.S. (2009). Invited review: reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science* *92*, 16–24.
- Veeramachaneni, D.N.R., Amann, R.P., and Jacobson, J.P. (2005). Testis and Antler Dysgenesis in Sitka Black-Tailed Deer on Kodiak Island, Alaska: Sequela of Environmental Endocrine Disruption? *Environmental Health Perspectives* *114*, 51–59.
- Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C. a, Holt, R. a, et al. (2001). The sequence of the human genome. *Science (New York, N.Y.)* *291*, 1304–1351.
- Vetro, A., Ciccone, R., Giorda, R., Patricelli, M.G., Della Mina, E., Forlino, A., and Zuffardi, O. (2011). XX males SRY negative: a confirmed cause of infertility. *Journal of Medical Genetics* *48*, 710–712.
- Vidal, V.P., Chaboissier, M.C., de Rooij, D.G., and Schedl, a (2001). Sox9 induces testis development in XX transgenic mice. *Nature Genetics* *28*, 216–217.
- Vincent, A.L., Tuggle, C.K., Rothschild, M.F., Evans, G., Short, T.H., Southwood, O.I., and Plastow, G.S. (1998). The Prolactin Receptor Gene is Associated with Increased Litter Size In Pigs. *Swine Research Report*, 1997. Paper 11.
- Vingborg, R.K.K., Gregersen, V.R., Zhan, B., Panitz, F., Høj, A., Sørensen, K.K., Madsen, L.B., Larsen, K., Hornshøj, H., Wang, X., et al. (2009). A robust linkage map of the porcine autosomes based on gene-associated SNPs. *BMC Genomics* *10*, 134.
- Virtanen, H.E., Tapanainen, A.E., Kaleva, M.M., Suomi, A.-M., Main, K.M., Skakkebaek, N.E., and Toppari, J. (2006). Mild gestational diabetes as a risk factor for congenital cryptorchidism. *The Journal of Clinical Endocrinology and Metabolism* *91*, 4862–4865.
- Voorrips, R.E. (2002). MapChart : Software for the Graphical Presentation of Linkage Maps and QTLs.
- Wagner, K.-D. (2003). The complex life of WT1. *Journal of Cell Science* *116*, 1653–1658.

Wagner, T., Wirth, J., Meyer, J., Zabel, B., Held, M., Zimmer, J., Pasantés, J., Bricarelli, F., Keutel, J., Hustert, E., et al. (1994). Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell* 79, 1111–1120.

Walters, J.R. (2010). Have we forgotten about inherited disease ? *UPB Genetic World* 79–86.

Walters, E.M., Wolf, E., Whyte, J.J., Mao, J., Renner, S., Nagashima, H., Kobayashi, E., Zhao, J., Wells, K.D., Critser, J.K., et al. (2012). Completion of the swine genome will simplify the production of swine as a large animal biomedical model. *BMC Medical Genomics* 5, 55.

Wang, J., and Wang, B. (2002). [Study on risk factors of cryptorchidism]. *Zhonghua Liu Xing Bing Xue Za Zhi = Zhonghua Liuxingbingxue Zazhi* 23, 190–193.

Wang, J., Jiang, J., Fu, W., Jiang, L., Ding, X., Liu, J.-F., and Zhang, Q. (2012). A genome-wide detection of copy number variations using SNP genotyping arrays in swine. *BMC Genomics* 13, 273.

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F.A., Hakonarson, H., and Bucan, M. (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research* 17, 1665–1674.

Wang, Z., Liu, X., Yang, B.-Z., and Gelernter, J. (2013). The Role and Challenges of Exome Sequencing in Studies of Human Diseases. *Frontiers in Genetics* 4, 160.

Warr, N., Carre, G.-A., Siggers, P., Vitos Faleato, J., Brixey, R., Pope, M., Bogani, D., Childers, M., Wells, S., Scudamore, C.L., et al. (2012). Gadd45y and Map3k4 interactions regulate mouse testis determination via p38 MAPK-Mediated control of Sry expression. *Developmental Cell* 1020–1031.

Warwick, B.L. (1926). A Study of Hernia in Swine.

Weiss, M.C., and Green, H. (1967). Human-mouse hybrid cell lines containing partial complements of human chromosomes and functioning human genes. *Proceedings of the National Academy of Sciences of the United States of America*.

Weissenbach, J., and Rouyer, F. (1989). Chromosome Y et détermination du sexe.

White, S., Ohnesorg, T., Notini, A., Roeszler, K., Hewitt, J., Daggag, H., Smith, C., Turbitt, E., Gustin, S., van den Bergen, J., et al. (2011). Copy number variation in patients with disorders of sex development due to 46,XY gonadal dysgenesis. *PloS One* 6, e17793.

Whitworth, D.J., Shaw, G., and Renfree, M.B. (1996). Gonadal sex reversal of the developing marsupial ovary in vivo and in vitro. *Development (Cambridge, England)* 122, 4057–4063.

Wiener, A.S. (1932). Method of measuring linkage in human genetics with special reference to blood groups\_Wiener\_1931. *Genetics*.

Wilhelm, D., and Englert, C. (2002). The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sf1. *Gene and Development* 1839–1851.

Wilhelm, D., Palmer, S., and Koopman, P. (2007). Sex Determination and Gonadal Development in Mammals. *Physiological Reviews* 1–28.

Wilkinson, S., Lu, Z.H., Megens, H.-J., Archibald, A.L., Haley, C., Jackson, I.J., Groenen, M. a M., Crooijmans, R.P.M. a, Ogden, R., and Wiener, P. (2013). Signatures of diversifying selection in European pig breeds. *PLoS Genetics* 9, e1003453.

- Wilson, A.S., Power, B.E., and Molloy, P.L. (2007). DNA hypomethylation and human diseases. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 1775, 138–162.
- Womack, J.E., Johnson, J.S., Owens, E.K., Rexroad, C.E., Schläpfer, J., and Yang, Y.P. (1997). A whole-genome radiation hybrid panel for bovine gene mapping. *Mammalian Genome : Official Journal of the International Mammalian Genome Society* 8, 854–856.
- Wong, A.K., Ruhe, A.L., Dumont, B.L., Robertson, K.R., Guerrero, G., Shull, S.M., Ziegler, J.S., Millon, L. V, Broman, K.W., Payseur, B. a, et al. (2010). A comprehensive linkage map of the dog genome. *Genetics* 184, 595–605.
- Worldwatch Institute (2011). Worldwatch Institute.
- Wright, C.F., Wei, Y., Higgins, J.P.T., and Sagoo, G.S. (2012). Non-invasive prenatal diagnostic test accuracy for fetal sex using cell-free DNA a review and meta-analysis. *BMC Research Notes* 5, 476.
- Wyman, A.R., and White, R.A.Y. (1980). A highly polymorphic locus in human DNA. 77, 6754–6758.
- Xiangdong, L., Yuanwu, L., Hua, Z., Liming, R., Qiuyan, L., and Ning, L. (2011). Animal models for the atherosclerosis research: a review. *Protein & Cell* 2, 189–201.
- Xiong, M., and Guo, S. (1997). Fine-Scale Genetic Mapping Based Theory and Applications Linkage Disequilibrium : *American Journal of Human Genetics* 1513–1531.
- Yao, H.H., Whoriskey, W., and Capel, B. (2002). Desert Hedgehog / Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. *Gene and Development* 1433–1440.
- Yao, H.H.C., Matzuk, M.M., Jorgez, C.J., Menke, D.B., Page, D.C., Swain, A., and Capel, B. (2004). Follistatin operates downstream of Wnt4 in mammalian ovary organogenesis. *Developmental Dynamics : an Official Publication of the American Association of Anatomists* 230, 210–215.
- Yeo, G.S.H. (2011). Where next for GWAS? *Briefings in Functional Genomics* 10, 51.
- Yerle, M., Echard, G., Robic, a, Mairal, a, Dubut-Fontana, C., Riquet, J., Pinton, P., Milan, D., Lahbib-Mansais, Y., and Gellin, J. (1996). A somatic cell hybrid panel for pig regional gene mapping characterized by molecular cytogenetics. *Cytogenetics and Cell Genetics* 73, 194–202.
- Yerle, M., Pinton, P., Robic, a, Alfonso, a, Palvadeau, Y., Delcros, C., Hawken, R., Alexander, L., Beattie, C., Schook, L., et al. (1998). Construction of a whole-genome radiation hybrid panel for high-resolution gene mapping in pigs. *Cytogenetics and Cell Genetics* 82, 182–188.
- Yerle, M., Pinton, P., Delcros, C., Arnal, N., Milan, D., and Robic, a (2002). Generation and characterization of a 12,000-rad radiation hybrid panel for fine mapping in pig. *Cytogenetic and Genome Research* 97, 219–228.
- Yoshida, M., Suda, Y., Matsuo, I., Miyamoto, N., Takeda, N., Kuratani, S., and Aizawa, S. (1997). Emx1 and Emx2 functions in development of dorsal telencephalon. *Development (Cambridge, England)* 124, 101–111.
- Young, A.C., Kirkness, E.F., and Breen, M. (2008). Tackling the characterization of canine chromosomal breakpoints with an integrated in-situ/in-silico approach: the canine PAR and PAB. *Chromosome Research* 16, 1193–1202.
- Yuan, F.P., Lin, D.X., Rao, C. V, and Lei, Z.M. (2006). Cryptorchidism in LhrKO animals and the effect of testosterone-replacement therapy. *Human Reproduction (Oxford, England)* 21, 936–942.

Zakaria, M., Azab, S., El Baz, M., Fawaz, L., and Bahagat, A. (2012). Cryptorchidism in Egyptian neonates. *Journal of Pediatric Urology* 5, 95.

Zeggini, E., Scott, L.J., Saxena, R., Voight, B.F., Marchini, L., Hu, T., Bakker, P.I.W. De, Abecasis, G.R., and Almgren, P. (2009). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *40*, 638–645.

Zhang, H., Wang, Z., Wang, S., and Li, H. (2012). Progress of genome wide association study in domestic animals. *Journal of Animal Science and Biotechnology* 3, 26.

Zhao, X. (2009). A candidate gene association study of cryptorchidism and scrotal hernia using canine and porcine models.

Zhao, X., Du, Z.-Q., Vukasinovic, N., Rodriguez, F., Clutter, A.C., and Rothschild, M.F. (2009). Association of HOXA10, ZFPM2, and MMP2 genes with scrotal hernias evaluated via biological candidate gene analyses in pigs. *American Journal of Veterinary Research* 70, 1006–1012.

Zhou, X., Carbonetto, P., and Stephens, M. (2013). Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genetics* 9, e1003264.

(1966). Loi sur l'élevage.

(2012). FAO food and agriculture organization of the United Nations.



# Discussion générale et conclusion

---

Les maladies congénitales, dites de naissance, sont responsables de pertes économiques majeures dans l'industrie porcine. Elles ont un impact économique aux différents stades de la chaîne de production, de la fécondation à l'âge adulte, et portent atteinte au bien-être des animaux. On estime aujourd'hui que les pertes financières dues aux anomalies génétiques s'élèvent à 17,9 millions de dollars par an dans le monde (Walters, 2010). En prenant en compte les pertes dues à la mortalité et aux sacrifices des animaux malades, les coûts pourraient être multipliés par 10.

Au delà des facteurs environnementaux et sanitaires associés à ces maladies, les études ont montré qu'une part de leur déterminisme était génétique. Depuis 2006, l'identification des mécanismes moléculaires sous-jacents est devenue une des priorités de la filière porcine. Dans ce cadre, le projet de recherche SwAn a été établi entre les OSP (association BIOPORC), l'IFIP (Institut Français de l'Industrie Porcine) et l'INRA, l'objectif étant d'identifier les variants génétiques prédisposant les animaux aux anomalies congénitales. Afin de disposer d'un grand nombre d'échantillons, un système de collection, ANOPORC, a été mis en place à partir de 2006 entre les OSP et l'INRA. L'objectif d'ANOPORC était de collecter le maximum d'échantillons d'animaux atteints, quelque soit l'anomalie congénitale présente et la race de l'animal atteint. Ce système a permis de collecter plus de 800 échantillons d'animaux atteints entre 2006 et 2008, correspond à quatorze anomalies (deux à 370 animaux atteints par anomalie). Afin de pouvoir entreprendre rapidement une analyse GWAS pour les anomalies les plus fréquentes, nous avons conjointement décidé de privilégier la collection de familles pour quelques races et trois anomalies, à savoir la hernie inguino-scrotale (en races Large White, et Landrace), la cryptorchidie (en races Piétrain et Large White), et l'intersexualité (en race Large White).

Parmi les autres défauts congénitaux représentés dans la collection, trois anomalies pourraient potentiellement être étudiées dans les années à venir en poursuivant la collecte d'échantillon. Il s'agit de la hernie ombilicale, des tremblements congénitaux et de l'hypoplasie myofibrillaire, encore appelée splay leg.



Pour ces trois anomalies, le nombre d'échantillons d'animaux atteints qui avaient été collectés en 2008 était comparable au nombre d'échantillons pour l'intersexualité (50 animaux environ). Cependant, aucune race pure n'était surreprésentée pour ces trois défauts.

Parmi ces trois défauts, les fréquences rapportées dans certaines études pour la hernie ombilicale sont plus importantes que celles des trois anomalies traitées dans cette thèse. Dans les travaux de Larzul et al. portant sur la prévalence des défauts de naissance majeurs chez trois populations issues d'OSP, 1,26% des animaux étaient atteints de hernie ombilicale, contre 1,09% pour la hernie inguino-scrotale, 0,37% pour la cryptorchidie et 0,36% pour l'intersexualité (Larzul et al., 2008). Enfin, l'héritabilité estimée pour la hernie ombilicale ( $h^2 = 0,37$ ) est supérieure à celle de la hernie inguino-scrotale (0,27) et la cryptorchidie (0,07) (Larzul et al., 2008). Des travaux de génétique moléculaire semblent confirmer ces estimations. Une étude portant sur la recherche de régions impliquées dans le déterminisme de la hernie ombilicale avait été réalisée en 2009 avec 19 ASP sino-européennes (Erhualian x duroc blanc) et 194 marqueurs (microsatellites et SNP) (Ding et al., 2009). Onze régions candidates localisées sur huit chromosomes avaient été identifiées à partir de deux méthodes d'analyse (tests de liaison non paramétriques, et analyses TDT). Mais la difficulté inhérente à ce caractère (et la raison pour laquelle cette anomalie n'a pas été retenue pour notre étude) est la fiabilité du phénotypage. Sur la base d'une simple observation morphologique externe, il est difficile de différencier une hernie ombilicale résultant d'une prédisposition héréditaire, d'un problème sanitaire (infection autour de l'ombilic). Cette difficulté réside également dans le diagnostic des hernies scrotales chez les porcelets, mais pour ce défaut le choix de ne collecter que des animaux non castrés nous a permis de contourner cette difficulté. Afin d'entreprendre une recherche du déterminisme génétique de ce caractère, il serait certainement plus sûr d'échantillonner et de génotyper des reproducteurs dont la valeur phénotypique serait estimée sur descendance (estimation d'une valeur génétique), plutôt que de génotyper des animaux atteints (mesure massale).

Les deux autres anomalies que sont le splay leg et les tremblements congénitaux ont un impact de façon récurrente dans les élevages. Le splay leg a fait l'objet de plusieurs études (Van Der Heyde et al., 1989; Papatsiros, 2012; Partlow et al., 1993) car sa fréquence peut atteindre 8%, et jusqu'à 4 porcelets par portée. Cependant, les facteurs environnementaux semblent avoir un effet fort sur la prévalence de ces deux maladies.



La fréquence d'apparition du splay leg est notamment augmentée chez les porcelets nés en Mai, Aout et Novembre, alors que la prévalence des tremblements congénitaux est plus importante en Automne, ou lors d'infections virales (Mattsson, 2011). L'importance des facteurs environnementaux nous a également été rapportée par les OSP, qui ont observé des variations importantes de la fréquence d'apparition de la maladie, d'un élevage à l'autre ou selon les saisons. Il sera donc essentiel pour ces deux anomalies de mettre en place des études épidémiologiques et d'estimer les paramètres génétiques de ces défauts, avant d'entreprendre une analyse moléculaire. Ces deux anomalies seront néanmoins très intéressantes à étudier à l'avenir et pourraient être des modèles d'études de défauts déterminés par des interactions génotype x milieu. Cependant, ces travaux seront difficiles à réaliser car ils nécessiteront un dispositif expérimental plus important, une caractérisation de l'environnement (température, mesures représentative de l'état sanitaire de l'élevage,...), et le développement de nouveaux modèles statistiques afin d'estimer les interactions entre données génotypiques et mesures « phénotypiques » de l'environnement.

En dehors de ces maladies et des trois anomalies étudiées dans le cadre de cette thèse, les autres défauts répertoriés depuis 2006 ne présentent pas de problème majeur pour la filière. Contrairement à l'espèce bovine, aucune tare grave ne survient subitement et voit sa fréquence rapidement augmentée dans la population. Cette différence de situation est inhérente à la différence de l'organisation de la sélection dans ces deux filières : le pouvoir de diffusion d'un taureau est très largement supérieur à celui d'un verrat. Un verrat est rapidement réformé et remplacé, et si une mutation survient chez un animal, ses descendants ne représenteront donc qu'un petit pourcentage de la population. La fréquence de cet allèle muté restera faible et par dérive pourra être plus facilement éliminé que lorsqu'une mutation est survenue chez un taureau, père de milliers de filles. L'autre différence majeure entre ces deux espèces est l'utilisation du croisement chez le porc pour la production du produit terminal. Les anomalies rapportées comme problématiques pour la filière sont les défauts qui affectent les animaux à l'étage de multiplication (défaut affectant la reproduction des deux sexes) et de production (défauts de conformation, hernies) entraînant un impact économique important. Dans le cas d'une anomalie dont le déterminisme serait monogénique récessif, le simple croisement entraîne la « non-apparition » du défaut (à moins que la même mutation ne soit en ségrégation dans les deux races ou lignées utilisées dans le croisement).



Il n'est donc pas surprenant que les défauts majeurs chez le porc semblent correspondre à des maladies oligo- ou poly-géniques. D'ailleurs, rares sont les exemples de mutations monogéniques caractérisées chez le porc.

Pour l'heure, des études d'association familiales ont été réalisées pour trois premières maladies prises en compte dans le schéma de collection ANOPORC. A partir d'un dispositif familial (pour le moment peu puissant), une analyse GWAS à l'aide de la puce porcine de génotypage haute densité a permis de mettre en évidence un gène (*SOX9*), impliqué dans l'intersexualité, et plusieurs régions qui pourraient, après confirmation, être impliquées dans le déterminisme de la hernie inguino-scrotale et la cryptorchidie. Actuellement, les données disponibles ne permettent pas d'envisager une contre-sélection sur marqueurs de ces défauts. Pour ces trois anomalies, il est essentiel, dans les mois à venir, de compléter ces analyses par l'ajout de familles supplémentaires, voire de combiner ce dispositif à d'autres dispositifs internationaux afin d'augmenter significativement la puissance. A terme, la prise en compte de ces résultats pourra être envisagée dans le cadre d'une sélection assistée par marqueurs (SAM) ou d'une sélection génomique.

L'utilisation de sélection assistée par marqueurs (SAM) permet de sélectionner les animaux selon l'allèle porté à la base mutée causale d'un gène responsable du caractère, ou à l'aide de marqueurs en très fort DL avec le caractère (Goddard and Hayes, 2009). Chez le porc, plusieurs gènes sont déjà intégrés dans des schémas de sélection. Ils concernent des caractères de production, de croissance, de reproduction et de résistance aux maladies. Le premier gène majeur identifié chez le porc a été le gène de la sensibilité à l'halothane, *Hal*, ou *RYR*, responsable de l'hyperthermie maligne (Fujii et al., 1991). Il est aujourd'hui utilisé en sélection contre la susceptibilité au stress et pour la qualité de la viande. Le gène *PRKAG3* ou *RN* a également un effet majeur sur la qualité de la viande (Milan et al., 2000b). D'autres gènes tels que *MC4R* (Kim et al., 2000) et *IGF2* (Van Laere et al., 2003) sont utilisés pour des caractères de composition des carcasses et de croissance des animaux (ainsi que la prise alimentaire pour *MC4R*). Enfin, plusieurs gènes sont associés à la résistance aux maladies, tels que *F18* (Meijerink et al., 2000) et *K88* (Jørgensen et al., 2003), ou aux caractères de reproduction (*ESR* (Rothschild et al., 1996), *PRLR* (Vincent et al., 1998) et *RBP4* (Rothschild et al., 2000)). Pour l'ensemble de ces loci, un test diagnostic est disponible afin de génotyper les individus pour la mutation. Le génotype obtenu est alors pris en compte pour l'estimation de la valeur génétique de l'animal pour les différents caractères d'intérêt.

Dans le cas des anomalies étudiées dans cette thèse, plusieurs régions, à terme, sont susceptibles d'être identifiées comme contribuant à la prédisposition des animaux à développer la maladie. Pour des caractères oligo ou polygéniques, il est possible de sélectionner sur plusieurs marqueurs dans le génome en affectant un poids relatif à chaque région en fonction de leur contribution à l'effet global. Une SAM sur plusieurs régions est déjà été mise en place chez le porc. L'objectif de cette SAM était de sélectionner des verrats croisés chinois, ¼ Meishan, combinant les effets favorables des races chinoises sur les qualités maternelles et la prolificité, et les effets favorables des races européennes sur les caractères de croissances et la composition de la carcasse (Schwob et al., 2009). Des marqueurs localisés dans quatre régions QTL (chromosomes 1, 2, 4 et 7) affectant la croissance et la composition corporelle (Bidanel et al., 2001) ont été utilisés afin de sélectionner dans ces régions les allèles d'origine Large White et Landrace.

Cependant, le nombre réduit de marqueurs utilisés en SAM limite souvent la capacité de prédiction du caractère, car seule une part très faible de la variance génétique du caractère est généralement expliquée (Goddard and Hayes, 2009). Pour remédier à ce problème, des modèles destinés à estimer simultanément les effets d'un très grand nombre de marqueurs répartis sur l'ensemble du génome et ainsi de prendre en compte la majorité des régions impliquées dans le caractère ont été développés. Cette méthode proposée au début des années 2000, porte le nom de sélection génomique (Meuwissen et al., 2001).

La sélection génomique consiste à sélectionner les animaux sur la base de leur valeur génétique, prédite à partir de leur génotype pour des milliers de marqueurs. Dans un premier temps, une population de référence de quelques milliers d'individus est phénotypée et génotypée, afin d'établir les relations statistiques entre génotype et phénotype. Ces valeurs d'effets estimés pour chaque marqueur sont ensuite appliquées aux candidats, ne disposant pas de performances, mais génotypés : chaque individu peut ainsi disposer d'une valeur génétique prédite. Cette méthode a révolutionné les schémas de sélection des bovins laitiers. Au milieu des années 90, l'estimation de la valeur génétique pour les caractères laitiers des candidats à la sélection était réalisée à l'aide d'un testage sur descendance (schéma filles ou schéma petite-filles) (Georges et al., 1995). Bien que très efficace, cette méthode était longue et coûteuse. La mise en place de la sélection génomique a permis de réduire très significativement ces deux paramètres, et donc d'accélérer l'amélioration génétique (Schaeffer, 2006).

Le succès de la sélection génomique chez les bovins, ainsi que le développement des puces de génotypage haute densité ont motivé son application chez les autres espèces domestiques. Cependant, disposer d'une puce de génotypage haute densité n'est pas suffisant pour mettre en place un schéma de sélection génomique. Il est nécessaire de disposer d'une grande population de référence à phénotyper et à génotyper pour obtenir une estimation très précise de l'effet des marqueurs sur le caractère étudié (VanRaden et al., 2009). De plus, l'estimation de la GEBV (*Genomic Breeding Value Estimator*) à partir d'une population de référence nécessite que le DL soit conservé entre la population sélectionnée et la population de référence. Chez les bovins, les animaux étudiés sont majoritairement de race pure, la constitution d'une grande population de référence est ainsi facilitée. Dans la filière porcine en revanche, les porcs charcutiers (produit terminal, résultat de la sélection) sont issus de croisements. Actuellement, différentes études (estimations par simulation et dispositifs d'évaluation) sont en cours chez le porc afin d'estimer l'efficacité de la sélection génomique, et la possibilité de ne constituer qu'une seule population de référence d'animaux croisés ou de différentes races, afin de réaliser une sélection au sein de noyaux de races pures. La combinaison d'animaux de races différentes nécessite de disposer de marqueurs dont la densité est suffisante pour capter le DL inter-races (Amaral et al., 2008). Des travaux récents, dans le cadre des travaux sur la sélection génomique, et portant sur les différences de DL entre quatre lignées américaines (Duroc, Hampshire, Landrace, et Yorkshire), ont montré qu'une précision maximale de 0,92 serait obtenue pour des marqueurs distants de moins de 10 kb, entre les populations Landrace et Yorkshire (Badke et al., 2012). Il sera donc certainement nécessaire de disposer prochainement d'un outil permettant d'obtenir une densité en marqueurs supérieure à celle de la puce porcine actuelle.

Pour les caractères semblables à ceux étudiés dans nos travaux, plusieurs paramètres sont en faveur d'une sélection assistée par marqueur plutôt que de leur prise en compte dans le cadre de la sélection génomique. De plus, la précision de l'estimation est fonction de la taille de la population de référence : pour un caractère très héritable ( $h^2 = 0,8$ ), une population de 5.000 animaux permet d'obtenir une précision de 0,7, alors que pour des caractères peu héritables, plusieurs dizaines de milliers d'animaux sont nécessaires (Goddard and Hayes, 2009). Les héritabilités estimées pour la hernie inguino-scrotale et de la cryptorchidie sont de 0,27 et 0,07, respectivement, il semble donc illusoire d'imaginer que la (ou les) populations de référence mises en place seront suffisamment informatives pour ces caractères.



L'hypothèse la plus probable est qu'à l'issue de ces travaux de GWAS les marqueurs des régions candidates et les haplotypes identifiés associés aux allèles mutés seront pris en compte dans le calcul de la valeur génétique de l'animal, combinant ainsi une sélection génomique pour la majorité des caractères, et une SAM pour les quelques régions candidates de prédisposition.

Enfin, les schémas de sélection utilisés actuellement pour un caractère supposent que l'effet des SNP est additif. Nos travaux ont permis d'identifier des régions potentiellement candidates qui semblent soumises à empreinte parentale, comme les régions sur les chromosomes 2 (origine maternelle), 7 et 13 (origine paternelle) dans le cas de la hernie. Si ces résultats sont confirmés dans l'avenir, cette particularité devra être prise en compte dans le schéma de sélection mis en place.

Au-delà de l'intérêt de ces travaux pour la filière porcine, les résultats obtenus pourraient également avoir des applications en génétique et médecine humaine. Le porc est un excellent modèle dans le domaine médical en raison de ses similitudes anatomiques, physiopathologiques et génétiques avec l'homme (Walters et al., 2012). D'un point de vue anatomique, les organes humains et porcins sont de taille comparable (Lunney, 2007), faisant du porc une espèce d'intérêt majeur pour la xénotransplantation. Mais la présence de rétrovirus dans le génome du porc potentiellement dangereux pour l'homme rend la xénotransplantation d'organes porcins à l'homme problématique. Cependant, la production depuis peu de lignées de porcs miniatures présentant un nombre réduit de rétrovirus ouvre de nouvelles perspectives à ce domaine d'application (Litten-Brown et al., 2010).

Du point de vue physiopathologique, de nombreuses similitudes existent entre les deux espèces dans le développement et la progression de maladies (Lunney, 2007). Le porc représente un modèle de choix pour les maladies infectieuses (Lunney, 2007), l'athérosclérose (Xiangdong et al., 2011), les pathologies gastro-intestinales et osseuses (Litten-Brown et al., 2010), et les maladies cardiovasculaires (Suzuki et al., 2011). De plus, certaines lignées ont été sélectionnées spécifiquement pour servir de modèle à quelques pathologies humaines comme la lignée MeLiM pour l'étude du développement et de la régression spontanée du mélanome cutané (Du et al., 2007b).

Enfin, beaucoup de maladies génétiques présentent des homologies fortes entre l'homme et le porc (Ollivier and Sellier, 1982), et des mutations responsables de pathologies équivalentes ont d'ores et déjà été identifiées dans le même gène chez ces deux espèces.

L'identification de la mutation causale du gène *RYRI*, associé à l'hyperthermie maligne porcine, a permis d'orienter les recherches de variants causaux pour la même maladie chez l'homme (Gillard et al., 1991). Le *Campus Syndrome*, caractérisé par une faiblesse musculaire et des tremblements intenses (Richter et al., 1995), est causé par un défaut du gène *MYH7* chez le porc (Murgiano et al., 2012). Ce gène est également impliqué dans le déterminisme d'une myopathie distale autosomique dominante chez l'homme. Des études récentes ont montré que plus de cent protéines présentes chez le porc (Groenen et al., 2012) étaient identiques à celles retrouvées dans des maladies humaines (obésité, diabète, maladies de Parkinson et Alzheimer).

Dans le cas des anomalies étudiées au cours de cette thèse, des applications pourraient être envisagées au vu des résultats chez le porc. Concernant la hernie scrotale, la compréhension des mécanismes sous-jacents pourrait aider au traitement. La hernie inguinale est une anomalie très fréquente chez l'homme (de 4 à 27% selon les études), et couramment traitée par opération chirurgicale (700.000 opérations de hernies inguino-scrotales aux Etats-Unis en 2008) (Matthews and Neumayer, 2008). Bien qu'elle soit généralement sans gravité, elle peut parfois conduire à des complications sévères, telles que l'étranglement de l'intestin, les dysfonctions gonadiques, les nécroses et perforations intestinales, les occlusions intestinales, l'atrophie ou la nécrose testiculaire (Jenkins and O'Dwyer, 2008). Il est donc important que le traitement soit efficace à long terme. Dans les cas les plus sévères, l'opération consiste à maintenir les organes herniés dans un filet dont la composition peut être adaptée aux mécanismes associés à la maladie. Les métalloprotéinases, enzymes de dégradation du collagène, pourraient être impliquées dans le déterminisme de la hernie (Antoniou et al., 2011; Jain et al., 2009). La confirmation de l'implication de ces mécanismes chez le porc permettrait d'envisager et de tester l'utilisation de filets contenant des inhibiteurs de métalloprotéinases (Pasternak, 2008).

La cryptorchidie, ou ectopie testiculaire, est également une anomalie fréquente chez l'homme, et les conséquences peuvent se révéler sévères et irréversibles. La cryptorchidie représente le défaut le plus fréquent en pédiatrie chez le garçon (Care Effective health program, 2012a), soit 30% des prématurés et 3-4% des nouveau-nés. Dans 70% des cas, la descente testiculaire se déroule après la naissance, entre le quatrième et le sixième mois. Dans 30% des cas, le défaut de migration persiste, augmentant ainsi le risque de complications chez

ces individus. Il est en effet établi que le risque de développer une tumeur testiculaire chez l'enfant cryptorchide est de cinq à dix fois plus important que chez l'enfant sain.

La cryptorchidie peut également être responsable d'infertilité, et un enfant cryptorchide sur deux présente des défauts de spermatogénèse lorsque l'ectopie testiculaire est unilatérale, et huit enfants sur dix dans les cas de cryptorchidie bilatérale (Hadziselimovic et al., 2007). L'identification précoce des nouveau-nés prédisposés à un défaut de migration testiculaire persistant permettrait d'anticiper la mise en place manuelle ou chirurgicale du testicule dans la bourse, afin d'éviter des complications ultérieures.

Concernant l'intersexualité, ou le désordre du développement sexuel, le nombre de cas s'avère plus difficile à estimer car le spectre des « malformations » sexuelles est important. On estime cependant qu'un individu sur 2.000 voire un individu sur 600 est intersexué (Cesalli and Farion, 2009). Le cas le plus fréquent d'intersexualité chez des individus XX est l'hyperplasie congénitale des surrénales. Cette « anomalie » est recherchée à la naissance dans le cadre du test de dépistage néo-natal réalisé systématiquement depuis 2002 en France chez les nouveau-nés de trois à quatre jours. Cinq maladies sont dépistées lors de ce test, et correspondent à des pathologies dont la prise en charge précoce permet l'amélioration des conditions de vie de l'individu atteint (Goizet and Lacombe, 2009). L'hyperplasie congénitale des surrénales, caractérisée par la sécrétion de testostérone chez le fœtus XX (Al-Agha et al., 2012), induit une masculinisation des organes génitaux, avec clitoris hypertrophié (Marumudi et al., 2013). Une prise en charge chirurgicale de « féminisation » est programmée chez les filles atteintes vers l'âge de trois à six mois, voire un an (Pienkowski and Cartlault, 2008). Elle concerne l'ouverture du vagin et la réduction du clitoris et des replis labioscrotaux. Cette opération, réalisée avec l'accord des parents, a pour but de favoriser l'intégration de l'enfant dans la société.

Au-delà des applications médicales, la compréhension des mécanismes moléculaires impliqués dans des pathologies permet une meilleure caractérisation de la maladie, et par voie de conséquence une meilleure acceptation par la société. Ce besoin de connaissance devient criant lorsque la pathologie est perçue péjorativement par la population, et peut entraîner des discriminations pour les individus atteints, comme c'est le cas pour l'intersexualité.

Depuis cinquante ans, les intersexués subissent des opérations chirurgicales de « normalisation », face à la pression sociale d'appartenir clairement à un sexe. Ces opérations sont très controversées par les intersexués. En effet, les individus opérés souffrent fréquemment de troubles de la sexualité au moment de l'adolescence, et il n'est pas rare que ces adolescents souhaitent changer de sexe.



Les associations des intersexués revendiquent le droit de choisir leur sexe et militent contre ces interventions qui sont pratiquées à la naissance. En Suisse, la commission nationale d'éthique dans le domaine de la médecine humaine estime qu'aucune opération d'assignation sexuelle ne devrait être entreprise avant que l'enfant ne puisse le décider (<http://www.tdg.ch>).

Cette notion d'appartenance à l'un ou l'autre des deux sexes est importante dans les sociétés occidentales, mais n'est pas généralisée à toutes les cultures. Par exemple, le « troisième genre » est très bien accepté socialement dans les sociétés bouddhistes, entre autres parce qu'elles ont échappé à la colonisation occidentale. C'est le cas en Thaïlande des *Lady Boy*, jeunes transsexuelles thaïlandaises bien intégrées à la société (Jackson, 2003).

La transsexualité est observée depuis toujours à travers le monde. Bien que souvent perçue comme une simple excentricité, des causes « biologiques » pourraient expliquer la mal-être ressenti par les personnes désirant changer de sexe. En effet, la différenciation sexuelle, sous l'action des hormones sexuelles, ne se déroule pas uniquement au niveau des organes génitaux, elle s'opère également dans le cerveau (Chung and Auger, 2013). La testostérone agit chez les mâles en « masculinisant » le cerveau. Les hormones sexuelles ont par conséquent un rôle de programmation de l'identité et de l'orientation sexuelle (Jašarević et al., 2012). Bien que n'étant activés qu'au moment de la puberté, les comportements sexuels sont donc organisés précocement. Ce mécanisme pourrait expliquer le mal-être perçu chez les enfants XX atteints d'hyperplasie congénitale des surrénales, dont le développement cérébral subit une action hormonale mâle au cours de la vie fœtale.

Au-delà des hormones sexuelles sécrétées par l'individu, il est possible que des facteurs environnementaux à effets endocriniens, comme le bisphénol A, puissent perturber les caractères sexuels à la fois au niveau cérébral chez les individus dont les organes génitaux sont déjà constitués (Jašarević et al., 2012), mais également chez le fœtus (Auger et al., 2013). L'action de ces hormones exogènes pourrait être à l'origine du désir chez certains individus « normaux » anatomiquement de changer d'identité sexuelle. Elle pourrait également expliquer l'augmentation du nombre de cas d'intersexués (Cesalli and Farion, 2009).

Il semble indispensable pour nos sociétés de mieux comprendre et d'accepter le troisième genre, plutôt que de continuer à opérer, peut-être à tort, des individus de plus en plus nombreux.



L'acceptation du troisième genre en Europe est déjà en route : l'Allemagne, sur les modèles australiens, néo-zélandais ou canadiens, a décidé qu'à partir de Novembre 2013, il serait possible de cocher une case "sexe indéterminé" sur le certificat de naissance des nouveau-nés, qui ne disposait jusqu'alors que des propositions "sexe féminin" ou "sexe masculin" (Le Blond, 2013).

Au cours des dernières années, l'évolution des méthodes et des outils en génétique animale a révolutionné la sélection des espèces domestiques. La mise en place et l'utilisation de la sélection génomique chez les premières espèces majeures ont permis de montrer que l'information moléculaire pouvait être efficacement intégrée à des schémas de sélection sans nécessairement connaître les différents gènes et mutations contribuant à la variabilité phénotypique des caractères ciblés. Si pour certains caractères, l'intérêt de la recherche des mutations est actuellement remis en question, beaucoup s'accordent sur l'idée que l'identification des gènes déterminant la prédisposition des animaux à des défauts congénitaux reste essentielle afin de pouvoir envisager une contre-sélection efficace de ces anomalies, la moins pénalisante possible pour les autres caractères pris en compte dans les schémas. Le second intérêt de la recherche des gènes déterminant ces caractères est les retombées possibles de ces résultats en génétique et médecine humaine. L'existence chez les espèces domestiques de maladies homologues à celles présentes chez l'homme peut permettre de mettre en évidence des gènes causaux que les seuls dispositifs familiaux humains n'auraient pas permis d'identifier. Alors que pour la sélection animale des marqueurs en DL sont suffisants, le développement de nouveaux traitements et une meilleure prise en charge des malades nécessitent en médecine humaine l'identification du gène responsable.

Nos travaux sur le déterminisme génétique de la hernie inguino-scrotale, de la cryptorchidie et de l'intersexualité ont permis d'apporter des connaissances nouvelles sur ces maladies. Au cours de cette thèse, un gène (*SOX9*) a d'ores et déjà été identifié comme impliqué dans le déterminisme de l'intersexualité chez le porc. Pour la hernie inguino-scrotale et la cryptorchidie, plusieurs régions ont été détectées et sont à confirmer. Pour la filière porcine, ces trois anomalies ont un impact économique et éthique non négligeable dans les élevages. Nos résultats devront être confirmés et des travaux complémentaires devront être réalisés ; mais à moyen terme ils pourraient permettre de réduire la prévalence de ces défauts dans les élevages. A plus long terme l'identification des gènes impliqués (et des mutations)

sera également essentielle pour mieux comprendre les mécanismes biologiques affectés chez le porc et chez d'autres espèces pour ce type de pathologie.

The End



# Références bibliographiques

---

- Adham, I.M., and AgoulNIK, A.I. (2004). Insulin-like 3 signalling in testicular descent. *International Journal of Andrology* 27, 257–265.
- Affara, N.A., Magenis, R.E., Tolmie, J.L., Boyd, E., Cooke, A., Jamieson, D., Kwok, K., Mitchell, M., and Snadden, L. (1987). Mapping the testis determinants by an analysis of Y-specific sequences in males with apparent. *Nucleic Acids Research* 15, 7325–7342.
- Ai, H., Huang, L., and Ren, J. (2013). Genetic diversity, linkage disequilibrium and selection signatures in chinese and Western pigs revealed by genome-wide SNP markers. *PLoS One* 8, e56001.
- Akhtar, N., and Karim, S. (2012). Varieties of foetal congenital abnormality. *Mymensingh Medical Journal: MMJ* 21, 662–667.
- Al-Agha, A.E., Ocheltree, A.H., and Al-Tamimi, M.D. (2012). Association between genotype, clinical presentation, and severity of congenital adrenal hyperplasia: a review. *The Turkish Journal of Pediatrics* 54, 323–332.
- Alexandrova, E. a, Olovnikov, I. a, Malakhova, G. V, Zabolotneva, A. a, Suntsova, M. V, Dmitriev, S.E., and Buzdin, A. a (2012). Sense transcripts originated from an internal part of the human retrotransposon LINE-1 5' UTR. *Gene* 511, 46–53.
- Altshuler, D., Daly, M.J., and Lander, E.S. (2008). Genetic Mapping in Human Disease. *Science* 322, 881–888.
- Amann, R.P., and Veeramachaneni, D.N.R. (2006). Cryptorchidism and associated problems in animals 1. *Animal Reproduction Science* 7, 108–120.
- Amaral, A.J., Megens, H.-J., Crooijmans, R.P.M. a, Heuven, H.C.M., and Groenen, M. a M. (2008). Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics* 179, 569–579.
- Andersson, L., Haley, C.S., Ellegren, H., Knott, S. a, Johansson, M., Andersson, K., Andersson-Eklund, L., Edfors-Lilja, I., Fredholm, M., and Hansson, I. (1994). Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science (New York, N.Y.)* 263, 1771–1774.
- Andresen, E. (1964). The Inheritance of the Blood Factors I a and I b in Pigs of the Duroc and Hampshire Breeds. *Vox Sanguinis* 9, 617–621.
- Andrieux, J. (2008). [Array-CGH for routine diagnosis of cryptic chromosomal imbalances]. *Pathologie-biologie* 56, 368–374.
- Antoniou, G.A., Tentes, I.K., Antoniou, S.A., Simopoulos, C., and Lazarides, M.K. (2011). Matrix Metalloproteinase Imbalance in Inguinal Hernia Formation. *Journal of Investigative Surgery* 24, 145–150.

- Arango, N.A., Lovell-badge, R., and Behringer, R.R. (1999). Targeted Mutagenesis of the Endogenous Mouse *Mis* Gene Promoter: In Vivo Definition of Genetic Pathways of Vertebrate Sexual Development. *Cell* *99*, 409–419.
- Archibald, A.L., Haley, C.S., Brown, J.F., Couperwhite, S., McQueen, H. a, Nicholson, D., Coppieters, W., Van de Weghe, A., Stratil, A., and Winterø, a K. (1995). The PiGMap consortium linkage map of the pig (*Sus scrofa*). *Mammalian Genome: Official Journal of the International Mammalian Genome Society* *6*, 157–175.
- Arias, J. a, Keehan, M., Fisher, P., Coppieters, W., and Spelman, R. (2009). A high density linkage map of the bovine genome. *BMC Genetics* *10*, 18.
- Armstrong, J.F., Pritchard-Jones, K., Bickmore, W. a, Hastie, N.D., and Bard, J.B. (1993). The expression of the Wilms' tumour gene, *WT1*, in the developing mammalian embryo. *Mechanisms of Development* *40*, 85–97.
- Asakawa, S., Abe, I., Kudoh, Y., Kishi, N., Wang, Y., Kubota, R., Kudoh, J., Kawasaki, K., Minoshima, S., and Shimizu, N. (1997). Human BAC library: construction and rapid screening. *Gene* *191*, 69–79.
- Auger, J., D, L.D., Berges, R., Doridot, L., Salmon, B., and Mc, C. (2013). Environmental levels of oestrogenic and antiandrogenic compounds feminize digit ratios in male rats and their unexposed male progeny. *Proceedings of the Royal Society B: Biological Sciences* *280*, 20131532.
- Badke, Y.M., Bates, R.O., Ernst, C.W., Schwab, C., and Steibel, J.P. (2012). Estimation of linkage disequilibrium in four US pig breeds. *BMC Genomics* *13*, 24.
- Bagheri-Fam, S., Barrionuevo, F., Dohrmann, U., Günther, T., Schüle, R., Kemler, R., Mallo, M., Kanzler, B., and Scherer, G. (2006). Long-range upstream and downstream enhancers control distinct subsets of the complex spatiotemporal *Sox9* expression pattern. *Developmental Biology* *297*, 382–397.
- Bagheri-Fam, S., Sinclair, A.H., Koopman, P., and Harley, V.R. (2010). Conserved regulatory modules in the *Sox9* testis-specific enhancer predict roles for SOX, TCF/LEF, Forkhead, DMRT, and GATA proteins in vertebrate sex determination. *The International Journal of Biochemistry & Cell Biology* *42*, 472–477.
- Bagheri-Fam, S., Sreenivasan, R., Bernard, P., Knowler, K.C., Sekido, R., Lovell-Badge, R., Just, W., and Harley, V.R. (2012). *Sox9* gene regulation and the loss of the XY/XX sex-determining mechanism in the mole vole *Ellobius lutescens*. *Chromosome Research: an International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* *20*, 191–199.
- Bailey, J. a, Carrel, L., Chakravarti, a, and Eichler, E.E. (2000). Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* *97*, 6634–6639.
- Bampton, P.R. (1994). The investigation of genetic and hereditary disorders of pigs. *Pig Journal*.
- Barreiro, L.B., Laval, G., Quach, H., Patin, E., and Quintana-Murci, L. (2008). Natural selection has driven population differentiation in modern humans. *Nature Genetics* *40*, 340–345.

Bashamboo, a, Ledig, S., Wieacker, P., Achermann, J.C., Achermann, J., and McElreavey, K. (2010). New technologies for the identification of novel genetic markers of disorders of sex development (DSD). *Sexual Development: Genetics, Molecular Biology, Evolution, Endocrinology, Embryology, and Pathology of Sex Determination and Differentiation* 4, 213–224.

Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G., and de Massy, B. (2010). PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* 327, 836–840.

Beck, J., Lanoi, W., Taeubert, H., Knorr, C., and Brenig, B. (2006a). Porcine Hernia – functional/positional candidate genes. In 1st European Conference on Pig Genomics, (Lodi), p. 512.

Beck, J., Bornemann-Kolatzki, K., Knorr, C., Taeubert, H., and Brenig, B. (2006b). Molecular characterization and exclusion of porcine GUSB as a candidate gene for congenital hernia inguinalis/scrotalis. *BMC Veterinary Research* 2, 14.

Behl, R., Sheoran, N., Behl, J., Tantia, M.S., and Vijn, R.K. (1999). Microsatellite Sequences of Mammals and Their Applications in Genome Analysis in Pigs - A Review. *Animal Genetics* 001, 1822–1830.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., and Haussler, D. (2004). Ultraconserved elements in the human genome. *Science* 304, 1321–1325.

Benko, S., Fantès, J. a, Amiel, J., Kleinjan, D.-J., Thomas, S., Ramsay, J., Jamshidi, N., Essafi, A., Heaney, S., Gordon, C.T., et al. (2009). Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. *Nature Genetics* 41, 359–364.

Bennett, S.N., Caporaso, N., Fitzpatrick, A.L., Agrawal, A., Boyd, H.A., Cornelis, M.C., Hansel, N.N., Heiss, G., Heit, A., Kang, J.H., et al. (2012). Phenotype harmonization and cross-study collaboration in GWAS consortia: the GENEVA experience. *Genetic Epidemiology* 35, 159–173.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* 27, 573–580.

Bernard, P., Ryan, J., Sim, H., Czech, D.P., Sinclair, A.H., Koopman, P., and Harley, V.R. (2012). Wnt signaling in ovarian development inhibits Sf1 activation of Sox9 via the Tesco enhancer. *Endocrinology* 153, 901–912.

Bersaglieri, T., Sabeti, P.C., Patterson, N., Vanderploeg, T., Schaffner, S.F., Drake, J. a, Rhodes, M., Reich, D.E., and Hirschhorn, J.N. (2004). Genetic signatures of strong recent positive selection at the lactase gene. *American Journal of Human Genetics* 74, 1111–1120.

Berta, P., Hawkins, J.B., Sinclair, A.H., Taylor, A., Griffiths, B.L., Goodfellow, N., and Fellous, M. (1990a). Genetic evidence equating SRY and the testis-determining factor. *450*, 348448.

Berta, P., Hawkins, J., Sinclair, A., Taylor, A., Griffiths, B., Goodfellow, P., and Fellous, M. (1990b). Genetic evidence equating SRY and the testis-determining factor. *Nature* 348, 448–450.

Beuermann, C., Beck, J., Schmelz, U., Dunkelberg, H., Schütz, E., Brenig, B., and Knorr, C. (2009). Tissue calcium content in piglets with inguinal or scrotal hernias or cryptorchidism. *Journal of Comparative Pathology* 140, 182–186.

Beverdam, A., and Koopman, P. (2006). Expression profiling of purified mouse gonadal somatic cells during the critical time window of sex determination reveals novel candidate genes for human sexual dysgenesis syndromes. *Human Molecular Genetics* 15, 417–431.

Biason-Lauber, A. (2010). Control of sex development. *Best Practice & Research. Clinical Endocrinology & Metabolism* 24, 163–186.

Biason-Lauber, A., Konrad, D., Navratil, F., and Schoenle, E.J. (2004). A WNT4 mutation associated with Müllerian-duct regression and virilization in a 46,XX woman. *The New England Journal of Medicine* 351, 792–798.

Biason-Lauber, A., Konrad, D., Meyer, M., DeBeaufort, C., and Schoenle, E.J. (2009). Ovaries and female phenotype in a girl with 46,XY karyotype and mutations in the CBX2 gene. *American Journal of Human Genetics* 84, 658–663.

Bidanel, J.P., Milan, D., Iannuccelli, N., Amigues, Y., Boscher, M.Y., Bourgeois, F., Caritez, J.C., Gruand, J., Le Roy, P., Lagant, H., et al. (2001). Detection of quantitative trait loci for growth and fatness in pigs. *Genetics, Selection, Evolution: GSE* 33, 289–309.

Bidanel, J.-P., Ducrocq, V., and Ollivier, L. (1990). Les nouvelles méthodes d'évaluation des reproducteurs - Intérêt et applications du modèle individuel chez le porc. *Journées Recherche Porcine*.

Birk, O.S., Casiano, D.E., Wassif, C. a, Cogliati, T., Zhao, L., Zhao, Y., Grinberg, a, Huang, S., Kreidberg, J. a, Parker, K.L., et al. (2000). The LIM homeobox gene Lhx9 is essential for mouse gonad formation. *Nature* 403, 909–913.

Bishop, C.E., Whitworth, D.J., Qin, Y., Agoulnik, a I., Agoulnik, I.U., Harrison, W.R., Behringer, R.R., and Overbeek, P. a (2000a). A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. *Nature Genetics* 26, 490–494.

Bishop, C.E., Whitworth, D.J., Qin, Y., Agoulnik, a I., Agoulnik, I.U., Harrison, W.R., Behringer, R.R., and Overbeek, P. a (2000b). A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. *Nature Genetics* 26, 490–494.

Bladon, B. (2002). Surgical management of cryptorchidism in the horse. *In Practice* 24, 126–134.

Le Blond, J. (2013). Germany becomes first European country to recognise "undetermined" sex. *The Guardian*.

Bodmer, W., and Bonilla, C. (2008). Europe PMC Funders Group Common and rare variants in multifactorial susceptibility to common diseases. *40*, 695–701.

Bogatcheva, N. V, Truong, A., Feng, S., Engel, W., Adham, I.M., and Agoulnik, A.I. (2003). GREAT/LGR8 is the only receptor for insulin-like 3 peptide. *Molecular Endocrinology (Baltimore, Md.)* 17, 2639–2646.

Bogatcheva, N. V, Ferlin, A., Feng, S., Truong, A., Giansello, L., Foresta, C., and Agoulnik, A.I. (2007). T222P mutation of the insulin-like 3 hormone receptor LGR8 is

associated with testicular maldescent and hinders receptor expression on the cell surface membrane. *American Journal of Physiology. Endocrinology and Metabolism* 292, E138–44.

Botstein, D., and Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature Genetics* 33 Suppl, 228–237.

BPEX (2010). 2010 PIG COST OF PRODUCTION IN SELECTED INTRODUCTION METHODOLOGY.

Brandão, M.P., Costa, E.M.F., Fukami, M., Gerdulo, M., Natalia P. Pereira, S., Domenice, O., Ogata, T., and Mendonca, B.B. (2011). MAMLD1 (Mastermind-Like Domain Containing 1) Homozygous Gain-of-Function Missense Mutation Causing 46,XX Disorder of Sex Development in a Virilized Female. In *Hormonal and Genetic Basis of Sexual Differentiation Disorders and Hot Topics in Endocrinology: Proceedings of the 2nd World Conference*,.

Breeuwsma, J. (1967). A case of XXY sex chromosome constitution in an intersex pig. *Journal of Reproduction and Fertility* 16, 119–120.

Bullejos, M., and Koopman, P. (2001). Spatially Dynamic Expression of Sry in Mouse. *Developmental Dynamics* 205, 201–205.

Burd, L., Kerbeshian, J., and Klug, M.G. (2001). Neuropsychiatric genetics: misclassification in linkage studies of phenotype-genotype research. *Journal of Child Neurology* 16, 499–504.

Cai, L., Taylor, J.F., Wing, R. a, Gallagher, D.S., Woo, S.S., and Davis, S.K. (1995). Construction and characterization of a bovine bacterial artificial chromosome library. *Genomics* 29, 413–425.

Camerino, G., Parma, P., Radi, O., and Valentini, S. (2006). Sex determination and sex reversal. *Current Opinion in Genetics & Development* 16, 289–292.

Cameron, F.J., and Sinclair, a H. (1997). Mutations in SRY and SOX9: testis-determining genes. *Human Mutation* 9, 388–395.

Campbell, C.D., Ogburn, E.L., Lunetta, K.L., Lyon, H.N., Freedman, M.L., Groop, L.C., Altshuler, D., Ardlie, K.G., and Hirschhorn, J.N. (2005). Demonstrating stratification in a European American population. *Nature Genetics* 37, 868–872.

Canada Pork International (2013). [www.canadapork.com](http://www.canadapork.com).

Canfield, M. a, Honein, M. a, Yuskiv, N., Xing, J., Mai, C.T., Collins, J.S., Devine, O., Petrini, J., Ramadhani, T. a, Hobbs, C. a, et al. (2006). National estimates and race/ethnic-specific variation of selected birth defects in the United States, 1999-2001. *Birth Defects Research. Part A, Clinical and Molecular Teratology* 76, 747–756.

Care Effective health program (2012a). Evaluation and treatment of cryptorchidism.

Care Effective health program (2012b). Evaluation and treatment of cryptorchidism.

Cesalli, F., and Farion, F. (2009). Entre deux sexes.

- Chaboissier, M.-C., Kobayashi, A., Vidal, V.I.P., Lützkendorf, S., van de Kant, H.J.G., Wegner, M., de Rooij, D.G., Behringer, R.R., and Schedl, A. (2004). Functional analysis of Sox8 and Sox9 during sex determination in the mouse. *Development (Cambridge, England)* *131*, 1891–1901.
- Chang, C., So, L., Rs, W., Yeh, S., and Tm, C. (2013). Androgen Receptor ( AR ) Physiological Roles in Male and Female Reproductive Systems : Lessons Learned from AR-Knockout Mice Lacking AR in Selective Cells . *89*, 109132.
- Charlier, C., Coppieters, W., Rollin, F., Desmecht, D., Agerholm, J.S., Cambisano, N., Carta, E., Dardano, S., Dive, M., Fasquelle, C., et al. (2008). Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nature Genetics* *40*, 449–454.
- Chassot, A., Bradford, S., Auguste, A., Gregoire, E., Pailhoux, E., De Rooij, D., Schedl, A., and Chaboissier, M. (2012). WNT4 and RSP01 together are required for cell proliferation in the early mouse gonad . *Development* *139*, 23095882.
- Chassot, A.-A., Ranc, F., Gregoire, E.P., Roepers-Gajadien, H.L., Taketo, M.M., Camerino, G., de Rooij, D.G., Schedl, A., and Chaboissier, M.-C. (2008). Activation of beta-catenin signaling by Rspo1 controls differentiation of the mammalian ovary. *Human Molecular Genetics* *17*, 1264–1277.
- Chen, C., Qiao, R., Wei, R., Guo, Y., Ai, H., Ma, J., Ren, J., and Huang, L. (2012). A comprehensive survey of copy number variation in 18 diverse pig populations and identification of candidate copy number variable genes associated with complex traits. *BMC Genomics* *13*, 733.
- Chen, S.-R., Zheng, Q.-S., Zhang, Y., Gao, F., and Liu, Y.-X. (2013). Disruption of genital ridge development causes aberrant primordial germ cell proliferation but does not affect their directional migration. *BMC Biology* *11*, 22.
- Cheung, C. (2005). Investigating Birth Defects.
- Chevillon, P., Mircovich, C., Fleho, S., and Dubroca, J. (2004). Euthanasie en élevage de porc. *Techniporc* *27*, 21–27.
- Choi, M., Scholl, U.I., Ji, W., Liu, T., Tikhonova, I.R., Zumbo, P., Nayir, A., Bakkaloğlu, A., Ozen, S., Sanjad, S., et al. (2009a). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* *106*, 19096–19101.
- Choi, M., Scholl, U.I., Ji, W., Liu, T., Tikhonova, I.R., Zumbo, P., Nayir, A., Bakkaloğlu, A., Ozen, S., Sanjad, S., et al. (2009b). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* *106*, 19096–19101.
- Chow, A., Purkayastha, S., Athanasiou, T., Tekkis, P., and Darzi, A. (2008). Inguinal hernia Digestive system disorders Inguinal hernia. *Clinical Evidence* 1–63.
- Chow, J., Ciaudo, C., Fazzari, M., Mise, N., Servant, N., Ji, G., Attreed, M., Avner, P., Barillot, E., Jm, G., et al. (2010). LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation . *Cell* *141*, 2010.

- Chow, J.C., Yen, Z., Ziesche, S.M., and Brown, C.J. (2005). Silencing of the mammalian X chromosome. *Annual Review of Genomics and Human Genetics* 6, 69–92.
- Chung, W.C.J., and Auger, A.P. (2013). Gender differences in neurodevelopment and epigenetics. *Pflügers Archiv: European Journal of Physiology* 465, 573–584.
- Ciobanu, D.C., Bastiaansen, J.W.M., Lonergan, S.M., Thomsen, H., Dekkers, J.C.M., Rothschild, M.F., and Plastow, G.S. (2004). New alleles in calpastatin gene are associated with meat quality traits in pigs. *Journal of Animal Science* 2829–2839.
- Cirulli, E.T., and Goldstein, D.B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Reviews. Genetics* 11, 415–425.
- Clancy, B.S., Ph, D., Shaw, K.M., and Education, P.D.N. (2008). DNA Deletion and Duplication and the Associated Genetic Disorders Clustering of Breakpoints: Recombination Hotspots. *Nature Education* 1.
- Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., et al. (2012). Origins and functional impact of copy number variation in the human genome. *464*, 704–712.
- Coop, G., Wen, X., Ober, C., Pritchard, J.K., and Przeworski, M. (2008). High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science* 319, 1395–1398.
- Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 16, 10881–10890.
- Courcelle, E., Beausse, Y., Letort, S., Stahl, O., Fremez, R., Ngom-Bru, C., Gouzy, J., and Faraut, T. (2008). Narcisse: a mirror view of conserved synteny. *Nucleic Acids Research* 36, D485–90.
- Cox, J.E., Edwards, G.B., and Neal, P. a (1979). An analysis of 500 cases of equine cryptorchidism. *Equine Veterinary Journal* 11, 113–116.
- Cox, J.J., Willatt, L., Homfray, T., and Woods, C.G. (2011). A SOX9 duplication and familial 46,XX developmental testicular disorder. *The New England Journal of Medicine* 364, 91–93.
- Damgaard, I.N., Jensen, T.K., Petersen, J.H., Skakkebaek, N.E., Toppari, J., and Main, K.M. (2007). Cryptorchidism and maternal alcohol consumption during pregnancy. *Environmental Health Perspectives* 115, 272–277.
- Daneau, I., Ethier, J.F., Lussier, J.G., and Silversides, D.W. (1996). Porcine SRY gene locus and genital ridge expression. *Biology of Reproduction* 55, 47–53.
- Das, P., Mishra, D., Ghosh, S., Avila, F., Johnson, G., Chowdhary, B., and Raudsepp, T. (2013). Comparative organization and gene expression profiles of the porcine pseudoautosomal region. *Cytogenetic and Genome Research* 2013.
- Das, P.J., Chowdhary, B.P., and Raudsepp, T. (2009). Characterization of the bovine pseudoautosomal region and comparison with sheep, goat, and other mammalian pseudoautosomal regions. *Cytogenetic and Genome Research* 126, 139–147.
- Davidoff, A.M. (2012). Wilms Tumor. *Advances in Pediatrics* 59, 247–267.

- Davis, E., Jensen, C.H., Schroder, H.D., Shay-hadfield, T., Kliem, A., Cockett, N., Georges, M., and Charlier, C. (2004). Ectopic Expression of DLK1 Protein in Skeletal Muscle of Padumnal Heterozygotes Causes the Callipyge Phenotype. *14*, 1858–1862.
- Dekkers, J.C.M. (2004). Commercial application of marker- and gene-assisted selection in livestock : Strategies and lessons. *Journal of Animal Science*.
- Desarda, M.P. (2003). Surgical physiology of inguinal hernia repair--a study of 200 cases. *BMC Surgery* *3*, 2.
- Dickson, S.P., Wang, K., Krantz, I., Hakonarson, H., and Goldstein, D.B. (2010). Rare variants create synthetic genome-wide associations. *PLoS Biology* *8*, e1000294.
- Ding, N.S., Mao, H.R., Guo, Y.M., Ren, J., Xiao, S.J., Wu, G.Z., Shen, H.Q., Wu, L.H., Ruan, G.F., Brenig, B., et al. (2009). A genome-wide scan reveals candidate susceptibility loci for pig hernias in an intercross between White Duroc and Erhualian. *Journal of Animal Science* *87*, 2469–2474.
- Dolgin, E.S., and Charlesworth, B. (2008). The effects of recombination rate on the distribution and abundance of transposable elements. *Genetics* *178*, 2169–2177.
- Du, F.-X., Mathialagan, N., Dyer, C.J., Grosz, M.D., Messer, L.A., Clutter, A.C., Lohuis, M.M., and Byatt, J.C. (2004). Discovery and mapping of a QTL affecting scrotal hernia incidence in chromosome 2 in domestic pigs. In *29th International Conference on Animal Genetics*, (Tokyo), p. 143.
- Du, F.-X., Clutter, A.C., and Lohuis, M.M. (2007a). Characterizing linkage disequilibrium in pig populations. *International Journal of Biological Sciences* *3*, 166–178.
- Du, Z.-Q., Vincent-Naulleau, S., Gilbert, H., Vignoles, F., Créchet, F., Shimogiri, T., Yasue, H., Leplat, J.-J., Bouet, S., Gruand, J., et al. (2007b). Detection of novel quantitative trait loci for cutaneous melanoma by genome-wide scan in the MeLiM swine model. *International Journal of Cancer. Journal International Du Cancer* *120*, 303–320.
- Du, Z.-Q., Zhao, X., Vukasinovic, N., Rodriguez, F., Clutter, A.C., and Rothschild, M.F. (2009). Association and haplotype analyses of positional candidate genes in five genomic regions linked to scrotal hernia in commercial pig lines. *PLoS One* *4*, e4837.
- Du, Z.-Q., Gorbach, D.M., Hu, Z., and Rothschild, M.F. (2010). Segmental Duplications In Pig Genome. In *Plant & Animal Genomes XVIII Conference*,.
- Ducos, a, Revay, T., Kovacs, a, Hidas, a, Pinton, a, Bonnet-Garnier, a, Molteni, L., Slota, E., Switonski, M., Arruga, M. V, et al. (2008). Cytogenetic screening of livestock populations in Europe: an overview. *Cytogenetic and Genome Research* *120*, 26–41.
- Ducos, A., Berland, H., Pinton, A., Séguéla, A., Blanc, M.F., Darré, A., Sans, P., Darré, R., Nationale, É., and Toulouse, V. De (1997). Les translocations réciproques chez le porc : état des lieux et perspectives. *Journées Recherche Porcine* *375–382*.
- Ducos, A., Pinton, A., Berland, H., Séguéla, A., Brun-baronnat, C., Bonnet, N., Darré, R., and Cytogénétique, U.M.R.I. De (2002). Contrôle chromosomique des populations porcines en France : bilan de cinq années d ' activité. *269–275*.
- Duijvesteijn, N., Knol, E.F., Merks, J.W.M., Crooijmans, R.P.M. a, Groenen, M. a M., Bovenhuis, H., and Harlizius, B. (2010). A genome-wide association study on

androstenedione levels in pigs reveals a cluster of candidate genes on chromosome 6. *BMC Genetics* 11, 42.

Dupuis, M.-C., Zhang, Z., Druet, T., Denoix, J.-M., Charlier, C., Lekeux, P., and Georges, M. (2011). Results of a haplotype-based GWAS for recurrent laryngeal neuropathy in the horse. *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 22, 613–620.

Duthie, C.-A., Simm, G., Pérez-Enciso, M., Doeschl-Wilson, A., Kalm, E., Knap, P.W., and Roehe, R. (2009). Genomic scan for quantitative trait loci of chemical and physical body composition and deposition on pig chromosome X including the pseudoautosomal region of males. *Genetics, Selection, Evolution: GSE* 41, 27.

Eggers, S., and Sinclair, A. (2012). Mammalian sex determination—insights from humans and mice. *Chromosome Research: an International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 20, 215–238.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J. a, Kawamoto, K., Buckler, E.S., and Mitchell, S.E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6, e19379.

Emile, J.-F., Leteurtre, E., and Guyétant, S. (2012). *Pathologie générale*.

Erichsen, H.C., and Chanock, S.J. (2004). SNPs in cancer research and treatment. *British Journal of Cancer* 90, 747–751.

Evans, P.D., Gilbert, S.L., Mekel-Bobrov, N., Vallender, E.J., Anderson, J.R., Vaez-Azizi, L.M., Tishkoff, S. a, Hudson, R.R., and Lahn, B.T. (2005). Microcephalin, a gene regulating brain size, continues to evolve adaptively in humans. *Science (New York, N.Y.)* 309, 1717–1720.

Evans, P.M., Chen, X., Zhang, W., and Liu, C. (2010). KLF4 interacts with beta-catenin/TCF4 and blocks p300/CBP recruitment by beta-catenin. *Molecular and Cellular Biology* 30, 372–381.

Fadista, J., Nygaard, M., Holm, L.-E., Thomsen, B., and Bendixen, C. (2008). A snapshot of CNVs in the pig genome. *PLoS One* 3, e3916.

Fan, B., Du, Z., Gorbach, D.M., and Rothschild, M.F. (2010). Development and Application of High-density SNP Arrays in Genomic Studies of Domestic Animals methodologies and technologies have been developed. *The Asian-Aust. J. Anim. Sci.* 23, 833–847.

Fan, B., Onteru, S.K., Du, Z.-Q., Garrick, D.J., Stalder, K.J., and Rothschild, M.F. (2011). Genome-wide association study identifies Loci for body composition and structural soundness traits in pigs. *PLoS One* 6, e14726.

Faraut, T., de Givry, S., Chabrier, P., Derrien, T., Galibert, F., Hitte, C., and Schiex, T. (2007). A comparative genome approach to marker ordering. *Bioinformatics (Oxford, England)* 23, e50–6.

Farce, M.-H. (2000). *Génétique moléculaire*.

Ferguson, L., and Agoulnik, A.I. (2013). Testicular cancer and cryptorchidism. *Frontiers in Endocrinology* 4, 32.

Ferlin, a, Bogatcheva, N. V, Giancesello, L., Pepe, a, Vinanzi, C., Agoulnik, a I., and Foresta, C. (2006). Insulin-like factor 3 gene mutations in testicular dysgenesis syndrome: clinical and functional characterization. *Molecular Human Reproduction* 12, 401–406.

Van Ferneij, J.P., Raynaud, F., and Rieu, M. (2012). Commerce international du porc : les principaux courants et leurs évolutions. *Journées Recherche Porcine* 229–234.

Fiegel, H.C., Rolle, U., Metzger, R., Gfroerer, S., and Kluth, D. (2011). Embryology of the testicular descent. *Seminars in Pediatric Surgery* 20, 170–175.

Fonseca, A.C.S., Bonaldi, A., Bertola, D.R., Kim, C. a, Otto, P. a, and Vianna-Morgante, A.M. (2013). The clinical impact of chromosomal rearrangements with breakpoints upstream of the SOX9 gene: two novel de novo balanced translocations associated with acampomelic campomelic dysplasia. *BMC Medical Genetics* 14, 50.

Food and Agriculture Organization of the United Nations (2008). L'état des ressources zoogénétiques pour l'alimentation et l'agriculture dans le monde.

Food and agriculture organization of the United Nations (2012). <http://faostat.fao.org/>.

Foresta, C., Zuccarello, D., Garolla, A., and Ferlin, A. (2008). Role of hormones, genes, and environment in human cryptorchidism. *Endocrine Reviews* 29, 560–580.

Foster, J., Dominguez-Steglich, M., Guioli, S., Kwok, C., Weller, P., Stevanović, M., Weissenbach, J., Mansour, S., Young, I., Goodfellow, P., et al. (1994). Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature* 8, 525–530.

Fraietta, R., Zylberstejn, D., and Esteves, S. (2013). Hypogonadotropic Hypogonadism Revisited. *Clinics* 68, 81–88.

Frazer, K. a, Murray, S.S., Schork, N.J., and Topol, E.J. (2009). Human genetic variation and its contribution to complex traits. *Nature Reviews. Genetics* 10, 241–251.

Freking, B.A., Murphy, S.K., Wylie, A.A., Rhodes, S.J., Keele, J.W., Leymaster, K.A., Jirtle, R.L., and Smith, T.P.L. (2002). Identification of the Single Base Change Causing the Callipyge Muscle Hypertrophy Phenotype , the Only Known Example of Polar Overdominance in Mammals. *Genome Research* 1496–1506.

Fujii, J., Otsu, K., Zorzato, F., de Leon, S., Khanna, V.K., Weiler, J.E., O'Brien, P.J., and MacLennan, D.H. (1991). Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. *Science (New York, N.Y.)* 253, 448–451.

Gaboreanu, a-M., Grapes, L., Ramos, a M., Kim, J.-J., and Rothschild, M.F. (2004). Characterization of an X-chromosome PCR-RFLP marker associated with fat deposition and growth in the pig. *Animal Genetics* 35, 401–403.

Galani, A., Kitsiou-Tzeli, S., Sofokleous, C., Kanavakis, E., and Kalpini-Mavrou, A. (2008). Androgen insensitivity syndrome: clinical features and molecular defects. *Hormones (Athens, Greece)* 7, 217–229.

Gall, J.G., and Pardue, M. Lou (1969). Formation and detection of rna-dna hybrid molecules in cytological preparations. *Proceedings of the National Academy of Sciences of the United States of America* 378–383.

Gasca, S., Canizares, J., De Santa Barbara, P., Mejean, C., Poulat, F., Berta, P., and Boizet-Bonhoure, B. (2002). A nuclear export signal within the high mobility group domain regulates the nucleocytoplasmic translocation of SOX9 during sexual determination. *Proceedings of the National Academy of Sciences of the United States of America* *99*, 11199–11204.

Georg, I., Bagheri-Fam, S., Knowler, K.C., Wieacker, P., Scherer, G., and Harley, V.R. (2010). Mutations of the SRY-responsive enhancer of SOX9 are uncommon in XY gonadal dysgenesis. *Sexual Development: Genetics, Molecular Biology, Evolution, Endocrinology, Embryology, and Pathology of Sex Determination and Differentiation* *4*, 321–325.

George, J.W., Dille, E. a, and Heckert, L.L. (2011). Current concepts of follicle-stimulating hormone receptor gene regulation. *Biology of Reproduction* *84*, 7–17.

Georges, M., Nielsen, D., Mackinnon, M., Mishra, a, Okimoto, R., Pasquino, a T., Sargeant, L.S., Sorensen, a, Steele, M.R., and Zhao, X. (1995). Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics* *139*, 907–920.

Ghazzal, a M.H. (2006). Inguinal hernias and genital abnormalities in young Jordanian males. *Eastern Mediterranean Health Journal = La Revue de Santé de La Méditerranée Orientale = al-Majallah Al-šihḥiyah Li-sharq Al-mutawassiṭ* *12*, 483–488.

Ghirri, P., Ciulli, C., Vuerich, M., Cuttano, a, Faraoni, M., Guerrini, L., Spinelli, C., Tognetti, S., and Boldrini, a (2002). Incidence at birth and natural history of cryptorchidism: a study of 10,730 consecutive male infants. *Journal of Endocrinological Investigation* *25*, 709–715.

Giannopoulos, M.F., Vlachakis, I.G., and Charissis, G.C. (2001). 13 Years' Experience with the Combined Hormonal Therapy of Cryptorchidism. *Hormone Research* *55*, 33–37.

Gierl, M.S., Gruhn, W.H., von Seggern, A., Maltry, N., and Niehrs, C. (2012). GADD45G functions in male sex determination by promoting p38 signaling and Sry expression. *Developmental Cell* *23*, 1032–1042.

Gilissen, C., Hoischen, A., Brunner, H.G., and Veltman, J. a (2011a). Unlocking Mendelian disease using exome sequencing. *Genome Biology* *12*, 228.

Gilissen, C., Hoischen, A., Brunner, H.G., and Veltman, J. a (2011b). Unlocking Mendelian disease using exome sequencing. *Genome Biology* *12*, 228.

Gillard, E.F., Otsu, K., Fujii, J., Khanna, V.K., de Leon, S., Derdemezi, J., Britt, B. a., Duff, C.L., Worton, R.G., and MacLennan, D.H. (1991). A substitution of cysteine for arginine 614 in the ryanodine receptor is potentially causative of human malignant hyperthermia. *Genomics* *11*, 751–755.

De Givry, S., Bouchez, M., Chabrier, P., Milan, D., and Schiex, T. (2005). CARHTA GENE: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics (Oxford, England)* *21*, 1703–1704.

Goddard, M.E., and Hayes, B.J. (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature Reviews. Genetics* *10*, 381–391.

- Godmann, M., Katz, J.P., Guillou, F., Simoni, M., Klaus, H., Behr, R., National, I., Recherche, D., Centre, A., and Universite, S. (2009). Krüppel-like factor 4 is involved in functional differentiation of testicular Sertoli cells. *Development Biology* 315, 552–566.
- Goizet, C., and Lacombe, D. (2009). Le dépistage néonatal.
- Gordo, I., and Charlesworth, B. (2001). The speed of Muller's ratchet with background selection, and the degeneration of Y chromosomes. *Genetics Research* 78, 2001.
- Gorlov, I.P., Kamat, A., Bogatcheva, N. V, Jones, E., Lamb, D.J., Truong, A., Bishop, C.E., McElreavey, K., and Agoulnik, A.I. (2002). Mutations of the GREAT gene cause cryptorchidism. *Human Molecular Genetics* 11, 2309–2318.
- Griffiths, A.J.F., Gelbart, W.M., Miller, J.H., and Lewontin, R.C. (2001). *Analyse génétique moderne*.
- Grindflek, E., Moe, M., Taubert, H., Simianer, H., Lien, S., and Moen, T. (2006). Genome-wide linkage analysis of inguinal hernia in pigs using affected sib pairs. *BMC Genetics* 7, 25.
- Grindflek, E., Meuwissen, T.H.E., Aasmundstad, T., Hamland, H., Hansen, M.H.S., Nome, T., Kent, M., Torjesen, P., and Lien, S. (2011a). Revealing genetic relationships between compounds affecting boar taint and reproduction in pigs. *Journal of Animal Science* 89, 680–692.
- Grindflek, E., Lien, S., Hamland, H., Hansen, M.H.S., Kent, M., van Son, M., and Meuwissen, T.H.E. (2011b). Large scale genome-wide association and LDLA mapping study identifies QTLs for boar taint and related sex steroids. *BMC Genomics* 12, 362.
- Grocock, C. a, Charlton, H.M., and Pike, M.C. (1988). Role of the fetal pituitary in cryptorchidism induced by exogenous maternal oestrogen during pregnancy in mice. *Journal of Reproduction and Fertility* 83, 295–300.
- Groenen, M. a M., Wahlberg, P., Foglio, M., Cheng, H.H., Megens, H.-J., Crooijmans, R.P.M. a, Besnier, F., Lathrop, M., Muir, W.M., Wong, G.K.-S., et al. (2009). A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Research* 19, 510–519.
- Groenen, M. a M., Archibald, A.L., Uenishi, H., Tuggle, C.K., Takeuchi, Y., Rothschild, M.F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H.-J., et al. (2012). Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491, 393–398.
- Groth, K. a, Skakkebæk, A., Høst, C., Gravholt, C.H., and Bojesen, A. (2013). Clinical review: Klinefelter syndrome--a clinical update. *The Journal of Clinical Endocrinology and Metabolism* 98, 20–30.
- Gyapay, G., Schmitt, K., Fizames, C., Jones, H., Vega-czarny, N., Spillet, D., Muselet, D., Homme, P., Dib, C., Auffray, C., et al. (1996). A radiation hybrid map of the human genome. *Human Molecular Genetics* 5, 339–346.
- Hacker, a, Capel, B., Goodfellow, P., and Lovell-Badge, R. (1995). Expression of Sry, the mouse sex determining gene. *Development (Cambridge, England)* 121, 1603–1614.

Hadziselimovic, F., Hocht, B., Herzog, B., and Buser, M.W. (2007). Infertility in cryptorchidism is linked to the stage of germ cell development at orchidopexy. *Hormone Research* 68, 46–52.

Haldane, J.B.S. (1919). The combination of linkage values, and the calculation of distances between the loci of linked factors. *Journal of Genetics* 8, 299–309.

Haldane, J.B.S. (1922). Sex-ratio and unisexual sterility in hybrid animal. *Journal of Genetics* 101–109.

Haldane, J.B.S., Sprunt, B.A., and Haldane, N.M. (1915). Reduplication in mice.

Haley, C.S., and Lee, G.J. (1993). Genetic basis of prolificacy in Meishan pigs. *Journal of Reproduction and Fertility. Supplement* 48, 247–259.

Hammes, a, Guo, J.K., Lutsch, G., Leheste, J.R., Landrock, D., Ziegler, U., Gubler, M.C., and Schedl, a (2001). Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. *Cell* 106, 319–329.

Han, J.I., and Na, K.J. (2011). Wnt /  $\beta$  Catenin Signaling Pathway in Canine Skin Melanoma and a Possibility as a Cancer Model for Human Skin Melanoma. In *Melanoma in the Clinic - Diagnosis, Management and Complications of Malignancy*,.

Hancock, J.L., and Daker, M.G. (1981). Testicular hypoplasia in a boar with abnormal sex chromosome constitution ( 39 XXY ). *Journal of Reproduction and Fertility* 61, 395–397.

Hanley, N.A., Hagan, D.M., Ball, S.G., Strachan, T., Mcelreavey, K., Lindsay, S., Robson, S., Bullen, P., Ostrer, H., and Wilson, D.I. (2000). SRY , SOX9 , and DAX1 expression patterns during human sex determination and gonadal development. 91, 403–407.

Hanrahan, J.P., Gregan, S.M., Mulsant, P., Mullen, M., Davis, G.H., Powell, R., and Galloway, S.M. (2004). Mutations in the genes for oocyte-derived growth factors GDF9 and BMP15 are associated with both increased ovulation rate and sterility in Cambridge and Belclare sheep (*Ovis aries*). *Biology of Reproduction* 70, 900–909.

Harmegnies, N., Farnir, F., Davin, F., Buys, N., Georges, M., and Coppieters, W. (2006). Measuring the extent of linkage disequilibrium in commercial pig populations. *Animal Genetics* 37, 225–231.

Hawken, R.J., Murtaugh, J., Flickinger, G.H., Yerle, M., Robic, a, Milan, D., Gellin, J., Beattie, C.W., Schook, L.B., and Alexander, L.J. (1999). A first-generation porcine whole-genome radiation hybrid map. *Mammalian Genome : Official Journal of the International Mammalian Genome Society* 10, 824–830.

Hayes, H. m. (1986). Epidemiological features of 5009 cases of equine. *Equine Veterinary Journal* 18, 5009.

Heard, E., Clerc, P., and Avner, P. (1997). X-chromosome inactivation in mammals. *Annual Review of Genetics* 31, 571–610.

Heid, I.M., Jackson, A.U., Randall, J.C., Winkler, T.W., Steinthorsdottir, V., Thorleifsson, G., and Zillikens, M.C. (2011). Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nature Genetics* 42, 949–960.

Helena Mangs, a, and Morris, B.J. (2007). The Human Pseudoautosomal Region (PAR): Origin, Function and Future. *Current Genomics* 8, 129–136.

Henricson, B., and Backstrom, L. (1964). Translocation heterozygosity in a boar.

Van Der Heyde, H., De Mets, J.P., Porreya, L., Henderickx, H., Calus, a., Bekaert, H., and Buysse, F. (1989). Influence of season, litter size, parity, gestation length, birth weight, sex and farrowing pen on frequency of congenital splayleg in piglets. *Livestock Production Science* 27, 143–155.

Hirasawa, R., and Feil, R. (2010). Genomic imprinting and human disease . *Essays in Biochemistry* 48, 20822494.

Hou, Y., Liu, G.E., Bickhart, D.M., Cardone, M.F., Wang, K., Kim, E.-S., Matukumalli, L.K., Ventura, M., Song, J., VanRaden, P.M., et al. (2011). Genomic characteristics of cattle copy number variations. *BMC Genomics* 12, 127.

El Houate, B., Rouba, H., Sibai, H., Barakat, A., Chafik, A., Chadli, E.B., Imken, L., Bogatcheva, N. V, Feng, S., Agoulnik, A.I., et al. (2007). Novel mutations involving the INSL3 gene associated with cryptorchidism. *The Journal of Urology* 177, 1947–1951.

Houmard, B., Small, C., Yang, L., Naluai-Cecchini, T., Cheng, E., Hassold, T., and Griswold, M. (2009). Global gene expression in the human fetal testis and ovary. *Biology of Reproduction* 81, 438–443.

Hu, Z.-L., Park, C. a, Wu, X.-L., and Reecy, J.M. (2013). Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Research* 41, D871–9.

Huang, B., Wang, S., Ning, Y., Lamb, a N., and Bartley, J. (1999). Autosomal XX sex reversal caused by duplication of SOX9. *American Journal of Medical Genetics* 87, 349–353.

Hughes, I. a (2001). Minireview: sex differentiation. *Endocrinology* 142, 3281–3287.

Hughes, I. a, Houk, C., Ahmed, S.F., and Lee, P. a (2006). Consensus statement on management of intersex disorders. *Journal of Pediatric Urology* 2, 148–162.

Hunter, R.H. (1996). Aetiology of intersexuality in female (XX) pigs, with novel molecular interpretations. *Molecular Reproduction and Development* 45, 392–402.

Husmann, D. a, and Levy, J.B. (1995). Current concepts in the pathophysiology of testicular undescend. *Urology* 46, 267–276.

Huyghe, J.R., Jackson, A.U., Fogarty, M.P., Buchkovich, M.L., Stringham, H.M., Sim, X., Yang, L., Cederberg, H., Chines, P.S., Teslovich, T.M., et al. (2013). Exome array analysis identifies novel loci and low-frequency variants for insulin processing and secretion. *Nature Genetics* 45, 197–201.

Irving, D.N. (2006). Chart of all 23 Stages , Detailed Descriptions of Stages 1 - 6 " Part One : Introduction. [Lifeissues.net](http://Lifeissues.net).

Ivell, R. (2003). The molecular basis of cryptorchidism. *Molecular Human Reproduction* 9, 175–181.

- Jääskeläinen, M., Prunskaitė-Hyyryläinen, R., Naillat, F., Parviainen, H., Anttonen, M., Heikinheimo, M., Liakka, A., Ola, R., Vainio, S., Vaskivuo, T.E., et al. (2010). WNT4 is expressed in human fetal and adult ovaries and its signaling contributes to ovarian cell survival. *Molecular and Cellular Endocrinology* 317, 106–111.
- Jackson, P. (2003). *Performative Genders, Perverse Desires: A Bio-History of Thailand's Same-Sex and Transgender Cultures*. Intersections: Gender, History and Culture in the Asian Context.
- Jacquet, P. (2004). Sensitivity of germ cells and embryos to ionizing radiation. *18*, 15471212.
- Jäger, R.J., Anvret, M., Hall, K., and Scherer, G. (1990). A human XY female with a frame shift mutation in the candidate testis-determining gene SRY. *Nature* 348, 452–454.
- Jain, V., Srivastava, R., Jha, S., Misra, S., Rawat, N.S., and Amla, D.V. (2009). Study of matrix metalloproteinase-2 in inguinal hernia. *Journal of Clinical Medicine Research* 1, 285–289.
- Jakob, S. (2010). FOXL2 represses the testis-specific enhancer of Sox9 to maintain ovary differentiation. University College London.
- Jakob, S., and Lovell-Badge, R. (2011). Sex determination and the control of Sox9 expression in mammals. *The FEBS Journal* 278, 1002–1009.
- Jakobsen, L.P., Ullmann, R., Christensen, S.B., Jensen, K.E., Mølsted, K., Henriksen, K.F., Hansen, C., Knudsen, M. a, Larsen, L. a, Tommerup, N., et al. (2007). Pierre Robin sequence may be caused by dysregulation of SOX9 and KCNJ2. *Journal of Medical Genetics* 44, 381–386.
- Jakubiczka, S., Schröder, C., Ullmann, R., Volleth, M., Ledig, S., Gilberg, E., Kroisel, P., and Wieacker, P. (2010). Translocation and deletion around SOX9 in a patient with acampomelic campomelic dysplasia and sex reversal. *Sex Dev* 4, 143–149.
- Jašarević, E., Geary, D.C., and Rosenfeld, C.S. (2012). Sexually selected traits: a fundamental framework for studies on behavioral epigenetics. *ILAR Journal / National Research Council, Institute of Laboratory Animal Resources* 53, 253–269.
- Jenkins, J.T., and O'Dwyer, P.J. (2008). Inguinal hernias. *BMJ (Clinical Research Ed.)* 336, 269–272.
- Jensen, M.S., Wilcox, A.J., Olsen, J., Bonde, J.P., Thulstrup, A.M., Ramlau-Hansen, C.H., and Henriksen, T.B. (2012). Cryptorchidism and hypospadias in a cohort of 934,538 Danish boys: the role of birth weight, gestational age, body dimensions, and fetal growth. *American Journal of Epidemiology* 175, 917–925.
- Jensen-Seaman, M.I., Furey, T.S., Payseur, B.A., Lu, Y., Roskin, K.M., Chen, C.-F., Thomas, M.A., Haussler, D., and Jacob, H.J. (2004). Comparative recombination rates in the rat, mouse, and human genomes. *Genome Research* 14, 528–538.
- Jiang, J., Ma, H., Lu, Y., Wang, Y., Wang, Y., Li, Q., and Ji, S. (2003). Transmission disequilibrium test for congenital dislocation of the hip and HOXB9 gene or COL1A1 gene. *Chinese Journal of Medical Genetics* 20, 193–195.

Jiang, T., Hou, C.-C., She, Z.-Y., and Yang, W.-X. (2012). The SOX gene family: function and regulation in testis determination and male fertility maintenance. *Molecular Biology Reports* 2187–2194.

Jiménez, R., Burgos, M., Sánchez, a, Sinclair, a H., Alarcón, F.J., Marín, J.J., Ortega, E., and Díaz de la Guardia, R. (1993). Fertile females of the mole *Talpa occidentalis* are phenotypic intersexes with ovotestes. *Development (Cambridge, England)* 118, 1303–1311.

Jørgensen, C.B., Cirera, S., Anderson, S.I., Archibald, a L., Raudsepp, T., Chowdhary, B., Edfors-Lilja, I., Andersson, L., and Fredholm, M. (2003). Linkage and comparative mapping of the locus controlling susceptibility towards *E. COLI* F4ab/ac diarrhoea in pigs. *Cytogenetic and Genome Research* 102, 157–162.

Jorgez, C.J., Klysik, M., Jamin, S.P., Behringer, R.R., and Matzuk, M.M. (2004). Granulosa cell-specific inactivation of follistatin causes female fertility defects. *Molecular Endocrinology (Baltimore, Md.)* 18, 953–967.

Jost, A. (1952). Investigation of hormonal control of genesis of the sex organs in rabbit and notes on certain malformations of the genital apparatus in man. *Ginecología y Obstetricia de México* 7, 13010521.

Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., and Kent, W.J. (2004). The UCSC Table Browser data retrieval tool. *Nucleic Acids Research* 32, D493–6.

Kashimada, K., Pelosi, E., Chen, H., Schlessinger, D., Wilhelm, D., and Koopman, P. (2011). FOXL2 and BMP2 act cooperatively to regulate follistatin gene expression during ovarian development. *Endocrinology* 152, 272–280.

Katoh-Fukui, Y., Owaki, A., Toyama, Y., Kusaka, M., Shinohara, Y., Maekawa, M., Toshimori, K., and Morohashi, K.-I. (2005). Mouse Polycomb M33 is required for splenic vascular and adrenal gland formation through regulating Ad4BP/SF1 expression. *Blood* 106, 1612–1620.

Katoh-Fukui, Y., Miyabayashi, K., Komatsu, T., Owaki, A., Baba, T., Shima, Y., Kidokoro, T., Kanai, Y., Schedl, A., Wilhelm, D., et al. (2012). Cbx2, a polycomb group gene, is required for Sry gene expression in mice. *Endocrinology* 153, 913–924.

Kauppi, L., Barchi, M., Baudat, F., Romanienko, P.J., Keeney, S., and Jasin, M. (2011). Distinct properties of the XY pseudoautosomal region crucial for male meiosis. *Science* 331, 916–920.

Kikuchi, K., and Hamaguchi, S. (2013). Novel sex-determining genes in fish and sex chromosome evolution. *Developmental Dynamics: an Official Publication of the American Association of Anatomists* 242, 339–353.

Kim, K.S., Larsen, N., Short, T., Plastow, G., and Rothschild, M.F. (2000). A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 11, 131–135.

Kim, Y., Kobayashi, A., Sekido, R., DiNapoli, L., Brennan, J., Chaboissier, M.-C., Poulat, F., Behringer, R.R., Lovell-Badge, R., and Capel, B. (2006). Fgf9 and Wnt4 act as antagonistic signals to regulate mammalian sex determination. *PLoS Biology* 4, e187.

- Klein, R.J., Zeiss, C., Chew, E.Y., Tsai, J., Sackler, R.S., Haynes, C., Henning, A.K., Sangiovanni, J.P., Mane, S.M., Susan, T., et al. (2005). NIH Public Access. *Science* *308*, 385–389.
- Knorr, C., Bornemann-Kolatzki, K., and Brenig, B. (2006). Genomic mismatch scanning (GMS) and a genome-wide microsatellite scan identify identical chromosome regions associated with porcine hernia inguinalis/scrotalis. In *Pig Genome I - An Overview of Cutting-edge Genomics with Emphasis on the Pig*.
- Knower, K.C., Kelly, S., Ludbrook, L.M., Bagheri-Fam, S., Sim, H., Bernard, P., Sekido, R., Lovell-Badge, R., and Harley, V.R. (2011). Failure of SOX9 regulation in 46XY disorders of sex development with SRY, SOX9 and SF1 mutations. *PLoS One* *6*, e17751.
- Koeleman, B., Al-Ali, A., van der Laan, S.W., and Asselbergs, F.C. (2013). A concise history of genome-wide association studies. *Saudi Journal of Medicine and Medical Sciences* *1*, 4.
- Kondrashov, A.S. (2003). Direct estimates of human per nucleotide mutation rates at 20 loci causing mendelian diseases. *Human Mutation* *21*, 12–27.
- Kong, A., Gudbjartsson, D.F., Sainz, J., Jonsdottir, G.M., Gudjonsson, S. a, Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., et al. (2002). A high-resolution recombination map of the human genome. *Nature Genetics* *31*, 241–247.
- Kong, A., Steinthorsdottir, V., Masson, G., Thorleifsson, G., Sulem, P., Besenbacher, S., Jonasdottir, A., Sigurdsson, A., Th, K., Jonasdottir, A., et al. (2013). Parental origin of sequence variants associated with complex diseases. *462*, 868–874.
- Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1991). Male development of chromosomally female mice transgenic for Sry. *Nature* *351*, 117–121.
- Kopera, I., Durlej, M., Hejmej, a., Knapczyk-Stwora, K., Duda, M., Slomczynska, M., Koziowski, M., and Bilinska, B. (2010). Effects of pre- and postnatal exposure to flutamide on connexin 43 expression in testes and ovaries of prepubertal pigs. *European Journal of Histochemistry* *54*, 67–73.
- Kosambi, D.D. (1944). The estimation of map distances from recombination values. *Annals of Eugenics* *12*, 172–175.
- Kozakowski, J., and Zgliczyński, W. (2013). Body composition, glucose metabolism markers and serum androgens - association in women with polycystic ovary syndrome. *Endokrynologia Polska* *64*, 94–100.
- Kramerov, D. a, and Vassetzky, N.S. (2005). Short retroposons in eukaryotic genomes. *International Review of Cytology* *247*, 165–221.
- Kreidberg, J.A., Sariola, H., Loring, J.M., Maeda, M., Pelletier, J., Housman, D., and Jaenisch, R. (1993). WT-1 is required for early kidney development. *Cell* *74*, 679–691.
- Kropatsch, R., Dekomien, G., Akkad, D. a, Gerding, W.M., Petrasch-Parwez, E., Young, N.D., Altmüller, J., Nürnberg, P., Gasser, R.B., and Epplen, J.T. (2013). SOX9 Duplication Linked to Intersex in Deer. *PLoS One* *8*, e73734.
- Kwok, C., Weller, P. a, Guioli, S., Foster, J.W., Mansour, S., Zuffardi, O., Punnett, H.H., Dominguez-Steglich, M. a, Brook, J.D., and Young, I.D. (1995). Mutations in SOX9, the

gene responsible for Campomelic dysplasia and autosomal sex reversal. *American Journal of Human Genetics* 57, 1028–1036.

Kyrönlahti, A., Euler, R., Bielinska, M., Schoeller, E.L., Moley, K.H., Toppari, J., Heikinheimo, M., and Wilson, D.B. (2012). GATA4 regulates Sertoli cell function and fertility in adult male mice. *Molecular and Cellular Endocrinology* 333, 85–95.

De la Chapelle, a (1987). The Y-chromosomal and autosomal testis-determining genes. *Development (Cambridge, England)* 101 Suppl, 33–38.

Van Laere, A.-S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald, A.L., Haley, C.S., Buys, N., Tally, M., et al. (2003). A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* 425, 832–836.

Laity, J.H., Dyson, H.J., and Wright, P.E. (2000). Molecular basis for modulation of biological function by alternate splicing of the Wilms' tumor suppressor protein. *Proceedings of the National Academy of Sciences of the United States of America* 97, 11932–11935.

Lander, E., and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet.* 11, 241–247.

Larzul, C., Delaunay, I., Schwob, S., Ur, I., Motte, L., and Rheu, L. (2008). Paramètres génétiques des principales anomalies congénitales porcines. 141–142.

Laus, E.J. (1984). inheritance of swine intersexuality (*Sus scrofa domestica*): an experimental study. *Brazilian Journal of Genetics*.

Lavery, R., Lardenois, A., Ranc-Jianmotamedi, F., Pauper, E., Gregoire, E.P., Vigier, C., Moreilhon, C., Primig, M., and Chaboissier, M.-C. (2011). XY Sox9 embryonic loss-of-function mouse mutants show complete sex reversal and produce partially fertile XY oocytes. *Developmental Biology* 354, 111–122.

Lear, T., and McGee, R. (2012). Disorders of sexual development in the domestic horse, *Equus caballus*. *Sex Dev* 6, 61–71.

Lee, J.H., and Jeon, J.T. (2008). Methods to detect and analyze copy number variations at the genome-wide and locus-specific levels. *Cytogenetic and Genome Research* 123, 333–342.

Lee, P. a, Houk, C.P., Ahmed, S.F., and Hughes, I. a (2006). Consensus statement on management of intersex disorders. *International Consensus Conference on Intersex. Pediatrics* 118, e488–500.

Lee, Y., Chao, A., Chen, C., Chou, T., Wang, S.M., and Wang, T. (2011). Analysis of human meiotic recombination events with a parent-sibling tracing approach. *BMC Genomics*.

Lequarré, A., Andersson, L., André, C., Fredholm, M., Hitte, C., Leeb, T., Lohi, H., Lindblad-toh, K., and Georges, M. (2011). LUPA: A European initiative taking advantage of the canine genome architecture for unravelling complex disorders in both human and dogs. *The Veterinary Journal* 189, 155–159.

- Li, Y., Mei, S., Zhang, X., Peng, X., Liu, G., Tao, H., Wu, H., Jiang, S., Xiong, Y., and Li, F. (2012). Identification of genome-wide copy number variations among diverse pig breeds by array CGH. *BMC Genomics* *13*, 725.
- Litten-Brown, J.C., Corson, a M., and Clarke, L. (2010). Porcine models for the metabolic syndrome, digestive and bone disorders: a general overview. *Animal: an International Journal of Animal Bioscience* *4*, 899–920.
- Liu, C.-F., Bingham, N., Parker, K., and Yao, H.H.-C. (2009). Sex-specific roles of beta-catenin in mouse gonadal development. *Human Molecular Genetics* *18*, 405–417.
- Livernois, a M., Graves, J. a M., and Waters, P.D. (2012). The origin and evolution of vertebrate sex chromosomes and dosage compensation. *Heredity* *108*, 50–58.
- Lourenço, D., Brauner, R., Rybczynska, M., Nihoul-Fékété, C., McElreavey, K., and Bashamboo, A. (2011). Loss-of-function mutation in GATA4 causes anomalies of human testicular development. *Proceedings of the National Academy of Sciences of the United States of America* *108*, 1597–1602.
- Lunney, J.K. (2007). Advances in swine biomedical model genomics. *International Journal of Biological Sciences* *3*, 179–184.
- Luo, W., Chen, S., Cheng, D., Wang, L., Li, Y., Ma, X., Song, X., Liu, X., Li, W., Liang, J., et al. (2012). Genome-wide association study of porcine hematological parameters in a Large White × Minzhu F2 resource population. *International Journal of Biological Sciences* *8*, 870–881.
- Luo, X., Ikeda, Y., and Parker, K.L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell* *77*, 481–490.
- Ma, J., Iannuccelli, N., Duan, Y., Huang, W., Guo, B., Riquet, J., Huang, L., and Milan, D. (2010). Recombinational landscape of porcine X chromosome and individual variation in female meiotic recombination associated with haplotypes of Chinese pigs. *BMC Genomics* *11*, 159.
- Ma, J., Gilbert, H., Iannuccelli, N., Duan, Y., Guo, B., Huang, W., Ma, H., Riquet, J., Bidanel, J.-P., Huang, L., et al. (2013). Fine mapping of fatness QTL on porcine chromosome X and analyses of three positional candidate genes. *BMC Genetics* *14*, 46.
- Magenis, R.E., Webb, M.J., McKean, R.S., Tomar, D., Allen, L.J., Kammer, H., Van Dyke, D.L., and Lovrien, E. (1982). Translocation(X;Y)(p22.33;p11.2) in XX males: etiology of male phenotype. *Human Genetics* *62*, 271–276.
- Mäkinen, a, Andersson, M., and Nikunen, S. (1998). Detection of the X chromosomes in a Klinefelter boar using a whole human X chromosome painting probe. *Animal Reproduction Science* *52*, 317–323.
- Mandel, H., Shemer, R., Borochowitz, Z.U., Okopnik, M., Knopf, C., Indelman, M., Drugan, A., Tiosano, D., Gershoni-baruch, R., and Choder, M. (2008). SERKAL Syndrome: An Autosomal-Recessive Disorder Caused by a Loss-of-Function Mutation in WNT4. *The American Journal of Human Genetics* 39–47.
- Manuylov, N.L., Zhou, B., Ma, Q., Fox, S.C., Pu, W.T., and Tevosian, S.G. (2011). Conditional Ablation of Gata4 and Fog2 Genes in Mice Reveals their Distinct Roles in Mammalian Sexual Differentiation. *Development Biology* *353*, 229–241.

- Marian, A.J. (2012). Molecular Genetic Studies of Complex Phenotypes. *Translational Research* 159, 64–79.
- Marklund, L., Winterö, a K., Thomsen, P.D., Johansson, M., Fredholm, M., Gustafsson, U., and Andersson, L. (1993). A linkage group on pig chromosome 4 comprising the loci for blood group L, GBA, ATP1B1 and three microsatellites. *Animal Genetics* 24, 333–338.
- Marklund, S., Kijas, J., Rodriguez-Martinez, H., Rönstrand, L., Funa, K., Moller, M., Lange, D., Edfors-Lilja, I., and Andersson, L. (1998). Molecular basis for the dominant white phenotype in the domestic pig. *Genome Research* 8, 826–833.
- Marumudi, E., Khadgawat, R., Surana, V., Shabir, I., Joseph, A., and Ammini, A.C. (2013). Diagnosis and management of classical congenital adrenal hyperplasia. *Steroids* 78, 741–746.
- Matson, C.K., Murphy, M.W., Sarver, A.L., Griswold, M.D., and Vivian, J. (2012). mammalian testis. *Nature* 476, 101–104.
- Matthews, R.D., and Neumayer, L. (2008). Inguinal hernia in the 21st century: an evidence-based review. *Current Problems in Surgery* 45, 261–312.
- Mattos, M.R.F., Somões-Mattos, L., and Domingues, S.F.S. (2000). Cryptorchidism in dog. *Ciência Animal* 2000 10, 61–70.
- Mattsson, P. (2011). Prevalence of congenital defects in Swedish Hampshire , Landrace and Yorkshire pig breeds and opinions on their prevalence in Swedish commercial herds.
- McCarroll, S. a, Kuruvilla, F.G., Korn, J.M., Cawley, S., Nemesh, J., Wysoker, A., Shapero, M.H., de Bakker, P.I.W., Maller, J.B., Kirby, A., et al. (2008a). Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature Genetics* 40, 1166–1174.
- McCarroll, S. a, Kuruvilla, F.G., Korn, J.M., Cawley, S., Nemesh, J., Wysoker, A., Shapero, M.H., de Bakker, P.I.W., Maller, J.B., Kirby, A., et al. (2008b). Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature Genetics* 40, 1166–1174.
- McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P. a, and Hirschhorn, J.N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews. Genetics* 9, 356–369.
- McClintock, B. (1983). The origin and behavior of matable loci in maize. *Genetics*.
- Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit : A MapReduce framework for analyzing next-generation DNA sequencing data. 1297–1303.
- McLaren, A. (1998). Gonad development: Assembling the mammalian testis. *Current Biology*.
- McLaren, A. (2003). Primordial germ cells in the mouse. *Developmental Biology* 262, 1–15.

McMahon, D.R., Kramer, S. a, and Husmann, D. a (1995). Antiandrogen induced cryptorchidism in the pig is associated with failed gubernacular regression and epididymal malformations. *The Journal of Urology* 154, 553–557.

Mead, T.J., Wang, Q., Bhattaram, P., Dy, P., Afelik, S., Jensen, J., and Lefebvre, V. (2013). A far-upstream (-70 kb) enhancer mediates Sox9 auto-regulation in somatic tissues during development and adult regeneration. *Nucleic Acids Research* gkt140–.

Meijerink, E., Neuenschwander, S., Fries, R., Dinter, a, Bertschinger, H.U., Stranzinger, G., and Vögeli, P. (2000). A DNA polymorphism influencing alpha(1,2)fucosyltransferase activity of the pig FUT1 enzyme determines susceptibility of small intestinal epithelium to *Escherichia coli* F18 adhesion. *Immunogenetics* 52, 129–136.

Mertes, F., Elsharawy, A., Sauer, S., van Helvoort, J.M.L.M., van der Zaag, P.J., Franke, A., Nilsson, M., Lehrach, H., and Brookes, A.J. (2011). Targeted enrichment of genomic DNA regions for next-generation sequencing. *Briefings in Functional Genomics* 10, 374–386.

Merveille, A.-C., Davis, E.E., Becker-Heck, A., Legendre, M., Amirav, I., Bataille, G., Belmont, J., Beydon, N., Billen, F., Clément, A., et al. (2012). CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs. 43, 72–78.

Meuwissen, T.H., Hayes, B.J., and Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829.

Meyer, A.L.M., Berger, E., Monteiro Jr, O., Alonso, P.A., Stavale, J.N., and Gonçalves, M.P.S. (2007). Quantitative and qualitative analysis of collagen types in the fascia transversalis of inguinal hernia patients. *Arq Gastroenterol* 230–234.

Meyers-Wallen, V.N. (2012). Gonadal and sex differentiation abnormalities of dogs and cats. *Sexual Development: Genetics, Molecular Biology, Evolution, Endocrinology, Embryology, and Pathology of Sex Determination and Differentiation* 6, 46–60.

Meyers-Wallen, V.N., Schlafer, D., Barr, I., Lovell-Badge, R., and Keyzner, a (1999). Sry-negative XX sex reversal in purebred dogs. *Molecular Reproduction and Development* 53, 266–273.

Miang-Lon Ng, P., and Lufkin, T. (2012). Embryonic stem cells: protein interaction networks. *Biomol Concepts* 2, 13–25.

Mikami, H., and Fredeen, H.T. (1979). A genetic study of cryptorchidism and scrotal hernia in pigs. *Canadian Journal of Genetics and Cytology. Journal Canadien de Génétique et de Cytologie* 21, 9–19.

Milan, D., Hawken, R., Cabau, C., Leroux, S., Genet, C., Lahbib, Y., Tosser, G., Robic, a, Hately, F., Alexander, L., et al. (2000a). IMpRH server: an RH mapping server available on the Web. *Bioinformatics (Oxford, England)* 16, 558–559.

Milan, D., Jeon, J.-T., Looflt, C., Amarger, V., Robic, A., Thelander, M., Rogel-Gaillard, C., Paul, S., Iannuccelli, N., Rask, L., et al. (2000b). A Mutation in PRKAG3 Associated with Excess Glycogen Content in Pig Skeletal Muscle. *Science* 288, 1248–1251.

Miller, N.A., Lue, S.J. Van, and Rawlings, C.A. (2004). Use of laparoscopic-assisted cryptorchidectomy in dogs and cats. *J Am Vet Med Assoc.* 875–878.

- Mills, R.E., Bennett, E.A., Iskow, R.C., and Devine, S.E. (2007). Which transposable elements are active in the human genome? *Trends in Genetics: TIG* 23, 183–191.
- Minozzi, G., Williams, J.L., Stella, A., Strozzi, F., Luini, M., Settles, M.L., Taylor, J.F., Whitlock, R.H., Zanella, R., and Neibergs, H.L. (2012). Meta-analysis of two genome-wide association studies of bovine paratuberculosis. *PLoS One* 7, e32578.
- Miyamoto, N., Yoshida, M., Kuratani, S., Matsuo, I., and Aizawa, S. (1997). Defects of urogenital development in mice lacking *Emx2*. *Development (Cambridge, England)* 124, 1653–1664.
- Moniot, B., Declosmenil, F., Barrionuevo, F., Scherer, G., Aritake, K., Malki, S., Marzi, L., Cohen-Solal, A., Georg, I., Klattig, J., et al. (2009). The *PGD2* pathway, independently of *FGF9*, amplifies *SOX9* activity in Sertoli cells during male sexual differentiation. *Development (Cambridge, England)* 136, 1813–1821.
- Montazer-Torbati, F., Kocer, A., Auguste, A., Renault, L., Charpigny, G., Pailhoux, E., and Pannetier, M. (2010). A study of goat *SRY* protein expression suggests putative new roles for this gene in the developing testis of a species with long-lasting *SRY* expression. *Developmental Dynamics: an Official Publication of the American Association of Anatomists* 239, 3324–3335.
- Morais da Silva, S., Hacker, A., Harley, V., Goodfellow, P., Swain, A., and Lovell-Badge, R. (1996a). *Sox9* expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. *Nature Genetics*.
- Morais da Silva, S., Hacker, A., Harley, V., Goodfellow, P., Swain, A., and Lovell-Badge, R. (1996b). *Sox9* expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. *Nature Genetics* 14, 62–68.
- Mormède, P., and Mercat, M.-J. (2002). génétiques sur les processus d'adaptation et le comportement alimentaire chez le porc. *INRA Productions Animales* 15, 349–356.
- Muller, H.J. (1964). the Relation of Recombination To Mutational Advance. *Mutation Research* 106, 2–9.
- Muller, C., Denis, M., Gentzbittel, L., and Faraut, T. (2004). The Iccare web server: an attempt to merge sequence and mapping information for plant and animal species. *Nucleic Acids Research* 32, W429–34.
- Mulley, R.C., and Edwards, M.J. (1984). Prevalence of congenital abnormalities in pigs. *Australian Veterinary Journal* 61, 116–120.
- Murgiano, L., Tammen, I., Harlizius, B., and Drögemüller, C. (2012). A de novo germline mutation in *MYH7* causes a progressive dominant myopathy in pigs. *BMC Genetics* 13, 99.
- Myers, S., Freeman, C., Auton, A., Donnelly, P., and G, M. (2008a). A common sequence motif associated with recombination hot spots and genome instability in humans. *Nature Genetics* 40, 2008.
- Myers, S., Freeman, C., Auton, A., Donnelly, P., and G, M. (2008b). A common sequence motif associated with recombination hot spots and genome instability in humans. *Nature Genetics* 40, 2008.

Nagamine, C.M., Morohashi, K., Carlisle, C., and Chang, D.K. (1999). Sex reversal caused by *Mus musculus domesticus* Y chromosomes linked to variant expression of the testis-determining gene *Sry*. *Developmental Biology* 216, 182–194.

Nagaraja, R., MacMillan, S., Kere, J., Jones, C., Griffin, S., Schmatz, M., Terrell, J., Shomaker, M., Jermak, C., Hott, C., et al. (1997). X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. *Genome Research* 7, 210–222.

Natoli, J.L., Ackerman, D.L., McDermott, S., and Edwards, J.G. (2012). Prenatal diagnosis of Down syndrome: a systematic review of termination rates (1995-2011). *Prenatal Diagnosis* 32, 142–153.

Nef, S., and Parada, L.F. (1999). Cryptorchidism in mice mutant for *Insl3*. *Nature Genetics* 22, 295–299.

Nef, S., Verma-Kurvari, S., Merenmies, J., Vassalli, J.-D., Efstratiadis, A., Accili, D., and Parada, L.F. (2003). Testis determination requires insulin receptor family function in mice. *Nature* 426, 291–295.

Newton-cheh, C., Johnson, T., Gateva, V., Tobin, M.D., Bochud, M., Coin, L., Najjar, S.S., Zhao, J.H., Heath, S.C., Eyheramendy, S., et al. (2009). Eight blood pressure loci identified by genome-wide association study of 34,433 people of European ancestry. *Nature Genetics* 41, 666–676.

Ninomiya, S., Isomura, M., Narahara, K., Seino, Y., and Nakamura, Y. (1996). Isolation of a testis-specific cDNA on chromosome 17q from a region adjacent to the breakpoint of t(12;17) observed in a patient with acampomelic campomelic dysplasia and sex reversal. *Human Molecular Genetics* 5, 69–72.

Nowacka, J., Nizanski, W., Klimowicz, M., Dzimira, S., and Switonski, M. (2005). Lack of the *SOX9* gene polymorphism in sex reversal dogs (78,XX; *SRY* negative). *The Journal of Heredity* 96, 797–802.

Nsengimana, J., Baret, P., Haley, C.S., and Visscher, P.M. (2004). Linkage disequilibrium in the domesticated pig. *Genetics* 166, 1395–1404.

Observatoire Bruxellois de la Consommation Durable (2007). Consommation de viande : un lourd tribut environnemental.

Öçal, G. (2011). Current concepts in disorders of sexual development. *Journal of Clinical Research in Pediatric Endocrinology* 3, 105–114.

Ollivier, L. (1979). Les anomalies héréditaires dans l'espèce porcine. *Journées Recherche Porcine* 371–382.

Ollivier, L., and Sellier, P. (1982). Pig genetics : a review.

OMS (2009). International statistical classification of diseases, 10th [tenth] revision ICD-10 (Genève).

OMS (2012). Congenital anomalies.

Ono, M., and Harley, V.R. (2013). Disorders of sex development: new genes, new concepts. *Nature Reviews Endocrinology*.

- Onteru, S.K., Fan, B., Nikkilä, M.T., Garrick, D.J., Stalder, K.J., and Rothschild, M.F. (2011). Whole-genome association analyses for lifetime reproductive traits in the pig. *Journal of Animal Science* *89*, 988–995.
- Orton, F., Rosivatz, E., Scholze, M., and Kortenkamp, A. (2012). Competitive Androgen Receptor Antagonism as a Factor Determining the Predictability of Cumulative Anti androgenic Effects of Widely Used Pesticides. 1578–1584.
- Ott, J. (1999). Analysis of human genetic linkage.
- Ottolenghi, C., Omari, S., Garcia-Ortiz, J.E., Uda, M., Crisponi, L., Forabosco, A., Pilia, G., and Schlessinger, D. (2005). Foxl2 is required for commitment to ovary differentiation. *Human Molecular Genetics* *14*, 2053–2062.
- Pailhoux, E., Pelliniemi, L., Barbosa, A., Parma, P., Kuopio, T., and Cotinot, C. (1997). Relevance of intersexuality to breeding and reproductive biotechnology programs XX sex reversal in pigs. *Theriogenology*.
- Pailhoux, E., Vaiman, D., Barboux, S., Crihiu, E., Cotinot, C., and Fellous, M. (1998). Etude de l'intersexualité chez l'homme et l'animal domestique. Apport en matière de différenciation sexuelle chez les mammifères.
- Pailhoux, E., Parma, P., Sundström, J., Vigier, B., Servel, N., Kuopio, T., Locatelli, a, Pelliniemi, L.J., and Cotinot, C. (2001a). Time course of female-to-male sex reversal in 38,XX fetal and postnatal pigs. *Developmental Dynamics: an Official Publication of the American Association of Anatomists* *222*, 328–340.
- Pailhoux, E., Vigier, B., Vaiman, D., Schibler, L., Vaiman, a, Crihiu, E., Nezer, C., Georges, M., Sundström, J., Pelliniemi, L.J., et al. (2001b). Contribution of domestic animals to the identification of new genes involved in sex determination. *The Journal of Experimental Zoology* *290*, 700–708.
- Pailhoux, E., Vigier, B., Chaffaux, S., Servel, N., Taourit, S., Furet, J.P., Fellous, M., Grosclaude, F., Crihiu, E.P., Cotinot, C., et al. (2001c). A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nature Genetics* *29*, 453–458.
- Pailhoux, E., Mandon-Pepin, B., and Cotinot, C. (2001d). Mammalian gonadal differentiation: the pig model. *Reprod Suppl* *58*, 65–80.
- Pannetier, M., Elzaiat, M., Thépot, D., and Pailhoux, E. (2012). Telling the story of XX sex reversal in the goat: highlighting the sex-crossroad in domestic mammals. *Sex Dev.* *6*, 33–45.
- Papatsiros, V.G. (2012). The Splay Leg Syndrome in Piglets : A Review. *American Journal of Animal and Veterinary Sciences* *7*, 80–83.
- Park, S.Y., and Jameson, J.L. (2005). Minireview: transcriptional regulation of gonadal development and differentiation. *Endocrinology* *146*, 1035–1042.
- Parma, P., Pailhoux, E., and Cotinot, C. (1999). Reverse transcription-polymerase chain reaction analysis of genes involved in gonadal differentiation in pigs. *Biology of Reproduction* *61*, 741–748.

Parma, P., Radi, O., Vidal, V., Chaboissier, M.C., Dellambra, E., Valentini, S., Guerra, L., Schedl, A., and Camerino, G. (2006). R-spondin1 is essential in sex determination, skin differentiation and malignancy. *Nature Genetics* 38, 1304–1309.

Partlow, G.D., Fisher, K.R., Page, P.D., MacMillan, K., and Walker, a F. (1993). Prevalence and types of birth defects in Ontario swine determined by mail survey. *Canadian Journal of Veterinary Research = Revue Canadienne de Recherche Vétérinaire* 57, 67–73.

Pascal, L., Rambourg Schepens, M.-O., and Lasalle, J.-L. (2001). Les malformations congénitales du petit garçon en Languedoc-Roussillon - Étude des cryptorchidies et des hypospadias opérés à partir des données du PMSI de 1998 à 2001.

Pascual, G., Corrales, C., Gómez-Gil, V., Buján, J., and Bellón, J.M. (2007). TGF-beta1 overexpression in the transversalis fascia of patients with direct inguinal hernia. *European Journal of Clinical Investigation* 37, 516–521.

Pasternak, B. (2008). Towards surgical use of matrix metalloproteinase biology.

Pathirana, I.N., Yamasaki, H., Kawate, N., Tsuji, M., Büllesbach, E.E., Takahashi, M., Hatoya, S., Inaba, T., and Tamada, H. (2012). Plasma insulin-like peptide 3 and testosterone concentrations in male dogs: changes with age and effects of cryptorchidism. *Theriogenology* 77, 2011–2012.

Paty, a.-C., Gomes do Espirito Santo, E., and Suzan, F. (2012). Étude des cryptorchidies et hypospadias opérés en France de 1998 à 2008 chez le petit garçon âgé de moins de sept ans à partir des données du Programme de médicalisation des systèmes d'information (PMSI). *Revue d'Épidémiologie et de Santé Publique* 60, S30.

Pearlman, A., Loke, J., Le Caignec, C., White, S., Chin, L., Friedman, A., Warr, N., Willan, J., Brauer, D., Farmer, C., et al. (2010). Mutations in MAP3K1 cause 46,XY disorders of sex development and implicate a common signal transduction pathway in human testis determination. *American Journal of Human Genetics* 87, 898–904.

Pereira, P.M.D.C.C., and Vicente, A.F.D.R.B. (2013). Meat nutritional composition and nutritive role in the human diet. *Meat Science* 93, 586–592.

Perry, G.H., Dominy, N.J., Claw, K.G., Lee, A.S., Fiegler, H., Redon, R., Werner, J., Villanea, F. a, Mountain, J.L., Misra, R., et al. (2007). Diet and the evolution of human amylase gene copy number variation. *Nature Genetics* 39, 1256–1260.

Pienkowskic, C., and Cartlault, A. (2008). Anomalie du développement sexuel du nouveau né.

Pilon, N., Daneau, I., Paradis, V., Hamel, F., Lussier, J.G., Viger, R.S., and Silversides, D.W. (2003). Porcine SRY promoter is a target for steroidogenic factor 1. *Biology of Reproduction* 68, 1098–1106.

Pinton, A., Barasc, H., Raymond Letron, I., Borededebat, M., Mary, N., Massip, K., Bonnet, N., Calgaro, A., Dudez, a M., Feve, K., et al. (2011). Meiotic studies of a 38,XY/39,XXY mosaic boar. *Cytogenetic and Genome Research* 133, 202–208.

Pinton, A., Calgaro, A., Bonnet, N., Mary, N., Dudez, A.M., Barasc, H., Plard, C., Yerle, M., and Ducos, A. (2012). Contrôle chromosomique des populations porcines en France : 43–44.

Pitetti, J.-L., Calvel, P., Romero, Y., Conne, B., Truong, V., Papaioannou, M.D., Schaad, O., Docquier, M., Herrera, P.L., Wilhelm, D., et al. (2013). Insulin and IGF1 receptors are essential for XX and XY gonadal differentiation and adrenal development in mice. *PLoS Genetics* 9, e1003160.

Polanco, J.C., and Koopman, P. (2007). Sry and the hesitant beginnings of male development. *Developmental Biology* 302, 13–24.

Polanco, J.C., Wilhelm, D., Davidson, T.-L., Knight, D., and Koopman, P. (2010). Sox10 gain-of-function causes XX sex reversal in mice: implications for human 22q-linked disorders of sex development. *Human Molecular Genetics* 19, 506–516.

Ponsuksili, S., Murani, E., Brand, B., Schwerin, M., and Wimmers, K. (2011). Integrating expression profiling and whole-genome association for dissection of fat traits in a porcine model. *Journal of Lipid Research* 52, 668–678.

Pritchard, J.K. (2001). Are rare variants responsible for susceptibility to complex diseases? *American Journal of Human Genetics* 69, 124–137.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. a R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81, 559–575.

Quigley, C., De Bellis, A., Marschke, K., El-Awady, M., Wilson, E., and French, F. (1995). Androgen receptor defects: historical, clinical, and molecular. *Endocr Rev.* 16, 7671849.

Quilter, C.R., Blott, S.C., Mileham, A.J., Nabeel, A.A., Sargent, C.A., and Gri, D.K. (2002). A mapping and evolutionary study of porcine sex chromosome genes. *Mammalian Genome* 594, 588–594.

Rainelli, P. (2001). L'image de la viande de Porc en France: attitudes des consommateurs.

Ramayo-Caldas, Y., Castelló, A., Pena, R.N., Alves, E., Mercadé, A., Souza, C. a, Fernández, A.I., Perez-Enciso, M., and Folch, J.M. (2010). Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. *BMC Genomics* 11, 593.

Ramos, A.M., Crooijmans, R.P.M. a, Affara, N. a, Amaral, A.J., Archibald, A.L., Beever, J.E., Bendixen, C., Churcher, C., Clark, R., Dehais, P., et al. (2009). Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS One* 4, e6524.

Ramos, A.M., Duijvesteijn, N., Knol, E.F., Merks, J.W.M., Bovenhuis, H., Crooijmans, R.P.M. a, Groenen, M. a M., and Harlizius, B. (2011). The distal end of porcine chromosome 6p is involved in the regulation of skatole levels in boars. *BMC Genetics* 12, 35.

Raudsepp, T., and Chowdhary, B.P. (2008). The horse pseudoautosomal region (PAR): characterization and comparison with the human, chimp and mouse PARs. *Cytogenetic and Genome Research* 127, 102–109.

Read, A., and Donnai, D. (2008). Génétique médicale ; de la biologie à la pratique clinique.

- Reich, D.E., and Lander, E.S. (2001). On the allelic spectrum of human disease. *Trends in Genetics* 17, 502–510.
- Reiner, G., Eckert, J., Peischl, T., Bochert, S., Jäkel, T., Mackenstedt, U., Joachim, a., Dauschies, a., and Geldermann, H. (2002). Variation in clinical and parasitological traits in Pietrain and Meishan pigs infected with *Sarcocystis miescheriana*. *Veterinary Parasitology* 106, 99–113.
- Ren, J., Mao, H., Zhang, Z., Xiao, S., Ding, N., and Huang, L. (2011). A 6-bp deletion in the TYRP1 gene causes the brown colouration phenotype in Chinese indigenous pigs. *Heredity* 106, 862–868.
- Rettenberger, G., Bruch, J., Fries, R., Archibald, a L., and Hameister, H. (1996). Assignment of 19 porcine type I loci by somatic cell hybrid analysis detects new regions of conserved synteny between human and pig. *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 7, 275–279.
- Richter, A., Wissel, J., Harlizius, B., Simon, D., Schelosky, L., Scholz, U., Poewe, W., and Löscher, W. (1995). The “Campus Syndrome” in Pigs: Neurological, Neurophysiological, and Neuropharmacological Characterization of a New Genetic Animal Model of High-Frequency Tremor. *Experimental Neurology*.
- Risch, N., and Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science* 273, 1516–1517.
- Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2012). Integrative genomics viewer. *Nature Biotechnology* 29, 24–26.
- Rodrigues Junior, A.J., Rodrigues, C.J., da Cunha, A.C.P., and Jin, Y. (2002). Quantitative analysis of collagen and elastic fibers in the transversalis fascia in direct and indirect inguinal hernia. *Revista Do Hospital Das Clínicas* 57, 265–270.
- Rohrer, G. a, Alexander, L.J., Hu, Z., Smith, T.P., Keele, J.W., and Beattie, C.W. (1996). A comprehensive map of the porcine genome. *Genome Research* 6, 371–391.
- Rosch, R., Junge, K., Lynen, P., Mertens, P.R., Klinge, U., and Schumpelick, V. (2003). Hernia - A Collagen Disease? *European Surgery-Acta Chirurgica Austriaca* 35, 11–15.
- Ross, D.G.F., Bowles, J., Koopman, P., and Lehnert, S. (2008). New insights into SRY regulation through identification of 5' conserved sequences. *BMC Molecular Biology* 9, 85.
- Rothschild, M.F. (2004). Porcine genomics delivers new tools and results: this little piggy did more than just go to market. *Genetical Research* 83, 1–6.
- Rothschild, M.F., and Ruvinsky, A. (2011). *The Genetics of the Pig*, 2nd Edition.
- Rothschild, M., Jacobson, C., Vaske, D., Tuggle, C., Wang, L., Short, T., Eckardt, G., Sasaki, S., Vincent, a, McLaren, D., et al. (1996). The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proceedings of the National Academy of Sciences of the United States of America* 93, 201–205.
- Rothschild, M.F., Christian, L.L., and Blanchard, W. (1988). Evidence for multigene control of cryptorchidism in swine. *The Journal of Heredity* 79, 313–314.

Rothschild, M.F., Messer, L., Day, a, Wales, R., Short, T., Southwood, O., and Plastow, G. (2000). Investigation of the retinol-binding protein 4 (RBP4) gene as a candidate gene for increased litter size in pigs. *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 11, 75–77.

Rouault, K., Scotet, V., Autret, S., Gaucher, F., Dubrana, F., Tanguy, D., Rassi, C.Y. El, Fenoll, B., and Férec, C. (2010). Evidence of association between GDF5 polymorphisms and congenital dislocation of the hip in a Caucasian population Objective. *Osteoarthritis and Cartilage* 18, 2010.

Roy, M. (2011). La problématique des hernies dans les élevages porcins.

Rudkin, G.T., and Stollar, B.D. (1977). High resolution detection of DNA-RNA hybrids in situ by indirect immunofluorescence. *Nature* 265, 5593.

Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z.P., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J. V, Patterson, N.J., Mcdonald, G.J., et al. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419.

Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., et al. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409, 928–933.

De Santa Barbara, P., Méjean, C., Moniot, B., Malclès, M.H., Berta, P., and Boizet-Bonhoure, B. (2001). Steroidogenic factor-1 contributes to the cyclic-adenosine monophosphate down-regulation of human SRY gene expression. *Biology of Reproduction* 64, 775–783.

Schaeffer, L.R. (2006). Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics = Zeitschrift Für Tierzüchtung Und Züchtungsbiologie* 123, 218–223.

Schibler, L., Vaiman, D., Oustry, a, Guinec, N., Dangy-Caye, a L., Billault, a, and Cribiu, E.P. (1998). Construction and extensive characterization of a goat bacterial artificial chromosome library with threefold genome coverage. *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 9, 119–124.

Schlotz, W., and Phillips, D.I.W. (2009). Fetal origins of mental health: evidence and mechanisms. *Brain, Behavior, and Immunity* 23, 905–916.

Schneider, J. a, Peto, T.E. a, Boone, R. a, Boyce, A.J., and Clegg, J.B. (2002). Direct measurement of the male recombination fraction in the human beta-globin hot spot. *Human Molecular Genetics* 11, 207–215.

Van der Schoot, P. (1992). Disturbed testicular descent in the rat after prenatal exposure to the antiandrogen flutamide. *Journal of Reproduction and Fertility* 96, 483–496.

Schull, W.J., and Otake, M. (1999). Cognitive function and prenatal exposure to ionizing radiation. *Teratology* 59, 222–226.

Schwob, S., Riquet, J., Bellec, T., Kernaleguen, L., Tribout, T., and Bidanel, J.-P. (2009). Mise en place d ' un programme de sélection assistée par marqueurs dans la population sino-européenne Duochan. *Journées Recherche Porcine* 1, 1–2.

Seeherunvong, T., Perera, E.M., Bao, Y., Benke, P.J., Benigno, A., Donahue, R.P., and Berkovitz, G.D. (2004). 46,XX sex reversal with partial duplication of chromosome arm 22q. *American Journal of Medical Genetics. Part A* 127A, 149–151.

Sekido, R., and Lovell-Badge, R. (2008). Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature* 453, 930–934.

Sekido, R., and Lovell-Badge, R. (2009). Sex determination and SRY: down to a wink and a nudge? *Trends in Genetics: TIG* 25, 19–29.

Sela, N., Kim, E., and Ast, G. (2010). The role of transposable elements in the evolution of non-mammalian vertebrates and invertebrates. *Genome Biology* 11, R59.

Servin, B., de Givry, S., and Faraut, T. (2010a). Statistical confidence measures for genome maps: application to the validation of genome assemblies. *Bioinformatics (Oxford, England)* 26, 3035–3042.

Servin, B., de Givry, S., and Faraut, T. (2010b). Statistical confidence measures for genome maps: application to the validation of genome assemblies. *Bioinformatics (Oxford, England)* 26, 3035–3042.

Servin, B., Faraut, T., Iannuccelli, N., Zelenika, D., and Milan, D. (2012). High-resolution autosomal radiation hybrid maps of the pig genome and their contribution to the genome sequence assembly. *BMC Genomics* 13, 585.

She, X., Cheng, Z., Zöllner, S., Church, D.M., and Eichler, E.E. (2008). Mouse segmental duplication and copy number variation. *Nature Genetics* 40, 909–914.

Shearin, A.L., and Ostrander, E. a (2010). Leading the way: canine models of genomics and disease. *Disease Models & Mechanisms* 3, 27–34.

Sherwood, L. (2006). *Physiologie humaine: A Human Perspective*.

Shifman, S., Bell, J.T., Copley, R.R., Taylor, M.S., Williams, R.W., Mott, R., and Flint, J. (2006). A high-resolution single nucleotide polymorphism genetic map of the mouse genome. *PLoS Biology* 4, e395.

Silva, K., Bastiaansen, J., Knol, E., Merks, J., Lopes, P., Guimarães, S., and van Arendonk, J. (2011). Meta-analysis of results from quantitative trait loci mapping studies on pig chromosome 4. *Animal Genetics* 42, 2145.

Sinclair, a H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, a M., Lovell-Badge, R., and Goodfellow, P.N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* 346, 240–244.

Sironen, a, Uimari, P., Iso-Touru, T., and Vilkki, J. (2012). L1 insertion within SPEF2 gene is associated with increased litter size in the Finnish Yorkshire population. *Journal of Animal Breeding and Genetics = Zeitschrift Für Tierzüchtung Und Züchtungsbiologie* 129, 92–97.

Sironen, A., Thomsen, B., Andersson, M., Ahola, V., and Vilkki, J. (2006). An intronic insertion in KPL2 results in aberrant splicing and causes the immotile short-tail sperm defect in the pig. *Proceedings of the National Academy of Sciences of the United States of America* 103, 5006–5011.

- Sironen, A., Vilkki, J., Bendixen, C., and Thomsen, B. (2007). Infertile Finnish Yorkshire boars carry a full-length LINE-1 retrotransposon within the KPL2 gene. *Molecular Genetics and Genomics: MGG* 278, 385–391.
- Sironen, A., Uimari, P., Nagy, S., Paku, S., Andersson, M., and Vilkki, J. (2010a). Knobbed acrosome defect is associated with a region containing the genes STK17b and HECW2 on porcine chromosome 15. *BMC Genomics* 11, 699.
- Sironen, A., Uimari, P., Nagy, S., Paku, S., Andersson, M., and Vilkki, J. (2010b). Knobbed acrosome defect is associated with a region containing the genes STK17b and HECW2 on porcine chromosome 15. *BMC Genomics* 11, 699.
- Skinner, B.M., Lachani, K., Sargent, C. a, and Affara, N. a (2013). Regions of XY homology in the pig X chromosome and the boundary of the pseudoautosomal region. *BMC Genetics* 14, 3.
- Smit, A., Hubley, R., and Green, P. (2010). RepeatMasker Open-3.2.8.
- Smith, J., Sones, K., Grace, D., MacMillan, S., Tarawali, S., and Herrero, M. (2012). Beyond milk, meat, and eggs: Role of livestock in food and nutrition security. *Animal Frontiers* 3, 6–13.
- Sotoodehnia, N., Isaacs, A., Bakker, P.I.W. de, Dörr, M., Newton-Cheh, C., Nolte, I.M., Harst, P. van der, Müller, M., Eijgelsheim, M., and Alonso, A. (2010). Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. *Nature Genetics* 42, 1068–1076.
- Speedy, A.W. (2003). Global production and consumption of animal source foods. *The Journal of Nutrition* 4048–4053.
- Speek, M. (2001). Antisense Promoter of Human L1 Retrotransposon Drives Transcription of Adjacent Cellular Genes Antisense Promoter of Human L1 Retrotransposon Drives Transcription of Adjacent Cellular Genes. *Molecular and Cellular Biology* 21.
- Stankiewicz, P., and Lupski, J.R. (2010). Structural variation in the human genome and its role in disease. *Annual Review of Medicine* 61, 437–455.
- Stemers, F.J., and Gunderson, K.L. (2007). Whole genome genotyping technologies on the BeadArray platform. *Biotechnology Journal* 2, 41–49.
- Stewart, E.A., Mckusick, K.B., Aggarwal, A., Brady, S., Chu, A., Fang, N., Hadley, D., Hussain, S., Lee, R., Maratukulam, A., et al. (1997). An STS-Based Radiation Hybrid Map of the Human Genome An STS-Based Radiation Hybrid Map of the Human Genome. *Genome Research* 422–433.
- Stinckens, A., Janssens, S., Spincemaille, G., and Buys, N. (2012). Promising loci for the susceptibility to both scrotal hernia and cryptorchidism in pigs. In ISAG, (Caims, Australia),.
- Straw, B., Bates, R., and May, G. (2009a). prevalence and pig performance. 17, 28–31.
- Straw, B., Bates, R., and May, G. (2009b). Anatomical abnormalities in a group of finishing pigs: prevalence and pig performance. *Journal of Swine Health and Production* 17, 28–31.

Sturtevant, A.H. (1913). The linear arrangement of six sex-linked factors in drosophila, as shown by their mode of association. *Journal of Experimental Zoology* 14, 43–59.

Sutton, E. (1943). Bar Eye in *Drosophila Melanogaster*: A Cytological Analysis of Some Mutations and Reverse Mutations. *Genetics* 28, 97–107.

Sutton, E., Hughes, J., White, S., Sekido, R., Tan, J., Arboleda, V., Rogers, N., Knowler, K., Rowley, L., Eyre, H., et al. (2011). Identification of SOX3 as an XX male sex reversal gene in mice and humans. *The Journal of Clinical Investigation* 121.

Suzuki, K., Asakawa, S., Iida, M., Shimanuki, S., Fujishima, N., Hiraiwa, H., Murakami, Y., Shimizu, N., and Yasue, H. (2000). Construction and evaluation of a porcine bacterial artificial chromosome library. *Animal Genetics* 31, 8–12.

Suzuki, Y., Yeung, A.C., and Ikeno, F. (2011). The representative porcine model for human cardiovascular disease. *Journal of Biomedicine & Biotechnology* 2011, 195483.

Tambasco, A.J. (1990). Cytogenetic study of intersex swine. *Brazilian Journal of Genetics*.

Tanyel, F.C., Ulusu, N.N., Tezcan, E.F., and Büyükpamukçu, N. (2003). Total Calcium Content of Sacs Associated with Inguinal Hernia, Hydrocele or Undescended Testis Reflects Differences Dictated by Programmed Cell Death. *Urologia Internationalis* 70, 211–215.

Teffene, O., Rieu, M., Dagorn, J., Mainsant, P., Marouby, H., and Porin, F. (1998). Trente ans d' évolution du secteur porcin en France: de l' autarcie à la compétitivité internationale. *Journées Recherche Porcine* 133–152.

Tennant, P.W.G., Pearce, M.S., Bythell, M., and Rankin, J. (2010). 20-Year Survival of Children Born With Congenital Anomalies: a Population-Based Study. *Lancet* 375, 649–656.

Teslovich, T.M., Musunuru, K., Smith, A. V., Edmondson, A.C., Stylianou, I.M., Koseki, M., and Pirruccello, J.P. (2011). Biological, Clinical, and Population Relevance of 95 Loci for Blood Lipids. *Nature* 466, 707–713.

Tevosian, S.G., Albrecht, K.H., Crispino, J.D., Fujiwara, Y., Eicher, E.M., and Orkin, S.H. (2002). Gonadal differentiation, sex determination and normal Sry expression in mice require direct interaction between transcription partners GATA4 and FOG2. *Development (Cambridge, England)* 129, 4627–4634.

Thaller, G., Dempfle, L., and Hoeschele, I. (1996). Investigation of the inheritance of birth defects in swine by complex segregation analysis. *Journal of Animal Breeding and Genetics* 113, 77–92.

The Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661–678.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* 14, 178–192.

Tomaselli, S., Megiorni, F., De Bernardo, C., Felici, A., Marrocco, G., Maggiulli, G., Grammatico, B., Remotti, D., Saccucci, P., Valentini, F., et al. (2008). Syndromic true hermaphroditism due to an R-spondin1 (RSPO1) homozygous mutation. *Human Mutation* *29*, 220–226.

Tomaselli, S., Megiorni, F., Lin, L., Mazzilli, M.C., Gerrelli, D., Majore, S., Grammatico, P., and Achermann, J.C. (2011). Human RSPO1/R-spondin1 is expressed during early ovary development and augments  $\beta$ -catenin signaling. *PLoS One* *6*, e16366.

Tomizuka, K., Horikoshi, K., Kitada, R., Sugawara, Y., Iba, Y., Kojima, A., Yoshitome, A., Yamawaki, K., Amagai, M., Inoue, A., et al. (2008). R-spondin1 plays an essential role in ovarian development through positively regulating Wnt-4 signaling. *Human Molecular Genetics* *17*, 1278–1291.

Torres, M., Gómez-Pardo, E., Dressler, G.R., and Gruss, P. (1995). Pax-2 controls multiple steps of urogenital development. *Development (Cambridge, England)* *121*, 4057–4065.

Tortereau, F., Servin, B., Frantz, L., Megens, H.-J., Milan, D., Rohrer, G., Wiedmann, R., Beever, J., Archibald, A.L., Schook, L.B., et al. (2012). A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content. *BMC Genomics* *13*, 586.

Turleau, C., and Prieur, M. (2000). Types, fréquences et mécanismes de formation des anomalies chromosomiques.

Uimari, P., and Tapio, M. (2011). Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. *Journal of Animal Science* *89*, 609–614.

Vainio, S., Heikkilä, M., Kispert, a, Chin, N., and McMahon, a P. (1999). Female development in mammals is regulated by Wnt-4 signalling. *Nature* *397*, 405–409.

Val, P., and Swain, A. (2010). Gene dosage effects and transcriptional regulation of early mammalian adrenal cortex development. *Molecular and Cellular Endocrinology* *323*, 105–114.

VanRaden, P.M., Van Tassell, C.P., Wiggans, G.R., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F., and Schenkel, F.S. (2009). Invited review: reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science* *92*, 16–24.

Veeramachaneni, D.N.R., Amann, R.P., and Jacobson, J.P. (2005). Testis and Antler Dysgenesis in Sitka Black-Tailed Deer on Kodiak Island, Alaska: Sequela of Environmental Endocrine Disruption? *Environmental Health Perspectives* *114*, 51–59.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C. a, Holt, R. a, et al. (2001). The sequence of the human genome. *Science (New York, N.Y.)* *291*, 1304–1351.

Vetro, A., Ciccone, R., Giorda, R., Patricelli, M.G., Della Mina, E., Forlino, A., and Zuffardi, O. (2011). XX males SRY negative: a confirmed cause of infertility. *Journal of Medical Genetics* *48*, 710–712.

Vidal, V.P., Chaboissier, M.C., de Rooij, D.G., and Schedl, a (2001). Sox9 induces testis development in XX transgenic mice. *Nature Genetics* *28*, 216–217.

Vincent, A.L., Tuggle, C.K., Rothschild, M.F., Evans, G., Short, T.H., Southwood, O.I., and Plastow, G.S. (1998). The Prolactin Receptor Gene is Associated with Increased Litter Size In Pigs. Swine Research Report, 1997. Paper 11.

Vingborg, R.K.K., Gregersen, V.R., Zhan, B., Panitz, F., Høj, A., Sørensen, K.K., Madsen, L.B., Larsen, K., Hornshøj, H., Wang, X., et al. (2009). A robust linkage map of the porcine autosomes based on gene-associated SNPs. *BMC Genomics* 10, 134.

Virtanen, H.E., Tapanainen, A.E., Kaleva, M.M., Suomi, A.-M., Main, K.M., Skakkebaek, N.E., and Toppari, J. (2006). Mild gestational diabetes as a risk factor for congenital cryptorchidism. *The Journal of Clinical Endocrinology and Metabolism* 91, 4862–4865.

Voorrips, R.E. (2002). MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs.

Wagner, K.-D. (2003). The complex life of WT1. *Journal of Cell Science* 116, 1653–1658.

Wagner, T., Wirth, J., Meyer, J., Zabel, B., Held, M., Zimmer, J., Pasantes, J., Bricarelli, F., Keutel, J., Hustert, E., et al. (1994). Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell* 79, 1111–1120.

Walters, J.R. (2010). Have we forgotten about inherited disease? *UPB Genetic World* 79–86.

Walters, E.M., Wolf, E., Whyte, J.J., Mao, J., Renner, S., Nagashima, H., Kobayashi, E., Zhao, J., Wells, K.D., Critser, J.K., et al. (2012). Completion of the swine genome will simplify the production of swine as a large animal biomedical model. *BMC Medical Genomics* 5, 55.

Wang, J., and Wang, B. (2002). [Study on risk factors of cryptorchidism]. *Zhonghua Liu Xing Bing Xue Za Zhi = Zhonghua Liuxingbingxue Zazhi* 23, 190–193.

Wang, J., Jiang, J., Fu, W., Jiang, L., Ding, X., Liu, J.-F., and Zhang, Q. (2012). A genome-wide detection of copy number variations using SNP genotyping arrays in swine. *BMC Genomics* 13, 273.

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F.A., Hakonarson, H., and Bucan, M. (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research* 17, 1665–1674.

Wang, Z., Liu, X., Yang, B.-Z., and Gelernter, J. (2013). The Role and Challenges of Exome Sequencing in Studies of Human Diseases. *Frontiers in Genetics* 4, 160.

Warr, N., Carre, G.-A., Siggers, P., Vitos Faleato, J., Brixey, R., Pope, M., Bogani, D., Childers, M., Wells, S., Scudamore, C.L., et al. (2012). Gadd45γ and Map3k4 interactions regulate mouse testis determination via p38 MAPK-Mediated control of Sry expression. *Developmental Cell* 1020–1031.

Warwick, B.L. (1926). A Study of Hernia in Swine.

Weiss, M.C., and Green, H. (1967). Human-mouse hybrid cell lines containing partial complements of human chromosomes and functioning human genes. *Proceedings of the National Academy of Sciences of the United States of America*.

- Weissenbach, J., and Rouyer, F. (1989). Chromosome Y et détermination du sexe.
- White, S., Ohnesorg, T., Notini, A., Roeszler, K., Hewitt, J., Daggag, H., Smith, C., Turbitt, E., Gustin, S., van den Bergen, J., et al. (2011). Copy number variation in patients with disorders of sex development due to 46,XY gonadal dysgenesis. *PloS One* 6, e17793.
- Whitworth, D.J., Shaw, G., and Renfree, M.B. (1996). Gonadal sex reversal of the developing marsupial ovary in vivo and in vitro. *Development (Cambridge, England)* 122, 4057–4063.
- Wiener, A.S. (1932). Method of measuring linkage in human genetics with special reference to blood groups\_Wiener\_1931. *Genetics*.
- Wilhelm, D., and Englert, C. (2002). The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sf1. *Gene and Development* 1839–1851.
- Wilhelm, D., Palmer, S., and Koopman, P. (2007). Sex Determination and Gonadal Development in Mammals. *Physiological Reviews* 1–28.
- Wilkinson, S., Lu, Z.H., Megens, H.-J., Archibald, A.L., Haley, C., Jackson, I.J., Groenen, M. a M., Crooijmans, R.P.M. a, Ogden, R., and Wiener, P. (2013). Signatures of diversifying selection in European pig breeds. *PLoS Genetics* 9, e1003453.
- Wilson, A.S., Power, B.E., and Molloy, P.L. (2007). DNA hypomethylation and human diseases. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 1775, 138–162.
- Womack, J.E., Johnson, J.S., Owens, E.K., Rexroad, C.E., Schläpfer, J., and Yang, Y.P. (1997). A whole-genome radiation hybrid panel for bovine gene mapping. *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 8, 854–856.
- Wong, A.K., Ruhe, A.L., Dumont, B.L., Robertson, K.R., Guerrero, G., Shull, S.M., Ziegler, J.S., Millon, L. V, Broman, K.W., Payseur, B. a, et al. (2010). A comprehensive linkage map of the dog genome. *Genetics* 184, 595–605.
- Worldwatch Institute (2011). Worldwatch Institute.
- Wright, C.F., Wei, Y., Higgins, J.P.T., and Sagoo, G.S. (2012). Non-invasive prenatal diagnostic test accuracy for fetal sex using cell-free DNA a review and meta-analysis. *BMC Research Notes* 5, 476.
- Wyman, A.R., and White, R.A.Y. (1980). A highly polymorphic locus in human DNA. 77, 6754–6758.
- Xiangdong, L., Yuanwu, L., Hua, Z., Liming, R., Qiuyan, L., and Ning, L. (2011). Animal models for the atherosclerosis research: a review. *Protein & Cell* 2, 189–201.
- Xiong, M., and Guo, S. (1997). Fine-Scale Genetic Mapping Based Theory and Applications Linkage Disequilibrium: *American Journal of Human Genetics* 1513–1531.
- Yao, H.H., Whoriskey, W., and Capel, B. (2002). Desert Hedgehog / Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. *Gene and Development* 1433–1440.

- Yao, H.H.C., Matzuk, M.M., Jorgez, C.J., Menke, D.B., Page, D.C., Swain, A., and Capel, B. (2004). Follistatin operates downstream of Wnt4 in mammalian ovary organogenesis. *Developmental Dynamics: an Official Publication of the American Association of Anatomists* 230, 210–215.
- Yeo, G.S.H. (2011). Where next for GWAS? *Briefings in Functional Genomics* 10, 51.
- Yerle, M., Echard, G., Robic, a, Mairal, a, Dubut-Fontana, C., Riquet, J., Pinton, P., Milan, D., Lahbib-Mansais, Y., and Gellin, J. (1996). A somatic cell hybrid panel for pig regional gene mapping characterized by molecular cytogenetics. *Cytogenetics and Cell Genetics* 73, 194–202.
- Yerle, M., Pinton, P., Robic, a, Alfonso, a, Palvadeau, Y., Delcros, C., Hawken, R., Alexander, L., Beattie, C., Schook, L., et al. (1998). Construction of a whole-genome radiation hybrid panel for high-resolution gene mapping in pigs. *Cytogenetics and Cell Genetics* 82, 182–188.
- Yerle, M., Pinton, P., Delcros, C., Arnal, N., Milan, D., and Robic, a (2002). Generation and characterization of a 12,000-rad radiation hybrid panel for fine mapping in pig. *Cytogenetic and Genome Research* 97, 219–228.
- Yoshida, M., Suda, Y., Matsuo, I., Miyamoto, N., Takeda, N., Kuratani, S., and Aizawa, S. (1997). Emx1 and Emx2 functions in development of dorsal telencephalon. *Development (Cambridge, England)* 124, 101–111.
- Young, A.C., Kirkness, E.F., and Breen, M. (2008). Tackling the characterization of canine chromosomal breakpoints with an integrated in-situ/in-silico approach: the canine PAR and PAB. *Chromosome Research* 16, 1193–1202.
- Yuan, F.P., Lin, D.X., Rao, C. V, and Lei, Z.M. (2006). Cryptorchidism in LhrKO animals and the effect of testosterone-replacement therapy. *Human Reproduction (Oxford, England)* 21, 936–942.
- Zakaria, M., Azab, S., El Baz, M., Fawaz, L., and Bahagat, A. (2012). Cryptorchidism in Egyptian neonates. *Journal of Pediatric Urology* 5, 95.
- Zeggini, E., Scott, L.J., Saxena, R., Voight, B.F., Marchini, L., Hu, T., Bakker, P.I.W. De, Abecasis, G.R., and Almgren, P. (2009). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *40*, 638–645.
- Zhang, H., Wang, Z., Wang, S., and Li, H. (2012). Progress of genome wide association study in domestic animals. *Journal of Animal Science and Biotechnology* 3, 26.
- Zhao, X. (2009). A candidate gene association study of cryptorchidism and scrotal hernia using canine and porcine models.
- Zhao, X., Du, Z.-Q., Vukasinovic, N., Rodriguez, F., Clutter, A.C., and Rothschild, M.F. (2009). Association of HOXA10, ZFPM2, and MMP2 genes with scrotal hernias evaluated via biological candidate gene analyses in pigs. *American Journal of Veterinary Research* 70, 1006–1012.
- Zhou, X., Carbonetto, P., and Stephens, M. (2013). Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genetics* 9, e1003264.

(1966). Loi sur l'élevage.

(2012). FAO food and agriculture organization of the United Nations.





# Annexe 1

## Chapitre XVII : Malformations congénitales et anomalies chromosomiques (Q00-Q99)

### Malformations congénitales du système nerveux (Q00-Q07)

#### Q00 Anencéphalie et malformations similaires

##### Q00.0 Anencéphalie

Acéphalie  
Acrânie  
Anencéphalomyélie  
Hémianencéphalie  
Hémicéphalie

##### Q00.1 Cranio-rachischisis

##### Q00.2 Inencéphalie

#### Q01 Encéphalocèle

##### *Inclus:*

encéphalomyélocèle  
hydro-encéphalocèle  
hydroméningocèle crânienne  
méningocèle cérébrale  
méningo-encéphalocèle

##### *Excl.:*

syndrome de Meckel-Gruber (Q61.9)

##### Q01.0 Encéphalocèle frontale

##### Q01.1 Encéphalocèle naso-frontale

##### Q01.2 Encéphalocèle occipitale

##### Q01.8 Encéphalocèle d'autres localisations

##### Q01.9 Encéphalocèle, sans précision

#### Q02 Microcéphalie

##### *Inclus:*

Hydromicrocéphalie  
Micro-encéphale

##### *Excl.:*

syndrome de Meckel-Gruber (Q61.9)

#### Q03 Hydrocéphalie congénitale

##### *Inclus:*

hydrocéphalie du nouveau-né

##### *Excl.:*

hydrocéphalie:

- acquise (G91.-)
- avec spina bifida (Q05.0-Q05.4)
- due à toxoplasmose congénitale (P37.1)

syndrome d'Arnold-Chiari (Q07.0)

##### Q03.0 Malformations de l'aqueduc de Sylvius

Anomalie	
Obstruction congénitale	de l'aqueduc de Sylvius
Sténose	

##### Q03.1 Atésie des fentes de Luschka et du foramen de Magendie

Syndrome de Dandy-Walker

##### Q03.8 Autres hydrocéphalies congénitales

##### Q03.9 Hydrocéphalie congénitale, sans précision

#### Q04 Autres malformations congénitales de l'encéphale

##### *Excl.:*

cyclopie (Q87.0)  
macrocéphalie (Q75.3)

**Q04.0 Malformations congénitales du corps calleux**

Agénésie du corps calleux

**Q04.1 Arhinencéphalie****Q04.2 Holoprosencéphalie****Q04.3 Autres anomalies localisées du développement de l'encéphale**

Absence		d'une partie de l'encéphale
Agénésie		
Aplasia		
Hypoplasie		

Agyrie  
Hydranencéphalie  
Lissencéphalie  
Microgyrie  
Pachygyrie

**Excl.:**

malformations congénitales du corps calleux (Q04.0)

**Q04.4 Dysplasie du septum et des voies optiques****Q04.5 Mégalencéphalie****Q04.6 Kystes cérébraux congénitaux**

Porencéphalie  
Schizencéphalie

**Excl.:**

kyste porencéphalique acquis (G93.0)

**Q04.8 Autres malformations congénitales précisées de l'encéphale**

Macrogyrie

**Q04.9 Malformation congénitale de l'encéphale, sans précision**

Anomalie congénitale		SAI de l'encéphale
Anomalies multiples congénitales		
Maladie ou lésion congénitale		

**Q05 Spina bifida****Inclus:**

hydroméningocèle (rachidienne)  
méningocèle (rachidienne)  
méningomyélocèle  
myélocèle  
myéломéningocèle  
rachischisis  
spina bifida (aperta) (kystique)  
syringomyélocèle

**Excl.:**

spina bifida occulta (Q76.0)  
syndrome d'Arnold-Chiari (Q07.0)

**Q05.0 Spina bifida cervical, avec hydrocéphalie****Q05.1 Spina bifida thoracique, avec hydrocéphalie**

Spina bifida:		avec hydrocéphalie
dorsal		
dorso-lombaire		

**Q05.2 Spina bifida lombaire, avec hydrocéphalie**

Spina bifida lombo-sacré avec hydrocéphalie

**Q05.3 Spina bifida sacré, avec hydrocéphalie****Q05.4 Spina bifida, avec hydrocéphalie, sans précision****Q05.5 Spina bifida cervical, sans hydrocéphalie****Q05.6 Spina bifida thoracique, sans hydrocéphalie**

Spina bifida:

- dorsal SAI
- dorso-lombaire SAI

**Q05.7 Spina bifida lombaire, sans hydrocéphalie**

Spina bifida lombo-sacré SAI

**Q05.8 Spina bifida sacré, sans hydrocéphalie****Q05.9 Spina bifida, sans précision****Q06 Autres malformations congénitales de la moelle épinière****Q06.0 Amyélie****Q06.1 Hypoplasie et dysplasie de la moelle épinière**

Atéломylie  
Myélatélie  
Myélodysplasie de la moelle épinière

**Q06.2 Diastématomyélie**

**Q06.3 Autres malformations congénitales de la queue de cheval**

**Q06.4 Hydromylie**

Hydrorachis

**Q06.8 Autres malformations congénitales précisées de la moelle épinière**

**Q06.9 Malformation congénitale de la moelle épinière, sans précision**

Anomalie congénitale

Maladie ou lésion congénitale

SAI de la moelle épinière et des méninges

---

**Q07 Autres malformations congénitales du système nerveux**

*Excl.:*

dysautonomie familiale [Riley-Day] (Q90.1)

neurofibromatose (non maligne) (Q85.0)

**Q07.0 Syndrome d'Arnold-Chiari**

**Q07.8 Autres malformations congénitales précisées du système nerveux**

Agénésie d'un nerf

Anomalie d'un plexus nerveux

Mâchoire à clignement

Phénomène de Marcus Gunn

**Q07.9 Malformation congénitale du système nerveux, sans précision**

Anomalie congénitale

Maladie ou lésion congénitale

SAI du système nerveux

---

**Malformations congénitales de l'œil, de l'oreille, de la face et du cou  
(Q10-Q18)**

*Excl.:*

anomalie morphologique congénitale du rachis (Q67.5)

fente labiale et fente palatine (Q35-Q37)

malformation congénitale (de):

- glande:
  - parathyroïde (Q89.2)
  - thyroïde (Q89.2)
- larynx (Q31.-)
- lèvre NCA (Q38.0)
- nez (Q30.-)
- rachis cervical (Q05.0, Q05.5, Q76.0-Q76.4)

---

**Q10 Malformations congénitales des paupières, de l'appareil lacrymal et de l'orbite**

*Excl.:*

cryptophtalmie SAI (Q11.2)

syndrome cryptophtalmique (Q87.0)

**Q10.0 Ptosis congénital**

**Q10.1 Ectropion congénital**

**Q10.2 Entropion congénital**

**Q10.3 Autres malformations congénitales des paupières**

Ablépharie

Absence ou agénésie de:

- cils
- paupière

Blépharophimosis congénital

Colobome de la paupière

Malformation congénitale de la paupière SAI

Muscle oculaire surnuméraire

Paupière surnuméraire

**Q10.4 Absence et agénésie de l'appareil lacrymal**

Absence d'orifice lacrymal

**Q10.5 Sténose ou rétrécissement congénital du canal lacrymal**

**Q10.6 Autres malformations congénitales de l'appareil lacrymal**

Malformation congénitale de l'appareil lacrymal SAI

**Q10.7 Malformation congénitale de l'orbite**

---

**Q11 Anophtalmie, microptalmie et macroptalmie**

**Q11.0 Œil kystique**

**Q11.1 Autres formes d'anophtalmie**

Agénésie

Aplasie

de l'œil

**Q11.2 Microptalmie**

Cryptophtalmie SAI  
Dysplasie de l'œil  
Hypoplasie de l'œil  
Œil rudimentaire

**Excl.:**

syndrome cryptophtalmique (Q87.0)

**Q11.3Macrophtalmie**

**Excl.:**

macrophtalmie au cours d'un glaucome congénital (Q15.0)

---

**Q12Malformations congénitales du cristallin**

**Q12.0Cataracte congénitale**

**Q12.1Luxation congénitale du cristallin**

**Q12.2Colobome du cristallin**

**Q12.3Aphakie congénitale**

**Q12.4Sphérophakie**

**Q12.8Autres malformations congénitales du cristallin**

**Q12.9Malformation congénitale du cristallin, sans précision**

---

**Q13Malformations congénitales de la chambre antérieure de l'œil**

**Q13.0Colobome de l'iris**

Colobome SAI

**Q13.1Absence d'iris**

Aniridie

**Q13.2Autres malformations congénitales de l'iris**

Anisocorie congénitale

Atrésie de la pupille

Corectopie

Ectopie pupillaire

Malformation congénitale de l'iris SAI

**Q13.3Opacité congénitale de la cornée**

**Q13.4Autres malformations congénitales de la cornée**

Anomalie de Peter

Malformation congénitale de la cornée SAI

Microcornée

**Q13.5Sclérotique bleue**

**Q13.8Autres malformations congénitales de la chambre antérieure de l'œil**

Anomalie de Rieger

**Q13.9Malformation congénitale de la chambre antérieure de l'œil, sans précision**

---

**Q14Malformations congénitales de la chambre postérieure de l'œil**

**Q14.0Malformation congénitale du corps vitré**

Opacité congénitale du corps vitré

**Q14.1Malformation congénitale de la rétine**

Anévrisme congénital de la rétine

**Q14.2Malformation congénitale de la papille optique**

Colobome congénital de la papille optique

**Q14.3Malformation congénitale de la choroïde**

**Q14.8Autres malformations congénitales de la chambre postérieure de l'œil**

Colobome du fond de l'œil

**Q14.9Malformation congénitale de la chambre postérieure de l'œil, sans précision**

---

**Q15Autres malformations congénitales de l'œil**

**Excl.:**

albinisme oculaire (E70.3)

nystagmus congénital (H55)

rétinite pigmentaire (H35.5)

**Q15.0Glaucome congénital**

Buphtalmie

Glaucome du nouveau-né

Hydrophthalmie

Kératoglobe congénital avec glaucome

Macrocornée ou mégalocornée avec glaucome

Macrophtalmie au cours d'un glaucome congénital

**Q15.8Autres malformations congénitales précisées de l'œil**

**Q15.9Malformation congénitale de l'œil, sans précision**

Anomalie congénitale SAI de l'œil

---

**Q16Malformations congénitales de l'oreille avec atteinte de l'audition**

**Excl.:**

surdit e cong enitale (H90.-)

**Q16.0Absence cong enitale du pavillon de l'oreille**

**Q16.1Absence, atr sie et r tr cissement cong enitaux du conduit auditif (externe)**

Atr sie ou r tr cissement du m at auditif osseux

**Q16.2Absence de trompe d'Eustache**

**Q16.3Malformation cong enitale des osselets**

Coalescence des osselets

**Q16.4Autres malformations congénitales de l'oreille moyenne**

Malformation congénitale de l'oreille moyenne SAI

**Q16.5Malformation congénitale de l'oreille interne**

Anomalie de:

- labyrinthe membraneux
- organe de Corti

**Q16.9Malformation congénitale de l'oreille avec atteinte de l'audition, sans précision**

Absence congénitale de l'oreille SAI

---

**Q17Autres malformations congénitales de l'oreille**

*Excl.:*

fistule préauriculaire (Q18.1)

**Q17.0Pavillon surnuméraire**

Appendice préauriculaire  
Lobule surnuméraire  
Oreille surnuméraire  
Polyotie  
Tragus accessoire

**Q17.1Macrotie**

**Q17.2Microtie**

**Q17.3Autres déformations de l'oreille**

Oreilles pointues

**Q17.4Anomalie de position de l'oreille**

Position basse de l'oreille

*Excl.:*

pavillon cervical (Q18.2)

**Q17.5Oreilles proéminentes**

Oreille "chauve-souris"

**Q17.8Autres malformations congénitales précisées de l'oreille**

Absence congénitale du lobe de l'oreille

**Q17.9Malformation congénitale de l'oreille, sans précision**

Anomalie congénitale de l'oreille SAI

---

**Q18Autres malformations congénitales de la face et du cou**

*Excl.:*

affections classées en Q67.0-Q67.4

anomalies dento-faciales [y compris les malocclusions] (K07.-)

cyclopie (Q87.0)

fente labiale et fente palatine (Q35-Q37)

malformations congénitales des os du crâne et de la face (Q75.-)

persistance du canal thyroïdienne (Q89.2)

syndromes congénitaux malformatifs affectant l'aspect de la face (Q87.0)

**Q18.0Fissure, fistule et kyste d'origine branchiale**

Vestiges branchiaux

**Q18.1Fistule et kyste préauriculaires**

Fistule (du):

- cervico-auriculaire
- pavillon, congénitale

**Q18.2Autres malformations d'origine branchiale**

Malformations de la fente branchiale SAI

Otocéphalie

Pavillon cervical

**Q18.3Cou palmé**

Pterygium colli

**Q18.4Macrostomie**

**Q18.5Microstomie**

**Q18.6Macrocheilie**

Hypertrophie congénitale de la lèvre

**Q18.7Microcheilie**

**Q18.8Autres malformations congénitales précisées de la face et du cou**

Fistule

Kyste

Sinus

médian(e) de la face et du cou

**Q18.9Malformation congénitale de la face et du cou, sans précision**

Anomalie congénitale de la face et du cou SAI

---

**Malformations congénitales de l'appareil circulatoire  
(Q20-Q28)**

**Q20Malformations congénitales des cavités et des orifices cardiaques**

*Excl.:*

dextrocardie avec situs inversus (Q89.3)

disposition auriculaire en miroir avec situs inversus (Q89.3)

**Q20.0Tronc artériel commun**

- Persistence du tronc artériel
- Q20.1 Ventricule droit à double issue**  
Syndrome de Taussig-Bing
- Q20.2 Ventricule gauche à double issue**
- Q20.3 Communication ventriculo-auriculaire discordante**  
Dextroposition de l'aorte  
Transposition des gros vaisseaux (complète)
- Q20.4 Ventricule à double issue**  
Cœur triloculaire biauriculaire  
Ventricule:
- commun
  - unique
- Q20.5 Communication auriculo-ventriculaire discordante**  
Inversion ventriculaire  
Transposition:
- corrigée
  - gauche
- Q20.6 Isomérisme des auricules cardiaques**  
Isomérisme des auricules cardiaques avec asplénie ou polysplénie
- Q20.8 Autres malformations congénitales des cavités et des orifices cardiaques**
- Q20.9 Malformation congénitale des cavités et des orifices cardiaques, sans précision**

---

**Q21 Malformations congénitales des cloisons cardiaques**

*Excl.:*

anomalie d'un septum cardiaque, acquise ([I51.0](#))

**Q21.0 Communication interventriculaire**

**Q21.1 Communication interauriculaire**

Malformation du sinus:

- coronaire
- veineux

Perméabilité ou persistance de:

- foramen ovale
- ostium secundum (type II)

**Q21.2 Communication auriculo-ventriculaire**

Canal auriculo-ventriculaire commun  
Malformation des replis endocardiques  
Persistence de l'ostium primum (type I)

**Q21.3 Tétralogie de Fallot**

Communication interventriculaire avec sténose ou atrésie pulmonaire, dextroposition de l'aorte et hypertrophie du ventricule droit.

**Q21.4 Communication aorto-pulmonaire**

Fistule aorto-pulmonaire  
Malformation du septum aortique

**Q21.8 Autres malformations congénitales des cloisons cardiaques**

Malformation cardiaque congénitale d'Eisenmenger  
Pentalogie de Fallot

*Excl.:*

complexe d'Eisenmenger ([I27.8](#))  
syndrome d'Eisenmenger ([I27.8](#))

**Q21.9 Malformation congénitale d'une cloison cardiaque, sans précision**

Communication (cardiaque) SAI

---

**Q22 Malformations congénitales de la valve pulmonaire et de la valvule tricuspide**

**Q22.0 Atrésie de la valve pulmonaire**

**Q22.1 Sténose congénitale de la valve pulmonaire**

**Q22.2 Insuffisance congénitale de la valve pulmonaire**

Reflux congénital au niveau de la valve pulmonaire

**Q22.3 Autres malformations congénitales de la valve pulmonaire**

Malformation congénitale de la valve pulmonaire SAI

**Q22.4 Sténose congénitale de la valvule tricuspide**

Atrésie de la valvule tricuspide

**Q22.5 Maladie d'Ebstein**

**Q22.6 Hypoplasie du cœur droit**

**Q22.8 Autres malformations congénitales de la valvule tricuspide**

**Q22.9 Malformation congénitale de la valvule tricuspide, sans précision**

---

**Q23 Malformations congénitales des valvules aortique et mitrale**

**Q23.0 Sténose congénitale de la valvule aortique**

Atrésie	aortique congénitale
Sténose	

*Excl.:*

au cours d'hypoplasie du cœur gauche ([Q23.4](#))  
sténose sous-aortique congénitale ([Q24.4](#))

**Q23.1 Insuffisance congénitale de la valvule aortique**

Insuffisance aortique congénitale  
Valvule aortique bicuspide

**Q23.2 Sténose mitrale congénitale**

Atrésie mitrale congénitale

**Q23.3 Insuffisance mitrale congénitale****Q23.4 Hypoplasie du cœur gauche**

Atrésie ou hypoplasie marquée de l'orifice ou de la valvule aortique, associée à une hypoplasie de l'aorte ascendante et à un défaut de développement du ventricule gauche (avec sténose ou atrésie de la valvule mitrale).

**Q23.8 Autres malformations congénitales des valvules aortique et mitrale****Q23.9 Malformation congénitale des valvules aortique et mitrale, sans précision****Q24 Autres malformations congénitales cardiaques****Excl.:**

fibroélastose endocardique ([I42.4](#))

**Q24.0 Dextrocardie****Excl.:**

dextrocardie avec situs inversus ([Q89.3](#))  
disposition auriculaire en miroir avec situs inversus ([Q89.3](#))  
isomérisme des auricules cardiaques (avec asplénie ou polysplénie) ([Q20.6](#))

**Q24.1 Lévocardié**

Cœur situé dans l'hémithorax gauche avec apex pointant à gauche, avec situs inversus des autres viscères et anomalie cardiaque ou correction d'une transposition des gros vaisseaux

**Q24.2 Cœur triatrial****Q24.3 Sténose de l'infundibulum pulmonaire****Q24.4 Sténose sous-aortique congénitale****Q24.5 Malformation des vaisseaux coronaires**

Anévrisme congénital coronaire (artère)

**Q24.6 Bloc congénital du cœur****Q24.8 Autres malformations cardiaques congénitales précisées**

Diverticule congénital du ventricule gauche  
Malformation congénitale du:

- myocarde
- péricarde

Malposition du cœur

Syndrome de Uhl

**Q24.9 Malformation cardiaque congénitale, sans précision**

Anomalie cardiaque		
Cardiopathie	congénital(e)	de l'aorte

**Q25 Malformations congénitales des gros vaisseaux****Q25.0 Perméabilité du canal artériel**

Perméabilité du canal de Botal  
Persistance du canal artériel

**Q25.1 Coarctation de l'aorte**

Coarctation de l'isthme aortique (avant le canal) (après le canal)

**Q25.2 Atrésie de l'aorte****Q25.3 Sténose de l'aorte**

Sténose aortique supra-avalvulaire

**Excl.:**

sténose aortique congénitale ([Q23.0](#))

**Q25.4 Autres malformations congénitales de l'aorte**

Absence			
Anévrisme		congénital(e)	de l'aorte
Dilatation			
Aplasia			

Anévrisme du sinus de Valsalva (rompu)

Double arc aortique [anneau vasculaire de l'aorte]

Hypoplasie de l'aorte

Persistance de:

- arc aortique droit
- convolution de l'arc aortique

**Excl.:**

hypoplasie aortique associée à une hypoplasie du cœur gauche ([Q23.4](#))

**Q25.5 Atrésie de l'artère pulmonaire****Q25.6 Sténose de l'artère pulmonaire****Q25.7 Autres malformations congénitales de l'artère pulmonaire**

Agénésie			
Anévrisme congénital		congénital(e)	de l'artère pulmonaire
Anomalie			

Hypoplasie  
 Anévrisme pulmonaire artério-veineux  
 Artère pulmonaire aberrante  
**Q25.8Autres malformations congénitales des gros vaisseaux**  
**Q25.9Malformation congénitale des gros vaisseaux, sans précision**

---

**Q26Malformations congénitales des grandes veines**

**Q26.0Sténose congénitale de la veine cave**  
 Sténose congénitale de la veine cave (inférieure) (supérieure)  
**Q26.1Persistance de la veine cave supérieure gauche**  
**Q26.2Retour veineux pulmonaire anormal total**  
**Q26.3Retour veineux pulmonaire anormal partiel**  
**Q26.4Retour veineux pulmonaire anormal, sans précision**  
**Q26.5Retour veineux portal anormal**  
**Q26.6Fistule entre la veine porte et l'artère hépatique**  
**Q26.8Autres malformations congénitales des grandes veines**  
 Absence de veine cave (inférieure) (supérieure)  
 Continuation de la veine cave inférieure dans la veine azygos  
 Persistance de la veine cardinale postérieure gauche  
 Syndrome du cimenterre  
**Q26.9Malformation congénitale d'une grande veine, sans précision**  
 Anomalie de la veine cave (inférieure) (supérieure) SAI

---

**Q27Autres malformations congénitales de l'appareil circulatoire périphérique**

**Excl.:**  
 anévrisme congénital de la rétine (Q14.1)  
 anomalies de:

- artère pulmonaire (Q25.5-Q25.7)
- vaisseaux cérébraux et précérébraux (Q28.0-Q28.3)
- vaisseaux coronaires (Q24.5)

hémangiome et lymphangiome (D18.-)  
**Q27.0Absence congénitale et hypoplasie de l'artère ombilicale**  
 Artère ombilicale unique  
**Q27.1Sténose congénitale de l'artère rénale**  
**Q27.2Autres malformations congénitales de l'artère rénale**  
 Artères rénales multiples  
 Malformation congénitale de l'artère rénale SAI  
**Q27.3Malformation artérioveineuse périphérique**  
 Anévrisme artérioveineux  
**Excl.:**  
 anévrisme artérioveineux acquis (I77.0)  
**Q27.4Phlébectasie congénitale**  
**Q27.8Autres malformations congénitales précisées de l'appareil circulatoire périphérique**

Absence Atrésie	d'une artère ou d'une veine NCA	
Anévrisme (périphérique) Rétrécissement artériel Varice		congénital(e)

Artère sous-clavière aberrante  
**Q27.9Malformation congénitale de l'appareil circulatoire périphérique, sans précision**  
 Anomalie artérielle ou veineuse SAI

---

**Q28Autres malformations congénitales de l'appareil circulatoire**

**Excl.:**  
 anévrisme congénital:

- SAI (Q27.8)
- coronaire (Q24.5)
- périphérique (Q27.8)
- pulmonaire (Q25.7)
- rétinien (Q14.1)

rupture de malformation (de):

- cérébrale (artérioveineuse) (I60.8)
- vaisseaux précérébraux (I72.-)

**Q28.0Malformation artérioveineuse de vaisseaux précérébraux**  
 Anévrisme artérioveineux précérébral congénital (non rompu)  
**Q28.1Autres malformations des vaisseaux précérébraux**  
 Anévrisme précérébral congénital (non rompu)  
 Malformation congénitale des vaisseaux précérébraux SAI  
**Q28.2Malformation artérioveineuse des vaisseaux cérébraux**  
 Anévrisme artérioveineux cérébral congénital (non rompu)

Malformation artérioveineuse du cerveau SAI  
**Q28.3Autres malformations des vaisseaux cérébraux**  
Anévrisme cérébral congénital (non rompu)  
Malformation congénitale des vaisseaux cérébraux SAI  
**Q28.8Autres malformations congénitales précisées de l'appareil circulatoire**  
Anévrisme congénital, localisation précisée NCA  
**Q28.9Malformation congénitale de l'appareil circulatoire, sans précision**

## Malformations congénitales de l'appareil respiratoire (Q30-Q34)

### Q30 Malformations congénitales du nez

*Excl.:*

déviations congénitales de la cloison nasale (Q67.4)

#### Q30.0 Atrésie des choanes

Atrésie

Sténose congénitale

d'un orifice nasal (antérieur) (postérieur)

#### Q30.1 Agénésie et hypoplasie du nez

Absence congénitale du nez

#### Q30.2 Fissure, échancrure et fente nasales

#### Q30.3 Perforation congénitale de la cloison nasale

#### Q30.8 Autres malformations congénitales du nez

Anomalie congénitale de la paroi des sinus de la face

Nez surnuméraire

#### Q30.9 Malformation congénitale du nez, sans précision

### Q31 Malformations congénitales du larynx

*Excl.:*

stridor laryngé congénital SAI (P28.8)

#### Q31.0 Ptérygion du larynx

Ptérygion du larynx:

- SAI
- glottique
- sous-glottique

#### Q31.1 Sténose sous-glottique congénitale

#### Q31.2 Hypoplasie du larynx

#### Q31.3 Laryngocèle

#### Q31.5 Laryngomalacie congénitale

#### Q31.8 Autres malformations congénitales du larynx

Absence

Agénésie

Atrésie

de cartilage cricoïde, de l'épiglotte, de la glotte, du larynx ou du cartilage thyroïdien

Fissure (de):

- cartilage thyroïdien
- épiglotte
- postérieure du cartilage cricoïde

Sténose congénitale du larynx NCA

#### Q31.9 Malformation congénitale du larynx, sans précision

### Q32 Malformations congénitales de la trachée et des bronches

*Excl.:*

bronchiectasie congénitale (Q33.4)

#### Q32.0 Trachéomalacie congénitale

#### Q32.1 Autres malformations congénitales de la trachée

Anomalie du cartilage trachéal

Atrésie de la trachée

Dilatation

Malformation

Sténose

congénitale de la trachée

Trachéocèle congénitale

#### Q32.2 Bronchomalacie congénitale

#### Q32.3 Sténose bronchique congénitale

#### Q32.4 Autres malformations congénitales des bronches

Absence

Agénésie

Atrésie

Malformation congénitale SAI

des bronches

Diverticule bronchique

### Q33 Malformations congénitales du poumon

#### Q33.0 Poumon kystique congénital

Maladie:

kystique

congénitale du poumon

polykystique |  
Poumon en nid d'abeille, congénital

**Excl.:**

affection kystique pulmonaire, acquise ou sans précision (J98.4)

**Q33.1 Lobe pulmonaire surnuméraire**

**Q33.2 Séquestration pulmonaire**

**Q33.3 Agénésie du poumon**

Absence de poumon (lobe)

**Q33.4 Bronchiectasie congénitale**

**Q33.5 Tissu ectopique intrapulmonaire**

**Q33.6 Hypoplasie et dysplasie du poumon**

**Excl.:**

hypoplasie pulmonaire associée à une brièveté de la gestation (P28.0)

**Q33.8 Autres malformations congénitales du poumon**

**Q33.9 Malformation congénitale du poumon, sans précision**

---

**Q34.4 Autres malformations congénitales de l'appareil respiratoire**

**Q34.0 Anomalie de la plèvre**

**Q34.1 Kyste congénital du médiastin**

**Q34.8 Autres malformations congénitales précisées de l'appareil respiratoire**

Atrésie du rhinopharynx

**Q34.9 Malformation congénitale de l'appareil respiratoire, sans précision**

Absence

Anomalie SAI

congénitale d'organes respiratoires

## Fente labiale et fente palatine

**(Q35-Q37)**

Utiliser, au besoin, un code supplémentaire (Q30.2) pour identifier les malformations du nez associées.

**Excl.:**

syndrome de Pierre Robin (Q87.0)

---

**Q35 Fente palatine**

**Inclus:**

fissure du palais

palatoschisis

**Excl.:**

fente labio-palatine (Q37.-)

**Q35.1 Fente de la voûte du palais**

**Q35.3 Fente du voile du palais**

**Q35.5 Fente de la voûte et du voile du palais**

**Q35.7 Fente de la luette**

**Q35.9 Fente du palais, sans précision**

Fente du palais SAI

---

**Q36 Fente labiale**

**Inclus:**

bec-de-lièvre

cheiloschisis

fissure labiale congénitale

**Excl.:**

fente labio-palatine (Q37.-)

**Q36.0 Fente labiale bilatérale**

**Q36.1 Fente labiale médiane**

**Q36.9 Fente labiale unilatérale**

Fente labiale SAI

---

**Q37 Fente labio-palatine**

**Q37.0 Fente labiale bilatérale avec fente de la voûte du palais**

**Q37.1 Fente labiale unilatérale avec fente de la voûte du palais**

Fente labiale avec fente de la voûte du palais SAI

**Q37.2 Fente labiale bilatérale avec fente du voile du palais**

**Q37.3 Fente labiale unilatérale avec fente du voile du palais**

Fente labiale avec fente du voile du palais SAI

**Q37.4 Fente labiale bilatérale avec fente du voile et de la voûte du palais**

**Q37.5 Fente labiale unilatérale avec fente du voile et de la voûte du palais**

Fente labiale avec fente du voile et de la voûte du palais SAI

**Q37.8 Fente labiale bilatérale avec fente palatine, sans précision**

**Q37.9 Fente labiale unilatérale avec fente palatine, sans précision**

Fente labio-palatine SAI

## Autres malformations congénitales de l'appareil digestif (Q38-Q45)

### **Q38** Autres malformations congénitales de la langue, de la bouche et du pharynx

*Excl.:*

- macrostomie (Q18.4)
- microstomie (Q18.5)

#### **Q38.0** Malformations congénitales des lèvres, non classées ailleurs

- Fistule congénitale de la lèvre
- Malformation congénitale de la lèvre SAI
- Syndrome de Van der Woude

*Excl.:*

- fente:
  - labiale (Q36.-)
  - labio-palatine (Q37.-)
- macrocheillie (Q18.6)
- microcheillie (Q18.7)

#### **Q38.1** Ankyloglossie

- Brièveté anormale du frein de la langue

#### **Q38.2** Macroglossie

#### **Q38.3** Autres malformations congénitales de la langue

- Adhérence
  - Fissure
  - Malformation SAI
- congénitale de la langue
- Aglossie
  - Hypoglossie
  - Hypoplasie de la langue
  - Langue bifide
  - Microglossie

#### **Q38.4** Malformations congénitales des glandes et des canaux salivaires

- Absence
  - Atrésie
- de glandes ou de canaux salivaires
- Fistule congénitale des glandes salivaires
  - Glandes ou canaux salivaires surnuméraires

#### **Q38.5** Malformations congénitales du palais, non classées ailleurs

- Absence de luette
- Malformation congénitale du palais SAI
- Palais ogival

*Excl.:*

- fente:
  - labio-palatine (Q37.-)
  - palatine (Q35.-)

#### **Q38.6** Autres malformations congénitales de la bouche

- Malformation congénitale de la bouche SAI

#### **Q38.7** Poche pharyngée

- Diverticule du pharynx

*Excl.:*

- syndrome de la poche branchiale (D82.1)

#### **Q38.8** Autres malformations congénitales du pharynx

- Malformation congénitale du pharynx SAI

---

### **Q39** Malformations congénitales de l'œsophage

#### **Q39.0** Atrésie de l'œsophage, sans fistule

- Atrésie de l'œsophage SAI

#### **Q39.1** Atrésie de l'œsophage avec fistule trachéo-œsophagienne

- Atrésie de l'œsophage avec fistule broncho-œsophagienne

#### **Q39.2** Fistule trachéo-œsophagienne congénitale, sans atrésie

- Fistule trachéo-œsophagienne congénitale SAI

#### **Q39.3** Sténose et rétrécissement congénitaux de l'œsophage

#### **Q39.4** Ptérygion de l'œsophage

#### **Q39.5** Dilatation congénitale de l'œsophage

#### **Q39.6** Diverticule de l'œsophage

- Poche œsophagienne

#### **Q39.8** Autres malformations congénitales de l'œsophage

- Absence
  - Dédoublement
  - Déplacement congénital
- de l'œsophage

#### **Q39.9** Malformation congénitale de l'œsophage, sans précision

---

### **Q40** Autres malformations congénitales des voies digestives supérieures

#### **Q40.0** Sténose hypertrophique congénitale du pylore

- Constriction
- congénital(e) ou infantile du pylore

Hypertrophie  
Rétrécissement  
Spasme  
Sténose

**Q40.1 Hernie hiatale congénitale**

Glissement du cardia à travers le hiatus œsophagien

**Excl.:**

hernie diaphragmatique congénitale (Q79.0)

**Q40.2 Autres malformations congénitales précisées de l'estomac**

Dédoublement de l'estomac  
Déplacement congénital de l'estomac  
Diverticule congénital de l'estomac  
Estomac en sablier, congénital  
Mégalogastrie  
Microgastrie

**Q40.3 Malformation congénitale de l'estomac, sans précision**

**Q40.8 Autres malformations congénitales précisées des voies digestives supérieures**

**Q40.9 Malformation congénitale des voies digestives supérieures, sans précision**

Anomalie congénitale SAI des voies digestives supérieures

---

**Q41 Absence, atrésie et sténose congénitales de l'intestin grêle**

**Inclus:**

obstruction, occlusion et rétrécissement congénitaux de l'intestin grêle ou de l'intestin SAI

**Excl.:**

iléus méconial (E84.1)

**Q41.0 Absence, atrésie et sténose congénitales du duodénum**

**Q41.1 Absence, atrésie et sténose congénitales du jéjunum**

Imperforation du jéjunum  
Syndrome de la pelure de pomme

**Q41.2 Absence, atrésie et sténose congénitales de l'iléon**

**Q41.8 Absence, atrésie et sténose congénitales d'autres localisations précisées de l'intestin grêle**

**Q41.9 Absence, atrésie et sténose congénitales de l'intestin grêle, sans précision de localisation**

Absence, atrésie et sténose congénitales de l'intestin SAI

---

**Q42 Absence, atrésie et sténose congénitales du côlon**

**Inclus:**

occlusion, obstruction et rétrécissement congénitaux du côlon

**Q42.0 Absence, atrésie et sténose congénitales du rectum, avec fistule**

**Q42.1 Absence, atrésie et sténose congénitales du rectum, sans fistule**

Imperforation du rectum

**Q42.2 Absence, atrésie et sténose congénitales de l'anus, avec fistule**

**Q42.3 Absence, atrésie et sténose congénitales de l'anus, sans fistule**

Imperforation anale

**Q42.8 Absence, atrésie et sténose congénitales d'autres localisations du côlon**

**Q42.9 Absence, atrésie et sténose congénitales du côlon, sans précision de localisation**

---

**Q43 Autres malformations congénitales de l'intestin**

**Q43.0 Diverticule de Meckel**

Persistance du canal:

- omphalo-mésentérique
- vitellin

**Q43.1 Maladie de Hirschsprung**

Aganglionose  
Mégacôlon congénital (sans ganglion)

**Q43.2 Autres anomalies fonctionnelles congénitales du côlon**

Dilatation congénitale du côlon

**Q43.3 Malformations congénitales de fixation de l'intestin**

Adhérences [brides] congénitales:

- épiploïques anormales
- péritonéales

Malrotation du côlon  
Membrane de Jackson  
Mésentère universel

Rotation:

absente	du cæcum et du côlon
incomplète	
insuffisante	

**Q43.4 Duplication de l'intestin**

**Q43.5 Ectopie de l'anus**

**Q43.6 Fistule congénitale du rectum et de l'anus**

**Excl.:**

avec absence, atrésie et sténose (Q42.0, Q42.2)

fistule congénitale:

- recto-urétrale (Q64.7)
- recto-vaginale (Q52.2)

fistule ou sinus pilonidal (L05.-)

#### Q43.7 Persistance du cloaque

Cloaque SAI

#### Q43.8 Autres malformations congénitales précisées de l'intestin

Diverticule congénital de l'intestin

Diverticulite congénitale du côlon

Dolichocôlon

Méga-appendice

Mégaduodénum

Microcôlon

Syndrome de l'anse borgne, congénital

Transposition de:

- appendice
- côlon
- intestin

#### Q43.9 Malformation congénitale de l'intestin, sans précision

---

### Q44 Malformations congénitales de la vésicule biliaire, des voies biliaires et du foie

#### Q44.0 Agénésie, aplasie et hypoplasie de la vésicule biliaire

Absence congénitale de la vésicule biliaire

#### Q44.1 Autres malformations congénitales de la vésicule biliaire

Malformation congénitale de la vésicule biliaire SAI

Vésicule biliaire intrahépatique

#### Q44.2 Atésie des voies biliaires

#### Q44.3 Sténose et rétrécissement congénitaux des voies biliaires

#### Q44.4 Kyste du cholédoque

#### Q44.5 Autres malformations congénitales des voies biliaires

Canal hépatique surnuméraire

Duplication du canal:

- biliaire
- cystique

Malformation congénitale des voies biliaires SAI

#### Q44.6 Maladie kystique du foie

Maladie fibrokystique du foie

#### Q44.7 Autres malformations congénitales du foie

Absence de foie

Hépatomégalie

Malformation du foie SAI

congénital(e)

Foie surnuméraire

Syndrome d'Alagille

---

### Q45 Autres malformations congénitales de l'appareil digestif

*Excl.:*

hernie:

- diaphragmatique congénitale (Q79.0)
- hiatale congénitale (Q40.1)

#### Q45.0 Agénésie, aplasie et hypoplasie du pancréas

Absence congénitale du pancréas

#### Q45.1 Pancréas annulaire

#### Q45.2 Kyste pancréatique congénital

#### Q45.3 Autres malformations congénitales du pancréas et du canal pancréatique

Malformation congénitale du pancréas et du canal pancréatique SAI

Pancréas surnuméraire

*Excl.:*

diabète sucré:

- congénital (E10.-)
- néonatal (P70.2)

maladie fibrokystique du pancréas (E84.-)

#### Q45.8 Autres malformations congénitales précisées de l'appareil digestif

Absence (complète) (partielle) des voies digestives SAI

Duplication

Malposition congénitale

des organes digestifs SAI

#### Q45.9 Malformation congénitale de l'appareil digestif, sans précision

Anomalie congénitale SAI de l'appareil digestif

### Malformations congénitales des organes génitaux (Q50-Q56)

*Excl.:*

syndrome de résistance aux androgènes (E34.5)

syndromes associés à des anomalies du nombre et de la forme des chromosomes (Q90-Q99)  
testicule féminisant (E34.5)

---

**Q50 Malformations congénitales des ovaires, des trompes de Fallope et des ligaments larges**

**Q50.0 Absence congénitale d'ovaire(s)**

*Excl.:*

syndrome de Turner (Q96.-)

**Q50.1 Kyste ovarien au cours du développement**

**Q50.2 Torsion congénitale de l'ovaire**

**Q50.3 Autres malformations congénitales de l'ovaire**

Aplasia ovarienne

Malformation congénitale de l'ovaire SAI

Ovaire surnuméraire

**Q50.4 Kyste embryonnaire de la trompe de Fallope**

Kyste de la frange ovarienne

**Q50.5 Kyste embryonnaire du ligament large**

Kyste (de):

- canal de Gartner
- épophoron
- paraovarien

**Q50.6 Autres malformations congénitales des trompes de Fallope et des ligaments larges**

Absence

Atrésie

Présence en surnombre

de la trompe de Fallope ou du ligament large

Malformation congénitale de la trompe de Fallope ou du ligament large SAI

---

**Q51 Malformations congénitales de l'utérus et du col de l'utérus**

**Q51.0 Agénésie et aplasie de l'utérus**

Absence congénitale de l'utérus

**Q51.1 Utérus double avec duplication du col et du vagin**

**Q51.2 Autres duplications de l'utérus**

Utérus double SAI

**Q51.3 Utérus bicorné**

**Q51.4 Utérus unicorne**

**Q51.5 Agénésie et aplasie du col de l'utérus**

Absence congénitale du col de l'utérus

**Q51.6 Kyste embryonnaire du col de l'utérus**

**Q51.7 Fistule congénitale utéro-digestive et utéro-urinaire**

**Q51.8 Autres malformations congénitales de l'utérus et du col de l'utérus**

Hypoplasie de l'utérus et du col de l'utérus

**Q51.9 Malformation congénitale de l'utérus et du col de l'utérus, sans précision**

---

**Q52 Autres malformations congénitales des organes génitaux de la femme**

**Q52.0 Absence congénitale de vagin**

**Q52.1 Dédoublement du vagin**

Vagin cloisonné

*Excl.:*

dédoublement du vagin avec utérus double et duplication du col (Q51.1)

**Q52.2 Fistule recto-vaginale congénitale**

*Excl.:*

cloaque (Q43.7)

**Q52.3 Imperforation de l'hymen**

**Q52.4 Autres malformations congénitales du vagin**

Kyste (du):

- canal de Nüch, congénital
- embryonnaire vaginal

Malformation congénitale du vagin SAI

**Q52.5 Fusion de la vulve**

**Q52.6 Malformation congénitale du clitoris**

**Q52.7 Autres malformations congénitales de la vulve**

Absence

Kyste

Malformation SAI

congénital(e) de la vulve

**Q52.8 Autres malformations congénitales précisées des organes génitaux de la femme**

**Q52.9 Malformation congénitale des organes génitaux de la femme, sans précision**

---

**Q53 Cryptorchidie**

**Q53.0 Testicule en situation anormale**

Ectopie testiculaire unilatérale ou bilatérale

**Q53.1 Cryptorchidie unilatérale**

**Q53.2 Cryptorchidie bilatérale**

**Q53.9Cryptorchidie, sans précision**  
Cryptorchidie SAI

---

**Q54Hypospadias**

*Excl.:*

épispadias (Q64.0)

**Q54.0Hypospadias balanique**

Hypospadias:

- coronal
- glandulaire

**Q54.1Hypospadias pénien**

**Q54.2Hypospadias pénoscrotal**

**Q54.3Hypospadias périnéal**

**Q54.4Chordée congénitale**

**Q54.8Autres hypospadias**

**Q54.9Hypospadias, sans précision**

---

**Q55Autres malformations congénitales des organes génitaux de l'homme**

*Excl.:*

hydrocèle congénitale (P83.5)

hypospadias (Q54.-)

**Q55.0Absence et aplasie du testicule**

Monorchidie

**Q55.1Hypoplasie du (des) testicule(s) et du scrotum**

Fusion des testicules

**Q55.2Autres malformations congénitales du (des) testicule(s) et du scrotum**

Malformation congénitale du (des) testicule(s) et du scrotum SAI

Migration du testicule

Polyorchidie

Testicule rétractile

**Q55.3Atrésie du canal déférent**

**Q55.4Autres malformations congénitales du canal déférent, de l'épididyme, des vésicules séminales et de la prostate**

Absence ou aplasie de:

- cordon spermatique
- prostate

Malformation congénitale du canal déférent, de l'épididyme, des vésicules séminales et de la prostate SAI

**Q55.5Absence et aplasie congénitales du pénis**

**Q55.6Autres malformations congénitales du pénis**

Courbure du pénis (latérale)

Hypoplasie du pénis

Malformation congénitale du pénis SAI

**Q55.8Autres malformations congénitales précisées des organes génitaux de l'homme**

**Q55.9Malformation congénitale des organes génitaux de l'homme, sans précision**

Anomalie congénitale SAI des organes génitaux de l'homme

---

**Q56Non-différenciation sexuelle et pseudo-hermaphrodisme**

*Excl.:*

pseudo-hermaphrodisme:

- avec anomalie chromosomique précisée (Q96-Q99)
- féminin, avec hyperadrénocorticisme (E25.-)
- masculin, avec résistance aux androgènes (E34.5)

**Q56.0Hermaphrodisme, non classé ailleurs**

Ovotestis

**Q56.1Pseudo-hermaphrodisme masculin, non classé ailleurs**

Pseudo-hermaphrodisme masculin SAI

**Q56.2Pseudo-hermaphrodisme féminin, non classé ailleurs**

Pseudo-hermaphrodisme féminin SAI

**Q56.3Pseudo-hermaphrodisme, sans précision**

**Q56.4Sexe indéterminé, sans précision**

Ambiguïté des organes génitaux

---

**Malformations congénitales de l'appareil urinaire**  
**(Q60-Q64)**

**Q60Agénésie rénale et autres défauts de développement rénal**

*Inclus:*

absence congénitale de rein(s)

atrophie rénale:

- congénitale
- infantile

**Q60.0Agénésie rénale, unilatérale**

**Q60.1Agénésie rénale, bilatérale**

**Q60.2Agénésie rénale, sans précision**

**Q60.3**Hypoplasie rénale unilatérale  
**Q60.4**Hypoplasie rénale bilatérale  
**Q60.5**Hypoplasie rénale, sans précision  
**Q60.6**Syndrome de Potter

---

**Q61**Maladies kystiques du rein

*Excl.:*

kyste acquis du rein (**N28.1**)  
syndrome de Potter (**Q60.6**)

**Q61.0**Kyste congénital unique du rein

Kyste du rein (congénital) (unique)

**Q61.1**Rein polykystique, autosomique récessif

Rein polykystique, type infantile

**Q61.2**Rein polykystique, autosomique dominant

Rein polykystique, type adulte

**Q61.3**Rein polykystique, sans précision

**Q61.4**Dysplasie rénale

**Q61.5**Kyste médullaire du rein

Rein en éponge SAI

**Q61.8**Autres maladies kystiques du rein

Dégénérescence ou maladie fibrokystique du rein  
Rein fibrokystique

**Q61.9**Maladie kystique du rein, sans précision

Syndrome de Meckel-Gruber

---

**Q62**Anomalies congénitales obstructives du bassin et malformations congénitales de l'uretère

**Q62.0**Hydronéphrose congénitale

**Q62.1**Atrésie et sténose de l'uretère

Imperméabilité de l'uretère

Occlusion congénitale de:

- jonction pyélo-urétérale
- orifice vésico-urétéral
- uretère

**Q62.2**Méga-uretère congénital

Dilatation congénitale de l'uretère

**Q62.3**Autres anomalies obstructives du bassin et de l'uretère

Urétérocèle congénitale

**Q62.4**Agénésie de l'uretère

Absence d'uretère

**Q62.5**Duplication de l'uretère

Uretère:

- double
- surnuméraire

**Q62.6**Malposition de l'uretère

Déplacement

Déviat

Ectopie

Implantation anormale

de l'uretère ou de l'orifice urétéral

**Q62.7**Reflux vésico-urétéro-rénal congénital

**Q62.8**Autres malformations congénitales de l'uretère

Anomalie de l'uretère SAI

---

**Q63**Autres malformations congénitales du rein

*Excl.:*

syndrome néphrotique congénital (**N04.-**)

**Q63.0**Rein surnuméraire

**Q63.1**Rein lobulé, soudé et en fer à cheval

**Q63.2**Rein ectopique

Déplacement congénital du rein

Malrotation du rein

**Q63.3**Rein hyperplasique et géant

**Q63.8**Autres malformations congénitales précisées du rein

Lithiase rénale congénitale

**Q63.9**Malformation congénitale du rein, sans précision

---

**Q64**Autres malformations congénitales de l'appareil urinaire

**Q64.0**Epispadias

*Excl.:*

hypospadias (**Q54.-**)

**Q64.1**Exstrophie vésicale

Ectopie vésicale

Extroversion de la vessie

**Q64.2**Valvules urétrales postérieures congénitales

**Q64.3Autres formes d'atrésie et de sténose de l'urètre et du col de la vessie**

Imperméabilité de l'urètre  
 Occlusion du col de la vessie  
 Rétrécissement de:  
 méat urinaire  
 orifice vésico-urétral  
 urètre  
 congénital(e)

**Q64.4Malformation de l'ouraque**

Kyste  
 Persistance  
 Prolapsus  
 de l'ouraque

**Q64.5Absence congénitale de la vessie et de l'urètre****Q64.6Diverticule congénital de la vessie****Q64.7Autres malformations congénitales de la vessie et de l'urètre**

Fistule recto-urétrale congénitale  
 Hernie congénitale de la vessie  
 Malformation congénitale de la vessie ou de l'urètre SAI  
 Méat urinaire double  
 Prolapsus congénital de:  
 • méat urinaire  
 • urètre  
 • vessie (muqueuse)  
 Urètre double

Urètre  
 Vessie  
 surnuméraire

**Q64.8Autres malformations congénitales précisées de l'appareil urinaire****Q64.9Malformation congénitale de l'appareil urinaire, sans précision**

Anomalie congénitale SAI de l'appareil urinaire

## Malformations congénitales du système ostéo-articulaire et des muscles (Q65-Q79)

**Q65Anomalies morphologiques congénitales de la hanche**

*Excl.:*

hanche à ressort (R29.4)

**Q65.0Luxation congénitale de la hanche, unilatérale****Q65.1Luxation congénitale de la hanche, bilatérale****Q65.2Luxation congénitale de la hanche, sans précision****Q65.3Subluxation congénitale de la hanche, unilatérale****Q65.4Subluxation congénitale de la hanche, bilatérale****Q65.5Subluxation congénitale de la hanche, sans précision****Q65.6Hanche instable**

Hanche:

- luxable
- subluxable

**Q65.8Autres malformations congénitales de la hanche**

Antéversion du col du fémur

Coxa:  
 valga  
 vara  
 congénitale

Dysplasie acétabulaire congénitale

**Q65.9Malformation congénitale de la hanche, sans précision****Q66Anomalies morphologiques congénitales du pied**

*Excl.:*

déformations en:

- valgus, acquises (M21.0)
- varus, acquises (M21.1)

malformations du pied avec raccourcissement (Q72.-)

**Q66.0Pied bot varus équin****Q66.1Pied bot talus varus****Q66.2Metatarsus varus****Q66.3Autres anomalies morphologiques congénitales du pied en varus**

Hallux varus congénital

**Q66.4Pied bot talus valgus****Q66.5Pied plat congénital**

Pied plat spastique (valgus)

**Q66.6Autres anomalies morphologiques congénitales du pied en valgus**

Metatarsus valgus

**Q66.7Pied creux**

**Q66.8Autres anomalies morphologiques congénitales du pied**

Astragale vertical  
Coalescence tarsale  
Orteil en marteau, congénital  
Pied bot:

- SAI
- asymétrique

**Q66.9Anomalie morphologique congénitale du pied, sans précision**

---

**Q67Anomalies morphologiques congénitales de la tête, de la face, du rachis et du thorax****Excl.:**

syndrome(s) (de):

- congénitaux malformatifs classés en Q87.-
- Potter ([Q60.6](#))

**Q67.0Asymétrie faciale****Q67.1Anomalie morphologique faciale par compression****Q67.2Dolichocéphalie****Q67.3Plagiocéphalie****Q67.4Autres anomalies morphologiques congénitales du crâne, de la face et de la mâchoire**

Atrophie ou hypertrophie hémifaciale  
Dépressions des os du crâne  
Déviation congénitale de la cloison nasale  
Ecrasement ou déviation congénital(e) du nez

**Excl.:**

anomalies dento-faciales [y compris les malocclusions] ([K07.-](#))  
nez en selle dû à la syphilis ([A50.5](#))

**Q67.5Anomalie morphologique congénitale du rachis**

Scoliose congénitale:

- SAI
- posturale

**Excl.:**

scoliose:

- due à une malformation osseuse congénitale ([Q76.3](#))
- idiopathique infantile ([M41.0](#))

**Q67.6Thorax en entonnoir congénital****Q67.7Thorax en carène congénital**

Thorax en bréchet congénital

**Q67.8Autres anomalies morphologiques congénitales du thorax**

Anomalie morphologique congénitale de la paroi thoracique SAI

---

**Q68Autres anomalies morphologiques congénitales ostéo-articulaires et des muscles****Excl.:**

malformations de membre(s) avec raccourcissement ([Q71-Q73](#))

**Q68.0Anomalie morphologique congénitale des muscles sterno-cléido-mastoldiens**

Contracture du muscle sterno-cléido-mastoldien  
Torticolis congénital (sterno-mastoldien)  
Tuméfaction sterno-mastoldienne (congénitale)

**Q68.1Anomalie morphologique congénitale de la main**

Hippocratisme digital congénital  
Main en bêche (congénitale)

**Q68.2Anomalie morphologique congénitale du genou**

Genu recurvatum congénital  
Luxation congénitale du genou

**Q68.3Incurvation congénitale du fémur****Excl.:**

antéversion du fémur (col) ([Q65.8](#))

**Q68.4Incurvation congénitale du tibia et du péroné****Q68.5Incurvation congénitale des os longs du membre inférieur, sans précision****Q68.8Autres anomalies morphologiques congénitales ostéo-articulaires et des muscles précisées**

Anomalie morphologique congénitale de:

- avant-bras
- clavicule
- coude
- omoplate

Luxation congénitale de:

- coude
- épaule

---

**Q69Polydactylie****Q69.0Doigt(s) surnuméraire(s)****Q69.1Pouce(s) surnuméraire(s)****Q69.2Orteil(s) surnuméraire(s)**

Gros orteil surnuméraire  
**Q69.9Polydactylie, sans précision**

---

**Q70Syndactylie**

**Q70.0Coalescence des doigts**

Syndactylie complexe des doigts avec synostose

**Q70.1Palmure des doigts**

Simple syndactylie des doigts sans synostose

**Q70.2Coalescence des orteils**

Syndactylie complexe des orteils avec synostose

**Q70.3Palmure des orteils**

Simple syndactylie des orteils sans synostose

**Q70.4Polysyndactylie**

**Q70.9Syndactylie, sans précision**

Coalescence des phalanges SAI

---

**Q71Raccourcissement du membre supérieur**

**Q71.0Absence congénitale complète d'un (des) membre(s) supérieur(s)**

**Q71.1Absence congénitale du bras et de l'avant-bras, main présente**

**Q71.2Absence congénitale de l'avant-bras et de la main**

**Q71.3Absence congénitale de la main et de(s) doigt(s)**

**Q71.4Raccourcissement longitudinal du radius**

Main bote (radiale)

**Q71.5Raccourcissement longitudinal du cubitus**

**Q71.6Main en pince de homard**

**Q71.8Autres raccourcissements d'un (des) membre(s) supérieur(s)**

**Q71.9Raccourcissement du membre supérieur, sans précision**

---

**Q72Raccourcissement du membre inférieur**

**Q72.0Absence congénitale complète d'un (des) membre(s) inférieur(s)**

**Q72.1Absence congénitale de la cuisse et de la jambe, pied présent**

**Q72.2Absence congénitale de la jambe et du pied**

**Q72.3Absence congénitale du pied et d'un (des) orteil(s)**

**Q72.4Raccourcissement longitudinal du fémur**

Anomalie du noyau céphalique fémoral

**Q72.5Raccourcissement longitudinal du tibia**

**Q72.6Raccourcissement longitudinal du péroné**

**Q72.7Pied bifide**

**Q72.8Autres raccourcissements d'un (des) membre(s) inférieur(s)**

**Q72.9Raccourcissement du membre inférieur, sans précision**

---

**Q73Raccourcissement d'un membre non précisé**

**Q73.0Absence congénitale d'un (des) membre(s) non précisé(s)**

Amélie SAI

**Q73.1Phocomélie, membre(s) non précisé(s)**

Phocomélie SAI

**Q73.8Autres raccourcissements d'un (des) membre(s) non précisé(s)**

Ectromélie

Hémimélie

SAI d'un (des) membre(s) SAI

Raccourcissement

Raccourcissement longitudinal d'un (des) membre(s) non précisé(s)

---

**Q74Autres malformations congénitales d'un (des) membre(s)**

**Excl.:**

polydactylie (Q69.-)

raccourcissement d'un membre (Q71-Q73)

syndactylie (Q70.-)

**Q74.0Autres malformations congénitales d'un (des) membre(s) supérieur(s), y compris la ceinture scapulaire**

Carpocypnose [maladie de Madelung]

Dysostose cléido-crânienne

Macroductylie (doigts)

Malformation de Sprengel

Os du carpe surnuméraire

Pouce à trois phalanges

Pseudarthrose congénitale de la clavicule

Synostose radio-cubitale

**Q74.1Malformation congénitale du genou**

Absence congénitale de la rotule

Genu valgum congénital

Genu varum congénital

Luxation congénitale de la rotule

Rotule rudimentaire

**Excl.:**

genu recurvatum congénital (Q68.2)

luxation congénitale du genou ([Q68.2](#))

ostéo-onychodysostose ([Q87.2](#))

**Q74.2 Autres malformations congénitales d'un (des) membre(s) inférieur(s), y compris la ceinture pelvienne**

Fusion de l'articulation sacro-iliaque	
Malformation de:	
articulation sacro-iliaque	congénitale
cheville	

**Excl.:**

antéversion du fémur (col) ([Q65.8](#))

**Q74.3 Arthrogrypose congénitale multiple**

**Q74.8 Autres malformations congénitales précisées d'un (des) membre(s) non précisé(s)**

**Q74.9 Malformation congénitale d'un (des) membre(s) non précisé(s), sans précision**

Anomalie congénitale d'un (des) membre(s) SAI

---

**Q75 Autres malformations congénitales des os du crâne et de la face**

**Excl.:**

anomalies:

- dento-faciales [y compris les malocclusions] ([K07.-](#))
- morphologiques congénitales de la tête et de la face ([Q67.0-Q67.4](#))

malformation(s) (du):

- crâne associées à des anomalies du cerveau telles que:
  - anencéphalie ([Q00.0](#))
  - encéphalocèle ([Q01.-](#))
  - hydrocéphalie ([Q03.-](#))
  - microcéphalie ([Q02](#))
- congénitale de la face SAI ([Q18.-](#))

syndromes congénitaux malformatifs classés en Q87.-

**Q75.0 Craniosynostose**

Acrocéphalie

Anomalie de suture des os du crâne

Oxycéphalie

Trigonocéphalie

**Q75.1 Dysostose crânio-faciale**

Maladie de Crouzon

**Q75.2 Hypertélorisme**

**Q75.3 Macrocéphalie**

**Q75.4 Dysostose mandibulo-faciale**

Syndrome de:

- Franceschetti
- Treacher Collins

**Q75.5 Dysostose oculo-mandibulaire**

**Q75.8 Autres malformations congénitales précisées des os du crâne et de la face**

Absence congénitale d'os du crâne

Malformation congénitale du front

Platybasie

**Q75.9 Malformation congénitale des os du crâne et de la face, sans précision**

Anomalie de:

- crâne SAI
- os de la face SAI

---

**Q76 Malformations congénitales du rachis et du thorax osseux**

**Excl.:**

anomalies morphologiques congénitales ostéo-articulaires et des muscles du rachis et du thorax ([Q67.5-Q67.8](#))

**Q76.0 Spina bifida occulta**

**Excl.:**

méningocèle (rachidienne) ([Q05.-](#))

spina bifida (aperta) (kystique) ([Q05.-](#))

**Q76.1 Syndrome de Klippel-Feil**

Synostose cervicale

**Q76.2 Spondylolisthésis congénital**

Spondylolyse congénitale

**Excl.:**

spondylolisthésis (acquis) ([M43.1](#))

spondylolyse (acquise) ([M43.0](#))

**Q76.3 Scoliose congénitale due à une malformation osseuse congénitale**

Synostose d'hémivertèbres ou défaut de segmentation avec scoliose

**Q76.4 Autres malformations congénitales du rachis, non associées à une scoliose**

Absence congénitale de vertèbres		sans précision ou non associée à une
Cyphose congénitale	scoliose	

Hémivertèbre [hémispondylie]  
Lordose congénitale  
Malformation congénitale (charnière) de la région lombo-  
sacrée  
Malformation du rachis  
Platyspondylie  
Synostose vertébrale congénitale  
Vertèbre surnuméraire

**Q76.5Côte cervicale**

Côte surnuméraire cervicale

**Q76.6Autres malformations congénitales des côtes**

Absence de côtes  
Coalescence des côtes | congénitale

Côte surnuméraire

Malformation congénitale d'une côte SAI

*Excl.:*

syndrome des côtes courtes (Q77.2)

**Q76.7Anomalies congénitales du sternum**

Absence congénitale de sternum

Sternum bifide

**Q76.8Autres malformations congénitales du thorax osseux**

**Q76.9Malformation congénitale du thorax osseux, sans précision**

---

**Q77Ostéochondrodysplasie avec anomalies de la croissance des os longs et du rachis**

*Excl.:*

mucopolysaccharidoses (E76.0-E76.3)

**Q77.0Achondrogénésie**

Hypocondrogénésie

**Q77.1Insuffisance staturale thanatophore**

**Q77.2Syndrome des côtes courtes**

Dysplasie thoracique asphyxiante [Jeune]

**Q77.3Chondrodysplasie ponctuée**

**Q77.4Achondroplasie**

Hypocondroplasie

Ostéosclérose congénitale

**Q77.5Dysplasie diastrophique**

**Q77.6Dysplasie chondro-ectodermique**

Syndrome d'Ellis-van Creveld

**Q77.7Dysplasie spondylo-épiphytaire**

**Q77.8Autres ostéochondrodysplasies avec anomalies de la croissance des os longs et du rachis**

**Q77.9Ostéochondrodysplasie avec anomalies de la croissance des os longs et du rachis, sans précision**

---

**Q78Autres ostéochondrodysplasies**

**Q78.0Ostéogénèse imparfaite**

Fragilité osseuse

Ostéopsathyrose

**Q78.1Dysplasie polyostotique fibreuse**

Syndrome d'Albright(-McCune-Sternberg)

**Q78.2Ostéopétrose**

Syndrome d'Albers-Schönberg

**Q78.3Dysplasie diaphysaire progressive**

Syndrome de Camurati-Engelmann

**Q78.4Enchondromatose**

Maladie d'Ollier

Syndrome de Maffucci

**Q78.5Dysplasie métaphysaire**

Syndrome de Pyle

**Q78.6Exostoses multiples congénitales**

Aclasia diaphysaire

**Q78.8Autres ostéochondrodysplasies précisées**

Ostéopoikilose

**Q78.9Ostéochondrodysplasie, sans précision**

Chondrodystrophie SAI

Ostéodystrophie SAI

---

**Q79Malformations congénitales du système ostéo-articulaire et des muscles, non classées ailleurs**

*Excl.:*

torticollis congénital (sterno-mastoidien) (Q68.0)

**Q79.0Hernie diaphragmatique congénitale**

*Excl.:*

hernie hiatale congénitale (Q40.1)

**Q79.1Autres malformations congénitales du diaphragme**

Absence de diaphragme

Eventration du diaphragme  
Malformation congénitale du diaphragme SAI

#### **Q79.2 Exomphale**

Omphalocèle

*Excl.:*

hernie ombilicale (K42.-)

#### **Q79.3 Gastroschisis**

#### **Q79.4 Aplasie congénitale de la musculature abdominale**

Syndrome d'Obrinsky

#### **Q79.5 Autres malformations congénitales de la paroi abdominale**

*Excl.:*

hernie ombilicale (K42.-)

#### **Q79.6 Syndrome d'Ehlers-Danlos**

#### **Q79.8 Autres malformations congénitales du système ostéo-articulaire et des muscles**

Absence de:

- muscle
- tendon

Amyotrophie congénitale

Brides congénitales

Brièveté congénitale d'un tendon

Muscle surnuméraire

Syndrome de Poland

#### **Q79.9 Malformation congénitale du système ostéo-articulaire et des muscles, sans précision**

Anomalie congénitale SAI du système ostéo-musculaire

### **Autres malformations congénitales (Q80-Q89)**

#### **Q80 Ichtyose congénitale**

*Excl.:*

maladie de Refsum (G60.1)

#### **Q80.0 Ichtyose vulgaire**

#### **Q80.1 Ichtyose liée au chromosome X**

#### **Q80.2 Ichtyose lamellaire**

Bébé collodion

#### **Q80.3 Erythrodermie ichtyosiforme bulleuse congénitale**

#### **Q80.4 Fœtus Arlequin**

#### **Q80.8 Autres ichtyoses congénitales**

#### **Q80.9 Ichtyose congénitale, sans précision**

#### **Q81 Epidermolyse bulleuse**

#### **Q81.0 Epidermolyse bulleuse simple**

*Excl.:*

syndrome de Cockayne (Q87.1)

#### **Q81.1 Epidermolyse bulleuse létale**

Syndrome de Herlitz

#### **Q81.2 Epidermolyse bulleuse dystrophique**

#### **Q81.8 Autres épidermolyses bulleuses**

#### **Q81.9 Epidermolyse bulleuse, sans précision**

#### **Q82 Autres malformations congénitales de la peau**

*Excl.:*

acrodermatitis enteropathica (E83.2)

kyste dermoïde pararectal ou pilonidal (L05.-)

porphyrie érythropoïétique congénitale (E80.0)

syndrome de Sturge-Weber(-Dimitri) (Q85.8)

#### **Q82.0 Lymphœdème héréditaire**

#### **Q82.1 Xeroderma pigmentosum**

#### **Q82.2 Mastocytose**

Urticaire pigmentée

*Excl.:*

mastocytose maligne (C96.2)

#### **Q82.3 Incontinentia pigmenti**

#### **Q82.4 Dysplasie ectodermique (anhidrotique)**

*Excl.:*

syndrome d'Ellis-van Creveld (Q77.6)

#### **Q82.5 Nævus congénital non néoplasique**

Envie SAI

Nævus:

- en tache de vin
- flammeus
- fraise
- sanguin
- vasculaire SAI

- verruqueux

**Excl.:**

lentigo (L81.4)

nævus (à):

- SAI (D22.-)
- arachnéen (I78.1)
- mélanocytes (D22.-)
- pigmentaire (D22.-)
- stellaire (I78.1)

taches café au lait (L81.3)

#### **Q82.8Autres malformations congénitales précisées de la peau**

Acrochordons

Anomalies des dermatoglyphes

Cutis laxa (hyperelastica)

Kératose:

- folliculaire [Darier-White]
- palmoplantaire héréditaire

Pemphigus bénin familial [Hailey-Hailey]

Plis palmaires anormaux

**Excl.:**

syndrome d'Ehlers-Danlos (Q79.6)

#### **Q82.9Malformation congénitale de la peau, sans précision**

#### **Q83Malformations congénitales du sein**

**Excl.:**

absence du muscle pectoral (Q79.8)

#### **Q83.0Absence congénitale de sein avec absence de mamelon**

#### **Q83.1Sein surnuméraire**

#### **Q83.2Absence de mamelon**

#### **Q83.3Mamelon accessoire**

Mamelon surnuméraire

#### **Q83.8Autres malformations congénitales du sein**

Hypoplasie mammaire

#### **Q83.9Malformation congénitale du sein, sans précision**

#### **Q84Autres malformations congénitales de la peau et des phanères**

#### **Q84.0Alopécie congénitale**

Atrichie congénitale

#### **Q84.1Anomalies congénitales de la morphologie des cheveux, non classées ailleurs**

Aplasia moniliforme

Monilethrix

Pili annulati

**Excl.:**

maladie des cheveux en fil de fer de Menkes (E83.0)

#### **Q84.2Autres anomalies congénitales des cheveux**

Hypertrichose congénitale

Malformation congénitale des cheveux SAI

Persistance du lanugo

#### **Q84.3Anonychie**

**Excl.:**

ostéo-onychodysostose (Q87.2)

#### **Q84.4Leuconychie congénitale**

#### **Q84.5Hypertrophie des ongles**

Onychose congénitale

Pachyonychie

#### **Q84.6Autres malformations congénitales des ongles**

Hippocratisme unguéal		congénital(e)
Koïlonychie		

Malformation congénitale des ongles SAI

#### **Q84.8Autres malformations congénitales précisées de la peau et des phanères**

Aplasia ectodermique congénitale

#### **Q84.9Malformation congénitale de la peau et des phanères, sans précision**

Anomalie congénitale des téguments SAI

#### **Q85Phacomatoses, non classées ailleurs**

**Excl.:**

ataxie télangectasique [Louis-Bar] (G11.3)

dysautonomie familiale [Riley-Day] (G90.1)

#### **Q85.0Neurofibromatose (non maligne)**

Maladie de von Recklinghausen

#### **Q85.1Sclérose tubéreuse**

Epiloïa

Maladie de Bourneville

**Q85.8Autres phacomatoses, non classées ailleurs**

Syndrome de:

- Peutz-Jeghers
- Sturge-Weber(-Dimitri)
- von Hippel-Lindau

**Excl.:**

syndrome de Meckel-Gruber (Q61.9)

**Q85.9Phacomatose, sans précision**

Hamartome SAI

---

**Q86Syndromes congénitaux malformatifs dus à des causes exogènes connues, non classés ailleurs****Excl.:**effets non tératogènes de substances transmises par voie transplacentaire ou par le lait maternel (P04.-)  
hypothyroïdie liée à une carence en iode (E00-E02)**Q86.0Syndrome d'alcoolisme fœtal (dysmorphique)****Q86.1Syndrome fœtal dû à l'hydantoïne**

Syndrome de Meadow

**Q86.2Dysmorphisme dû à la warfarine****Q86.8Autres syndromes congénitaux malformatifs dus à des causes exogènes connues**

---

**Q87Autres syndromes congénitaux malformatifs précisés atteignant plusieurs systèmes****Q87.0Syndromes congénitaux malformatifs atteignant principalement l'aspect de la face**

Acrocéphalopolysyndactylie

Acrocéphalosyndactylie [Apert]

Cyclopie

Syndrome (de):

- cryptophtalmique
- Goldenhar
- Mœbius
- oro-facio-digital
- Pierre Robin

Visage de siffleur

**Q87.1Syndromes congénitaux malformatifs associés principalement à une petite taille**

Syndrome de:

- Aarskog
- Cockayne
- De Lange
- Dubowitz
- Noonan
- Prader-Willi
- Robinow-Silverman-Smith
- Russell-Silver
- Seckel
- Smith-Lemli-Opitz

**Excl.:**

syndrome d'Ellis-van Creveld (Q77.6)

**Q87.2Syndromes congénitaux malformatifs impliquant principalement les membres**

Ostéo-onychodysostose

Syndrome de:

- Holt-Oram
- Klippel-Trénaunay-Weber
- onycho-arthro-ostéodysplasie
- Rubinstein-Taybi
- sirénomélie
- thrombopénie avec absence de radius
- VATER

**Q87.3Syndromes congénitaux malformatifs comprenant un gigantisme néonatal**

Syndrome de:

- Beckwith-Wiedemann
- Sotos
- Weaver

**Q87.4Syndrome de Marfan****Q87.5Autres syndromes congénitaux malformatifs avec d'autres modifications du squelette****Q87.8Autres syndromes congénitaux malformatifs précisés, non classés ailleurs**

Syndrome de:

- Alport
- Laurence-Moon-Bardet-Biedl
- Zellweger

---

**Q89**Autres malformations congénitales, non classées ailleurs**Q89.0**Malformations congénitales de la rate

Asplénie (congénitale)  
Splénomégalie congénitale

**Excl.:**

isomérisme des auricules cardiaques (avec asplénie ou polysplénie) ([Q20.6](#))

**Q89.1**Malformations congénitales des surrénales**Excl.:**

hyperplasie surrénale congénitale ([E25.0](#))

**Q89.2**Malformations congénitales des autres glandes endocrines

Kyste thyroglosse  
Malformation congénitale de la glande parathyroïde ou thyroïde  
Persistance du canal thyroglosse

**Q89.3**Situs inversus

Dextrocardie avec situs inversus  
Disposition auriculaire en miroir avec situs inversus  
Situs inversus ou transversus:

- abdominal
- thoracique

Transposition des viscères:

- abdominale
- thoracique

**Excl.:**

dextrocardie SAI ([Q24.0](#))

lévocardie ([Q24.1](#))

**Q89.4**Jumeaux soudés

Craniopage  
Dicéphale  
Monstre double  
Pygopage  
Thoracopage

**Q89.7**Malformations congénitales multiples, non classées ailleurs

Anomalies multiples congénitales SAI  
Monstre SAI

**Excl.:**

syndromes congénitaux malformatifs atteignant plusieurs systèmes ([Q87.-](#))

**Q89.8**Autres malformations congénitales précisées**Q89.9**Malformation congénitale, sans précision

Anomalie congénitale SAI

---

**Anomalies chromosomiques, non classées ailleurs****(Q90-Q99)****Q90**Syndrome de Down**Q90.0**Trisomie 21, non-disjonction méiotique**Q90.1**Trisomie 21, mosaïque chromosomique (non-disjonction mitotique)**Q90.2**Trisomie 21, translocation**Q90.9**Syndrome de Down, sans précision

Trisomie 21 SAI

---

**Q91**Syndrome d'Edwards et syndrome de Patau**Q91.0**Trisomie 18, non-disjonction méiotique**Q91.1**Trisomie 18, mosaïque chromosomique (non-disjonction mitotique)**Q91.2**Trisomie 18, translocation**Q91.3**Syndrome d'Edwards, sans précision**Q91.4**Trisomie 13, non-disjonction méiotique**Q91.5**Trisomie 13, mosaïque chromosomique (non-disjonction mitotique)**Q91.6**Trisomie 13, translocation**Q91.7**Syndrome de Patau, sans précision

---

**Q92**Autres trisomies et trisomies partielles des autosomes, non classées ailleurs**Inclus:**

translocations et insertions non équilibrées

**Excl.:**

trisomies des chromosomes 13, 18 et 21 ([Q90-Q91](#))

**Q92.0**Trisomie d'un chromosome entier, non-disjonction méiotique**Q92.1**Trisomie d'un chromosome entier, mosaïque chromosomique (non-disjonction mitotique)**Q92.2**Trisomie partielle majeure

Duplication complète d'au moins un bras du chromosome.

**Q92.3**Trisomie partielle mineure

Duplication de moins qu'un bras du chromosome.

**Q92.4**Duplications vues seulement à la prométaphase**Q92.5**Duplications avec autres réarrangements complexes**Q92.6**Chromosomes marqueurs supplémentaires**Q92.7**Tripléidie et polyploïdie

Q92.8Autres trisomies et trisomies partielles précisées des autosomes  
Q92.9Trisomie et trisomie partielle des autosomes, sans précision

---

**Q93**Monosomies et délétions des autosomes, non classées ailleurs  
Q93.0Monosomie d'un chromosome entier, non-disjonction méiotique  
Q93.1Monosomie d'un chromosome entier, mosaïque chromosomique (non-disjonction mitotique)  
Q93.2Chromosome remplacé par un anneau ou dicentrique  
Q93.3Délétion du bras court du chromosome 4  
    Syndrome de Wolff-Hirschorn  
Q93.4Délétion du bras court du chromosome 5  
    Syndrome du cri du chat  
Q93.5Autres délétions partielles d'un chromosome  
    Syndrome d'Angelman  
Q93.6Délétions vues seulement à la prométaphase  
Q93.7Délétions avec autres réarrangements complexes  
Q93.8Autres délétions des autosomes  
Q93.9Délétion des autosomes, sans précision

---

**Q95**Réarrangements équilibrés et marqueurs structuraux, non classés ailleurs

*Inclus:*

translocations et insertions réciproques robertsoniennes et équilibrées  
Q95.0Translocation ou insertion équilibrée chez un sujet normal  
Q95.1Inversion chromosomique chez un sujet normal  
Q95.2Réarrangement autosomique équilibré chez un sujet anormal  
Q95.3Réarrangement sexuel/autosomique équilibré chez un sujet anormal  
Q95.4Sujet avec marqueur d'hétérochromatine  
Q95.5Sujet avec site autosomique fragile  
Q95.8Autres réarrangements équilibrés et marqueurs structurels  
Q95.9Réarrangement équilibré et marqueur structurel, sans précision

---

**Q96**Syndrome de Turner

*Excl.:*

syndrome de Noonan (Q87.1)  
Q96.0Caryotype 45,X  
Q96.1Caryotype 46,X iso (Xq)  
Q96.2Caryotype 46,X avec chromosome sexuel anormal, sauf iso (Xq)  
Q96.3Mosaïque chromosomique, 45,X/46,XX ou XY  
Q96.4Mosaïque chromosomique, 45,X/autre(s) lignée(s) cellulaire(s) avec chromosome sexuel anormal  
Q96.8Autres formes du syndrome de Turner  
Q96.9Syndrome de Turner, sans précision

---

**Q97**Autres anomalies des chromosomes sexuels, phénotype féminin, non classées ailleurs

*Excl.:*

syndrome de Turner (Q96.-)  
Q97.0Caryotype 47,XXX  
Q97.1Femme avec plus de trois chromosomes X  
Q97.2Mosaïque chromosomique, lignées avec divers nombres de chromosomes X  
Q97.3Femme avec un caryotype 46,XY  
Q97.8Autres anomalies précisées des chromosomes sexuels, phénotype féminin  
Q97.9Anomalie des chromosomes sexuels, phénotype féminin, sans précision

---

**Q98**Autres anomalies des chromosomes sexuels, phénotype masculin, non classées ailleurs

Q98.0Syndrome de Klinefelter, caryotype 47,XXY  
Q98.1Syndrome de Klinefelter, homme avec plus de deux chromosomes X  
Q98.2Syndrome de Klinefelter, homme avec caryotype 46,XX  
Q98.3Autre phénotype masculin avec caryotype 46,XX  
Q98.4Syndrome de Klinefelter, sans précision  
Q98.5Caryotype 47,XYY  
Q98.6Homme avec chromosomes sexuels de structure anormale  
Q98.7Homme avec mosaïque des chromosomes sexuels  
Q98.8Autres anomalies précisées des chromosomes sexuels, phénotype masculin  
Q98.9Anomalie des chromosomes sexuels, phénotype masculin, sans précision

---

**Q99**Autres anomalies des chromosomes, non classées ailleurs

Q99.0Chimère 46,XX/46,XY  
    Chimère 46,XX/46,XY hermaphrodite vrai  
Q99.1Hermaphrodite vrai 46,XX  
    46,XX avec ovaire en bandelettes  
    46,XY avec ovaire en bandelettes  
    Dysgénése gonadique pure  
Q99.2Chromosome X fragile  
    Syndrome de l'X fragile  
Q99.8Autres anomalies précisées des chromosomes  
Q99.9Anomalie chromosomique, sans précision







# Annexe 2

## Fiche hernies scrotale et/ou inguinale

### Définition

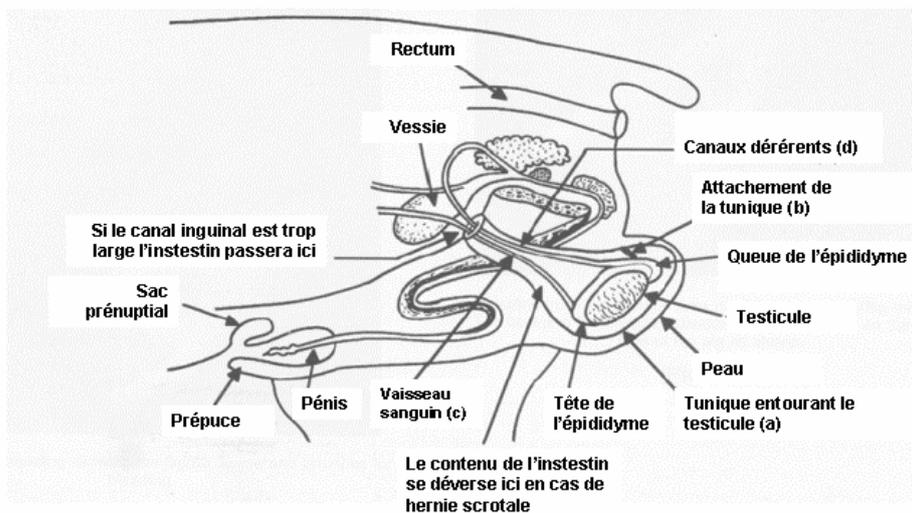
Les hernies scrotale et inguinale sont des renflements souples au niveau de l'aîne ou du scrotum. Elles correspondent à un passage de l'intestin grêle dans le scrotum (hernie scrotale) ou dans le canal inguinal (hernie inguinale). Elles peuvent ne pas être visibles à la naissance et se développer quelques semaines plus tard.

C'est une anomalie de la paroi abdominale et non du tube digestif en lui-même : on pense que les hernies sont dues à une faiblesse musculaire ou à un canal inguinal anormalement large.

Le canal inguinal est une cavité située dans l'aîne entre les muscles de l'abdomen au travers de laquelle passe, chez la femelle, le ligament qui soutient l'utérus et chez le mâle, le cordon spermatique. Le cordon spermatique est constitué des canaux déférents (petits tubes qui sortent des testicules), de nerfs et de vaisseaux sanguins ; il va du testicule à l'abdomen.

La hernie scrotale est évidemment limitée au sexe mâle tandis que la hernie inguinale peut affecter les deux sexes.

Les hernies scrotales peuvent être unilatérales (elles apparaissent le plus souvent du côté gauche) ou bilatérales.



### Description

Distinction des différents cas :

- Hernie scrotale/inguinale unilatérale
  - o gauche

- droite
- Hernie scrotale/inguinale bilatérale

## Photos

Hernies scrotale gauche



Hernie inguinale gauche



## Prélèvements

- Tubes : tubes sous vide Terumo Venosafe 6 ml EDTA K2, référence VF076SDK (fournisseur, CML [www.cml.fr](http://www.cml.fr))
- Quantité de sang : 5 ml de sang par animal : attention, ne pas trop remplir le tube (laisser 1 ml de vide), risque d'éclatement lors de la congélation.
- Portées faisant l'objet de prélèvements : portées avec au moins 2 animaux atteints à condition de pouvoir prélever :
  - Les animaux atteints de la portée (2 minimum)
  - Même nombre d'animaux sains que d'animaux atteints
  - Mère de la portée
  - Père de la portée : sauf si son ADN est déjà stocké
- Les prélèvements peuvent être conservés quelques jours (5 jours) au réfrigérateur avant d'être expédiés. Ils ne doivent en aucun cas être congelés.
- Envoyer les échantillons à Nathalie Iannuccelli ; Laboratoire de Génétique Cellulaire

INRA ; BP 52627 ; 31326 Castanet-Tolosan (envoi de type Chronopost)

## Fichier

- Renseigner le fichier « saisie anomalies »
  - Nom de l'anomalie : Hernie scrotale ou bien Hernie inguinale
  - Description de l'anomalie :
    - ✓ Si unilatérale : indiquer également si gauche ou si droite. Il n'ai pas aisé de distinguer une hernie scrotale d'une hernie inguinale chez un mâle, en cas de doute, préciser hernie scrotale/inguinale.
    - ✓ Sinon : indiquer bilatérale
- Envoyer le fichier à [nathalie.iannuccelli@toulouse.inra.fr](mailto:nathalie.iannuccelli@toulouse.inra.fr) et copie à [marie-jose.mercat@itp.asso.fr](mailto:marie-jose.mercat@itp.asso.fr)
- Pour toute question : MJ Mercat : 02.99.60.98.51

## Fiche cryptorchidie

### Définition

La cryptorchidie ou ectopie testiculaire est une anomalie du développement de l'appareil génital mâle qui se caractérise par l'absence de descente ou la migration incomplète d'un ou des deux testicules dans le scrotum. L'ectopie peut être unilatérale (mono-cryptorchidie) ou bilatérale. On parle d'animaux « PIF ».

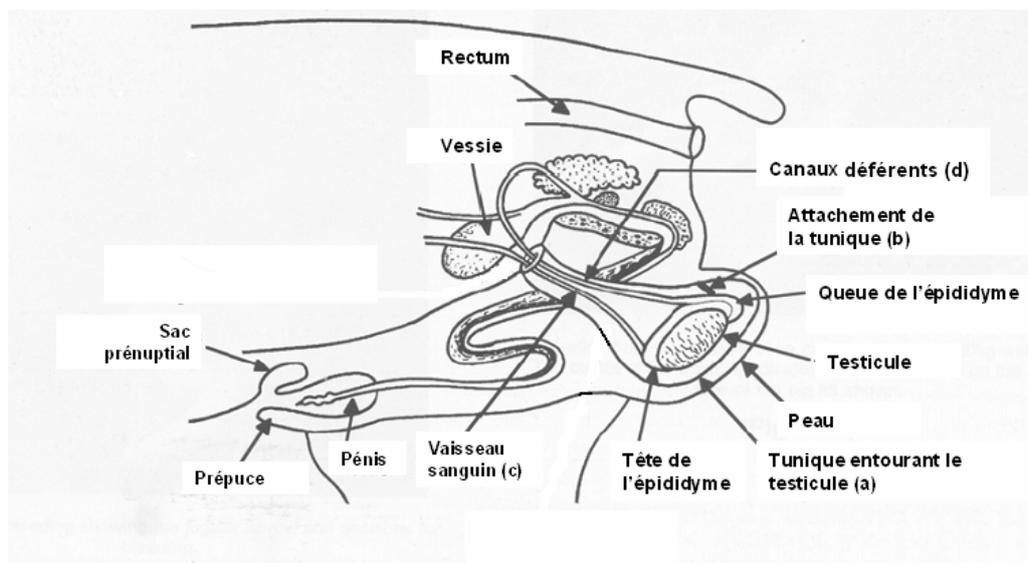
Par une simple palpation des bourses, il est facile de se rendre compte de l'existence de cette anomalie:

- Si un seul testicule est palpable, on parle de **cryptorchidie unilatérale** et le verrat est monorchide.
- Si les deux testicules sont impalpables, on parle de **cryptorchidie bilatérale**.

On peut mettre en évidence une cryptorchidie très précocement.

Selon que le testicule est resté dans l'abdomen ou qu'il est engagé dans le canal inguinal, on distingue des cryptorchidies abdominales et des cryptorchidies inguinales. Chez le Porc, l'ectopie est le plus souvent abdominale qu'inguinale, située à gauche qu'à droite.

Schéma :



### Description

Distinction des différents cas :

- Monorchidie
  - o Gauche
  - o Droite
- Cryptorchidie bilatérale

## Prélèvements

- Tubes : tubes sous vide Terumo Venosafe 6 ml EDTA K2, référence VF076SDK (fournisseur, CML [www.cml.fr](http://www.cml.fr))
- Quantité de sang : 5 ml de sang par animal : attention, ne pas trop remplir le tube (laisser 1 ml de vide), risque d'éclatement lors de la congélation.
- Portées faisant l'objet de prélèvements : portées avec au moins 1 animal atteint à condition de pouvoir prélever :
  - o Les animaux atteints de la portée (1 minimum)
  - o Au moins deux d'animaux mâles sains de la portée (plus si plus de 2 animaux atteints prélevés)
  - o Mère de la portée
  - o Père de la portée : sauf si son ADN est déjà stocké
- Les prélèvements peuvent être conservés quelques jours (5 jours) au réfrigérateur avant d'être expédiés. Ils ne doivent en aucun cas être congelés.
- Envoyer les échantillons à Nathalie Iannuccelli ; Laboratoire de Génétique Cellulaire INRA ; BP 52627 ; 31326 Castanet-Tolosan (envoi du type Chronopost)

## Fichier

- Renseigner le fichier « saisie anomalies » :
  - Nom de l'anomalie : Cryptorchidie
  - Description de l'anomalie :
    - ✓ Mono-cryptorchidie: indiquer également gauche ou droite
    - ✓ Sinon : indiquer bilatérale
- Envoyer le fichier à [nathalie.iannuccelli@toulouse.inra.fr](mailto:nathalie.iannuccelli@toulouse.inra.fr) et copie à [marie-jose.mercat@itp.asso.fr](mailto:marie-jose.mercat@itp.asso.fr)
- Pour toute question : MJ Mercat : 02.99.60.98.51

## Fiche intersexualité

### Définition

Certains animaux sont dits intersexués car ils présentent en même temps des caractéristiques mâles et femelles.

L'hermaphrodisme est un cas particulier d'intersexualité. Ce terme devrait être réservé aux animaux chez lesquels on retrouve à la fois des tissus testiculaires et ovariens isolés (testicule + ovaire) ou réunis dans une gonade indifférenciée (ovotestis). Le développement du tractus génital est habituellement typiquement femelle. Le plus souvent les vrais hermaphrodites ont une apparence externe de type femelle.

Les pseudo-hermaphrodites ont des gonades d'un seul sexe (testicules ou ovaires) et des voies génitales ou organes génitaux des deux sexes ou du sexe opposé. On parle de pseudo-hermaphrodisme femelle si les gonades sont des ovaires et de pseudo-hermaphrodisme mâle quand les gonades sont des testicules. Les pseudo-hermaphrodites femelles sont rares chez le porc. Les pseudo-hermaphrodites mâles ont des testicules qui sont généralement retenus dans la cavité abdominale. Ils ont généralement un utérus, un cervix, un vagin. Un développement certain des glandes mâles accessoires de l'épididyme et des canaux déférents peut aussi être observé.

### Description

Le simple examen clinique d'un animal intersexué ne permet pas de distinguer entre tous les cas d'intersexualité car seules les caractéristiques extérieures sont observables. L'analyse histologique des gonades est souvent le seul moyen précis de diagnostic.

Chaque partie extérieure de l'appareil génital doit être décrite :

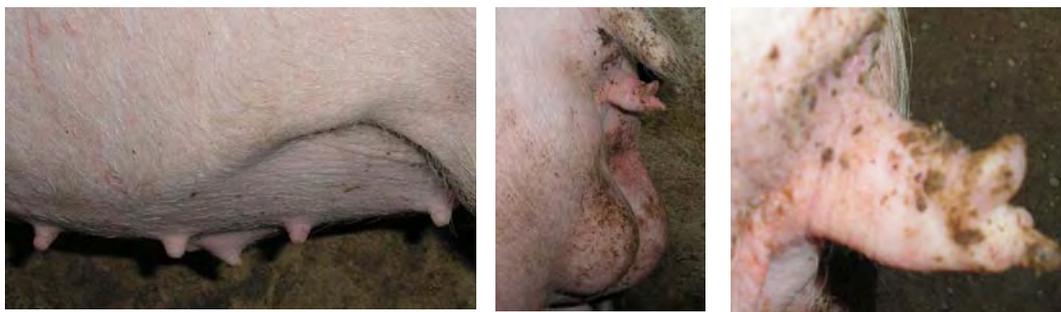
- Vulve
- Clitoris
- Testicules
- Fourreau – verge

Ce sont généralement des anomalies congénitales, donc théoriquement visibles à la naissance. Mais plus l'âge avance, plus l'observation (la description) est facilitée.

## Photos

		
Vulve atrophiée – Clitoris : non visible - Testicules normaux – Fourreau et verge non visibles sur la photo	Vulve atrophiée – Clitoris : non visible - Testicules : petite taille – Fourreau et verge non visibles sur la photo	Vulve absente – Clitoris péniforme- Testicules : grande taille – Fourreau et verge non visibles sur la photo

Les photographies ci-après ont été prises sur le même animal : on observe un «prémisse» de fourreau, un clitoris péniforme, des testicules déséquilibrés (le droit plus volumineux que le gauche), la vulve est absente



## Prélèvements

- Tubes : tubes sous vide Terumo Venosafe 6 ml EDTA K2, référence VF076SDK (fournisseur, CML [www.cml.fr](http://www.cml.fr))
- Quantité de sang : 5 ml de sang par animal : attention, ne pas trop remplir le tube (laisser 1 ml de vide), risque d'éclatement lors de la congélation.
- Portées faisant l'objet de prélèvements : portées avec au moins 1 animal atteint à condition de pouvoir prélever :
  - o Les animaux atteints de la portée (1 minimum)
  - o Au moins deux animaux sains de la portée (plus si plus de 2 animaux atteints prélevés)
  - o Mère de la portée
  - o Père de la portée : sauf si son ADN est déjà stocké
- Les prélèvements peuvent être conservés quelques jours (5 jours) au réfrigérateur avant d'être expédiés. Ils ne doivent en aucun cas être congelés.
- Envoyer les échantillons à Nathalie Iannuccelli ; Laboratoire de Génétique Cellulaire INRA ; BP 52627 ; 31326 Castanet-Tolosan (envoi du type Chronopost

## Fichier

- Renseigner le fichier « saisie anomalies »
  - Nom de l'anomalie : Intersexualité
  - Description de l'anomalie :
    - ✓ Vulve : Absence / Présence : normale – atrophiée – hypertrophiée
    - ✓ Clitoris : Normal/hyper développé etc.
    - ✓ Testicules : Absence/Présence (1 – 2) – information sur la taille
    - ✓ Fourreau - Verge : Absence/Présence – information sur la taille, la position sur le corps
    - ✓ Autre : exemple Présence d'un petit « pénis » au niveau de la vulve (clitoris péniforme) etc.
  
- Envoyer le fichier à [nathalie.iannuccelli@toulouse.inra.fr](mailto:nathalie.iannuccelli@toulouse.inra.fr) et copie à [marie-jose.mercat@itp.asso.fr](mailto:marie-jose.mercat@itp.asso.fr)
- Pour toute question : MJ Mercat : 02.99.60.98.51



Les anomalies congénitales sont définies comme des défauts de structure ou de fonction présents à la naissance. Chez le porc, les anomalies les plus fréquentes sont de type urogénital (hernie inguino-scrotale congénitale, cryptorchidie et intersexualité) et représentent un problème majeur pour l'industrie porcine (pertes économiques, atteinte au bien être des animaux). Les études ayant montré qu'une part de leur déterminisme était génétique, l'identification des gènes sous-jacents est devenue une priorité pour la filière porcine. Les outils récents de génomique offrent la possibilité de rechercher, à l'échelle du génome, des régions chromosomiques associées à un caractère.

Des analyses TDT (*Transmission Disequilibrium Test*) ont été réalisées à partir d'un dispositif de 200 familles nucléaires de descendants atteints, et ont permis d'identifier deux régions significativement associées aux maladies : une région localisée en extrémité du chromosome X, significative pour les trois maladies étudiées, et une région sur le chromosome 12 associée à l'intersexualité. D'autres régions suggestives ont également été mises en évidence pour la hernie et la cryptorchidie.

Des travaux de cartographie physique et génétique sur le chromosome X ont montré que le signal significatif était localisé sur la région pseudo-autosomale porcine et qu'il correspondait à un biais de l'analyse, dû au mode de ségrégation très particulier de la région. L'absence de recombinaison a également été observée pour un tiers du chromosome X, riche en motifs répétés.

La région du chromosome 12 associée à l'intersexualité est une excellente région candidate car un seul gène y est annoté et correspond à un des acteurs majeurs de la différenciation sexuelle chez le mâle. La cartographie fine de la région n'a pas permis la détection de mutation causale ou de CNV (*Copie Number Variation*) mais son implication dans le déterminisme de l'intersexualité a été confirmée.

Les loci identifiés pour les trois principales anomalies congénitales permettront de prévenir leur apparition dans les élevages porcins par sélection assistée par marqueurs. De plus, ces résultats apportent des données nouvelles en génétique humaine.

Mots clés: génétique, anomalies congénitales, porc

Congenital anomalies are defined as physical or functional defects which are present at birth. The most defects occurring in piglets are urogenital (congenital inguino-scrotal hernia, cryptorchidism, intersexuality) and have a serious impact on the pig industry (severe economic loss and poor animal welfare). Studies have shown that a part of determinism is genetic, thus the identification of underlying genes is a priority for the pig industry. Recently the developed genomics tools offer possibilities to underlie pan-genomics chromosomal regions associated with traits.

TDT analyses (*Transmission Disequilibrium Test*) performed on 200 nuclear families with affected offspring allowed to identify two significant regions associated with diseases: one on chromosome X for the three diseases and the other on chromosome 12 involved in intersexuality. Other suggestive regions are highlighted for hernia and cryptorchidism.

Genetic and physical mapping of chromosome X subsequently showed that the significant signal was located on the porcine pseudo-autosomal region. Further analyses showed it was an analysis bias, due to the specific segregation of the region. Absence of recombination was observed for one third of chromosome X, enriched in repeated motifs.

The region on chromosome 12 involved in intersexuality is an excellent candidate region because the only annotated gene in is one of the key genes for sexual differentiation in males. Fine mapping did not allow detection of a causative mutation or CNV (*Copies Number Variation*) but its implication was confirmed.

For the three major defects, the loci identified can be used in pig breeding for effective marker assisted selection against diseases. Moreover, these results provide new data for human genetics.

Keywords : genetics, congenital anomalies, pig