

# Université Paris Descartes

Ecole doctorale Bio Sorbonne Paris CitéBioSPC ED157

## **Relation entre la méthylation des promoteurs du gène *IGF1* et les variations de la croissance des enfants**

Par Meriem Ouni

Thèse de doctorat de Génétique

*Equipe de recherche "Thérapie Génique, Génétique, Epigénétique en neurologie,  
Endocrinologie et Développement de l'enfant" - UMR1169  
Hôpital de Bicêtre, 80 rue du G<sup>al</sup> Leclerc, 94270 Le Kremlin-Bicêtre, France*

Présentée et soutenue publiquement le 2 Juillet 2015

Devant un jury composé de :

Pr Marc DELPECH, PU-PH, Université Paris Descartes (Paris 5), Président

Pr Irène NETCHINE, PU-PH, Université Pierre et Marie Curie (Paris 6), rapporteur

Pr Pierre CHATELAIN, PU-PH, Université Claude Bernard (Lyon 1), rapporteur

Pr Michel VIDAUD, PU-PH, Université Paris Descartes (Paris 5), examinateur

Pr Agnès LINGLART, PU-PH, Université Paris Sud (Paris 11), examinateur

Pr Pierre Bougnères, PU-PH, Université Paris Sud (Paris 11), directeur de thèse

## Résumé

A l'interface de la génétique et de l'environnement, l'épigénétique contribue à la diversité phénotypique. Déterminer l'impact de la variation épigénétique sur les caractères quantitatifs (QT) est un nouveau défi. La croissance staturale fournit l'opportunité d'étudier la variabilité de plusieurs traits phénotypiques liés entre eux : des QT cliniques (la taille, l'accélération de la vitesse de croissance en réponse à l'hormone de croissance, GH) et des QT biologiques tels que la concentration d'IGF1 et la réponse de cette concentration à la GH.

L'« Insulin-like Growth Factor 1 » (*IGF1*) contrôle la croissance postnatale chez les mammifères, y compris l'homme. Nous l'avons choisi comme locus candidat pour nos études épigénétiques. Nous avons quantifié la méthylation des deux promoteurs P1 et P2 de ce gène, qui régulent son expression. Notre objectif était d'évaluer la contribution de la méthylation d'ADN de ces promoteurs i) à la taille des enfants en croissance, ii) à l'IGF1 circulant, iii) et à la réponse de ces paramètres à un traitement par la GH.

**Taille et IGF1 circulant.** La relation entre la méthylation des promoteurs d'*IGF1* et la taille a été étudiée au sein de deux cohortes du service d'endocrinologie pédiatrique, totalisant 216 enfants prépubères de différentes statures. Nous avons montré que la méthylation d'un groupe de six CGs situés dans la partie proximale du promoteur P2 du gène *IGF1* présentait une corrélation inverse avec la croissance et l'IGF1 circulant. Les enfants les plus grands sont ainsi moins méthylés sur ces CGs que les enfants de petite taille. La contribution de la méthylation à la variance de la taille a été évaluée à environ 13%, et à 10% pour la variance de l'IGF1 sérique.

Pour montrer que l'association observée reflète une causalité biologique, nous avons étudié le lien entre la méthylation des promoteurs P1 et P2 et l'activité transcriptionnelle du gène *IGF1* *in vivo* et *in vitro*. Nous avons montré que les quantités de transcrits de classe II, issus du promoteur P2, sont inversement corrélés à la méthylation du promoteur P2 dans les cellules sanguines mononucléées. *In vitro*, nous avons cloné le promoteur P2 déméthylé ou méthylé dans un plasmide rapporteur (luciférase) transfecté dans la lignée HEK293 : le promoteur déméthylé s'est révélé nettement plus actif (+57%).

Finalement, nous suggérons que l'hyperméthylation de certains CGs du P1 et du P2 d'*IGF1* pourrait être un des nombreux mécanismes moléculaires responsables d'une moindre expression du gène et d'un phénotype de petite taille.

**La réponse au traitement par la GH.** Une fraction des enfants de petite taille est traitée par l'hormone de croissance (GH) pour accélérer sa croissance, mais l'efficacité du traitement est très variable entre les individus. Les causes de cette variabilité sont partiellement comprises : la génétique joue un rôle, mais il reste une place possible pour la variabilité épigénétique. Dans ce but, nous avons étudié l'effet direct de la variabilité épigénétique sur la transcription du gène *IGF1* et l'IGF1 circulant, dans un test aigu d'administration de GH, puis sur la réponse thérapeutique à un traitement d'un an par la GH.

Après une injection de GH, nous avons constaté une augmentation variable du nombre de transcrits d'IGF1 chez les enfants étudiés. L'augmentation des transcrits de la classe II était inversement corrélée à la méthylation des CGs du P2. La variabilité de méthylation au CG-137 contribuait pour 20% à 67% de l'expression d'*IGF1* en réponse à la GH.

Chez 136 enfants de petite taille, nous avons montré que la méthylation de l'ADN du promoteur P2 était associée à la réponse au traitement par la GH au cours de la première année. Cette association est observée pour l'augmentation de la vitesse de croissance et pour les taux d'IGF1. Le niveau de méthylation du CG-137 du promoteur P2 contribuerait à lui seul à 25% de la variance de la réponse de la croissance à la GH. Le lien entre la méthylation de l'ADN et la réponse à un traitement illustre ici le rôle de certaines marques épigénétiques comme biomarqueurs de prédiction thérapeutique « pharmacoépigénétique ».

Notre travail peut trouver d'autres applications : i) dans la physiologie IGF1-dépendante de la croissance et dans la physiologie individuelle de l'hormone de croissance, ii) dans des études sur le vieillissement, la longévité ou le cancer où la production d'IGF1 joue un rôle primordial.

## Remerciements

*Tout d'abord, je remercie vivement mon directeur de thèse, Pierre Bougnères, pour m'avoir donné l'opportunité de travailler au sein de son équipe et de son laboratoire, pour m'avoir confié un sujet qu'il avait à cœur, m'avoir encadrée au quotidien, merci pour votre disponibilité et votre présence régulière.*

*Je le remercie particulièrement pour son soutien et son aide scientifique et humain tout au long de mon parcours durant ces quatre années de thèse ainsi que pour ses critiques toujours très instructives. Je le remercie de m'avoir appris l'exploitation des données cliniques, sans oublier tous ses efforts de corrections des données cliniques. Je le remercie enfin de m'avoir facilité l'entrée dans le monde des cliniciens pour une meilleure réalisation de ce travail.*

*Je remercie les Professeurs Irène Netchine et Pierre Châtelain pour le temps qu'ils ont bien voulu m'accorder en étant rapporteurs de cette thèse. Je remercie également mes examinateurs, les Professeurs Agnès Linglart, Marc Delpech et Michel Vidaud pour l'intérêt qu'ils ont porté à mon travail en acceptant de faire partie de mon jury.*

*Un grand merci à Delphine Fradin, chercheuse dans notre équipe, de m'avoir formée et aidée dans mes réflexions sur l'épigénétique, et de m'avoir soutenue scientifiquement et moralement durant ma thèse. Merci pour ton encouragement, pour tes conseils pertinents, pour tes corrections et ta disponibilité. Merci pour ton sourire et ta gentillesse.*

*Je remercie également Marie-Pierre Belot, post-doctorante, de m'avoir formée et soutenue pendant toute ma thèse. Merci également pour ta collaboration, ton investissement et ton apport scientifique dans ce projet. Merci pour tes corrections de mes fautes de français dans mon document de thèse, et même à l'oral. Merci pour ta bonne humeur pendant ces quatre ans, pour les moments de joie et pour nos rires partagés dans notre bureau.*

*Les filles, Delphine et Marie-Pierre, je dois à Augustin, Côme, Morgane et Solène, des excuses de les avoir privés de leurs mamans pendant des heures pour compter avec moi des cellules, ou corriger ce manuscrit.*

*Un merci tout particulier à Anne-laure Castell, Chef de clinique au service d'endocrinologie pédiatrique. Merci pour ta précieuse collaboration et tes conseils cliniques. Merci pour ta belle compagnie. Je n'oublierai jamais nos discussions, nos rires et notre petit goûter de 18h.*

*Je remercie également Yasemin Gunes pour sa participation à mon projet de thèse, à un moment où le temps qu'il me restait ne me permettait pas de tout faire...*

*Merci Sophie, « ma camarade du soir », pour ton grand soutien et ta compagnie précieuse, pour tes conseils de gestion de ma base de données cliniques, et pour ton aide pour le document de ma thèse. Merci pour ton encouragement pendant les moments difficiles que j'ai vécus.*

*Je remercie toutes les personnes avec lesquelles j'ai pu collaborer dans le service d'endocrinologie pédiatrique, notamment les infirmières qui ont réalisé pour moi tant de prélèvements avec le sourire. Merci Cathy, Florie, les Aurélie, Imène, Anne-Laure... J'espère que je n'ai oublié personne.*

*Je remercie tous mes collègues de l'unité INSERM U1169 (Ex U986). Merci à vous pour votre présence, toutes nos conversations scientifiques et constructives toujours dans la bonne humeur. Merci Virginie pour ton grand soutien moral et nos discussions scientifiques enrichissantes. Merci Clémence, Dalila, Catherine, Nadia, Christine, Cyndie, Alain, Gaël, Adeline, Nga, Sofia, Sophie, Muriel.*

*Je remercie infiniment ma famille de m'avoir soutenue et de me pousser toujours vers l'avant en toutes circonstances, vous êtes toujours à mes côtés.*

*A mes chers parents, je ne peux pas exprimer ma gratitude ni mes remerciements. Sans vous je ne serai pas arrivée ici. Merci de me faire confiance, merci de me supporter. Merci pour votre soutien malgré les distances qui nous séparent. Ce travail est le votre avant qu'il soit le mien.*

*A ma mère qui m'a appris que les longues études sont une fierté. Merci de vivre avec moi tous les moments difficiles avant les moments de bonheur. Je te dois tout ce que je suis.*

*A mon père qui n'a pas cessé de me soutenir moralement et financièrement. Merci de me faire confiance et de compter sur moi. J'espère que je suis digne de ta confiance.*

*A mon frère, ma sœur et ma belle-soeur.*

*A ma tante Monia et mon oncle Mahjoub qui m'ont appris la persévérance et la passion pour la science et la recherche. Merci pour vos conseils et vos encouragements. Merci d'être un exemple dans la famille que j'ai suivi avec fierté.*

*Un merci particulier à mon oncle Mokhtar pour son soutien et son encouragement, ses coups de fil sympathiques, merci pour nos rires qui me soulagent aux moments les plus difficiles.*

*Merci à tout le reste de ma famille... sans exception.*

*A Hanene, depuis le lycée on ne s'est pas séparées bien que chacune ait choisi un monde différent. Merci pour ton soutien, ta présence, merci d'être ce que tu es, une amie très proche... Merci pour tout.*

*Merci à ma deuxième famille à Paris, la famille Chaibi. Vous m'avez fait oublier les distances et le mal du pays. Latifa, tu es une deuxième mère pour moi, merci pour tout, Nourredine merci pour ton soutien, ton aide et ta gentillesse, Marwa, Manel, Safo, les filles vous étiez une source de joie et de bonne humeur, vous êtes de vraies sœurs...*

*Je remercie l'école doctorale de Paris Descartes pour m'avoir fait confiance en me donnant une bourse de thèse.*



## Table des matières

<b>Table des illustrations .....</b>	<b>7</b>
<b>Liste des abréviations.....</b>	<b>8</b>
<b>1 Les traits quantitatifs, la diversité phénotypique .....</b>	<b>9</b>
<b>1.1 Qu'appelle-t-on un trait quantitatif ?.....</b>	<b>9</b>
<b>1.2 La diversité phénotypique humaine .....</b>	<b>9</b>
<b>1.3 Les facteurs de la diversité.....</b>	<b>11</b>
1.3.1 Les causes génétiques de la diversité phénotypique .....	11
1.3.2 Les causes épigénétiques de la diversité phénotypique.....	14
1.3.3 Les causes environnementales de la diversité phénotypique .....	15
1.3.4 Comment les facteurs génétiques, épigénétiques et environnementaux créent-ils la diversité phénotypique ? .....	15
1.3.5 Apports des études de jumeaux dans l'évaluation de la part de la génétique, de l'épigénétique et de l'environnement dans la diversité phénotypique .....	18
<b>1.4 L'héritabilité des traits quantitatifs .....</b>	<b>19</b>
<b>2 Croissance staturale .....</b>	<b>20</b>
<b>2.1 Généralités .....</b>	<b>20</b>
<b>2.2 La diversité de la taille entre individus .....</b>	<b>20</b>
<b>2.3 La diversité de la taille entre les groupes humains .....</b>	<b>21</b>
2.3.1 Un groupe humain isolé dans les îles Adaman .....	22
2.3.2 Les Pygmées .....	22
<b>2.4 L'héritabilité de la taille .....</b>	<b>22</b>
2.4.1 La taille des jumeaux à la naissance et pendant l'enfance .....	23
2.4.2 La taille des jumeaux à l'âge adulte.....	23
<b>2.5 Les facteurs responsables de la variabilité de la taille .....</b>	<b>24</b>
2.5.1 Les gènes responsables de la variabilité de la taille humaine .....	24
2.5.2 L'exemple de la taille d'une espèce domestique, le chien, qui dépend de la variation génétique au locus IGF1.....	25
2.5.3 La part prise par l'environnement et l'accélération séculaire de la croissance.....	26
2.5.4 La part de l'épigénétique dans la variabilité de la taille .....	27
<b>3 Rappels sur l'épigénétique .....</b>	<b>28</b>
<b>3.1 Définitions.....</b>	<b>28</b>
<b>3.2 Mécanismes et acteurs épigénétiques .....</b>	<b>31</b>
3.2.1 La chromatine et les histones.....	32
3.2.2 Les ARNs non codants .....	34
3.2.3 La méthylation de l'ADN .....	35
3.2.4 Relations entre la méthylation de l'ADN et les histones .....	49
3.2.5 Les variations de méthylation de l'ADN.....	50
<b>4 La croissance du squelette dans l'enfance .....</b>	<b>58</b>
<b>4.1 La taille .....</b>	<b>58</b>
<b>4.2 Historique et généralités .....</b>	<b>58</b>
<b>4.3 Physiologie de l'axe GH/IGF1 .....</b>	<b>58</b>
<b>4.4 Les maladies génétiques mendéliennes (monogéniques) de la production d'IGF1.....</b>	<b>60</b>
4.4.1 Défauts de production de GH .....	60
4.4.2 Les insensibilités (résistances) génétiques à l'hormone de croissance (GHI : GH insensitivity) .....	61
<b>4.5 IGF1 dans le foie et le cartilage .....</b>	<b>63</b>

4.6	Le transport d'IGF1 dans la circulation .....	64
4.7	Effets paracrines d'IGF1 dans le cartilage et signalisation <i>via</i> son récepteur (IGF1R) .....	64
4.7.1	Effets paracrines .....	64
4.7.2	Signalisation <i>via</i> le récepteur d'IGF1R .....	64
4.8	Les concentrations circulantes d'IGF1 .....	65
4.9	La corrélation IGF1-taille en déviation standard .....	66
5	Le gène <i>IGF1</i> .....	68
5.1	La structure du gène .....	68
5.2	Les polymorphismes du gène <i>IGF1</i> .....	69
5.2.1	Les SNPs .....	69
5.2.2	Les CNVs .....	69
5.3	Les transcrits d' <i>IGF1</i> .....	69
5.3.1	La régulation de l'expression du gène <i>IGF1</i> .....	70
5.3.2	La régulation épigénétique du locus <i>IGF1</i> .....	72
5.4	La protéine <i>IGF1</i> .....	72
5.4.1	Structure du prépropeptide .....	72
5.4.2	Modification post traductionnelle du prépropeptide à la protéine mature .....	74
6	Objectifs généraux de la thèse : épigénétique de la croissance .....	75
6.1	Choix de l'approche épigénétique .....	75
6.1.1	Le principe d'analyse de méthylation .....	75
6.1.2	Choix de notre stratégie .....	75
6.1.3	Choix du gène candidat .....	77
7	Les questions au début de la thèse .....	78
8	Les outils de la thèse .....	79
8.1	Le service, les médecins .....	79
8.2	Sujets étudiés .....	79
9	Article 1 : « The <i>IGF1</i> P2 promoter is an epigenetic QTL for circulating <i>IGF1</i> and human growth » .....	80
9.1	Objectifs de l'article 1 .....	80
9.2	Conclusions de l'article 1 .....	81
10	Article 2 : « The P2 promoter of the <i>IGF1</i> gene is a major epigenetic locus for GH responsiveness » .....	82
10.1	Objectifs de l'article 2 .....	82
10.2	Conclusions de l'article 2 .....	83
11	Article 3 : « Genetic and epigenetic modulation of growth hormone sensitivity studied with the <i>IGF1</i> generation test » .....	85
11.1	Objectifs de l'article 3 .....	85
11.2	Conclusions de l'article 3 .....	86
12	Article 4 : « Higher methylation of the <i>IGF1</i> P2 promoter is associated with idiopathic short stature » .....	87
12.1	Objectifs de l'article 4 .....	87
12.2	Conclusions de l'article 4 .....	88
	Références bibliographiques .....	98

## TABLE DES ILLUSTRATIONS

Figure 1 : Distribution infinitésimale de l'effet des gènes sur les traits quantitatifs .....	12
Figure 2 : Représentation schématique d'un événement stochastique.....	17
Figure 3 : Représentation schématique des interactions entre les facteurs génétiques, épigénétiques et environnementaux, et comment ils contribuent aux phénotypes complexes .....	18
Figure 4 : Distribution normale de la taille selon Galton .....	20
Figure 5 : Répartitions des conscrits français mesurés en 1907 par intervalles de taille de 1 cm .....	21
Figure 6 : Représentation graphique de la régression statistique de la taille des parents sur la taille des enfants selon Galton .....	23
Figure 7 : Représentation graphique de l'épigénétique selon Waddington.....	28
Figure 8 : Représentation imagée des interactions sous-jacentes du paysage épigénétique .....	29
Figure 9: Représentation imagée de l'épigénétique moléculaire .....	30
Figure 10 : Représentation schématique de la structure de la chromatine .....	32
Figure 11 : Structure chimique d'un dinucléotide CG méthylé.....	35
Figure 12 : Principaux fournisseurs en groupements méthyles chez l'homme .....	36
Figure 13 : Modèle illustrant le maintien de la méthylation lors de la réplication .....	37
Figure 14 : Réaction de déméthylation de l'ADN par la voie des enzymes TET.....	39
Figure 15 : Classification et distribution des promoteurs sur le génome humain .....	41
Figure 16 : Cycle de l'épigénome.....	42
Figure 17: Exemple de la persistance de certaines marques épigénétiques au cours de la reprogrammation fœtale chez l'homme .....	44
Figure 18: la reproduction du profil de méthylation pendant la réplication cellulaire .....	45
Figure 19: Représentation schématique de la transmission intergénérationnelle des marques épigénétiques .....	46
Figure 20 : Impact du délaissement maternel sur la méthylation du récepteur aux glucocorticoïdes ( <i>GR</i> ) .....	46
Figure 21: Influence des facteurs environnementaux sur l'épigénome de plusieurs générations .....	47
Figure 22: Schématisation de l'ensemble de processus épigénétiques modulant l'expression des gènes .....	49
Figure 23 : Exemple d'action coordonnée des marques épigénétiques de la chromatine et de l'ADN .....	50
Figure 24 : Exemples d'épiallèles métastables chez l'homme .....	52
Figure 25: Influence du comportement social précoce sur l'épigénome .....	56
Figure 26 : Voie de signalisation de l'hormone de croissance .....	59
Figure 27: Courbes de référence de la concentration en IGF1 chez des enfants français, garçons et filles, en fonction de l'âge .....	65
Figure 28 : Structure du gène <i>IGF1</i> chez les mammifères et les étapes aboutissant à la production d'ARNm et de protéines.....	68
Figure 29 : Carte de déséquilibre de liaison du locus <i>IGF1</i> .....	69
Figure 30 : Distribution des CGs sur les deux promoteurs P1 et P2 du gène <i>IGF1</i> .....	72
Figure 31 : La structure du gène <i>IGF1</i> et ses différents peptides.....	73
Figure 32: Relation entre la réponse à l'hormone de croissance dans un tissu cible (foie) et dans un tissu non cible (les cellules du sang (PBMC)).....	84
Figure 33 : Proposition d'un mécanisme de régulation transcriptionnelle au locus <i>IGF1</i> .....	91
Figure 34 : Déméthylation du CG-137 du promoteur d' <i>IGF1</i> en fonction de la dose du traitement GH et de la durée du traitement.....	92
Figure 35: La durée et la dose du traitement GH sont associées à la déméthylation du CG-137 .....	93
Figure 36 : Représentation schématique du promoteur P2 et des CGs étudiés chez l'homme et le macaque. 94	
Tableau 1 : Moyennes (en déviation standard) des âges, des niveaux d'IGF1 et d'anthropométrie à 5, 7-8 et 9-10 ans parmi les enfants inclus dans les analyses de croissance pour la période intermédiaire.....	66
Tableau 2: Associations transversales de IGF-1 à 5 et 7-8 ans avec la taille, les tailles de la cage thoracique et de longueur de jambe, l'IMC et le poids de naissance.....	67

## LISTE DES ABBREVIATIONS

AA : Acide aminé  
ALS : Acid-labile subunit  
ARE : AU rich element  
ASM : Allele specific methylation  
CG : Dinucléotide cytosine-guanosine  
CGI : Ilot CG  
CNV : Copy number variation  
DMR : Differentially methylation region  
DNMT : DNA methyl transferase  
DZ : Dizygote  
EWAS : Epigenetic-wide association study  
GH : Growth hormone (hormone de croissance)  
GHDI : Growth hormone deficiency initiation  
GHI : Growth hormone insensitivity  
GHIH : Growth hormone-inhibiting hormone  
GHR : Growth hormone receptor  
GHRH : Growth Hormone-releasing hormone  
Grb2 : growth factor réception bound protéine 2  
GWAS : Genome-wide association study  
HAT : Histone acetyl transferase  
HDAC : Histone deacetylase  
HMT : Histone methyl transferase  
HNF : Hepatocyte nuclear facteur  
IBD : Identical-by-descent  
ICR : imprinting control region  
IGF1 : Insulin-like growth factor 1  
IGFALS : Insulin-like growth factor binding protein, acid labile subunit  
IGHD : Isolated growth hormone deficiency  
IRS : Insulin Receptor Substrate  
ISS : Idiopathic short stature  
JAK : Janus kinase  
LCP : Low CG promoter  
MAP : Mitogen-Activated Protein  
MBD : Methyl binding domain  
MZ : Monozygote  
PBMC : Peripheral blood mononuclear cell  
PGC : Primordial germ cell  
PI3K : Phosphatidylinositol-3-Kinase  
QT : Quantitative trait  
QTL : Quantitative trait locus  
rhGH : Hormone de croissance recombinante  
SAM : S-adénosyl-méthionine  
SH2 : Src Homology 2  
SHC : Src homology collagen  
SNP : Single-nucleotide polymorphism  
SRIF : somatotropin release-inhibiting factor  
SOCS : Suppressor Of Cytokine Signaling  
STAT : Signal transducers and activators of transcription  
TET : Ten eleven translocation  
TSS : Transcription Start Site  
VMR : Variably methylated region  
VNTR : Variable number of tandem repeat

## 1 LES TRAITS QUANTITATIFS, LA DIVERSITE PHENOTYPIQUE

### 1.1 QU'APPELLE-T-ON UN TRAIT QUANTITATIF ?

Un trait quantitatif (*quantitative trait*, QT) est un phénotype mesurable. Un QT dépend des effets cumulés de nombreux facteurs, il est donc dit multifactoriel ou « complexe ». Selon le théorème de la limite centrale et son application à la biologie (Dacunha-Castelle D, 1996; Lange, 1978), les QT suivent une loi normale dans la nature : de nombreux phénomènes quantitatifs sont dus aux effets cumulés d'un grand nombre de petites perturbations plus ou moins aléatoires, dont chacune joue pour un petit effet, qui diffère d'un individu à l'autre. Ces traits quantitatifs, comme la taille, le poids de naissance, la pression artérielle, le rythme cardiaque, l'âge des premières règles, la glycémie, le taux de cholestérol, et d'autres nombreux paramètres biologiques, ont une distribution gaussienne car leur déterminisme répond à une multiplicité de facteurs. Cette distribution est caractérisée par une valeur moyenne et par une gamme de variation plus ou moins large reflétée par la variance ou l'écart-type.

### 1.2 LA DIVERSITE PHENOTYPIQUE HUMAINE

Selon Dobjansky: « *Nothing in biology makes sense except in the light of evolution* ». Au cours de leur évolution historique et géographique, les hommes ont développé des caractéristiques physiques et biologiques diverses leur permettant de survivre et de se reproduire dans les environnements variés qu'ils ont rencontrés. A l'issue de leurs déplacements et de leur reproduction dans les groupes sociaux qu'ils constituaient de façon durable, les groupes humains se sont distingués par de nombreux traits : couleur de la peau, taille et poids adultes, forme des yeux, texture et couleur des cheveux ou pilosité corporelle, sont les plus apparents et les plus connus. Un des premiers naturalistes à étudier les variations humaines, Buffon (1707-1788), fut pour ainsi dire un évolutionniste avant la lettre. Il exprima sa conviction que les êtres humains sont une seule espèce :

*«Tout concourt à prouver que le genre humain n'est pas composé d'espèces essentiellement différentes entre elles ; au contraire. Il n'y a eu initialement qu'une seule espèce d'hommes, qui, s'étant multipliée et répandue sur toute la surface de la terre, a subi différents changements. Les uns sont d'ordre climatique, les autres concernent les différences de nourriture ou de façons de vivre ; il y a aussi les maladies épidémiques; et le mélange varié, à l'infini, des individus plus au moins ressemblants; que ces altérations n'étaient pas si marquées, et ne produisaient que des variétés individuelles; qu'elles sont ensuite devenues variétés de l'espèce, parce qu'elles sont devenues plus générales, plus sensibles et plus constantes par l'action continue de ces mêmes causes; qu'elles se sont perpétuées et qu'elles se perpétuent de génération en génération, comme les difformités ou les maladies que les parents transmettent à leurs enfants; et qu'enfin, comme elles n'ont été produites ordinairement que par le concours de causes extérieures et accidentelles, qu'elles n'ont été confirmées et rendues constantes que par le temps et l'action continue de ces mêmes causes, il est très probable qu'elles disparaîtraient aussi, peu à peu, avec le temps, ou même qu'elles deviendraient différentes de ce qu'elles sont aujourd'hui, si ces mêmes causes ne subsistaient plus, ou si elles venaient à varier dans d'autres circonstances et par d'autres combinaisons ».*

Citons Piazza qui écrivait en 1992 (Piazza, 1992) :

*« L'explication évolutionniste est simple. Il y a une grande variation génétique dans toutes les populations, même dans les petites. Cette variation individuelle s'est accumulée pendant des périodes*

*très longues, car la plupart des polymorphismes sont antérieurs à la séparation des continents. Peut-être même se situent-ils à l'origine de notre espèce, il y a quelques 700,000 années. Les mêmes polymorphismes sont présents dans presque toutes les populations, mais avec des fréquences différentes dans chacune d'elles, car la différenciation géographique des groupes humains est récente: elle s'inscrit dans une période représentant peut-être 20 % à peine de la durée d'existence de l'espèce. Voilà pourquoi il n'y a pas eu assez de temps pour l'accumulation d'une divergence considérable. Voilà pourquoi la différence entre les groupes est petite en comparaison avec la différence à l'intérieur des groupes. Notre espèce et son prédécesseur immédiat, Homo erectus, ont fait montre d'une activité migratoire considérable, dans toutes les directions. Quelques-unes de ces migrations ont probablement donné lieu à des mélanges entre branches qui s'étaient séparées antérieurement, longtemps avant. »*

La diversité est assez forte pour les traits cités, mais beaucoup moins forte pour des traits tels que poids et taille de naissance, pour lesquels il existe une contrainte physiologique évidente limitant la variation et que partagent tous les humains. Il n'est pas toujours facile d'imaginer ce en quoi la variation d'un trait peut apporter des avantages à un groupe humain, par exemple pour la grande ou la petite taille moyenne d'une population. Comme le soulignent Piazza et de nombreux autres auteurs, la diversité entre les individus est très grande au sein du même groupe humain, rendant hasardeux de définir des groupes aux caractères bien tranchés.

La diversité phénotypique interindividuelle est une chose, la transmission transgénérationnelle de traits en est une autre. Pour se transmettre au sein d'un groupe humain, la diversité des traits peut utiliser divers canaux, dont chacun fonctionne à sa manière et de façon plus ou moins durable : i) la transmission d'allèles de gènes en un ensemble remaniés à chaque union par la recombinaison ; ii) la transmission de marques épigénétiques effacées en large partie lors de la méiose et lors du développement embryonnaire ; iii) la transmission de cultures et de modes de vie. Les hommes changent, la diversité demeure, exceptée lorsque la population devient très petite et très refermée sur elle-même, avec un pool très réduit d'allèles ou d'autres facteurs transmissibles, obligeant à une diversité diminuée.

La diversité de la taille humaine, le centre d'intérêt de cette thèse, est un trait quantitatif classique représentatif de la diversité phénotypique. Les individus d'une population donnée ont une grande variété de tailles dans un intervalle qui caractérise cette population, de part et d'autre d'une valeur moyenne. Dans notre travail, nous nous sommes intéressés seulement à la variabilité phénotypique interindividuelle. Nous n'avons pas abordé la question de la transmission de la variabilité ou des facteurs qui conditionnent cette transmission d'une génération à l'autre.

Nous commencerons par une brève réflexion sur la causalité de la diversité individuelle des traits physiologiques, qui sera suivie d'une réflexion plus ciblée sur la croissance et la diversité de la taille humaine.

### 1.3 LES FACTEURS DE LA DIVERSITE

Trois grandes catégories de facteurs entrent en jeu pour créer la diversité des phénotypes. Ces facteurs interagissent selon des combinaisons fonctionnelles qui restent actuellement inconnues. Il s'agit :

- des variations génétiques (allèles) qui sont des dizaines de millions à travers le génome, et que seuls deux jumeaux monozygotes possèdent pour l'essentiel en commun ;
- des variations environnementales au sens large, incluant les facteurs physiques et chimiques comme les différences climatiques ou dans l'alimentation par exemple, mais aussi les facteurs culturels et sociaux qui modifient la vie des individus et peuvent influencer certains traits de leur phénotype ;
- des facteurs épigénétiques, à l'interface de la génétique et de l'environnement, qui peuvent modifier l'expression des gènes. Ces facteurs peuvent différer entre les tissus et entre les individus, créant un champ immense de variations susceptibles d'influencer l'expression d'une majorité de gènes, et donc d'agir sur les phénotypes. Cette thèse aborde la question des facteurs épigénétiques qui gouvernent la variabilité de la taille.

Reprenons une à une ces trois catégories.

#### 1.3.1 *Les causes génétiques de la diversité phénotypique*

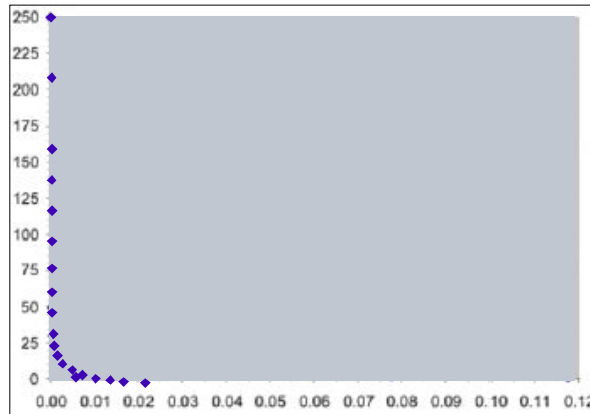
##### 1.3.1.1 Généralités

Les génotypes conditionnent en partie les phénotypes humains, et la diversité des génotypes la diversité des phénotypes. La variation génétique repose sur la présence d'une grande diversité allélique individuelle dans une population, et d'une population à l'autre.

Certains allèles sont fréquents dans la population humaine de nos jours, considérée dans son ensemble ou dans ses différentes composantes géographiques et sociétales. Cela indique que ces allèles sont anciens et/ou profitables, et qu'ils se sont fixés au cours de l'évolution de notre espèce, dans son ensemble ou dans les différentes populations.

Les traits quantitatifs n'échappent qu'en apparence aux lois de Mendel, car ils résultent de la combinaison des effets mendéliens de si nombreux gènes que leur transmission n'est pas régie par une arithmétique simple. La génétique des traits quantitatifs repose donc sur l'intervention de dizaines ou centaines de gènes (ou loci régulateurs), dont la variation n'est pas suffisante pour créer à elle seule le phénotype et sa variation. La génétique entre pour une part dans la causalité de la diversité phénotypique, mais pour une part seulement. Celle-ci varie selon les traits et il n'est pas si facile de l'évaluer (cf paragraphe 1.3.5).

En accord avec la théorie Darwinienne infinitésimale de l'évolution (Figure 1), chacun des allèles les plus déterminants pour un trait donné ne détermine en fait qu'une petite fraction de la variabilité de ce trait, en général moins de 1% de la variance (Farrall, 2004). Au total l'ensemble des gènes qui pèsent sur la variation phénotypique ne représente qu'une fraction de celle-ci. La proportion d'un trait qui est attribuée à la variation génétique est issue de la comparaison des concordances entre jumeaux monozygotes et jumeaux dizygotes. Ce postulat n'est cependant pas aussi évident qu'il pourrait paraître, car les jumeaux monozygotes peuvent partager plus de facteurs non génétiques que les jumeaux dizygotes.



**Figure 1 : Distribution infinitésimale de l'effet des gènes sur les traits quantitatifs.** En abscisse, la proportion de la variance d'un trait expliquée par l'allèle, en ordonnée, l'ordre gène par gène, de l'effet allélique. Le gène le plus important dans la variance a le numéro 1 et contribue pour 2% à la variance totale. Le gène le moins puissant a le numéro et contribue pour moins de 0,05% à la variance totale du trait. *Tiré de (Farrall, 2004).*

Le rôle des facteurs génétiques dans les variations d'un trait quantitatif a été modélisé par la théorie infinitésimale en 1918. Fisher a fait l'hypothèse que la variance des traits quantitatifs était due à la somme d'un très grand nombre de gènes présentant chacun des effets modestes sur le phénotype (Fisher, 1919). Selon cette théorie, une mutation qui entraîne une variation extrême du trait ne persiste pas et a tendance à disparaître. Fisher a établi cette hypothèse en réfléchissant à la taille humaine. Cette théorie a été remise en question en découvrant des gènes capables d'exercer des effets importants sur la variation d'un trait. Il s'agit donc d'une répartition des effets alléliques plus importante que prévue par la théorie de Fisher, selon la théorie exponentielle d'Allan Robertson. Selon cette théorie, les variations phénotypiques sont dues à la fois à des gènes exerçant des effets majeurs et à des gènes exerçant des effets mineurs (Falconer DS and Mackay Trudy F.C, 1996a, 1996b; Robertson A, 1967).

Deux autres adaptations de l'hypothèse de Fisher ont été proposées : Kimura a proposé que les mutations ayant des variations phénotypiques extrêmes puissent être conservées à condition qu'elles soient bénéfiques et qu'elles soient éliminées dans le cas contraire, et Orr a proposé que les variations génétiques des traits soient issues d'un ensemble de mutations successives ayant des effets de moins en moins importants sur le phénotype (Farrall, 2004; Kimura M, 1983; Orr, 1998).

### 1.3.1.2 Les variations de séquence du génome humain

Il existe plusieurs formes de polymorphismes génomiques qui touchent i) soit la séquence nucléotidique dans un endroit donné (changement de la séquence nucléotidique), ii) soit le nombre de copies de certains motifs qui se répètent à l'identique (une répétition en tandem).

Tous les polymorphismes n'ont pas de conséquence fonctionnelle : beaucoup sont « neutres ». Le fondement génétique de la théorie de la sélection naturelle de Darwin postule la perpétuation favorisée des polymorphismes qui confèrent des avantages à l'individu qui les porte. Un polymorphisme est d'autant plus fréquent dans une population qu'il est apparu il y a plus longtemps et qu'il a trouvé à se perpétuer dans cette population. Les polymorphismes de la lactase en sont un



exemple. Mais un polymorphisme peut apparaître et disparaître au sein d'une population sans obéir à la sélection naturelle, ne serait ce que parce que ses porteurs peuvent avoir un destin (survie, fertilité) non régi par la sélection naturelle.

#### 1.3.1.2.1 *Les polymorphismes ponctuels (Single Nucleotide Polymorphism : SNP)*

Les mutations sont des événements moléculaires qui semblent survenir « au hasard » et qui entraînent des modifications de la séquence nucléotidique, créant des « polymorphismes ». Ces mutations peuvent se produire dans les cellules germinales et dans les cellules somatiques. Chez l'homme, la fréquence moyenne d'une mutation est  $10^{-6}$ . Par convention et usage, on appelle « polymorphisme » ou « SNP » (Single Nucleotide Polymorphism) des variations non exceptionnelles de la séquence nucléotidiques (>1%), par distinction avec les mutations rares.

Les SNPs sont des polymorphismes touchant un seul nucléotide, par exemple le remplacement d'un A par un G. Le génome humain compte 3.4 milliards de nucléotides et environ 10 millions de SNPs ont été répertoriés (<http://www.ncbi.nlm.nih.gov/snp/>, <http://hapmap.ncbi.nlm.nih.gov/>). Ceci représente en moyenne un SNP tous les 100-1000 pb. Beaucoup de SNPs sont considérés comme étant non fonctionnels, car ils n'ont pas d'effet qualitatif sur le produit du gène, ni sur la transcription du gène. Mais cette vision évolue : les SNPs peuvent s'associer à des variations épigénomiques locales, incluant la méthylation de l'ADN et les modifications des histones.

#### 1.3.1.2.2 *Les variations du nombre de copies (Copy Number Variation : CNV)*

La séquence du génome humain peut aussi varier dans son organisation séquentielle. Il existe deux principaux types de polymorphisme de répétition :

- les microsatellites sont des polymorphismes de répétition di-, tri- ou tétra-nucléotidiques, pouvant aller jusqu'à 10 nucléotides répétés. La taille des microsatellites est inférieure à 1,5 kb. La distance entre deux microsatellites est en moyenne de 50 kb avec une répartition assez uniforme sur les chromosomes, à l'exception des régions centromériques (Griffiths A.J.F et al., 1999).
- les minisatellites possèdent des unités de répétition plus grandes que les microsatellites (entre 10 et 30 nucléotides), leur taille varie de 1 à 30 kb. Les minisatellites tendent à être plus fréquemment localisés vers les télomères, régions sensibles à la recombinaison.

La variabilité allélique dans ce type de polymorphisme est portée par le nombre de répétitions du motif constituant le micro ou le minisatellite (Griffiths A.J.F et al., 1999).

#### 1.3.1.2.3 *Micro-délétions et micro-insertions*

Les micro-délétions et micro-insertions sont respectivement des événements de suppression et d'addition de quelques bases pouvant avoir des conséquences fonctionnelles :

- elles peuvent décaler le cadre de lecture et générer ainsi un changement d'acide aminé dans la protéine correspondante ou créer un codon stop ;
- elles peuvent affecter des sites d'épissage et générer des transcrits anormaux.

### 1.3.1.3 Les combinaisons individuelles de polymorphismes

La reproduction sexuée est une source de diversité génétique chez l'homme. En effet, lors d'une reproduction sexuée, le nouvel individu va hériter, pour chaque gène, d'un des deux allèles de son premier parent - sélectionné aléatoirement - et d'un des deux allèles de son autre parent. Le patrimoine génétique du nouvel individu est donc ainsi composé aléatoirement d'une partie du patrimoine de chacun de ses deux parents. Ceci permet au matériel génétique de se répandre au sein de l'espèce. Il existe deux principaux types de brassage génétique, qui ont lieu lors de la méiose :

- les recombinaisons intra-chromosomiques : les deux chromosomes d'une même paire portent des allèles différents à un certain nombre de locus. Au cours de la prophase de la première division méiotique, les chromosomes homologues s'apparient : un allèle d'un chromosome peut être échangé avec l'allèle porté par le chromosome homologue, c'est le crossing-over. Tous les gènes situés sur une paire de chromosomes peuvent être « brassés » grâce à ce crossing-over, ce qui modifie l'association d'allèles portés par chacun des chromosomes. Il existe des « hotspots » de recombinaison, correspondant à des régions recombinées fréquemment ; au contraire, certaines régions ne subissent des recombinaisons que très rarement, conduisant à la formation d'haplotypes, c'est-à-dire de groupes d'allèles transmis ensemble.
- les recombinaisons inter-chromosomiques : elles ont lieu pendant la méiose, en anaphase ou lorsque la cellule se sépare en deux ; il y a une séparation aléatoire des chromosomes.

La dernière étape du brassage génétique est la fécondation: la rencontre des gamètes guidée par des facteurs sociétaux.

### 1.3.1.4 Qu'appelle-t-on un QTL ?

Un QTL (« *Quantitative Trait Locus* »), est une région chromosomique contenant un ou plusieurs gènes, capable d'influencer un trait QT. Ces régions chromosomiques contiennent des allèles qui peuvent contribuer à la variation d'un trait. Les traits quantitatifs sont influencés par plusieurs QTLs (Kenneth M Weiss, 1993; Falconer DS and Mackay Trudy F.C, 1996c).

## 1.3.2 Les causes épigénétiques de la diversité phénotypique

Les composants épigénétiques sont capables de faire varier les phénotypes indépendamment de la séquence d'ADN en agissant sur l'expression des gènes. Mais les variations épigénétiques sont également capables d'interagir avec la séquence ADN. En effet, Irizarry et Feinberg (Feinberg and Irizarry, 2010) ont suggéré que des variants génétiques, même s'ils ne changent pas le phénotype moyen, pourraient changer la variabilité de certains phénotypes, par un mécanisme épigénétique associé à ces allèles. Ils rapportent des exemples où les variations de la méthylation à un locus sont associées à des variants génétiques en *cis* ou en *trans*. Ce modèle de variation stochastique héritée fournirait un mécanisme pour expliquer le rôle de l'épigénétique i) dans la biologie du développement ; ii) dans la variation phénotypique sélectionnable ; et iii) dans la variation génétique héréditaire, largement inexpliquée dans les maladies complexes. L'augmentation de la gamme de variation épigénétique, en partie liée à la génétique, en partie stochastique, pourrait faciliter l'adaptation à un environnement variable.

L'analyse de certaines marques épigénétiques, telles que la méthylation de l'ADN, donne comme estimation quantitative le pourcentage de cellules qui portent une marque épigénétique donnée. L'épigénétique offre donc une gamme très importante de variabilité phénotypique.

Les marques épigénétiques peuvent varier: i) d'un tissu à l'autre, chez un même individu ; ii) d'un individu à l'autre dans un même type cellulaire ; et iii) au cours du temps chez un même individu et dans un même tissu.

La variabilité épigénétique se rencontre chez les jumeaux monozygotes, même si les constituants épigénétiques des jumeaux monozygotes sont plus comparables que ceux partagés par les jumeaux issus de deux zygotés différents.

En outre, les erreurs de réplication des marques épigénétiques sont plus fréquentes que les mutations géniques, constituant une autre source de variation (Bennett-Baker et al., 2003). Ces événements de variation épigénétique touchent les cellules capables de se diviser et peuvent survenir à n'importe quel moment de la vie.

### **1.3.3 Les causes environnementales de la diversité phénotypique**

L'environnement est défini comme « *l'ensemble des éléments biotiques et abiotiques qui entourent un individu ou une espèce et dont certains contribuent directement à subvenir à ses besoins* » (dictionnaire Larousse). L'environnement varie selon les générations et au cours de la vie d'un individu. Il est quasiment impossible de le caractériser précisément dans son ensemble.

On peut distinguer deux catégories majeures des composants environnementaux : naturels et non naturels. Les composants naturels englobent les conditions physico-chimiques telles que la lumière du soleil, l'air respiré, la température extérieure, l'oxygène, les êtres vivants (faune et flore) et leurs multiples interactions. Les composants non naturels sont créés par la société et les interventions de l'homme. L'alimentation et le microbiome font aussi partie de notre environnement.

### **1.3.4 Comment les facteurs génétiques, épigénétiques et environnementaux créent-ils la diversité phénotypique ?**

Dans son célèbre livre « *Le Hasard et la Nécessité* » (Jacques Monod, 1970), Monod affirme que l'évolution naturelle est basée sur l'interaction entre le hasard et la nécessité d'amener un changement évolutif adaptatif. Pour autant, la conception de Monod sur la chance est-elle pertinente dans la biologie contemporaine ?

« *God does not play dice* » est une citation célèbre d'A. Einstein appliquée à la physique et à l'univers. On peut certainement dire que le développement ne joue pas aux dés non plus. Nous discuterons à ce propos deux concepts qui se situent à mi-chemin du hasard et de la nécessité : la canalisation et la stochasticité.

#### **1.3.4.1 La canalisation**

La canalisation et la stabilité du développement se réfèrent à la tendance des processus de développement à suivre des trajectoires particulières, malgré des perturbations externes ou internes (voir Figure 8). La canalisation est la tendance au développement d'un génotype particulier de suivre

la même trajectoire dans des conditions différentes (différents environnements, ou contextes génétiques différents), tandis que la stabilité développementale est la tendance d'un génotype particulier à suivre la même trajectoire dans les mêmes conditions. Hallgrímsson B. et ses collaborateurs ont trouvé que les mécanismes sous-jacents à la canalisation et à la stabilité du développement étaient liés (Hallgrímsson et al., 2002).

Les microARNs (cf paragraphe 3) jouent un rôle de canalisation dans le développement animal. En effet, les microARNs confèrent une robustesse aux programmes génétiques du développement, notamment lorsqu'un microARN et ses cibles sont exprimés dans des domaines mutuellement exclusifs. La robustesse du développement se manifeste aussi dans sa réponse à l'environnement et toutes ses perturbations. De nombreux exemples décrivent l'impact de la variation de l'environnement sur les mécanismes épigénétiques (Posadas and Carthew, 2014) .

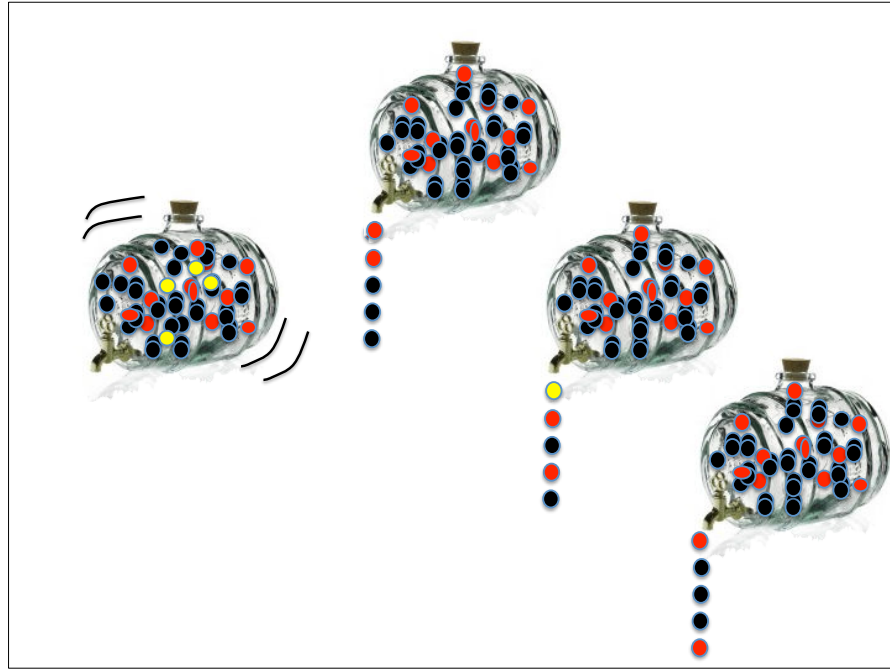
#### 1.3.4.2 La stochasticité

Un processus stochastique est un enchaînement ni entièrement pré-déterminé, ni entièrement aléatoire de phénomènes (Figure 2). Cashmore a décrit une trinité de forces qui gouvernent la biologie : i) les gènes ; ii) l'environnement et iii) la stochasticité (GES) (Cashmore, 2010).

Dans certains cas, ce qui à première vue peut sembler être stochastique pourrait simplement refléter des différences micro-environnementales. Certains physiciens, par exemple Hooft ('t Hooft et al., 2005), n'acceptent pas l'imprévisibilité apparente associée à la mécanique quantique (cette imprévisibilité a incité Einstein à dire, comme nous l'avons déjà dit, «*Dieu ne joue pas aux dés*»).

L'activation des promoteurs, en réponse à un événement stochastique, entraîne des fluctuations aléatoires des niveaux des transcrits (Ford and Thanos, 2009). Ce «bruit» peut perturber la fonction nominale du programme génétique et donc être préjudiciable au phénotype. Les microARNs sont capables de réarranger les niveaux d'expression perturbés. Ce mécanisme est observé dans certains tissus où le transcrit et son microARN sont coexprimés avec des niveaux intermédiaires. Son rôle est de « tamponner » ou de corriger ces fluctuations d'expression. Ces mécanismes, appelés « Noise-buffer », existent afin de réparer et conférer une robustesse à ces voies génétiques (Arias and Hayward, 2006).

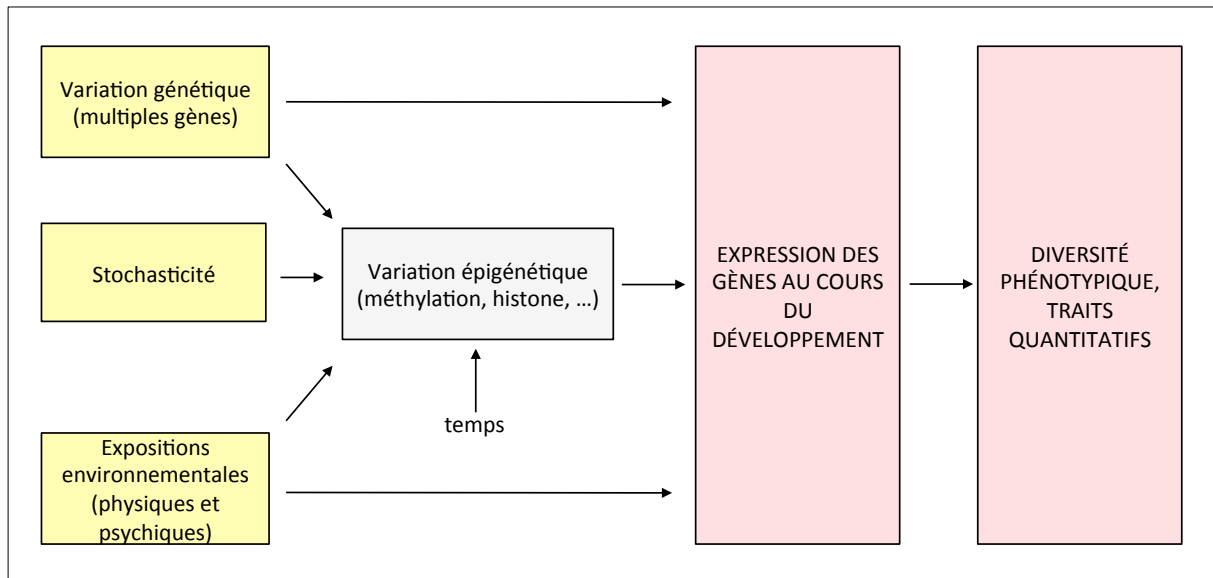
Dans certains cas, la nature stochastique d'un phénomène développemental pourrait refléter des différences micro-environnementales ou des différences spatiales et temporelles subtiles dans de nombreux phénomènes de la biologie moléculaire et cellulaire. Le fait que des déterminants identiques conduisent à des enchaînements différents est symbolisé par deux représentations : les conteneurs de billes (Figure 3) et le paysage de Waddington (Figure 8).



**Figure 2 : Représentation schématique d'un événement stochastique.** Les conteneurs de billes contiennent tous le même nombre de billes jaunes, noires et rouges. Les billes représentent les facteurs ou déterminants. Après avoir secoué le conteneur, ils peuvent laisser échapper différentes combinaisons de billes : noire – noire – noire – rouge – rouge pour le conteneur du haut contre noire – rouge – noire – rouge – jaune pour celui du milieu. Chaque ordre des billes est différent, ni entièrement déterminé, ni dû au seul hasard. Il peut figurer un phénomène développemental.

#### 1.3.4.3 Les interactions

Ces trois facteurs (génétique, épigénétique et environnementaux) peuvent créer la diversité phénotypique de façon indépendante ou conjointe (Figure 3). En effet, comme nous l'avons dit précédemment (cf paragraphe 1.3.2), la génétique peut influencer l'épigénétique modifiant de ce fait l'expression des gènes et pouvant engendrer un nouveau phénotype. L'environnement est également capable de modifier l'épigénétique. L'impact de la génétique sur l'épigénétique et notamment sur la méthylation de l'ADN, ainsi que l'impact de l'environnement sur cette dernière seront expliqués plus en détails dans le paragraphe 3.



**Figure 3 : Représentation schématique des interactions entre les facteurs génétiques, épigénétiques et environnementaux, et comment ils contribuent aux phénotypes complexes.** Les variations génétiques peuvent induire à elles seules des modifications de l'expression des gènes lorsqu'elles touchent des sites clés de la régulation ou de la partie codante du gène, mais elles peuvent également exercer leurs effets *via* des changements épigénétiques. De même, certains facteurs environnementaux agissent directement sur l'expression des gènes tandis que d'autres exercent leurs effets *via* l'épigénétique.

### **1.3.5 Apports des études de jumeaux dans l'évaluation de la part de la génétique, de l'épigénétique et de l'environnement dans la diversité phénotypique**

#### **1.3.5.1 Détermination de la part de la génétique**

La concordance phénotypique entre les jumeaux monozygotes (MZ) sont issus d'une seule fécondation et partagent ainsi la très grande majorité de leur matériel génétique. En revanche, les jumeaux dizygotes (DZ) sont issus de deux fécondations différentes et leur ressemblance phénotypique comme leur degré de partage génétique est la même que celle des frères et sœurs non jumeaux. Le rôle de la génétique dans la ressemblance phénotypique des jumeaux est essentielle, mais n'est pas le seul déterminant héritable ou non qu'ils partagent. Dans le cas des maladies complexes où l'environnement intervient, telles que le diabète de type 2, le taux de concordance entre les jumeaux MZ et DZ est respectivement de 63% et 43%.

#### **1.3.5.2 Détermination de la part de l'épigénétique**

Les jumeaux MZ sont plus ressemblants épigénétiquement non seulement parce qu'ils partagent le même fond génétique, parce qu'ils se sont développés dans un environnement très comparable (Petronis, 2010). Ils présentent cependant des variations individuelles de leur méthylome.

Fraga *et al.* ont montré que les variations individuelles de méthylation dans les lymphocytes de 40 jumeaux MZ augmentaient avec l'âge (Fraga *et al.*, 2005). L'exposition des individus à différents facteurs environnementaux (comme le tabac, la nourriture, les polluants) au cours de la vie pourrait donc induire des modifications épigénétiques (Fraga *et al.*, 2005). Les auteurs de ce travail n'avaient à leur disposition que deux jumeaux âgés de 3 ans, leurs observations peuvent donc être limitées par un manque d'effectif et une seule tranche d'âge. Une autre équipe a comparé le profil de méthylation des cellules buccales de 11 jumeaux MZ âgés de 5 ans. Ils ont focalisé leur étude sur le

locus catechol-o-methyltransferase. Les variations interindividuelles reportées dans cette étude montrent que les variations épigénétiques peuvent être présentes dès l'enfance, avant même une exposition longue et répétée à différents facteurs environnementaux (Mill et al., 2006; Waterland and Michels, 2007). L'origine de ces variations peut donc être stochastique, notamment lorsqu'il s'agit de jumeaux MZ monochorionique monoamniotique (Mill et al., 2006). Dans les autres cas, il peut aussi s'agir d'une exposition différentielle à certains facteurs environnementaux, y compris *in utero*, comme nous le décrivons plus tard dans le chapitre traitant des épiallèles métastables (cf paragraphe 3.2.5.1.4).

#### 1.3.5.3 Détermination de la part de l'environnement

La discordance phénotypique entre les jumeaux MZ est généralement utilisée pour déterminer la part de l'environnement dans la variance phénotypique. L'environnement non partagé entre les jumeaux pourrait donc être à la base des différences phénotypiques observées. Mais, comme on l'a discuté, la part d'épigénétique existe dans le déterminisme de ces différences. On ne peut donc pas trancher entre stochasticité de l'épigénétique et effets de l'environnement (Figure 8).

### 1.4 L'HERITABILITE DES TRAITS QUANTITATIFS

L'héritabilité *stricto sensu* est la fraction de la variabilité phénotypique VP qui est héritable.

L'hérédité *stricto sensu* (VG) est la part de l'héritabilité qui peut être attribuée à la génétique (on dit parfois hérédité génétique ou a tort, par raccourci l'héritabilité tout court).

L'hérédité est classiquement définie par l'équation suivante :

$$H^2 = VG / VP$$

avec VG, la variance génétique, et VP, la variance phénotypique.

La variable IBD (Identical By Descent) correspond à la proportion d'allèles partagés qui sont hérités d'un ancêtre commun (Falconer DS and Mackay Trudy F.C, 1996d).

L'héritabilité inclut également les facteurs non génétiques transmis des parents aux enfants, incluant certains marques épigénétiques, certains composants de l'environnement, de la culture, du comportement...

Classiquement, l'héritabilité ne peut pas encore prendre quantitativement en considération la variation épigénétique, puisque les limites de l'héritabilité épigénétique chez les mammifères sont encore largement discutées (cf paragraphe 3).

## 2 CROISSANCE STATURALE

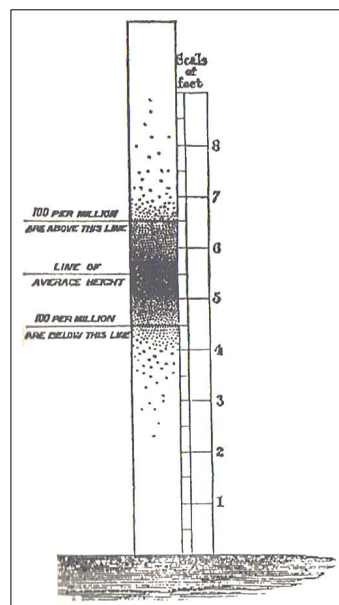
Pour qui s'intéresse au déterminisme des traits quantitatifs humains, la croissance staturale offre un bon domaine de recherche. Facile à mesurer aux différents âges de l'enfance, répondant à une physiologie assez bien établie, et à un contrôle hormonal bien étudié chez l'homme et les animaux de laboratoire, elle permet d'étudier de nombreux traits.

### 2.1 GENERALITES

La croissance humaine n'est pas homogène: elle est la somme de la croissance de la tête, du tronc/bassin et des membres inférieurs, qui ont chacun leur patron de croissance spécifique (Bogin, 1999). La croissance dépend d'un grand nombre de facteurs génétiques et environnementaux. L'interaction complexe, stochastique, entre ces facteurs au cours du développement de l'organisme détermine la croissance squelettique prénatale et postnatale. Les facteurs génétiques participent à la détermination de la taille à la naissance, et deviennent de plus en plus importants pour déterminer la taille après l'âge de deux ans. Ainsi, la croissance prénatale est déterminée par un plus grand nombre de facteurs environnementaux que la croissance postnatale, y compris la parité, l'espacement des grossesses, l'âge de la mère, la taille, la pression artérielle, l'origine ethnique, l'état de santé des mères, le tabagisme, la consommation d'alcool, la contrainte de la nutrition intra-utérine, etc.

### 2.2 LA DIVERSITE DE LA TAILLE ENTRE INDIVIDUS

La taille est un trait mesurable (en centimètre ou en déviation standard). Elle est très variable entre individus de même âge (Figures 4 et 5) et suit une distribution normale à l'échelle de la population, décrite pour la première fois par Galton (cité dans Lange 1978).



**Figure 4 : Distribution normale de la taille selon Galton.** Un point représente un conscrit. *Tiré du livre de Nicholas Wright Gilham, 2001 ; consacré à la biographie de Galton.*



La taille humaine est un trait quantitatif typique qui présente une large gamme de variation au sein d'une même population ou d'une population à l'autre (Figures 4 et 5). La taille humaine varie aussi selon l'âge et le sexe.

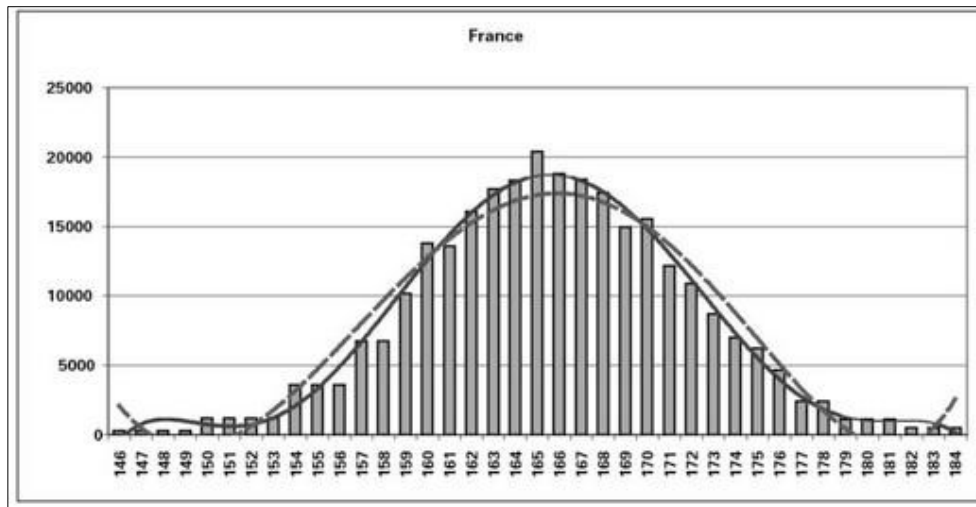


Figure 5 : Répartitions des conscrits français mesurés en 1907 par intervalles de taille de 1 cm. En abscisses, la taille en cm, en ordonnées, l'effectif. Tiré de (Bassino JP Dormois JP, 2009).

### 2.3 LA DIVERSITE DE LA TAILLE ENTRE LES GROUPES HUMAINS

La taille humaine est un trait quantitatif typique qui présente une large gamme de variation au sein d'une même population ou d'une population à l'autre (Figures 4 et 5). La taille humaine varie aussi selon l'âge et le sexe.

La variation de moyenne staturale dans différentes populations est remarquable. La moyenne de taille d'une population congolaise est de 150 cm, et nettement inférieure à celle d'une population hollandaise qui peut atteindre 35 cm de plus. La taille moyenne est généralement plus élevée au Nord de l'Europe qu'au Sud (Cavelaars et al., 2000; Schmidt et al., 1995). La variation frappante de la taille des populations humaines a conduit les chercheurs à se demander si cela représentait une réponse adaptative aux conditions environnementales. Au cours de leur histoire, les hommes ont colonisé presque tous les habitats, et fait face à plusieurs facteurs de stress écologiques.

Dans chaque endroit colonisé, les individus ont interagi entre eux et avec leur milieu : facteurs nutritionnels, facteurs climatiques. D'autre part, les mutations se produisant au hasard, elles peuvent toucher une population donnée dans un endroit donné, ou devenir plus fréquentes dans cette population, tandis qu'elle sera rare ou absente dans d'autres populations. Cette mutation est transmise aux générations suivantes de cette population, sauf en cas de sélection naturelle ou de mort accidentelle de ceux qui la portent (dérive génétique). Le choix du partenaire de reproduction diminue la transmission aléatoire de la mutation chez les humains. Par exemple, les conjoints s'apparient assez fidèlement selon leur taille respective (« *assortative mating* »), comme remarqué par Galton (Galton, 1876; Nicholas Wright Gilham, 2001) et Pearson au 19<sup>ème</sup> siècle (Pearson, 1895).

Citons deux exemples de populations humaines soumises à un isolement géographique.

### **2.3.1 Un groupe humain isolé dans les îles Adaman**

Les îles Adaman, situées dans la baie du Bengale entre l'Inde, la Birmanie et l'Indonésie, sont habitées par une population de chasseurs caractérisés par une peau foncée et une petite taille, très différents de leurs voisins géographiques. Ils sont restés isolés jusqu'à récemment, en partie à cause de leur répulsion vigoureuse des intrus, mais aussi parce qu'ils sont très peu nombreux, seulement 500-600 individus au total. Il a été suggéré qu'ils pouvaient être les descendants des premiers migrants d'Afrique (Jobling M.A et al.).

### **2.3.2 Les Pygmées**

Les pygmées vivent en Afrique centrale. Ils ont plusieurs caractéristiques en commun dont la très petite taille. Les Mbuti, un sous-groupe de pygmées, sont les hommes les plus petits au monde avec une taille moyenne de 150 cm. Généralement, la taille des populations vivant dans les forêts tropicales est plus petite comparée aux populations vivant dans un autre endroit dans le monde. L'explication proposée serait la petite taille est une adaptation aux conditions climatiques, essentiellement la chaleur et l'humidité, qui entraînent des défis particuliers, tels que la thermorégulation et les déplacements dans les forêts denses et la famine (Cavalli-Sforza et al., 1994).

## **2.4 L'HERITABILITE DE LA TAILLE**

En observant la taille des parents et celles des enfants, tout le monde a constaté depuis toujours une tendance pour les parents de taille élevée à avoir des enfants de taille élevée, et des parents de petite taille à avoir des enfants de petite taille (Figure 6). Le premier à notre connaissance, Francis Galton a proposé une relation statistique la régression linéaire (Rodgers and Nicewander, 1988; Stigler, 1989) et le coefficient de corrélation qui décrit le lien entre la taille des parents et celle de leurs descendants, ainsi que la haute héritabilité de ce trait (Galton, 1876; Galton and Dickson, 1886).

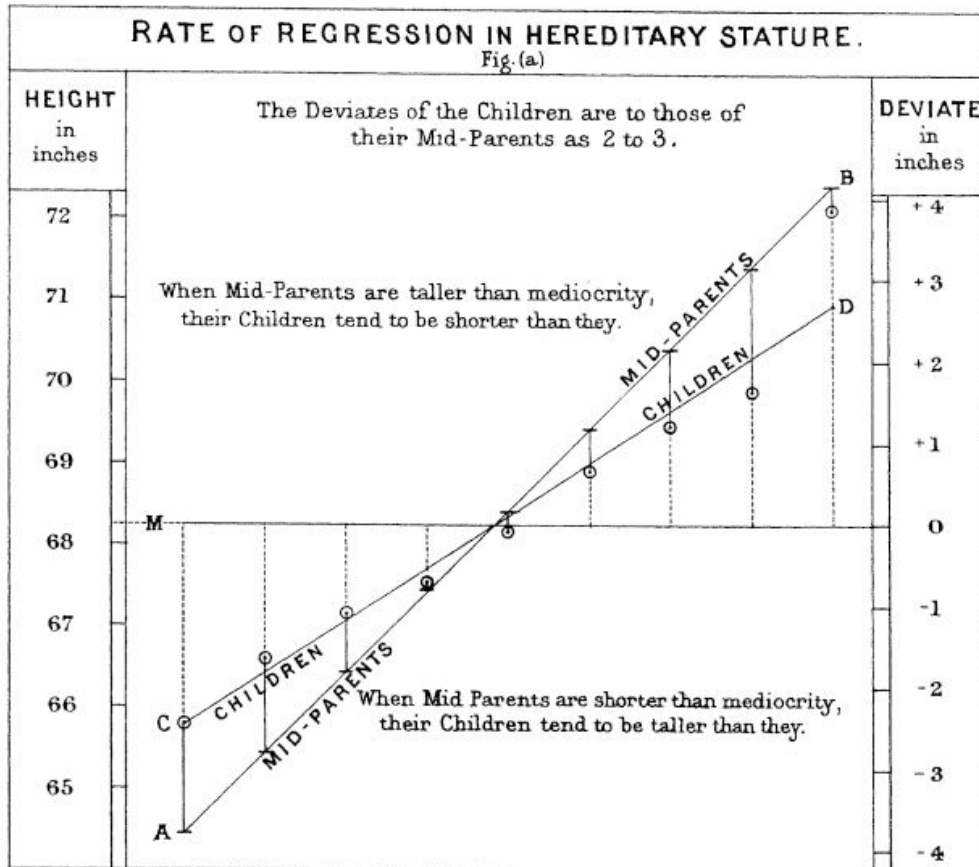


Figure 6 : Représentation graphique de la régression statistique de la taille des parents sur la taille des enfants selon Galton. La taille des enfants nés de parents particulièrement grands ou petits (ligne CD) se rapproche de la taille moyenne de la population (ligne M). Tiré de (Galton, 1886) .

Les études de ressemblance entre apparentés et la comparaison des jumeaux MZ et DZ ont permis d'évaluer l'héritabilité de la taille à environ 0.8. Cette forte héritabilité a été confirmée par plusieurs études dans différentes populations (Silventoinen et al., 2003).

#### 2.4.1 La taille des jumeaux à la naissance et pendant l'enfance

La taille a été largement étudiée chez les jumeaux. Chez les jumeaux MZ, le coefficient de corrélation est de 0.62 à la naissance et augmente au cours de la croissance postnatale pour atteindre 0.94 à l'âge de 8 ans (Bogin, 1999).

Chez les jumeaux DZ le coefficient de corrélation chez les jumeaux de même sexe est plus élevé que chez les jumeaux de sexe opposé. Contrairement aux jumeaux MZ, les jumeaux DZ de même sexe naissent avec des tailles plus semblables que leurs tailles à 8 ans ( $r^2=0.49$  à 8 ans contre 0.79 à la naissance). En revanche, les coefficients de corrélation des jumeaux de sexe différent sont assez stables à 12, 24 et 36 mois, et baissent de la naissance à 8 ans (Bogin, 1999).

#### 2.4.2 La taille des jumeaux à l'âge adulte

L'étude de la taille adulte par Silventoinen a été menée sur 30,111 paires de jumeaux MZ et DZ d'origine caucasienne. La taille des jumeaux MZ est fortement corrélée, le coefficient de corrélation

variant de 0.87 à 0.94 selon la population étudiée. Cette corrélation diminue de moitié chez les jumeaux DZ allant de 0.42 à 0.57 chez les hommes, de 0.49 à 0.56 chez les femmes, et de 0.30 à 0.50 chez les jumeaux de sexe opposé. Ceci suggère une haute héritabilité de la taille humaine avec une forte composante génétique.

## 2.5 LES FACTEURS RESPONSABLES DE LA VARIABILITE DE LA TAILLE

### 2.5.1 Les gènes responsables de la variabilité de la taille humaine

Plusieurs études d'association « génome entier » (GWAS) ont porté sur la taille adulte :

- Lettre et collaborateurs ont génotypé 2,2 millions de SNPs chez 15,821 individus d'origine caucasienne. Ils ont ainsi mis en évidence 12 loci associés à la stature, expliquant à peine 2% de la variation de la taille dans la population étudiée (Lettre et al., 2008).
- Weedon a ensuite estimé la contribution génétique de 20 polymorphismes dans la variance de la taille à seulement 3%, à partir d'un échantillon de 13,665 individus d'origine européenne, génotypés sur 402,951 SNPs (Weedon et al., 2007).
- En portant le nombre d'individus à 29,820 (25,174 Islandais, 2,876 Néerlandais, 1,770 Américains d'origine européenne et 1,148 Afro-Américains), Gudbjartsson n'a pu attribuer à des variants génétiques que 3,7 % de la variance phénotypique (Gudbjartsson et al., 2008).

Ces études ont détecté des SNPs dans des régions de gènes impliqués dans le développement squelettique, le développement embryonnaire et la mitose en général. Les loci candidats les plus associés à la taille adulte sont *HMG2* (High Mobility Group AT-Hook 2) et *ZBTB38* (zinc finger and BTB domain containing 38) ; chacun d'entre eux explique environ 0.3% de la variance phénotypique de la taille, une contribution là encore très minime (Gudbjartsson et al., 2008; Weedon et al., 2007). Ces résultats concordent très bien avec l'hypothèse de Fisher selon laquelle la taille humaine dépend de nombreux loci, chaque locus exerçant un effet minime sur la variance phénotypique globale.

Les analyses de Lango Allen en 2010 ont permis d'identifier un plus grand nombre de polymorphismes impliqués dans la variance de la taille et de proposer de nouveaux gènes candidats. Ces loci se trouvaient à proximité de gènes impliqués dans la croissance squelettique ou de gènes dont la fonction était liée à celle des gènes de la croissance (Lango Allen et al., 2010).

L'héritabilité de la taille étant estimée à 80% (Fisher, 1919; Visscher, 2008; Visscher et al., 2006), les SNPs décrits comme associés à la taille n'expliquant qu'au maximum 10% de la variance phénotypique (Lango Allen et al., 2010), la question de l'« héritabilité manquante » (« *missing heritability* ») a été soulevée. Trois hypothèses ont été proposées pour expliquer cette « héritabilité manquante » :

- les études généalogiques ont surestimé l'héritabilité
- les variants impliqués expliquent individuellement une faible quantité de la variation. Ainsi, leurs effets n'atteignent pas la signification statistique dans les études GWAS ;
- des variants de causalité ne sont pas en déséquilibre de liaison avec les SNPs génotypés dans les puces commercialisées et ne peuvent donc pas être détectés (Visscher et al., 2010; Yang et al., 2010).

Afin de résoudre l'énigme de la « *missing heritability* », Yang *et al.* ont proposé une nouvelle stratégie d'analyse des GWAS. La méthode consiste à estimer la variance totale détectée par l'ensemble des SNPs sans se focaliser sur les SNPs individuellement. Par conséquent, des SNPs individuels n'atteignant pas le seuil de significativité entrent dans le calcul. Dans cette étude, la variance de la taille humaine a été étudiée par 294,831 SNPs chez 3,925 individus non apparentés (Yang *et al.*, 2010). La méthode d'estimation a été validée par des simulations basées sur les données génotypiques observées. L'étude suggère que 45% de la variance pourraient ainsi être expliqués en considérant tous les SNPs simultanément. La différence de 35% entre les études de Allen et de Yang s'explique par l'effet conjoint de ces SNPs ayant individuellement de petits effets non significatifs. Il reste cependant à expliquer les 50% restants de la variance.

On peut suspecter que l'héritabilité restante soit due au déséquilibre de liaison incomplet entre les variants de causalité, ceux qui affectent la taille du fait d'effets fonctionnels, et les SNPs génotypés qui ne sont que des marqueurs. Ce déséquilibre de liaison limité est accru lorsque les variants causaux ont une moindre fréquence de l'allèle mineur (*Minor Allele Frequency*, MAF) que le SNP marqueur de la puce utilisée (Yang *et al.*, 2010).

La méta-analyse la plus récente, utilisant les résultats de 79 GWAS (regroupant 253,288 individus d'origine européenne), a révélé un nombre encore plus élevé de SNPs associés à la taille. En utilisant les données de l'ensemble du génome de ces participants, 697 variants ont été identifiés ( $p < 5 \times 10^{-8}$ ), pouvant expliquer 20% de l'héritabilité de la taille adulte. En diminuant les critères de stringence statistique ( $p < 5 \times 10^{-3}$ ), environ 10,000 SNP semblent contribuer à 36% de l'héritabilité de la taille adulte. En testant différents nombres de variants dans des études indépendantes, les auteurs ont montré que les 2,000, 3,700 et 9,500 SNPs les plus fortement associés expliquent respectivement 21%, 24% et 29% de la variance de la taille adulte (Wood *et al.*, 2014).

Les résultats de Wood et Lango Allen suggèrent une architecture génétique pour la taille humaine caractérisée par un très grand nombre de variants de causalité (Wood *et al.*, 2014). Environ 1% de la totalité des gènes humains contribuerait à la variation de la taille de façon directe ou indirecte (Lango Allen *et al.*, 2010).

Parmi les loci associés à la taille et identifiés par GWAS, une proportion est impliquée i) dans la formation de l'os, ii) dans la prolifération et la différenciation des chondrocytes, et iii) dans la voie de signalisation de l'hormone de croissance (Lettre, 2011; Wood *et al.*, 2014).

### **2.5.2 L'exemple de la taille d'une espèce domestique, le chien, qui dépend de la variation génétique au locus IGF1**

La domestication du chien, à partir de plusieurs souches de loups (Wayne and Ostrander, 2007; Wayne and vonHoldt, 2012), offre un exemple intéressant de réflexion sur le déterminisme de la taille. Aucune espèce n'offre d'aussi grandes différences de taille que les chiens. Les hommes ont généré des races très variées de chiens, dont les traits phénotypiques varient fortement, en croisant entre eux des individus porteurs des traits qu'ils souhaitaient sélectionner. Ceci rappelle les expériences horticoles de Gregor Mendel, qui croisait des pois selon leur couleur, surface, hauteur de hampe... de façon à sélectionner des traits et à les rendre prévisibles.

Par une approche génome entier, plusieurs variants génomiques ont été associés à la taille des chiens et notamment un QTL situé sur le chromosome 15. L'analyse étendue de ce QTL (+/-15 Mb) a permis de mettre en évidence le rôle du gène *IGF1* (insulin-like growth factor 1) dans le déterminisme de la taille chez le chien (Sutter et al., 2007).

### **2.5.3 La part prise par l'environnement et l'accélération séculaire de la croissance**

La faible pression en oxygène liée à l'altitude est un facteur classiquement associé à la petite taille (Bogin, 2013; Chamla, 1983; Jobling M.A et al.). D'autres facteurs liés à l'environnement, moins évidents à discerner, jouent un rôle direct ou indirect sur la croissance staturale en influençant la nutrition et lors du développement *in utero*, de l'enfance (taille des enfants) et de l'âge adulte (influence du pic pubertaire de croissance). Ces facteurs sont certainement multiples et complexes et sont rendus évidents par de nombreuses observations des historiens, des anthropologues, des sociologues et des biologistes à propos de l'évolution de la taille humaine au cours des siècles dans différents groupes humains.

L'analyse des données de 12 pays d'Europe occidentale de 1880 à 1960 a montré une augmentation staturale de 2,4 cm à 10,8 cm chez des conscrits âgés de 20 ans, variable selon les pays. En 80 ans, la moyenne de la taille de la population hollandaise a augmenté de 10,8 cm, celle des espagnols de 2,4 cm seulement. Certaines populations avaient ainsi une moyenne de taille très similaire en 1880 et sont par la suite devenues très divergentes : la moyenne de la taille de la population portugaise est passée de 163,4 à 167,1 cm, alors que la taille de la population hollandaise est passée de 165,2 à 180,3 cm (Bogin, 2013; Chamla, 1983).

L'accélération d'accroissement de la stature est discontinue, il y a des périodes d'arrêt, d'accélération ou même de régression de la taille. Les explications possibles de ces régressions peuvent être dues au biais d'estimation de la taille, ou au manque d'information. Mais il est important de noter que les périodes de décélération sont marquées par des événements particuliers, comme la crise économique mondiale ou la guerre mondiale (Chamla, 1983).

L'étude de la taille chez les Mayas illustre bien l'impact des conditions économiques sur le développement et la croissance. Bogin a comparé deux groupes de Mayas, qui vivent dans des conditions environnementales complètement différentes : un groupe vit dans son pays d'origine, le Guatemala, et l'autre groupe a immigré aux Etats-Unis (Floride et Los Angeles, Californie). La taille et le poids des enfants et des jeunes Mayas vivant au Guatemala sont nettement plus bas comparés à ceux qui vivent aux Etats-Unis. Une différence de 5,5 cm en moyenne pour la taille a été observée. Cette augmentation de stature a été observée en une seule génération. Ces analyses de Bogin réfutent donc l'hypothèse disant que la petite taille des Mayas est essentiellement due à une adaptation génétique à leur environnement très pauvre en apports nutritionnels (Bogin, 1999).

#### *Comment expliquer le changement de la taille humaine au cours du temps ?*

L'accroissement de la taille est lié aux facteurs médico-socio-économiques (Martin, 1958). De manière générale, il a été associé à l'amélioration des conditions de vie. On peut décomposer ces facteurs en trois groupes :

- les facteurs nutritionnels influencent la croissance des enfants. Ainsi, un apport suffisant en acides aminés est essentiel et la sous-alimentation est associée à des baisses de la taille moyenne (Variot, 1904).
- le pic de croissance pubertaire est un déterminant majeur de la taille adulte d'un individu et des différences entre individus. Le lien entre la taille et la précocité actuelle de la maturation physique est prouvé par l'analyse des tailles des conscrits : la taille maximale chez l'homme est atteinte à 20-21 ans ce qui n'était pas le cas pour les hommes du siècle précédent. L'âge auquel la taille adulte est atteinte a considérablement baissé depuis le siècle dernier, mais les données sur l'âge et la taille maximale sont contradictoires (Chamla, 1983; Morant, 1949).
- les facteurs sociaux sont des paramètres environnementaux très importants. En 1950, Trémolières et Boulenger ont montré un lien entre le niveau d'instruction ou l'activité professionnelle et la taille. Ils ont constaté des différences de 4 cm dans la moyenne de la taille entre les niveaux professionnels les plus élevés et les plus bas. Cependant, le niveau d'instruction doit être considéré avec précaution et dans son contexte général sans le séparer des autres conditions sociales. Le travail précoce et la fatigue physique influencent la croissance des individus et entraînent des tassements des cartilages intervertébraux. Les enfants recrutés pour un travail fatiguant ont des tailles moyennes plus petites que les enfants non travailleurs. L'amélioration des conditions d'hygiène et l'élimination des foyers endémiques ont été corrélées avec le phénomène d'accroissement statural. Les infections répétées et les pathologies chroniques ralentissent la croissance. Enfin, l'endroit où vivent les individus, ruraux ou urbains, sont susceptibles d'influer l'augmentation de la taille au cours du temps. En 1951, Giot a proposé une association entre l'accroissement de la stature et le degré d'urbanisme en se basant sur les données de différents départements français. L'augmentation de la taille était de 6,4 cm à Quimper en trois quarts de siècle alors que dans le département du Finistère, l'augmentation n'a été que de 2,9 cm et 0,7 cm dans le Morbihan (Chamla, 1983).

#### **2.5.4 La part de l'épigénétique dans la variabilité de la taille**

La génétique quantitative reste majoritairement une affaire de gènes et d'allèles. La transmission des caractères, l'héritabilité, ne prend encore guère en considération la variation épigénétique. Même avec l'amélioration des études « pangénomiques », la contribution des variants génétiques ne dépasse pas 45% de la variation phénotypique de la taille humaine. Il est donc possible que des facteurs non génétiques, transmis entre les générations, expliquent, au moins en partie, l'héritabilité manquante de la stature chez l'homme.

#### *Comment introduire la variabilité épigénétique dans l'analyse des traits quantitatifs ?*

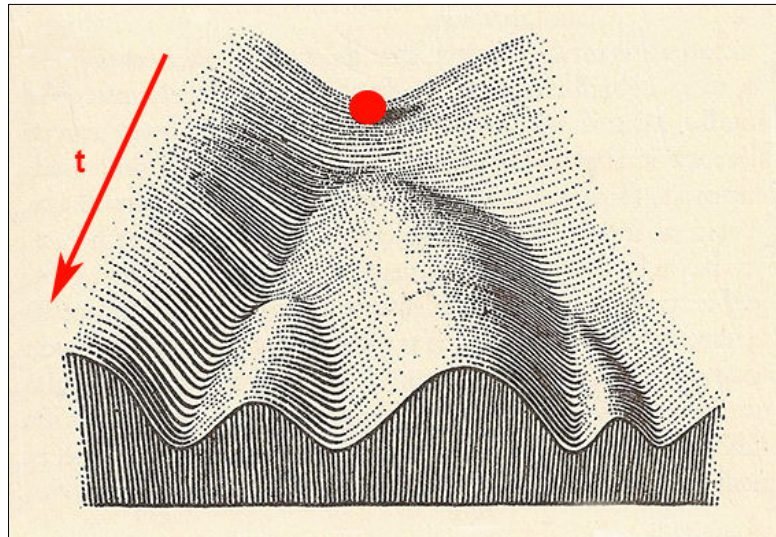
Il n'est pas possible d'étudier la variation épigénétique d'un trait phénotypique indépendamment de la variation génétique en dehors des populations expérimentales isogéniques, ou chez les jumeaux MZ. Dans ces conditions, on peut par exemple analyser des marques épigénétiques à l'échelle du génome entier et détecter des épiallèles impliqués dans la variation phénotypique (phQTL<sup>epi</sup> selon

la nomenclature de Richards) (Richards, 2006). Ces variations épigénétiques sont supposées agir en modulant principalement l'expression génique. Ce genre d'étude reste donc difficile et peu applicable aux QT « naturels » (Kakutani et al., 1999).

### 3 RAPPELS SUR L'ÉPIGÉNÉTIQUE

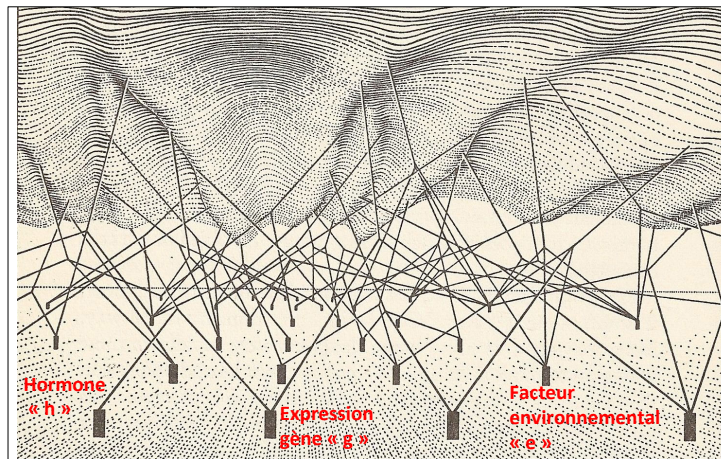
#### 3.1 DEFINITIONS

Il est difficile de définir le terme épigénétique sans un bref historique sur le contexte de son apparition. On attribue généralement à Conrad Waddington l'ébauche d'une nouvelle discipline en biologie qui s'appelle épigénétique. Waddington était un « développementaliste » qui cherchait à comprendre comment une aile de mouche peut se constituer avec ses différentes composantes en partant d'un stock d'allèles identique dans toutes les parties de l'aile (Waddington, 1940; Waddington, 1942, 2012). Il a été très influencé par son séjour dans le laboratoire du drosophiliste Thomas Morgan. En 1947, Waddington a défini l'épigénétique : « *causal interactions between genes and their products which bring the phenotype into being* ». A l'aide du fameux dessin d'un paysage (Figure 7), Waddington a illustré sa façon de voir l'épigénétique (Waddington, 1957). Selon le chemin emprunté par l'œuf fécondé et ses différents tissus dans le paysage épigénétique, le développement s'effectue selon une certaine « canalisation » (cf paragraphe 1.3.4.1) pour donner un individu complet avec ses organes, ses tissus, ses cellules organisées en ordre de marche pour la vie biologique (Waddington, 1942, 1957, 2012).



**Figure 7 : Représentation graphique de l'épigénétique selon Waddington.** La boule rouge représente une cellule ou un embryon pluricellulaire ou un tissu embryonnaire ou fœtal. Les différentes voies de développement/différenciation qui lui sont offertes sont représentées par les vallées. Elle s'orientera vers l'une ou l'autre des vallées selon les facteurs indiqués pour exemples dans la Figure 8, qui seront présents au moment du choix. La flèche rouge indique le fil du temps afin de devenir une cellule différenciée.





**Figure 8 : Représentation imagée des interactions sous-jacentes du paysage épigénétique.** Chaque bloc noir représente un facteur (génétique, environnemental, etc), chaque trait représente le réseau dans lequel ce facteur est impliqué pour créer le paysage épigénétique représenté au dessus. *Inspiré de Waddington, 1957.*

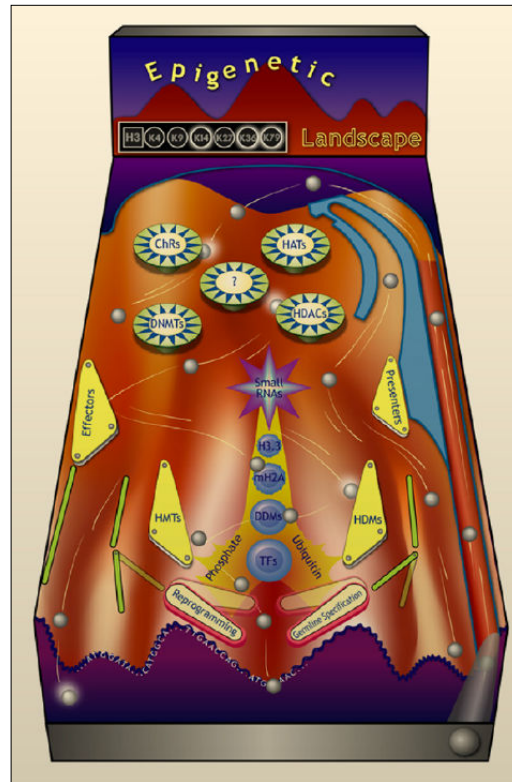
Après les travaux de Waddington, l'épigénétique semble être abandonnée par les biologistes à l'exception d'une citation de S. Lovtrup dans son livre « Theoretical Biology » (Waddington, 1974). Cet auteur introduit le mot « événement épigénétique » et utilise le terme « épigénétique » comme synonyme de biologie du développement.

Au début des années 90, grâce aux progrès de la biologie cellulaire et moléculaire et de la génétique, l'épigénétique est citée de nouveau avec une définition adaptée aux nouvelles connaissances. En 1992, Hall définit l'épigénétique par l'ensemble des facteurs génétiques et non génétiques qui agissent de façon sélective sur les cellules, régulent l'expression génique et augmentent le degré de compréhension de la complexité phénotypique (Jablonka and Lamb, 2002).

Le changement du sens donné à la définition d'épigénétique suit ensuite les nouvelles connaissances dans la régulation de l'expression des gènes et la transmission des phénotypes cellulaires. Les travaux d'Holliday sur la mémoire cellulaire et son lien avec la méthylation de l'ADN participent à dévier la définition de l'épigénétique de son acceptation initiale de biologie de développement. Holliday définit en effet l'épigénétique par « *the study of the mechanism of temporal and spatial control of gene activity during the development of complex organisms* » (Holliday, 1990). Il souligne que les mécanismes épigénétiques doivent inclure la transmission d'un spectre d'activités des gènes dans chaque cellule différenciée.

Quelques années plus tard, et en intégrant les mécanismes moléculaires contrôlant l'expression des gènes tels que les interactions ADN-protéines, Holliday redéfinit l'épigénétique comme « *the study of the changes in gene expression, which occur in organisms with differentiated cells, and the mitotic inheritance of given patterns of gene expression* ». L'auteur ne mentionne pas spécifiquement les mécanismes de développement embryonnaire (Holliday, 1994).

L'apparition de nouveaux mécanismes biologiques jusque-là non explicables a fait évoluer la définition de l'épigénétique, notamment les changements mitotiques et méiotiques dans la fonction des gènes qui ne peuvent pas être expliqués par des changements de séquence de l'ADN. Une revue consacrée aux mécanismes moléculaires de l'épigénétique a remplacé « le paysage épigénétique » de Waddington par un billard électrique moléculaire (Goldberg et al., 2007) dont la complexité est croissante avec les progrès de l'épigénétique moléculaire (Figure 9).



**Figure 9: Représentation imagée de l'épigénétique moléculaire.** Dans cette représentation, les composants de l'épigénétique sont disposés de manière à localiser spécifiquement certains effecteurs dans certaines régions. L'idée est d'imager les rencontres successives de la cellule ES qui vont conduire à sa différenciation. Ainsi, selon l'endroit où va atterrir la bille (représentant la cellule), elle va rencontrer telle ou telle enzyme ou effecteur, qui à son tour va « l'envoyer » vers une autre enzyme, etc, jusqu'à sa destination finale. Tiré de (Goldberg et al., 2007)

L'épigénétique prend sa véritable ampleur en biologie et se définit en 2001 par « *the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence* ». Wu.C et al réduit ainsi l'épigénétique à l'hérédité épigénétique sans intégrer réellement son rôle primordial dans le développement embryonnaire (Wu Ct and Morris, 2001).

Jablonka a introduit l'épigénétique dans les théories de l'évolution. Plus précisément, Jablonka a incorporé l'épigénétique moderne dans la théorie « néo-darwinienne ». Elle propose d'intégrer les différentes composantes épigénétiques dans la plasticité phénotypique. Elle pense que les causes des changements évolutifs pourraient être basées uniquement sur des variations non génétiques transmises par les systèmes de transmission cellulaire (Eva Jablonka and Marion J. Lamb, 2005; Jablonka and Lamb, 2002).

Plus récemment, l'épigénétique a été définie par l'ensemble des modifications de l'expression génique transmissibles au cours des divisions cellulaires, et réversibles (Bonasio et al., 2010), sans modification de la séquence nucléotidique (Jaenisch and Bird, 2003; Mill et al., 2008). Cette définition n'intègre pas la transmission méiotique (transgénérationnelle) des marques épigénétiques encore assez discutée à ce jour chez les mammifères.

La définition qui nous paraît pouvoir résumer les précédentes est la suivante :

***Aujourd'hui, l'épigénétique peut être définie comme un ensemble de mécanismes :***

- 1. qui peuvent influencer ou agir sur des phénotypes cellulaires sans faire intervenir des modifications/variations de la séquence de l'ADN, mais en agissant sur l'expression des gènes ou sur d'autres mécanismes de régulation de la biologie cellulaire ;***
- 2. qui se transmettent d'une cellule-mère à ses filles, et ainsi de suite (hérédité somatique, hérédité mitotique, soft inheritance), assurant ainsi une marque durable au cours de la vie d'un individu, et qui peut se retrouver dans sa descendance proche ;***
- 3. qui sont susceptibles d'un modeste degré de variation et de réversibilité (être plus ou moins effacés et reprogrammés) à différents stades du développement, au cours de la vie de l'individu et selon certains facteurs environnementaux rencontrés ;***
- 4. qui peuvent varier selon les tissus (chaque cellule possède son épigénome), composant une gamme considérable de variation structurale et fonctionnelle (l'épigénotype d'un individu) ;***
- 5. qui peuvent être influencés par la variation génétique en cis- et en trans- (notion de méthylation dépendante d'allèles) indépendamment de l'origine parentale de l'allèle ;***
- 6. dont une petite partie est conditionnée par l'origine parentale du locus (gènes dont la fonction est soumise à l'empreinte parentale) ;***
- 7. dont on discute depuis peu, dans certains cas, la probabilité d'une transmission à la descendance (hérédité transgénérationnelle).***

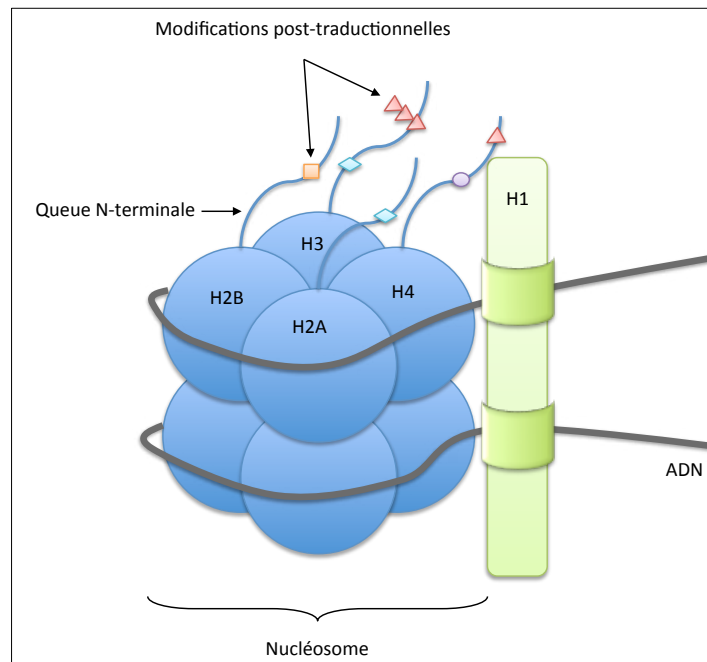
L'épigénétique participe ainsi à la canalisation nécessaire au bon déroulement du développement. Elle constitue également une des composantes de la stochasticité, à l'interface de processus biologiques bien déterminés et d'un certain degré d'influences aléatoires jalonnant la vie de l'individu, depuis le zygote, l'embryon jusqu'à la mort. Cette vue du développement, de la canalisation et de la stochasticité fait aussi partie de la représentation dessinée par Waddington (Figure 7).

### **3.2 MECANISMES ET ACTEURS EPIGENETIQUES**

Les principaux acteurs épigénétiques sont la méthylation de l'ADN, les modifications post-traductionnelles des protéines histones et les ARN non codant.

### 3.2.1 La chromatine et les histones

La chromatine est l'état dans lequel est organisé l'ADN dans le noyau. C'est la substance de base qui forme un chromosome. L'ADN génomique est enroulé autour des protéines « histones ». La structure de la chromatine est restée inconnue jusqu'aux années 1970. Les premières observations de la structure de la chromatine ont été réalisées par microscopie électronique mais elle est restée mal caractérisée jusqu'à la fin des années 90 (Luger et al., 1997) (Figure 10).



**Figure 10 : Représentation schématique de la structure de la chromatine.** Chaque rond bleu représente une protéine histone, sauf l'histone H1 qui est représentée en vert. L'ADN (en gris) est enroulé autour de l'octamère d'histones. Chaque histone peut subir des modifications post-traductionnelles variées (représentées par les triangles rouges, le rond violet, les losanges bleus et le carré orange) sur sa queue N-terminale.

#### 3.2.1.1 Structure de la chromatine

La chromatine est composée de l'ADN, l'ARN et les protéines histones et non histones contenues dans le noyau. Le nucléosome est constitué d'environ 147-165 paires de bases (pb) enroulées autour de huit protéines histone appelées « octamère ». Un octamère est généralement composé d'un tétramère H3-H4 entouré par un dimère H2A-H2B (Figure 10). Les histones sont des protéines très riches en acides aminés basiques chargés positivement ce qui facilite leur liaison à la molécule d'ADN qui porte une charge négative sur son groupement phosphate.

Dans le noyau, la chromatine est présente sous deux états, l'euchromatine et l'hétérochromatine, qui dépendent du degré d'enroulement et d'interaction au sein du noyau :

- l'euchromatine correspond à un état décondensé, comportant des régions génomiques accessibles et des gènes actifs qui s'expriment de façon permanente comme les gènes ménagers (« *Housekeeping gene* ») ;

- l'hétérochromatine correspond à un état condensé et moins accessible. L'hétérochromatine peut être constitutive ou facultative. L'hétérochromatine constitutive correspond à des régions génomiques « muettes » qui ne s'expriment pas. Parmi ces régions, on peut citer les régions répétées, les régions centromériques et télomériques. En revanche, l'hétérochromatine facultative peut contenir des régions capables de s'exprimer occasionnellement dans certains types cellulaires ou selon l'état de différenciation d'une cellule donnée. En effet, l'hétérochromatine facultative correspond à des régions essentiellement compactées et qui peuvent ponctuellement se présenter sous forme accessible.

### 3.2.1.2 Les protéines histones

Il existe 5 classes de protéine histones nommées H3, H2A, H2B, H4, H1. La position centrale est occupée par les histones H3, H2A, H2B, H4 au niveau du complexe nucléoprotéique (« *core histone* »). L'histone H1 est considérée comme protéine de liaison (Figure 10). Les histones sont formées par une partie globulaire (octamère) et une partie flexible et communicante avec l'extérieur à leur extrémité NH<sub>2</sub>-terminale. Ces extrémités subissent des modifications chimiques comme la méthylation, l'acétylation, la phosphatation, l'ubiquitination et la sumoylation. Ces modifications sont essentiellement des liaisons covalentes entre un acide aminé (lysine K ou arginine R) et le groupement chimique en question (méthyl, acétyl ou autre).

#### 3.2.1.2.1 *Le code histone*

Il correspond à la modification chimique retrouvée sur les histones qui permet de les classer en marque active ou répressive. H3K9, H3K27, H3K4, H4K20, H2AZ, H3K36 sont les positions des histones les plus connues et les plus étudiées. Généralement, une acétylation des histones 3 et 4 est associée à un état transcriptionnel actif lorsqu'ils sont localisés au niveau des promoteurs comme H3K9ac, H3K27ac et H4K20ac. La lysine 4 de l'histone 3 fait exception puisqu'elle est plutôt activée sous forme tri-méthylée (Shahbazian and Grunstein, 2007; Wang et al., 2008). La méthylation au niveau de H3K9 et H3K27 est une marque de répression. La lysine 4 est soit mono-méthylée, soit di-méthylée. Parfois, la même marque peut être classée à la fois active et répressive selon le contexte génomique. Par exemple, H3K4me<sub>3</sub> est une marque active quand elle est retrouvée dans les promoteurs, mais répressive quand elle est située au niveau d'un enhancer.

#### 3.2.1.2.2 *Machinerie responsable des modifications des histones*

Il existe de nombreuses enzymes impliquées dans l'établissement des modifications post-traductionnelles des histones, nous en citons quelques unes ci-dessous (Kouzarides, 2007):

- l'Histone Déacétylase (HDAC) qui est l'enzyme catalysant l'élimination du groupement acétyl (Haberland et al., 2009) ;
- l'Histone Acétyl transférase (HAT) qui est l'enzyme catalysant l'acétylation des histones à partir des acétyl co-enzyme A (Kouzarides, 2007) ;
- l'Histone Methyl transférase (HMT) qui est l'enzyme catalysant la réaction de méthylation des protéines histones. A partir de la S-Adénosyl-Méthionine (SAM), cette enzyme est capable de fixer un, deux ou trois groupements méthyle sur l'acide aminé en question (une lysine ou arginine) (Haberland et al., 2009; Kouzarides, 2007).

### 3.2.1.2.3 *Rôle des modifications post-traductionnelles des protéines histones*

Effet sur la transcription : L'ADN doit être décondensé et accessible aux facteurs de transcription et à toute la machinerie transcriptionnelle pour pouvoir s'exprimer. La conformation et l'accessibilité de la chromatine sont modulées par les modifications covalentes des histones. La charge apportée par la modification chimique conditionne le lien entre l'ADN et l'histone modifiée. Par exemple, l'acétylation de la lysine neutralise la charge basique de la protéine histone et diminue l'affinité électrique entre l'histone et l'ADN.

Rôle dans la réplication : Une enzyme impliquée dans la machinerie de mise en place des modifications des histones est impliquée dans le mécanisme de la réplication. L'acétyltransférase (HBO1) a été identifiée comme un acteur important de la réplication de l'ADN. HBO1 est capable de catalyser la réaction d'acétylation de l'histone H4 et de se lier à l'assemblage protéique de pré-réplication. La délétion du HBO1 entraîne une réduction de la synthèse d'ADN et la dérégulation de la phase S du cycle cellulaire. HBO1 est donc importante pour l'initiation de la phase S et la fixation à l'origine de réplication.

Rôle dans l'organisation chromosomique : La structure condensée de la chromatine au niveau des régions télomériques et centromériques semble protéger la structure du chromosome. En effet, l'existence des éléments génomiques répétés dans un contexte chromatinien compacté est un moyen d'empêcher l'expression de ces séquences capables de déréguler l'expression de tout le génome.

## 3.2.2 *Les ARNs non codants*

Certaines régions génomiques sont transcrites sans être traduites, ce sont les ARNs non codant (ARNnc). Ce sont des acteurs épigénétiques importants car ils sont capables de réguler l'expression des gènes. Ils sont transmis lors des divisions cellulaires et des gamètes au zygote.

### 3.2.2.1 Différents types d'ARNs non codants

Les ARNs non codants sont très hétérogènes, on en distingue environ 9 types différents. Les principaux ARNnc sont : les ARNs ribosomiques, les ARNs de transfert, les microARNs, les petits ARNs interférents (« *Small interfering RNA* » : siRNA), les long ARNs codés par les séquences intergéniques (LncARN), etc. On peut plus simplement les classer en ARNs non codants longs et courts.

### 3.2.2.2 Mécanisme d'action des ARNs non codants et machinerie

Les mécanismes d'action des ARNs non codants sont des interactions ARN-ARN, ARN-protéine et ARN-ADN. Prenons l'exemple des petits ARNs interférents. Les petits ARNs interférents sont de petits ARNs double brin de 21 à 24 nucléotides. Les petits ARNs interférents à l'état de double brin sont reconnus dans le cytoplasme de la cellule par un complexe de guidage protéique nommé complexe RISC2 (*RNA induce silencing complex*). Celui-ci s'active en libérant le brin complémentaire de l'ARN. Le complexe ainsi activé va reconnaître son transcrite cible, un ARN messager, par complémentarité des bases nucléiques. Ce système de reconnaissance assure la haute spécificité de ce mécanisme.

Une fois la cible liée, la protéine Argonaute, faisant partie du complexe RISC, va couper le transcrit au niveau du site de reconnaissance. Les deux morceaux du transcrit clivé par Argo vont être rapidement dégradés *via* leurs extrémités par des exonucléases. Les ARNs non codants peuvent aussi interagir avec les ARNs de transfert (ARNt) ou les ARNs ribosomiaux (ARNr) et perturber ainsi la traduction du gène.

En outre, les ARNnc peuvent modifier la stabilité des ARNm ou bloquer la traduction du gène. Les recherches récentes dans le domaine des ARNnc prouvent qu'ils interagissent à la fois avec des protéines de la chromatine et avec la machinerie de la méthylation de l'ADN pour réprimer l'expression. Par exemple H19 interagit avec MBD1. Les ARNnc peuvent également exercer leurs effets après la transcription et la traduction en inhibant le transport protéique (Filipowicz et al., 2005; Tijsterman et al., 2002).

### 3.2.2.3 Rôle des ARNs non codants

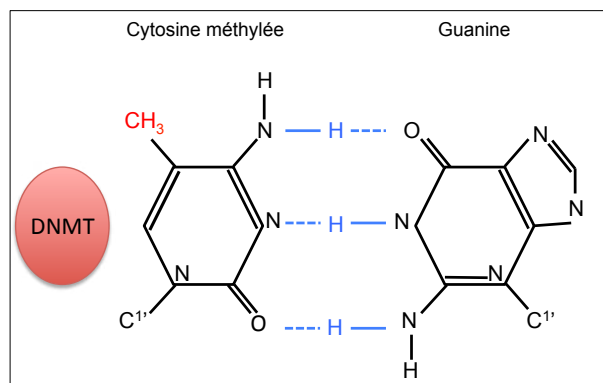
L'action majeure des ARNnc est la répression des gènes. Ils sont notamment indispensables pour la régulation des gènes des régions soumises à l'empreinte parentale (Fitzpatrick et al., 2002; Williamson et al., 2006). De même, l'inactivation du chromosome X est assurée par l'ARN non codant Xist qui recouvre la région chromosomique réprimée.

## 3.2.3 La méthylation de l'ADN

La méthylation de l'ADN est la marque épigénétique la plus connue et la plus étudiée depuis 1975. Elle est relativement stable et c'est la plus facile à étudier dans un contexte d'études épidémiologiques portant sur un grand nombre d'individus.

### 3.2.3.1 L'ADN méthylé

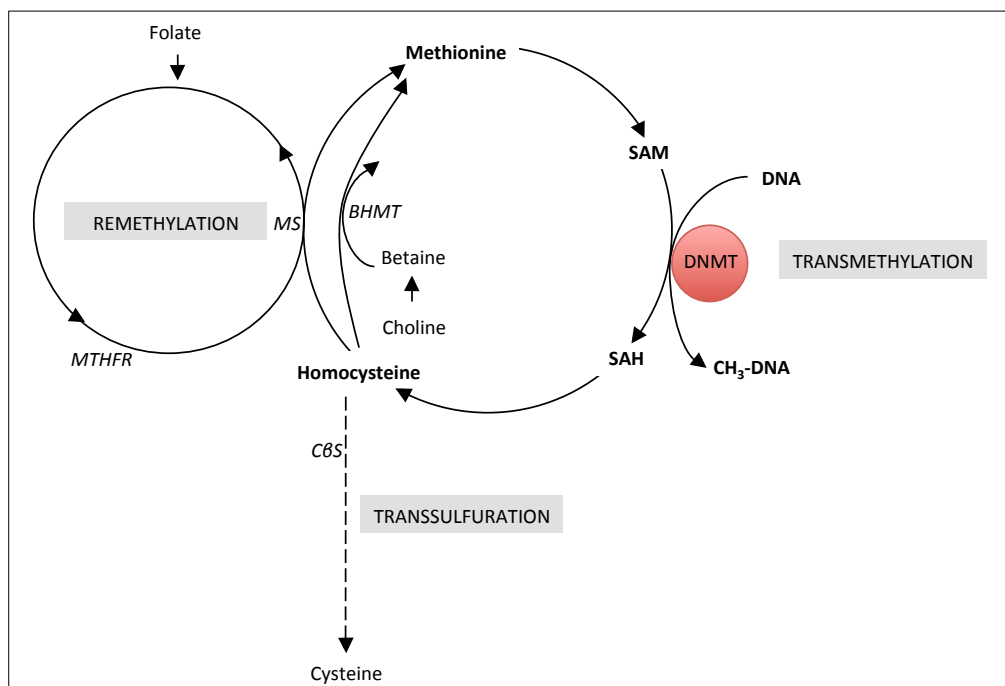
La méthylation de l'ADN est une marque épigénétique apposée directement sur les cytosines de la molécule d'ADN (Figure 11). La méthylation des cytosines a été identifiée chez les mammifères dans les années 60 par Daskočil et Šorm (Daskočil and Šorm, 1962). Cette modification est présente sur les deux brins de la double hélice d'ADN (symétrique) et ne perturbe pas leur hybridation. La méthylation CG est répliquée de façon semi-conservative lors de la réplication de l'ADN dans la cellule.



**Figure 11 : Structure chimique d'un dinucléotide CG méthylé.** En rouge, le groupement CH<sub>3</sub> ou groupement méthyle. En bleu les liaisons entre les 2 bases. L'ovale rouge représente les protéines DNMTs responsables du transfert du groupement méthyle du donneur à la molécule d'ADN.

### 3.2.3.2 La réaction chimique

La méthylation de l'ADN est une modification covalente par laquelle un groupement méthyle est transféré de la S-adenosylmethionine (SAM) au carbone 5 d'une cytosine d'un dinucléotide CG, par l'intermédiaire d'une DNA méthyltransférase (DNMT). La méthionine interagit avec l'ATP pour former la SAM. Ce substrat (SAM) est le donneur principal du groupement méthyle (CH<sub>3</sub>) dans la cellule. La DNMT fixe le groupement CH<sub>3</sub> sur le carbone 5 de la cytosine tout en transformant le S-adenosylméthionine (SAM) en S-adenosylhomocystéine (SAH). Ce dernier est hydrolysé par la suite en homocystéine. L'homocystéine néoformée peut subir une méthylation, recréant ainsi de la méthionine (Figure 12).



**Figure 12 : Principaux fournisseurs en groupements méthyles chez l'homme.** La méthionine est transformée en SAM qui lors de sa dégradation en SAH va libérer un groupement méthyle qui pourra être transféré sur l'ADN. La SAH pourra ensuite être « reméthylée » en méthionine *via* l'homocystéine. Cette reméthylation est dépendante des apports en acide folique ou en choline/bétaine. L'ovale rouge représente les protéines DNMTs responsables du transfert du groupement méthyle du donneur à la molécule d'ADN. SAM: S-adenosyl-méthionine. DNMT: DNA méthyltransférase, SAH: S-adenosyl-homocystéine, CBS: Cystathionine β synthase, BHMT: Bétaine homocystéine S méthyltransférase, MS: Méthionine synthase, MTHFR: Méthylène tétrahydrofolate reductase, IIB, IIF: Initiation factors, TBP: TATA binding protein, PolII: RNA polymérase II, HAT: Histone acétyltransférase, HDAC: Histone désacétylase, MeCP, MBD: Methyl Binding Protein. *Tiré de (Fradin et Bougnères, 2012).*

### 3.2.3.3 Principaux acteurs de la réaction de méthylation

#### 3.2.3.3.1 Source de groupements méthyles (CH<sub>3</sub>)

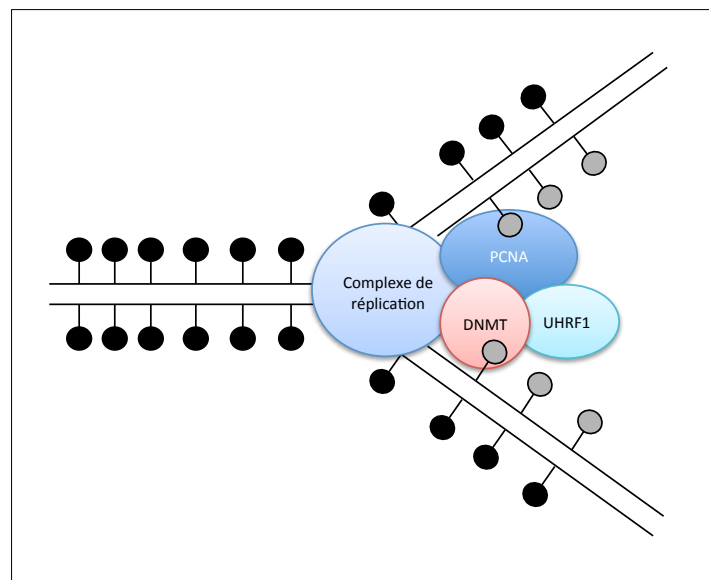
La méthionine, acide aminé apporté par l'alimentation, est la source principale de CH<sub>3</sub> dans la réaction de méthylation de l'ADN. La reméthylation de l'homocystéine en méthionine est dépendante des apports en acides foliques par l'alimentation ou en choline/bétaine (Figure 12).



### 3.2.3.3.2 Les DNMTs

Les DNMTs jouent un rôle clé dans l'établissement de la méthylation d'ADN. Elles sont très dépendantes de l'apport en SAM de la cellule. Le rapport SAM/SAH est parfois utilisé comme indice du taux de méthylation dans les études épidémiologiques (Dominguez-Salas et al., 2014). Chez les mammifères, trois familles de DNMTs ont été décrites (Cheng and Blumenthal, 2008), nous détaillerons ici les 2 principales :

- DNMT1 a été la première protéine DNMT caractérisée. Elle possède un domaine C terminal, responsable de l'activité catalytique de l'enzyme et un domaine N terminal régulateur, permettant d'interagir avec d'autres protéines et guidant sa localisation nucléaire. DNMT1 est active sous forme d'homodimère. Son activité est régulée par divers mécanismes : i) sa quantité d'ARNm est contrôlée par les micro-ARNs ; ii) des modifications post-traductionnelles comprenant la phosphorylation, la méthylation, l'acétylation ou la sumoylation (Denis et al., 2011; Huang et al., 2010; Svedružić, 2011; Zhao et al., 2011). DNMT1 présente une forte affinité pour l'ADN hémiméthylé, c'est pourquoi DNMT1 a longtemps été décrite comme la DNMT impliquée dans le maintien de la méthylation durant la réplication cellulaire. De plus, DNMT1 est exprimée principalement durant la phase S du cycle cellulaire, phase pendant laquelle l'ADN est répliqué et la méthylation des cytosines également. DNMT1 est également présente au niveau de la fourche de réplication, via son interaction avec des protéines telles que PCNA et UHRF1 (Figure 13) (Leonhardt et al., 1992). L'inactivation de DNMT1 dans les cellules souches embryonnaires (ES) ou les embryons murins aboutit à une déméthylation globale du génome et une léthalité embryonnaire (en milieu de gestation).



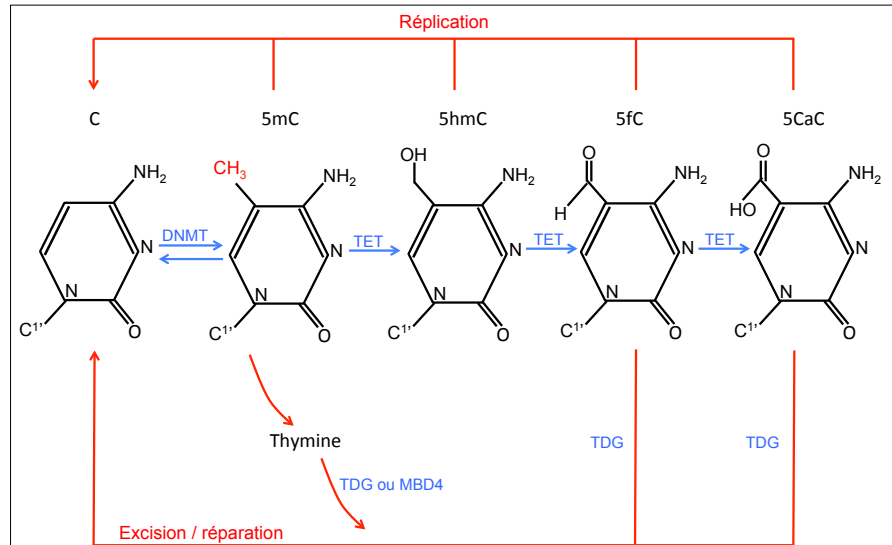
**Figure 13 : Modèle illustrant le maintien de la méthylation lors de la réplication.** La méthyltransférase 1 (DNMT, rond rouge) est recrutée au niveau de la fourche de réplication par interactions avec la protéine UHRF1 (ovale turquoise). Elle interagit également avec l'ADN hémiméthylé et avec l'antigène nucléaire de prolifération cellulaire (PCNA, ovale bleu). Les cercles noirs et gris représentent respectivement les cytosines méthylées sur le brin parental et celles à méthyler sur le brin en cours de réplication. *Inspiré de (Law et Jacobsen, 2010).*

- DNMT3 : Cette classe comprend plusieurs membres : DNMT3A, B, 3L, 3o... Initialement cette classe de DNMTs a été décrite comme responsable de la méthylation *de novo*, même si elles sont désormais aussi associées au maintien de la méthylation (Kaneda et al., 2004; Okano et al., 1998). Les protéines DNMT3A et B ont une forte homologie de séquence et présentent des domaines protéiques communs : i) un domaine catalytique (C terminal), ii) un domaine putatif d'interaction avec les protéines (N terminale) (Dhayalan et al., 2010). Ainsi, ce domaine joue un rôle dans l'interaction entre les marques des histones et la méthylation de l'ADN. Les sites d'interaction de DNMT3A et 3B avec l'ADN restent à définir mais semblent nécessiter la présence de certains nucléotides donnés en amont et en aval du CG à méthyler et une périodicité de 8 à 10 nucléotides entre deux CpG méthylés (Jia et al., 2007; Lister et al., 2009). Elles sont très exprimées au stade blastocyste et dans les PGCs (*Primordial Germ cells*), deux stades subissant d'intenses méthylations *de novo*. Leurs expressions somatiques restent très faibles à partir du milieu de la gestation sauf dans certains tissus incluant la rate, le thymus et le cerveau.  
DNMT3L a été découverte par homologie de séquence avec les autres DNMT3. Elle est dépourvue de site actif, se lie mal à l'ADN et ne possède pas de site de fixation du S-adenosyl-méthionine, mais elle est un cofacteur essentiel au fonctionnement des DNMT3A et 3B (Cheng and Blumenthal, 2008).

#### 3.2.3.4 Le processus de déméthylation

Il existe deux voies de déméthylation de l'ADN, l'une passive, liée à la non réplication de la méthylation au cours des divisions cellulaires et l'autre active, résultant d'un ou plusieurs processus enzymatiques conduisant au retrait du groupement méthyle sur la cytosine.

La déméthylation active a été mise en évidence sur divers loci, tel que le locus de l'interleukine 2, qui voit son gène se déméthyle en 20 minutes sans réplication de l'ADN (Bruniquel and Schwartz, 2003) ou sur le locus *Bdnf* (Brain-derived neurotrophic factor) (Kangaspeska et al., 2008; Métivier et al., 2008). Cette déméthylation active passe principalement par l'oxydation par les protéines TET des cytosines méthylées (5mC) en hydroxyméthylcytosine (5hmC). Ces mêmes protéines sont aussi capables de modifier la 5hmC en 5-formylcytosine (5fC) puis en 5-carboxylcytosine (5-caC), *in vitro* et *in vivo* (He et al., 2011; Ito et al., 2011). La transformation de la 5mC en 5fC ou 5caC conduit à l'excision de cette base par les thymine-DNA-glycosylases (TDG, (He et al., 2011; Maiti and Drohat, 2011), appartenant à la famille des protéines de réparation d'ADN (Figure 14).



**Figure 14 : Réaction de déméthylation de l'ADN par la voie des enzymes TET.** La méthylation de l'ADN (5mC) est établie et maintenue par les DNMT. La 5mC peut être oxydée par la famille des TET dioxygénase pour générer la 5-hydroxyméthylcytosine (5hmC), la 5-formylcytosine (5fC) et la 5-carboxylcytosine (5CaC). Les dérivés 5mC oxydés (5hmC, 5fC, 5CaC) ne peuvent pas servir de substrats pour DNMT1, ils peuvent être perdus par la réplication via la déméthylation passive. La 5hmC peut être désaminée par AID/APOBEC et devenir une 5-hydroxyméthyluracil (5hmU) qui, avec la 5fC et la 5CaC, peuvent être excisées par une glycosylase telle que la TDG (Thymine DNA Glycosylase), excision suivie par la réparation de l'ADN pour générer une cytosine (C). En variante, une décarboxylase putative peut convertir la 5CaC en C.

#### 3.2.3.4.1 L'hydroxyméthylation de l'ADN

L'ADN des mammifères contient des formes oxydées de la 5-méthylcytosine (5mC). La plus abondante est la 5-hydroxyméthylcytosine (5hmC). La 5hmC est produite à partir de la 5mC dans une voie enzymatique impliquant trois oxydases TET (*Ten eleven translocation*) : TET1, TET2 et TET3 (Figure 14). La découverte des propriétés catalytiques des protéines Tet et de leur implication dans la réaction de l'oxydation 5mC a permis de lier ce processus à la déméthylation de l'ADN. Bien que le rôle de 5hmC ne soit pas entièrement élucidé, elle pourrait être un intermédiaire dans le processus de déméthylation de l'ADN actif ou passif. Ce processus a été bien documenté dans le contexte de reprogrammation de l'œuf fécondé pendant le développement embryonnaire des mammifères (Szulwach et al., 2011; Tahiliani et al., 2009). L'existence de la 5hmc dans des cellules différenciées et dans plusieurs tissus, notamment le cerveau, suggère qu'elle n'est pas qu'une simple transition d'un état méthylé à un état non méthylé mais qu'elle existe en tant que telle (Kriaucionis and Heintz, 2009). Aujourd'hui il est connu que la 5hmc est une marque épigénétique qui intervient dans l'expression génique en interagissant avec les protéines de remodelage de la chromatine (Jin et al., 2010).

#### 3.2.3.5 Distribution de la méthylation CG sur le génome humain

La fréquence des CGs dans le génome des mammifères est faible comparée à celle des dinucléotides AT, et surtout moins fréquent que ne le voudrait le hasard. Cette faible fréquence est peut-être due au fait que la méthylcytosine mute facilement en thymine. La méthylation de l'ADN est présente sur tout le génome et 70% à 80% des CGs sont méthylés.

### 3.2.3.5.1 *Les régions génomiques denses en CG : les « îlots CG »*

Un îlot CG (*CG Island* : CGI) est défini par : i) sa composition en C+G supérieure à 55% ; ii) un ratio de C+G observé sur théorique qui doit être supérieur à 0.6 ; et iii) une taille supérieure ou égale à 200 bp (Law and Jacobsen, 2010; Portela and Esteller, 2010; Suzuki and Bird, 2008).

Les îlots CG sont retrouvés, entre autres, dans les promoteurs et l'exon 1 des gènes. Les CGIs sont généralement déméthylés, indépendamment de l'état transcriptionnel du gène. Ces régions semblent protégées de la méthylation *de novo* par un mécanisme encore mal décrit. Bien qu'ils soient déméthylés, les promoteurs à CGI présentent un enrichissement en H3K4me2, marque inactive, qui leur assure un état transcriptionnel inactif (Weber et al., 2007).

En dehors des régions soumises à l'empreinte parentale, dans les cellules somatiques 3% des îlots CGs sont méthylés (Weber et al., 2007). Les îlots CG hyperméthylés sont localisés surtout dans les promoteurs des gènes de régulation du développement embryonnaire comme homeobox et PAX (Illingworth et al., 2008; Oda et al., 2006). Les CGI ont été la cible privilégiée des études EWAS pendant de nombreuses années du fait de leur association avec divers phénotypes cancéreux.

### 3.2.3.5.2 *Les régions à basse densité en CG*

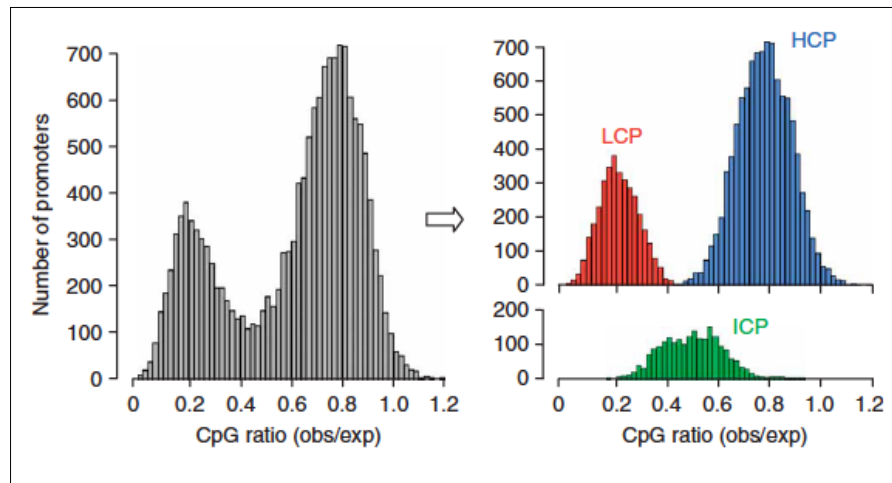
Une faible densité en CG est trouvée principalement au niveau des régions inter-géniques et des jonctions exon-intron du génome des mammifères.

### 3.2.3.5.3 *La composition CG des promoteurs*

Les promoteurs peuvent être classés selon leur contenu en CG. Weber et ses collaborateurs (Weber et al., 2007) ont ainsi défini 3 classes de promoteurs : des promoteurs de haute fréquence en CGs, des promoteurs à faible fréquence en CGs et entre les deux une classe de promoteurs avec un niveau intermédiaire de CGs (Figure 15).

Promoteurs à haute densité en CGs : Chez les mammifères, au moins 60 % des promoteurs contiennent des îlots CG. Les gènes à expression ubiquitaire possèdent généralement un promoteur de haute fréquence en CGs.

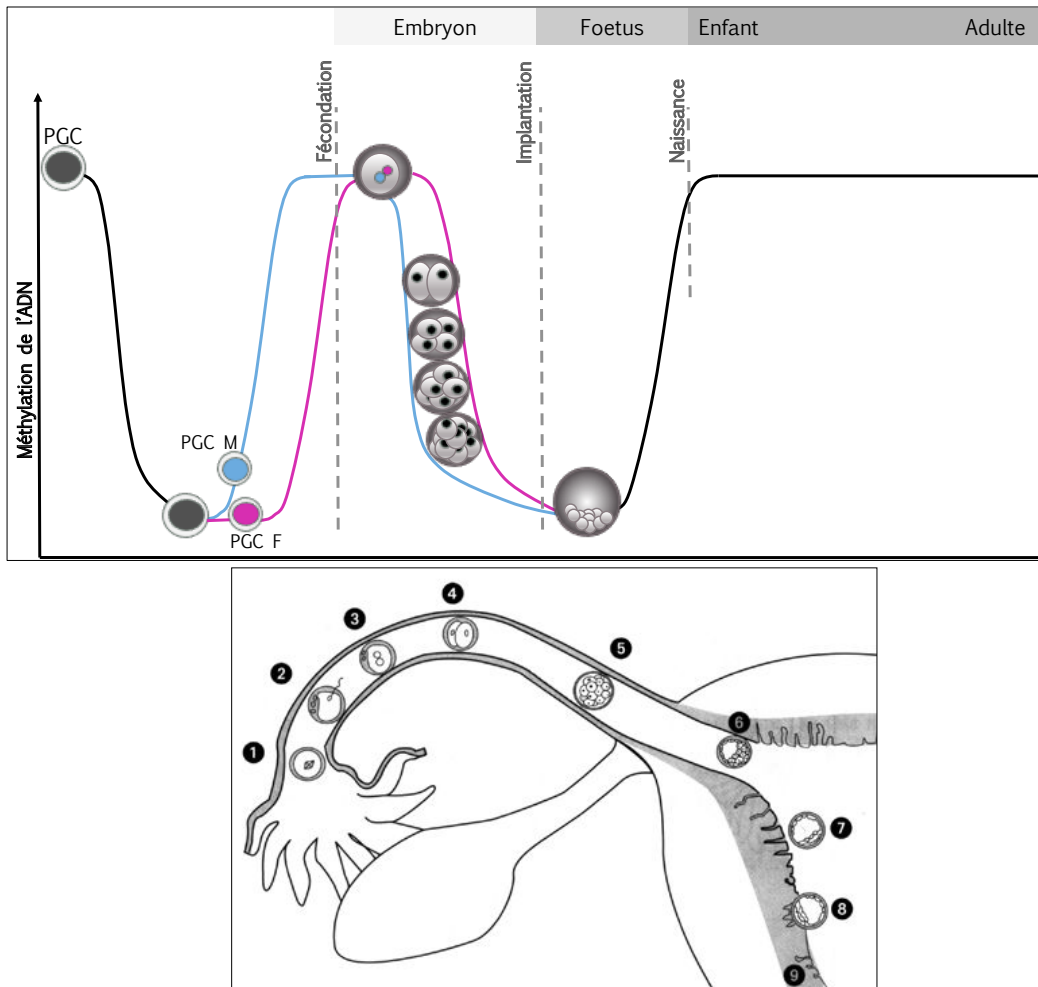
Promoteurs à faible densité en CGs (Low CG Promoter : LCP) : Cette classe de promoteur possède peu de dinucléotides CG, mais leur fréquence varie selon les nomenclatures. Ils sont définis comme des séquences contenant entre 10 à 20 CGs par kilobase (kb) (Weber et al., 2007). Contrairement aux CGI, les CGs de ces promoteurs sont généralement méthylés. Certains travaux suggèrent que la méthylation des CGs des LCP n'empêche pas la transcription du gène (Ball et al., 2009; Meissner et al., 2008). Cependant, l'étude du profil de méthylation dans les cellules différenciées montre que la plupart des changements d'expression de gène est associée à des modifications de méthylation au niveau des CGs des promoteurs LCPs (Rakyan et al., 2008; Yagi et al., 2008). Le rôle de la méthylation des promoteurs LCP reste à définir. Jusqu'à présent, les LCP n'ont pas suffisamment étudiés pour déterminer leurs rôles dans la régulation de la transcription. En effet, comme nous le disions, les outils techniques développés pour les études EWAS ciblaient jusqu'à récemment presque exclusivement les CGIs.



**Figure 15 : Classification et distribution des promoteurs sur le génome humain.** Distribution bimodale de 15,609 promoteurs selon le ratio CG du génome humain. LCP : promoteur à faible densité en dinucléotides CGs. ICP : promoteurs à densité intermédiaire en CGs. HCP : promoteurs à haute densité en CGs. *Tiré de (Weber, 2007).*

### 3.2.3.6 Cycle de la méthylation de l'ADN au cours du développement

Le génome des mammifères subit des changements dynamiques de la méthylation de l'ADN dans le cycle de la vie de chaque individu. L'établissement de la méthylation, où ce qu'on appelle la méthylation *de novo*, est un processus dynamique qui s'effectue principalement pendant i) la gamétogenèse et ii) les stades précoces du développement embryonnaire (Figure 16).



**Figure 16 : Cycle de l'épigénome.** Représentation schématique du niveau global de méthylation de l'ADN dans les cellules germinales primordiales (PGC) à gauche et de la fécondation à la vie adulte (à droite). La ligne rose représente la méthylation dans les gamètes femelles puis dans le pronucléus femelle, de même la ligne bleue représente la méthylation dans les gamètes et le pronucléus mâle. En bas à gauche, représentation schématique du trajet de l'ovocyte à l'implantation du zygote dans l'utérus : 1) l'ovocyte, 2) la fécondation, 3) la fusion, 4-5-6-7) les divisions successives jusqu'à l'implantation en 8.

### 3.2.3.6.1 Du zygote à l'embryon

Quelques heures après la fécondation, les génomes parentaux subissent des modifications épigénétiques majeures (Mayer et al., 2000; Oswald et al., 2000; Reik, 2007).

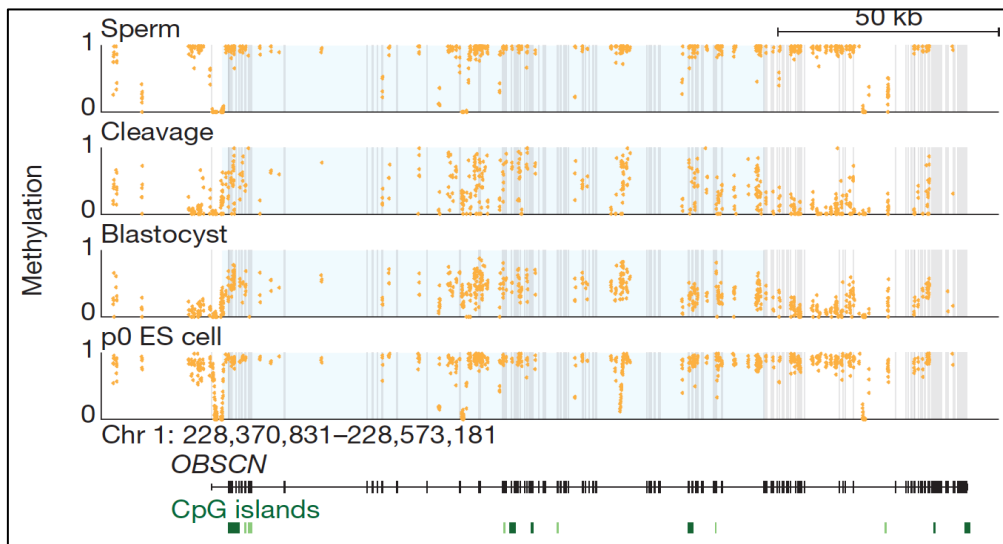
La déméthylation du pronucléus mâle est dite « active ». Celle du génome maternel, plus tardive, est dite passive. La déméthylation est achevée au niveau du stade blastocyste (E3.5 chez la souris) (Reik, 2007). Pendant cette première vague de déméthylation, certaines régions génomiques échappent à la déméthylation et conservent leurs marques initialement établies dans les gamètes, ce sont les régions soumises à l'empreinte parentale. L'empreinte parentale est un mécanisme par lequel un seul des deux allèles parentaux s'exprime, en fonction de son origine parentale. Après la nidation, le profil de méthylation est progressivement mis en place par les DNMT3A et B, en fonction du type cellulaire (Guibert et al., 2009).

### 3.2.3.6.2 *Dans les gamètes*

La deuxième vague de déméthylation est effectuée au niveau des PGCs (*Primordial germ cell*). Sept jours après la fécondation, les PGCs se différencient. Les PGCs migrent alors vers les crêtes génitales correspondant aux gonades en formation. Pendant les différents stades de formation des gamètes, le génome de ces cellules subit une reprogrammation de ses marques épigénétiques. La reprogrammation épigénétique touche à la fois la méthylation de l'ADN et les marques de la chromatine. Contrairement à la première vague de déméthylation, la deuxième touche l'ensemble du génome y compris les régions soumises à l'empreinte. Cependant, les régions répétées sont partiellement déméthylées afin d'éviter des événements de délétion aléatoire ou de réactivation des éléments mobiles dans les cellules germinales (Seisenberger et al., 2013). A partir du stade embryonnaire (E12.5), le génome se reméthyle progressivement notamment sur les ICR (Imprinting Control Region) et les régions répétées. Un nouveau profil de méthylation s'établit selon le sexe au niveau des régions soumises à l'empreinte parentale. Dans les cellules germinales mâles, l'empreinte paternelle s'établit entre le stade embryonnaire 14.5 (souris) et la naissance. Dans les lignées germinales femelles, la méthylation des régions ICR ne commence qu'après la naissance, pendant la période correspondante à la multiplication des ovocytes (Chotalia et al., 2009; Lucifero et al., 2004).

### 3.2.3.6.3 *L'embryon et le fœtus*

La formation des gamètes puis celle de l'embryon sont accompagnées chacune par des événements d'effacement des marques épigénétiques : cette "remise à zéro" est nécessaire à la spécialisation des gamètes puis à la pluripotence des toutes premières cellules de l'embryon. La remise à zéro des marques épigénétiques provenant des parents n'est cependant pas totale dans l'embryon. Certaines marques échappent à l'effacement (Figure 17). Au sein de l'embryon, les cellules sont au départ (presque) identiques. Elles vont rapidement recevoir des signaux très orchestrés les conduisant à activer ou inactiver certains de leurs gènes pour se différencier en telle ou telle lignée cellulaire et construire l'organisme (Figure 7, Waddington, 1942). Les marques épigénétiques alors mises en place doivent se transmettre au cours des divisions cellulaires, pour qu'une cellule de foie reste une cellule de foie, une cellule nerveuse une cellule nerveuse. C'est ce qu'on peut appeler une forme de mémoire cellulaire.



**Figure 17: Exemple de la persistance de certaines marques épigénétiques au cours de la reprogrammation fœtale chez l'homme.** Le gène *OBSCN* est pris comme exemple. Dans l'embryon dès les premiers clivages, une région de 130 kb, surlignée en bleu, reste spécifiquement méthylée tandis que les régions flanquantes, en blanc, sont déméthylées. Chaque point représente un CG capturé par RRBS (Reduce Representation Bisulfite Sequencing). L'axe des ordonnées représente la méthylation de l'ADN mesurée par la fréquence des CGs méthylés capturés (comprise entre 0 et 1). Tiré de (Smith et al., 2014)

La différenciation cellulaire de l'embryon commence de façon concomitante avec la reméthylation de l'ADN. Le rôle de la méthylation de l'ADN dans la spéciation cellulaire a pris son ampleur avec la méthylation *de novo* des gènes marqueurs de pluripotence. Les travaux étudiant la reprogrammation de cellules différenciées en cellules pluripotentes ont montré que la méthylation de l'ADN pouvait être une barrière contre la dédifférenciation cellulaire. En effet, l'inactivation des marqueurs de pluripotence cellulaire tels que NANOG et OCT4 coïncide avec la déméthylation des promoteurs correspondant à ces deux facteurs (Bhutani et al., 2010).

*In vivo* chez la souris, après la nidation, le profil de méthylation est progressivement mis en place par les DNMT3A et B, en fonction du type cellulaire (Guibert et al., 2009; Seisenberger et al., 2013). La méthylation *de novo* est responsable de la répression des gènes de pluripotence, des gènes spécifiques de la lignée germinale et des gènes spécifiques selon la lignée cellulaire. La méthylation de l'ADN augmente de façon globale dans l'embryon et s'achève vers le stade embryonnaire E6.5 (Borgel et al., 2010). A partir de la phase embryonnaire E7.5 caractérisée par une activité mitotique intense, le maintien de la méthylation se fait grâce à la DNMT1 tandis que les protéines DNMT3A et 3B commencent à disparaître. Une fois la méthylation de l'ADN établie dans l'embryon, cette marque est répliquée à chaque division cellulaire de façon identique à la cellule mère.

Bien que l'ensemble des cellules de l'embryon avance vers un état de plus en plus différencié, certaines cellules individuelles gardent une certaine plasticité dont l'ampleur dépend du stade de développement. La restriction progressive de plasticité cellulaire s'accompagne de modifications épigénétiques qui imposent une mémoire et un destin cellulaire stable (Hemberger et al., 2009).



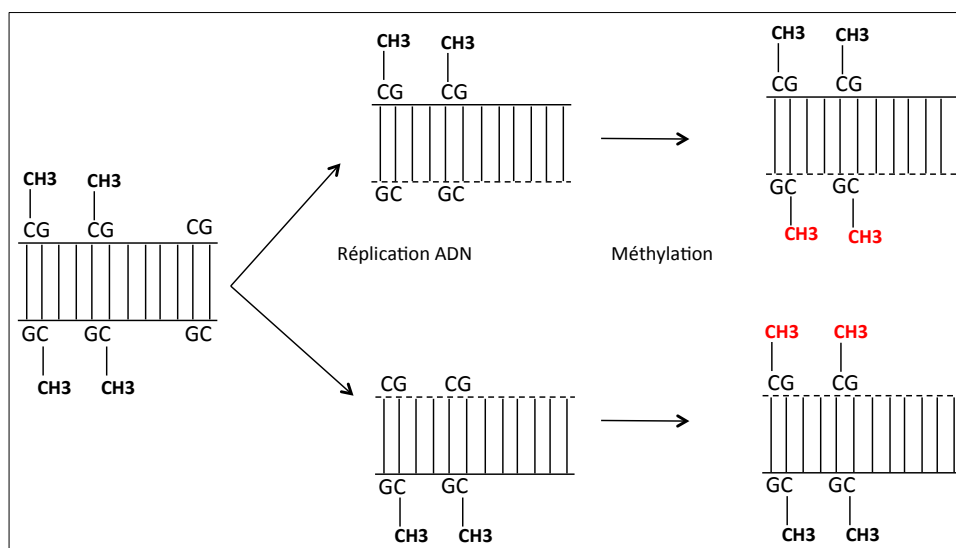
### 3.2.3.6.4 Au cours de la vie postnatale

Au cours de la vie postnatale, les cellules sont différenciées et continuent à se diviser pour certaines. Au cours de ces divisions mitotique les marques épigénétiques sont maintenues grâce à la DNMT1 (Figure 13), mais, comme nous l'avons déjà dit, des erreurs peuvent survenir.

### 3.2.3.7 L'héritabilité de la méthylation de l'ADN

#### 3.2.3.7.1 Héritabilité mitotique

Elle repose sur la réplication de la méthylation de l'ADN qui reproduit les marques de la cellule-mère dans ses deux cellules-filles selon le principe figuré ici (Figure 18) grâce à des mécanismes moléculaires figurés (Figure 14). Cette transmission des marques assure le maintien des régulations de l'expression des gènes et la stabilité du phénotype cellulaire différencié au sein d'un tissu.

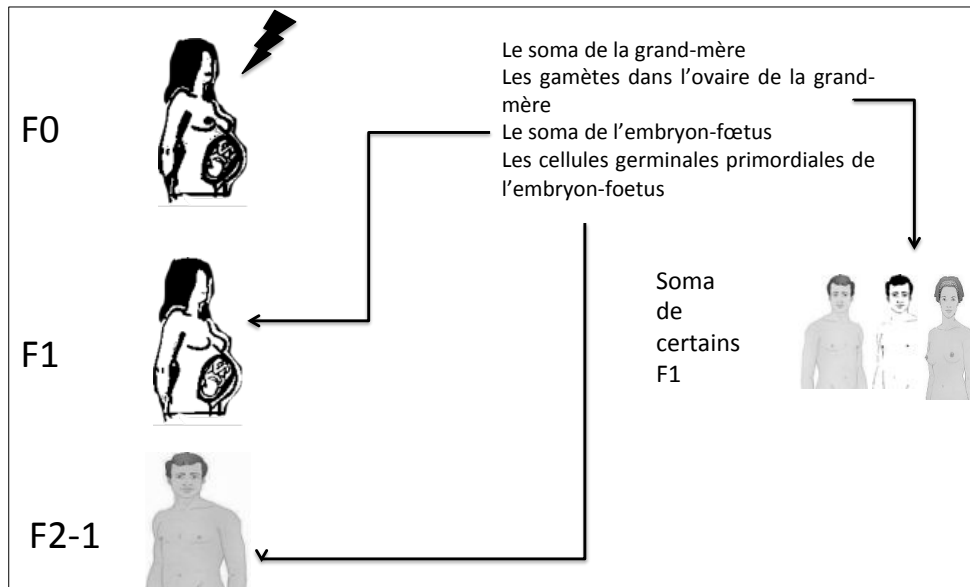


**Figure 18: la reproduction du profil de méthylation pendant la réplication cellulaire.** Tiré de (Eva Jablonka and Marion J. Lamb, 2005) trait continu brin d'ADN parental, trait discontinue brin d'ADN néosynthétisé.

Du fait de la transmission mitotique, une modification d'une marque épigénétique imprimée à moment donné dans une cellule peut être durablement transmise à la descendance cellulaire ou persister dans la cellule si celle-ci est post-mitotique comme la majorité des neurones, par exemple. C'est la transmission mitotique qui permet la transmission inter-générationnelle, à ne pas confondre avec la transmission transgénérationnelle.

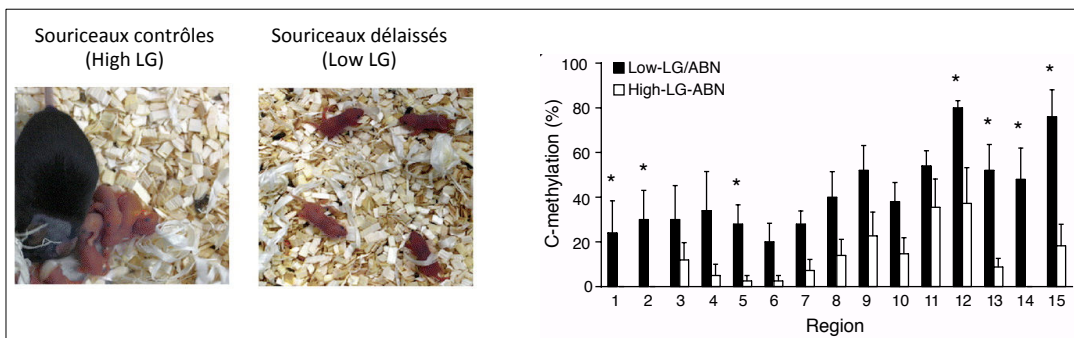
#### 3.2.3.7.2 Héritabilité intergénérationnelle

Elle désigne la trace laissée par une génération dans la génération suivante et même celle d'après, comme représenté sur la figure « des trois générations F0, F1, F2 » (Figure 19).



**Figure 19: Représentation schématique de la transmission intergénérationnelle des marques épigénétiques.** L'exposition *in utero* d'une femme (F0) à un phénomène environnemental (l'éclair noir) peut provoquer des changements épigénétiques sur l'embryon F1 en développement et sur sa lignée germinale. L'exposition touche ainsi trois générations: la mère (F0), l'embryon (F1), et l'individu issu des cellules germinales (F2). Schéma de Pierre Bougnères.

La transmission intergénérationnelle est également représentée par les expériences de Meaney chez le rat nouveau-né. Dans ce cas le comportement maternel modifie l'épigénétique de la fille, qui influence en retour le comportement en un cycle « comportement-épigénétique » qui passe les générations (Figure 20).

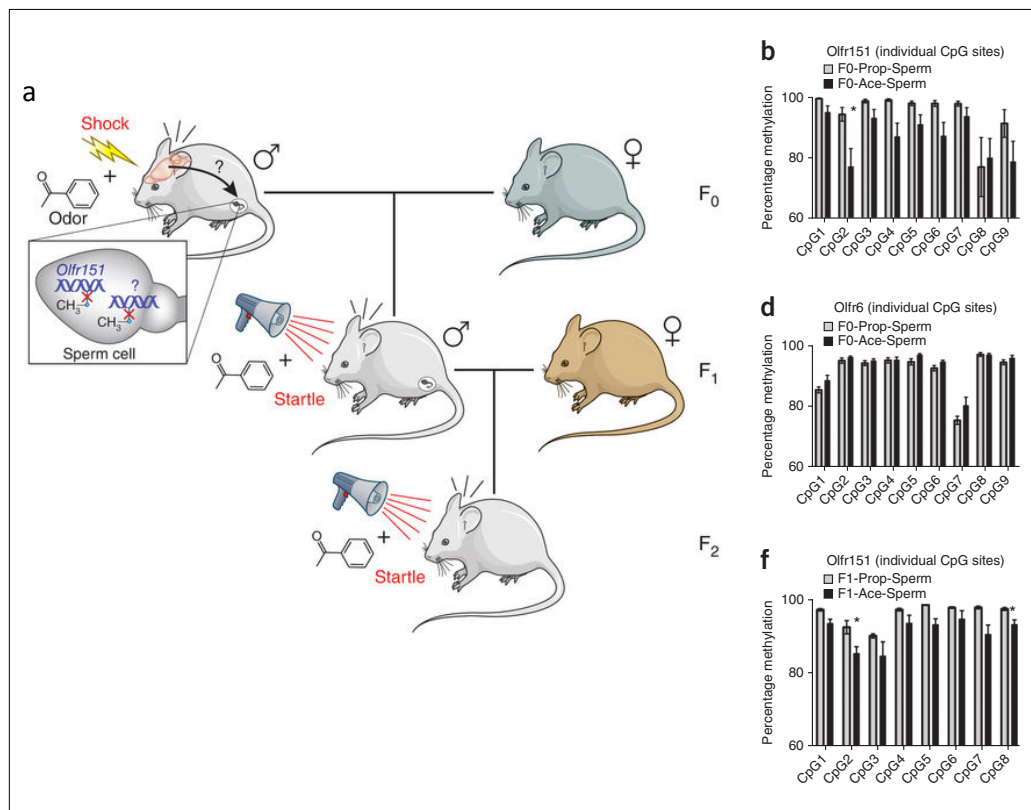


**Figure 20 : Impact du délaissement maternel sur la méthylation du récepteur aux glucocorticoïdes (GR).** Les souriceaux contrôles (High LG, en blanc sur l'histogramme) passent les premières semaines de leur vie à proximité immédiate de leur mère qui les lèche régulièrement. A l'inverse, les souriceaux délaissés (Low LG, en noir sur l'histogramme) vont être séparés quotidiennement de leur mère, empêchant de ce fait les soins maternels. Ces différences de traitement conduisent à des variations de méthylation de l'ADN sur le promoteur du gène codant pour le récepteur aux glucocorticoïdes (GR). L'histogramme de gauche représente en abscisse 15 CGs du promoteur du gène GR et en ordonnées leur pourcentage de méthylation. Sur certains CG, nous observons globalement une déméthylation du promoteur chez les souriceaux délaissés comparés aux souriceaux contrôles. \* différence significative.

### 3.2.3.7.3 Héritabilité transgénérationnelle

Son existence est débattue, mais semble de plus en plus probable, dans certains cas très particuliers et en petite proportion. C'est la transmission de marques épigénétiques des parents à leur enfant, rendue possible par le non effaçage de ces marques, pourtant la règle habituelle décrite précédemment dans le cycle de l'épigénome. Certaines marques de méthylation, en petit nombre, persistent inchangées ou presque dans l'embryon alors que les autres, en grande majorité sont effacées puis reprogrammées à partir de « zéro ».

Citons deux exemples d'une telle transmission. Après exposition à la vinclozoline, des anomalies urogénitales se rencontrent dans les générations suivantes, du fait des effets épigénétiques de ce pesticide (Anway et al., 2005). Un autre exemple, tout récent, est la transmission à la descendance sur deux générations de marques de méthylation engendrées par l'exposition à un conditionnement de peur via les gamètes paternels. De sorte que les petits enfants gardent une mémoire portée par des traces épigénétiques héritées du grand père rat (Figure 21).



**Figure 21: Influence des facteurs environnementaux sur l'épigénome de plusieurs générations. (a)** Le père de la génération F<sub>0</sub>, en haut à gauche, a été conditionné pour associer deux événements : l'odeur de l'acétophénone ou du propanol et un choc électrique au niveau de la patte. Il se crée ainsi un conditionnement de peur qui peut être induit uniquement par l'odeur sans choc électrique, une fois les souris entraînées. Ces mâles seront ensuite croisés avec des femelles non conditionnées pour donner naissance à la génération F<sub>1</sub>, qui donnera elle-même naissance à la génération F<sub>2</sub> après croisement avec des souris non conditionnées. Les mâles des générations F<sub>1</sub> et F<sub>2</sub> ont été exposés à l'acétophénone ou au propanol mais sans choc électrique associé. Bien qu'ils n'aient pas été exposés à un choc électrique, ils ont montré une réaction de peur exacerbée à l'odeur d'acétophénone lorsque leur père y avait été conditionné. Les souris issues de pères conditionnés au propanol n'ont pas montré de différence significative quant à elles, dans leur comportement, comparées aux souris issues de pères non conditionnés. Afin de tester comment ce comportement de peur pouvait être transmis de génération en génération suite à une exposition à l'acétophénone, les auteurs de cette recherche ont mesuré la méthylation de gènes codant pour des récepteurs olfactifs dans le sperme des souris F<sub>0</sub> (**b et d**) et F<sub>1</sub> (**f**). Ils ont ainsi montré des différences de méthylation significatives sur certains CGs entre les souris exposées à l'acétophénone (en noir) et celles exposées au propanol (en gris). Il s'agit de l'un des premiers exemples de transmission d'un comportement *via* l'épigénétique. Tiré de (Szyf, 2014)

### 3.2.3.8 Les fonctions de la méthylation de l'ADN

#### 3.2.3.8.1 *Contrôle de l'expression des gènes*

L'importance de cette marque située « au-dessus » du génome a été mise en évidence par les travaux de Riggs et Holliday en 1975 qui décrivent pour la première fois son implication dans la régulation de l'expression génique (Riggs, 1975). Son action inhibitrice a été démontrée d'abord dans un contexte artificiel. Les études des années 80 ont montré l'inhibition de la transcription d'un gène plasmidique sous le contrôle d'un promoteur artificiellement méthylé dans une lignée cellulaire (Stein et al., 1982). La réactivation de la transcription des gènes faisant suite à l'action d'un agent démethylant (5-aza-cytidine) a ainsi permis de montrer le rôle causal de la méthylation dans l'inhibition de la transcription (Mohandas et al., 1981).

La poursuite des études sur la méthylation de l'ADN a permis de déterminer les mécanismes moléculaires intervenant dans la régulation de l'expression génique (Figure 22). Ces mécanismes sont décrits dans les paragraphes suivants.

- En bloquant la fixation de certains facteurs de transcription sensibles à la méthylation d'ADN. Les séquences consensus de liaison de certains facteurs contiennent un ou plusieurs CG qui lorsqu'ils sont méthylés empêchent leur fixation. Par exemple le facteur SP1 ne peut pas se fixer sur son élément cis s'il est méthylé. D'autres exemples de facteurs de transcription obéissant au même mécanisme sont CREB, la famille des facteurs E2F (Campanero et al., 2000; Iguchi-Arigo and Ariga, 1989), etc.

- En recrutant des répresseurs et des co-répresseurs. Certaines protéines sont capables de se lier à l'ADN méthylé ce sont les « Methyl CpG binding Protein » telles que Mecp2 (Bogdanović and Veenstra, 2009). Mecp2 interagit avec des co-répresseurs complexes tel que Sina3a, NCo et c-Ski au niveau de leurs sites de fixations à l'ADN (Jones et al., 1998; Kokura et al., 2001; Nan et al., 1998). Ces co-répresseurs renforcent ainsi l'action de compaction de la chromatine et entraînent une inhibition totale de la transcription. Mais Mecp2 peut également interagir avec d'autres protéines afin de réprimer, activer, épisser ou modifier l'organisation nucléaire. Ce type de répression par des protéines ayant un domaine de fixation à l'ADN méthylé (MBD, Methyl Binding Domain) permet par exemple d'inhiber les éléments mobiles du génome humain.

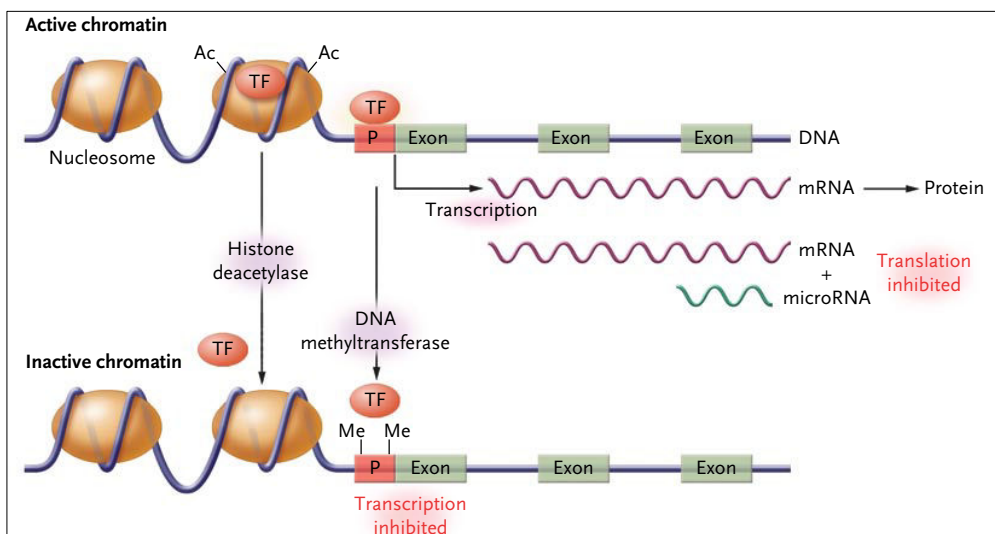
- La méthylation de l'ADN permet aussi une transcription efficace. La méthylation de l'ADN au niveau du corps du gène est associée à une transcription active. Les gènes hautement exprimés présentent un profil de méthylation particulier, une hypo-méthylation de leurs promoteurs et une hyper-méthylation intragénique (Ball et al., 2009). Les facteurs de transcription et l'ARN polymérase peuvent se fixer tout le long du génome. Leurs séquences consensus sont des séquences nucléotidiques courtes (6-20 nucléotides). Ces motifs en dehors des promoteurs sont méthylés, ils sont ainsi non accessibles à l'ARN Pol II et à la machinerie de transcription.

- L'établissement de l'empreinte génomique et l'inactivation du chromosome X. L'expression des gènes soumis à l'empreinte parentale est régulée par une méthylation allèle dépendante (*Allele Specific Methylation*, ASM) au niveau des ICR (Meaburn et al., 2010). L'inactivation du chromosome X est un mécanisme qui permet d'égaliser le dosage génique du chromosome X chez les mâles et les femelles, en réprimant la transcription d'un des deux chromosomes X dans les cellules femelle. Dans

ce mécanisme la méthylation de l'ADN stabilise un état condensé de la chromatine (Guibert and Weber, 2013).

- Méthylation et épissage alternatif. La méthylation des jonctions exon-intron peut jouer un rôle dans l'épissage alternatif. Les changements de méthylation de l'ADN dans ou à proximité des sites d'épissage sont associés aux choix de l'exon (Bonasio, 2012; Foret et al., 2012; Zhou et al., 2012). Le mécanisme qui lie la méthylation à l'épissage n'est pas clair. Le seul exemple décrit dans ce contexte est celui du facteur CTCF qui pourrait dans ce cas stopper l'action de l'ARN pol II (Shukla et al., 2011).

- Stabilité du génome. La méthylation d'ADN réprime la transcription des éléments transposables considérés comme « parasites » pour le génome. Leurs abondances et leurs fortes homologies de séquence donnent la faculté de changer de localisation dans le génome et d'augmenter les risques mutagènes. En effet, l'hypométhylation augmente la possibilité de recombinaison entre les régions homologues répétées ainsi que les régions télomériques (Bourc'his and Bestor, 2004). L'hypométhylation entraîne une instabilité génomique globale créant des aberrations et des réarrangements chromosomiques (Dodge et al., 2005) comme il l'a été montré chez l'homme dans le syndrome ICF (Immuno-déficience combinée) ou dans certains cancers (Xu et al., 1999).

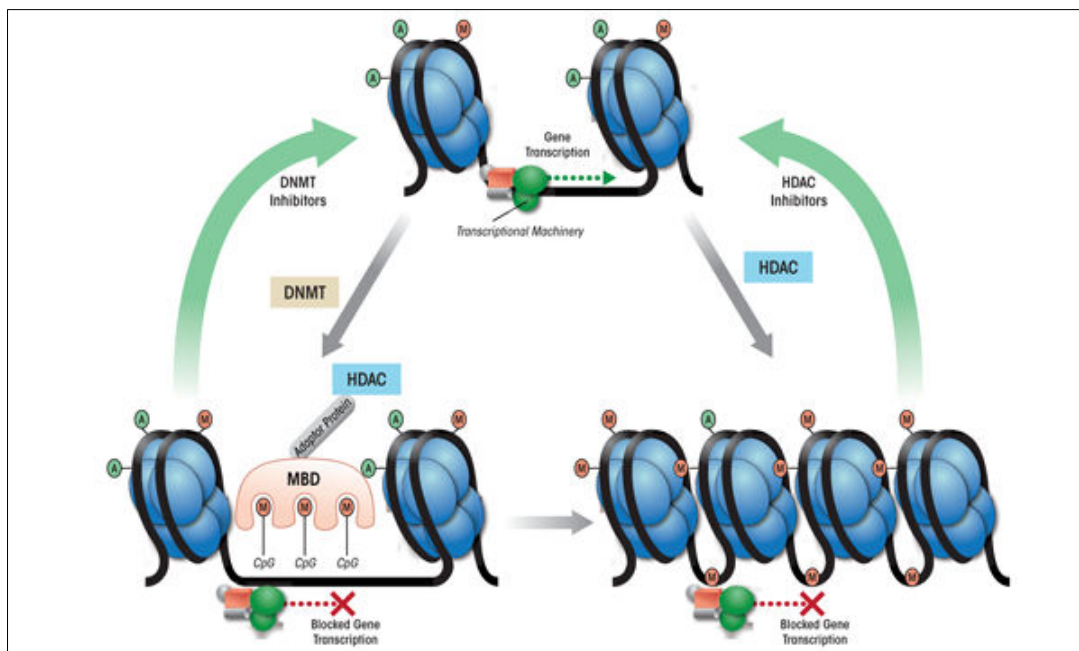


**Figure 22: Schématisation de l'ensemble de processus épigénétiques modulant l'expression des gènes.** L'acétylation des histones est associée à l'expression des gènes les facteurs de transcription noté (TF) sont capable de se fixer sur le promoteur du gène, la méthylation de l'ADN et les micro ARN inhibe la transcription. La méthylation dans ce cas bloque la fixation des facteurs de transcription sur le promoteurs en revanche les micro ARN sont capable de cibler les ARNm synthétisé et les dégrader et inhiber ainsi la traduction du gène.

### 3.2.4 Relations entre la méthylation de l'ADN et les histones

La méthylation de l'ADN est généralement inversement corrélée à l'acétylation des histones (H3K9ac et 27ac). En revanche, la méthylation de la lysine 9 de l'histone 3 est le plus souvent corrélé avec la méthylation d'ADN, assurant ainsi un état transcriptionnel inactif retrouvé généralement dans

l'hétérochromatine. Ceci prouve qu'il y a des mécanismes moléculaires reliant les deux marques épigénétiques. Aujourd'hui il est bien établi que les modifications des histones et la méthylation de l'ADN interagissent entre elles pour réguler l'expression des gènes. Cette interaction met en jeu divers complexes protéiques: i) les methyl binding Protein domain MBD (Bird and Wolffe, 1999) ; ii) les enzymes responsables des modifications post-traductionnelles des histones (méthyltransférase, déméthylases, acétylases et désacétylase...) (Ciccarone et al., 2012; Jelinic et al., 2006). En dehors des MBD, les enzymes DNMT peuvent également se lier à l'ADN méthylé et aux protéines de la chromatine. Ainsi l'interaction entre Dnmt1 et la méthyl-transférase spécifique de la marque H3K9me2 Ga/GLP est importante pour la méthylation *de novo* des rétrotransposons dans les cellules ES (Estève et al., 2006). Finalement, il est important de souligner que la méthylation de l'ADN et les modifications chimiques de la chromatine sont deux acteurs épigénétiques qui participent ensemble au maintien de l'état transcriptionnel d'une cellule donnée (Figure 23).



**Figure 23 : Exemple d'action coordonnée des marques épigénétiques de la chromatine et de l'ADN.** En haut, un gène est transcrit lorsque l'ADN est déméthylé et les histones acétylés. En bas à gauche le gène est réprimé par la méthylation de l'ADN qui a permis le recrutement des protéines MBD, la désacétylation des histones par les HDACs et la compaction de la chromatine à droite. Les ronds bleus représentent les histones, les ronds rouges la méthylation de l'ADN et les ronds verts l'acétylation des histones. Tiré de (Moonat et Pandey, 2012).

### 3.2.5 Les variations de méthylation de l'ADN

#### 3.2.5.1 Les régions génomiques qui présentent des différences de méthylation

Une région différenciellement méthylée (*Differentially methylated region*, DMR) peut se définir comme une région génomique ayant une méthylation CG différente entre divers échantillons. Il existe différents types de DMRs décrits dans les paragraphes suivants.

#### 3.2.5.1.1 *aDMRs (age DMRs)*

Dans un même tissu et chez un même individu, la méthylation de certaines régions génomiques peut varier au fil du temps (Talens et al., 2010; Wong et al., 2010). Des changements de méthylation liés à l'âge ont été identifiés dans plusieurs tissus qu'on appelle « age DMRs » ou aDMR (Bird and Wolffe, 1999; Bjornsson et al., 2008 ; Rakyan et al., 2010). Les changements de méthylation de l'ADN liés à l'âge sont associés à la sénescence cellulaire et à la longévité, mais les causes et les conséquences fonctionnelles de ces variations de méthylation restent inconnues. Certains auteurs postulent que ces changements de méthylation au fil du temps répondent à une accumulation d'exposition à des facteurs environnementaux capables d'agir sur la méthylation de l'ADN (Wong et al., 2010). Ils seraient également responsables en partie de l'apparition des maladies dites de l'adulte.

#### 3.2.5.1.2 *cDMRs (cancer DMRs)*

Les recherches dans le domaine du cancer montrent l'importance des défauts de la méthylation de l'ADN tels que l'hypométhylation des oncogènes et/ou l'hyperméthylation de gènes suppresseurs de tumeurs (Feinberg, 2004). Ces cDMRs sont généralement localisés dans les promoteurs à CGI et dans les régions génomiques entourant les CGI appelés CG Island shores (Weber et al., 2005).

#### 3.2.5.1.3 *tDMRs (tissu DMRs)*

Depuis sa découverte, la méthylation de l'ADN a été considérée comme le mécanisme responsable de la différence d'expression des gènes observée entre les différents types cellulaires pendant le développement embryonnaire (Riggs, 1975). Les tDMRs sont des régions génomiques différemment méthylées d'un tissu à l'autre et sont généralement responsables de l'expression tissu spécifique de certains gènes (Song et al., 2005, 2009). En 2005, les expériences de Song F sur six tissus différents de souris âgées de 12 semaines ont permis l'identification des premiers tDMRs. En 2009, la même équipe a établi le profil de méthylation des tDMRs à partir de différents tissus murins au cours du développement embryonnaire et des cellules embryonnaires. La méthylation observée dans les cellules ES est complètement différente de celle de cellules différenciées. La méthylation des tDMRs change au cours des stades embryonnaires. Les gènes déméthylés dans les tissus adultes, sont généralement méthylés durant les stades embryonnaires précoces puis perdent progressivement leur méthylation dans les stades embryonnaires avancés. L'établissement des tDMRs est un processus dynamique qui se poursuit jusqu'au stade embryonnaire tardif (Song et al., 2005).

#### 3.2.5.1.4 *Les variations individuelles de la méthylation*

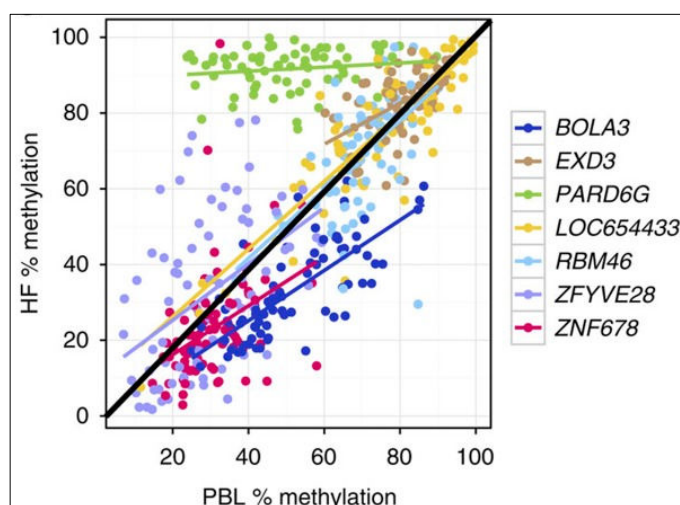
Dans le cadre d'un projet d'analyse de l'épigénome (Human Epigenome Project HEP : <http://www.epigenome.org/index.php?page=pilotproject>), une variabilité dans la méthylation de l'ADN entre individus a été mise en évidence (Rakyan et al., 2004). Contrairement aux autres régions différenciellement méthylées déjà citées dans les paragraphes précédents, la terminologie des variations individuelles de méthylation n'est pas unique. Chaque auteur a utilisé sa propre nomenclature.

**Épiallèles.** Un épiallèle est une région génomique subissant des modifications épigénétiques variables entre les individus au sein d'une population naturelle. Ces modifications épigénétiques peuvent toucher la méthylation de l'ADN ou les modifications des histones. Les épiallèles sont



stables et maintenus au cours des divisions cellulaires mitotiques. Les épiallèles sont capables d'influencer un phénotype en modulant l'expression des gènes. Les exemples illustrant ce mécanisme sont nombreux dans le monde végétal où les épiallèles sont largement décrits (Zilberman et al., 2007). Un épiallèle peut être une conséquence d'un événement génétique, environnementale ou stochastique. Les épiallèles peuvent donc être classés selon l'évènement impliqué dans son apparition : on peut distinguer des épiallèles « obligatoires », « facilités » et « purs » (« Obligatory » « Facilitated » and « Pur » epiallele) (Richards, 2006). Les épiallèles « obligatoires » et « facilités » sont issus d'un évènement génétique, contrairement aux épiallèles « purs » dont l'apparition est liée à un évènement stochastique ou environnemental indépendamment des variations génétiques. Ces évènements seront détaillés dans un paragraphe suivant sur les causes de la variabilité de la méthylation de l'ADN.

**Epiallèles métastables.** Le mécanisme moléculaire qui régule l'expression du gène *Agouti* chez la souris et ses conséquences phénotypiques ont introduit la notion d'épiallèle dans le monde animal (Argeson et al., 1996; Siracusa et al., 1995). L'insertion spontanée d'un rétrotransposon IAP non méthylé en amont du gène *Agouti* entraîne une couleur jaune du pelage et une prédisposition au diabète. La méthylation de cet IAP restaure le phénotype et donne une couleur brune au pelage. Plus tard, Rakyan et son équipe ont proposé le terme d'épiallèle métastable (« metastable epiallele » : ME) pour caractériser les épiallèles mis en place pendant les stades embryonnaires précoces et de ce fait présents dans toutes les cellules de l'organisme (Rakyan et al., 2002). Ces épiallèles métastables peuvent être définis également comme des régions génomiques subissant des modifications épigénétiques variables et réversibles, se produisant dans des cellules génétiquement identiques, et entraînant des changements phénotypiques (Jirtle and Skinner, 2007). Des études récentes du méthylome ont permis l'identification d'épiallèles métastables chez l'homme (Figure 24) (Dominguez-Salas et al., 2014; Harris et al., 2013; Waterland et al., 2006). Ces régions ME ont été validées dans deux tissus d'origine embryonnaire différente. Ils peuvent être en dehors des éléments répétés et des IAP, régions décrites pour les premiers épiallèles, et sont présents tout le long du génome (5'UTR, 3'UTR, premier exon, intron, exons) (Harris et al., 2013).



**Figure 24 : Exemples d'épiallèles métastables chez l'homme.** Ces épiallèles métastables (ME) ont été mis en évidence dans une population gambienne soumise à des variations saisonnières d'apport en acide folique. Ce graphique représente la méthylation des MEs dans 2 tissus : les cellules sanguines (PBL) d'origine mésodermique et les cellules folliculaires des cheveux (HF) d'origine ectodermique. La méthylation pour chacun de ces MEs dans ces tissus est fortement corrélée. Chaque couleur représente un ME.



**VMR (Variably methylated region).** Feinberg et collaborateurs ont étudié 74 personnes avec une puce « CHARM » analysant 4,5 millions de CGs et incluant des îlots CG et des régions de moindre densité CG provenant d'échantillons d'ADN extraits des lymphocytes circulants. Ils ont identifié 227 régions du génome humain qui montraient une extrême variabilité individuelle. Chaque VMR était définie comme une région où au moins 10 sondes consécutives atteignaient des valeurs de  $s_{jt} < 99^{\text{ème}}$  percentile de tous les  $s_{jt}$  et un  $s_{jt}$  moyen  $> 0.125$ . La valeur de  $s_{jt}$  est calculée à partir de la médiane de  $|M_{ijt} - m_{jt}|$  des sujets, avec  $m_{jt}$ , la médiane de  $M_{ijt}$  parmi les sujets  $i$ ,  $M_{ijt}$  étant le pourcentage de méthylation du sujet  $i$  pour le paramètre de la puce au temps  $t$  (Feinberg and Irizarry, 2010). La moitié de ces VMR était stable sur une période de 11 ans, les sujets ayant été étudiés durant cet intervalle de temps. Cet article représente une des premières définitions d'une signature épigénomique personnalisée chez l'homme, et de sa plasticité au fil du temps, fondée sur les variations les plus extrêmes (et les plus indiscutables de ce fait) de la méthylation des CGs de notre génome. Mais ces variations extrêmes sont loin de résumer toute la gamme des variations individuelles, qui peuvent être moins extrêmes et néanmoins biologiquement importantes. La méthode ne prend pas non plus en compte la variation individuelle distincte de chaque CG. Les auteurs rapportent une association de 4/119 VMRs stables avec l'adiposité des sujets, évaluée avec l'index de masse corporelle (IMC). L'étude ne dit pas si la variabilité de ces 4 VMRs était une cause de la variation de l'adiposité ou une conséquence de celle-ci (ou de facteurs associés à l'obésité comme la nutrition). C'est souvent une des grandes limites des observations de l'épidémiologie épigénétique.

**iiDMR.** Gemma et ses collaborateurs ont étudié 3 vrais jumeaux avec une puce Illumina 450K, interrogeant 369,908 CGs contenus dans les promoteurs, les enhancers et les CGI provenant d'ADN extrait, dans des monocytes CD14+. Les auteurs ont défini les iiDMRs par une différence de méthylation supérieure à 5 % dans au moins 2 CGs adjacents à moins de 500 pb l'un de l'autre. Ils ont identifié 2,528 à 5,077 régions différentiellement méthylées intra-jumeau et 9,062 à 10,526 régions différentiellement méthylées inter-jumeaux, comparé deux à deux. Les iiDMRs retrouvés au niveau des promoteurs sont inversement corrélés avec la marque de chromatine active H2AZ (Barski et al., 2007; Thambirajah et al., 2006). Les iiDMR identifiés sont stables sur une période de deux ans. D'après les auteurs, les iiDMRs sont situées au niveau des promoteurs des gènes non exprimés ou faiblement exprimés dans les CD14+. Ce travail représente la première étude qui a intégré à la fois des analyses de l'épigénome, du génome et le niveau d'expression génique. Grâce à cette stratégie, Gemma et ses collaborateurs ont mis en évidence des régions différentiellement méthylées indépendamment des effets en cis des variants génétiques. Ils suggèrent de classer les iiDMR en épiallèle, en se basant sur le fait que les iiDMR sont des régions exposées à des perturbations chromatiniennes. En revanche, ils n'ont pas réussi à montrer un lien entre des iiDMRs situés dans les promoteurs et le niveau d'expression ni avec un phénotype. Il est possible que le niveau de couverture des puces 450K ne soit pas suffisant pour détecter toutes les régions différentiellement méthylées notamment dans les LCP. Les iiDMRs retrouvés entre jumeaux monozygotes peuvent être classés dans la catégorie des épiallèles « purs » d'origine stochastique ou environnementale.

### 3.2.5.2 Causes de la variation de la méthylation de l'ADN entre individus

Nous avons développé ce paragraphe car la variation de la méthylation entre individus, objet d'études épidémiologiques, reste largement inexplicée.

### 3.2.5.2.1 *La variation génétique*

Une mutation ou un polymorphisme génétique peut faire varier la méthylation de l'ADN. Les exemples les plus connus sont cités ci-dessous.

**Une délétion.** Des délétions, peuvent se produire dans les ICRs des loci soumis à l'empreinte parentale (*IGF2-H19*, *GNAS*) et entraîner des changements de méthylation dans les régions voisines (Feinberg, 2007).

**Une mutation.** La mutation d'un gène affectant la machinerie de méthylation peut avoir de lourdes conséquences. L'exemple le mieux décrit est celui de la mutation inactivatrice de la protéine *Mecp2*, qui devient alors incapable de reconnaître les cytosines méthylées ou les complexes protéiques co-répresseurs et perd ainsi son action répressive. Les conséquences phénotypiques associées aux mutations du gène *Mecp2* sont généralement des défauts dans le développement et des troubles neurologiques sévères, connus sous le nom de « Syndrome de Rett » (Bienvenu and Chelly, 2006; Feinberg, 2007).

**Un polymorphisme.** Une mutation peut avoir lieu au niveau d'un dinucléotide CG créant ainsi une perte de la méthylation puisque ce polymorphisme conduit de fait à la perte du CG, principale cible de la méthylation chez les mammifères.

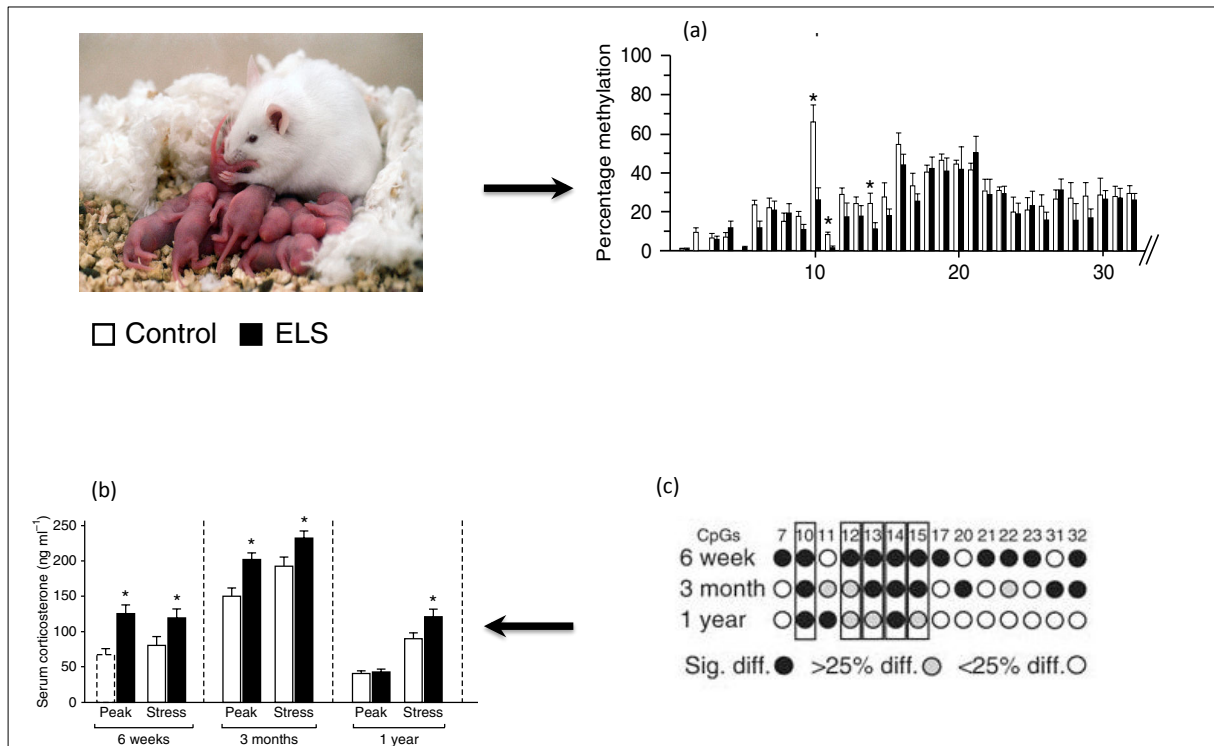
Un polymorphisme peut aussi présenter une association préférentielle avec la méthylation d'un ou plusieurs CGs, appelés ASM (*Allele Specific Methylation*). Dans ce cas, la méthylation est déterminée par la séquence d'ADN en *cis* et suit de fait un mode de transmission mendélien. Il existe ainsi des associations préférentielles entre les allèles de certains SNPs et la méthylation de certains CGs (Fradin et al., 2012). Des études récentes ont montré que ces ASMs étaient répartis sur l'ensemble du génome et pouvaient présenter des différences entre les tissus. L'ASM est un modèle qui illustre l'interaction entre la méthylation et les variants génétiques. Cette interaction existe sous forme *cis* ou *trans*. Les ASMs en *cis* sont les plus répandus dans le génome humain. La méthylation d'ADN peut varier selon une action *cis* de certains polymorphismes au niveau de dinucléotides CGs distants d'environ 1 à 3 kb (Banovich et al., 2014). Par conséquent, certains génotypes peuvent s'associer à un état de méthylation différent (Bell et al., 2011; Kerkel et al., 2008; Gibbs et al., 2010). Les ASM sont de plus en plus étudiés et au moins 3,500 sites ont ainsi été identifiés sur le génome humain (Schalkwyk et al., 2010). Une minorité des ASMs est issue des effets *trans* de facteurs tels que les événements stochastiques ou des loci soumis à empreinte. Manolio et ses collaborateurs pensent que ce type d'ASM pourrait être responsable de l'héritabilité manquante de certaines maladies multifactorielles (Manolio et al., 2009). Le mécanisme moléculaire qui lie les variants génétiques à la méthylation de l'ADN n'est pas encore connu. En revanche, les ASMs semblent associées à des variations de l'accessibilité de la chromatine et de l'expression des gènes (Banovich et al., 2014).

### 3.2.5.2.2 *Les facteurs environnementaux*

Des facteurs chimiques, physiques ou nutritionnels sont capables de modifier l'expression des gènes et le phénotype correspondant (Anway et al., 2005; Sutherland and Costa, 2003; Waterland, 2003). L'action de ces facteurs environnementaux sur la transcription est médiée par des modifications de la méthylation de l'ADN.

L'établissement de la méthylation est un processus dynamique pendant lequel les modificateurs environnementaux sont potentiellement capables de changer l'état de méthylation dans certaines régions génomiques comme les épiallèles. Chez les mammifères, les deux vagues de déméthylation et reméthylation du génome se font de façon programmées et prévisibles. En revanche, le rétablissement des marques épigénétiques des épiallèles métastables est un événement non régulier. En effet, la marque épigénétique réétablie après la déméthylation globale peut être différente de celle du début. On ne sait pas pourquoi ces marques épigénétiques sont capables de dériver ou de changer. Il est possible que des fluctuations de facteurs environnementaux soient à l'origine de ces changements épigénétiques. Le régime alimentaire pendant la gestation, par exemple, est un facteur déterminant de l'établissement de la méthylation de l'ADN du fœtus. En effet, dans une cellule de mammifère, la disponibilité en méthionine est fortement modulée par l'alimentation (Figure 12). Plusieurs études chez la souris ont montré qu'un régime alimentaire riche en acide folique, donneur de méthionine, permet une méthylation efficace de l'IAP du gène *Agouti* chez le fœtus, se traduisant par un pelage brun de la progéniture, tandis qu'un régime alimentaire pauvre en acide folique conduit à un pelage jaune (Waterland, 2014). Des études épidémiologiques chez l'homme ont également mis en évidence l'influence du régime alimentaire durant la grossesse sur le profil de méthylation d'ADN de l'enfant (Dominguez-Salas et al., 2014). Certains substrats (l'acide folique, la méthionine, l'homocystéine, la choline... Figure 12) représentent à la fois un indice de méthylation et un indice d'apport nutritionnel. Les niveaux de ces marqueurs mesurés dans le sang des femmes gambiennes enceintes changent selon l'apport nutritionnel de la saison. Ces changements semblent affecter la méthylation des épiallèles métastables. Les enfants conçus dans la saison pluvieuse sont plus méthylés que ceux conçus dans la saison de sécheresse. Les différences de méthylation sont retrouvées dans deux tissus différents : les cellules du sang et les cellules de follicules pileux (Figure 24).

L'environnement postnatal est aussi capable de faire varier la méthylation de l'ADN en certains sites. Les études de Meaney et Spengler montrent que des événements stressants pendant la vie foetale peuvent laisser «une trace » sur l'épigénome des individus. Ainsi les premières expériences sociales peuvent influencer de façon plus ou moins pérennes la manière dont les gènes sont exprimés, et donner lieu à une mémoire biologique pouvant avoir des répercussions à vie (Figure 25) (Murgatroyd et al., 2009). Meaney et ses collaborateurs ont ainsi montré que le délaissement maternel pendant les premières semaines de vie modifiait les marques épigénétiques du promoteur du gène codant pour le récepteur aux glucocorticoïdes (Figure 20). Ces changements touchent à la fois la méthylation de l'ADN et les modifications post-traductionnelles des histones, entraînant ainsi des conséquences phénotypiques importantes se manifestant par un défaut de réponse au stress durant la vie adulte. Notons toutefois que ces modifications épigénétiques observées sont réversibles par une injection d'histone déacétylase. Les marques épigénétiques associées à ces facteurs environnementaux peuvent affecter des tissus et des organes multiples (Teh et al., 2014).



**Figure 25: Influence du comportement social précoce sur l'épigénome.** Le comportement maternel imprime durablement le génome de sa progéniture. Sont comparées des souris contrôles qui auront des soins maternels classiques et des souris qui seront séparées quotidiennement et aléatoirement de leur mère, créant ainsi une carence de soins maternels engendrant un stress (« ELS » pour Early Life Stress,) (a) Histogramme de la méthylation du promoteur du gène AVP (Arginine Vasopressin) mesurée dans le noyau PVN de l'hypothalamus. Les CGs testés sont numérotés de 1 à 33 en abscisses. Les souris ELS présentent une hypométhylation significative de certains CGs. (b) Persistance des effets sur la réponse au stress. Les souris ELS restent hyper-stressables à l'âge adulte. Le comportement social précoce de la mère laisse donc une empreinte durable sur sa descendance (c) Mesure de la méthylation des CGs significativement différents en (a) un an après la fin de l'exposition au stress précoce. Les différences de méthylation persistent entre les souris contrôles et les souris ELS. Tiré de (Murgatroyd et al., 2009)

Il est possible également que pendant la vie adulte, l'environnement fasse varier la méthylation de l'ADN en certains sites. Le mécanisme reste inconnu mais certaines études épidémiologiques montrent des associations très significatives entre la méthylation et un facteur environnemental comme le tabagisme (Breitling et al., 2011).

### 3.2.5.3 Epidémiologie de la variation de la méthylation

En 2013, Mill a redéfini l'épidémiologie-épigénétique par l'intégration des analyses épigénétiques dans la recherche épidémiologique, dans le but d'identifier à la fois les causes et les conséquences phénotypiques des variations épigénomiques.

Actuellement, la majorité des études épidémiologiques se basent sur l'identification de DMRs entre un groupe de patients et un groupe d'individus sains. Les DMRs commencent à apparaître dans la littérature médicale. Les études qui visent à associer les différences du méthylome à des pathologies humaines sont dites « épidémiologie-épigénétique ». Ces études restent étroitement liées aux études épidémiologie-génétiques à cause des interactions entre les variants génétiques et variants épigénétiques.

Avec le progrès technologique et la diversification des stratégies d'étude du méthylome, le débat concernant l'utilité et l'interprétation des données issues de ces études est actif.

Citons deux exemples d'études épidémio-épigénétiques :

- dans une étude publiée en 2013, Delgado-Calle et ses collègues ont comparé les profils de méthylation du génome entier de femmes atteintes d'ostéoporose et de femmes souffrant d'arthrose. Ils ont détecté environ 250 DMRs. Ces DMRs sont situées principalement dans des régions codants pour des facteurs de transcription impliqués dans la différenciation cellulaire et la formation du squelette (Delgado-Calle et al., 2013).
- Jeffries et ses collaborateurs ont détecté des différences de méthylation dans les lymphocytes T CD4+ de patients sains et de patients atteints de lupus, une maladie auto-immune. Ils ont identifié 341 DMRs, dont 236 étaient moins méthylés et 105 plus méthylés chez les patients. De nombreux sites étaient localisés dans des gènes ayant un rôle dans l'auto-immunité (Jeffries et al., 2011).

Le problème majeur de ces études reste de savoir si ces DMRs sont une conséquence ou une cause de la maladie.

## 4 LA CROISSANCE DU SQUELETTE DANS L'ENFANCE

### 4.1 LA TAILLE

La croissance staturale humaine reflète le développement du squelette de différents segments du corps : la tête, le tronc, le bassin et les membres inférieurs.

La croissance postnatale de l'homme se compose de 3 phases: une croissance très rapide de 0 à 4 ans, une vitesse de croissance intermédiaire de 4 ans jusqu'à la prépuberté et enfin une poussée dite pubertaire jusqu'à la taille finale.

### 4.2 HISTORIQUE ET GENERALITES

L'insulin-like growth factor 1 (IGF1) a été identifié pour la première fois par Salmon et Daughaday en 1957 (Salmon and Daughaday, 1957). Ils ont montré que cette protéine stimule l'incorporation du sulfate par le cartilage de croissance du rat. Pour cette raison la protéine a été nommée « facteur de sulfatation ». Au début des années 1970, les travaux menés sur cette protéine ont montré sa physiologie pléiotrope et sa capacité de médier les effets de l'hormone de croissance (Growth Hormone, GH), d'où son nom « somatomédine » (Laron, 2001; Salmon and Daughaday, 1990). La somatomédine a été rangée dans la famille des facteurs de croissance. En 1978, l'équipe de Rinderknecht a caractérisé sa structure protéique et montré sa forte homologie avec la proinsuline. En effet, les domaines B, C et A sont identiques à ceux de la proinsuline, (Rinderknecht and Humbel, 1978) d'où sa nomenclature « insulin-like growth factor 1 ».

Aujourd'hui, les rôles de la protéine IGF1 dans les processus de biologie cellulaire sont bien caractérisés : l'IGF1 intervient dans la prolifération, la différenciation, la migration et la survie de nombreux types cellulaires (Jones and Clemmons, 1995; Koutsilieris et al., 2000; Siddle et al., 2001). Selon le tissu, l'IGF1 exerce ses effets selon un mode endocrine, autocrine et/ou paracrine.

### 4.3 PHYSIOLOGIE DE L'AXE GH/IGF1

L'axe GH/IGF est une voie endocrine clé de la croissance postnatale. La physiologie de cette voie obéit à une boucle de plusieurs rétrocontrôles. L'hormone de croissance (GH) est produite par les cellules somatotropes de la partie antérieure de l'hypophyse et libérée par pulses dans la circulation. Les pulses de GH sont générés sous l'effet de deux peptides hypothalamiques. La somatolibérine ou GHRH (Growth Hormone releasing hormone) induit la sécrétion de GH. La somatostatine GHIH (Growth hormone inhibiting hormone) ou SRIF (Somatotropin release-inhibiting factor) l'inhibe. La sécrétion pulsatile de la GH est due à l'alternance des effets de GHIH et GHRH, eux mêmes sécrétés de façon pulsatile (Bick et al., 1992; Tannenbaum and Ling, 1984). Des informations venant du cerveau tels que le sommeil et le stress ou de la périphérie tels que l'hypoglycémie et le taux des acides aminés sont intégrées par des neuromédiateurs pour stimuler la production du GHRH. D'autres facteurs agissent sur la production des hormones hypothalamiques comme l'exercice physique, la nutrition.

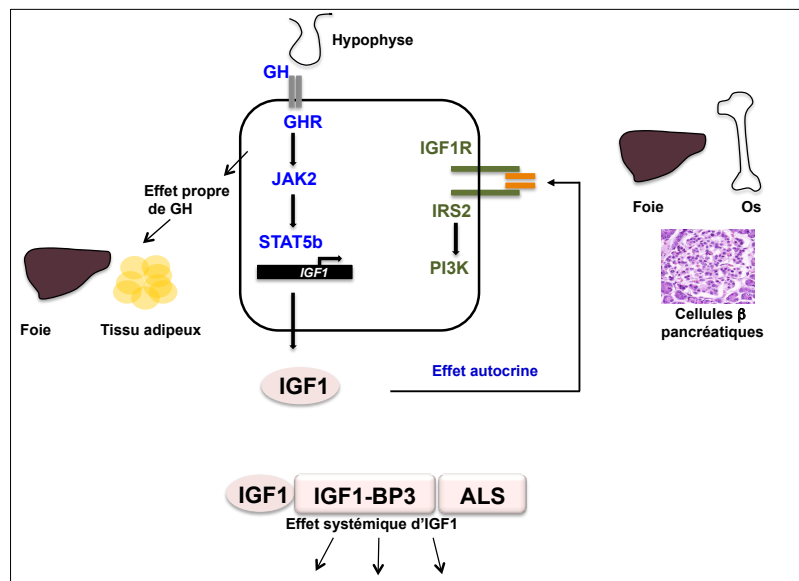
Une fois l'hormone de croissance libérée dans la circulation, elle est transportée vers ses tissus cibles, presque toutes les cellules de l'organisme, notamment le foie et les cartilages. La GH exerce des

effets pléiotropes sur la différenciation et la prolifération cellulaires et le métabolisme (Isaksson et al., 1985, 1987). Quelques uns de ces effets répondent à une action directe de la GH, d'autres effets sont induits par d'autres facteurs médiateurs de la GH comme l'IGF1.

La GH augmente la production d'IGF1 dans de nombreux tissus et par suite dans la circulation. Le transport d'IGF1 dans la circulation sera détaillé au paragraphe 4.6. L'augmentation du taux d'IGF1 circulant inhibe la sécrétion hypophysaire de GH (Clark et al., 1996). L'administration d'IGF1 réduit la sécrétion de GH. L'absence d'IGF1 circulant chez les souris KO mutante IGF1-/- entraine une augmentation de la sécrétion de GH (Liu and LeRoith, 1999). La GH exerce directement un rétrocontrôle négatif sur GHRH-GHIH ou indirectement *via* l'action d'IGF1 libre, non liée à ses protéines porteuses (Wallenius et al., 2001).

### **Signalisation cellulaire de la GH**

Les actions cellulaires de la GH sont déclenchées par la liaison à son récepteur transmembranaire, le GHR (GH Receptor). L'activation du GHR déclenche plusieurs cascades de signalisation impliquant les voies JAK/STAT (STAT pour Signal transducers and activators of transcription), MAP (pour Mitogen-Activated Protein) kinases et IRS/PI3K (pour Insulin Receptor Substrate et pour Phosphatidylinositol-3- Kinase). Ces trois voies permettent à l'hormone de croissance d'induire à la fois ses effets transcriptionnels et métaboliques (Figure 26).



**Figure 26 : Voie de signalisation de l'hormone de croissance.** L'hormone de croissance (GH) se fixe sur son récepteur pour exercer directement ses effets sur le foie et le tissu adipeux. Elle agit également via la synthèse d'IGF1 qu'elle stimule, sur le tissu osseux et cartilagineux, le foie et les cellules beta pancréatiques. IGF1 dans le sang se lie aux protéines de transport IGF1-BP3 et ALS.

### **La voie de signalisation JAK/STAT5**

Une fois fixée sur son récepteur, la GH entraîne un changement conformationnel du GHR et un rapprochement physique des molécules JAK2 conduisant à leur phosphorylation. Ainsi activé, JAK2 phosphoryle à son tour plusieurs tyrosines du GHR, formant des sites d'ancrage pour plusieurs

protéines, dont les Stats (Argetsinger and Carter-Su, 1996). STAT5B est un des principaux effecteurs de la signalisation de la GH. Il est présent sous forme inactive monomérique dans le cytoplasme. STAT5b se fixe sur ses sites d'ancrage spécifiques entraînant ainsi sa propre phosphorylation. STAT5b phosphorylé se dimérise et est transporté vers le noyau où il active la transcription de plusieurs gènes, notamment IGF1 et ALS (Acid Labile Subunit). La voie JAK/STAT induit l'expression de la famille des gènes *SOCS* (Suppressor Of Cytokine Signaling), qui vont contribuer à la régulation négative de cette même voie. En effet, les protéines *SOCS* exercent un rétrocontrôle négatif à deux niveaux en interférant avec l'activation de la kinase JAK2 et avec l'interaction des STATs avec le GHR (Greenhalgh et al., 2005; Strous et al., 2004).

#### **4.4 LES MALADIES GENETIQUES MENDELIENNES (MONOGENIQUES) DE LA PRODUCTION D'IGF1**

Nous ne parlerons pas des causes acquises (tumeurs, infiltrations) de défaut de sécrétion de l'hormone de croissance, ni des causes acquises de résistance à l'hormone de croissance (dénutrition, maladies inflammatoires, etc). Nous citerons seulement les causes génétiques monogéniques altérant la production d'IGF1 et la croissance.

##### **4.4.1 Défaits de production de GH**

Le déficit en hormone de croissance peut être isolé ou associé à d'autres déficits hormonaux anté-hypophysaires.

###### **4.4.1.1 Déficit en GH isolé (GHDI)**

Son incidence est de 1/4000 à 1/10000 naissances. La majorité des cas sont sporadiques. Les cas familiaux représentent, selon les séries, 3 à 30%. Les mutations responsables touchent les gènes codant pour la GH (*GH1*) ou le récepteur de la *GHRH* (growth hormone releasing hormone, *GHRH-R*). 21 mutations du *GHRH-R* et une soixantaine de mutations de *GH1* sont connues (Alatzoglou and Dattani, 2010). Aucune mutation n'a été décrite au niveau du gène *GHRH*. Les GHDI sont classés selon la mutation, son mode de transmission et le phénotype clinique des patients.

###### **4.4.1.1.1 GHDI type I A et B**

Les patients du groupe A ont tous un retard statural très sévère  $\leq -4,5$  SDS depuis l'âge de 6 mois et une concentration sérique de GH non détectable. Les patients du type B ont des caractéristiques phénotypiques moins accentués, un retard statural moins sévère, et un niveau de GH détectable. Les patients de type IB ont une bonne réponse au traitement par l'hormone de croissance recombinante (rhGH) contrairement à certains patients de type IA qui développent des anticorps contre la rhGH.

###### **4.4.1.1.2 GHDI type II**

Les patients présentent des mutations faux-sens, introniques et des mutations des sites d'épissage uniquement sur le gène *GH1* (Alatzoglou and Dattani, 2010).



#### 4.4.1.1.3 *GHDI type III*

Il s'agit de mutations récessives au niveau du chromosome X. Seulement quelques mutations ont été caractérisées dans ce groupe de patients, comme par exemple le saut d'exon du gène *BTK* (Bruton Tyrosine Kinase), régulateur clé dans le développement des lymphocytes B, entraînant une agammaglobulinémie liée à l'X (Stewart et al., 2008). Une duplication de 33 nucléotides entraîne une expansion de 11 polyalanines. Phénotypiquement la mutation entraîne un retard mental et des anomalies crânio-faciales associées à la déficience en GH (Alatzoglou and Dattani, 2010).

#### 4.4.1.2 Déficits en GH associés à d'autres déficits hormonaux

Des mutations des gènes codant pour les facteurs de transcription hypophysaires tels que *PIT-1*, *PROP1*, *LHX3*, *HESX1* peuvent entraîner des déficits de sécrétion de GH associés à d'autres déficits hormonaux antéhypophysaires (TSH, LH, FSH, ACTH).

### 4.4.2 **Les insensibilités (résistances) génétiques à l'hormone de croissance (GHI : GH insensitivity)**

Elles résultent d'un défaut de signalisation de GH. Les causes de l'insensibilité à la GH sont multiples. Le déficit en IGF1 circulant associé à une sécrétion normale ou élevée d'hormone de croissance est une caractéristique commune aux patients insensibles à la GH. La taille varie de -2DS à -9DS (Burren et al., 2001). Les défauts moléculaires décrits sont des mutations des gènes de la voie de signalisation de GH : *GHR*, *IGF1*, *STAT5B*, *IGF1R*, *PTPN11* (David et al., 2011).

#### 4.4.2.1 Défaut du gène *GHR*

Le premier décrit est le syndrome de Laron, autosomique récessif, dû à une mutation du *GHR*, qui entraîne un nanisme sévère (-4 à -12 DS), avec une GH normale voire élevée et une IGF1 basse (Laron, 2001; Laron, 2004). Plus de 70 mutations homo- et hétérozygotes composites du gène *GHR* ont été rapportées (David et al., 2011). L'ensemble des mutations du gène *GHR* altèrent soit i) l'expression du *GHR*, ii) sa dimérisation fonctionnelle iii) son site de liaison avec la GH, iiiii) son domaine de transmission du signal. Le phénotype des porteurs de ces mutations varie d'un retard statural modéré jusqu'aux formes extrêmes du syndrome de Laron (-12 DS) (Burren et al., 2001; Diniz et al., 2008). La variabilité phénotypique existe même chez les porteurs de la même mutation. Par exemple, pour la mutation du site d'épissage E180 du *GHR*, la taille adulte varie de -5,3 à -12 DS (Rosenbloom, 1999). De même, les tailles varient de -3,3 à -6 DS chez les porteurs d'une mutation intronique ( $A_{-1} \rightarrow G_{-1}$ ) du gène *GHR* provoquant l'apparition d'un pseudo-exon supplémentaire entre l'exon 6 et 7 (Metherell et al., 2001). Les taux d'IGF1 et d'IGFBP3 sériques sont corrélés avec les tailles des patients (SDS) (Woods et al., 1997).

#### 4.4.2.2 Défaut du gène *STAT5B*

*STAT5B* est un facteur de transcription appartenant à la cascade de signalisation intracellulaire du récepteur de la GH. Les patients ayant un défaut de *STAT5B* présentent un déficit sévère de la croissance postnatale avec une résistance totale à la GH, associée à un déficit immunitaire (Kofoed et al., 2003). La taille varie entre -5 et -10 DS. Les patients ont des taux de GH élevés, GHBP normale, des taux d'IGF1, IGFBP-3 et d'ALS très faibles et n'augmentant pas après injection de rhGH. En 2003, la première mutation homozygote affectant *STAT5B* a été décrite. Une mutation faux sens affecte un

résidu du domaine structural appelé SH2 (pour Src Homology 2) et empêche les STATs d'interagir avec les tyrosines phosphorylées du récepteur. Il s'agit d'une substitution d'un acide aminé (Alanine en Proline) dans le domaine SH2 très conservé et critique pour la phosphorylation de STAT5B sous stimulus GH. Ainsi, la cascade JAK2/STAT5 est interrompue (Kofoed et al., 2003).

Sept autres mutations homozygotes dans différents domaines de la protéine, comprenant des mutations faux sens, non sens, indel et des petites délétions de *STAT5B* sont décrites (David et al., 2011). Les défauts indel changent le cadre de lecture et génèrent une protéine tronquée. Certaines mutations indel affectent le domaine de liaison avec l'ADN de la protéine STAT5B (David et al., 2011).

#### 4.4.2.3 Défaut du gène *IGF1*

Un retard sévère de la croissance fœtale est associé à un retard de croissance postnatal, un retard mental sévère, une microcéphalie, une surdité, un retard pubertaire et une adiposité (David et al., 2011; Netchine et al., 2009). Le taux de GH est élevé, l'IGF1 est indétectable (Woods et al., 1996) ou au contraire très augmenté dans le cas particulier d'un IGF1 bio-inactif (Denley, A., 2005). Les taux d'IGFBP-3, l'ALS et l'IGF2 sont normaux ou augmentés (Denley et al., 2005; Woods et al., 1996).

Six défauts du gène *IGF1* ont été identifiés : des délétions, des mutations non sens ou alors une duplication de nucléotides (David et al., 2011). La délétion de l'exon 4 et 5 fut la première mutation identifiée dans le gène *IGF1* par (Woods et al., 1995). La traduction du gène muté donne une protéine tronquée dépourvue de 45 acides aminés. En dehors de la région codante pour la protéine E3-4, les mutations qui touchent le domaine E du peptide et les sites d'épissage entraînent des protéines tronquées (ou bio-inactives). C'est le cas de la transversion d'un A en T de la région non traduite de l'exon 6, affectant le domaine E du précurseur IGF1 et de la mutation faux sens (IGF1V44M) à l'état homozygote, découverte par l'équipe de Walenkamp, il s'agit d'une substitution d'une valine par une méthionine (Walenkamp et al., 2013). Ce changement d'acide aminé affecte l'affinité de l'IGF1 à son récepteur. En revanche cette protéine mutée est capable de se lier normalement avec l'IGFBP3. Il s'agit donc d'une molécule bio-inactive mais détectable par le dosage. En 2009, une mutation hétérozygote identifiée par l'équipe de Rosenfeld altère le site d'épissage entre l'intron 4 et l'exon 4 donnant une protéine tronquée. Une année plus tard, une autre étude décrit une duplication de 4 nucléotides qui génère un codon de terminaison prématuré à l'état hétérozygote.

#### 4.4.2.4 Les défauts de transport de l'IGF1

La déficience en ALS est associée à des troubles de croissance modérés -2 à -3 SDS et une déficience sévère d'IGF1, IGFBP3 et ALS liée à leurs incapacités de former le complexe ternaire circulant. Les taux d'ALS et d'IGFBP3 restent effondrés même après injection d'hormone de croissance (Domené et al., 2009; Domené et al., 2004). Certains patients présentent une résistance à l'insuline (Fofanova-Gambetti et al., 2009).

Environ 60 mutations autosomales récessives ont été identifiées (Domené et al., 2004). Ces mutations touchent à peu près tous les domaines de la protéine IGFALS et sont à l'état homozygote ou hétérozygote composite. A peu près tous les types de mutations ont été décrits dans ce gène (faux sens, non sens, délétion, duplication) (David et al., 2011).

#### 4.4.2.5 Les défauts du récepteur IGF1 (IGF1R)

Le récepteur d'IGF1 est le médiateur des effets mitotiques et métaboliques de la protéine IGF1. Son activation entraîne l'activation de plusieurs voies de signalisation. N'importe quel défaut touchant ce récepteur empêche la transmission du signal cellulaire de l'IGF1. Les défauts du récepteur IGF1R sont associés à une réduction de croissance intra-utérine, postnatale et des taux sériques d'IGF1 et de GH normaux ou même parfois élevés. Les porteurs de ce défaut naissent avec une taille inférieure à la valeur normale et qui varie ensuite de -2 à -4,9 DS, une dysmorphie, microcéphalie, un retard mental et parfois des problèmes psychomoteurs. Les mutations responsables des défauts du récepteur *IGF1R* existent toutes à l'état hétérozygote à l'exception d'une seule mutation hétérozygote composite. 8 mutations sont identifiées dans le gène *IGF1R*, de type faux sens, non sens et micro-duplication :

- l'unique cas d'hétérozygote composite était deux mutations ponctuelles faux sens dans l'exon 2 du gène qui ont modifié la séquence d'acides aminés à p.R108Q dans un allèle et p.K115N dans l'autre ;
- une mutation non-sens hétérozygote (p.R59X) qui réduit le nombre de IGF1R sur les fibroblastes ;
- une mutation du site de clivage du précurseur du récepteur (proreceptor) IGF1R ;
- une mutation non sens dans la partie intracellulaire de IGF1R ;
- une duplication de 19 nucléotides dans l'exon 18 (Fang et al., 2009). Les études fonctionnelles de cette mutation ont montré que l'ARNm exprimé à partir de l'allèle mutant est dégradé ;
- une mutation faux-sens hétérozygote dans le domaine N-terminal (fibronectine de type III) de IGF1R altère la signalisation IGF1 (Inagaki et al., 2007) ;
- p.G1125 A réduction autophosphorylation du récepteur muté (Kruis et al., 2010).

## 4.5 IGF1 DANS LE FOIE ET LE CARTILAGE

Afin de comprendre le rôle d'IGF1 produit par le foie, plusieurs groupes ont utilisé des modèles de souris KO.

L'étude de 2009 a montré que l'inactivation du gène codant pour le récepteur de la GH dans le foie entraîne la diminution de 95 % de l'IGF1 sérique, de 80 % d'IGFBP3, une diminution de la densité osseuse et des changements du poids de certains organes : diminution du poids des reins et augmentation de celui du foie. En revanche, la longueur tibiale et le volume de la plaque de croissance restent inchangés chez les souris dont le GHR a été invalidé par KO dans le foie (Fan et al., 2009).

Une deuxième étude réplique les résultats observés pour l'IGF1 circulant mais les souris avaient aussi un poids plus réduit que les souris contrôles (List et al., 2014). Selon les auteurs, cette discordance de résultats est expliquée par les différences de fond génétique des souris KO.

Afin de déterminer l'action paracrine et autocrine du gène *IGF1*, des modèles KO tissu-spécifiques ont été obtenus. L'inactivation du gène *IGF1* dans les chondrocytes entraîne une réduction de taille de 7% par rapport aux souris témoins. L'inactivation d'IGF1 au niveau des ostéoblastes réduit la

croissance de 14% par rapport aux souris sauvages. Le rôle d'IGF1 dans la croissance est donc plus important s'il est produit *in situ* dans le cartilage que sécrété par le foie et transporté secondairement au cartilage pour y exercer un rôle endocrine.

#### 4.6 LE TRANSPORT D'IGF1 DANS LA CIRCULATION

Les protéines de transport stabilisent la protéine IGF1 et modulent ainsi sa biodisponibilité. Il existe 6 protéines de transport d'IGFs de type IGFBP1-6 et la protéine ALS. Ces protéines sont capables de se lier aux protéines IGFs dès qu'elles sont relarguées dans circulation. L'intégrité du complexe protéique IGF1-IGFBP3-ALS est très importante pour la fonction de l'axe somatotrope et notamment la fonction d'IGF1. Ce complexe stabilise la protéine IGF1 et régule la biodisponibilité d'IGF1. Le principal ligand d'IGF1 est IGFBP3. 80% d'IGF1 plasmatique est liée à la protéine de transport BP3. Le temps de demi-vie du complexe IGF1-IGFBP3-ALS est de 12h-15h.

#### 4.7 EFFETS PARACRINES D'IGF1 DANS LE CARTILAGE ET SIGNALISATION VIA SON RECEPTEUR (IGF1R)

##### 4.7.1 Effets paracrines

L'injection de GH dans le cartilage du rat hypophysectomisé stimule l'expression d'*Igf1* et entraîne une stimulation de croissance en longueur plus importante que celle des rats non injectés. En revanche, des injections de l'hormone de croissance chez les souris *Igf1*<sup>-/-</sup> n'augmente pas leurs croissances. Il est proposé, par conséquent, que la GH peut agir directement sur le cartilage et produire l'IGF1 au niveau local sans passer par l'IGF1 sérique produit par le foie (Isaksson et al., 1987). Ceci fut confirmé par les expériences de l'inactivation d'*Igf1* au niveau du foie qui entraîne une diminution sévère dans le niveau d'IGF1 circulant sans affecter la viabilité et la croissance des souris (Sjögren et al., 1999). Ces études montrent que l'IGF1 circulant issu du foie n'est pas indispensable pour une croissance normale. L'action d'IGF1 local, au niveau du cartilage, suffit pour rétablir une croissance normale. Par contre, d'autres auteurs ont expliqué la diminution de certaines protéines de liaison d'IGF1 par l'augmentation de la biodisponibilité des 25% d'IGF1 circulant. Cette proportion d'IGF1 provenant des tissus non hépatiques peut être beaucoup plus biodisponible et suffira pour une croissance normale (Sjögren et al., 1999, 2002; Yakar et al., 1999).

##### 4.7.2 Signalisation via le récepteur d'IGF1R

Les effets cellulaires d'IGF1 sont médiés par son récepteur IGF1R. Le récepteur IGF1R est capable de lier IGF1, IGF2 et l'insuline mais il a une plus grande affinité pour l'IGF1 par rapport aux deux autres protéines. La fixation d'IGF1 sur son récepteur conduit à l'activation de nombreuses voies de signalisation : la voie PI3K/AKT, la principale activée, ainsi que RAS/MAP kinase ou NF-κB. La liaison d'IGF1 à IGF1R entraîne une autophosphorylation des chaînes beta intracellulaires au niveau des tyrosines 1131, 1135 et 1136, induisant la phosphorylation des résidus tyrosines et serines adjacents en extrémité carboxy-terminale. Le résidu Tyr950 phosphorylé peut alors lier les protéines de type IRS/SHC. Ces substrats déclenchent les voies des PI3K/AKT et la voie RAS/RAF/MEK/ERK (Haisa, 2013). IRS-1 phosphorylé recrute PI3K qui est alors activée. L'activation de la PI3K génère le phosphatidylinositol et l'activation d'AKT par phosphorylation. AKT phosphorylé active la synthèse protéique nécessaire pour la prolifération cellulaire et déclenche des effets anti-apoptotiques. Parallèlement, l'activation du récepteur IGF1R déclenche le recrutement de la protéine GRB2 (growth factor récepteur bound protéine 2) par IRS ou SHC activé. Ainsi un complexe GRB2 et SOS se forme et

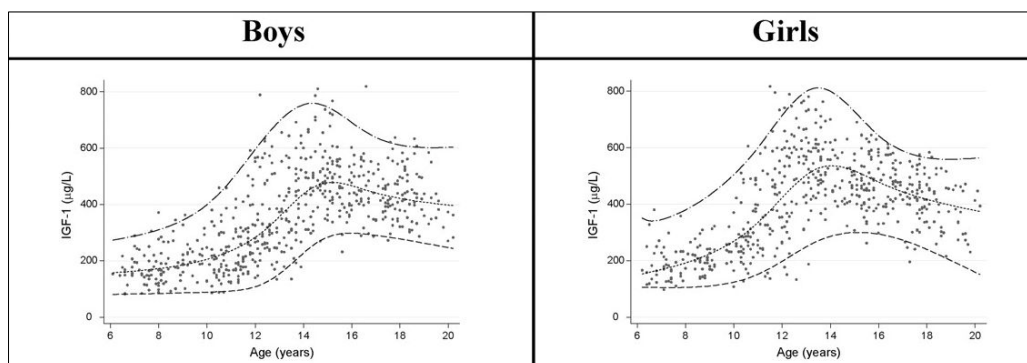
déclenche la voie de signalisation RAS/RAF/MEK/ERK. ERK activé, l'effecteur final de cette cascade de signalisation, phosphoryle des facteurs de transcription nucléaires tel que Elk-1, c-fos.

Le cycle cellulaire est étroitement régulé par la signalisation du récepteur IGF1R. En effet, IGF1R augmente l'activité des ribosomes en activant p70S6K et S6 pour faciliter la transition entre la phase G0 et G1 du cycle cellulaire (Sunami et al., 2010). D'autre part, il augmente l'expression de la cycline D1 afin de faciliter la transition G1/S. Il peut aussi réguler la transition G2/M en augmentant la synthèse des cyclines A, B et cdc2 (Mahmoudabady et al., 2009). En revanche, le rôle majeur du IGF1R dans le cycle cellulaire est la transition entre G1/S.

#### 4.8 LES CONCENTRATIONS CIRCULANTES D'IGF1

La concentration d'IGF1 dans la circulation varie en fonction de l'âge. Elle est faible chez le nouveau-né et augmente progressivement jusqu'à la puberté pour y atteindre son maximum et baisse progressivement à l'âge adulte.

Nous prenons l'exemple d'une étude française qui a permis de mettre au point des normes standard pour le dosage d'IGF1 et IGFBP3. Cette étude a été menée sur 1119 témoins français de 6 à 20 ans. Chez un groupe d'enfants de même âge et de même sexe, les auteurs ont noté une variabilité interindividuelle importante dans les concentrations d'IGF1. Malgré cette variabilité, une tendance générale d'augmentation des concentrations d'IGF1 au moment de la puberté est évidente. Entre 6 et 10 ans, la médiane des concentrations n'augmente pas chez les garçons. Chez les filles le taux d'IGF1 commence à augmenter à partir de 8 ans. Ensuite les concentrations d'IGF1 augmentent de façon linéaire et plus rapidement. Les pics de concentration d'IGF1 sont concomitants du pic pubertaire de croissance, entre 15 et 16 ans pour les garçons et de 13 à 14 ans pour les filles (Figure 27). L'allure des courbes établies à partir des concentrations d'IGF1 varie selon le stade pubertaire chez les filles. Ceci montre que la production de la protéine IGF1 est très dépendante de l'action des hormones sexuelles. Il a déjà été démontré que l'oestradiol est capable de se fixer sur le récepteur GHR et de stimuler la voie JAK2/STAT5 pour induire l'expression du gène *IGF1*. En effet, pendant les stades pubertaires, l'augmentation des concentrations d'IGF1 circulant n'est pas due uniquement à l'hormone de croissance mais également aux hormones sexuelles (Callewaert et al., 2010). Les concentrations d'IGF1 diminuent progressivement à partir de 14 ans chez les filles et de 16 ans chez les garçons (Alberti et al., 2011).



**Figure 27: Courbes de référence de la concentration en IGF1 chez des enfants français, garçons et filles, en fonction de l'âge.** Les courbes représentent le 2,5<sup>ème</sup> percentile (---), le 50<sup>ème</sup> percentile ( . . . ) et le 97,5<sup>ème</sup> percentile (— — —). En abscisses, l'âge en années, et en ordonnées, la concentration d'IGF1 circulant en µg/L. Tiré de (Alberti et al., 2011).

#### 4.9 LA CORRELATION IGF1-TAILLE EN DEVIATION STANDARD

Chez les enfants en phase de croissance, une corrélation significative entre la taille et le taux d'IGF1 circulant en SDS a été montré dans l'étude de Rogers en 2006 (Rogers et al., 2006a). Les auteurs ont analysé les données anthropométriques ainsi que les taux d'IGF1 circulant de 675 enfants impubères. Les tranches d'âge analysées dans ce travail sont de 5 à 10 ans. Les corrélations observées dans cette étude entre la taille des enfants et le taux d'IGF1 sont variables selon la tranche d'âge étudiée (Rogers et al., 2006b). A 5 ans, le taux d'IGF1 est positivement corrélé à la longueur du tronc, des jambes et à la stature des enfants. Entre 7-8 ans, il est positivement associé à la taille des enfants et à la longueur du tronc chez les deux sexes. En revanche, la longueur des jambes est associée au taux d'IGF1 chez les garçons, et faiblement associée chez les filles (Tableau 1) (Rogers et al., 2006b). Cette association entre la croissance en longueur et les taux d'IGF1 avait été montrée dans une précédente étude (Juul et al., 1994).

**Tableau 1 : Moyennes (en déviation standard) des âges, des niveaux d'IGF1 et d'anthropométrie à 5, 7-8 et 9-10 ans parmi les enfants inclus dans les analyses de croissance pour la période intermédiaire. Tiré de (Rogers et al., 2006b).**

	5 ans		De 7 à 8 ans	
	Garçons	Filles	Garçons	Filles
Enfants inclus dans les analyses 5 ans et de 7 à 8 ans				
Nombre	247	203	247	203
Age (années)	5.16 (0.05)	5.17 (0.06)	7.49 (0.13)	7.49 (0.11)
IGF-1 (ng/ml)*	150.1 (64.6)	175.2 (68.7)	141.7 (53.6)	148.9 (50.5)
Taille (cm)	110.9 (4.2)	110.2 (4.4)	125.9 (5.2)	125.2 (5.1)
Longueur de jambe (cm)	48.9 (2.5)	48.8 (2.4)	57.8 (3.1)	57.5 (2.9)
Taille assise (cm)	62.0 (2.3)	61.4 (2.5)	68.2 (2.7)	67.7 (2.8)
IMC (kg/m <sup>2</sup> )	16.0 (1.3)	16.0 (1.5)	16.1 (1.8)	16.2 (2.0)
Comparaison aux standards de croissance britannique de 1990 (x score)				
Taille	0.05 (0.93)	0.02 (0.97)	0.21 (0.98)	0.18 (0.96)
IMC	0.25 (0.92)	0.24 (0.88)	0.16 (1.00)	0.05 (0.96)
Croissance entre 5 et 7-8 ans (mm/semaine)				
Taille			1.24 (0.13)	1.25 (0.12)
Longueur de jambe			0.73 (0.11)	0.72 (0.1)
Taille assise			0.51 (0.1)	0.52 (0.1)
	7 à 8 ans		9 à 10 ans	
Enfants inclus dans l'analyse des 7 à 9 ans				
Nombre	298	257	298	257
Age (années)	7.48 (0.13)	7.49 (0.11)	9.9 (0.17)	9.89 (0.16)
IGF-1 (ng/ml)*	142.8 (51.4)	154.2 (49.6)		
IGFBP-3 (ng/ml)*	4763 (1407)	4933 (1452)		
Taille (cm)	125.7 (5.1)	125.4 (5.0)	139.5 (5.7)	139.6 (6.0)
Longueur de jambe (cm)	57.6 (3.1)	57.7 (3.0)	66.1 (3.5)	66.3 (3.5)

Taille assise (cm)	68.1 (2.6)	67.7 (2.6)	73.4 (2.8)	73.2 (3.1)
IMC (kg/m <sup>2</sup> )	16.2 (1.9)	16.4 (2.3)	17.7 (2.7)	17.9 (2.9)
Comparaison aux standards de croissance britannique de 1990 (x score)				
Taille	0.18 (0.95)	0.22 (0.93)	0.26 (0.92)	0.28 (0.94)
IMC	0.18 (1.06)	0.15 (0.99)	0.41 (1.11)	0.27 (1.04)
Croissance entre 7-8 et 9-10 ans (mm/semaine)				
Taille			1.09 (0.11)	1.13 (0.15)
Longueur de jambe			0.67 (0.1)	0.69 (0.11)
Taille assise			0.42 (0.09)	0.44 (0.12)

\* Ajusté à l'âge

**Tableau 2: Associations transversales de IGF-1 à 5 et 7-8 ans avec la taille, les tailles de la cage thoracique et de longueur de jambe, l'IMC et le poids de naissance. Tiré de (Rogers et al., 2006a).**

	Garçons		Filles	
	Moyenne (95%, IC)	P		P
5 ans	n = 247		n = 203	
Taille (cm)	2.71 (1.51 à 3.92)	<0.001	2.76 (1.35 à 4.18)	<0.001
Largeur thoracique (cm)	1.16 (0.82 à 2.29)	<0.001	1.34 (0.55 à 2.13)	0.001
Longueur de jambe (cm)	1.56 (0.48 à 1.84)	0.001	1.42 (0.6 à 2.24)	0.001
IMC (kg/m <sup>2</sup> )	0.4 (0.12 à 0.69)	0.006	0.36 (0.04 à 0.69)	0.028
Poids de naissance (g)	-0.44 (-0.75 à -0.13)	0.006	-0.57 (-0.89 à -0.24)	0.001
7-8 ans	n = 298		n = 257	
Taille (cm)	4.21 (2.8 à 5.62)	<0.001	2.74 (0.98 à 4.5)	0.002
Largeur thoracique (cm)	1.65 (0.91 à 2.39)	<0.001	1.71 (0.82 à 2.61)	<0.001
Longueur de jambe (cm)	2.56 (1.65 to 3.47)	<0.001	1.02 (-0.1 à 2.15)	0.073
IMC (kg/m <sup>2</sup> )	0.47 (0.15 à 0.78)	0.004	0.09 (-0.3 à 0.48)	0.66
Poids de naissance (g)	-0.55 (-0.86 à -0.23)	0.001	-0.58 (-0.95 à 0.21)	0.002

IC, Intervalle de confiance

<sup>a</sup> Ajusté pour l'IMC, poids de naissance, le logement et l'éducation maternelle

<sup>b</sup> Ajusté pour la taille, poids de naissance, le logement et l'éducation maternelle

<sup>c</sup> Ajusté pour la taille, l'IMC, le logement et l'éducation maternelle

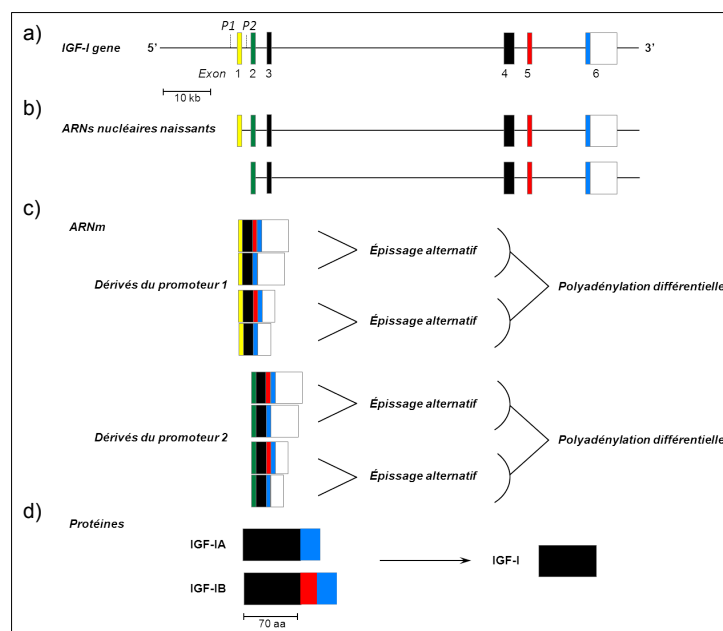
## 5 LE GENE *IGF1*

### 5.1 LA STRUCTURE DU GENE

La structure du gène *IGF1* humain a été décrite pour la première fois par Rotwein et son équipe en 1986 (Rotwein et al., 1986). Dans ce travail, la taille précise du gène n'a pu être déterminée en raison de difficultés rencontrées pour cartographier la région 5' et certaines régions introniques. Ce travail a cependant permis d'avoir une estimation de la taille génomique, au moins 45 Kb, et de proposer une structure génique. Rotwein a considéré que le gène contenait 5 exons et que l'exon 1 était divisé en deux parties A et B. Quatre ans plus tard, l'organisation de la structure du gène *IGF1* a été révisée suite à la mise en évidence d'un second exon situé entre l'exon 1 et 2 selon la première nomenclature. Actuellement, la structure du gène *IGF1* est clairement définie (Figure 28a).

*IGF1* se situe sur le bras long du chromosome 12, en position 12q23.2, il contient 6 exons et 5 introns couvrant une distance génomique de 80 kb. Seuls les exons 3 et 4 codent pour la protéine mature. Les exons 1 et 2 sont des séquences « leader » et codent pour le peptide signal. Les exons 5 et 6 codent pour l'extrémité E de la protéine native. Ces derniers exons sont éliminés par des modifications posttraductionnelles (Figure 28c).

*IGF1* possède plusieurs particularités au niveau de sa structure et dans son mode de régulation. Jusqu'à présent, le mécanisme de régulation de l'expression du gène *IGF1* demeure énigmatique. Ce mécanisme a été l'objet de plusieurs études *in vivo* et *in vitro* : *In vitro*, les modèles les plus utilisés pour ces travaux sont des lignées dérivées de certaines cibles de l'hormone de croissance ou des tissus finement régulés par la protéine IGF1, par exemple les lignées musculaires, les ostéoblastes, les lignées myocardiques ou hépatiques. Les chondrocytes ont aussi été utilisés mais uniquement chez la souris, des problèmes de dédifférenciation ayant été rencontrés chez l'homme. *In vivo*, plusieurs modèles animaux ont été étudiés, de la souris jusqu'aux modèles plus complexes comme le macaque.



**Figure 28 : Structure du gène *IGF1* chez les mammifères et les étapes aboutissant à la production d'ARNm et de protéines.** a) gène *IGF1* avec les exons (De 1 à 6) et les deux promoteurs (P1 et P2); b) ARNs nucléaires naissants issus de P1 et P2 et c) ARNm correspondants; d) La protéine IGF-I mature est dérivée des deux formes de pré-pro-IGF-I, IGF-IA et IGF-IB, exprimées en conséquence de l'épissage alternatif des ARNm. Tiré de la revue de (Rotwein, 2012).



## 5.2 LES POLYMORPHISMES DU GENE *IGF1*

Deux types de polymorphismes sont présents dans le locus *IGF1*, les polymorphismes ponctuels (SNP) et les polymorphismes de répétition (CNV).

### 5.2.1 Les SNPs

Dans le locus *IGF1*, le nombre de SNPs identifiés est de 445. Ces SNPs sont retrouvés tout le long du gène *IGF1*, depuis les régions intergéniques, en amont des promoteurs, jusqu'à la région 3' UTR (1014421894-101276092) selon la database NCBI (Figure 29).

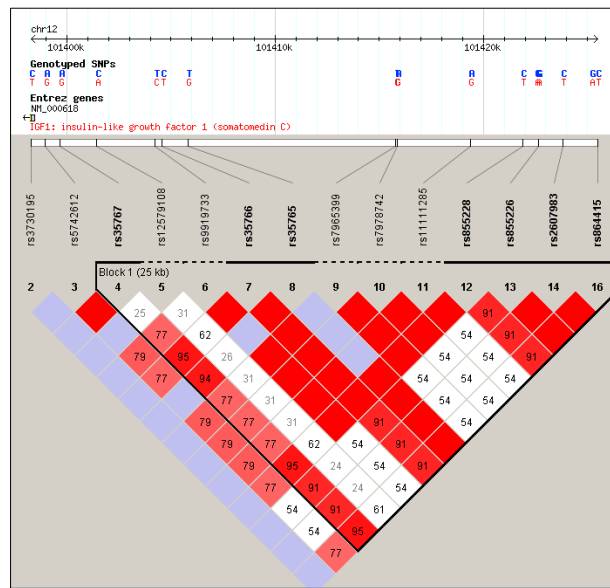


Figure 29 : Carte de déséquilibre de liaison du locus *IGF1*. Données issues de HapMap (chr12 : 1014421894-101276092).

### 5.2.2 Les CNVs

Le locus *IGF1* contient trois microsatellites, situés dans le promoteur 1, l'intron 2 et l'extrémité 3'UTR. Malgré le rôle principal d'*IGF1* dans la croissance postnatale, aucun polymorphisme de ce gène n'a été associé avec la taille dans la population caucasienne (Lettre et al., 2007). En revanche dans la population asiatique, le rs17032362 ( $P= 8.1 \times 10^{-9}$ ) est associé à la taille humaine (Okada et al., 2010). Cette association a été identifiée par GWAS chez 19.633 japonais. En raison de la faible fréquence de l'allèle mineur, ou parce qu'il n'est pas impliqué, le rs17032362 n'a pas été détecté dans les GWAS réalisées sur la population caucasienne. Les variants du locus *IGF1* n'influencent pas non plus les taux circulants d'*IGF1* (Palles et al., 2008).

## 5.3 LES TRANSCRITS D'*IGF1*

Le gène *IGF1* humain est transcrit en plusieurs préARNs dont les tailles varient de 0,95 à 7,5 kb et séparés en deux classes. La classe I englobe les transcrits issus du promoteur 1 et la classe II contient les transcrits du promoteur 2. Au sein de chacune des deux classes existent différents variants issus d'épissages alternatifs en 5' et en 3'. Chez les mammifères, le gène *IGF1* est exprimé dans la majorité des tissus mais en quantité variable : *IGF1* est fortement exprimé dans le tissu adipeux, l'utérus, les

testicules, le muscle, le cartilage et le foie. Une quantité de transcrits relativement faible est détectée dans les autres tissus.

### **5.3.1 La régulation de l'expression du gène IGF1**

#### **5.3.1.1 Les éléments *cis* de la régulation de la transcription du gène IGF1**

##### **5.3.1.1.1 Promoteur**

L'expression du gène *IGF1* est commandée par deux promoteurs en tandem. Les deux promoteurs du gène *IGF1* humain ont été identifiés pour la première fois en 1992 par Jansen et ses collaborateurs qui ont cloné la séquence minimale en amont du TSS capable d'induire une activité transcriptionnelle d'un gène rapporteur (Jansen et al., 1992). Les deux promoteurs, promoteur 1 (P1) et promoteur 2 (P2), sont considérés comme atypiques car dépourvus des éléments *cis* généralement retrouvés dans les promoteurs comme une boîte TATA ou un îlot CGs. Le Promoteur 2 est plus petit et structurellement plus simple que le P1 qui contient un VNTR et une Ebox (Hall et al., 1992).

L'équipe de Le Roith a étudié l'activité des promoteurs P1 et P2 du gène *Igf1* et l'apparition de leurs transcrits pendant les stades embryonnaires jusqu'à l'âge adulte, chez la souris. Le promoteur P1 est actif depuis les stades embryonnaires précoces au niveau du foie et reste actif jusqu'à la naissance. Le promoteur P2 est complètement inactif pendant la vie embryonnaire et commence à s'activer à partir de la vie postnatale. Le début de l'activité du promoteur P2 est concomitant avec la sensibilité du gène *IGF1* à l'hormone de croissance (Shemer et al., 1992).

La complexité de la régulation de l'expression d'*IGF1* repose sur l'existence de plusieurs sites d'initiation transcriptionnelle (Transcription Start Site, TSS) au sein de chaque promoteur. P1 contient 4 TSS tandis que P2 en possède 3.

En se basant sur les études réalisées sur le modèle murin, le promoteur P2 a été considéré exclusivement actif dans le foie. Des études réalisées sur d'autres modèles animaux ont mis en évidence des transcrits issus du promoteur 2 en dehors du tissu hépatique (dans le muscle, l'intestin, le cerveau, les reins ou les poumons) chez l'espèce bovine (Wang et al., 2003)

##### **5.3.1.1.2 Epissage alternatif en 5'**

Les 4 TSS sont capables de donner des transcrits d'extrémité 5' différente. En revanche, l'initiation de la transcription à partir du P1 se fait majoritairement à partir d'un site commun pour tous les tissus (« TSS majeur») (Lowe et al., 1987). La transcription peut parfois commencer à partir des autres TSS (Figure 28a). L'abondance des transcrits issus de ces TSS est faible et variable d'un tissu à un autre (Adamo et al., 1991; Jansen et al., 1991). De même, l'initiation de la transcription à partir du P2 peut se faire à partir de 3 TSS dont un est majoritaire (Adamo et al., 1991).

##### **5.3.1.1.3 Epissage alternatif en 3'**

Les différences de taille des ARNm du gène *IGF1*, variant de 0,95 à 7,5 kb, sont dues à un épissage alternatif en 3'. 4 sites de polyadénylation génèrent des ARNm d'extrémité 3'UTR variable (Lund et al., 1989; Shimatsu and Rotwein, 1987). Chez l'homme, l'extrémité 3' des transcrits *IGF1* inclus soit

l'exon 5 (classe Eb) soit l'exon 6 (classe Ea) ou les deux exons 5 et 6 ensemble dans les variants Ec (Figure 28a et c).

Selon les combinaisons de l'épissage alternatif en 5' et 3' les différents transcrits d'*IGF1* sont classés en : classe IEa (exons 1-3-4-6), classe IEb (exons 1-3-4-5), classe IIEa (exons 2-3-4-6), classe IIEb (exons 2-3-4-5) et classe IIEc (exons 1-3-4-5-6) (Figure 28c).

Le produit final de l'ensemble de ces transcrits est une protéine mature unique. Quel est l'intérêt de générer tous ces types de transcrits ?

Le rôle et les caractéristiques des variants d'*IGF1* ont fait l'objet de travaux depuis les années 90 et jusqu'à ce jour.

Des expériences de traduction *in vitro* ont montré que les ARN longs et courts sont capables de former des polysomes en se liant aux ribosomes en 5'UTR (Foyt et al., 1991). En revanche, les transcrits courts sont préférentiellement situés dans les ribosomes. Les ribosomes se fixent sur l'extrémité 5' de tous les transcrits et en 3' les ribosomes sont enrichis uniquement en ARN court (0,8-1,3 kb), suggérant ainsi que les extrémités 3'UTR des ARNm long sont moins accessibles à la machinerie de la traduction que les ARNm courts.

D'autre part, la stabilité des transcrits d'*IGF1* est différente. La demi-vie des ARNms d'*IGF1* est inversement corrélée avec leur taille. Les formes courtes des transcrits d'*IGF1* sont plus stables que les formes longues (Hepler et al., 1990). Les transcrits de classe II sont beaucoup plus stables que ceux de la classe I. L'instabilité des ARNs longs d'*IGF1* est probablement due à une dégradation par des micro ARNs tels mir1 ou mir206 (Xie et al., 2005) ou par déadénylation. Les transcrits de 7 et 7,5 kb contiennent des séquences riches en AU en 3' (environ 18 répétitions de -AUUUUUA-) appelées « AU rich element » (ARE). Les ARE sont capables de fixer des protéines qui reconnaissent spécifiquement ces séquences et capables de modifier la stabilité de l'ARNm ou l'efficacité de la traduction (Hao and Baltimore, 2009; Myer et al., 1997).

### 5.3.1.2 Les éléments *trans* de la régulation de la transcription du gène *IGF1*

STAT5B est l'élément *trans* clé de la transcription du gène *IGF1*. Les transcrits d'*IGF1* sont faibles voir indétectables dans le foie des souris KO *Stat5b*<sup>-/-</sup> ainsi que le niveau de la protéine IGF1 circulante (Teglund et al., 1998).

En 2003, les travaux de l'équipe de Rotwein ont démontré le rôle de STAT5B dans l'activation de la transcription du gène *IGF1 in vitro*. Dans un modèle cellulaire, l'inactivation de la protéine STAT5B bloque la transcription du gène *IGF1* tandis que l'activation constitutive de la protéine STAT5B active l'expression d'*IGF1* même en absence de stimulation GH (Woelfle et al., 2003).

Généralement, STAT5B se fixe sur le promoteur proximal des gènes sensibles à la GH tel que IGFALS, SOCS2 (Chia and Rotwein, 2010). En revanche, les études d'immunoprecipitation n'ont pas pu détecter STAT5B sur un des deux promoteurs du gène *IGF1*, qui ne contiennent aucun motif de fixation de STAT5B (Chia et al., 2010).

D'autres facteurs de transcription se fixent sur le promoteur 1 comme les protéine de famille HNF (hepatocyte nuclear facteur) (Nolten et al., 1996), le facteur C/EBP (CCAAT/enhancer binding protein) (Nolten et al., 1994), GATA-binding protein motifs (Wang et al., 2000) et le facteur MyoD (ayant un grand rôle dans l'expression d'*IGF1* dans le tissu musculaire). Ces éléments *trans* semblent importants pour la transcription basale d'*IGF1*, indépendamment de la stimulation par l'hormone de croissance.

### 5.3.2 La régulation épigénétique du locus *IGF1*

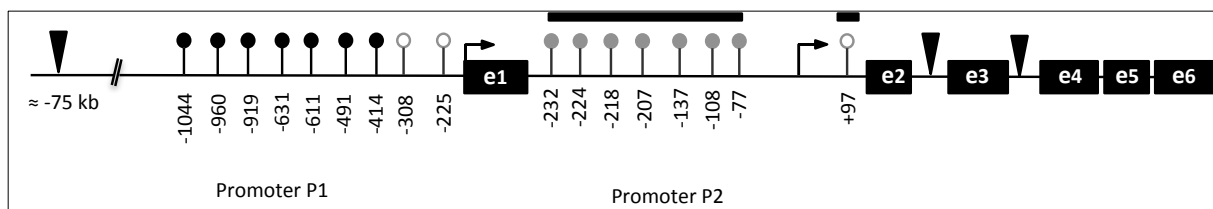
#### 5.3.2.1 Chez le rat

Dans le foie de rat hypophysectomisé, l'expression GH-dépendante du gène *IGF1* est associée à des modifications de l'accessibilité de la chromatine. La GH induit des changements dans la conformation de la chromatine au niveau des deux promoteurs et active la transcription. Après 60 minutes de stimulation, l'acétylation des histones H3 et H4 rend la chromatine plus accessible au niveau des deux promoteurs (Chia and Rotwein, 2010). La stimulation par la GH entraîne aussi un enrichissement en ARN Pol II et en coactivateur transcriptionnel P300. Au niveau du promoteur P1, GH conduit à un changement de localisation de l'ARN Pol II qui est « mise en pause » à proximité du TSS. En revanche, au niveau du promoteur P2, la GH facilite le recrutement puis l'activation de l'ARN Pol II pour initier la transcription. Donc l'arrivée de l'ARN Pol II sur le promoteur 2 semble être conditionnée par la stimulation GH (Chia and Rotwein, 2010).

#### 5.3.2.2 Chez l'homme

Les deux promoteurs d'*IGF1* peuvent être classés dans les promoteurs à faible densité de CG (cf paragraphe 3).

Le promoteur proximal 1 contient 9 CGs répartis sur environ 1000 pb et le promoteur P2 contient 10 CGs répartis sur 700 pb (Figure 30). A partir des données Encode, nous constatons que les deux promoteurs d'*IGF1* colocalisent avec des marques de chromatine active et un enrichissement en ARN pol II. Les deux promoteurs sont surtout enrichis en marque H3K27ac, marque active pour la transcription, dans les 7 lignées cellulaires humaines étudiées.



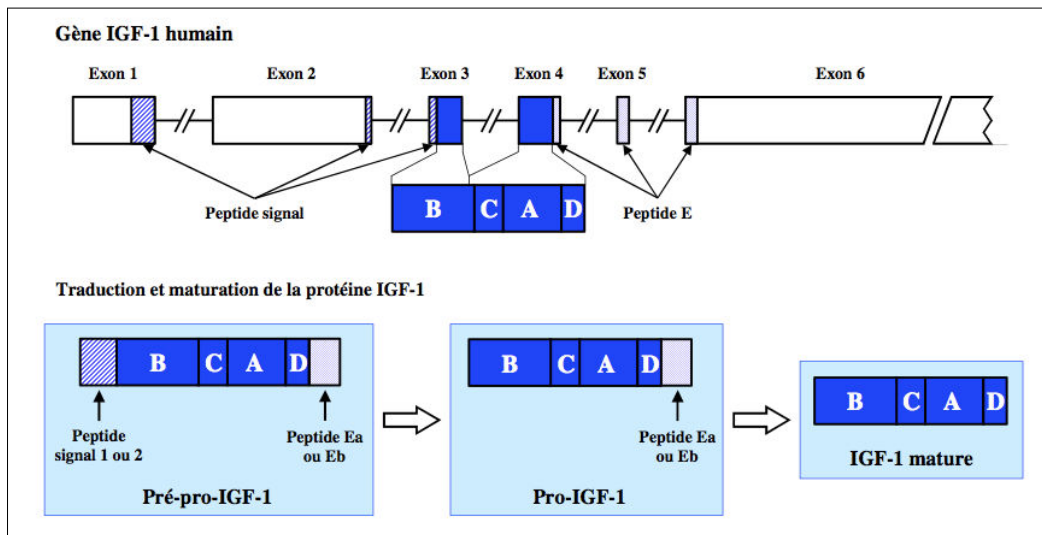
**Figure 30 : Distribution des CGs sur les deux promoteurs P1 et P2 du gène *IGF1*.** Les sites de liaison de STAT5B sont figurés par des triangles noirs (▼). Les TSS sont présentés sous forme de flèches (↗), et les exons sous forme de rectangles noirs. Les CGs étudiés sont représentés par des cercles blancs, gris ou noirs selon leur degré de méthylation (respectivement hypo-, inter- et hyper-méthylés). Voir Article 2.

## 5.4 LA PROTEINE IGF1

La traduction des différents transcrits donnent des pré-peptides légèrement différents de la structure de la protéine IGF1 mature.

### 5.4.1 Structure du prépropeptide

La traduction de différents variants du transcrit *IGF1* donne plusieurs protéines natives (ou pré-peptides) différentes. Elles sont composées de : i) un peptide signal codé par l'exon 1 ou 2, ii) les domaines BCAD codés par les exons 3 et 4 iii) le domaine E ou peptide E codé par l'exon 5 et/ou 6.



**Figure 31 : La structure du gène *IGF1* et ses différents peptides : de la pré-pro-protéine IGF1 à la protéine IGF1 mature. Figure issue de la thèse de Armelle SEROSE, 2007.**

#### 5.4.1.1 Peptide signal

La traduction des différents variants d'*IGF1* se fait à partir de deux codons d'initiation alternative AUG présents dans les exons 1 et 2. Les deux premiers exons du gène *IGF1* codent pour 5 à 21 acides aminés qui correspondent aux extrémités N-terminales différentes du peptide signal. Son extrémité C terminale est codée par le début de l'exon 3 (Figure 31).

#### 5.4.1.2 Domaine BCAD

Ces quatre domaines caractéristiques de la protéine mature sont codés par les exons 3 et 4. 25 acides aminés du domaine B sont issus de l'exon 3 et le reste du domaine B et les domaines C, A, D correspondent à l'exon 4. L'exon 4 code pour les 2/3 de la protéine mature. C'est le domaine B de la protéine mature qui interagit avec le récepteur IGF1R (Gauguin et al., 2008) ainsi qu'avec les protéines de transport IGFBP3 (Magee et al., 1999) (Figure 31).

#### 5.4.1.3 Domaine E

Le domaine E correspond à l'extrémité 3' du prépropeptide IGF1 distingué par deux parties. Une partie de 16 codons identiques pour tous les pro-pré-peptides sont codés par la fin de l'exon 4. La partie centrale du peptide E est variable et dépend de l'exon inclus en position 3' (l'exon 5 ou l'exon 6, Figure 31). Ainsi, l'extrémité C terminale du peptide E et la région 3'UTR sont aussi variables. Les deux formes de peptides E générées sont Ea et Eb. Ea est codée par les séquences des exons 4 et 6 et le peptide Eb est codé par les séquences des exons 4, 5 et 6. En outre, Ea contient deux sites de glycosylation qui ne sont pas retrouvés dans la forme Eb. Cependant, le rôle des sites de glycosylation n'est pas connu (LeRoith and Roberts, 1991). D'un point de vue fonctionnel, le peptide Eb semble être plus associé avec l'activité endocrine de l'IGF1 que le peptide Ea. Enfin, la fonction du domaine peptidique E demeure inconnue. Sachant que certaines études ont montré que les Eb et Ea n'ont pas la même capacité à traverser la membrane cellulaire, il est possible que ces différences du domaine E, distingue les formes endocrines des formes autocrines et paracrines (Oberbauer, 2013).

#### **5.4.2 Modification post traductionnelle du prépropeptide à la protéine mature**

Les précurseurs protéiques subissent des modifications post traductionnelles. Dans un premier temps, la forme « pré pro IGF1 » perd son peptide signal et devient « pro IGF1 ». Un troisième clivage protéique d'un triplet d'acides aminés glycine-proline-glutamate, peut s'effectuer au niveau de l'extrémité N terminal du domaine B. Ce dernier clivage améliore l'activité de la protéine mature ainsi que sa circulation. En effet, l'absence de ces trois résidus diminue l'affinité de l'IGF1 pour les IGFBP et permet donc d'augmenter celle pour son récepteur. La forme clivée est retrouvée dans les tissus alors que la forme complète est circulante. La protéine tissulaire clivée représente donc la forme autocrine/paracrine de l'IGF1 et la protéine circulante complète, la forme endocrine (LeRoith and Roberts, 1991).

##### **5.4.2.1 Structure de la protéine mature**

La protéine IGF1 est composée de 70 acides aminés (pour un poids moléculaire de 7,6kDa) et se trouve très conservée entre les espèces (Rinderknecht and Humbel, 1978). Elle est constituée d'une chaîne A et d'une chaîne B, reliée par une chaîne C de 12 aa et formant des ponts disulfures entre les domaines A et B. Sous forme endocrine, IGF1 est sécrétée majoritairement par le foie et véhiculée par le sang. En outre, la protéine IGF1 exerce ses effets de façon autocrine ou paracrine.

## 6 OBJECTIFS GENERAUX DE LA THESE : EPIGENETIQUE DE LA CROISSANCE

La croissance staturale est assurée par la prolifération et la différenciation des chondrocytes, notamment au niveau des plaques de croissance épiphysaires des membres inférieurs. Ces processus cellulaires obéissent à des interactions complexes de signaux hormonaux et de facteurs de croissance, eux mêmes dépendants de facteurs génétiques, nutritionnels, environnementaux.

L'épigénétique prend actuellement une grande ampleur au sein des traits complexes, car elle est un des carrefours d'interactions entre l'environnement et la génétique.

Notre équipe s'intéresse à la variation phénotypique entre les individus et plus spécifiquement à identifier des variations épigénétiques pouvant expliquer une partie des différences individuelles.

Le domaine de la croissance squelettique est particulièrement fertile en traits quantitatifs. Ceux que nous avons étudiés dans cette thèse sont la taille des enfants, le taux circulant d'IGF1, et les réponses de ces deux paramètres à l'administration d'hormone de croissance. Notre travail peut donc avoir une portée physiologique (*régulation des traits, et sensibilité à l'hormone de croissance*) ainsi qu'une portée *thérapeutique* dans le champ encore nouveau de la « *pharmaco-épigénétique* ».

### 6.1 CHOIX DE L'APPROCHE EPIGENETIQUE

#### 6.1.1 *Le principe d'analyse de méthylation*

Les techniques d'analyse de la méthylation sont plus simples à mettre en œuvre en recherche clinique que l'étude des modifications des histones. Elles sont basées sur la détection des méthyl-cytosines par diverses techniques i) l'utilisation d'enzymes sensibles à la méthylation, ii) l'utilisation du bisulfite de sodium qui transforme les cytosines non méthylées en uracile sans modifier les cytosines méthylées iii) la détection des cytosines méthylées par des anticorps anti-méthyl-cytosines. Ces techniques peuvent être appliquées sur une région bien précise ou sur des régions génomiques vastes *via* l'utilisation de puces à ADN ou des techniques de séquençage haut débit.

#### 6.1.2 *Choix de notre stratégie*

Le choix était entre une approche gène candidat *versus* génome entier. Dans le paragraphe suivant nous discutons brièvement le pour et le contre de chaque stratégie et citons les arguments à l'origine de notre choix.

##### 6.1.2.1 Les approches « génome entier »

L'approche génome entier (appelée aussi EWAS) a été mise en place pour l'étude des changements épigénétiques dans les phénotypes complexes. Les EWAS visent à détecter des variations épigénétiques et les associer à des maladies complexes en comparant un groupe de patients à un groupe de témoins. Les EWAS génèrent généralement une masse très importante d'informations qui nécessitent ensuite une analyse bioinformatique pointue pour comparer avec rigueur les deux échantillons d'individus. La robustesse statistique des analyses utilisées pour interpréter les résultats des EWAS confère une certaine confiance dans les différences observées entre un petit nombre de patients et un petit nombre de contrôles. Cependant, si ces différences ont une réalité, elles peuvent

ne représenter qu'une partie des différences qui seraient trouvées à partir de plus grand échantillons (Des études sont publiées avec de petits effectifs : n=10-25) (Bell and Saffery, 2012).

Les techniques d'études de méthylome à l'échelle du génome entier sont i) les puces, précédées d'une méthode d'enrichissement (immunoprécipitation de l'ADN méthylé, Medip) ou d'une digestion enzymatique ; ii) le séquençage bisulfite à l'échelle du génome entier.

Ce travail de thèse ne compare pas des patients et des contrôles, mais étudie des cohortes présentant une variabilité phénotypique continue de la taille ou de l'IGF1. Pour utiliser des puces « génome entier », nous aurions peut être pu comparer deux populations composées l'une de « patients » de petite taille, l'autre de « contrôles » de taille normale ou grande. Cela aurait peut être permis d'identifier des loci où la méthylation se serait révélée différente. Il aurait ensuite été nécessaire de vérifier ces différences en mesurant par bisulfite-pyroséquençage la méthylation précise des CG correspondants. Une revue détaillée des conditions d'utilisation et d'interprétation des EWAS a été rédigée par Michel Binder et ses collaborateurs avec des recommandations méthodologiques (Michels et al., 2013).

Les puces sont de plus en plus utilisées. Les premières puces 27K étaient conçues principalement pour l'étude épigénétique des cancers. Ces puces avaient un biais de couverture car elles ciblaient surtout les CGI. Les puces les plus utilisées actuellement sont les puces Illumina Infinium HumanMethylation450 Illumina 450K, (Dedeurwaerder et al., 2011). Elles ciblent 485,000 CGs répartis sur l'ensemble des promoteurs, les CGI, les « rives » (shores) des îlots et les régions de faible densité en CG. Malgré cette plus large couverture, les sites CG ciblés par les sondes de la puce 450K n'étaient pas bien adaptés à notre étude. Par exemple, la puce 450K ne cible qu'un seul CG (CG-108) dans le promoteur 2 du gène *IGF1*.

Les puces dessinées à façon peuvent être une bonne alternative puisque les régions étudiées sont choisies. Mais parfois des contraintes techniques comme la température d'hybridation des sondes peuvent être un facteur limitant dans la résolution de la puce et sa couverture.

Hormis le problème de couverture déjà évoqué, il ne faut pas perdre de vue que le signal détecté dans chaque région génomique n'est qu'une estimation de la méthylation de l'ensemble des CGs couverts par dans la sonde correspondante. Cette dernière est dessinée pour lire plusieurs CGs à la fois, au minimum 2 CGs, et le signal lumineux émis suite à la réaction d'hybridation reflète la méthylation moyenne de ces CGs. Ainsi l'étude de méthylation à l'échelle du génome entier ne permet pas d'établir des profils de méthylation CG site spécifique.

La stratégie du génome entier nécessite une étape de validation par pyroséquençage de tous les DMRs identifiés.

#### 6.1.2.2 L'approche « Gène candidat »

L'approche « gène candidat » se fonde sur les associations entre la variation épigénétique dans un gène dit d'intérêt et le phénotype étudié. La sélection des gènes candidats se fait selon leurs fonctions et leur implication dans le phénotype étudié. La stratégie gène candidat permet de focaliser sur une région génomique précise et de mesurer la méthylation de chaque dinucléotide CG choisi en utilisant la méthode du pyroséquençage. Contrairement à l'analyse des puces de méthylation, les résultats sont directement analysés sans avoir besoin d'analyses bioinformatiques préalables. Les valeurs de méthylation données par le logiciel du pyroséquençeur peuvent être utilisées pour i) identifier des DMR, ii) comparer deux groupes d'individus (cas, témoins) et ainsi identifier des DMR dues à la maladie, et iii) des recherches de corrélation avec les phénotypes.



### 6.1.3 Choix du gène candidat

1. Nous avons choisi le gène *IGF1* comme gène candidat d'abord au vu de son rôle de médiateur principal des effets de l'hormone de croissance et de son rôle dans la physiologie de la croissance.
2. Parmi les gènes qui régulent l'axe somatotrope, seuls trois gènes ont une transcription stimulée par les effets de la GH : *IGF1*, *IGFBP3*, *IGFALS*. Les autres gènes codent soit pour des facteurs de transcription soit pour des récepteurs et interagissent pendant la signalisation de GH uniquement par des réactions de phosphorylation protéique.
3. La transcription d'*IGF1* est directement activée par l'hormone de croissance. *IGF1* est le seul gène sensible à la GH exprimé dans les cellules du sang que nous pouvions étudier dans une large cohorte.
4. Bien que sa régulation transcriptionnelle reste imparfaitement connue, les facteurs épigénétiques commencent à être déchiffrés. La transcription du gène *IGF1* qui dépend de la GH est contrôlée par des modifications de la chromatine (Chia and Rotwein, 2010).
5. La faible densité en CG des deux promoteurs rend la région attractive pour l'étude de variabilité des marques épigénétiques car ces promoteurs peu riches en CG ont montré exercer un rôle sur la régulation de la transcription (Bock et al., 2008). Par exemple, les promoteurs des autres gènes de l'axe somatotrope, comme *GHR*, *STAT5B*, *IGF1R*, *IGFBP3*, *IGFALS* contiennent tous des îlots CG qui sont généralement déméthylés dans la majorité des tissus différenciés et ne sont pas variables entre les individus. Ils étaient pour nous de mauvais candidats à tester dans notre recherche.
6. Le produit du gène *IGF1* est une protéine circulante dans le sang, facile à mesurer comme endophénotype assez direct reflétant la transcription avec une gamme de variabilité individuelle très large et une corrélation significative à la taille.

## 7 LES QUESTIONS AU DEBUT DE LA THESE

Notre hypothèse était que la variation des marques épigénétiques dans les promoteurs d'*IGF1* est capable de faire varier l'expression de ce gène et d'influencer la croissance staturale, participant ainsi à la variabilité individuelle de ce trait quantitatif.

Nos questions étaient :

1. Décrire la méthylation des promoteurs du gène *IGF1* en étudiant chacun des CG qu'ils contiennent
2. Est ce que les promoteurs d'*IGF1* humain portent des marques épigénétiques (méthylation CG) variables d'un sujet à l'autre?
3. Quels sont les profils de méthylation dans les cellules du sang, le foie, le cartilage de croissance ?
4. Les profils individuels de méthylation de ces promoteurs sont ils corrélés à la variabilité de la taille des enfants et à leurs taux d'IGF1 circulant ? Si oui, quels CG portent cette corrélation ?
5. Des variations de niveau d'expression d'*IGF1* dans les PBMCs sont elles associées à la variation individuelle de la méthylation d'ADN?
6. La méthylation des promoteurs d'*IGF1* peut elle moduler la sensibilité de la vitesse de croissance à l'hormone de croissance ?
7. Est elle corrélée à la réponse de l'IGF1 circulant au traitement par l'hormone de croissance ?
8. La méthylation d'ADN est elle capable de moduler l'expression du gène *IGF1* en réponse à la GH dans des cellules PBMC ?
9. La méthylation des promoteurs a-t-elle un effet fonctionnel direct sur la transcription du gène *IGF1* dans un système artificiel de plasmide transfecté dans une lignée de cellules hépatiques

## 8 LES OUTILS DE LA THESE

### 8.1 LE SERVICE, LES MEDECINS

Ce travail a étudié la variation de différents traits cliniques, habituellement mesurés dans le service d'endocrino-pédiatrie de l'hôpital du Kremlin-Bicêtre dirigé par Pierre Bougnères et Agnès Linglart, deux spécialistes de la croissance des enfants. Plusieurs autres médecins de ce service ont aidé à la réalisation de cette thèse. Agnès Linglart, Anya Rothenbuhler et Anne-Laure Castell ont veillé à l'inclusion des participants, au recueil des données cliniques et biologiques. Toutes les données cliniques recueillies pour les cohortes étudiées ont été vérifiées et corrigées par Pierre Bougnères.

### 8.2 SUJETS ETUDIÉS

Nos observations ont porté sur des groupes d'enfants suivis dans le service d'endocrinologie pédiatrique ou recrutés dans le cadre de cohortes. Parmi les sujets étudiés, des enfants de très petite taille non pathologique traités par l'hormone de croissance ont permis l'analyse de la réponse à cette hormone.

**Définition clinique de la petite taille idiopathique.** La petite taille idiopathique (ISS, idiopathic short stature) est définie arbitrairement par convention comme étant une taille inférieure à  $-2$  déviations standard (SDscore) par rapport à la moyenne de la population normalisée pour l'âge et le sexe, sans aucune cause pathologique. Ces enfants ont une taille et un poids de naissance normaux et un niveau de GH normal. Une partie de ces enfants ont des taux d'IGF1 plutôt bas (Attie et al., 1995; Goddard et al., 1997). Les enfants présentant des défauts de sécrétion de la GH, des dysmorphies, des dysplasies squelettiques ou un syndrome de Turner ont été attentivement exclus de notre étude, de même que toutes les causes de petite taille: par exemple, la maladie coeliaque, les maladies inflammatoires de l'intestin, l'arthrite juvénile chronique, l'hypothyroïdie, la maladie de Cushing. Beaucoup des enfants ISS étudiés ont des parents de petite taille (Hermanussen and Cole, 2003; Luo et al., 1998).

**Le traitement par GH.** Certains des enfants ISS sont traités par GH pour accélérer leur vitesse de croissance et augmenter leur taille adulte finale. L'efficacité de ce traitement varie entre les individus. La distribution de la réponse au traitement suit une distribution normale tronquée chez les enfants de petite taille idiopathique. De multiples facteurs interviennent : l'âge du début de traitement, le mode de traitement, les doses de GH, la compliance. La variation de séquence du gène GHR joue elle aussi un rôle significatif dans cette variabilité de réponse individuelle.

## 9 ARTICLE 1 : « THE IGF1 P2 PROMOTER IS AN EPIGENETIC QTL FOR CIRCULATING IGF1 AND HUMAN GROWTH »

### 9.1 OBJECTIFS DE L'ARTICLE 1

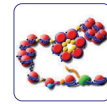
Bien que la taille humaine soit un trait hautement héritable, la variance génétique n'explique que 45 % de la variance de ce trait. L'héritabilité manquante dans ce trait peut être due à des facteurs non génétiques et parmi eux les facteurs épigénétiques. Encore faudrait-il pour le démontrer trouver des déterminants épigénétiques entrant dans la variabilité individuelle de la taille. Sans montrer que l'épigénétique contribue à la diversité individuelle, l'attribution de l'héritabilité manquante à des marques épigénétiques transmises ne pourrait se faire.

A notre connaissance, aucun travail n'avait identifié de facteur épigénétique impliqué dans la variation de la taille humaine, ni d'ailleurs dans celle d'un autre trait phénotypique.

Nous avons choisi d'étudier la croissance staturale chez l'enfant car i) elle offre une large gamme de variation phénotypique ii) elle est corrélée physiologiquement à l'effet d'IGF1 et à la concentration circulante d'IGF1 iii) elle échappe au facteur de confusion majeur qui est la croissance pubertaire déterminée en majeure partie par la maturation sexuelle et ses variations propres iv) un échantillon approprié à nos questions était disponible.

Nous nous sommes aussi intéressés à la concentration d'IGF1 sérique, d'origine essentiellement hépatique. Selon les études portant sur des jumeaux, la proportion de la variance de l'IGF1 circulant due à des effets héréditaires (certains diraient génétiques) est estimée entre 38% et 80%. Et justement, les études d'association variants génétiques-IGF1 circulant sont contradictoires et aucun variant ne semble vraiment contribuer à la variabilité des taux d'IGF1.

Notre hypothèse de travail était que la variation des profils de méthylation des promoteurs du gène *IGF1* peuvent jouer un rôle dans la régulation transcriptionnelle du gène (dans le foie et les cartilages de croissance), contribuant ainsi à la variation individuelle de la production de la protéine IGF1 circulante et de la croissance de l'enfant. Pour le démontrer, nous avons travaillé dans des cellules sanguines, le seul succédané possible des tissus physiologiques en clinique pédiatrique.



RESEARCH

Open Access

# The *IGF1* P2 promoter is an epigenetic QTL for circulating IGF1 and human growth

Meriem Ouni<sup>1</sup>, Yasemin Gunes<sup>1</sup>, Marie-Pierre Belot<sup>1</sup>, Anne-Laure Castell<sup>2</sup>, Delphine Fradin<sup>1</sup> and Pierre Bougnères<sup>1,2\*</sup>

## Abstract

**Background:** Even if genetics play an important role, individual variation in stature remains unexplained at the molecular level. Indeed, genome-wide association study (GWAS) have revealed hundreds of variants that contribute to the variability of height but could explain only a limited part of it, and no single variant accounts for more than 0.3% of height variance. At the interface of genetics and environment, epigenetics contributes to phenotypic diversity. Quantifying the impact of epigenetic variation on quantitative traits, an emerging challenge in humans, has not been attempted for height. Since insulin-like growth factor 1 (IGF1) controls postnatal growth, we tested whether the CG methylation of the two promoters (P1 and P2) of the *IGF1* gene is a potential epigenetic contributor to the individual variation in circulating IGF1 and stature in growing children.

**Results:** Child height was closely correlated with serum IGF1. The methylation of a cluster of six CGs located within the proximal part of the *IGF1* P2 promoter showed a strong negative association with serum IGF1 and growth. The highest association was for CG-137 methylation, which contributed 13% to the variance of height and 10% to serum IGF1. CG methylation (studied in children undergoing surgery) was approximately 50% lower in liver and growth plates, indicating that the *IGF1* promoters are tissue-differentially methylated regions (t-DMR). CG methylation was inversely correlated with the transcriptional activity of the P2 promoter in mononuclear blood cells and in transfection experiments, suggesting that the observed association of methylation with the studied traits reflects true biological causality.

**Conclusions:** Our observations introduce epigenetics among the individual determinants of child growth and serum IGF1. The P2 promoter of the *IGF1* gene is the first epigenetic quantitative trait locus (QTL<sup>epi</sup>) reported in humans. The CG methylation of the P2 promoter takes place among the multifactorial factors explaining the variation in human stature.

**Keywords:** QTL<sup>epi</sup>, Growth, Short stature, Height, IGF1, Epigenetics, DNA methylation, t-DMR

## Background

Although defined as a variant of normal, 'idiopathic' short stature in a child is a common source of medical investigation and a potential indication for treatment with growth hormone [1]. The molecular causes of idiopathic short stature are multiple [2], as expected for a multifactorial trait that is influenced both by individual genotypes and environmental factors [3].

First studied by F. Galton at the end of the nineteenth century [4], height heritability and variability remain partially understood [5,6]. Twin and family studies have

consistently estimated that the additive genetic contribution to normal variation in adult height approximates 80% in a given population at a given time [7-9]. As predicted by R.A Fisher [7], the genetic heritability of height is explained by many variants of individually small effect. A recent genome-wide association study (GWAS) study identified variants at 180 loci that together explain approximately 12% of the heritable variation in height [10]. Another GWAS suggested that half of the heritability of height can be accounted for by additive effects of a large number of common variants [11]. The most recent meta-analysis using the summary statistics from 79 studies totalling 253,288 individuals of European ancestry showed that the most strongly associated 2,000, 3,700, and 9,500 SNPs account for 21%, 24%, and 29% of height variance, respectively [9]. No single variant identified in these GWAS explained more than 0.3% of height variance.

\* Correspondence: pierre.bougneres@inserm.fr

<sup>1</sup>Institut National de la Santé et de la Recherche Médicale U986, Bichêtre Hospital, Paris Sud University, 80 rue du Général Leclerc Le Kremlin-Bicêtre, Paris 94276, France

<sup>2</sup>Department of Pediatric Endocrinology and Diabetes, I3E Pole, Bichêtre Hospital, Paris Sud University, rue du Général Leclerc Le Kremlin-Bicêtre, Paris 94276, France

Genetics thus seems to explain a major but limited part of height variability across individuals.

Classically, quantitative geneticists envision DNA sequence variants as the only source of heritable phenotypes. This view should be revised in light of accumulating evidence for widespread epigenetic variation in natural populations [12,13]. Indeed, the current sequence-based quantitative trait locus (QTL) approaches for dissecting complex traits miss important phenotypic effects exerted by epigenetic variants [13]. While epigenetics is a considerable source of inherited and acquired variability among humans [14], its contribution to height variability has not been studied.

To explore the individual variability of child growth and search for new causes of short stature, we selected the *insulin-like growth factor 1 (IGF1)* gene as a prominent physiological candidate. IGF1 is a key player in postnatal growth and GH signaling. Inactivating mutations in the *IGF1* gene alter postnatal growth in humans [15] and mice [16]. Clinical studies show a strong correlation between height and circulating IGF1 in childhood [17], and IGF1 production disorders are a source of short stature [18]. However, common genetic variation in the *IGF1* gene sequence does not contribute significantly to adult height variation in populations of Caucasians [19,20] while it does so in Asians whose *IGF1* allele frequency is different [21,22].

Twin studies indicate that the genetically determined proportion of circulating IGF1 variance ranges between 38% and 80% [23–25]. The association of circulating IGF1 with several genetic variants is debated [26–28], but variants at the *IGF1* locus do not seem to influence circulating IGF1 in Caucasian adults [26,29] except, maybe, the commonest Z allele of the microsatellite located 1 kb upstream of the *IGF1* gene [19,30]. Overall, the genetic basis for serum IGF1 variability remains unknown in adults and has not been studied in growing children.

Most epigenetic effects on phenotypes result from effects on gene expression, particularly for the methylation of CG located within gene promoters. Our hypothesis was that variation in epigenetic marks located in the promoters that regulate *IGF1* gene expression [31] might play a role in modulating *IGF1* gene expression, thus contribute to the individual variation of IGF1 production and child growth. In addition, these promoters are CG-poor and expected to exhibit inter-individual variation [32]. Instead of using a whole-methylome array to test associations with height or serum IGF1, our candidate gene approach enabled us to quantify the methylation of each CG of the *IGF1* promoters accurately. Indeed, an individual CG may have a significant functional role different from its CG neighbors [33,34], and this effect cannot be unraveled if the specific CG is not part of the commercial array. In addition, quantifying the methylation of each CG provides a mean to test its correlation with height or circulating IGF1.

Among the various categories of *IGF1* transcripts, class I transcripts have their initiation sites on exon 1 and are driven by P1 promoter, while class II transcripts use exon 2 as a leader exon and are driven by P2 promoter [35,36]. *IGF1* transcripts initiating at P1 are constitutively expressed in many tissues, transcripts initiating at P2 are expressed primarily but not exclusively in the liver [31,37].

The current study explores the relation between the methylation of the CGs located within *IGF1* gene promoters and child height and circulating IGF1 and tests how promoter methylation affects transcriptional activity in subjects' blood cells and in transfected cells.

## Results

### Patterns of methylation in the two IGF1 gene promoters

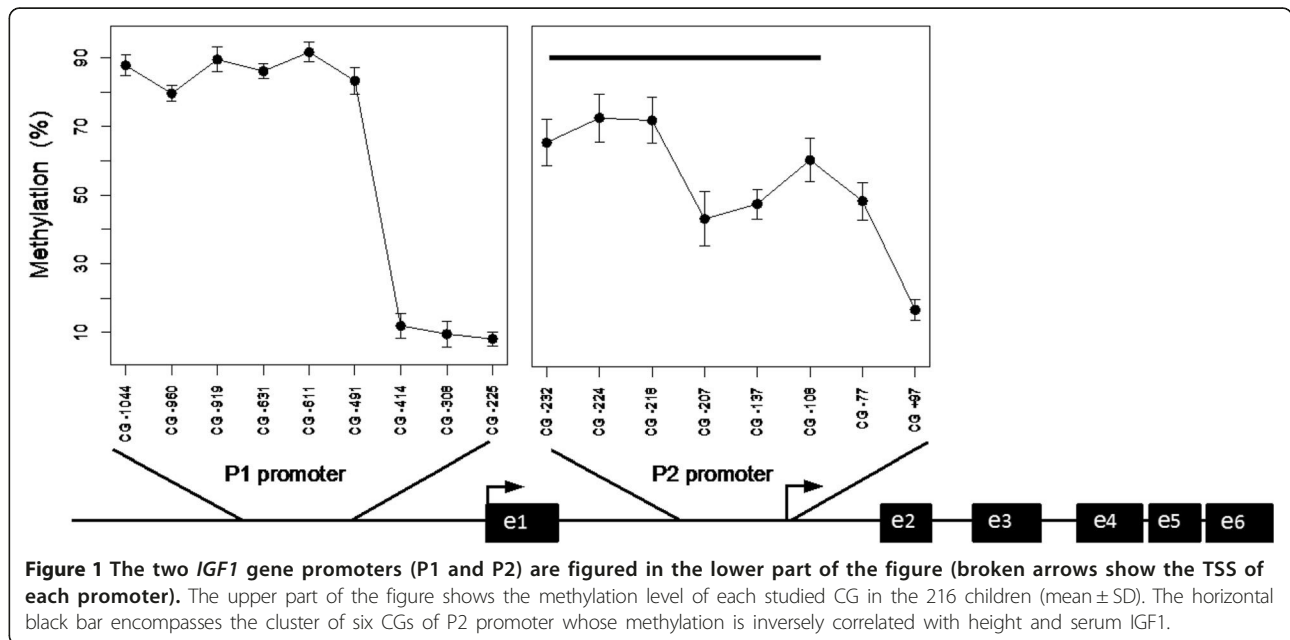
The patterns of CG methylation within the two promoters of *IGF1* are depicted in Figure 1. Methylation in the proximal part of P2 promoter showed an important individual variability (Figure 1 and Additional file 1: Table S2). Strong intra-promoter correlations ( $10^{-8} < P < 10^{-3}$ ) were observed between methylation of the CGs in P2 promoter, while there was no correlation within P1 or between CGs of P1 and P2. The methylation at the P2 and P1 promoters does not seem to be significantly influenced by sex or age (Additional file 1: Table S3).

Methylation levels were comparable in white blood cells (WBC), peripheral blood mononuclear cells (PBMC), and CD4<sup>+</sup> T lymphocytes (Figure 2). The methylation levels were approximately 50% lower for most P1 and P2 CGs in liver and growth plates than in blood cells (Figure 2, Additional file 1: Table S4). The *IGF1* promoters can thus be considered as tissue differentially methylated regions (t-DMR).

### Height and serum IGF1 are inversely correlated with CG methylation of the P2 promoter

In the discovery cohort of 110 children (Table 1), height was found to be inversely correlated with the methylation of 4/8 CGs of the P2 promoter ( $10^{-3} < P < 0.01$  after Bonferroni correction) (Table 1). These observations were replicated in a second cohort of 106 children (Table 1). After merging the two cohorts, the inverse correlation with height was confirmed for 6/8 CGs of the P2 promoter and was maximal for CG-137 methylation ( $P = 4 \times 10^{-7}$ ) (Figure 3A and Table 1), so that CG-137 methylation accounted for 13% of height variability (Figure 3A). Since 6/8 CGs in P2 trend to an association and are physically located very close together, the average methylation was calculated for these 6 CGs and showed a strong inverse correlation with height ( $P = 7 \times 10^{-5}$ ) (Table 1) and with only 1/9 CGs (CG-611) of the P1 promoter ( $P = 4.2 \times 10^{-7}$ ) (Table 1).

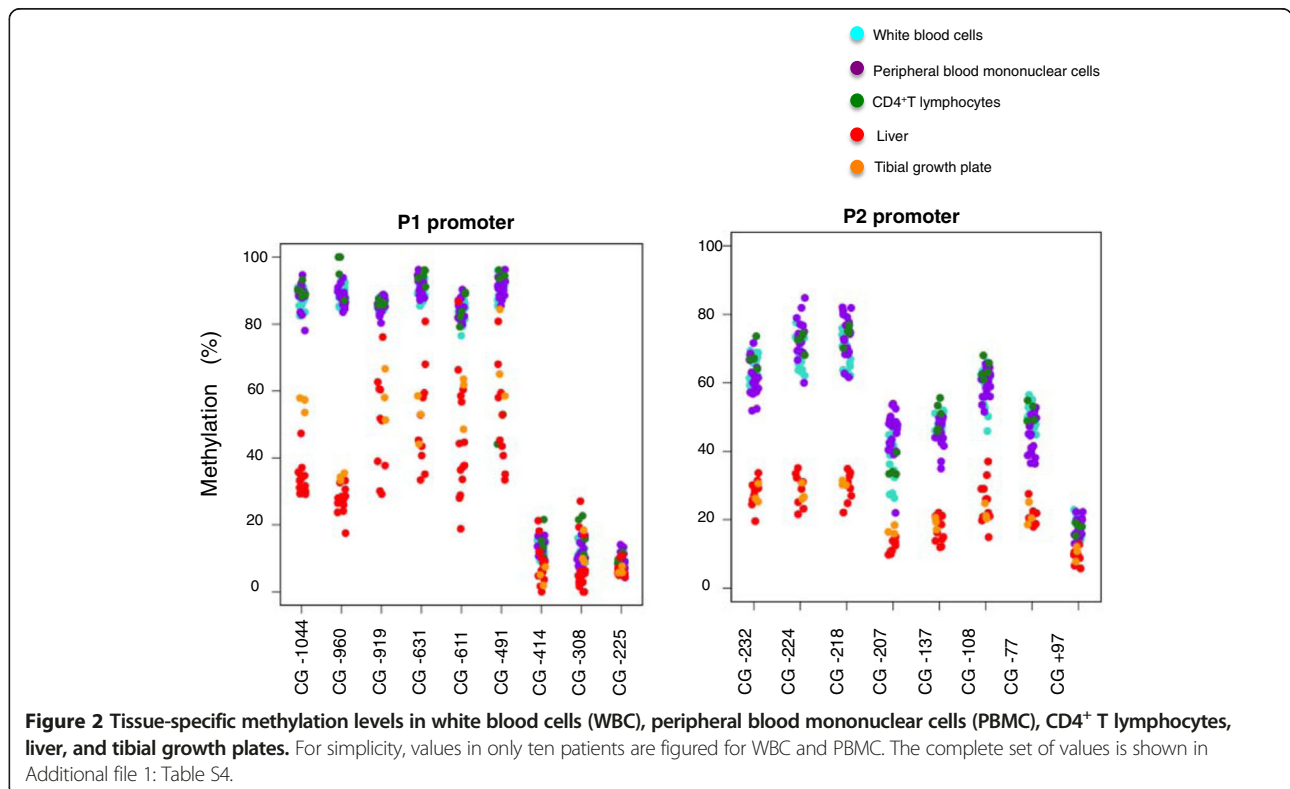
Height was closely correlated with serum IGF1 ( $P = 6 \times 10^{-15}$ ) (Figure 3B), so that the variability of circulating IGF1 accounts for 27% of height variability in the current cohort. As previously observed for height, strong inverse



correlations were found between serum IGF1 and the methylation of 4/8 CGs in the P2 promoter ( $10^{-4} < P < 0.01$  after Bonferroni correction), notably CG-137 (Figure 3C and Table 2). CG-137 methylation contributed 10% to serum IGF1 variation. Again, CG-611 was the only CG of the P1 promoter to show an inverse correlation with serum IGF1 ( $P = 0.009$ ) (Table 2).

**Relationship between methylation and genetic variation at the IGF1 locus**

We found no relationship between the common CA repeat variant located  $-822$  bp from the P2 transcription start site (TSS) and methylation of P1 or P2 promoters (Figure 4) or between this repeat variant and height or circulating IGF1 (Figure 4).



**Table 1 Correlation between the percent methylation of each studied CG and child height (N = 216)**

		Height (SDS)					
		Discovery cohort		Replication cohort		Total	
		<i>R</i>	<i>P</i>	<i>R</i>	<i>P</i>	<i>R</i>	<i>P</i>
P1 promoter	CG-1044	0.02	1	0.02	1	0.03	1
	CG-960	0.05	1	0.06	1	0.06	1
	CG-919	0.02	1	-0.13	1	0.04	1
	CG-631	-0.04	1	-0.04	1	-0.09	1
	<i>CG-611</i>	<i>-0.25</i>	<i>0.1</i>	<i>-0.48</i>	<i>2.2.10<sup>-5</sup></i>	<i>-0.39</i>	<i>4.10<sup>-7</sup></i>
	CG-491	-0.06	1	-0.31	0.03	-0.10	1
	CG-414	0.00	1	-0.19	0.4	0.09	1
	CG-308	0.02	1	0.02	1	0.00	1
	CG-225	0.08	1	0.012	1	-0.10	1
P2 promoter	CG-232	-0.08	1	-0.28	<i>0.03</i>	<i>-0.21</i>	<i>0.01</i>
	<i>CG-224</i>	<i>-0.32</i>	<i>10<sup>-2</sup></i>	<i>-0.17</i>	<i>0.7</i>	<i>-0.24</i>	<i>0.003</i>
	<i>CG-218</i>	<i>-0.36</i>	<i>10<sup>-3</sup></i>	<i>-0.30</i>	<i>0.02</i>	<i>-0.33</i>	<i>1.2.10<sup>-5</sup></i>
	<i>CG-207</i>	<i>-0.35</i>	<i>2.10<sup>-3</sup></i>	<i>-0.12</i>	<i>1</i>	<i>-0.24</i>	<i>4.10<sup>-3</sup></i>
	<i>CG-137</i>	<i>-0.30</i>	<i>10<sup>-2</sup></i>	<i>-0.40</i>	<i>1.9.10<sup>-4</sup></i>	<i>-0.36</i>	<i>4.10<sup>-7</sup></i>
	CG-108	-0.25	0.10	-0.19	0.5	-0.23	0.12
	<i>Average*</i>	<i>-0.31</i>	<i>0.007</i>	<i>-0.27</i>	<i>0.004</i>	<i>-0.3</i>	<i>7.10<sup>-5</sup></i>
	CG-77	-0.09	1	-0.04	1	-0.04	1
	CG + 97	-0.08	1	-0.03	1	0.07	1

\*Values for the averaged 6 CGs from -108 to -232. To account for multiple CG testing, we used Bonferroni correction of the *P* values. *P* values greater than 1 are equated to 1.

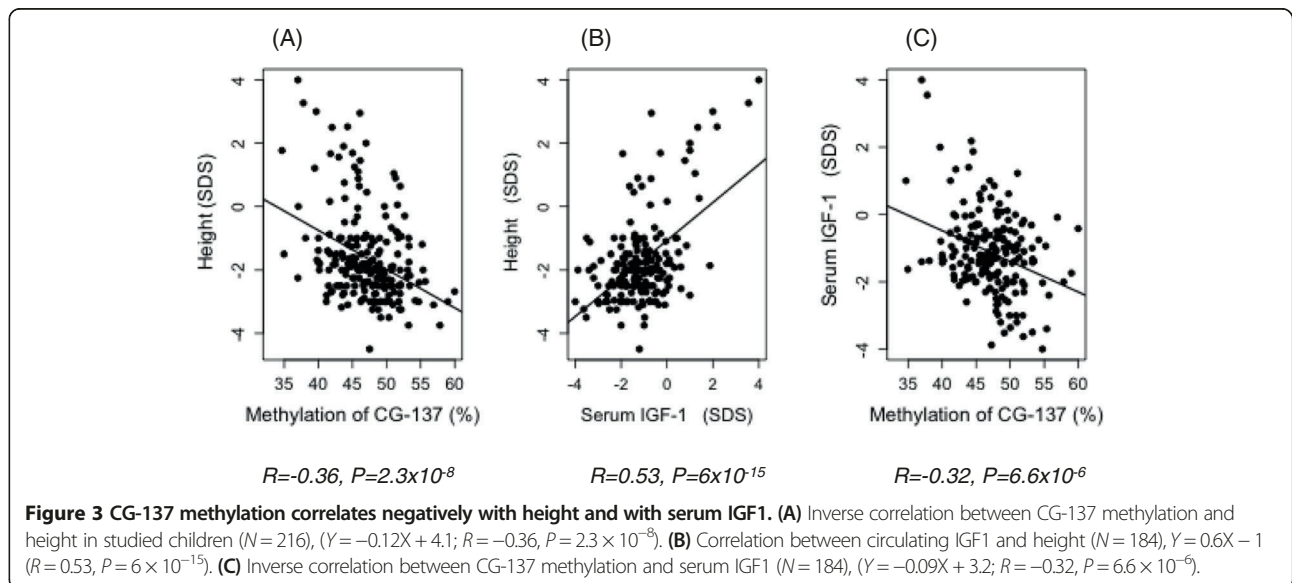
Italicized letters and numbers indicate significant correlation.

**Relationship between methylation and IGF1 transcripts in children PBMC**

Class I transcripts accounted for 73% of the total *IGF1* transcripts and class II for the remaining 27% in the PBMC of the studied children (Figure 5A). Within an

individual, the levels of class I and class II transcripts were highly correlated ( $R = 0.66; P = 1.4 \times 10^{-7}$ ) (Figure 5B).

Class II transcripts were inversely correlated with CG-137 methylation ( $R = -0.67; P = 7 \times 10^{-8}$ ) (Figure 5C), which accounted for 45% of the variation in class II





**Table 2 Correlation between methylation of all studied CGs and serum IGF1 concentration (SDS) (N = 184)**

		Serum IGF1 (SDS)					
		Discovery cohort		Replication cohort		Total	
		R	P	R	P	R	P
P1 promoter	CG-1044	-0.05	1	0.08	1	0.03	1
	CG-960	0.08	1	0.07	1	0.01	1
	CG-919	0.07	1	-0.22	0.6	-0.10	1
	CG-631	-0.18	0.7	-0.07	1	-0.10	1
	<i>CG-611</i>	<i>-0.27</i>	<i>0.06</i>	<i>-0.33</i>	<i>0.02</i>	<i>-0.25</i>	<i>9.10<sup>-3</sup></i>
	CG-491	-0.05	1	-0.11	1	-0.07	1
	CG-414	0.06	1	0.21	0.6	-0.15	0.5
	CG-308	0.00	1	0.04	1	0.03	1
	CG-225	0.09	1	0.09	1	0.04	1
P2 promoter	CG-232	-0.19	0.46	-0.29	0.08	-0.24	8.10 <sup>-3</sup>
	CG-224	-0.24	0.14	-0.05	1	-0.19	9.10 <sup>-2</sup>
	CG-218	-0.23	0.17	-0.22	0.5	-0.24	0.01
	CG-207	-0.14	1	0.09	1	-0.14	0.5
	<i>CG-137</i>	<i>-0.34</i>	<i>5.10<sup>-3</sup></i>	<i>-0.28</i>	<i>0.11</i>	<i>-0.32</i>	<i>10<sup>-4</sup></i>
	CG-108	-0.29	0.08	-0.28	0.05	-0.30	4.10 <sup>-4</sup>
	<i>Average*</i>	<i>-0.30</i>	<i>0.017</i>	<i>-0.25</i>	<i>0.17</i>	<i>-0.29</i>	<i>6.10<sup>-4</sup></i>
	CG-77	-0.23	0.25	-0.09	1	-0.17	0.3
	CG + 97	-0.22	0.29	-0.04	1	-0.12	1

To account for multiple CG testing, we used Bonferroni correction of the P values studied. P values greater than 1 are equated to 1. \*Values for the averaged 6 CGs from -108 to -232.

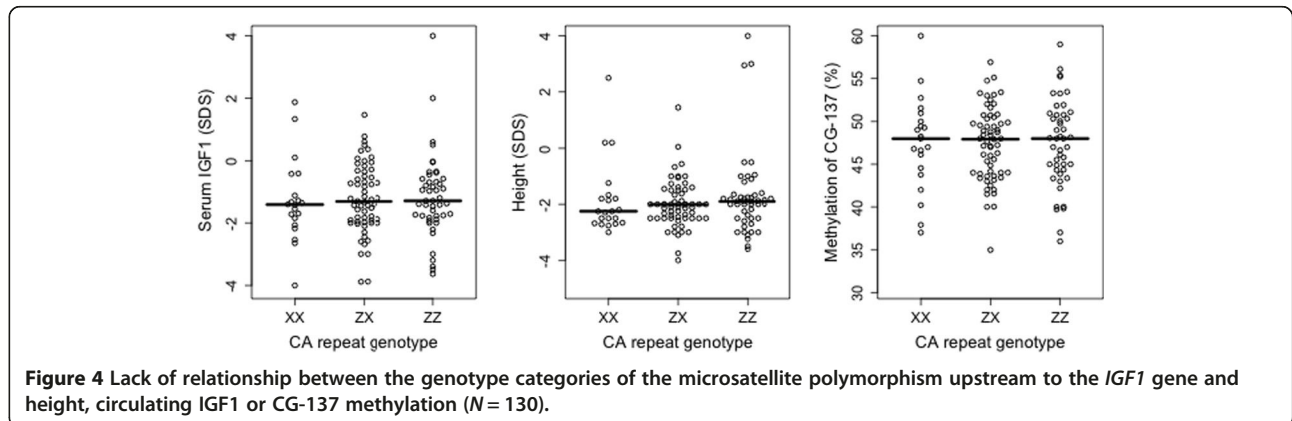
Italicized letters and numbers indicate significant correlation.

transcripts in PBMC. CG-137 methylation showed only a weak inverse correlation with class I transcripts (Figure 5D) and a highly significant correlation with total IGF1 transcripts (Figure 5E).

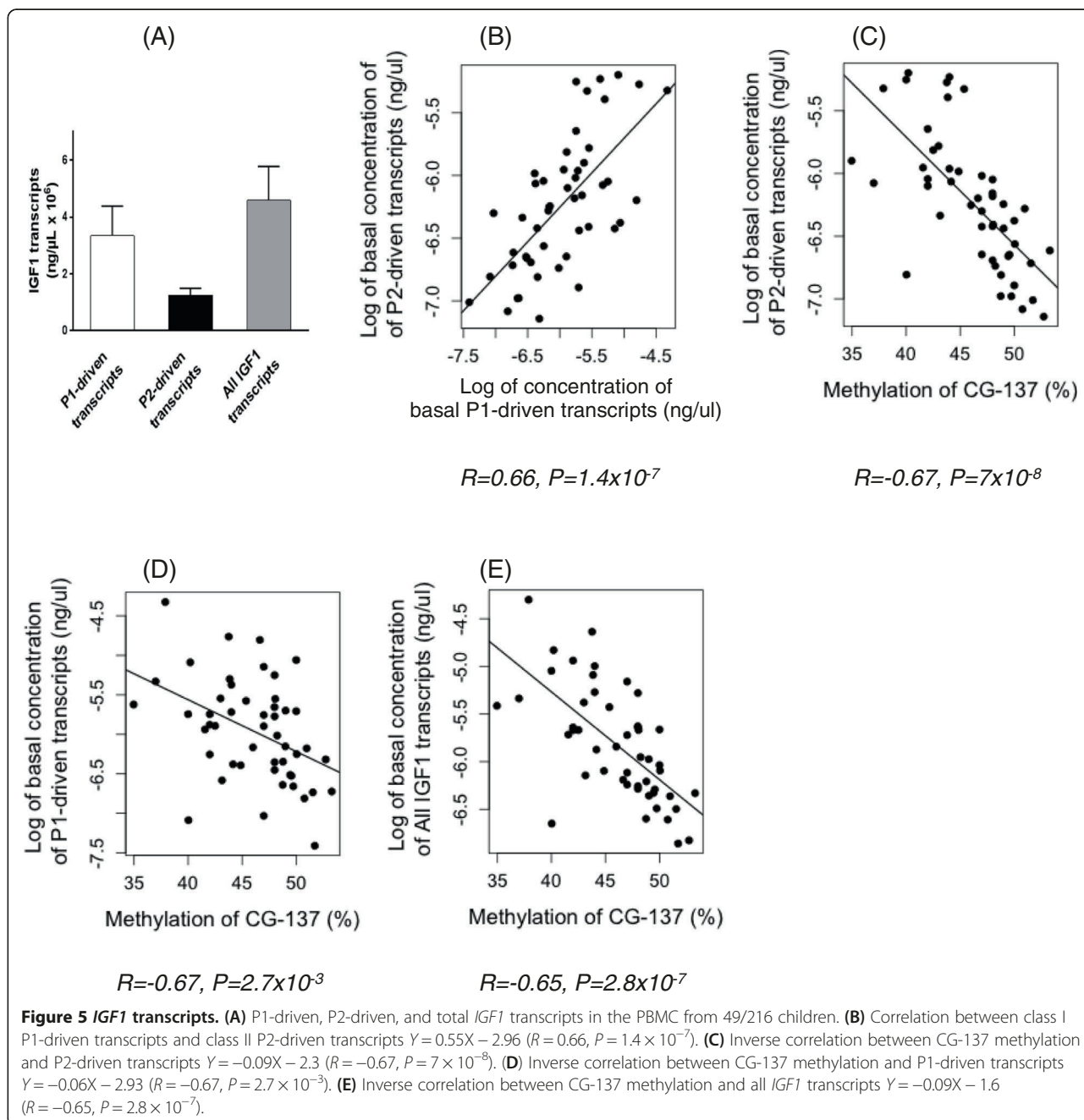
**Artificial demethylation of P2 promoter increases transcription in a reporter gene assay**

To test *in vitro* whether the methylation of the P2 promoter affects IGF1 gene expression, we used a reporter gene construct made by inserting a fragment of IGF1 P2

promoter extending from -1,014 to +64 bp of the TSS - thus containing 7 of the studied CGs - into a promoterless CpG-free luciferase expression plasmid. Constructed plasmid was mock-methylated or methylated with M. SssI CpG methyltransferase, then transiently transfected into HEK 293 cells in which basal luciferase gene expression was measured. Demethylation of the human IGF1 P2 promoter increased luciferase reporter gene expression by 57% (Figure 6), indicating that the methylation status of the P2 promoter is a significant regulator



**Figure 4 Lack of relationship between the genotype categories of the microsatellite polymorphism upstream to the IGF1 gene and height, circulating IGF1 or CG-137 methylation (N = 130).**



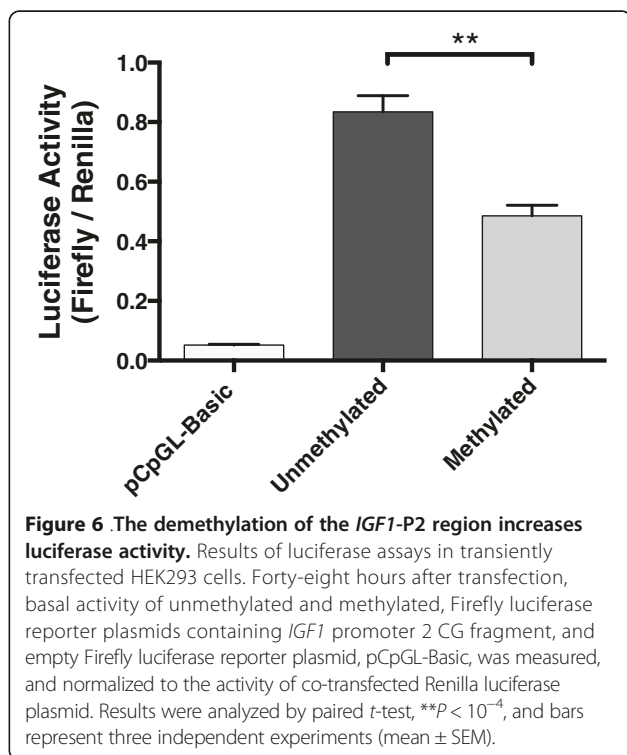
of transcriptional activity of this promoter in a plasmid environment devoid of chromatin.

### Discussion

Our observations indicate that the methylation of the P2 promoter of *IGF1* in PBMC is strongly and negatively associated with serum IGF1 and child growth. The P2 promoter can thus be considered an epigenetic quantitative trait locus (QTL<sup>epi</sup>) [13] for these traits. To our knowledge, QTL<sup>epi</sup> have yet been identified only in plants [12,13]. It is remarkable that P2 promoter methylation

contributes 13% to individual height variance, a much greater contribution than ‘major’ genetic variants that account for less than 0.3% each [10,11,20,38-40]. This underscores the potential quantitative importance of epigenetics in the variation of phenotypes.

The molecular causes of idiopathic short stature are multiple [2], as for any multifactorial trait influenced by individual genotypes and environmental factors. High levels of CG methylation at the P2 promoter now appear as one of the many molecular mechanisms responsible for ‘idiopathic’ short stature.



Another potential application of our observation is the secular increase of height [41]. One could speculate that *IGF1* P2 promoter methylation might be an epigenetic link between genetic and environmental determinants of height and contribute to the secular increase in growth observed in genetically stable populations. Indeed, modifications of DNA methylation patterns can be inherited trans-generationally, through incomplete erasure of epigenetic patterning in the germline. An accumulation of epigenetic changes through generations would then provide a valuable, reversible mechanism of adaptation to progressively changing environments. In this respect, the known interaction between energy metabolism, body composition, and IGF1 physiology [42] may stimulate the research of epigenetic links between evolving nutritional factors and height changes throughout generations.

The mechanisms generating variation of CG methylation across humans are not known [43], and this lack of knowledge applies to the individual variations observed in the CG methylation of the *IGF1* P2 promoter. Variation in methylation can arise through *cis* [44-46] or *trans* genetic effects [47,48] or as a result of stochastic events, developmental changes, or exposure to environmental cues [49,50]. The variation in DNA methylation patterns observed in MZ twins suggests that maternal environment affects the fetal epigenome [49,51].

The inverse correlation of P2 promoter methylation with both height and circulating IGF1 and the strong positive correlation between serum IGF1 and height

suggest that the epigenetic effects of P2 methylation on child growth are mediated by primary effects on IGF1 production. The *IGF1* gene is expressed in most, if not all, tissues of the body, but the liver and the chondrocytes of the growth plates are the main regulators of systemic IGF1 production and skeletal growth, respectively. Mice models have revealed that GH effects on skeletal growth are mediated by IGF1 produced *in situ* by chondrocytes, not by circulating IGF1 produced by the liver [52,53]. The close correlation of P2 methylation with circulating IGF1 and height supports that the correlation between P2 methylation and *IGF1* gene expression observed in PBMC has a biological relevance to IGF1 and growth physiology. However, our observation should be interpreted with caution because of the lack of data in growth-related tissues. Given the tissue specificity of cell epigenomes, the lack of analysis of physiological tissues and the use of blood cells as surrogates [54,55] are common but major weaknesses of epigenetic epidemiology [14,56]. Measurement of CG methylation in liver and growth plates in a limited number of children undergoing surgical procedures showed that methylation was much lower in these tissues than in blood cells, indicating that the P2 and P1 promoters are tissue-specific differentially methylated regions (t-DMR). One further caveat of using DNA extracted from WBC is the relative contributions made by the respective cell types [57], although the overall impact of blood cell composition across individuals is not considered to be substantial [55]. Herein, CG methylation levels were comparable in WBC, PBMC, and CD4<sup>+</sup> T lymphocytes.

The molecular mechanisms that link CG methylation of the P2 promoter and *IGF1* gene expression are unknown. In general, gene expression is regulated epigenetically by DNA methylation, histone modifications, and nucleosome positioning. The P2 promoter of *IGF1* belongs to the category of non-CG island promoters, more precisely to the family of low CG promoters, where methylation is known to contribute to the regulation of gene expression [58]. Despite the fact that 45% of all human gene promoters, particularly those controlling the expression of tissue-specific genes, do not lie within CG islands [59], little is known about their regulation and the potential role of methylation as a transcriptional control mechanism [60]. Among the few studies that have investigated the correlation between DNA methylation and expression of genes having non-CG island promoters, the majority do not support the view that DNA methylation directly leads to transcription silencing of these genes [60]. A number of genes with non-CG island promoters display a tissue-specific methylation pattern, suggesting that CG methylation may play a role in the establishment and maintenance of tissue-specific expression patterns [61]. Some studies have shown that there

is an inverse correlation between methylation and gene expression [61], whereas other studies have reported that CG-poor promoters could be still expressed when they are methylated [62]. However, when promoters were categorized into low, intermediate, and high CG density, it was found that the inverse correlation between methylation and expression holds in promoters with low CG density [63] and we suspect this could apply to the CG-poor P2 promoter of the *IGF1* gene.

To assess whether DNA methylation can directly regulate the P2 promoter, we utilized a CG-less luciferase-gene containing vector to perform the luciferase reporter assay. This CG-less vector overcomes the problems previously associated with testing methylation-sensitive promoter activity *in vitro*, notably the existence of CG sites in the coding region of the luciferase gene and vector backbone, which potentially contributes to the repression of the promoter [64]. The Stat5b binding sites being located outside of the transfected P2 promoter [65], Stat5b could not contribute to the observed regulation of gene expression in our transfection experiments. These experiments demonstrate that DNA methylation of the non-CG island P2 promoter of *IGF1* can directly silence gene expression, as previously shown when plasmid constructs containing the non-CG island promoters of *LAMB3* and *RUNX3* genes were transfected into HaCaT cells and 623 melanoma cells [60].

The main physiological regulator of IGF1 is GH. The importance of epigenetic mechanisms in the regulation of *IGF1* gene expression by GH starts to be unraveled in the liver of hypophysectomised rats, where GH induces dramatic changes in chromatin at the *IGF1* locus and activates *IGF1* transcription by distinct promoter-specific epigenetic mechanisms [66,67]. The proximal part of rat P2 is an important site of transcriptional regulation by GH via Stat5b [65]. In the rat liver, GH induces rapid and dramatic changes in chromatin at the P2 promoter and activates *IGF1* transcription by specific epigenetic mechanisms [66,67]. At promoter P2, GH facilitates recruitment then activation of RNA Pol II to initiate transcription, whereas at promoter P1, GH causes RNA Pol II to be released from a previously recruited poised and paused pre-initiation complex [66]. CG locations are different in human and rat *IGF1* promoters, and the pattern of methylation or transcriptional effects of these CGs are unknown in rats. Taken together, our observations in human PBMC and transfected cells support that P2 promoter methylation is a strong negative and direct modulator of P2-driven transcription of the human *IGF1* gene, through mechanisms that do not involve Stat5b intervention.

Many questions are left unanswered by the present study, including the epigenetic participation of the P2 promoter to the 'missing heritability/causality' of human

height variability [68,69] and to the phenotypic plasticity of growth physiology [70]. Another potential application of our finding comes from the pleiotropic role of IGF1 in many developmental and physiological or pathological processes related to aging, longevity, energy metabolism, brain functions, and cancer.

## Conclusions

The P2 promoter of the *IGF1* gene is a significant QTL<sup>epi</sup> for skeletal growth and serum IGF1. The relationship between P2 methylation and these growth-related traits seems to be mediated by primary transcriptional effects taking place at the P2 promoter.

## Methods

### Participants

To explore the relation between *IGF1* promoter methylation and height, we recruited children who have not yet entered puberty. Mean age was 9.7 years in boys and 9.6 in girls. Testosterone levels were lower than 0.1 ng/ml in all boys, and no girl had any breast development at time of study. This criterion of selection avoids the confounding effect of the variable tempo of sexual maturation, which adds to the variability of adolescent growth, adult height, and serum IGF1. Our discovery cohort was formed with 110 Caucasian white children (58 M/52 F) who had been recruited by the EpiGrowth Consortium. They either had a history of short stature or were control children of various statures. For our replication cohort, we used 106 unrelated healthy white children (63 M/43 F) of the GH-Pharmacogenetic Cohort, which include children with short stature [71]. Children were all healthy, had normal clinical examination and no signs of puberty. GH deficiency was excluded with a stimulated GH peak >15 ng/ml. All had normal TSH levels; subtle chondrodysplasia were excluded by forearm, pelvis, and spine radiographs.

Trained nurses performed height measurements in duplicate using the Harpenden stadiometer. Ten milliliters of peripheral blood samples was obtained, from which WBC and/or PBMC or CD4<sup>+</sup> T lymphocytes were purified immediately. WBC and PBMC were counted at time of sampling. Samples were obtained from liver and growth plates of children undergoing surgical procedures for reasons independent from the current study.

Parents gave written informed consent for the current study and for using surgical specimens, according to the French rules of bioethics in biomedical research. The research protocol received the agreement of our Institutional Review Board of Paris Sud University.

### Serum IGF1 measurements

Serum IGF1 concentration was measured around 7.00 to 8.00 am before breakfast in 184/216 children using an



immune-radiometric assay after ethanol-acid extraction using DSL-5600 Active (Diagnostic System Laboratories, Webster, TX, USA) or Cisbio reagents (Cisbio International, Codolet, France). Intra- and inter-series coefficients of variation were 1.5% and 3.7% at 260 ng/ml and 3.9% at 760 ng/ml. The sensitivity was 4 ng/ml. IGF1 SDS were calculated using the norms of Alberti *et al.* in French children [72].

#### Pyrosequencing-based bisulfite PCR analysis

For promoter P1, we studied 9 CGs located over a 800-bp distance, the closest CG being 225 bp upstream from the corresponding major TSS [73]. For promoter P2, we studied 7/8 CGs located upstream from the major TSS within the proximal part of the promoter and 1 CG located 97 bp downstream this TSS. CGs are denominated according to their position versus each promoter TSS (Figure 1). The methylation of CG-22 could not be measured for technical reasons. Nucleic acids were extracted from WBC or PBMC using Gentra Puregene blood kit (Qiagen, Hilden, Germany). In a subset of 20 patients, WBC and PBMC were measured in the same samples. Methylation was measured in other patients using PBMC. A bisulfite-PCR-pyrosequencing technique [74] was used to measure the methylation of the CGs. We improved the resolution of this method from a handful of bases to up to 100 nucleotides, with the ability to quantify methylation in the same sample of blood with a coefficient of variation (SD/mean) as little as 1% to 5%. Briefly, 400 nanograms of genomic DNA were treated with EZ DNA Methylation-Gold Kit, according to manufacturer's protocol (Zymo Research Corporation, Irvine, CA, USA). The bisulfite-treated genomic DNA was PCR-amplified using unbiased *IGF1* primers (see Additional file 1: Table S1) and performed quantitative pyrosequencing using a PyroMark Q96 ID Pyrosequencing instrument (Qiagen, Hilden, Germany). Pyrosequencing assays were designed using MethPrimer (<http://www.urogene.org/methprimer/index1.html>). Biotin-labeled single-stranded amplicons were isolated according to protocol using the Qiagen Pyromark Q96 Work Station (Qiagen, Hilden, Germany) and underwent pyrosequencing with 0.5  $\mu$ M primer. The percent methylation for each of the CGs within the target sequence was calculated using PyroQ CpG Software (Qiagen, Hilden, Germany).

Liver and tibial growth plate samples were washed with PBS and dissected to remove conjunctive tissue and blood. The biopsies were ground in cell lysis solution (Qiagen, Hilden, Germany) with Minilys® (Bertin Technologies, Montigny-le-Bretonneux) according to manufacturer's recommendation. Nucleic acids were extracted using Gentra Puregene tissue kit (Qiagen, Hilden, Germany).

#### Study of *IGF1* transcripts in PBMC

PBMC were almost immediately isolated from fresh blood in 49 children using a density gradient. Four milliliters of fresh blood was mixed with 4 ml of NaCl 154 mM, and then, 4 ml of Lymphoprep solution (Eurobio, Paris, France) was added to diluted blood and centrifuged for 20 min at room temperature at 800g. After centrifugation, the interphase containing PBMC was carefully aspirated and the cells were mixed with NaCl. The cell suspension was centrifuged at 300g, and the cell pellet washed with PBS before 700  $\mu$ l of Qiazol (Qiagen, Hilden, Germany) was added and the pellet frozen in  $-80^{\circ}\text{C}$ . Total RNA was isolated using miRNA mini kit (Qiagen), according to manufacturer's protocol. Genomic DNA was removed using DNase treatment. RNA integrity was assessed by bioanalyser (Agilent 2100, Agilent Technologies, Palo Alto, CA, USA) and expressed as RNA integrity number (RIN) considered acceptable within the range of 7 to 10. One microgram RNA was reverse-transcribed with random hexamers and oligo(dT) in a final volume of 18  $\mu$ l using PrimeScript RT reagent kit (TAKARA, Tokyo, Japan). Parallel reactions without reverse transcriptase enzyme were prepared as negative controls.

qPCR for samples were prepared in triplicate and run on ABI 7500 fast (Applied Biosystems, Foster City, CA, USA). We amplified P1-driven, P2-driven, and whole transcripts of *IGF1* transcripts using predesigned Custom TaqMan gene expression assay (Integrated DNA Technologies IDT, Coralville, IA, USA) (Additional file 1: Table S1). To detect P1-driven transcripts, P2-driven transcripts, and global *IGF1* transcripts separately, TaqMan assays were derived respectively from exons 1 to 3, exons 2 to 3, and exons 3 to 4. We measured P1-driven transcripts, P2-driven transcripts, and global transcripts of *IGF1* gene in PBMC using absolute quantification with an external calibration curve model.

A calibration standard curve was established using a cDNA clone containing qPCR product targeted by each *IGF1* TaqMan assay. A series of seven concentrations of *IGF1* cDNA clone was prepared by serial dilution. The standard curves were constructed by plotting the cycle threshold *vs.* the concentration of DNA ( $\log_{10}$  scale). The equation, slope, and correlation coefficient for each curve are presented in Additional file 2: Figure S1.

#### Construction of the reporter gene plasmid

Fragment of human *IGF1*-P2,  $-1,014$  to  $+64$ , including CGs located within the proximal part of the promoter 2, was PCR amplified from the human genomic DNA using SpeI restriction site added forward primer (Additional file 1: Table S1) and NcoI restriction site added reverse primer with Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA). PCR

product cloned into upstream of the firefly luciferase gene in the promoter-less CpG-free pCpGL-Basic vector (gift from Rehli's Lab, Regensburg, Germany) between SpeI-NcoI sites and transformed into PIR1 competent cells (Invitrogen, Carlsbad, CA, USA) for plasmid production.

*In vitro* methylation protocol was adapted from Klug and Rehli [64]. *IGF1*-P2, -1,014 to +64, in pCpGL-Basic was incubated at 37°C for 6 h with M.SssI CpG Methyltransferase (New England Biolabs, Ipswich, MA, USA) (2.5 U/μg DNA), in the presence of a methyl group donor 160 μM S-adenosylmethionine (SAM). One hundred sixty micromolar SAM was added every 2 h during incubation. Unmethylated plasmid was treated as above but without the M.SssI CpG Methyltransferase. Plasmid DNA was phenol/chloroform extracted and ethanol precipitated and quantified using a NanoDrop Spectrophotometer (ND-1000, NanoDrop Technologies, Wilmington, DE, USA). The completeness of the methylation for both methylated and unmethylated plasmids were first confirmed with methylation-sensitive restriction enzyme HpaII and methylation-insensitive (MspI) restriction enzyme.

#### Cell culture, transient transfection, and luciferase reporter assays

Human embryonic kidney 293 (HEK293) cell line (American Type Culture Collection (ATCC)) was maintained in Dulbecco's Modified Eagle's medium (DMEM; Life Technologies, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (HyClone, Thermo Fisher Scientific, Waltham, MA, USA), penicillin/streptomycin, and glutamine (PAA Laboratories, Pasing, Austria). Cells were maintained at 37°C in a humidified atmosphere containing 95% air and 5% CO<sub>2</sub>. HEK293 cells in 96-well plates were co-transfected using X-tremeGENE HP DNA Transfection Reagent (Roche, Indianapolis, IN, USA) with 80 ng/well methylated or unmethylated promoter-less CpG-free *Firefly* Luciferase expressing plasmid containing *IGF1*-P2 fragment with the 8 studied CGs, and 1 ng/well pRL-TK expressing plasmid (Promega, Madison, WI, USA), containing a constitutively expressed *Renilla* luciferase reporter gene used as an internal control for transfection efficiency. Luminescence was measured 48 h after transfection using dual-luciferase reporter assay system (Promega, Madison, WI, USA), and Centro LB 960 luminometer (Berthold Technologies, Oak Ridge, TN, USA). Results from three independent transfection experiments were averaged for comparison.

#### Genotyping distribution of (CA)<sub>n</sub> *IGF1* repeat

Genotyping of the *IGF1* gene repeats, located 1 kb upstream from the major P1 transcription start site, was adapted from Cleveland *et al.* [75] and Arends *et al.* [76]

protocols. PCR were carried out in 25-μl volumes with 30 ng of genomic DNA using 2 μM of each primer (see Additional file 1: Table S1), 100 μM dNTPs, 1.5 mM MgCl<sub>2</sub> 25 mM, 2.5 μl of the manufacturer's standard buffer and 1U of Taq RED polymerase (Eurogentec, Angers). Samples were processed through one cycle of 3 min at 94°C, 35 temperature cycles consisting of 30 s at 94°C (denaturation), 30 s at 55°C (annealing), and 30 s at 72°C (elongation), with the last elongation step lengthened to 10 min. The forward primer was 5' labeled with FAM fluorescent dye for automated fragment analysis (ABI Applied Biosystems, Foster City, CA, USA). The PCR amplified products were diluted 1/10, and 1 μl of the diluted PCR product is added to 10 μl of formamide mix and 0.5 μl of Gen Scan 500 marker. GeneMapper 4 was used to analyze the length of samples.

#### Calculations and statistical methods

IGF1 levels and height were expressed as SDS to adjust for age and sex. Pearson correlations were calculated as adjusted *R*-squared. The fraction of explained variance across children was calculated under the linear regression model, using the usual definition:  $r^2 \times 100$ . Wilcoxon rank tests and unpaired Student's *t*-tests were both used to compare methylation levels in studied tissues and the results of transfections with the demethylated or methylated P2 promoter. The Bonferroni correction was used to account for the number of CG tested. All statistics were computed using R 2.10.1. The results are expressed as mean ± SD.

#### Additional files

**Additional file 1:** List of primers used in our study. Sequences are given from 5 to 3'.

**Additional file 2:** Standard curves used for the measurement of *IGF1* gene transcript content in PBMC.

#### Abbreviations

IGF1: insulin-like growth factor 1; PBMC: peripheral blood mononuclear cells; QTL: quantitative trait locus; QTL<sup>ep</sup>: epigenetic quantitative trait locus; WBC: whole blood cells.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

PB designed the study, recruited the patients, interpreted the data, and wrote the manuscript. MO did the experiments and interpreted the results. YG did the plasmid experiment. MPB studied the P1 promoter. DF supervised the pyrosequencing platform. All authors read and approved the final manuscript.

#### Acknowledgements

We thank Rehli's lab for the gift of the promoter-less vector used in the reporter gene assay experiment. Pfizer France supported this project with research grants. MO is a doctoral fellow at Paris Descartes doctoral school. MPB was supported by the Association de Recherche sur le Diabète.

Received: 1 December 2014 Accepted: 19 February 2015

Published online: 13 March 2015

## References

- Cohen P, Rogol AD, Deal CL, Saenger P, Reiter EO, Ross JL, et al. Consensus statement on the diagnosis and treatment of children with idiopathic short stature: a summary of the Growth Hormone Research Society, the Lawson Wilkins Pediatric Endocrine Society, and the European Society for Paediatric Endocrinology Workshop. *J Clin Endocrinol Metab.* 2007;2008(93):4210–7.
- Rosenfeld RG. The molecular basis of idiopathic short stature. *Growth Horm IGF Res.* 2005;15(Suppl A):S3–5.
- Cole TJ. The secular trend in human physical growth: a biological view. *Econ Hum Biol.* 2003;1:161–8.
- Galton F, Dickson JDH. Family likeness in stature. *Proc R Soc Lond.* 1886;40:42–73.
- Visscher PM, McEvoy B, Yang J. From Galton to GWAS: quantitative genetics of human height. *Genet Res.* 2010;92:371–9.
- Hirschhorn JN, Gajdos ZKZ. Genome-wide association studies: results from the first few years and potential implications for clinical medicine. *Annu Rev Med.* 2011;62:11–24.
- Fisher RA. XV. - The correlation between relatives on the supposition of Mendelian inheritance. *Earth Environ Sci Trans R Soc Edinb.* 1919;52:399–433.
- Silventoinen K, Sammalisto S, Perola M, Boomsma DI, Cornes BK, Davis C, et al. Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Res.* 2003;6:399–408.
- Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet.* 2014;46:1173–86.
- Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature.* 2010;467:832–8.
- Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet.* 2010;42:565–9.
- Murrell A, Rakan VK, Beck S. From genome to epigenome. *Hum Mol Genet.* 2005;14(Spec No 1):R3–10.
- Johannes F, Colot V, Jansen RC. Epigenome dynamics: a quantitative genetics perspective. *Nat Rev Genet.* 2008;9:883–90.
- Rakan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet.* 2011;12:529–41.
- Woods KA, Camacho-Hübner C, Savage MO, Clark AJ. Intrauterine growth retardation and postnatal growth failure associated with deletion of the insulin-like growth factor I gene. *N Engl J Med.* 1996;335:1363–7.
- Lupu F, Terwilliger JD, Lee K, Segre GV, Efstratiadis A. Roles of growth hormone and insulin-like growth factor 1 in mouse postnatal growth. *Dev Biol.* 2001;229:141–62.
- Rogers I, Metcalfe C, Gunnell D, Emmett P, Dunger D, Holly J, et al. Insulin-like growth factor-I and growth in height, leg length, and trunk length between ages 5 and 10 years. *J Clin Endocrinol Metab.* 2006;91:2514–9.
- Cohen P. Controversy in clinical endocrinology: problems with reclassification of insulin-like growth factor I production and action disorders. *J Clin Endocrinol Metab.* 2006;91:4235–6.
- Frayling TM, Hattersley AT, McCarthy A, Holly J, Mitchell SMS, Gloyne AL, et al. A putative functional polymorphism in the IGF-I gene association studies with type 2 diabetes, adult height, glucose tolerance, and fetal growth in U.K. Populations. *Diabetes.* 2002;51:2313–6.
- Lettre G, Butler JL, Ardlie KG, Hirschhorn JN. Common genetic variation in eight genes of the GH/IGF1 axis does not contribute to adult height variation. *Hum Genet.* 2007;122:129–39.
- Kim J-J, Lee H-I, Park T, Kim K, Lee J-E, Cho NH, et al. Identification of 15 loci influencing height in a Korean population. *J Hum Genet.* 2009;55:27–31.
- Okada Y, Kamatani Y, Takahashi A, Matsuda K, Hosono N, Ohmiya H, et al. A genome-wide association study in 19 633 Japanese subjects identified LHX3-QSOX2 and IGF1 as adult height loci. *Hum Mol Genet.* 2010;19:2303–12.
- Harrela M, Koistinen H, Kaprio J, Lehtovirta M, Tuomilehto J, Eriksson J, et al. Genetic and environmental components of interindividual variation in circulating levels of IGF-I, IGF-II, IGFBP-1, and IGFBP-3. *J Clin Invest.* 1996;98:2612–5.
- Kao PC, Matheny AP, Lang CA. Insulin-like growth factor-I comparisons in healthy twin children. *J Clin Endocrinol Metab.* 1994;78:310–2.
- Verhaeghe J, Loos R, Vlietinck R, Van Herck E, van Bree R, DSchutter AM. C-peptide, insulin-like growth factors I and II, and insulin-like growth factor binding protein-1 in cord serum of twins: genetic versus environmental regulation. *Am J Obstet Gynecol.* 1996;175:1180–8.
- Palles C, Johnson N, Coupland B, Taylor C, Carvajal J, Holly J, et al. Identification of genetic variants that influence circulating IGF1 levels: a targeted search strategy. *Hum Mol Genet.* 2008;17:1457–64.
- Al-Zahrani A, Sandhu MS, Luben RN, Thompson D, Baynes C, Pooley KA, et al. IGF1 and IGFBP3 tagging polymorphisms are associated with circulating levels of IGF1, IGFBP3 and risk of breast cancer. *Hum Mol Genet.* 2006;15:1–10.
- Wong H-L, DeLellis K, Probst-Hensch N, Koh W-P, Berg DVD, Lee H-P, et al. A new single nucleotide polymorphism in the insulin-like growth factor I regulatory region associates with colorectal cancer risk in Singapore Chinese. *Cancer Epidemiol Biomarkers Prev.* 2005;14:144–51.
- Ester WA, Hokken-Koelega ACS. Polymorphisms in the IGF1 and IGF1R genes and children born small for gestational age: results of large population studies. *Best Pract Res Clin Endocrinol Metab.* 2008;22:415–31.
- Vaessen N, Janssen JA, Heutink P, Hofman A, Lamberts SWJ, Oostra BA, et al. Association between genetic variation in the gene for insulin-like growth factor-I and low birthweight. *Lancet.* 2002;359:1036–7.
- Rotwein P. Mapping the growth hormone - Stats5b - IGF-I transcriptional circuit. *Trends Endocrinol Metab.* 2012;23:186–93.
- Bock C, Walter J, Paulsen M, Lengauer T. Inter-individual variation of DNA methylation and its implications for large-scale epigenome mapping. *Nucleic Acids Res.* 2008;36:e55.
- Kuroda A, Rauch TA, Todorov I, Ku HT, Al-Abdullah IH, Kandeel F, et al. Insulin gene expression is regulated by DNA methylation. *PLoS ONE.* 2009;4:e6953.
- Murgatroyd C, Patchev AV, Wu Y, Micala V, Bockmühl Y, Fischer D, et al. Dynamic DNA methylation programs persistent adverse effects of early-life stress. *Nat Neurosci.* 2009;12:1559–66.
- Yang H, Adamo ML, Koval AP, McGuinness MC, Ben-Hur H, Yang Y, et al. Alternative leader sequences in insulin-like growth factor I mRNAs modulate translational efficiency and encode multiple signal peptides. *Mol Endocrinol.* 1995;9:1380–95.
- Adamo ML, Ben-Hur H, Roberts CT, LeRoith D. Regulation of start site usage in the leader exons of the rat insulin-like growth factor-I gene by development, fasting, and diabetes. *Mol Endocrinol.* 1991;5:1677–86.
- Oberbauer AM. The regulation of IGF-1 gene transcription and splicing during development and aging. *Front Endocrinol.* 2013;4:39.
- Weedon MN, Lettre G, Freathy RM, Lindgren CM, Voight BF, Perry JRB, et al. A common variant of HMG2 is associated with adult and childhood height in the general population. *Nat Genet.* 2007;39:1245–50.
- Sanna S, Jackson AU, Nagaraja R, Willer CJ, Chen W-M, Bonnycastle LL, et al. Common variants in the GDF5-UQCC region are associated with variation in human height. *Nat Genet.* 2008;40:198–203.
- Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, et al. Genome-wide association analysis identifies 20 loci that influence adult height. *Nat Genet.* 2008;40:575–83.
- Bogin B. Secular changes in childhood, adolescent and adult stature. *Nestlé Nutr Inst Workshop Ser.* 2013;71:115–26.
- Clemmons DR, Underwood LE. Nutritional regulation of IGF-I and IGF binding proteins. *Annu Rev Nutr.* 1991;11:393–412.
- Bird A. Putting the DNA back into DNA methylation. *Nat Genet.* 2011;43:1050–1.
- Knight JC. Allele-specific gene expression uncovered. *Trends Genet TIG.* 2004;20:113–6.
- Meaburn EL, Schalkwyk LC, Mill J. Allele-specific methylation in the human genome: implications for genetic studies of complex disease. *Epigenetics.* 2010;5:578–82.
- Fradin D, Le Fur S, Mille C, Naoui N, Groves C, Zelenika D, et al. Association of the CpG methylation pattern of the proximal insulin gene promoter with type 1 diabetes. *PLoS ONE.* 2012;7:e36278.
- Richards EJ. Population epigenetics. *Curr Opin Genet Dev.* 2008;18:221–6 [Chromosomes and Expression Mechanisms].
- Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* 2011;12:R10.

49. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A*. 2005;102:10604–9.
50. Kaminsky ZA, Tang T, Wang S-C, Ptak C, Oh GHT, Wong AHC, et al. DNA methylation profiles in monozygotic and dizygotic twins. *Nat Genet*. 2009;41:240–5.
51. Petronis A. Epigenetics and twins: three variations on the theme. *Trends Genet TIG*. 2006;22:347–50.
52. Yakar S, Pennisi P, Wu Y, Zhao H, LeRoith D. Clinical relevance of systemic and local IGF-I. In: Cianfarani S, Clemmons DR, Savage MO, editors. *Endocrine Development*. Basel: KARGER; 2005. p. 11–6.
53. Liu JL, Yakar S, LeRoith D. Conditional knockout of mouse insulin-like growth factor-1 gene using the Cre/loxP system. *Proc Soc Exp Biol Med*. 2000;223:344–51.
54. Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T, et al. Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci Transl Med*. 2010;2:49a67.
55. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. *FASEB J*. 2010;24:3135–44.
56. Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nat Rev Genet*. 2013;14:585–94.
57. Wu H-C, Delgado-Cruzata L, Flom JD, Kappil M, Ferris JS, Liao Y, et al. Global methylation profiles in DNA from different blood cell types. *Epigenetics*. 2011;6:76–85.
58. Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, et al. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell*. 2013;153:1134–48.
59. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A*. 2002;99:3740–5.
60. Han H, Cortez CC, Yang X, Nichols PW, Jones PA, Liang G. DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter. *Hum Mol Genet*. 2011;20:4299–310.
61. Eckhardt F, Lewin J, Cortese R, Rzykany VK, Attwood J, Burger M, et al. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet*. 2006;38:1378–85.
62. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet*. 2007;39:457–66.
63. Gal-Yam EN, Egger G, Iniguez L, Holster H, Einarsson S, Zhang X, et al. Frequent switching of polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. *Proc Natl Acad Sci U S A*. 2008;105:12979–84.
64. Klug M, Rehli M. Functional analysis of promoter CpG methylation using a CpG-free luciferase reporter vector. *Epigenetics*. 2006;1:127–30.
65. Varco-Merth B, Mirza K, Alzhanov DT, Chia DJ, Rotwein P. Biochemical characterization of diverse Stat5b-binding enhancers that mediate growth hormone-activated insulin-like growth factor-I gene transcription. *PLoS ONE*. 2012;7:e50278.
66. Chia DJ, Young JJ, Mertens AR, Rotwein P. Distinct alterations in chromatin organization of the two IGF-I promoters precede growth hormone-induced activation of IGF-I gene transcription. *Mol Endocrinol*. 2010;24:779–89.
67. Chia DJ, Rotwein P. Defining the epigenetic actions of growth hormone: acute chromatin changes accompany GH-activated gene transcription. *Mol Endocrinol*. 2010;24:2038–49.
68. Slatkin M. Epigenetic inheritance and the missing heritability problem. *Genetics*. 2009;182:845–50.
69. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461:747–53.
70. Atchley WR, Zhu J. Developmental quantitative genetics, conditional epigenetic variability and growth in mice. *Genetics*. 1997;147:765–76.
71. Dos Santos C, Essioux L, Teinturier C, Tauber M, Goffin V, Bougnères P. A common polymorphism of the growth hormone receptor is associated with increased responsiveness to growth hormone. *Nat Genet*. 2004;36:720–4.
72. Alberti C, Chevenne D, Mercat I, Josserand E, Armoogum-Boizeau P, Tichet J, et al. Serum concentrations of insulin-like growth factor (IGF)-1 and IGF binding protein-3 (IGFBP-3), IGF-1/IGFBP-3 ratio, and markers of bone turnover: reference values for French children and adolescents and z-score comparability with other references. *Clin Chem*. 2011;57:1424–35.
73. Jansen E, Steenbergh PH, LeRoith D, Roberts CT, Sussenbach JS. Identification of multiple transcription start sites in the human insulin-like growth factor-I gene. *Mol Cell Endocrinol*. 1991;78:115–25.
74. Tost J, Gut IG. DNA methylation analysis by pyrosequencing. *Nat Protoc*. 2007;2:2265–75.
75. Cleveland RJ, Gammon MD, Edmiston SN, Teitelbaum SL, Britton JA, Terry MB, et al. IGF1 CA repeat polymorphisms, lifestyle factors and breast cancer risk in the Long Island Breast Cancer Study Project. *Carcinogenesis*. 2006;27:758–65.
76. Arends N, Johnston L, Hokken-Koelega A, van Duijn C, de Ridder M, Savage M, et al. Polymorphism in the IGF-I gene: clinical relevance for short children born small for gestational age (SGA). *J Clin Endocrinol Metab*. 2002;87:2720.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)





**Additional file 1: Table S1.** List of primers used in our study. Sequences are given from 5' to 3'.

<b>Region</b>	<b>Primers</b>
<b>Pyrosequencing</b>	
CGs -1044 to -919	F- TATTGTTTTATAAAAATTAAGGGAA R- ATTCTAAATAACACCAACTAACTAAC
primary for CGs -631. to -308	F- TGATAGGTAGTTTAGTAGAAGAATAA R- AAAACTAAAACACAAAAACATTTTT
nested 1 for CGs -631 and -611	F- GTTTTATTTTAGTTGGGTTTTATAG R- CTCTACAATTTTAACCCTAAAAT
nested 2 for CGs -491 to -308	F- GTAATTGGGATAAGGGGTATT R- AACTTTCTACTAAACATAAAAAACACAAAC
CG -225	F- TTTTAAATATTGGTAAGGTGTATTAG R- TTTCCCTTAAAAAATATATATTAAT
Sequencing CG -1044	F- AGGGAAATAGGTATAAAT
Sequencing CG -960	F- AGTTGGTTTGGATTATGT
Sequencing CG -919	F- GGGATTTGAATGATATTA
Sequencing CGs -631 and -611	R- CTATCTCATACTTTTTCT
Sequencing CG -491	F- TAATTGGGATAAGGGG
Sequencing CG -414	F- ATTTTAGGGTTAAAATTGTAGAGT
Sequencing CG -308	F- ATATTGGTAAGGTGTATT
Sequencing CG -225	F- GGTAGTATAAGTATTTTA
CGs -232 to 108	F- AATTTGGTTGTTGTTGTTAGTGTAT R- AATTAAACCCTCAAACAATTAATC
CGs -77 and +97	F- AATTGTTTGAGGGTTTAATTTATAAGA R- ACTACTAAATCACATAATATATATACACA
Sequencing CGs -232, -224, 218 and -207	F- TGGTTGTTGTTGTTAGTGTAT
Sequencing CG -137	R- ACCAATAACAACAACCTTAA
Sequencing CG -108	F- AAGTTGTTGTTATTGGTT
Sequencing CG -77	R- ACTCCCTCAAACCACTTCCTACT

Sequencing CG +97

R- AATTGTTTGAGGGTTTAATTTATAAGA

---

### **Quantitative PCR**

*HPRT1* Hs.PT.5821454446 Assay IDT

Total *IGF1* Hs.PT.56a.21022358.g IDT

P1-driven *IGF1* N001111285.1.pt.IGF1 IDT

P2-driven *IGF1* N001111284.1.pt.IGF1 IDT

---

### **Plasmid**

Forward TCAACTAGTATGTGTACTGTTTGCTTCTGCCTAGA

Reverse TCACCATGGTGGCATCTCAGGAACAGCAG

---

### **Genotypage *IGF1* CAn**

Forward GCTAGCCAGCTGGTGTATT

Reverse ACCACTCTGGGAGAAGGGTA

---

**Additional file 1: Table S2.** CG methylation in the P1 and P2 promoters of the *IGF1* gene. The position of each CG is given as bp from the TSS of the corresponding promoter.

		Percent Methylation			
		Mean $\pm$ SD	Extreme Values%	Range interval	SD/mean
Promoter P1	CG -1044	88 $\pm$ 3	78 - 95	17	0.03
	CG -960	79 $\pm$ 4	75 - 87	12	0.05
	CG -919	89 $\pm$ 6	80 - 98	18	0.07
	CG -631	86 $\pm$ 2	75 - 91	16	0.02
	CG -611	92 $\pm$ 3	85 - 99	14	0.03
	CG -491	84 $\pm$ 4	69 - 93	24	0.05
	CG -414	12 $\pm$ 4	5 - 36	31	0.33
	CG -308	10 $\pm$ 4	3 - 28	25	0.40
	CG -225	8 $\pm$ 2	4 - 14	10	0.25

Promoter P2	CG -232	66 $\pm$ 7	46 - 87	41	0.11
	CG -224	73 $\pm$ 7	58 - 92	34	0.10
	CG -218	73 $\pm$ 6	61 - 89	28	0.08
	CG -207	44 $\pm$ 8	20 - 60	40	0.18
	CG -137	47 $\pm$ 4	35 - 60	25	0.09
	CG -108	61 $\pm$ 6	46 - 74	28	0.10
	CG -77	48 $\pm$ 5	35 - 60	25	0.10
	CG +97	16 $\pm$ 3	9 - 24	15	0.19

**Additional file 1: Table S3.** CG methylation is not influenced by age. Adult controls from our biobank were used for comparison.

Age (yrs)	<b>&lt; 11</b>	<b>11-20</b>	<b>20-60</b>	<b>&gt;60</b>
<i>N</i>	<i>185</i>	<i>120</i>	<i>47</i>	<i>46</i>
CG-137 methylation %	46 ± 5	45 ± 5	44 ± 4	45 ± 4

**Additional file 1: Table S4.** Methylation levels in white blood cells (WBC), peripheral blood mononuclear cells (PBMC), CD4<sup>+</sup>T lymphocytes, liver and tibial growth plates.

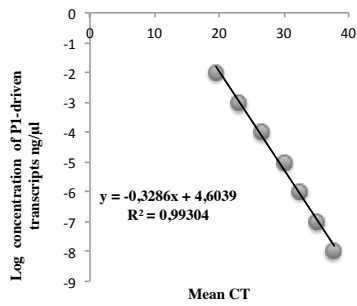
		<b>WBC</b>	<b>PBMC</b>	<b>CD4<sup>+</sup>T lymphocytes</b>	<b>Liver</b>	<b>Tibial growth plate</b>
		<i>N=16</i>	<i>N=20</i>	<i>N=5</i>	<i>N=10</i>	<i>N=3</i>
Promoter P1	CG -1044	87 ± 2	88 ± 4	90 ± 2	34 ± 5*	58 ± 3*
	CG -960	80 ± 2	79 ± 2	87 ± 4	26 ± 4*	21 ± 1*
	CG -919	90 ± 2	88 ± 3	96 ± 6	27 ± 4*	34 ± 2*
	CG -631	86 ± 2	86 ± 2	86 ± 1	50 ± 15*	60 ± 7*
	CG -611	90 ± 3	92 ± 3	94 ± 2	52 ± 15*	55 ± 8*
	CG -491	83 ± 3	84 ± 3	85 ± 5	46 ± 18*	59 ± 7*
	CG -414	12 ± 3	12 ± 2	15 ± 5	9 ± 6	5 ± 2
	CG -308	10 ± 3	11 ± 3	16 ± 6	7 ± 8	11 ± 4
	CG -225	8 ± 1	8 ± 3	10 ± 1	7 ± 2	7 ± 1

Promoter P2	CG -232	64 ± 4	61 ± 5	68 ± 4	28 ± 4*	27 ± 3*
	CG -224	68 ± 5	73 ± 6	72 ± 3	29 ± 5*	28 ± 3*
	CG -218	66 ± 5	73 ± 6	74 ± 3	30 ± 4*	31 ± 1*
	CG -207	38 ± 7	45 ± 7	35 ± 3	12 ± 2*	17 ± 1*
	CG -137	47 ± 3	45 ± 4	50 ± 4	18 ± 4*	19 ± 2*
	CG -108	58 ± 6	60 ± 4	64 ± 3	25 ± 6*	22 ± 2*
	CG -77	50 ± 3	44 ± 5	51 ± 3	21 ± 3*	21 ± 3*
	CG +97	18 ± 2	17 ± 2	18 ± 1	9 ± 2*	10 ± 2*

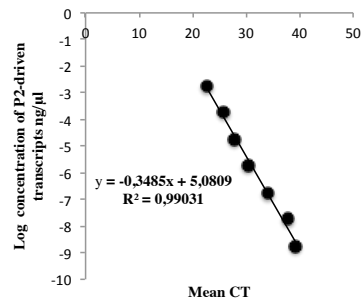
*\*10<sup>-6</sup> < P < 0.03 for comparison with CG methylation in PBMC.*

**Additional file 2: Figure S1.** Standard curves used for the measurement of *IGF1* gene transcript content in PBMC.

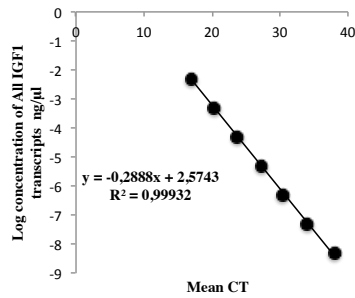
**(A)** Standard Curve of P1-driven transcripts



**(B)** Standard Curve of P2-driven transcripts



**(C)** Standard Curve of global IGF1 transcripts



## 9.2 CONCLUSIONS DE L'ARTICLE 1

La méthylation du gène *IGF1* n'a jamais été étudié chez l'homme jusqu'à très récemment (Desgagné et al., 2014)

Ce travail a permis plusieurs observations cohérentes. Citons :

1. Une variabilité interindividuelle de la méthylation d'ADN au niveau du promoteur 2 beaucoup plus importante que celle du promoteur 1. Nous pouvons donc considérer ce promoteur comme un VMR (Variably Methylated Region).
2. La méthylation du promoteur P2 n'est pas affectée par l'âge ni par le sexe. Le promoteur P2 du gène *IGF1* n'est pas un aDMR.
3. Nous avons montré que les deux promoteurs du gène *IGF1* sont des t-DMR. Le niveau de méthylation dans des chondrocytes tibiaux et dans le foie est beaucoup plus faible dans ces tissus que dans les cellules du sang.
4. Parmi les cellules sanguines, les niveaux de méthylation étaient comparables dans les WBC, PBMCs, et les lymphocytes T CD4+.
5. La méthylation des CGs est corrélée au sein de chacun des promoteurs, mais pas entre les promoteurs.
6. La méthylation du promoteur P2 est inversement corrélée à la taille des enfants et l'IGF1 circulant. Il est remarquable que la méthylation du promoteur P2 contribue pour 13% à la variance de la taille sachant qu'au maximum la contribution d'un variant génétique est de 0.3% (Lang.Allen.H,2010,Yang.J,2010).
7. Un CG du promoteur P1 (CG-611) est inversement corrélé avec la taille et le taux d'IGF1 circulant.
8. La méthylation du même groupe du promoteur P2 est inversement corrélée à l'IGF1 circulant. Il est remarquable que la méthylation du CG-137 du promoteur P2 contribue pour 10% à la variance de l'IGF1 circulant.
9. Nous avons confirmé dans notre échantillon de sujets qu'IGF1 circulant et taille sont positivement corrélés.
10. Dans les cellules sanguines PBMCs, les transcrits du gène *IGF1* issus du promoteur P2 sont inversement corrélés avec la méthylation du promoteur 2.
11. La méthylation artificielle du promoteur P2 d'IGF1 dans une construction plasmidique bloque la transcription. Ces expériences ont été conduites par Yasemin Gunes, ingénieur Inserm dans notre équipe.

Nous proposons que le promoteur P2 soit considéré comme le premier « épIQTL » décrit à notre connaissance chez l'homme (Johannes.F, 2008).

Une application potentielle de notre observation est l'augmentation séculaire de la taille (Bogin B, 2013). Nous pourrions spéculer que la méthylation du promoteur P2 du gène *IGF1* crée un lien épigénétique entre des facteurs génétiques et environnementaux influençant la taille humaine et puisse ainsi contribuer à l'augmentation séculaire de la croissance observée dans certaines populations génétiquement stables. En effet, les modifications de profil de méthylation d'ADN peuvent être héritées de façon intergénérationnelle et même trans-générationnelle. Une accumulation de modifications épigénétiques au fil des générations pourrait contribuer, si elle se poursuivait dans le même sens à l'adaptation de la taille à des changements environnementaux.

## 10 ARTICLE 2 : « THE P2 PROMOTER OF THE IGF1 GENE IS A MAJOR EPIGENETIC LOCUS FOR GH RESPONSIVENESS »

### 10.1 OBJECTIFS DE L'ARTICLE 2

Des centaines de milliers d'enfants de petite taille sont traités par l'hormone de croissance recombinante (GH) pour améliorer leur taille finale adulte. Mais l'efficacité de ce traitement est très variable d'un individu à un autre. Cette variabilité peut être due à l'étiologie de la petite taille ou aux régimes et la compliance au traitement (Rosenfeld and Bakker, 2008) ainsi qu'à certains polymorphismes génétiques, que notre équipe a identifiés il y a dix ans (Dos Santos et al., 2004).

La variation individuelle de production d'IGF1 pendant le traitement par l'hormone de croissance est le mécanisme central de l'augmentation de la taille sous traitement, à laquelle elle est corrélée, mais est très variable selon les enfants (Cohen et al., 2010).

Chez les enfants de petite taille idiopathique, les augmentations corrélées de la taille et du taux d'IGF1 induites par le traitement GH sont normalement distribuées et peuvent donc être modélisées comme un caractère quantitatif continu.

Nous nous sommes intéressés à la réponse au traitement par l'hormone de croissance et sa grande variabilité individuelle. Les facteurs épigénétiques sont susceptibles de faire varier l'efficacité d'un traitement mais les résultats publiés dans ce domaine étaient encore quasi inexistants.

Etant donnée que l'hormone de croissance active directement l'expression du gène *IGF1*, l'hypothèse épigénétique nous paraissait fondée.

Notre hypothèse de travail était que la variation individuelle de la méthylation des CGs situés dans les deux promoteurs du gène *IGF1* module son expression sous GH, et pourraient donc contribuer à la variation individuelle de l'augmentation de la production de l'IGF1 et de la croissance de l'enfant.

Nous avons sélectionné un groupe d'enfants très petits et impubères.

Pour explorer la relation entre la méthylation des promoteurs l'*IGF1* et la réponse à la GH, nous avons étudié :

1. l'effet d'un an de traitement par la GH sur la vitesse de croissance ;
2. l'effet de ce traitement sur la concentration circulante d'IGF1 ;
3. la relation entre les deux réponses ;
4. l'effet direct de la GH sur la transcription du gène *IGF1* dans les cellules de patients exposés à l'effet aigu de la GH.



## ORIGINAL ARTICLE

The P2 promoter of the *IGF1* gene is a major epigenetic locus for GH responsivenessM Ouni<sup>1</sup>, MP Belot<sup>1</sup>, AL Castell<sup>2</sup>, D Fradin<sup>1</sup> and P Bougnères<sup>1,2</sup>

Short children using growth hormone (GH) to accelerate their growth respond to this treatment with a variable efficacy. The causes of this individual variability are multifactorial and could involve epigenetics. Quantifying the impact of epigenetic variation on response to treatments is an emerging challenge. Here we show that methylation of a cluster of CGs located within the P2 promoter of the insulin-like growth factor 1 (*IGF1*) gene, notably CG-137, is inversely closely correlated with the response of growth and circulating IGF1 to GH administration. For example, variability in CG-137 methylation contributes 25% to variance of growth response to GH. Methylation of CGs in the P2 promoter is negatively associated with the increased transcriptional activity of P2 promoter in patients' mononuclear blood cells following GH administration. Our observation indicates that epigenetics is a major determinant of GH signaling (physiology) and of individual responsiveness to GH treatment (pharmacoeigenetics).

*The Pharmacogenomics Journal* advance online publication, 14 April 2015; doi:10.1038/tpj.2015.26

## INTRODUCTION

Tens of thousands of children affected by various causes of short stature currently receive recombinant growth hormone (GH) to improve their final height. However, the high variability of individual GH responsiveness results in unequal growth improvement. The large individual variability of the therapeutic response to GH has puzzled pediatric endocrinologists for decades. The causes underlying such variability have until now been searched in the aetiologies of short stature, treatment regimens, patients' compliance<sup>1</sup> and genetic polymorphisms.<sup>2</sup> This individual variation is partly due to the variable insulin-like growth factor 1 (*IGF1*) production under GH treatment, reflected by circulating IGF1 concentration,<sup>3</sup> thus serum IGF1 measurement can be used to adjust GH dosage in treated children.<sup>3</sup> Despite its importance to therapeutics, the variation of GH responsiveness across treated patients has prompted few studies in search of biological mechanisms.<sup>4</sup> The increment in growth rate induced by GH treatment in healthy children with 'idiopathic' short stature is normally distributed and can thus be modeled as a continuous quantitative trait. Genetic factors certainly have a role.<sup>4</sup> Notably, the deletion of exon 3 within the GH receptor (*GHRd3*) gene has been recognized as a significant pharmacogenetic predictor of GH growth-promoting effects in children with idiopathic short stature.<sup>2</sup> Children carrying *GHRd3* also show higher circulating IGF1 in response to GH injection.<sup>5</sup> Despite this first advance, the variability of GH responsiveness in children with 'idiopathic' short stature has yet received limited molecular explanation.

Pharmacoeigenomics is a nascent field of clinical medicine that holds many promises, but has not yet produced tangible results.<sup>6</sup> The methylation of the cytosine within CG dinucleotides is the simplest component of DNA epigenetics that can be studied in patients receiving a treatment. Among the millions of CG residues in the human genome sequence, a yet unknown number of regions were found to show individual variation of CG

methylation in a given cell population. These regions contain CG residues that are expected to contribute to the individual variability of human phenotypes, provided that such CGs are significantly involved in the regulation of neighboring gene transcription. In many experimental studies, the methylation of CG residues located within low CG-rich promoters has been recognized as a potentially major factor for gene regulation.<sup>7</sup>

To explore the epigenetic component of the individual variability of growth and circulating IGF1 responses to treatment with GH, we thought there was no better physiological candidate than the *IGF1* gene, a key player in postnatal growth, and GH signaling. Inactivating mutations in *IGF1* alter postnatal growth in humans<sup>8</sup> and mice.<sup>9</sup> In contrast, common genetic variation in *IGF1* gene does not contribute significantly to adult height variation in Caucasians<sup>10</sup> but do so in Asians<sup>11</sup> in whom minor allele frequency is greater. Estimates of the proportion of variance in circulating IGF1 that is genetically determined vary between 38% and >80% according to twin studies.<sup>12</sup> The association of circulating IGF1 with several genetic variants is debated,<sup>13</sup> but variants at the *IGF1* locus do not seem to influence circulating IGF1 levels in Caucasians.<sup>13</sup>

Our working hypothesis was that epigenetic marks located in regulatory regions might have a role in modulating *IGF1* gene expression, as observed for many genes, and could therefore contribute to the individual variation of IGF1 production and child's growth. We therefore focused on the two promoters that are directly involved in the regulation of *IGF1* gene expression<sup>14</sup> and are CG-poor promoters expected to exhibit inter-individual variation.<sup>15</sup> The choice of a candidate gene approach instead of commercially available arrays made us able to quantify the methylation of each CG of the *IGF1* promoters. Indeed, individual CG may have a significant functional role possibly different from its CG neighbors. GH has been shown to directly stimulate transcription of the *IGF1* gene in rats and mice.<sup>16,17</sup> GH exerts its effects through the JAK/Stat pathway with translocation of

<sup>1</sup>Institut National de la Santé et de la Recherche Médicale U986, Paris Sud University, Bicêtre Hospital, Le Kremlin-Bicêtre, France and <sup>2</sup>Department of Pediatric Endocrinology and Diabetes, Paris Sud University, Bicêtre Hospital, Le Kremlin-Bicêtre, France. Correspondence: Professor P Bougnères, Department of Pediatric Endocrinology and Diabetes, Paris Sud University, Bicêtre Hospital, 80 avenue du Général Leclerc, Le Kremlin-Bicêtre 94276, France.

E-mail: pierre.bougneres@inserm.fr

Received 24 November 2014; revised 2 February 2015; accepted 9 February 2015

activated Stat5b transcription factor to the nucleus where it regulates *IGF1* transcription.<sup>18</sup> GH-induced transcription promotes accumulation of all classes of *IGF1* mRNA.<sup>16,17,19</sup> Class 1 transcripts have their initiation sites on exon 1 and are driven by P1 promoter, whereas class 2 transcripts use exon 2 as a leader exon (P2) and are driven by P2 promoter.<sup>20,21</sup>

In growing children, GH responsiveness is important to both physiology and therapeutics. To explore the relation between *IGF1* promoter methylation and response to GH in growing children, we selected children who have not entered puberty to avoid the confounding effect of the variable tempo of sexual maturation, which adds to the variability of growth and circulating IGF1. We used the long-studied 'generation test'<sup>5</sup> to test the direct effect of GH on circulating IGF1 and transcription of *IGF1* gene in blood cells of 40 children with idiopathic short stature yet naive to GH treatment. Whether P2 CG methylation could influence the therapeutic efficacy of GH was our next question. To study whether the therapeutic response to GH differs across the various levels of promoter P2 methylation, 136 children with so called 'idiopathic' short stature were studied during their first year of GH administration for both increment in growth rate and in circulating IGF1.

## MATERIALS AND METHODS

### Participants

136 children who had varying degrees of 'idiopathic short stature' were treated with recombinant GH (Table 1). All of them were healthy, had normal clinical examination and no signs of puberty (girls showed no breast development and boys had unmeasurable testosterone levels). GH deficiency was excluded with a stimulated GH peak  $>15 \text{ ng ml}^{-1}$ . All subjects had normal TSH levels. Subtle chondrodysplasia were excluded by forearm, pelvis and spine radiographs. Trained nurses performed height measurements in duplicate using the Harpenden stadiometer. Blood samples were obtained before onset of GH treatment. We did not carry out this study using GH deficient children because the causes for this deficiency are highly heterogeneous (mutations of pituitary transcription factors, irradiation for cancer, hypothalamic tumors and so on) and are often associated with other hormone deficits or medical problems.

At onset of treatment, an acute test was performed by injecting  $100 \mu\text{g kg}^{-1}$  of recombinant GH intramuscularly to 40 children and sampling blood before and 12 h after the injection. Thereafter, children were all followed by a pediatric endocrinologist for the management of GH treatment. Height (Harpenden stadiometer) and serum IGF1 measurements were performed at 6, 9 and 12 months of treatment.

For methylation and transcript measurements, 10 ml peripheral blood samples were obtained, from which white blood cells and/or PBMC (peripheral blood mononuclear cells) were purified immediately. White and mononuclear cell were counted at the time of sampling. For measuring

transcripts, PBMC were collected at the clinical center close to the laboratory and mRNA was extracted immediately.

Parents of all studied children gave their written informed consent for the current study and for using surgical specimens, according to the French rules of bioethics in biomedical research checked by our Institutional Review Board.

### Serum IGF1 concentrations

Serum IGF1 concentration was measured around 0700 to 0800 hours before breakfast in 136 children using an immune-radiometric assay after ethanol-acid extraction using DSL-5600 Active (Diagnostic System Laboratories, Webster, TX, USA) or Cisbio reagents. Intra- and inter-series coefficients of variation were 1.5% and 3.7% at  $260 \text{ ng ml}^{-1}$  and 3.9% at  $760 \text{ ng ml}^{-1}$ . The sensitivity was  $4 \text{ ng ml}^{-1}$ . IGF1 SDS were calculated using the norms of Alberti *et al.*<sup>22</sup> in French children.

### DNA methylation at CpG resolution in IGF1 promoters 1 and 2

For promoter P1, we studied nine CGs located over a 800 bp distance, the closest CG being 225 bp upstream from the corresponding major transcription start site<sup>23</sup> (Figure 1). For promoter P2, we studied 7/8 CGs located upstream from the major transcription start site within the proximal part of the promoter and one CG located 97 bp downstream this transcription start site (Figure 1). CGs are denominated according to their position versus each promoter transcription start site. The methylation of CG-22 could not be measured for technical reasons. Nucleic acids were extracted from white blood cells or PBMC using Genra Puregene blood kit (Qiagen, Hilden, Germany). A bisulfite-PCR-pyrosequencing technique<sup>24</sup> was used to measure the methylation of the CGs. We improved the resolution of this method from a handful of bases to up to 100 nucleotides, with the ability to quantify methylation in the same sample of blood with a coefficient of variation (s.d./mean) as little as 1–5%. Briefly, 400 ng of genomic DNA were treated with EZ DNA Methylation-Gold Kit, according to manufacturer's protocol (Zymo Research Corporation, CA, USA). The bisulfite-treated genomic DNA was PCR-amplified using unbiased *IGF1* primers (Supplementary Methods Table 1) and performed quantitative pyrosequencing using a PyroMark Q96 ID Pyrosequencing instrument (Qiagen). Pyrosequencing assays were designed using MethPrimer (<http://www.urogene.org/methprimer/index1.html>). Biotin-labeled single stranded amplicons were isolated according to protocol using the Qiagen Pyromark Q96 Work Station and underwent pyrosequencing with  $0.5 \mu\text{m}$  primer. The percent methylation for each of the CGs within the target sequence was calculated using PyroQ CpG Software (Qiagen).

### Study of IGF1 transcripts in PBMC

Methods are detailed in Supplementary Methods Table 2.

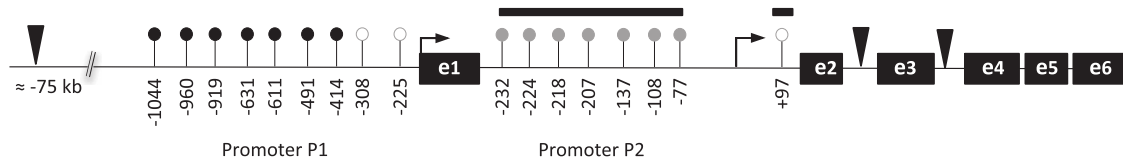
### Calculations and statistical methods

IGF1 levels and height were expressed as SDS to adjust for age and sex. The growth rate response to GH administration was expressed as increment in growth rate, the difference between growth rate during GH treatment (in cm per year) and previous growth rate (evaluated during the whole year before onset of GH administration). We chose this quantitative criterion because spontaneous growth in children with idiopathic short stature is linear during this period of childhood. Correlations were calculated as adjusted R square that measures the proportion of the variation in the dependent variable accounted for by the explanatory variables. The fraction of explained variance is calculated under the linear regression model, using the usual definition:  $r^2 \times 100$ . We fitted a multivariate linear model to the data to estimate the proper effect of CG methylation on response to GH, adjusted for the effect of the other covariates contributing to the growth under treatment, such as age at diagnosis, sex and the received dose of GH. This approach is suitable for estimating the association between the variable of interest, here the methylation level, and the trait in the presence of correlation between the covariates. We carried out tests of independence of each covariate one at a time, keeping the others in the model. Statistics and estimations of effect given in the tables are thus adjusted for the others whenever appropriate and are not subject to marginal association. We checked the normality of the residuals, and the residuals versus the fitted values did not show any trend, indicating that there was no noticeable deviation from the assumption of the linear model. All statistics and linear model were computed using R 2.10.1. Results are expressed as mean  $\pm$  s.d.

**Table 1.** Main baseline characteristics of the studied children.

	Acute GH Injection test	GH-treated Children
N	40	136
Sex (M/F)	27/13	72/64
Age (years)	$10.8 \pm 1.8$	$8.6 \pm 3$
Height (SDS)	$-1.8 \pm 0.7$	$-2.2 \pm 0.7$
Growth rate before GH (SDS)	$-1.3 \pm 1.2$	$-1.3 \pm 1.2$
Serum IGF 1 before GH (SDS)	$-1.2 \pm 0.8$	$-1.3 \pm 1$
rhGH treatment dose ( $\text{mg kg}^{-1}$ per week)	—	$0.47 \pm 0.12$
Growth rate at 1 year GH (SDS)	—	$4.6 \pm 1.9$
Serum IGF 1 at 1 year GH (SDS)	—	$0.6 \pm 1.5$

Abbreviations: F, female; GH, growth hormone; IGF1, insulin-like growth factor 1; M, male; SDS, standard deviation score.



**Figure 1.** Schematic representation of the human *IGF1* gene with its two promoters (P1, P2). The three closest Stat5b binding sites are figured as black triangles (▼). The studied CGs are shown as lollypops within the two promoters. Mean methylation levels are figured as (● >75%; ◐ 40-75%; ○ <20%). TSS are shown as broken arrows. The horizontal bar encompasses the eight CGs that show the strongest association with response to GH.

**Table 2.** General linear model for regression of sex, age, GH dose and CG-137 methylation on increment in growth rate (cm per year) induced by GH treatment

	Estimates	s.e.	t-value	Pr(> t )
Intercept	13.21	1.7	7.8	$2 \times 10^{-12}$
Age at GH onset (years)	0.13	0.05	2.46	0.015
Sex	0.03	0.28	0.12	0.91
GH dose (mg kg <sup>-1</sup> per week)	0.014	$6.7 \times 10^{-3}$	2.04	0.04
CG-137 methylation (%)	-0.22	0.032	-6.8	$4 \times 10^{-10}$

## RESULTS

Relationship between CG methylation and response of circulating IGF1 and IGF1 transcripts to a first GH injection

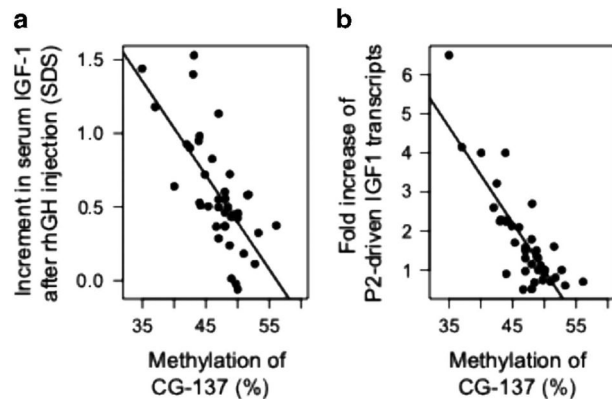
The methylation levels for the CGs located within the two promoters of *IGF1* is given in Supplementary Table S1.

Following the acute injection of GH, the increase in circulating IGF1 was inversely related with CG-137 methylation ( $P = 3 \times 10^{-7}$ ), which accounted for as much as ~49% of the variability of this increase (Figure 2a and Supplementary Table 2) and with the methylation of 3/8 other P2 CGs (Figure 1 and Supplementary Table 2).

In order to test whether methylation at the P2 promoter affects the transcriptional response of the *IGF1* gene to GH, we measured P1-driven, P2-driven and total *IGF1* transcripts in PBMC before and 12 h after the GH injection. We found a variable increase in *IGF1* transcripts across the studied children. The increase in P2-driven transcripts showed a very strong inverse correlation with CG-137 methylation ( $P = 5 \times 10^{-11}$ ; Figure 2b) and with 4/8 other P2 CGs (Supplementary Table 3). Among the CGs of P1 promoter, only CG-611 showed an inverse correlation with P1-driven transcripts (Supplementary Table 3). Methylation was unchanged by this brief exposition to GH (not shown). We were not able to assess direct GH effects on cultured PBMC from the patients, because when these cells are submitted to short culture conditions, they show an unreliable response of *IGF1* expression to stimulation by GH.

### IGF1 and growth response to GH treatment

As expected, GH-induced growth responses were variable, closely fitting the normal distribution (Supplementary Figure 1a) and were correlated with circulating IGF1 ( $P = 3 \times 10^{-5}$ ; Supplementary Figure 2). A strong inverse relationship was observed between the methylation of 8/8 CGs of the P2 promoter and growth acceleration in response to GH administration (Supplementary Table 4). Again, the correlation was maximal for CG-137 methylation ( $P = 2.7 \times 10^{-10}$ ; Figure 3a and Supplementary Figure 1a), which accounted for ~25% of the variability in the response to GH. We confirmed this finding by building a general linear model for regression of age, sex, GH dose and CG-137 methylation on growth rate increment (Table 2). Comparable patterns of correlation were found for the increase in serum IGF1 ( $P = 1.7 \times 10^{-10}$  for



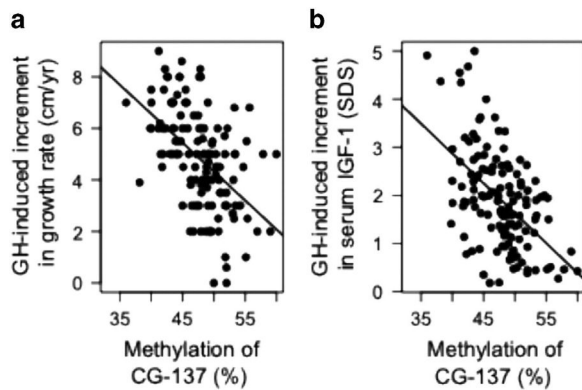
**Figure 2.** (a) CG-137 methylation correlates negatively with the increment in serum IGF1 after the first GH injection ( $N = 40$ ;  $R = -0.7$ ,  $P = 3.4 \times 10^{-7}$ ); (b) CG-137 methylation correlates negatively with the increment in P2-driven *IGF1* transcripts in PBMC in response to acute GH injection in 40 children ( $R = -0.82$ ,  $P = 5 \times 10^{-11}$ ).

CG-137; Figure 3b and Supplementary Figure 1b, Supplementary Table 5).

## DISCUSSION

The consistency and statistical strength of the observed correlations support that the CG methylation of the P2 promoter of the *IGF1* gene is a major determinant of the individual response to GH treatment across children with idiopathic short stature. The effect size of this association is comparable with the association previously reported between growth rate response and the common GH receptor d3 deletion variant.<sup>2</sup> Furthermore, our observations suggest that the epigenetic association reported herein has functional relevance since methylation of CG-137, as well as methylation of the other CGs of the P2 promoter, is associated with the transcriptional effects initiated at the P2 promoter following GH injection. Mice models have revealed that GH effects on skeletal growth are mediated by IGF1 produced in situ by growth plate chondrocytes, not by circulating IGF1.<sup>25</sup> Like other studies before,<sup>26</sup> we found a strong correlation





**Figure 3.** (a) CG-137 methylation correlates negatively with the GH-induced increment in growth rate in 136 children treated with GH ( $R = -0.50$ ,  $P = 2.7 \times 10^{-10}$ ); (b) CG-137 methylation correlates negatively with the GH-induced increment in serum IGF1 concentration in children treated with GH ( $R = -0.52$ ,  $P = 1.7 \times 10^{-10}$ ).

between the increase in height and the increase in circulating IGF1. As circulating IGF1 originates mostly in the liver,<sup>27</sup> the latter correlation suggests that the GH-induced levels of *IGF1* expression in liver and chondrocytes are proportional in a given individual, and could share some of the epigenetic regulation reported in the current study.

Clearly however, a weakness of our study is that we could not study the association of P2 methylation with transcription or GH responsiveness in liver and growth plates, the physiological tissues regulating GH effects on IGF1 production and skeletal growth. Hepatocytes and chondrocytes have their own epigenome. Such lack of availability and analysis of epigenomes in specific cells of physiological tissues is a common but major limitation of epigenetic epidemiology,<sup>28–30</sup> which most often has to rely on blood cells. Herein the observations made in PBMC can only be extrapolated to liver and growth plate through physiological speculation.

Other questions left unanswered include the molecular mechanisms taking place at the P2 promoter locus under GH effect and correlating CG methylation with the regulation of gene transcription. This correlation was observed in the mononuclear blood cells of our patients following the first GH injection. Liver expression of the *IGF1* gene is mainly controlled at the transcriptional level by GH from the pituitary.<sup>31</sup> In mammals including humans, the *IGF1* gene is composed of six exons and five introns that span >80 kb of chromosomal DNA<sup>32,33</sup> which are located in 12q23.2 in humans. Tandem promoters direct *IGF1* gene transcription through unique leader exons. Promoter 1, which uses heterogeneous transcription initiation sites, is active in multiple animal tissues,<sup>34</sup> while the smaller and simpler promoter 2 is primarily but not exclusively active in the liver of cattle,<sup>19</sup> unlike in rodents where promoter 2 activity seems exclusively hepatic.<sup>14</sup> Although the biochemical mechanisms responsible for different tissue-specific patterns of IGF1 promoter activity are unknown, the DNA sequences of both proximal promoters are relatively well-conserved in mammals based on analyses of available genomic databases (74% over 420 nucleotides for promoter 1, 58% over 404 nucleotides for promoter 2 between rat and human *IGF1*),<sup>14</sup> suggesting that functional properties of each promoter have been maintained during speciation as essential aspects of the biology of *IGF1* gene regulation.<sup>14</sup> GH exerts its effects through the JAK/Stat pathway with the translocation of activated Stat5b transcription factor to the nucleus where it regulates IGF1 transcription.<sup>18</sup> Recent results suggest that GH-induced Stat5 activation of *IGF1* gene expression in mouse liver might be collectively mediated by at least eight Stat5 binding sites located in distal intronic and

5'-flanking regions of the *IGF1* gene, distantly from the IGF1 promoter.<sup>35</sup> The identification of multiple distal Stat5 binding sites underscores the complexity of the mechanism that mediates GH regulation of *IGF1* gene expression. Active Stat5b interacts with multiple DNA binding sites in chromatin within the *IGF1* locus, and through mechanisms not yet characterized, promotes the rapid transmission of information to the two *IGF1* promoters, culminating in induction of *IGF1* gene transcription and production of IGF1 mRNAs and protein.<sup>14</sup> In the liver of hypophysectomised rats, GH induces dramatic changes in chromatin at the *IGF1* locus and activates *IGF1* transcription by distinct promoter-specific epigenetic mechanisms.<sup>17,36</sup> The proximal part of rat P2 is an important site of transcriptional regulation by GH via Stat5b.<sup>14</sup> In rat liver, GH induces rapid and dramatic changes in chromatin at the P2 promoter and activates *IGF1* transcription by specific epigenetic mechanisms.<sup>17,36</sup> At promoter P2, GH facilitates recruitment then activation of RNA Pol II to initiate transcription, whereas at promoter P1, GH causes RNA Pol II to be released from a previously recruited poised and paused pre-initiation complex.<sup>17</sup> These recent advances on epigenetic mechanisms involving chromatin landscape in rodent liver,<sup>17</sup> as well as our observation of a relationship between DNA methylation variation and *IGF1* transcripts in PBMC, provide an impetus to address fundamental mechanistic questions that will help decipher the epigenetic regulation of *IGF1* gene expression in baseline conditions and in response to GH. Yet, not only CG location and composition are different in human and rat *IGF1* promoters (personal observation) but the pattern of methylation and its transcriptional effects on rat promoters are still unknown.

To our knowledge, the current study offers the first clinical evidence of a link between DNA methylation and the response to a treatment in humans, illustrating the role of epigenetic variation as a potent contributor to personalized therapeutics.

## CONFLICT OF INTEREST

The authors declare no conflict of interest

## ACKNOWLEDGMENTS

We thank Pfizer France for supporting the GH pharmacoepigenomics program. We are grateful to their colleagues of the Department of Pediatric Endocrinology for allowing the study of their GH-treated patients, and to the nurses who help carry clinical research.

## REFERENCES

- Rosenfeld RG, Bakker B. Compliance and persistence in pediatric and adult patients receiving growth hormone therapy. *Endocr Pract* 2008; **14**: 143–154.
- Dos Santos C, Essieux L, Teinturier C, Tauber M, Goffin V, Bougnères P et al. A common polymorphism of the growth hormone receptor is associated with increased responsiveness to growth hormone. *Nat Genet* 2004; **36**: 720–724.
- Cohen P, Germak J, Rogol AD, Weng W, Kappelgaard AM, Rosenfeld RG et al. Variable degree of growth hormone (GH) and insulin-like growth factor (IGF) sensitivity in children with idiopathic short stature compared with GH-deficient patients: evidence from an IGF-based dosing study of short children. *J Clin Endocrinol Metab* 2010; **95**: 2089–2098.
- Rosenfeld RG. The pharmacogenomics of human growth. *J Clin Endocrinol Metab* 2006; **91**: 795–796.
- Toyoshima MTK, Castroneves LA, Costalonga EF, Mendonca BB, Arnhold JJ, Jorge AA et al. Exon 3-deleted genotype of growth hormone receptor (GHRD3) positively influences IGF-1 increase at generation test in children with idiopathic short stature. *Clin Endocrinol (Oxf)* 2007; **67**: 500–504.
- Ivanov M, Kacevska M, Ingelman-Sundberg M. Epigenomics and interindividual differences in drug response. *Clin Pharmacol Ther* 2012; **92**: 727–736.
- Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* 2007; **39**: 457–466.

- 8 Woods KA, Camacho-Hübner C, Savage MO, Clark AJ. Intrauterine growth retardation and postnatal growth failure associated with deletion of the insulin-like growth factor I gene. *N Engl J Med* 1996; **335**: 1363–1367.
- 9 Lupu F, Terwilliger JD, Lee K, Segre GV, Efstratiadis A. Roles of growth hormone and insulin-like growth factor 1 in mouse postnatal growth. *Dev Biol* 2001; **229**: 141–162.
- 10 Lettre G, Butler JL, Ardlie KG, Hirschhorn JN. Common genetic variation in eight genes of the GH/IGF1 axis does not contribute to adult height variation. *Hum Genet* 2007; **122**: 129–139.
- 11 Okada Y, Kamatani Y, Takahashi A, Matsuda K, Hosono N, Ohmiya H *et al*. A genome-wide association study in 19 633 Japanese subjects identified LHX3-QSOX2 and IGF1 as adult height loci. *Hum Mol Genet* 2010; **19**: 2303–2312.
- 12 Kao PC, Matheny AP, Lang CA. Insulin-like growth factor-I comparisons in healthy twin children. *J Clin Endocrinol Metab* 1994; **78**: 310–312.
- 13 Palles C, Johnson N, Coupland B, Taylor C, Carvajal J, Holly J *et al*. Identification of genetic variants that influence circulating IGF1 levels: a targeted search strategy. *Hum Mol Genet* 2008; **17**: 1457–1464.
- 14 Rotwein P. Mapping the growth hormone–Stat5b–IGF-I transcriptional circuit. *Trends Endocrinol Metab* 2012; **23**: 186–193.
- 15 Bock C, Walter J, Paulsen M, Lengauer T. Inter-individual variation of DNA methylation and its implications for large-scale epigenome mapping. *Nucleic Acids Res* 2008; **36**: e55.
- 16 Bichell DP, Kikuchi K, Rotwein P. Growth hormone rapidly activates insulin-like growth factor I gene transcription *in vivo*. *Mol Endocrinol* 1992; **6**: 1899–1908.
- 17 Chia DJ, Young JJ, Mertens AR, Rotwein P. Distinct alterations in chromatin organization of the two IGF-I promoters precede growth hormone-induced activation of IGF-I gene transcription. *Mol Endocrinol* 2010; **24**: 779–789.
- 18 Heim MH. The Jak-Stat pathway: cytokine signalling from the receptor to the nucleus. *J Recept Signal Transduct* 1999; **19**: 75–120.
- 19 Wang Y, Price SE, Jiang H. Cloning and characterization of the bovine class 1 and class 2 insulin-like growth factor-I mRNAs. *Domest Anim Endocrinol* 2003; **25**: 315–328.
- 20 Yang H, Adamo ML, Koval AP, McGuinness MC, Ben-Hur H, Yang Y *et al*. Alternative leader sequences in insulin-like growth factor I mRNAs modulate translational efficiency and encode multiple signal peptides. *Mol Endocrinol Baltim Md* 1995; **9**: 1380–1395.
- 21 Adamo ML, Ben-Hur H, Roberts CT, LeRoith D. Regulation of start site usage in the leader exons of the rat insulin-like growth factor-I gene by development, fasting, and diabetes. *Mol Endocrinol* 1991; **5**: 1677–1686.
- 22 Alberti C, Chevenne D, Mercat I, Josserand E, Armoogum-Boizeau P, Tichet J *et al*. Serum concentrations of insulin-like growth factor (IGF)-1 and IGF binding protein-3 (IGFBP-3), IGF-1/IGFBP-3 ratio, and markers of bone turnover: reference values for French children and adolescents and z-score comparability with other references. *Clin Chem* 2011; **57**: 1424–1435.
- 23 Jansen E, Steenbergh PH, LeRoith D, Roberts CT, Sussenbach JS. Identification of multiple transcription start sites in the human insulin-like growth factor-I gene. *Mol Cell Endocrinol* 1991; **78**: 115–125.
- 24 Tost J, Gut IG. DNA methylation analysis by pyrosequencing. *Nat Protoc* 2007; **2**: 2265–2275.
- 25 Liu JL, Yakar S, LeRoith D. Conditional knockout of mouse insulin-like growth factor-1 gene using the Cre/loxP system. *Proc Soc Exp Biol Med* 2000; **223**: 344–351.
- 26 Kriström B, Lundberg E, Jonsson B, Albertsson-Wikland K. & study group. IGF-1 and growth response to adult height in a randomized GH treatment trial in short non-GH-deficient children. *J Clin Endocrinol Metab* 2014; **99**: 2917–2924.
- 27 Yakar S, Pennisi P, Wu Y, Zhao H, LeRoith D in *Endocrine Development*, (eds) Cianfarani S, Clemmons DR, & Savage MO 11–16. KARGER: S Karger AG, Basel, 2005 at < <http://www.karger.com/gate2.inist.fr/Article/FullText/85718> > .
- 28 Rakyant VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet* 2011; **12**: 529–541.
- 29 Murrell A, Rakyant VK, Beck S. From genome to epigenome. *Hum Mol Genet* 2005; **14 Spec No 1**: R3–R10.
- 30 Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nat Rev Genet* 2013; **14**: 585–594.
- 31 Daughaday WH, Rotwein P. Insulin-like growth factors I and II. Peptide, messenger ribonucleic acid and gene structures, serum, and tissue concentrations. *Endocr Rev* 1989; **10**: 68–91.
- 32 Rotwein P, Pollock KM, Didier DK, Krivi GG. Organization and sequence of the human insulin-like growth factor I gene. Alternative RNA processing produces two insulin-like growth factor I precursor peptides. *J Biol Chem* 1986; **261**: 4828–4832.
- 33 Rotwein P. Structure, evolution, expression and regulation of insulin-like growth factors I and II. *Growth Factors* 1991; **5**: 3–18.
- 34 Hall LJ, Kajimoto Y, Bichell D, Kim SW, James PL, Counts D *et al*. Functional analysis of the rat insulin-like growth factor I gene and identification of an IGF-I gene promoter. *DNA Cell Biol* 1992; **11**: 301–313.
- 35 Eleswarapu S, Gu Z, Jiang H. Growth hormone regulation of insulin-like growth factor-I gene expression may be mediated by multiple distal signal transducer and activator of transcription 5 binding sites. *Endocrinology* 2008; **149**: 2230–2240.
- 36 Chia DJ, Rotwein P. Defining the epigenetic actions of growth hormone: acute chromatin changes accompany GH-activated gene transcription. *Mol Endocrinol* 2010; **24**: 2038–2049.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Supplementary Information accompanies the paper on the The Pharmacogenomics Journal website (<http://www.nature.com/tpj>)

## **Supplementary Information**

### **The IGF1 P2 promoter is a major epigenetic locus for GH responsiveness.**

Ouni M<sup>1</sup>, Belot MP<sup>1</sup>, Castell AL<sup>2</sup>, Fradin D<sup>1</sup>, Bougnères P<sup>1,2</sup>.

1. Institut National de la Santé et de la Recherche Médicale, Paris Sud University, Bicêtre Hospital, Le Kremlin-Bicêtre

2. Dept of Pediatric Endocrinology and Diabetes, Paris Sud University, Bicêtre Hospital, Le Kremlin-Bicêtre

<sup>1</sup>Institut National de la Santé et de la Recherche Médicale U986, Paris Sud University, Bicêtre Hospital, Le Kremlin-Bicêtre, France.

<sup>2</sup>Dept of Pediatric Endocrinology and Diabetes, Paris Sud University, Bicêtre Hospital, Le Kremlin-Bicêtre, France.

## **Legends of Supplementary Figures.**

**Supplementary Figure 1:** Frequency histogram showing the distribution of individual increment in growth rate and serum IGF1 during the first year of growth hormone treatment. The distribution of values was near normal. the increment of growth rate and IGF1 during GH treatment are smaller in children in the higher quartile of CG-137 methylation (figured as grey bars) .

**Supplementary Figure 2:** In response to GH, increments in IGF1 and in growth rate were correlated (N=136) ( $R=0.35$ ,  $P=3.10^{-5}$ ).

**Supplementary Table 1:** CG methylation in the P1 and P2 promoters of the *IGF1* gene. The position of each CG is given as bp from the TSS of the corresponding promoter.

		Percent Methylation mean $\pm$ SD	Methylation range		
			Extreme Values%	Range %	Range/Mean
Promoter P1	CG -1044	88 $\pm$ 3	(78 - 95)	17	0.19
	CG -960	79 $\pm$ 4	(75 - 87)	12	0.15
	CG -919	89 $\pm$ 6	(80 - 98)	18	0.20
	CG -631	86 $\pm$ 2	(75 - 91)	16	0.18
	CG -611	92 $\pm$ 3	(85 - 99)	14	0.15
	CG -491	84 $\pm$ 4	(69 - 93)	24	0.29
	CG -414	12 $\pm$ 4	( 5 - 36)	31	2.58
	CG -308	10 $\pm$ 4	( 3 - 28)	25	2.5
	CG -225	08 $\pm$ 2	( 4 - 14)	10	1.25

Promoter P2	CG -232	66 $\pm$ 7	(46 - 87)	41	0.62
	CG -224	73 $\pm$ 7	(58 - 92)	34	0.46
	CG -218	73 $\pm$ 6	(61 - 89)	28	0.38
	CG -207	44 $\pm$ 8	(20 - 60)	40	0.91
	CG -137	47 $\pm$ 4	(35 - 60)	25	0.53
	CG -108	61 $\pm$ 6	(46 - 74)	28	0.45
	CG -77	48 $\pm$ 5	(35 - 60)	25	0.52
	CG +97	16 $\pm$ 3	( 9 - 24)	15	0.94



**Supplementary Table 2:** Correlation between the methylation of the studied CGs and the increase of serum IGF1 in response to an acute injection of GH (N=40).

		Increment in Serum IGF1 (SDS)	
		R	P
Promoter P1	CG-1044	-0.10	0.46
	CG-960	-0.05	0.74
	CG-919	-0.12	0.22
	CG-631	0.18	0.15
	CG-611	0.03	0.88
	CG-491	0.44	$5.10^{-3}$
	CG-414	0.09	0.58
	CG-308	0.08	0.62
	CG-225	0.03	0.84

Promoter P2	CG-232	-0.26	0.06
	CG-224	-0.42	$4.4.10^{-3}$
	CG-218	-0.24	0.07
	CG-207	-0.46	$1.5.10^{-3}$
	CG-137	-0.7	$3.3.10^{-7}$
	CG-108	-0.33	0.022
	CG-77	0.00	0.99
	CG+97	-0.08	0.61

**Supplementary Table 3:** Correlation between the methylation of the studied CGs and the increase of P1-driven, P2-driven, and total *IGF1* transcripts (N=40). Note that CG-207 in the P2 promoters is associated with the increase in P1-driven transcripts.

		P1-driven Transcripts (Fold Increase)		P2-driven Transcripts (Fold Increase)		ALL IGF1 Transcripts (Fold Increase)	
		R	P	R	P	R	P
Promoter P1	CG -1044	-0.15	0.36	-0.3	0.032	-0.3	0.033
	CG -960	-0.07	0.65	-0.07	0.64	-0.01	0.95
	CG -919	-0.05	0.78	-0.15	0.38	-0.06	0.70
	CG -631	0.07	0.69	0.04	0.80	0.10	0.54
	CG -611	-0.48	0.0011	-0.05	0.78	-0.16	0.18
	CG -491	-0.11	0.51	0.18	0.28	0.08	0.61
	CG -414	-0.06	0.74	0.25	0.14	0.09	0.60
	CG -308	-0.14	0.38	0.05	0.78	-0.03	0.84
	CG -225	-0.08	0.64	-0.04	0.80	-0.11	0.51

Promoter P2	CG -232	0.15	0.33	0.14	0.36	-0.10	0.45
	CG -224	-0.18	0.15	-0.61	$10^{-5}$	-0.60	$2.10^{-5}$
	CG -218	-0.09	0.41	-0.52	$3.6.10^{-4}$	-0.45	$2.2.10^{-3}$
	CG -207	-0.46	0.002	-0.66	$2.10^{-6}$	-0.82	$6.10^{-11}$
	CG -137	0.03	0.86	-0.82	$5.3.10^{-11}$	-0.69	$6.5.10^{-7}$
	CG -108	-0.10	0.23	-0.12	0.22	-0.05	0.72
	CG -77	-0.08	0.61	-0.04	0.83	-0.07	0.67
	CG +97	-0.15	0.17	0.02	0.91	-0.06	0.28

**Supplementary Table 4:** Correlation of the methylation of CGs of promoter P2 and the increase in growth rate induced by GH treatment in 136 children with ISS. Note that CG-611 methylation previously found to correlate with height, showed no correlation with the response to GH.

		Increment in Growth Rate			
		(cm/yr)		(SDS)	
		R	P	R	P
Promoter P1	CG -1044	0.00	0.95	-0.06	0.44
	CG -960	-0.01	0.81	-0.01	0.85
	CG -919	-0.09	0.18	-0.04	0.67
	CG -631	-0.07	0.51	-0.03	0.68
	CG -611	-0.07	0.54	-0.15	0.047
	CG -491	-0.10	0.13	-0.07	0.21
	CG -414	-0.04	0.28	-0.11	0.12
	CG -308	-0.14	0.07	-0.13	0.08
	CG -225	-0.11	0.11	-0.11	0.10

Promoter P2	CG -232	-0.36	$1.45 \cdot 10^{-5}$	-0.38	$3.6 \cdot 10^{-6}$
	CG -224	-0.26	$1.5 \cdot 10^{-3}$	-0.33	$6 \cdot 10^{-5}$
	CG -218	-0.19	0.016	-0.26	$1.6 \cdot 10^{-3}$
	CG -207	-0.20	0.012	-0.22	$8 \cdot 10^{-3}$
	CG -137	-0.50	$2.7 \cdot 10^{-10}$	-0.57	$2.5 \cdot 10^{-13}$
	CG -108	-0.32	$1.7 \cdot 10^{-4}$	-0.36	$1.7 \cdot 10^{-5}$
	CG -77	-0.23	$4 \cdot 10^{-3}$	-0.27	$8 \cdot 10^{-4}$
	CG +97	-0.30	$2 \cdot 10^{-4}$	-0.40	$7 \cdot 10^{-7}$

**Supplementary Table 5.** Correlation of the methylation of CGs of promoter P2 with the increment in circulating IGF1 induced by GH treatment in 136 children with ISS.

		Increment in Serum IGF1 (SDS)	
		R	P
Promoter P1	CG -1044	-0.24	4.10 <sup>-3</sup>
	CG -960	-0.04	0.38
	CG -919	-0.08	0.17
	CG -631	-0.08	0.18
	CG -611	-0.07	0.21
	CG -491	-0.16	0.04
	CG -414	-0.09	0.16
	CG -308	-0.02	0.85
	CG -225	-0.05	0.60

Promoter P2	CG -232	-0.30	5.2.10 <sup>-4</sup>
	CG -224	-0.31	2.6.10 <sup>-4</sup>
	CG -218	-0.24	3.10 <sup>-3</sup>
	CG -207	-0.10	0.12
	CG -137	-0.52	1.7.10 <sup>-10</sup>
	CG -108	-0.28	10 <sup>-3</sup>
	CG -77	-0.17	0.03
	CG +97	-0.17	0.03

**Supplementary Methods Table 1: PCR conditions for methylation studies of *IGF1* Promoters.**

PCR conditions for methylation studies of *IGF1* Promoter 1.

hg18 : chr12 : 101398637 – 101101399558

CpG sites	Taq Polymerase	PCR conditions	Amplification Primers	Amplicon size
-1044, -960, -919	EpiMark® Hot Start Taq DNA Polymerase (NEB)	95°C for 30s [95°C, 30s; 53°C, 30s, 68°C, 30s] x 45 68°C for 5 minutes	F-TATTGTTTTATAAAATTAAGGGAA R- ATTCTAAATAACACCAACTAACTAAC	208 bp
PRIMARY for -631, -611, -491, -414, -308	Taq DNA Polymerase (recombinant) (Thermoscientific) with (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> buffer	95°C for 2 minutes [95°C, 30s; 53°C, 30s, 72°C, 1 minute] x 40 72°C for 7 minutes	F-TGATAGGTAGTTTGTAGTAGAAGAATAA R- AAAACTAAAACACAAAAACATTTTT	528 bp
NESTED 1 -631, -611	Taq DNA Polymerase (recombinant) (Thermoscientific) with (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> buffer	95°C for 2 minutes [95°C, 30s ; 53°C, 30s, 72°C, 30s] x 40 72°C for 7 minutes	F- GTTTTATTTTGTAGTTGGGTTTTATAG R- CTCTACAATTTTAACCCTAAAAT	224 bp
NESTED 2 -491, -414, -308	Taq DNA Polymerase (recombinant) (Thermoscientific) with KCl buffer	95°C for 2 minutes [95°C, 30s; 57°C, 30s, 72°C, 30s] x 40 72°C for 7 minutes	F- GTAATTGGGATAAGGGTTATT R- AACTTTCTACTAAACATAAAAACACAAAC	235 bp

-225	EpiMark® Hot Start Taq DNA Polymerase (NEB)	95°C for 30s [95°C, 30s; 53°C, 30s, 68°C, 30s] x 45 68°C for 5 minutes	F-TTTTAAATATTGGTAAGGTGTATTAG R- TTTCCCTTAAAAAATATATATTAAT	150 bp
------	---	--	--	--------

Primers for pyrosequencing

CpG sites	Primers
-1044	F- AGGGAAATAGGTATAAAT
-960	F- AGTTGGTTTGGATTATGT
-919	F- GGGATTTGAATGATATTA
-631 and -611	R- CTATCTCATACTTTTCT
-491	F- TAATTGGGATAAGGGG
-414	F- ATTTTAGGGTTAAAATTGTAGAGT
-308	F- ATATTGGTAAGGTGTATT
-225	F- GGTAGTATAAGTATTTTA

PCR conditions for methylation studies of *IGF1* Promoter 2 :

hg18 : chr12 : 101396974-101397434

CpG sites	Taq Polymerase	PCR conditions	Amplification Primers	Amplicon size
-232,-224, -218,-207, -137,-108	Methyltaq Diagenode	95°C for 10 minutes [95°C, 30s; 54°C, 1minute, 73°C, 1min] x 45 73°C, 5 minutes	F-AATTGGTTGTTGTTAGTGTAT R- AATTAAACCCTCAAACAATTAAATC	175 pb
-77, +97	Taq'Ozyme HS (Ozyme )	95°C for 1 minute [95°C, 15s; 48°C, 30s 72°C, 30s] x 45 72°C, 5 min	F- AATTGTTGAGGGTTAATTTATAAGA R- ACTACTAAATCACATAATATATACAC A	243 pb

Primers for pyrosequencing

CpG sites	Primers
-232, -224, 218 and -207	F- TGGTTGTTGTTAGTGTAT
-137	R-ACCAATAACAACAACCTAA
-108	F- AAGTTGTTGTTATTGGTT
-77	R- ACTCCCTCAAACCACTTCTACT
+97	R- AATTGTTTGGAGGGTTAATTTATAAGA

**Supplementary Methods Table 2:** Table Real time RT-PCR information.

<b>Experimental design</b>	
Tissue	PBMC
Definition of experimental group	Children with idiopathic short stature
Number	N = 40
<b>Sample</b>	
Description	Four ml of fresh blood were diluted with 4 ml of NaCl 154 mM and carefully overlaid on 4 ml of lymphoprep solution (Eurobio) and centrifuged for 20 minutes at room temperature at 800 g with the break off. After centrifugation, the interface containing PBMC was carefully collected and cells were washed with NaCl. The cell suspension was centrifuged at 300 g and the cell pellet washed with PBS before RNA extraction
<b>Nucleic acid extraction</b>	
Procedure/Kit	RNA was extracted with miRNA mini kit (Qiagen): 700 $\mu$ l of Qiazol (Qiagen) were added and the pellet frozen at -80°C.
DNase treatment	DNase treatment was performed with RNase-free DNase set (Qiagen) (30 min at room temperature)
Contamination	Absence of genomic DNA contamination in the RNA samples was tested by amplification of RNA (without reverse transcription), with the primers used for real time RT-PCR ( <i>HPRT1</i> , <i>IGF1</i> ).
Quantification	Nanodrop 1000 Spectrophotometer
Purity	Purity of total RNA was determined by the A260/280 and A260/A230 ratio. All the samples had A260/A280 ratio between 1.8 and 2 and a A260/A230 ratio superior to 2. Bioanalyser Agilent
Integrity	RNA integrity was assessed by bioanalyser (Agilent 2100) and expressed as RNA Integrity Number (RIN) and considered acceptable within the range of 7-10.
<b>Reverse transcription</b>	
Procedure	One $\mu$ g total RNA was reverse-transcribed to cDNA with random hexamers and oligos dT
Kit	PrimerScript™ RT reagent kit (perfect real-Time, TAKARA )
Amount of RNA	1 $\mu$ g of total RNA
Reaction volume	20 $\mu$ l
Temperature and time	37°C for 15 minutes 85°C for 5 seconds
Negative controls	Parallel reactions without reverse transcriptase enzyme were prepared as negative controls.
Storage condition of cDNA	-20°C

<b>Real Time RT-PCR target information</b>	
Housekeeping gene choice	<i>HPRT1</i> was chosen as housekeeping gene because it is constantly expressed in samples with or without GH treatment and is appropriate for normalisation of RT-qPCR.
Sequence accession number	<i>HPRT1</i> (NM_000194) Total <i>IGF1</i> and P1-driven <i>IGF1</i> (NM_001111283) P2-driven <i>IGF1</i> (NM_001111284.1)
Amplicon length	<i>HPRT1</i> : 115 pb; Total <i>IGF1</i> =139 pb P1- <i>IGF1</i> =124 pb, P2- <i>IGF1</i> =138 pb
<b>Real Time RT-qPCR taq man assay</b>	
Assay ID from IDT (INTEGRATED DNA TECHNOLOGIES)	<i>HPRT1</i> : (Hs.PT.5821454446) Total <i>IGF1</i> : (Hs.PT.56a.21022358.g ) P1-driven <i>IGF1</i> : (N001111285.1.pt.IGF1) P2-driven <i>IGF1</i> : (N001111284.1.pt.IGF1)
<b>Real Time RT-PCR protocol</b>	
Reaction volume	12 $\mu$ l
Reaction volume amount of cDNA/probes/Master Mix/ H <sub>2</sub> O	2 $\mu$ l/0.5 $\mu$ l/6 $\mu$ l/3.5 $\mu$ l
Complete thermo cycling parameters	50°C 20 secondes 95°C for 10 minutes [95°C 15s; 60°C 1minutes] x 45
Real time RT-PCR instrument	ABI 7500 fast (Applied Biosystems)
<b>Real Time RT-PCR validation</b>	
PCR efficiency	<i>HPRT1</i> (E=1.87, R <sup>2</sup> =0.99) All <i>IGF1</i> transcripts (E=1.95, R <sup>2</sup> =0.99) <i>IGF1</i> P1-driven transcripts (E=1.99, R <sup>2</sup> =0.99) <i>IGF1</i> P2-driven transcripts (E=2, R <sup>2</sup> =0.99)
Reproductibility	Each sample was measured in triplicate for each experiment and in 2 to 3 different experiments. Samples with low copy number were measured in 3 different experiments. Samples with low copy number were measured in 3 different experiments. Replicates were within a 0.3 SD range.
<b>Data analysis</b>	
Relative Quantification Method	Threshold cycles (Ct values), the cycle number at which the fluorescence signal of the sample exceeds background fluorescence, were determined for quantitative comparison of the amplification of the candidate genes.

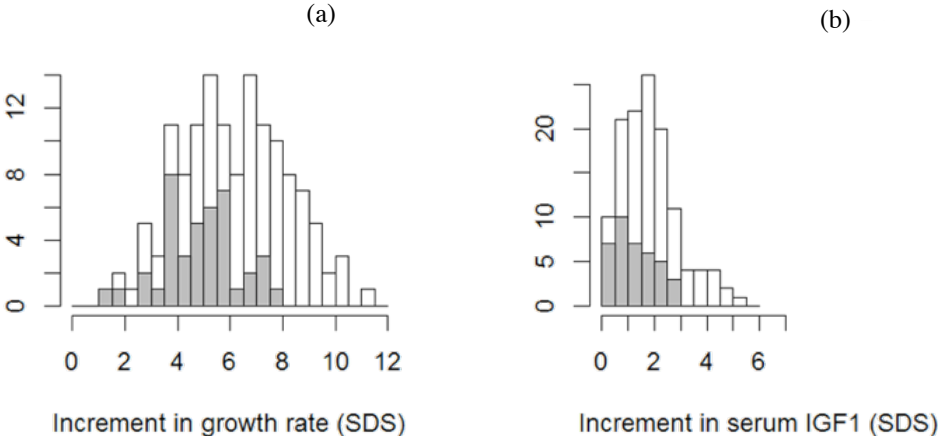


	<p>Relative expression levels were measured by the comparative Ct method (<math>\Delta\Delta Ct</math>), using a GH-unresponsive housekeeping gene <i>HPRT1</i> as endogenous control. Fold changes of the difference between reference and the target gene were calculated with the <math>2^{-\Delta\Delta Ct}</math> equation, according to Livak and Schmittgen<sup>1</sup>.</p> <p><math>\Delta\Delta Ct = \Delta Ct (\text{treated sample}) - \Delta Ct (\text{untreated sample})</math>  where <math>\Delta Ct(\text{sample}) = Ct(\text{target}) - Ct(\text{reference})</math></p> <p>The ratio of target gene in treated sample relative to untreated sample is calculate with the <math>2^{-\Delta\Delta Ct}</math> equation</p>
--	---

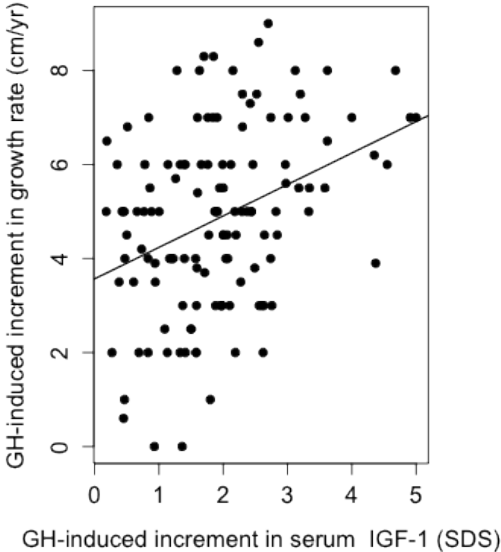
**Supporting Information References.**

1. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-(\Delta\Delta C(T))}$  Method. *Methods San Diego Calif* **25**, 402–408 (2001).

**Supplementary Figure 1**



Supplementary Figure 2



## 10.2 CONCLUSIONS DE L'ARTICLE 2

Après l'injection aiguë de GH, l'augmentation d'IGF1 en circulation était inversement corrélée à la méthylation d'ADN du promoteur P2, qui représentait environ 49% de la variabilité de cette augmentation. Donc nous observons un lien assez robuste entre le produit final du gène *IGF1*, la protéine circulante, qui est produite principalement par le foie et la méthylation d'ADN dans un autre tissu, les PBMCs. Et selon notre hypothèse de travail, ce lien peut être extrapolé sur le tissu hépatique, le tissu cible. Il est connu qu'une stimulation GH augmente la transcription d'*IGF1* dans les tissus cibles et que cette stimulation est réglée par des facteurs épigénétiques. La stimulation de l'expression est décrite dans l'introduction (Chia et al., 2010; Rotwein, 2012). L'accès aux tissus cibles dans des études menées chez des cohortes de patients est très difficile. La majorité des cellules de l'organisme répondent à l'hormone de croissance car elles possèdent toutes le récepteur de l'hormone de croissance (GHR). Nous avons donc cherché à montrer que l'activité transcriptionnelle du gène *IGF1* GH-dépendante, est modulée par la méthylation d'ADN dans les PBMCs.

Nous avons montré pour la première fois que les PBMCs sont des cellules capables de répondre à l'hormone de croissance *in vivo*. L'augmentation des deux classes de transcrits est variable, lors du test aigu. L'augmentation des transcrits de classe II est fortement corrélée avec la méthylation du P2 notamment avec le CG-137. La GH augmente la transcription du gène *IGF1* chez les individus les moins méthylés sur 4/8 CG du promoteur P2. Parmi les dinucléotides CG du promoteur P1, un seul CG (CG-611) était inversement corrélé avec ses transcrits (classe I).

La contribution de la méthylation du CG-137 est de 67% de la variabilité de l'augmentation des transcrits de la classe II dans les PBMC. Nous avons donc prouvé le rôle de la méthylation d'ADN dans l'expression du gène *IGF1* GH-dépendante *in vivo*. Le gène *IGF1* est faiblement exprimé dans les PBMC. En effet, il était difficile de valider nos résultats *in vitro*. La quantification des transcrits d'un gène faiblement exprimé dans une lignée primaire telle que les PBMC est difficile étant donnée la grande mortalité cellulaire *in vitro*.

L'augmentation des transcrits de la classe II des PBMC est proportionnelle au gain de la protéine IGF1 en circulation issue du foie (Figure 32).

### **Réponse thérapeutique GH**

Chez 130 enfants traités par l'hormone de croissance, l'accélération de la vitesse de croissance sous traitement était inversement corrélée avec la méthylation du promoteur P2. La méthylation du CG-137, par exemple, participe de 25% à la variabilité de la réponse à la GH. Nous avons confirmé ce résultat en construisant un modèle de régression multiple comprenant l'âge, le sexe, la dose de GH et la méthylation du CG-137.

L'augmentation d'IGF1 circulant sous GH, montre aussi de fortes corrélations avec différents dinucléotides CGs du promoteur P2. La méthylation du CG-137 contribue pour 27% dans la variance de l'augmentation des taux sériques d'IGF1 sous traitement.

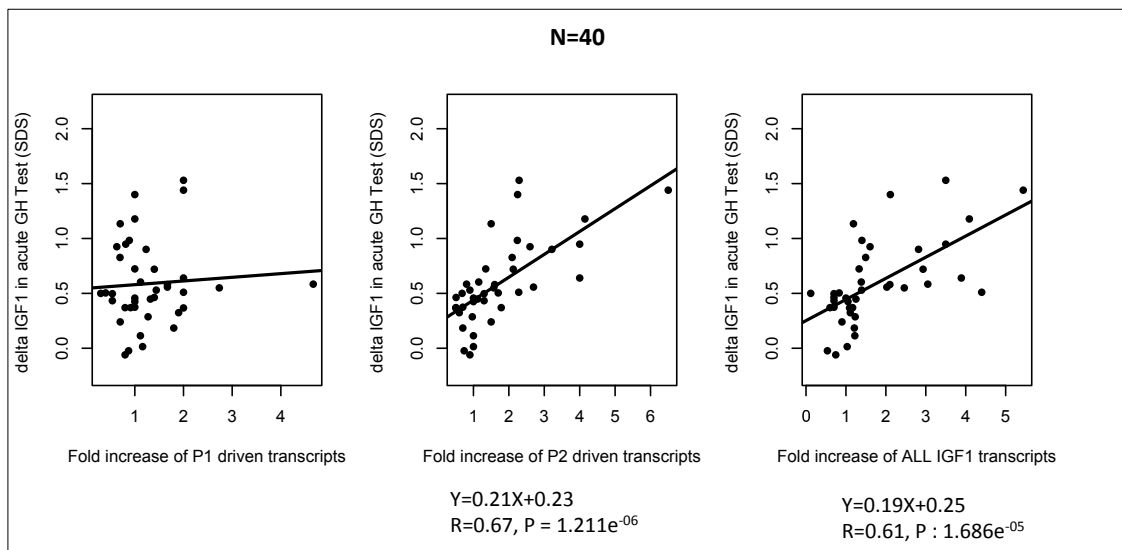
La cohérence et la force statistique des corrélations observées soutiennent que la méthylation du promoteur P2 du gène *IGF1* est un déterminant majeur de la réponse individuelle au traitement GH entre les enfants ayant une petite taille idiopathique. Nous proposons ainsi que l'association épigénétique observée montre une pertinence fonctionnelle, appuyée sur des effets de la transcription initiés au niveau du promoteur P2 après l'injection de GH.

Comme d'autres études avant (Kiström et al., 2014), nous avons trouvé une forte corrélation entre l'augmentation de la taille et de l'augmentation de la circulation d'IGF1. Comme l'IGF1 circulant

provient principalement du foie (Yakar et al., 2005), cette dernière corrélation suggère que les niveaux GH-induits de l'expression d'*IGF1* dans le foie et les chondrocytes sont proportionnelles chez un individu donné, et pourraient partager une partie de la régulation épigénétique rapportée par notre étude.

D'autres questions sont laissées sans réponse comme les mécanismes moléculaires qui se produisent au niveau du promoteur P2 sous l'effet GH et ainsi que la corrélation entre la méthylation des CGs et la régulation de la transcription du gène.

A notre connaissance, la présente étude offre la première preuve clinique d'un lien entre la méthylation de l'ADN et la réponse à un traitement chez l'homme, illustrant le rôle de la variation épigénétique comme un contributeur à la thérapeutique puissant et personnalisé.



**Figure 32: Relation entre la réponse à l'hormone de croissance dans un tissu cible (foie) et dans un tissu non cible (les cellules du sang (PBMC)).** Corrélation entre l'augmentation de l'IGF1 en circulation et les différentes classes de transcrits du gène *IGF1* dans les PBMC après le test de génération d'IGF1 (12 heures de stimulation par 100µg/kg de GH), N=40.

## **11 ARTICLE 3 : « GENETIC AND EPIGENETIC MODULATION OF GROWTH HORMONE SENSITIVITY STUDIED WITH THE IGF1 GENERATION TEST »**

### **11.1 OBJECTIFS DE L'ARTICLE 3**

La taille pendant l'enfance est étroitement dépendante de l'axe GH-IGF1. Nous avons utilisé un test aigu de stimulation de la production d'IGF1 par la GH comme une façon de tester :

- la variabilité individuelle de la réponse de l'IGF1 circulant à ce test selon les enfants ;
- l'influence du génotype du récepteur GH sur cette réponse, un complément d'information aux observations réalisées en 2004 ;
- l'influence de la méthylation du promoteur P2 du gène IGF1 sur cette réponse ;
- l'effet additif du génotype de GHR et de l'épigénotype d'IGF1.

Le test de génération d'IGF1 (IGFGT) est un test de stimulation endocrinien classique utilisé depuis longtemps pour l'évaluation de la sensibilité à l'hormone de croissance (GH) (Daughaday, 1969). Ce test sert ainsi i) à l'évaluation des diverses formes de petite taille, notamment celles qui sont associées à un défaut de la sensibilité à la GH, ii) à prévoir la réponse au traitement GH (Gretner JM, 1984 ; Plotnick LP, Rudman.D, 1981 ; Van Vliet.G, 1983). Cependant, l'utilité du test est controversée, sa capacité à discriminer différentes étiologies de retard statural n'est pas claire et a été mise en doute. Ceci vient en partie du fait que l'augmentation d'IGF1 circulant lors du test est très variable d'un patient à un autre au sein de la même catégorie clinique, ce qui rendait attractif de rechercher les causes individuelles de cette variabilité.

Il existe plusieurs protocoles cliniques utilisés par les endocrinologues selon i) la durée du test et ii) la dose de GH. Le test de génération a été simplifié pour notre étude. Nous avons utilisé un protocole unique pour 72 enfants pré-pubères qui ont reçu une première injection d'hormone de croissance. La durée de stimulation a été réduite à 12h. Une dose standardisée de 100 µg.kg a été utilisée pour tous les participants.

## Genetic and Epigenetic Modulation of Growth Hormone Sensitivity Studied With the IGF-1 Generation Test

Meriem Ouni, Anne-Laure Castell, Agnès Linglart, and Pierre Bougnères

Institut National de la Santé et de la Recherche Médicale Unité 986 (M.O., A.-L.C., A.L., P.B.) and Department of Pediatric Endocrinology and Diabetes (A.-L.C., A.L., P.B.), Paris Sud University, Bicêtre Hospital, 94275 Le Kremlin-Bicêtre, France

**Context:** Like all hormones, GH has variable physiological effects across people. Many of these effects initiated by the binding of GH to its receptor (GHR) in target tissues are mediated by the expression of the *IGF1* gene. Genetic as well as epigenetic variation is known to contribute to the individual diversity of GH-dependent phenotypes through two mechanisms. The first one is the genetic polymorphism of the *GHR* gene due to the common deletion of exon 3. The second, more recently reported, is the epigenetic variation in the methylation of a cluster of CGs dinucleotides located within the proximal part of the P2 promoter of the IGF-1 (*IGF1*) gene, notably CG-137.

**Objective:** The current study evaluates the relative contribution of these two factors controlling individual GH sensitivity by measuring the response of serum IGF-1 to a GH injection (IGF-1 generation test) in a sample of 72 children with idiopathic short stature.

**Results:** Although the d3 polymorphism of the *GHR* contributed 19% to the variance of the IGF-1 response, CG-137 methylation in the *IGF-1* promoter contributed 30%, the combined contribution of the two factors totaling 43%.

**Conclusion:** Our observation indicates that genetic and epigenetic variation at the *GHR* and *IGF-1* loci play a major role as independent modulators of individual GH sensitivity. (*J Clin Endocrinol Metab* 100: E919–E925, 2015)

Tens of thousands of children affected by various causes of short stature currently receive recombinant GH to improve their final height. There is, however, a highly variable individual GH response to this treatment, which does not allow pediatric endocrinologists to predict final height reliably, although clinical predictive models have been developed (1) as well as biological models based on GH/IGF-1 hormones measured before treatment (2). The multiple causes underlying the variability of the therapeutic response to GH include the etiology of short stature (3, 4), GH dose (5), child age (1), treatment regimens (6), and therapeutic compliance (7). Because IGF-1 is the mediator

of GH effects, the individual variation in growth responses is in large part due to the variable IGF-1 production under GH treatment, reflected by circulating IGF-1 concentration (8–10). On the other hand, the personal capacity to respond to GH has prompted few studies in search of biological mechanisms able to vary across individuals (11, 12). The increment in IGF-1 induced by GH administration in healthy children with idiopathic short stature is normally distributed (8, 10, 13, 14) and can thus be modeled as a continuous quantitative trait. Among the molecular determinants involved in the variability of IGF-1 responses to GH, it has become clear that genetic variation

ISSN Print 0021-972X ISSN Online 1945-7197

Printed in USA

This article has been published under the terms of the Creative Commons Attribution License (CC-BY; <https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Copyright for this article is retained by the author(s).

Received February 13, 2015. Accepted March 30, 2015.

First Published Online April 2, 2015

Abbreviations: Ct, cycle threshold; GHR, GH receptor; IGFBP3, IGF binding protein-3; rhGH, recombinant human growth hormone; SDS, SD score.

plays a role (11). Notably, the deletion of exon 3 within the GH receptor (*GHRd3*) gene has been recognized as a significant predictor of GH growth-promoting effects in children with idiopathic short stature (15–18) or other etiologies of short stature (11, 12, 19).

Estimates of the proportion of variance in circulating IGF-1 that is genetically determined vary between 38% and 80% (20, 21). The association of serum IGF-1 concentration with several genetic variants is debated (22–24), but variants at the *IGF-1* locus do not seem to influence circulating IGF-1 in Caucasian adults (22, 25) except, perhaps, the most common Z allele of the microsatellite located 1 kb upstream the *IGF-1* gene (26, 27). Overall, the genetic basis for serum IGF-1 variability remains largely unknown in adults and has not been studied during the period of physiological growth in normal children. In children with idiopathic short stature, considered to be a variant category of normal children, the *GHRd3* is associated with higher circulating IGF-1 in response to GH injection (28).

Unlike pharmacogenetics, pharmacoepigenetics is a nascent field of clinical medicine (29), and the epigenetics of growth and IGF-1 responses have only recently started to be investigated (30). A recent study showed that the methylation of a cluster of dinucleotides (CGs) located within the P2 promoter of the *IGF-1* (*IGF-1*) gene, notably CG-137, is inversely closely correlated with the response of growth and circulating IGF-1 to GH treatment (30).

In growing children, GH responsiveness is important to physiology and in some of them to therapeutics. The current study investigates the individual response of children to a GH injection with the objective of evaluating the respective role of genetic polymorphism of the *GHR* and the degree of methylation of the P2 promoter of the *IGF-1* gene. We selected children who have not entered puberty to avoid the confounding effect of the variable tempo of sexual maturation, which adds to the variability of basal (31–33) and GH-stimulated (34) circulating IGF-1. We used the long-studied generation test (31, 34–37) to study the GH responsiveness of circulating IGF-1 in 72 children. We studied the GH receptor (*GHR*) d3 genotype and the methylation of the *IGF-1* P2 promoter in these children to study how these factors contribute to the individual variability of the GH response to the test.

## Materials and Methods

### Participants

Seventy-two children with various degrees of short stature (–1.1 SD to –3.2 SD) belonging to the Epigrowth cohort (30) had venous blood sampling at 8:00 PM before dinner, 10 minutes before receiving an injection of 100  $\mu$ g/kg body weight of re-

combinant human growth hormone (rhGH) im in the left thigh, had a normocaloric standard dinner at 8:10 PM, fasted overnight, and had a second blood sampling at 8:00 AM. Their main characteristics are depicted in Table 1. All children were healthy and had normal clinical examination. Thirty children had criteria of idiopathic short stature, as defined in (11). In these children, GH deficiency was excluded with a stimulated GH peak greater than 15 ng/mL, TSH levels were normal. Subtle chondrodysplasia were excluded by forearm, pelvis, and spine radiographs. Pubertal stages were estimated using the Tanner definition.

Parents of all studied children gave their written informed consent for the study according to the French rules of bioethics in biomedical research checked by our institutional review board.

### Serum IGF-1 and IGF binding protein-3 (IGFBP3) concentrations

Serum IGF-1 concentration was measured at approximately 7:00–8:00 AM before breakfast in 136 children using an immune-radiometric assay after ethanol-acid extraction using Cis-bio reagents. Intra- and interseries coefficients of variation were 1.5% and 3.7% at 260 ng/mL and 3.9% at 760 ng/mL. The sensitivity was 4 ng/mL. IGF-1 SD score (SDS) was calculated using the norms of Alberti et al (33) in French children. Serum IGFBP3 was measured using an immune-radiometric assay after

**Table 1.** Main characteristics of the studied children, mean  $\pm$  SD

	Acute rhGH Injection test
n	72
Sex, M/F	45/27
Age, y	11 $\pm$ 2.5
Height (SDS)	–1.8 $\pm$ 0.7
Tanner stages, n	
1	41 (57%)
2	20 (28%)
3	10 (14%)
4	1 (1%)
Serum IGF-1 before test	
Nanograms per milliliter	236 $\pm$ 149
SDS	–1.16 $\pm$ 0.8
Serum IGF-1 12 h after rhGH	
Nanograms per milliliter	326 $\pm$ 150
SDS	–0.55 $\pm$ 0.84
<i>GHR</i> genotype	
fl/fl	33 (46%)
fl/d3	32 (44%)
d3/d3	7 (10%)
<i>IGF-1</i> P2 promoter methylation, %	
CG-232	62 $\pm$ 6
CG-224	74 $\pm$ 7
CG-218	72 $\pm$ 6
CG-207	45 $\pm$ 7
CG-137	47 $\pm$ 4
CG-108	60 $\pm$ 6
Average <sup>a</sup>	58 $\pm$ 4

Abbreviations: F, female; M, male.

<sup>a</sup> Mean value for the six studied CGs.



ethanol-acid extraction with Diagnostic Systems Laboratory reagents (Beckman-Coulter). Intra- and interseries coefficient of variation was 10.4 and 14% and the sensitivity limit was 6.2 ng/mL.

### GHR genotype

Analysis of the *GHR* exon 3 polymorphism was carried out in patients by quantitative PCR performed on ABI 7500 fast (Applied Biosystems). Genotyping protocol was adapted from Bernabeu et al (38). Primer/probe sets were targeted to exon 3 and exon 10 of *GHR*, which was used as an internal positive control. We used oligonucleotide primers and probes previously described in (38).

For each analysis, 50 ng genomic DNA was quantitative PCR amplified in 96-well plates in a volume of 12  $\mu$ L using Gotaq probe quantitative PCR master mix (Promega), 1.25  $\mu$ L *GHR* exon 3 primer/probe set (3 pm/ $\mu$ L each), 1.25  $\mu$ L *GHR* exon 10 primer/probe set (9 pm/ $\mu$ L each), and 2.75  $\mu$ L sterile water. Cycle conditions were 50°C for 2 minutes and 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Differences in cycle threshold (Ct) between the exon 3 and exon 10 amplicons were used to determine the exon 3 copy number for each sample. A  $\delta$ Ct value of 1 indicates two exon 3 copies (genotype fl/fl), a  $\delta$ Ct value of 2 indicates one exon 3 copy (genotype fl/d3), and no signal for exon 3 in the presence of a normal exon 10 signal indicates an absent exon 3 (genotype d3/d3).

### DNA methylation

For methylation measurements, 6 mL peripheral blood samples were obtained, from which peripheral blood mononuclear cells were purified immediately.

For promoter P2, we studied the six CGs located upstream from the major transcription start site within the proximal part of the P2 promoter. CGs are denominated according to position vs each promoter transcription start site. Nucleic acids were extracted from peripheral blood mononuclear cells using Gentra Puregene blood kit (QIAGEN). A bisulfite-PCR-pyrosequencing technique (39) was used to measure the methylation of the CGs. We improved the resolution of this method from a handful of bases to up to 100 nucleotides, with the ability to quantify methylation in the same sample of blood with a coefficient of variation (SD/mean) as little as 1%–5%. Briefly, 400 ng of genomic DNA was treated with an EZ DNA Methylation-Gold kit, according to manufacturer's protocol (Zymo Research Corp). The bisulfite-treated genomic DNA was PCR amplified using unbiased *IGF-1* primers (30) and performed quantitative pyrosequencing using a PyroMark Q96 ID Pyrosequencing instrument (QIAGEN). Pyrosequencing assays were designed using MethPrimer (<http://www.urogene.org/methprimer/index1.html>). Biotin-labeled single-stranded amplicons were isolated according to protocol using the QIAGEN Pyromark Q96 Work Station and underwent pyrosequencing with 0.5  $\mu$ M primer. The percentage methylation for each of the CGs within the target sequence was calculated using PyroQ CpG Software (QIAGEN).

### Calculations and statistics

The IGF-1 response to GH administration was expressed as increment in serum IGF-1 concentration, ie, the difference between serum IGF-1 12 hours after injecting GH and serum IGF-1 concentration before GH injection. Pearson correlations were

calculated as adjusted R square that measures the proportion of the variation in the dependent variable accounted for by the explanatory variables. The fraction of explained variance is calculated under the linear regression model, using the usual definition:  $r^2 \times 100$ . We fitted a multivariate linear model to the data to estimate the proper effect of *GHR* genotype and CG-137 methylation on response to GH, adjusted for the effect of the other covariates contributing to the IGF-1 response, such as age and sex. We carried out tests of independence of each covariate one at a time, keeping the others in the model. Statistics and estimations of effect given in the tables are thus adjusted for the others whenever appropriate and are not subject to marginal association. We checked the normality of the residuals, and the residuals vs the fitted values did not show any trend, indicating that there was no noticeable deviation from the assumption of the linear model. All statistics and linear model were computed using R 2.10.1. Results are expressed as mean  $\pm$  SD.

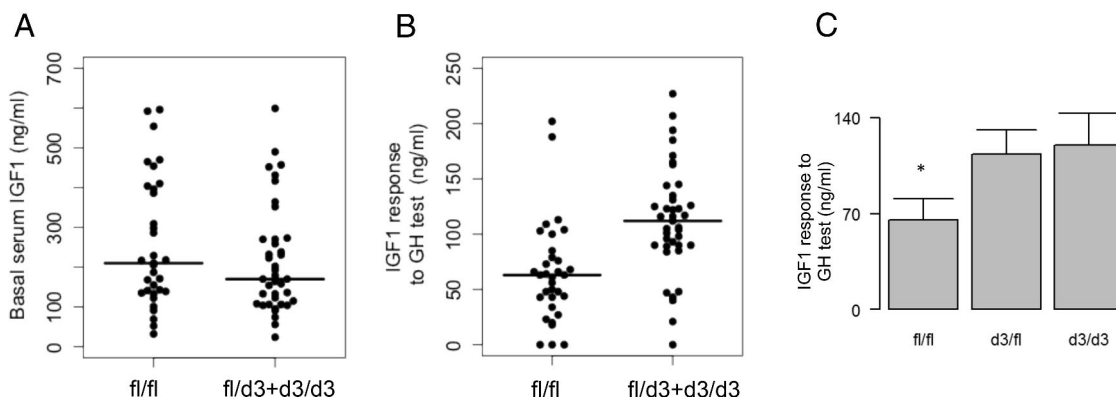
### Results

The response of serum IGF-1 concentration to the GH test averaged  $90 \pm 54$  ng/mL, an increase of  $38 \pm 36\%$  from the basal IGF-1 value. There was no correlation between basal IGF-1 and the increase of IGF-1 during the test. Within the studied group of children, the IGF-1 response to GH showed a nonsignificant trend of positive association with age ( $P = .11$ ) and Tanner pubertal stages ( $P = .45$ ).

Basal IGF-1 concentration was comparable in children of various *GHR* genotypes (Figure 1A), but a significant difference appeared after the GH injection (Figure 1B). Children carrying the dominant *GHR*-allele d3 (15) had a 56% higher IGF-1 response to GH than the fl/fl homozygotes. Because d3/3 children and fl/d3 showed comparable IGF-1 responses to GH, they were merged in a single group for analysis.

The serum IGF-1 response to the GH test showed a negative correlation with the methylation of three of six CGs of the IGF-1 P2 promoter (Supplemental Table 1), notably with CG-137 methylation ( $R = -0.54$ ,  $P = 4.10^{-7}$ ) as shown in Figure 2A. Children in the highest tertile for CG-137 methylation showed a 142% increase in IGF-1 response compared with those in the lowest tertile (Figure 2B). CG methylation was comparable in the *GHR* genotype groups.

Once combined in a multivariate analysis with age, pubertal stages, and sex, both *GHR* d3 genotype and the methylation of CG-137 showed a significant association with the serum IGF-1 response to the GH test (Table 2). The combined effect of *GHR* genotype and IGF-1 epigenotype is illustrated by the three-dimensional histograms of Figure 3. Overall, the *GHR* d3 genotype contributed 19% to the variance of the IGF-1 response and the *IGF-1* P2 methylation contributed 30%, resulting in a total con-



**Figure 1.** Relationship between GHR-d3 genotype and serum IGF-1 concentration in the studied children. A, Basal IGF-1 shows no difference across the genotypic groups. B, The response of serum IGF-1 to GH test is greater in children carrying a d3 allele ( $P = 2.10^{-5}$ ). C, Mean values of IGF-1 response in the three GHR genotypic groups show that d3/fl and d3/d3 classes have a comparable response to GH.

tribution of 43% to the individual GH sensitivity evaluated by the IGF-1 generation test. Strikingly, the children who had both a GHR fl/fl genotype and a CG-137 methylation in the highest tertile ( $n = 18$ ) had an IGF-1 response to the GH test 3.4 times smaller than children carrying one or two GHR d3 alleles combined with CG-137 methylation in the lowest tertile ( $n = 11$ ) (Figure 3).

We found no correlation of serum IGFBP3 with the GHR genotype or with the IGF-1 P2 promoter methylation, in the basal state as well as in response to the GH test.

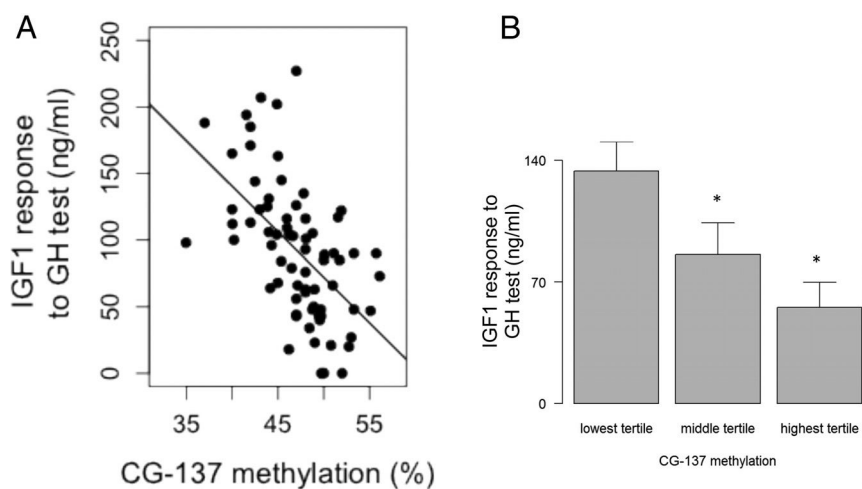
## Discussion

Our study supports that both the GHR genotype and the CG methylation of the P2 promoter of the *IGF-1* gene are major determinants of the individual IGF-1 response to GH in childhood.

IGF-1 generation tests were developed more than 20 years ago (31, 34–37) and are currently used in evaluating

GH sensitivity in children with unexplained short stature, notably when characterized by low serum IGF-I. Major limitations have included variability in protocols for administration of GH, timing of samples, differences in IGF assay methodologies, and lack of adequate normative data (40). The latter is particularly problematic, given the well documented age-related variability in IGF-I concentrations (12, 32) as well as gender-related differences in responsiveness to exogenous GH (41, 42). The results of the generation test in children with idiopathic short stature are of particular interest, as previously reported (36), to have serum IGF-I concentrations in the lower portion of the normal range or below the lower limits of normal. Interestingly, these subjects also failed, in general, to raise their serum IGF-1 concentrations in response to GH; many did not even attain levels within the baseline normal range. According to another report (34), this was previously considered to result from mild mutations of the *GHR* gene or from subtle postreceptor mechanisms (36, 43, 44). The current results establish the GHR d3 genotype and the *IGF-1* P2 epigenotype as major sources of individual variation of the IGF-1 generation test. This can be of importance for understanding the phenotypic diversity in GH sensitivity at the individual level.

The vast majority of the increase in serum IGF-1 concentration occurring in response to GH is considered to result from GH effects on *IGF-1* expression in child liver, mediated by GH binding to its hepatic receptor. A weakness of our study is that we could not study the association of liver P2 methylation with GH responsiveness. Such lack of availability and analysis of epigenomes in spe-



**Figure 2.** Relationship between CG-137 methylation and IGF-1 response to the GH test. A, Correlation between serum IGF-1 and CG-137 methylation ( $R = -0.54$ ,  $P = 4.10^{-7}$ ,  $y = -6.8x + 413$ ). B, IGF-1 response across the three tertiles of CG-137 methylation ( $P = 10^{-6}$ ).

**Table 2.** General linear model for regression of age, sex, Tanner stage, *GHR* d3 polymorphism, and methylation of CG-137 on IGF-1 response to the GH test

	Estimates	SE	t Value	Pr (> t)
Intercept	275	58	4.7	10 <sup>-5</sup>
Age, y	6.3	2.74	2.30	0.025
Sex	15.03	10.3	1.46	0.15
Tanner stage (1–4)	–12.84	9.3	–1.40	0.17
<i>GHR</i> d3 polymorphism	35.7	9.20	3.87	2.10 <sup>-4</sup>
CG-137 methylation, %	–5.60	1.13	–4.95	5.6.10 <sup>-6</sup>

Pr (> t), probability of a value > t.

cific cells of physiological tissues is a common but major limitation of epigenetic epidemiology (45, 46), which most often has to rely on blood cells (47, 48).

The expression of the *IGF-1* gene in liver is controlled by GH at a transcriptional level. In mammals including humans, the *IGF-1* gene is composed of six exons and five introns that span greater than 80 kb of chromosomal DNA (49, 50). Tandem promoters direct *IGF-1* gene transcription through unique leader exons. Promoter 1, which uses heterogeneous transcription initiation sites, is active in multiple animal tissues (51), whereas the smaller and simpler promoter 2 is primarily but not exclusively active in the liver of cattle (52), unlike in rodents in which promoter 2 activity seems exclusively hepatic (14). GH exerts its effects through the Janus kinase/signal transducer and activator of transcription pathway with translocation of activated signal transducer and activator of transcription-5b transcription factor to the nucleus in which it regulates *IGF-1* transcription (53). In the liver of hypophysectomized rats, GH induces dramatic changes in chromatin at the *IGF-1* locus and activates *IGF-1* transcription by dis-

tinct promoter-specific epigenetic mechanisms (54, 55). Yet not only CG location and composition are different in human and rat *IGF-1* promoters, but also the pattern of methylation and its transcriptional effects on rat promoters are still unknown.

Although the d3 polymorphism at the *GHR* locus and the variation of methylation at the *IGF-1* P2 promoter account for a total 43% of the individual variability of GH effects on serum IGF-1, the remaining 57% of variance are yet to be explained by other factors. Age, puberty, or sex could be important but do not play a significant role in the current group of children recruited within a relatively narrow age range. Because serum IGF-1 concentration mostly reflects IGF-1 production by the liver in response to GH, it is possible that some of the variance comes from the factors controlling GH pharmacokinetics or from the molecular pathway involved in the hepatic signaling of GH upstream IGF-1 production. This pathway encompasses the *GHR*, *JAK2*, and *STAT5b* gene products whose expression can vary, depending on individual genotypes and epigenotypes. We are not aware that genome-wide association studies or methylome-wide association studies for serum IGF-1 in children (basal or in response to GH) have yet been carried out. The challenge of understanding the remaining part of individual GH sensitivity thus remains entire.

Our observation indicates that genetic and epigenetic variation at the *GHR* and *IGF-1* loci play a major role as independent modulators of individual GH sensitivity.

## Acknowledgments

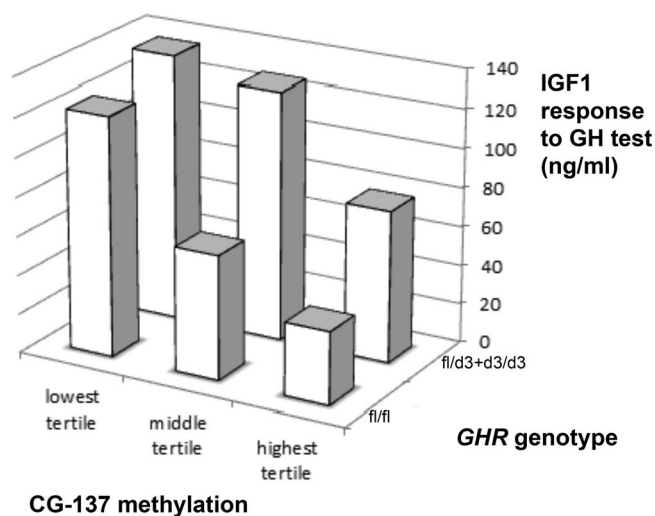
We thank Pfizer France for supporting the GH Pharmacogenomics (“GH-EPiGx”) program.

Address all correspondence and requests for reprints to: Professor Pierre Bougnères, INSERM Unité 986 and Department of Pediatric Endocrinology, Bicêtre Hospital, 80, Rue du Général Leclerc, 94275 Le Kremlin Bicêtre, France. E-mail: [pierre.bougneres@inserm.fr](mailto:pierre.bougneres@inserm.fr).

Disclosure Summary: The authors have nothing to disclose.

## References

1. Ranke MB, Lindberg A. Predicting growth in response to growth hormone treatment. *Growth Horm IGF Res.* 2009;19(1):1–11.
2. Dahlgren J, Krström B, Niklasson A, Nierop AFM, Rosberg S, Albertsson-Wikland K. Models predicting the growth response to growth hormone treatment in short children independent of GH status, birth size and gestational age. *BMC Med Inform Decis Mak.* 2007;7:40.
3. Lee PA, Säwendahl L, Oliver I, et al. Comparison of response to 2-years' growth hormone treatment in children with isolated growth hormone deficiency, born small for gestational age, idiopathic short



**Figure 3.** 3-D relationship between IGF-1 response to GH test and both *GHR*-d3 genotype and CG-137 methylation showing higher responses in children carrying a d3 allele and having low levels of CG-137 methylation ( $P = 10^{-9}$ ).



- stature, or multiple pituitary hormone deficiency: combined results from two large observational studies. *Int J Pediatr Endocrinol*. 2012;2012(1):22.
4. Collett-Solberg PF. Update in growth hormone therapy of children. *J Clin Endocrinol Metab*. 2011;96(3):573–579.
  5. Albertsson-Wikland K, Aronson AS, Gustafsson J, et al. Dose-dependent effect of growth hormone on final height in children with short stature without growth hormone deficiency. *J Clin Endocrinol Metab*. 2008;93(11):4342–4350.
  6. Sotos JF, Tokar NJ. Growth hormone significantly increases the adult height of children with idiopathic short stature: comparison of subgroups and benefit. *Int J Pediatr Endocrinol*. 2014;2014(1):15.
  7. Rosenfeld RG, Bakker B. Compliance and persistence in pediatric and adult patients receiving growth hormone therapy. *Endocr Pract*. 2008;14(2):143–154.
  8. Cohen P, Germak J, Rogol AD, et al. Variable degree of growth hormone (GH) and insulin-like growth factor (IGF) sensitivity in children with idiopathic short stature compared with GH-deficient patients: evidence from an IGF-based dosing study of short children. *J Clin Endocrinol Metab*. 2010;95(5):2089–2098.
  9. Cohen P, Rogol AD, Weng W, et al. Efficacy of IGF-based growth hormone (GH) dosing in nonGH-deficient (nonGHD) short stature children with low IGF-I is not related to basal IGF-I levels. *Clin Endocrinol (Oxf)*. 2013;78(3):405–414.
  10. Cohen P, Rogol AD, Howard CP, et al. Insulin growth factor-based dosing of growth hormone therapy in children: a randomized, controlled study. *J Clin Endocrinol Metab*. 2007;92(7):2480–2486.
  11. Rosenfeld RG. The Pharmacogenomics of Human Growth. *J Clin Endocrinol Metab*. 2006;91(3):795–796.
  12. Ranke MB. Treatment of children and adolescents with idiopathic short stature. *Nat Rev Endocrinol*. 2013;9(6):325–334.
  13. Park P, Cohen P. Insulin-like growth factor I (IGF-I) measurements in growth hormone (GH) therapy of idiopathic short stature (ISS). *Growth Horm IGF Res*. 2005;15(suppl A):S13–S20.
  14. Krström B, Lundberg E, Jonsson B, Albertsson-Wikland K, Study Group. IGF-1 and growth response to adult height in a randomized GH treatment trial in short non-GH-deficient children. *J Clin Endocrinol Metab*. 2014;99(8):2917–2924.
  15. Dos Santos C, Essioux L, Teinturier C, Tauber M, Goffin V, Bougnères P. A common polymorphism of the growth hormone receptor is associated with increased responsiveness to growth hormone. *Nat Genet*. 2004;36(7):720–724.
  16. Bougnères P. The exon-3 deletion of the growth hormone receptor (GHR) gene still has a limited impact in clinical endocrinology. *J Clin Endocrinol Metab*. 2010;95(1):56–59.
  17. Renehan AG, Solomon M, Zwahlen M, et al. Growth hormone receptor polymorphism and growth hormone therapy response in children: a Bayesian meta-analysis. *Am J Epidemiol*. 2012;175(9):867–875.
  18. Wassenaar MJE, Dekkers OM, Pereira AM, et al. Impact of the exon 3-deleted growth hormone (GH) receptor polymorphism on baseline height and the growth response to recombinant human GH therapy in GH-deficient (GHD) and non-GHD children with short stature: a systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2009;94(10):3721–3730.
  19. Rosenfeld RG. Pharmacogenomics and pharmacoproteomics in the evaluation and management of short stature. *Eur J Endocrinol*. 2007;157(suppl 1):S27–S31.
  20. Kao PC, Matheny AP, Lang CA. Insulin-like growth factor-I comparisons in healthy twin children. *J Clin Endocrinol Metab*. 1994;78(2):310–321.
  21. Harrela M, Koistinen H, Kaprio J, et al. Genetic and environmental components of interindividual variation in circulating levels of IGF-I, IGF-II, IGFBP-1, and IGFBP-3. *J Clin Invest*. 1996;98(11):2612–2651.
  22. Palles C, Johnson N, Coupland B, et al. Identification of genetic variants that influence circulating IGF-1 levels: a targeted search strategy. *Hum Mol Genet*. 2008;17(10):1457–1415.
  23. Al-Zahrani A, Sandhu MS, Luben RN, et al. IGF-1 and IGFBP3 tagging polymorphisms are associated with circulating levels of IGF-1, IGFBP3 and risk of breast cancer. *Hum Mol Genet*. 2006;15(1):1–101.
  24. Wong H-L, DeLellis K, Probst-Hensch N, et al. A new single nucleotide polymorphism in the insulin-like growth factor I regulatory region associates with colorectal cancer risk in Singapore Chinese. *Cancer Epidemiol Biomarkers Prev*. 2005;14(1):144–151.
  25. Ester WA, Hokken-Koelega ACS. Polymorphisms in the IGF-1 and IGF-1R genes and children born small for gestational age: results of large population studies. *Best Pract Res Clin Endocrinol Metab*. 2008;22(3):415–431.
  26. Vaessen N, Janssen JA, Heutink P, et al. Association between genetic variation in the gene for insulin-like growth factor-I and low birth-weight. *Lancet*. 2002;359(9311):1036–1037.
  27. Frayling TM, Hattersley AT, McCarthy A, et al. A putative functional polymorphism in the IGF-I gene association studies with type 2 diabetes, adult height, glucose tolerance, and fetal growth in UK populations. *Diabetes*. 2002;51(7):2313–2316.
  28. Toyoshima MTK, Castroneves LA, Costalonga EF, Mendonca BB, Arnhold IJP, Jorge AAL. Exon 3-deleted genotype of growth hormone receptor (GHRd3) positively influences IGF-1 increase at generation test in children with idiopathic short stature. *Clin Endocrinol (Oxf)*. 2007;67(4):500–504.
  29. Ivanov M, Kacevska M, Ingelman-Sundberg M. Epigenomics and interindividual differences in drug response. *Clin Pharmacol Ther*. 2012;92(6):727–736.
  30. Ouni M, Belot MP, Castell AL, Fradin D, Bougnères P. The IGF-1 P2 promoter is a major epigenetic locus for GH responsiveness. *Pharmacogenomics*. Published online ahead of print March 13, 2015. doi:10.1038/tpj.2015.26.
  31. Coutant R, Dörr H-G, Gleeson H, Argente J. Diagnosis of endocrine disease: limitations of the IGF-1 generation test in children with short stature. *Eur J Endocrinol*. 2012;166(3):351–373.
  32. Cole T j., Ahmed M I., Preece M a., Hindmarsh P, Dunger D b. The relationship between Insulin-like Growth Factor 1, sex steroids and timing of the pubertal growth spurt. *Clin Endocrinol (Oxf)*. Published online ahead of print November 23, 2014. doi: 10.1111/cen.12682.
  33. Alberti C, Chevenne D, Mercat I, et al. Serum concentrations of insulin-like growth factor (IGF)-1 and IGF binding protein-3 (IGFBP-3), IGF-1/IGFBP-3 ratio, and markers of bone turnover: reference values for French children and adolescents and z-score comparability with other references. *Clin Chem*. 2011;57(10):1424–1435.
  34. Buckway CK, Guevara-Aguirre J, Pratt KL, Burren CP, Rosenfeld RG. The IGF-I generation test revisited: a marker of GH sensitivity. *J Clin Endocrinol Metab*. 2001;86(11):5176–5183.
  35. Copeland KC, Underwood LE, Van Wyk JJ. Induction of immunoreactive somatomedin C human serum by growth hormone: dose-response relationships and effect on chromatographic profiles. *J Clin Endocrinol Metab*. 1980;50(4):690–697.
  36. Attie KM, Carlsson LMS, Rundle AC, Sherman BM. Evidence for partial growth hormone insensitivity among patients with idiopathic short stature. *J Pediatr*. 1995;127(2):244–250.
  37. Rudman D, Kutner MH, Blackston RD, Cushman RA, Bain RP, Patterson JH. Children with normal-variant short stature: treatment with human growth hormone for six months. *N Engl J Med*. 1981;305(3):123–131.
  38. Bernabeu I, Alvarez-Escolá C, Quinteiro C, et al. The exon 3-deleted growth hormone receptor is associated with better response to pegvisomant therapy in acromegaly. *J Clin Endocrinol Metab*. 2010;95(1):222–291.
  39. Tost J, Gut IG. DNA methylation analysis by pyrosequencing. *Nat Protoc*. 2007;2(9):2265–2275.
  40. Schwarze CP, Wollmann HA, Binder G, Ranke MB. Short-term increments of insulin-like growth factor I (IGF-I) and IGF-binding protein-3 predict the growth response to growth hormone (GH)

- therapy in GH-sensitive children. *Acta Paediatr Suppl (Oslo Nor)*. 1999;88(428):200–208.
41. Cook DM, Ludlam WH, Cook MB. Route of estrogen administration helps to determine growth hormone (GH) replacement dose in GH-deficient adults. *J Clin Endocrinol Metab*. 1999;84(11):3956–3601.
  42. Bouhours-Nouet N, Gatelais F, Boux de Casson F, Rouleau S, Coutant R. The insulin-like growth factor-I response to growth hormone is increased in prepubertal children with obesity and tall stature. *J Clin Endocrinol Metab*. 2007;92(2):629–635.
  43. Clayton PE, Freeth JS, Whatmore AJ, Ayling RM, Norman MR, Silva CM. Signal transduction defects in growth hormone insensitivity. *Acta Paediatr Suppl (Oslo Nor)*. 1999;88(428):174–178; discussion 179.
  44. López-Bermejo A, Buckway CK, Rosenfeld RG. Genetic defects of the growth hormone-insulin-like growth factor axis. *Trends Endocrinol Metab TEM*. 2000;11(2):39–49.
  45. Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nat Rev Genet*. 2013;14(8):585–594.
  46. Rakan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet*. 2011;12(8):529–541.
  47. Feinberg AP, Irizarry RA, Fradin D, et al. Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci Transl Med*. 2010;2(49):49ra67.
  48. Talens RP, Boomsma DI, Tobi EW, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. *FASEB J*. 2010;24(9):3135–3144.
  49. Rotwein P, Pollock KM, Didier DK, Krivi GG. Organization and sequence of the human insulin-like growth factor I gene. Alternative RNA processing produces two insulin-like growth factor I precursor peptides. *J Biol Chem*. 1986;261(11):4828–4832.
  50. Rotwein P. Structure, evolution, expression and regulation of insulin-like growth factors I and II. *Growth Factors*. 1991;5(1):3–181.
  51. Hall LJ, Kajimoto Y, Bichell D, et al. Functional analysis of the rat insulin-like growth factor I gene and identification of an IGF-I gene promoter. *DNA Cell Biol*. 1992;11(4):301–313.
  52. Wang Y, Price SE, Jiang H. Cloning and characterization of the bovine class 1 and class 2 insulin-like growth factor-I mRNAs. *Domest Anim Endocrinol*. 2003;25(4):315–328.
  53. Heim MH. The Jak-Stat pathway: cytokine signalling from the receptor to the nucleus. *J Recept Signal Transduct*. 1999;19(1–4):75–120.
  54. Chia DJ, Young JJ, Mertens AR, Rotwein P. Distinct alterations in chromatin organization of the two IGF-I promoters precede growth hormone-induced activation of IGF-I gene transcription. *Mol Endocrinol*. 2010;24(4):779–789.
  55. Chia DJ, Rotwein P. Defining the epigenetic actions of growth hormone: acute chromatin changes accompany GH-activated gene transcription. *Mol Endocrinol*. 2010;24(10):2038–2049.

## 11.2 CONCLUSIONS DE L'ARTICLE 3

La contribution du polymorphisme GHR à la réponse aigüe à la GH est claire.

Nos résultats concernant la contribution du polymorphisme GHR dans la réponse au test sont plus consistante que celle reporté dans l'étude de Toyoshima et ses collaborateurs en 2007. Une réponse meilleure au test est montrée chez les individus porteurs de l'allèle d3 et aussi chez les individus hypométhylés sur leur P2 d'*IGF1*. La méthylation était comparable dans les 3 groupes génotypiques GHR. Le polymorphisme GHR d3 contribue par 19% et la méthylation de l'ADN du P2 (notamment CG-137) par 30% de la variance de la réponse de l'*IGF1*.

D'autre part, les individus hypométhylés sur leur P2 d'*IGF1* ont une meilleure réponse. La méthylation du CG-137 du P2, notamment, contribue à elle seule à 30% de la variance de la réponse de l'*IGF1*.

Combinés, ces deux facteurs montrent l'additivité de leurs effets et rendent compte au total de 43% de la sensibilité individuelle à la GH évaluée par le test.

## **12 ARTICLE 4 : « HIGHER METHYLATION OF THE IGF1 P2 PROMOTER IS ASSOCIATED WITH IDIOPATHIC SHORT STATURE »**

*En cours de revue à Clinical Endocrinology*

### **12.1 OBJECTIFS DE L'ARTICLE 4**

La petite taille idiopathique, celle qui n'a pas de cause pathologique, est une cause fréquente de consultation en endocrinologie pédiatrique. Nous avons mesuré la méthylation du promoteur P2 dans une cohorte de tels enfants et l'avons comparée à celle d'enfants de taille normale. Nous avons ainsi cherché à montrer que les régions promotrices du gène *IGF1* sont des régions différenciellement méthylées associées à la petite taille. Ce travail est l'application clinique des observations effectuées dans l'article 1.

**Higher methylation of the IGF1 P2 promoter is associated with idiopathic short stature.**

Journal:	<i>Clinical Endocrinology</i>
Manuscript ID:	CEN-2015-000342
Manuscript Type:	1 Original Article - UK, Europe
Date Submitted by the Author:	10-Jun-2015
Complete List of Authors:	OUNI, Meriem; INSERM1069, Pediatric endocrinology Castell, Anne-Laure; INSERM, Department of Pediatric Endocrinology and Diabetes, I3E Pole Rothenbuhler, Anya; Hospital Kremlin Bicetre, Department of Pediatric Endocrinology and Diabetes, I3E Pole Linglart, agnes; INSERM, Department of Pediatric Endocrinology and Diabetes, I3E Pole Bougneres, Pierre; INSERM U986, Bicetre Hospital, Pole I3E, Paris Sud University; Bicetre Hospital, Department of Pediatric Endocrinology, Pole I3E, Paris Sud University
Key Words:	Growth disorders < Conditions: < Paediatric endocrinology, Growth hormone < Hormones/related: < Pituitary, Hormone replacement therapy < Investigations & Rx: < Pituitary



1  
2  
3  
4  
5  
6  
7 **Higher methylation of the *IGF1* P2 promoter is associated with idiopathic short stature.**  
8  
9

10  
11 Ouni Meriem<sup>1</sup>, Castell Anne-Laure<sup>2</sup>, Anya Rothenbuhler<sup>2</sup>, Linglart Agnès<sup>2</sup>, Bougnères  
12 Pierre<sup>1,2</sup>.  
13  
14  
15

16  
17  
18  
19  
20 <sup>1</sup>Institut National de la Santé et de la Recherche Médicale U986, Paris Sud University, Bicêtre  
21 Hospital, Le Kremlin-Bicêtre, France.  
22

23 <sup>2</sup>Department of Pediatric Endocrinology and Diabetes, I3E Pole, Paris Sud University, Bicêtre  
24 Hospital, Le Kremlin-Bicêtre, France.  
25  
26  
27  
28  
29

30  
31 Corresponding author : Pr Pierre Bougnères  
32 Tel : 01 45 21 78 54  
33 Fax : 01 45 21 78 50  
34 E-mail: pierre.bougneres@inserm.fr  
35  
36

37 Short title: IGF1 promoter methylation and human growth  
38

39  
40 Word Count: 2057 (including Abstract and Methods)  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Abstract

**Background:** Idiopathic short stature (ISS) has a strong familial component but genetics explains only part of it. Indeed, environmental factors act on human growth either directly or through epigenetic factors that remain to be determined. Given the importance of the GH/IGF1 axis for child growth, we suspected that such epigenetic factors could involve the CG methylation at the *IGF1* gene P2 promoter, which was recently shown to be a transcriptional regulator for *IGF1* gene and a major contributor to GH sensitivity.

**Objective.** Explore if the methylation of the two *IGF1* low-CG rich promoters (P1 and P2) is associated with ISS.

**Subjects and methods:** 94 children with ISS were compared with 119 age-matched children of normal height for the methylation of CGs located within the *IGF1* promoters measured with bisulfite-PCR-pyrosequencing.

**Results.** The methylation of 5 CGs of the P2 promoter was higher in ISS children, notably CG-137 ( $49 \pm 4\%$  in ISS versus  $46 \pm 4\%$  in control children,  $P=9 \times 10^{-5}$ ). This was also true for CG-611 of the P1 promoter ( $93 \pm 3\%$  vs  $91 \pm 3\%$   $P=10^{-4}$ ). The CG methylation of the *IGF1* promoters thus takes place among the multifactorial factors that are associated with ISS.

**Keywords:** Idiopathic short stature, epigenetics, IGF1 promoters

## INTRODUCTION

Although “idiopathic” short stature (ISS) is a variant of normal growth <sup>1</sup>, it has become a common source of medical investigation <sup>2</sup> and can be an indication for treatment with growth hormone <sup>3</sup>. The molecular causes of idiopathic short stature are multiple <sup>4</sup>, since height variation across individual is influenced both by genotypes <sup>5</sup> and environmental factors <sup>5</sup>. The genetic heritability of height is explained by many variants of individually small effect <sup>6</sup>. The recent meta-analysis using the summary statistics from 79 studies totalling 253,288 European individuals showed that the most strongly associated 9,500 SNPs account for 29% of height variance <sup>7</sup>. No single variant identified in these GWAS explained more than 0.3% of height variance. Copy number variations are other components of height genetics <sup>8</sup>.

Classically, quantitative geneticists have envisioned DNA sequence variants as the only source of heritable phenotypes. However, there is accumulating evidence for widespread epigenetic variation in natural populations and epigenetics is now recognized as a considerable source of inherited and acquired phenotypic diversity across humans.

To search for new causes of ISS, we selected the *IGF1* gene as a prominent physiological candidate. Most epigenetic effects on phenotypes result from effects on gene expression, particularly for the methylation of CGs located within gene promoters. We have recently shown that variation in *IGF1* P2 promoter methylation regulates *IGF1* gene expression in mononuclear blood cells <sup>9</sup>. *IGF1* class I transcripts have their initiation sites on exon 1 and are driven by P1 promoter, while class II transcripts use exon 2 as a leader exon and are driven by P2 promoter <sup>10</sup>. *IGF1* transcripts initiating at P1 are constitutively expressed in many tissues, while transcripts initiating at P2 are expressed primarily but not exclusively in the liver <sup>11,12</sup>. Following the demonstration of functional epigenetic effects of the *IGF1* P2 promoter methylation <sup>9,13</sup>, the current study explores the association between P2 promoter methylation and ISS.

## METHODS

### Subjects.

We recruited a group of 94 children (55M/39F) with strict criteria of ISS <sup>1</sup> whose clinical and biological characteristics are presented in Table 1. All were healthy and had normal clinical examination. GH deficiency was excluded with a stimulated GH peak >10 ng/ml. All children with ISS had normal TSH levels. Chondrodysplasia, notably dyschondrosteosis were excluded by forearm, pelvis and spine radiographs. A control group was formed with 119 unrelated healthy European white children (68M/51F) recruited within the EpiGrowth Cohort <sup>9,13</sup>. Trained nurses performed height measurements in duplicate using the Harpenden stadiometer. Ten ml peripheral blood samples were obtained, from which PBMC (peripheral blood mononuclear cells) were purified immediately. Parents gave written informed consent for the current study and for using surgical specimens, according to the French rules of bioethics in biomedical research. The research protocol received the agreement from Bicêtre I3E Institutional Review Board.

### Serum IGF1 measurements

Serum IGF1 concentration was measured around 7.00-8.00 am before breakfast in 84/94 ISS and 95/119 controls using an immune-radiometric assay after ethanol-acid extraction using DSL-5600 Active (Diagnostic System Laboratories, Webster, TX) or Cisbio reagents. Intra- and inter-series coefficients of variation were 1.5 and 3.7% at 260 ng/ml and 3.9% at 760 ng/ml. The sensitivity was 4 ng/ml. IGF1 SDS were calculated using norms recently established in French children <sup>14</sup>.

### Pyrosequencing-based bisulfite PCR analysis

For promoter P1, we studied 9 CGs located over a 800 bp distance, the closest CG being 225 bp upstream from the corresponding major transcription start site (TSS) <sup>15</sup>. For promoter P2, we studied 7/8 CGs located upstream from the major TSS within the proximal part of the promoter and 1 CG located 97 bp downstream this TSS. CGs were denominated according to their position versus each promoter TSS (Fig. 1). The methylation of CG-22 could not be measured for technical reasons. Nucleic acids were extracted from WBC or PBMC using Gentra Puregene blood kit (Qiagen). In a subset of 20 patients, WBC and PBMC were measured in the same samples <sup>9</sup>. Methylation was measured in other patients using PBMC. A bisulfite-PCR-pyrosequencing technique <sup>16</sup> was used to measure the methylation of the CGs. We improved the resolution of this method from a handful of bases to up to 100 nucleotides, with the ability to quantify methylation in the same sample of blood with a coefficient of variation (SD/mean) as little as 1-5%. Briefly, 400 nanograms of genomic DNA were treated with EZ DNA Methylation-Gold Kit, according to manufacturer's protocol (Zymo Research Corporation). The bisulfite-treated genomic DNA was PCR-amplified using unbiased *IGF1* primers <sup>9</sup> and performed quantitative pyrosequencing using a PyroMark Q96 ID Pyrosequencing instrument (Qiagen). Pyrosequencing assays were designed using MethPrimer (<http://www.urogene.org/methprimer/index1.html>). Biotin-labeled single stranded amplicons were isolated according to protocol using the Qiagen Pyromark Q96 Work Station and underwent pyrosequencing with 0.5  $\mu$ M primer. The percent methylation for each of the CGs within the target sequence was calculated using PyroQ CpG Software (Qiagen).

### Calculations and statistical methods

Serum IGF1 levels and height values were expressed as SDS to adjust for age and sex. Unpaired Student's *t*-tests were used to compare methylation levels in the two groups. Bonferroni correction was used to adjust for the number of CGs tested. All statistics were computed using R 2.10.1. Results are expressed as mean  $\pm$  SD.

## RESULTS

Serum IGF1 concentration was lower in the ISS group ( $P=10^{-7}$ ) (Table 1, Fig. 3A).

The patterns of CG methylation within the two promoters of *IGF1* are depicted in Fig. 1 and Table 2. Methylation in the proximal part of P2 promoter showed a strong correlation between several neighboring CGs (Fig. 2). In contrast, there was limited correlation between CGs within the P1 promoter. No correlation of methylation levels was detected between the CGs of the two promoters (Fig. 2 and supplemental Fig. 1).

The CG methylation within the P2 promoter was consistently higher in children with ISS (Fig.1 and Table 2). This was true for 5/8 of the studied CGs (Table 2 and Fig.1), notably for CG-137 ( $49 \pm 4\%$  in children with ISS vs  $46 \pm 4\%$  in controls,  $P=9 \times 10^{-5}$  after Bonferroni correction) (Fig. 3B and Table 2). In contrast, the level of CG methylation within the P1 promoter was comparable in ISS children and controls, except for CG-611 ( $P=10^{-4}$ ), which showed higher methylation in children with ISS (Table 2).

Methylation levels showed neither significant relationship nor any association trend with sex or age in the children with ISS and in the control children.

## DISCUSSION

1  
2  
3 IGF1 production disorders are a cause of short stature, as shown by inactivating mutations in  
4 the *IGF1* gene in humans<sup>17,18</sup>, and by the correlation between child height and circulating  
5 IGF1<sup>19</sup>. Common sequence polymorphisms at the *IGF1* locus, however, do not seem to  
6 contribute to height variation in Caucasians<sup>20</sup> while it does so in Asians in whom *IGF1* minor  
7 allele frequencies are different<sup>21,22</sup>.

8 The current observations show that epigenetic variation contributes to the multifactorial  
9 complexity of height variability among children, by showing that a higher CG methylation  
10 within the *IGF1* P2 promoter is associated with ISS. It is common to say that association does  
11 not mean causality, unless a relevant biological effect can be attributed to the associated  
12 marker. We have previously established that CG methylation at the P2 promoter influences  
13 *IGF1* transcription<sup>9</sup>. Although this observation has been made in PBMC, not in the growth  
14 plates of growing children, it is tempting to consider high CG methylation within the P2  
15 promoter as one of the multiple molecular mechanisms responsible for “idiopathic” short  
16 stature.

17  
18 The mechanisms generating individual variation of CG methylation across humans are not  
19 known. It is interesting in this respect that several of the studied CGs show a high degree of  
20 correlation between their methylation levels at the P2 promoter, which suggests a shared  
21 mechanism governing the variability of the P2 methylation status of these CGs in a given  
22 individual. In contrast, other CGs of the *IGF1* locus showed no correlation of their  
23 methylation level. Notably CG methylation at P1 and P2 seemed independent from each  
24 other. In particular, the higher methylation observed at CG-611 in children with ISS showed  
25 no statistical relationship with CG-137 or other CGs of the P2 promoter. In theory, variation  
26 in methylation can arise through stochastic events, developmental changes, or exposure to  
27 environmental cues and is also known to be influenced by *cis* or *trans* genomic variation.  
28 Previous data<sup>9</sup> have shown the lack of relationship between common *cis* genomic variation at  
29 the *IGF1* locus and methylation status of the two promoters. As observed in twins, the  
30 maternal environment can affect specific sites of the fetal epigenome<sup>23,24</sup>. This could in  
31 theory apply to the current observations in short children, although this remains a pure  
32 speculation in the absence of pregnancy data for the studied children.

33  
34 Another question raised by our observation is about the epigenetic inheritance (“soft  
35 inheritance”) of short stature. Indeed, modifications of DNA methylation patterns can be  
36 inherited trans-generationally, through incomplete erasure of epigenetic patterning in the  
37 germline and the early embryo<sup>25</sup>. This has not been shown, however, for the *IGF1* locus, and  
38 will be the matter of future familial and embryonic studies of *IGF1* epigenetics.

39  
40 Mice models have revealed that skeletal growth is mediated by IGF1 produced in situ by  
41 chondrocytes, not by circulating IGF1 produced by the liver<sup>26</sup>. The main physiological  
42 regulator of IGF1-mediated skeletal growth in chondrocytes is growth hormone (GH). The  
43 importance of epigenetic mechanisms in the regulation of *IGF1* gene expression by GH is  
44 unknown in growth plates but starts to be unraveled in the liver of hypophysectomised rats,  
45 where GH induces dramatic changes in chromatin at the *IGF1* locus and activates *IGF1*  
46 transcription by distinct promoter-specific epigenetic mechanisms at the P2 promoter<sup>27,28</sup>.

47  
48 The proximal part of rat P2 promoter regulates the transcriptional activity of the *IGF1* gene in  
49 response to GH via Stat5b<sup>28,29</sup>. We have previously shown that DNA methylation of the non-  
50 CG island P2 promoter of *IGF1* can directly silence gene expression<sup>9</sup>.

51  
52 The association of higher P2 promoter methylation with ISS suggests that the effects of P2  
53 methylation on growth are mediated through primary effects on IGF1 production. However,  
54 our observation should be interpreted with caution because of the lack of data in the growth  
55 plate, given the tissue specificity of DNA methylation levels. The lack of analysis of  
56 physiological tissues and the use of blood cells as surrogates<sup>30,31</sup> are a common weakness of  
57 epigenetic epidemiological studies<sup>32-34</sup>.

1  
2  
3 In summary, the highest methylation observed in several CGs of the *IGF1* P2 promoter and  
4 one CG of the P1 promoter in the PBMC of children with ISS supports that epigenetics plays  
5 a major role in the causality of short stature.  
6  
7

### 8 **Competing interests**

9 The authors declare that they have no competing interests.  
10

### 11 **Authors' contributions**

12 PB designed the study, recruited the patients, interpreted the data and wrote the ms. MO did  
13 the experiments, analyzed the data and interpreted the results. ALC, AR and AL contributed  
14 to the clinical and epigenetic observations.  
15  
16

### 17 **Acknowledgements**

18 NovoNordisk France and Pfizer France supported this project with research grants. MO is  
19 supported by Association de Recherche sur le Diabète. We thank MP Belot for the  
20 measurement of CG methylation in P1 promoter.  
21  
22  
23  
24  
25  
26

### 27 **References**

- 28 1 Cohen P, Rogol AD, Deal CL, Saenger P, Reiter EO, Ross JL *et al.* Consensus statement  
29 on the diagnosis and treatment of children with idiopathic short stature: a summary of the  
30 Growth Hormone Research Society, the Lawson Wilkins Pediatric Endocrine Society, and  
31 the European Society for Paediatric Endocrinology Workshop. *J Clin Endocrinol Metab*  
32 2008; **93**: 4210–4217.  
33
- 34 2 Wit JM, Clayton PE, Rogol AD, Savage MO, Saenger PH, Cohen P. Idiopathic short  
35 stature: definition, epidemiology, and diagnostic evaluation. *Growth Horm IGF Res Off J*  
36 *Growth Horm Res Soc Int IGF Res Soc* 2008; **18**: 89–110.  
37
- 38 3 Sotos JF, Tokar NJ. Growth hormone significantly increases the adult height of children  
39 with idiopathic short stature: comparison of subgroups and benefit. *Int J Pediatr*  
40 *Endocrinol* 2014; **2014**: 15.  
41
- 42 4 Dauber A, Rosenfeld RG, Hirschhorn JN. Genetic evaluation of short stature. *J Clin*  
43 *Endocrinol Metab* 2014; **99**: 3080–3092.  
44
- 45 5 Visscher PM, McEVOY B, Yang J. From Galton to GWAS: quantitative genetics of  
46 human height. *Genet Res* 2010; **92**: 371–379.  
47
- 48 6 Fisher RA. XV.—The Correlation between Relatives on the Supposition of Mendelian  
49 Inheritance. *Earth Environ Sci Trans R Soc Edinb* 1919; **52**: 399–433.  
50
- 51 7 Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S *et al.* Defining the role of  
52 common variation in the genomic and biological architecture of adult human height. *Nat*  
53 *Genet* 2014. doi:10.1038/ng.3097.  
54  
55  
56  
57  
58  
59  
60



- 1  
2  
3 8 Van Duyvenvoorde HA, Lui JC, Kant SG, Oostdijk W, Gijsbers ACJ, Hoffer MJV *et al.* Copy number variants in patients with short stature. *Eur J Hum Genet EJHG* 2014; **22**: 602–609.
- 4  
5  
6  
7 9 Ouni M, Gunes Y, Belot M-P, Castell A-L, Fradin D, Bougnères P. The IGF1 P2 promoter is an epigenetic QTL for circulating IGF1 and human growth. *Clin Epigenetics* 2015; **7**: 22.
- 8  
9  
10  
11 10 Yang H, Adamo ML, Koval AP, McGuinness MC, Ben-Hur H, Yang Y *et al.* Alternative leader sequences in insulin-like growth factor I mRNAs modulate translational efficiency and encode multiple signal peptides. *Mol Endocrinol Baltim Md* 1995; **9**: 1380–1395.
- 12  
13  
14  
15 11 Rotwein P. Mapping the growth hormone--Stat5b--IGF-I transcriptional circuit. *Trends Endocrinol Metab TEM* 2012; **23**: 186–193.
- 16  
17  
18  
19 12 Oberbauer AM. The Regulation of IGF-1 Gene Transcription and Splicing during Development and Aging. *Front Endocrinol* 2013; **4**: 39.
- 20  
21  
22  
23 13 Ouni M, Belot MP, Castell AL, Fradin D, Bougnères P. The P2 promoter of the IGF1 gene is a major epigenetic locus for GH responsiveness. *Pharmacogenomics J* 2015. doi:10.1038/tpj.2015.26.
- 24  
25  
26  
27 14 Alberti C, Chevenne D, Mercat I, Josserand E, Armoogum-Boizeau P, Tichet J *et al.* Serum concentrations of insulin-like growth factor (IGF)-1 and IGF binding protein-3 (IGFBP-3), IGF-1/IGFBP-3 ratio, and markers of bone turnover: reference values for French children and adolescents and z-score comparability with other references. *Clin Chem* 2011; **57**: 1424–1435.
- 28  
29  
30  
31 15 Jansen E, Steenbergh PH, LeRoith D, Roberts CT, Sussenbach JS. Identification of multiple transcription start sites in the human insulin-like growth factor-I gene. *Mol Cell Endocrinol* 1991; **78**: 115–125.
- 32  
33  
34  
35 16 Tost J, Gut IG. DNA methylation analysis by pyrosequencing. *Nat Protoc* 2007; **2**: 2265–2275.
- 36  
37  
38  
39 17 Woods KA, Camacho-Hübner C, Barter D, Clark AJ, Savage MO. Insulin-like growth factor I gene deletion causing intrauterine growth retardation and severe short stature. *Acta Paediatr Oslo Nor 1992 Suppl* 1997; **423**: 39–45.
- 40  
41  
42  
43 18 Lupu F, Terwilliger JD, Lee K, Segre GV, Efstratiadis A. Roles of growth hormone and insulin-like growth factor 1 in mouse postnatal growth. *Dev Biol* 2001; **229**: 141–162.
- 44  
45  
46  
47 19 Rogers I, Metcalfe C, Gunnell D, Emmett P, Dunger D, Holly J *et al.* Insulin-like growth factor-I and growth in height, leg length, and trunk length between ages 5 and 10 years. *J Clin Endocrinol Metab* 2006; **91**: 2514–2519.
- 48  
49  
50  
51 20 Lettre G, Butler JL, Ardlie KG, Hirschhorn JN. Common genetic variation in eight genes of the GH/IGF1 axis does not contribute to adult height variation. *Hum Genet* 2007; **122**: 129–139.
- 52  
53  
54  
55 21 Kim J-J, Lee H-I, Park T, Kim K, Lee J-E, Cho NH *et al.* Identification of 15 loci influencing height in a Korean population. *J Hum Genet* 2009; **55**: 27–31.
- 56  
57  
58  
59  
60

- 1  
2  
3 22 Okada Y, Kamatani Y, Takahashi A, Matsuda K, Hosono N, Ohmiya H *et al.* A genome-  
4 wide association study in 19 633 Japanese subjects identified LHX3-QSOX2 and IGF1 as  
5 adult height loci. *Hum Mol Genet* 2010; **19**: 2303–2312.  
6  
7 23 Petronis A. Epigenetics and twins: three variations on the theme. *Trends Genet TIG* 2006;  
8 **22**: 347–350.  
9  
10 24 Kaminsky ZA, Tang T, Wang S-C, Ptak C, Oh GHT, Wong AHC *et al.* DNA methylation  
11 profiles in monozygotic and dizygotic twins. *Nat Genet* 2009; **41**: 240–245.  
12  
13 25 Smith ZD, Chan MM, Humm KC, Karnik R, Mekhoubad S, Regev A *et al.* DNA  
14 methylation dynamics of the human preimplantation embryo. *Nature* 2014; **511**: 611–615.  
15  
16 26 Liu JL, Yakar S, LeRoith D. Conditional knockout of mouse insulin-like growth factor-I  
17 gene using the Cre/loxP system. *Proc Soc Exp Biol Med Soc Exp Biol Med N Y N* 2000;  
18 **223**: 344–351.  
19  
20 27 Chia DJ, Young JJ, Mertens AR, Rotwein P. Distinct alterations in chromatin  
21 organization of the two IGF-I promoters precede growth hormone-induced activation of  
22 IGF-I gene transcription. *Mol Endocrinol Baltim Md* 2010; **24**: 779–789.  
23  
24 28 Varco-Merth B, Mirza K, Alzhanov DT, Chia DJ, Rotwein P. Biochemical  
25 characterization of diverse Stat5b-binding enhancers that mediate growth hormone-  
26 activated insulin-like growth factor-I gene transcription. *PloS One* 2012; **7**: e50278.  
27  
28 29 Chia DJ, Rotwein P. Defining the epigenetic actions of growth hormone: acute chromatin  
29 changes accompany GH-activated gene transcription. *Mol Endocrinol Baltim Md* 2010;  
30 **24**: 2038–2049.  
31  
32 30 Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T *et al.*  
33 Personalized epigenomic signatures that are stable over time and covary with body mass  
34 index. *Sci Transl Med* 2010; **2**: 49ra67.  
35  
36 31 Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G *et al.* Variation,  
37 patterns, and temporal stability of DNA methylation: considerations for epigenetic  
38 epidemiology. *FASEB J Off Publ Fed Am Soc Exp Biol* 2010; **24**: 3135–3144.  
39  
40 32 Richards EJ. Population epigenetics. *Curr Opin Genet Dev* 2008; **18**: 221–226.  
41  
42 33 Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for  
43 common human diseases. *Nat Rev Genet* 2011; **12**: 529–541.  
44  
45 34 Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology.  
46 *Nat Rev Genet* 2013; **14**: 585–594.  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



## Tables

**Table 1.** Main clinical and biological characteristics of the studied children (mean  $\pm$  SD).

**Table 2.** Comparison of CG methylation in the two promoters of *IGF1* gene in the studied groups of children (mean  $\pm$  SD).

## Fig. Legends

**Figure 1.** The two *IGF1* gene promoters (P1 and P2) are figured in the lower part of the Figure. (broken arrows show the TSS of each promoter). The upper part of the figure shows the mean methylation level of each studied CG in the studied children. Children with ISS are figured with black barplot and children of normal height with white barplot.

\*  $p < 0.05$  \*\*  $p \leq 0.01$  \*\*\*  $\leq 0.0001$  (Bonferroni correction).

**Figure 2.** Correlation matrix of methylation levels among the CGs located in the P2 and P1 promoters of the *IGF1* gene. The two groups of children have been merged for analysis. To simplify visualization, black boxes indicate only the strongest correlations, defined by  $R > 0.30$  and  $P < 10^{-3}$ . Values for R coefficients and P values are presented in supplemental Figure 1.

**Figure 3.** Comparison of the studied children (A) for serum IGF1 concentration ( $P=1.5 \times 10^{-7}$ ); (B) for CG-137 methylation ( $P=7 \times 10^{-6}$ ).

**Supplemental Figure 1.** Correlation matrix of methylation levels in the P2 and P1 promoters of the *IGF1* gene. Values for R coefficient and P value for each pair of studied CGs are presented as R / P.

**Table 1.** Main clinical and biological characteristics of the studied children (mean±SD).

	Children with idiopathic short stature	Controls of normal height	P
<i>N</i>	94	119	-
Sex (M/F)	55/39	68/51	NS
Age (yrs)	9 ± 3	10 ± 3	NS
Height (SDS)	-2.9 ± 0.5	-0.6 ± 1	<2.10 <sup>-16</sup>
Tanner stage (N)			
1	72 (77%)	86 (72%)	NS
2	12 (13%)	12 (10%)	NS
3	4 (4%)	10 (8%)	NS
4	4 (4%)	11 (10%)	NS
Birth length (SDS)	-2 ± 1	-1 ± 1	10 <sup>-7</sup>
Parental height			
Mother (SDS)	-1.3 ± 1	-0.6 ± 1	10 <sup>-2</sup>
Father (SDS)	-0.9 ± 1	-0.6 ± 1	NS
Serum IGF1			
(ng/ml)	129 ± 82	227 ± 132	10 <sup>-8</sup>
(SDS)	-1.7 ± 1	-0.7 ± 1	10 <sup>-7</sup>

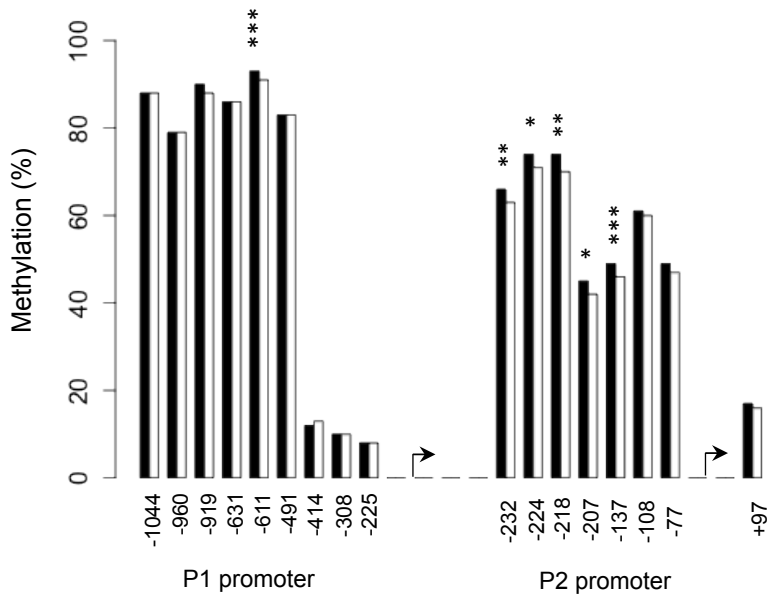
**Table 2.** Comparison of CG methylation in the two promoters of the *IGF1* gene in the studied groups of children (mean  $\pm$  SD).

		Children with idiopathic short stature	Controls of normal height	p*
P1 promoter	CG -1044	88 $\pm$ 3	88 $\pm$ 3	1
	CG -960	79 $\pm$ 2	79 $\pm$ 5	1
	CG -919	90 $\pm$ 4	88 $\pm$ 6	1
	CG -631	86 $\pm$ 2	86 $\pm$ 2	1
	CG -611	93 $\pm$ 3	91 $\pm$ 3	10 <sup>-4</sup>
	CG -491	83 $\pm$ 4	83 $\pm$ 4	1
	CG -414	12 $\pm$ 3	13 $\pm$ 5	1
	CG -308	10 $\pm$ 4	10 $\pm$ 4	1
	CG -225	8 $\pm$ 2	8 $\pm$ 2	1
P2 promoter	CG -232	66 $\pm$ 7	63 $\pm$ 7	0.005
	CG -224	74 $\pm$ 7	71 $\pm$ 7	0.026
	CG -218	74 $\pm$ 7	70 $\pm$ 6	0.008
	CG -207	45 $\pm$ 7	42 $\pm$ 7	0.04
	CG -137	49 $\pm$ 4	46 $\pm$ 4	9.10 <sup>-5</sup>
	CG -108	61 $\pm$ 7	60 $\pm$ 6	1
	CG -77	49 $\pm$ 6	47 $\pm$ 5	1
	CG +97	17 $\pm$ 3	16 $\pm$ 3	1

\*after Bonferroni correction for the number of CGs tested (non significant values equated to 1)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58

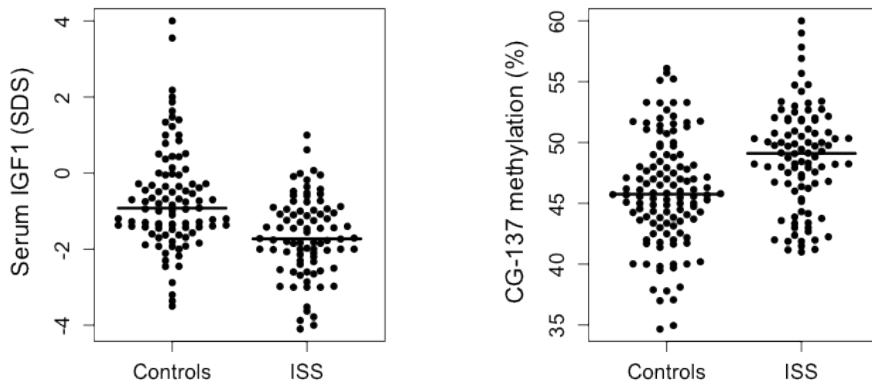
Figure 1.





1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58

Figure 3.



Supplemental Figure 1.

P1 promoter									P1 promoter									
-960	-919	-631	-611	-491	-414	-308	-225	-232	-224	-218	-207	-137	-108	-77	97			
0.14	0.22	0.00	0.00	-0.03	-0.05	0.06	0.17	-0.06	0.11	0.19	0.20	0.06	0.10	0.04	-0.05	-1044		
0.02	5.10 <sup>-3</sup>	0.98	0.90	0.76	0.24	0.18	8.10 <sup>-3</sup>	0.94	0.05	3.10 <sup>-3</sup>	10 <sup>-3</sup>	0.87	0.07	0.34	0.55	-960		
	0.49	-0.11	-0.04	0.26	0.05	0.08	0.01	-0.13	0.02	-0.07	-0.03	-0.12	0.13	-0.01	0.01	-919		
	<10 <sup>-3</sup>	0.70	0.68	0.56	0.56	0.13	0.80	0.15	0.80	0.42	0.74	0.19	0.16	0.87	0.90			
		-0.07	-0.01	0.07	-0.04	0.02	0.04	-0.09	0.08	0.06	-0.03	-0.10	0.24	-0.08	0.08			
		0.89	0.20	0.67	0.67	0.80	0.25	0.31	0.39	0.49	0.75	0.26	9.10 <sup>-3</sup>	0.38	0.40			
			0.42	0.05	-0.08	0.16	-0.03	-0.10	-0.13	-0.05	-0.11	0.10	0.07	0.03	-0.17	-631		
			<10 <sup>-3</sup>	0.37	0.37	0.01	0.74	0.26	0.15	0.58	0.22	0.27	0.44	0.75	0.06			
				0.07	-0.10	0.18	0.00	0.21	0.01	0.14	0.06	0.05	0.10	0.10	-0.06	-611		
				0.44	0.30	5.10 <sup>-3</sup>	0.97	10 <sup>-3</sup>	0.29	0.02	0.18	0.22	0.08	0.07	0.50			
					-0.06	0.17	-0.32	0.05	0.11	0.14	-0.08	0.01	0.17	0.04	-0.11	-491		
					0.53	8.10 <sup>-3</sup>	4.10 <sup>-4</sup>	0.60	0.22	0.13	0.41	0.92	0.06	0.66	0.24			
						0.36	0.07	-0.13	-0.19	-0.24	-0.03	-0.02	-0.16	-0.03	0.06	-414		
						<10 <sup>-3</sup>	0.43	0.16	0.04	0.01	0.75	0.83	0.08	0.76	0.53			
							0.34	-0.02	-0.03	0.04	-0.05	-0.01	0.09	-0.08	-0.05	-308		
							<10 <sup>-3</sup>	0.80	0.75	0.67	0.60	0.92	0.33	0.41	0.55			
								0.00	-0.11	-0.04	0.15	0.06	-0.26	0.03	-0.01	-225		
								0.99	0.24	0.64	0.11	0.48	5.10 <sup>-3</sup>	0.76	0.95			
									0.67	0.66	0.30	0.55	0.62	0.43	0.25	-232		
									<10 <sup>-3</sup>	<10 <sup>-3</sup>	<10 <sup>-3</sup>	<10 <sup>-3</sup>	<10 <sup>-3</sup>	<10 <sup>-3</sup>	6.10 <sup>-3</sup>			
										0.85	0.57	0.50	0.63	0.21	0.12	-224		
										<10 <sup>-3</sup>	<10 <sup>-3</sup>	<10 <sup>-3</sup>	0.02	0.18	0.18			
											0.64	0.44	0.63	0.13	0.02	-218		
											<10 <sup>-3</sup>	<10 <sup>-3</sup>	<10 <sup>-3</sup>	0.16	0.80			
												0.21	0.35	0.09	0.12	-207		
												10 <sup>-3</sup>	<10 <sup>-3</sup>	0.33	0.20			
													0.53	0.47	0.36	-137		
													<10 <sup>-3</sup>	<10 <sup>-3</sup>	<10 <sup>-3</sup>			
														0.33	0.25	-108		
														5.10 <sup>-3</sup>	5.10 <sup>-3</sup>			
															0.59	-77		
															<10 <sup>-3</sup>			
																97		

P1 promoter

P2 promoter

## **12.2 CONCLUSIONS DE L'ARTICLE 4**

Les taux d'IGF1 circulant sont significativement plus bas chez les enfants de petite taille. En comparant le profil de méthylation des deux groupes d'individus, nous avons montré de légères différences de méthylation qui sont néanmoins très significatives dans les deux promoteurs. Un seul CG (CG-611) du promoteur 1 présente des différences significatives entre patients et contrôles. Dans le promoteur 2, un groupe de 5 CG diffèrent entre les deux groupes testés.



### 13 DISCUSSION GENERALE

Nous avons abordé durant cette thèse le rôle de l'épigénétique dans le domaine de la croissance staturale. Dans ce travail, nous avons montré la contribution de la méthylation de l'ADN dans deux types de traits quantitatifs liés à la croissance :

- des traits quantitatifs dits naturels, la taille et le taux d'IGF1 circulant ;
- des traits quantitatifs modulés par un facteur externe, à savoir des doses thérapeutiques d'hormone de croissance.

En 2011, quand nous avons commencé ce travail, la régulation épigénétique du gène *IGF1* chez l'homme était inconnue et aucune étude n'avait été faite sur la méthylation de ce locus. Nous avons donc dû faire le design de l'approche par bisulfite/pyroséquençage et réaliser toutes les mises au point. Malgré l'absence de CGI dans ces promoteurs, généralement difficile à amplifier et séquencer, les mises au point ont été fastidieuses.

La méthylation du gène *IGF1* chez l'homme commence à apparaître dans les données récentes de la littérature. En revanche, les régions investiguées varient d'une étude à l'autre et aucune d'entre elles ne s'est intéressée aux régions régulatrices que nous avons ciblées. De plus, la méthylation a été mesurée dans différents types cellulaires, tels que placenta dans l'étude de Desgagné et collaborateurs (Desgagné et al., 2014) ou dans le sang de cordon pour l'étude de Straughen et collaborateurs (Straughen et al., 2015). Ces travaux portent sur la croissance prénatale dans deux ethnies et montrent une corrélation entre le poids de naissance et la méthylation du gène *IGF1*. Aucune variabilité interindividuelle n'a été observée cependant, comme nous l'avons déjà dit la méthylation de l'ADN au niveau du promoteur 2, là où sont concentrés la majorité de nos résultats, n'a pas été étudiée.

La description complète des profils de méthylation des deux promoteurs du gène *IGF1* n'a été obtenue que dans notre étude. Ces deux promoteurs présentent une variabilité individuelle de méthylation de l'ADN et peuvent être classés dans les catégories de iDMRs, VMRs (Feinberg et al., 2010) ou d'épiallèles (Finer et al., 2011). Nous spéculons que les différences individuelles observées sont mises en place précocement au stade embryonnaire, au moment de la programmation individuelle de la méthylation. Nous ne pouvons cependant proposer de mécanisme pour l'expliquer, même s'il ne s'agit probablement pas d'un mécanisme global touchant par exemple la disponibilité en donneurs de groupements méthyles ou l'activité des DNMTs. En revanche, un complexe protéique promoteur P2 spécifique pourrait probablement inscrire une méthylation différente entre individus selon la disponibilité des groupes méthyles. Une origine génétique ne semble pas non plus expliquer cette variabilité puisque nous n'avons pu corrélérer les allèles du microsatellite du promoteur 1 à un profil de méthylation particulier. Classiquement dans ce genre d'étude, la recherche d'une cause génétique dans la variation de méthylation cible les SNPs plutôt que les microsatellites. Nous avons toutefois choisi d'analyser ce microsatellite pour 3 raisons :

- les SNPs qui sont dans les deux promoteurs d'*IGF1* ou à proximité sont très rares ;
- la position du microsatellite est en plein milieu des CGs analysés dans le promoteur donc probablement capable d'établir des liens avec la méthylation ;

- nous avons spéculé que le nombre de répétition du motif CA dans cette région pourrait agir sur la structure tridimensionnelle de la chromatine au niveau du locus *IGF1* et ainsi sur son accessibilité. Cet exemple existe déjà dans la littérature des VNTRs (Melas et al., 2013).

Le gène *IGF1* possède une régulation transcriptionnelle atypique, nous ne pouvons donc exclure la présence de variants génétiques au niveau des enhanceurs du gène pouvant moduler la méthylation des promoteurs.

D'autre part, nous avons constaté une forte corrélation entre la méthylation des CGs du même promoteur, notamment pour le promoteur 2. Ceci n'était pas particulièrement nouveau dans l'analyse de méthylation de gènes candidats (Fradin et al., 2012) mais les corrélations entre les CGs du promoteur P2 du gène *IGF1* sont plus fortes que les observations précédentes. Sachant que toutes les vérifications techniques ont été réalisées, notre observation ne semble pas dûe à un phénomène de sélection d'amorces. Nous spéculons que ce « déséquilibre de liaison » entre les CGs pourraient s'expliquer par un mécanisme commun et conjoint lors de l'établissement de la méthylation.

En conclusion, le promoteur P2 du gène *IGF1* serait le premier épIQTL décrit chez l'homme (Johannes et al., 2008). Cette thèse présente cependant quelques limites que nous discuterons dans les paragraphes suivants. Nous parlerons également des questions et hypothèses soulevées par nos résultats.

### **Rôle de la méthylation sur la transcription du gène *IGF1* ?**

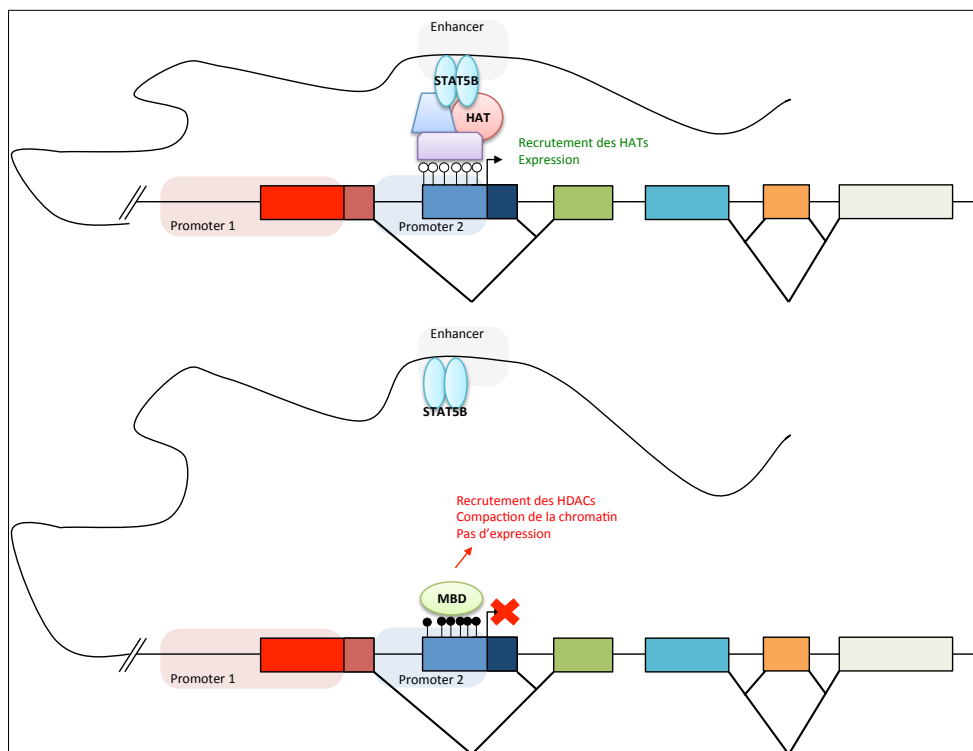
Nous avons montré des corrélations significatives entre les phénotypes étudiés et la méthylation de l'ADN du P2 du gène *IGF1*. Il est courant de dire qu'une association ne signifie pas obligatoirement causalité, sauf si un effet biologique pertinent du paramètre associé peut être démontré.

*IGF1* possède deux promoteurs de faible densité en CG, selon la classification de (Weber et al., 2007). Le rôle régulateur de la méthylation de ces promoteurs est controversé, certains ont montré que la méthylation n'empêchait pas la transcription des gènes tandis que d'autres au contraire montraient une corrélation inverse entre méthylation et expression (Ball et al., 2009; Meissner et al., 2008; Shen et al., 2013). Nos résultats sont en cohérence avec ceux de Weber puisque nous montrons que la méthylation du promoteur P2 bloque la transcription du gène dans un contexte plasmidique. Mais cette structure moléculaire est dépourvue du contexte chromatinien qui existe dans une cellule dans les conditions physiologiques normales. Nous n'avons pas pu aborder cette question durant cette thèse, nous avons donc émis des hypothèses pour expliquer le mode de régulation de la transcription du gène *IGF1* :

1) Les facteurs de transcription qui se fixent sur les deux promoteurs proximaux d'*IGF1* ne contiennent pas de dinucléotide CG dans leur séquence cible, une régulation par accessibilité directe des facteurs par l'état de méthylation semble donc exclue (Campanero et al., 2000; Iguchi-Arigo and Ariga, 1989).

2) Il est possible que la méthylation du P2 interagisse avec des complexes répresseurs et/ou des complexes de remodelage de la chromatine afin de moduler l'expression du gène *IGF1*. A notre connaissance, aucun répresseur n'a été identifié pour le gène *IGF1* humain.

3) La transcription du gène *IGF1* est dépendante en partie de la GH. La GH exerce ses effets *via* la protéine STAT5b. Chez l'homme, il n'y a aucun site STAT5b dans le P1 ou le P2. Il en existe cependant dans les introns ainsi que dans les enhancers situés en amont du gène (voir la partie Introduction). Il est possible que la méthylation empêche les interactions entre le P2 et l'enhancer fixant les protéines Stat5b (Figure 33). Nous proposons un modèle hypothétique d'interaction entre la méthylation des promoteurs et les facteurs STAT5b. Dans ce modèle, nous pensons que la méthylation du promoteur P2 fait intervenir des complexes incluant les MBDs et les HDACs pour empêcher la fixation de l'ARN pol II et inhiber la transcription du gène *IGF1*. En l'absence de méthylation, la région du promoteur P2 et les enhancers peuvent interagir en recrutant les HATs et permettre la transcription. Nous avons proposé un modèle qui fait intervenir les complexes HAT et HDAC, en se basant sur l'unique étude réalisée sur la régulation épigénétique du locus *IGF1* (Voir détail dans l'introduction, (Chia and Rotwein, 2010; Chia et al., 2010). Le type d'interaction que nous proposons a été démontré dans la régulation transcriptionnelle de certains gènes (Ciccone et al., 2009).



**Figure 33 : Proposition d'un mécanisme de régulation transcriptionnelle au locus *IGF1*.** En haut, l'activation du gène *IGF1* faisant intervenir l'acétylation des histones par les protéines HAT et l'interaction entre les protéines STAT5B liées à l'enhancer et le P2 déméthylé. En bas, l'état inactif du gène avec l'intervention des protéines MBDs et la désacétylation des histones par les HDACs.

En dehors de cette construction plasmidique, les corrélations retrouvées dans ce travail sont plus ou moins fortes selon les dinucléotides CGs étudiés. Le CG-137 du promoteur 2 est retrouvé le plus associé à l'expression des gènes et aux phénotypes étudiés. Le rôle majoritaire d'un seul CG dans

l'expression d'un gène a déjà été décrit (ins par (Kuroda et al., 2009), *rarb2* (Le May et al., 2012) et autres...). Potentiellement, les facteurs de remodelage et corépresseurs de la chromatine tel que les MBD pourraient se fixer de manière site spécifique sur le promoteur 2 et peut être même de façon ciblée sur certains CGs. Il est possible de vérifier l'importance de chaque CG du promoteur 2 dans la régulation de la transcription d'*IGF1 in vitro* en adaptant notre construction plasmidique contenant le promoteur P2 et en réalisant une méthylation CG spécifique. Ce travail est assez fastidieux et n'était pas réalisable durant cette thèse. De plus, comme nous l'avons déjà dit, ce type d'expérience ne prend pas en compte la conformation chromatinienne dans les conditions physiologiques.

### La méthylation du locus *IGF1* est une cause ou une conséquence ?

Nous avons montré qu'il existait un lien entre la méthylation et la transcription du gène *IGF1* mais des questions demeurent quant à savoir si la méthylation est causale dans l'ensemble des phénotypes étudiés.

Nous avons émis 2 hypothèses :

1) La variation de la méthylation des promoteurs du gène *IGF1* est une marque épigénétique pré-établie pour chaque individu durant l'embryogénèse et reste inchangée tout au long de la vie. Ces marques pré-définissent la réponse à l'hormone de croissance et ne peuvent être modifiées. Cette hypothèse est étayée par nos tests de génération d'*IGF1*. Quand 40% de l'ensemble de cellules sont méthylés, ce pourcentage reste le même après un dose aigüe de GH. Ainsi s'explique l'incapacité des individus hyperméthylés à induire une expression du gène *IGF1*.

2) La méthylation des promoteurs du gène *IGF1* est modulée par l'hormone de croissance. Nous postulons que la méthylation au niveau des deux promoteurs d'*IGF1* est variable entre individus depuis son établissement, mais capable d'être modulée par l'hormone de croissance pendant la vie postnatale. L'hormone de croissance pourrait déméthyliser le promoteur P2 pour induire l'expression du gène *IGF1*. Pour étayer cette hypothèse, nous avons étudié la méthylation du P2 après une exposition prolongée à la GH (au moins un an) et observé une hypo-méthylation proportionnelle à la dose de GH et la durée de l'exposition (Figures 34 et 35).

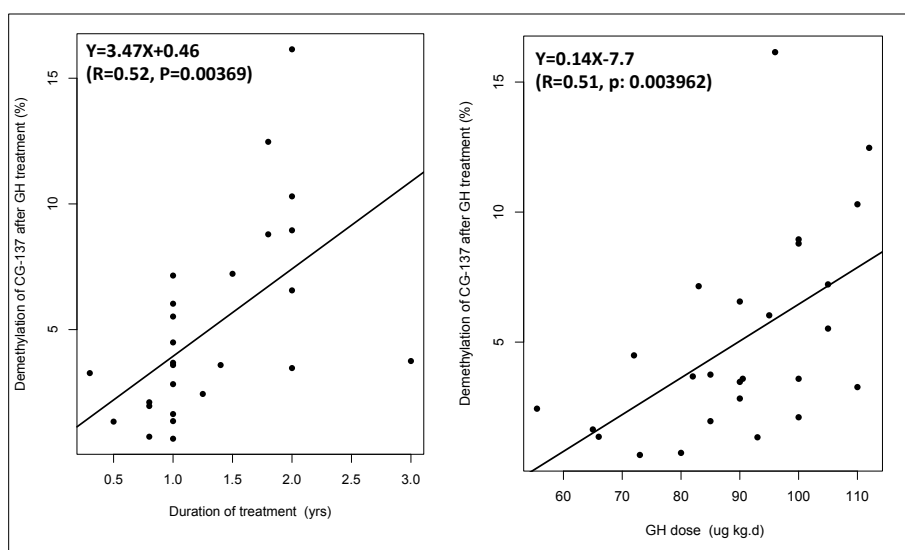


Figure 34 : Déméthylation du CG-137 du promoteur d'*IGF1* en fonction de la dose du traitement GH et de la durée du traitement. (N=25)

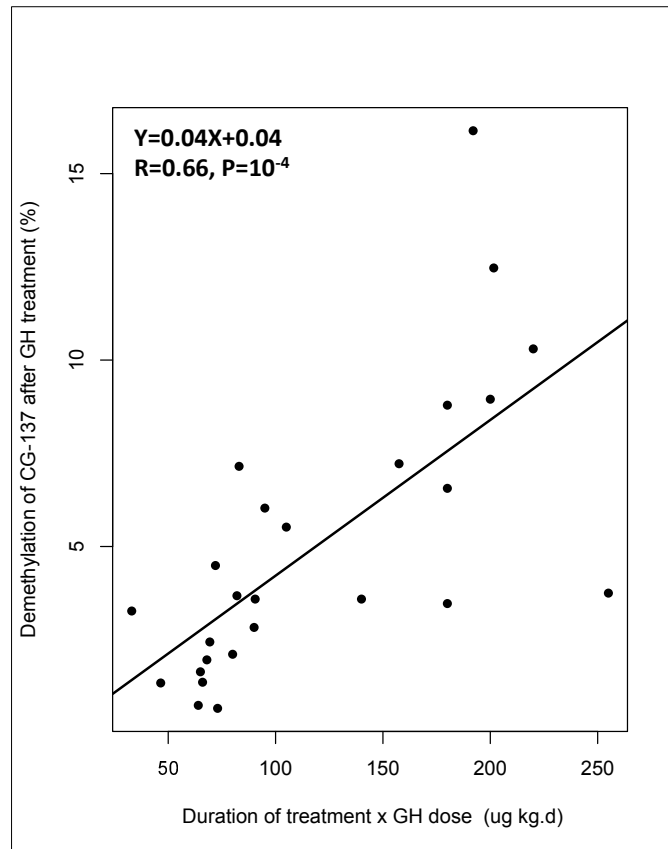


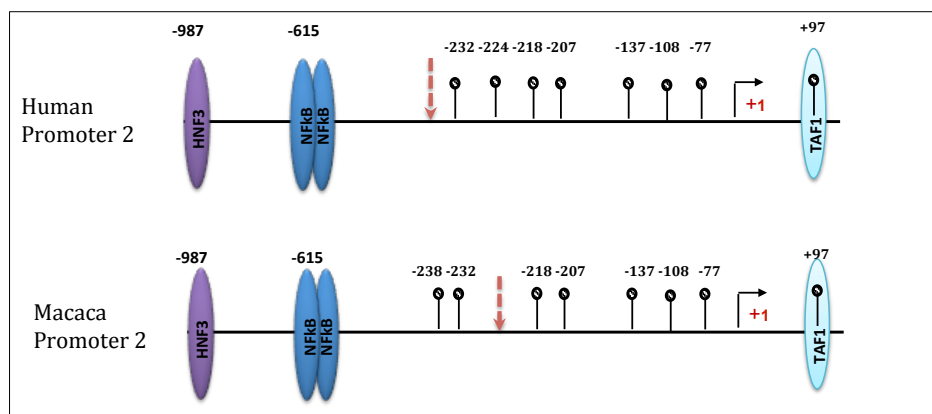
Figure 35: La durée et la dose du traitement GH sont associées à la déméthylation du CG-137.

Afin d'explorer plus en détails cette hypothèse nous nous sommes intéressés aux lignées cellulaires issues des tissus cibles de la GH, cependant nous n'avons pu aller plus loin car ces lignées n'expriment que très peu l'*IGF1*. La lignée de foie humain HepG2, par exemple, exprime des niveaux d'*IGF1* plus faibles que les PBMCs. Nous avons donc essayé de stimuler des PBMCs de patients *in vitro* avec des doses de GH, mais l'expression du gène *IGF1* dans des cultures primaires de PBMCs baisse beaucoup et très vite au fil du temps. Notre seule explication pour cette baisse d'expression est que la mise en culture induit une forte mortalité cellulaire, comme pour toutes les lignées primaires, et entraîne donc une difficulté supplémentaire pour détecter les transcrits d'*IGF1* déjà rares. Ainsi notre question concernant la capacité de l'hormone de croissance à déméthyler reste sans réponse claire.

Les exemples de déméthylation active spécifique d'un locus en réponse à une hormone dans les cellules différenciées sont limités. Nous citons ici un exemple typique de méthylation cyclique de l'ADN au niveau d'un loci cible du récepteur aux estrogènes, pS2. Après une stimulation par les estrogènes, le promoteur de pS2 entre dans un cycle de méthylation-déméthylation prouvant que la méthylation peut être réversible mais aussi dynamique (Schomacher, 2013). Les exemples de déméthylation dans les cellules différenciées sont aussi connus dans la régulation de la transcription des gènes du système immunitaire, tel que le locus de l'interleukine 2, qui se déméthylye en 20 minutes sans réplication de l'ADN (Bruniquel and Schwartz, 2003). En comparant avec ces exemples de déméthylation, nous proposons deux hypothèses :

- La GH est capable de déméthyle le promoteur P2 très vite suite à l'injection mais après 12h de stimulation, le promoteur P2 se reméthyle de nouveau comme le gène pS2.
  - En se basant sur les exemples de déméthylation des cellules différenciées dans la littérature, il est possible qu'une déméthylation active se réalise au niveau du promoteur P2 par oxydation et ensuite par excision basique (BER). Selon cette hypothèse, nous ne serions pas en mesure de voir cette déméthylation car le traitement au bisulfite ne permet pas de distinguer la méthylation de l'hydroxyméthylation. A notre connaissance, l'hydroxy-méthyl-cytosine est un état transitoire dans ce mécanisme de déméthylation mais la durée de cette transition est inconnue.
- Notre hypothèse 1 stipulant que la méthylation du locus IGF1 est stable n'est plus valable dans ces conditions, le fait que nous ne voyions pas la méthylation changer après une dose aiguë de GH serait uniquement dû à un problème de « timing ».

Le modèle animal serait l'idéal pour étudier l'action de la GH sur la méthylation du promoteur P2 *in vivo*. Afin de trouver l'animal le plus approprié, nous avons comparé les séquences du promoteur 2 entre différentes espèces : la souris (*Mus musculus*), le rat (*Rattus norvegicus*), le macaque (*Macaca mulatta*) et l'homme. Aucun dinucléotide CG n'est conservé entre l'homme et la souris ou le rat selon un alignement de séquence réalisé sur NCBI (<https://blast.ncbi.nlm.nih.gov/Blast>). Une forte homologie de séquence a été observée entre les séquences du promoteur P2 chez l'homme et le macaque puisque seul le CG-224 n'est pas conservé (Figure 36). Nous avons réalisé des études préliminaires dans ce modèle animal, mais aucun résultat n'est pour l'instant exploitable.



**Figure 36 : Représentation schématique du promoteur P2 et des CGs étudiés chez l'homme et le macaque.** Les ronds hachurés correspondent aux CGs méthylés ou non. La flèche rouge indique les CGs manquants dans l'une ou l'autre des espèces. Les ovales représentent des facteurs de transcription et la position de leur site de fixation sur le locus *IGF1*. La numérotation des CGs a été faite à partir du TSS noté +1.

### Limites du tissu étudié et pertinence ?

La plupart des études épigénétiques en cohorte ont été réalisées sur de l'ADN extrait des cellules nucléées du sang, pour une raison évidente qui est leur accessibilité, bien que souvent ce ne soit pas le tissu physiologique où le trait prend naissance. Ces études sont réalisées soit à partir i) de la totalité des cellules du sang (WBC, whole blood cell) ; ii) des cellules mononucléées du sang (PBMCs, Peripheral Blood Mononuclear Cells) ou iii) d'une population cellulaire triée (par exemple les lymphocyte T CD4+). Comme nous l'avons déjà dit, l'épigénétique contrôle l'expression des gènes, et participe de ce fait à la spécificité cellulaire. En effet, certaines marques épigénétiques sont

spécifiques du tissu considéré (voir le paragraphe sur les T-DMRs). Il existe cependant des modifications présentes uniformément dans toutes les cellules de l'organisme et d'autres qui, bien que différentes, présentent une forte corrélation entre les différents tissus (Dominguez-Salas et al., 2014). Nous avons étudié la méthylation des promoteurs du gène *IGF1* dans des cellules hépatiques (principal tissu d'expression du gène *IGF1*) et montré qu'il s'agissait d'un T-DMR (voir article 1). Nous n'avons toutefois pas pu relier la méthylation mesurée dans le sang à celle mesurée dans le foie, compte tenu du petit nombre d'individus pour lesquels nous disposions de ces 2 tissus.

Il nous semble important de rappeler ici les propriétés des cellules sanguines afin de pouvoir interpréter correctement nos résultats. Les leucocytes sont des cellules assez hétérogènes, composées de 40-30% de cellules mononucléaires (monocytes : lymphocytes B, T et NK (Natural Killer)) et 60-70% polynucléaires (granulocytes : basophile, neutrophile, éosinophile). La méthylation mesurée dans le sang total est une moyenne de l'ensemble des cellules citées précédemment. L'identification des DMRs en comparant le profil de méthylation des leucocytes d'un groupe de patients et de témoins est la méthode la plus répandue, dans les études épidémiologique-épigénétique. L'interprétation de ces DMRs doit se faire avec beaucoup de précaution car les différences de méthylation peuvent être dûes simplement à des différences dans la contingence cellulaire. Ce problème a été largement évoqué notamment dans les études épigénétiques des maladies immunitaires et inflammatoires qui affectent le nombre de sous-populations cellulaires des cellules du sang (Reinius et al., 2012).

Reinius et ses collaborateurs (Reinius et al., 2012) ont ainsi apporté une preuve de l'importance de la contingence cellulaire dans l'interprétation des études de méthylation basées sur l'ADN de sang total. Ils ont montré qu'il existait une variation importante des profils de méthylation entre les granulocytes et sept sous-populations cellulaires purifiées : les lymphocytes T ( $CD4^+$ ,  $CD8^+$ ), les monocytes ( $CD14^+$ ), les Natural Killer ( $CD56^+$ ), les lymphocytes B ( $CD19^+$ ), les neutrophiles et les éosinophiles. Ces variations de méthylation sont surtout observées dans les gènes précédemment associés aux troubles du système immunitaires (Reinius et al., 2012).

La description du profil de méthylation de chaque type cellulaire a permis de caractériser un ensemble de DMRs typiques de chaque sous-population servant de « biomarqueurs d'identité ». Ces DMRs sont à la base d'un algorithme permettant de corriger les résultats de méthylation *a posteriori* (Houseman et al., 2015).

Nous avons intégré toutes ces informations pour choisir le tissu le plus approprié pour notre étude. Pour cette raison, nous avons comparé le profil de méthylation des promoteurs du gène *IGF1* dans le sang total, les PBMCs et la population lymphocytaire majoritaire : les T  $CD4^+$ . Nous avons montré que la méthylation de l'ADN était similaire dans le sang total, les PBMCs et les lymphocytes T. La méthylation des promoteurs du gène *IGF1* ne semble donc pas différente entre les sous-populations du sang, elle ne sera donc pas *a priori* affectée par les variations de contingent cellulaire entre les individus. Nous avons donc mené la majorité de nos expériences sur les PBMCs pour plusieurs raisons :

- accessibilité sur un grand nombre d'individus
- permettent d'obtenir une plus grande quantité d'ADN que sur une sous-population triée
- présence de récepteurs à la GH sur les cellules mononucléées, permettant de « voir » la GH

Les cellules sanguines ont une durée de vie variable de quelques jours à des années selon leurs fonctions. Contrairement aux autres cellules différenciées, les cellules sanguines en circulation se renouvellent et se différencient constamment à partir de leurs progéniteurs dans la moelle osseuse. Par conséquent, la reprogrammation épigénétique réalisée pendant l'hématopoïèse donne des cellules sanguines avec des marques épigénétiques nouvellement établies. Les leucocytes en circulation sont un mélange de cellules contenant des cellules qui viennent d'être différenciées de la moelle osseuse et d'autres qui sont en circulation depuis des années créant ainsi un mélange de cellules ayant gardées une mémoire de son environnement et d'autres naïves.

Le profil de méthylation du promoteur P2 dans les PBMCs semble non affecté par l'âge. Dans cette interprétation il faut toutefois garder à l'esprit ce renouvellement « fréquent » des cellules étudiées.

### **Y-a-t-il une transmission des marques de méthylation du promoteur P2 entre les parents et les enfants ?**

La taille présente une forte héritabilité (voir la partie Introduction). Par ailleurs, il apparaît de plus en plus clair que certaines marques épigénétiques peuvent être transmises de génération en génération. Nous pouvons donc interroger quant à la transmission trans-générationnelle de la variation de méthylation des promoteurs d'*IGF1*. Pour la démontrer, il faudrait établir que des modifications de méthylation de l'ADN au niveau des promoteurs d'*IGF1* échappent aux deux vagues de déméthylation et persistent dans les lignées germinales et dans l'embryon précoce. Cela demeure hypothétique et difficilement envisageable chez l'homme. Une des solutions serait probablement là encore d'utiliser le modèle macaque, si nous arrivons à mettre en évidence une variabilité interindividuelle de cette région chez cet animal.

### **Limites de la stratégie gène candidat**

Nous avons évoqué les raisons pour lesquelles nous avons choisi la stratégie du gène candidat dans l'introduction. Cette stratégie nous a permis de répondre à plusieurs questions mais reste une vision incomplète sur une région chromosomique donnée.

Récemment, les associations établies dans les études de méthylation d'un gène candidat ont été critiquées par Shabalin et ses collaborateurs (Shabalin et al., 2015). Il a souligné que les associations statistiques avec des sites dinucléotides spécifiques peuvent être surestimées à cause du nombre réduit de variables utilisées (Shabalin et al., 2015). Selon eux, 30 à 50% des CGs testés peuvent présenter des associations significatives dans une étude gène candidat.

Pour essayer de limiter ces faux positifs, nous avons appliqué des corrections pour les tests multiples en utilisant (Bonferroni) et toutes les P-values sont fortement significatives même après correction ( $10^{-7} < P < 10^{-2}$ ). Nous avons essayé de limiter les probabilités d'avoir un faux positif mais ce risque n'est pas impossible.

### **L'apport de notre étude dans la compréhension de l'efficacité du traitement GH**

Nous avons montré que la méthylation du promoteur P2 expliquait environ 25% de la variabilité de la réponse au traitement par l'hormone de croissance. Nous montrons pour la première fois la



contribution d'un facteur épigénétique dans l'efficacité d'un médicament et nous proposons la méthylation du promoteur P2 du gène IGF1 comme un déterminant majeur de la réponse individuelle au traitement GH entre les enfants ayant une petite taille idiopathique. Dans ces conditions, le profil de méthylation du promoteur P2 pourrait prévoir la réponse à un traitement par l'hormone de croissance chez ces enfants. A notre connaissance, nos observations au niveau du promoteur P2 sont parmi les premiers exemples de pharmaco-epigénomics.

A l'occasion de notre étude, la simplification du protocole du test de génération d'IGF1 nous a permis de faire des constatations cliniques intéressantes. Après une année de traitement les données cliniques des participants à ce test sont déjà disponibles pour un petit effectif (N=20). Nous constatons qu'il y a une tendance à une meilleure réponse au traitement chez les bons répondeurs au test de génération d'IGF1. Sachant que la capacité de ce test à prévoir la réponse thérapeutique par l'hormone de croissance est très controversée (Buckway et al., 2001), le test simplifié semble mieux prévoir la réponse au traitement GH que le test classique.

## REFERENCES BIBLIOGRAPHIQUES

- Adamo, M.L., Ben-Hur, H., Roberts, C.T., and LeRoith, D. (1991). Regulation of start site usage in the leader exons of the rat insulin-like growth factor-I gene by development, fasting, and diabetes. *Mol. Endocrinol. Baltim. Md* 5, 1677–1686.
- Alatzoglou, K.S., and Dattani, M.T. (2010). Genetic causes and treatment of isolated growth hormone deficiency-an update. *Nat. Rev. Endocrinol.* 6, 562–576.
- Alberti, C., Chevenne, D., Mercat, I., Josserand, E., Armoogum-Boizeau, P., Tichet, J., and Léger, J. (2011). Serum concentrations of insulin-like growth factor (IGF)-1 and IGF binding protein-3 (IGFBP-3), IGF-1/IGFBP-3 ratio, and markers of bone turnover: reference values for French children and adolescents and z-score comparability with other references. *Clin. Chem.* 57, 1424–1435.
- Anway, M.D., Cupp, A.S., Uzumcu, M., and Skinner, M.K. (2005). Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 308, 1466–1469.
- Argeson, A.C., Nelson, K.K., and Siracusa, L.D. (1996). Molecular Basis of the Pleiotropic Phenotype of Mice Carrying the Hypervariable Yellow (A(hvy)) Mutation at the Agouti Locus. *Genetics* 142, 557–567.
- Argetsinger, L.S., and Carter-Su, C. (1996). Mechanism of signaling by growth hormone receptor. *Physiol. Rev.* 76, 1089–1107.
- Arias, A.M., and Hayward, P. (2006). Filtering transcriptional noise during development: concepts and mechanisms. *Nat. Rev. Genet.* 7, 34–44.
- Attie, K.M., Carlsson, L.M.S., Rundle, A.C., and Sherman, B.M. (1995). Evidence for partial growth hormone insensitivity among patients with idiopathic short stature. *J. Pediatr.* 127, 244–250.
- Ball, M.P., Li, J.B., Gao, Y., Lee, J.-H., LeProust, E.M., Park, I.-H., Xie, B., Daley, G.Q., and Church, G.M. (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat. Biotechnol.* 27, 361–368.
- Banovich, N.E., Lan, X., McVicker, G., van de Geijn, B., Degner, J.F., Blischak, J.D., Roux, J., Pritchard, J.K., and Gilad, Y. (2014). Methylation QTLs are associated with coordinated changes in transcription factor binding, histone modifications, and gene expression levels. *PLoS Genet.* 10, e1004663.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823–837.
- Bassino JP Dormois JP (2009). Comment tenir compte des erreurs de mesure dans l'estimation de la stature des conscrits français ?
- Bell, J.T., and Saffery, R. (2012). The value of twins in epigenetic epidemiology. *Int. J. Epidemiol.* 41, 140–150.

- Bell, J.T., Pai, A.A., Pickrell, J.K., Gaffney, D.J., Pique-Regi, R., Degner, J.F., Gilad, Y., and Pritchard, J.K. (2011). DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* *12*, R10.
- Bennett-Baker, P.E., Wilkowski, J., and Burke, D.T. (2003). Age-Associated Activation of Epigenetically Repressed Genes in the Mouse. *Genetics* *165*, 2055–2062.
- Bhutani, N., Brady, J.J., Damian, M., Sacco, A., Corbel, S.Y., and Blau, H.M. (2010). Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* *463*, 1042–1047.
- Bick, T., Hochberg, Z., Amit, T., Isaksson, O.G., and Jansson, J.O. (1992). Roles of pulsatility and continuity of growth hormone (GH) administration in the regulation of hepatic GH-receptors, and circulating GH-binding protein and insulin-like growth factor-I. *Endocrinology* *131*, 423–429.
- Bienvvenu, T., and Chelly, J. (2006). Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. *Nat. Rev. Genet.* *7*, 415–426.
- Bird, A.P., and Wolffe, A.P. (1999). Methylation-induced repression--belts, braces, and chromatin. *Cell* *99*, 451–454.
- Bock, C., Walter, J., Paulsen, M., and Lengauer, T. (2008). Inter-individual variation of DNA methylation and its implications for large-scale epigenome mapping. *Nucleic Acids Res.* *36*, e55.
- Bogdanović, O., and Veenstra, G.J.C. (2009). DNA methylation and methyl-CpG binding proteins: developmental requirements and function. *Chromosoma* *118*, 549–565.
- Bogin, B. (1999). *Patterns of Human Growth* (Cambridge University Press). *23*, 599.9–dc21.
- Bogin, B. (2013). Secular changes in childhood, adolescent and adult stature. *Nestlé Nutr. Inst. Workshop Ser.* *71*, 115–126.
- Bonasio, R. (2012). Emerging topics in epigenetics: ants, brains, and noncoding RNAs. *Ann. N. Y. Acad. Sci.* *1260*, 14–23.
- Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular signals of epigenetic states. *Science* *330*, 612–616.
- Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., Forné, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. *Nat. Genet.* *42*, 1093–1100.
- Bourc'his, D., and Bestor, T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* *431*, 96–99.
- Breitling, L.P., Yang, R., Korn, B., Burwinkel, B., and Brenner, H. (2011). Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am. J. Hum. Genet.* *88*, 450–457.
- Bruniquel, D., and Schwartz, R.H. (2003). Selective, stable demethylation of the interleukin-2

gene enhances transcription by an active process. *Nat. Immunol.* *4*, 235–240.

Buckway, C.K., Guevara-Aguirre, J., Pratt, K.L., Burren, C.P., and Rosenfeld, R.G. (2001). The IGF-I generation test revisited: a marker of GH sensitivity. *J. Clin. Endocrinol. Metab.* *86*, 5176–5183.

Burren, C.P., Woods, K.A., Rose, S.J., Tauber, M., Price, D.A., Heinrich, U., Gilli, G., Razzaghy-Azar, M., Al-Ashwal, A., Crock, P.A., et al. (2001). Clinical and endocrine characteristics in atypical and classical growth hormone insensitivity syndrome. *Horm. Res.* *55*, 125–130.

Callewaert, F., Sinnesael, M., Gielen, E., Boonen, S., and Vanderschueren, D. (2010). Skeletal sexual dimorphism: relative contribution of sex steroids, GH-IGF1, and mechanical loading. *J. Endocrinol.* *207*, 127–134.

Campanero, M.R., Armstrong, M.I., and Flemington, E.K. (2000). CpG methylation as a mechanism for the regulation of E2F activity. *Proc. Natl. Acad. Sci. U. S. A.* *97*, 6481–6486.

Cashmore, A.R. (2010). The Lucretian swerve: The biological basis of human behavior and the criminal justice system. *Proc. Natl. Acad. Sci.* *107*, 4499–4504.

Cavalli-Sforza, L.L., Menozzi, P., and Piazza, A. (1994). *The History and Geography of Human Genes* (Princeton University Press).

Cavelaars, A.E., Kunst, A.E., Geurts, J.J., Crialesi, R., Grötvedt, L., Helmert, U., Lahelma, E., Lundberg, O., Mielck, A., Rasmussen, N.K., et al. (2000). Persistent variations in average height between countries and between socio-economic groups: an overview of 10 European countries. *Ann. Hum. Biol.* *27*, 407–421.

Chamla, M.C. (1983). L'évolution récente de la stature en Europe occidentale (Période 1960-1980). *Bull. Mém. Société Anthropol. Paris* *10*, 195–224.

Cheng, X., and Blumenthal, R.M. (2008). Mammalian DNA methyltransferases: a structural perspective. *Struct. Lond. Engl.* *1993* *16*, 341–350.

Chia, D.J., and Rotwein, P. (2010). Defining the epigenetic actions of growth hormone: acute chromatin changes accompany GH-activated gene transcription. *Mol. Endocrinol. Baltim. Md* *24*, 2038–2049.

Chia, D.J., Varco-Merth, B., and Rotwein, P. (2010). Dispersed Chromosomal Stat5b-binding elements mediate growth hormone-activated insulin-like growth factor-I gene transcription. *J. Biol. Chem.* *285*, 17636–17647.

Chotalia, M., Smallwood, S.A., Ruf, N., Dawson, C., Lucifero, D., Frontera, M., James, K., Dean, W., and Kelsey, G. (2009). Transcription is required for establishment of germline methylation marks at imprinted genes. *Genes Dev.* *23*, 105–117.

Ciccarone, F., Klinger, F.G., Catizone, A., Calabrese, R., Zampieri, M., Bacalini, M.G., De Felici, M., and Caiafa, P. (2012). Poly(ADP-ribosyl)ation Acts in the DNA Demethylation of Mouse Primordial Germ Cells Also with DNA Damage-Independent Roles. *PLoS ONE* *7*, e46927.

Cicccone, D.N., Su, H., Hevi, S., Gay, F., Lei, H., Bajko, J., Xu, G., Li, E., and Chen, T. (2009). KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints.

Nature 461, 415–418.

Clark, R.G., Mortensen, D.L., Carlsson, L.M., Spencer, S.A., McKay, P., Mulkerrin, M., Moore, J., and Cunningham, B.C. (1996). Recombinant human growth hormone (GH)-binding protein enhances the growth-promoting activity of human GH in the rat. *Endocrinology* 137, 4308–4315.

Cohen, P., Germak, J., Rogol, A.D., Weng, W., Kappelgaard, A.-M., Rosenfeld, R.G., and American Norditropin Study Group (2010). Variable degree of growth hormone (GH) and insulin-like growth factor (IGF) sensitivity in children with idiopathic short stature compared with GH-deficient patients: evidence from an IGF-based dosing study of short children. *J. Clin. Endocrinol. Metab.* 95, 2089–2098.

Dacunha-Castelle D (1996). *Chemins de l'aléatoire : le hasard et le risque dans la société moderne.*

Daughaday, W.H. (1969). Growth hormone assay in acromegaly, gigantism, dwarfism and hypopituitarism. *Postgrad. Med.* 46, 84–91.

David, A., Hwa, V., Metherell, L.A., Netchine, I., Camacho-Hübner, C., Clark, A.J.L., Rosenfeld, R.G., and Savage, M.O. (2011). Evidence for a continuum of genetic, phenotypic, and biochemical abnormalities in children with growth hormone insensitivity. *Endocr. Rev.* 32, 472–497.

Dedeurwaerder, S., Defrance, M., Calonne, E., Denis, H., Sotiriou, C., and Fuks, F. (2011). Evaluation of the Infinium Methylation 450K technology. *Epigenomics* 3, 771–784.

Delgado-Calle, J., Fernández, A.F., Sainz, J., Zarrabeitia, M.T., Sañudo, C., García-Renedo, R., Pérez-Núñez, M.I., García-Ibarbia, C., Fraga, M.F., and Riancho, J.A. (2013). Genome-wide profiling of bone reveals differentially methylated regions in osteoporosis and osteoarthritis. *Arthritis Rheum.* 65, 197–205.

Denis, H., Ndlovu, 'Matladi N., and Fuks, F. (2011). Regulation of mammalian DNA methyltransferases: a route to new mechanisms. *EMBO Rep.* 12, 647–656.

Denley, A., Wang, C.C., McNeil, K.A., Walenkamp, M.J.E., van Duyvenvoorde, H., Wit, J.M., Wallace, J.C., Norton, R.S., Karperien, M., and Forbes, B.E. (2005). Structural and functional characteristics of the Val44Met insulin-like growth factor I missense mutation: correlation with effects on growth and development. *Mol. Endocrinol. Baltim. Md* 19, 711–721.

Desgagné, V., Hivert, M.-F., St-Pierre, J., Guay, S.-P., Baillargeon, J.-P., Perron, P., Gaudet, D., Brisson, D., and Bouchard, L. (2014). Epigenetic dysregulation of the IGF system in placenta of newborns exposed to maternal impaired glucose tolerance. *Epigenomics* 6, 193–207.

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J. Biol. Chem.* 285, 26114–26120.

Diniz, E.T., Jorge, A.A.L., Arnhold, I.J.P., Rosenbloom, A.L., and Bandeira, F. (2008). [Novel nonsense mutation (p.Y113X) in the human growth hormone receptor gene in a Brazilian patient with Laron syndrome]. *Arq. Bras. Endocrinol. Metabol.* 52, 1264–1271.

- Dodge, J.E., Okano, M., Dick, F., Tsujimoto, N., Chen, T., Wang, S., Ueda, Y., Dyson, N., and Li, E. (2005). Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. *J. Biol. Chem.* **280**, 17986–17991.
- Domené, H.M., Bengolea, S.V., Martínez, A.S., Ropelato, M.G., Pennisi, P., Scaglia, P., Heinrich, J.J., and Jasper, H.G. (2004). Deficiency of the circulating insulin-like growth factor system associated with inactivation of the acid-labile subunit gene. *N. Engl. J. Med.* **350**, 570–577.
- Domené, H.M., Hwa, V., Argente, J., Wit, J.M., Wit, J.M., Camacho-Hübner, C., Jasper, H.G., Pozo, J., van Duyvenvoorde, H.A., Yakar, S., et al. (2009). Human acid-labile subunit deficiency: clinical, endocrine and metabolic consequences. *Horm. Res.* **72**, 129–141.
- Dominguez-Salas, P., Moore, S.E., Baker, M.S., Bergen, A.W., Cox, S.E., Dyer, R.A., Fulford, A.J., Guan, Y., Laritsky, E., Silver, M.J., et al. (2014). Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5**, 3746.
- Doskočil, J., and Šorm, F. (1962). Distribution of 5-methylcytosine in pyrimidine sequences of deoxyribonucleic acids. *Biochim. Biophys. Acta BBA - Spec. Sect. Nucleic Acids Relat. Subj.* **55**, 953–959.
- Estève, P.-O., Chin, H.G., Smallwood, A., Feehery, G.R., Gangisetty, O., Karpf, A.R., Carey, M.F., and Pradhan, S. (2006). Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev.* **20**, 3089–3103.
- Eva Jablonka, and Marion J. Lamb (2005). *Evolution in Four Dimensions Genetic, Epigenetic, Behavioral, and symbolic Variation in the History of the life.*
- Falconer DS, and Mackay Trudy F.C (1996a). Variance. In *Introduction to Quantitative Genetics*, pp. 122–143.
- Falconer DS, and Mackay Trudy F.C (1996b). *Introduction to quantitative genetics chapitre 20 : Metric characters under naturel selection 335-351 Longman.*
- Falconer DS, and Mackay Trudy F.C (1996c). chapitre 21 *Quantitative trait loci. In Introduction to Quantitative Genetics*, pp. 356–375.
- Falconer DS, and Mackay Trudy F.C (1996d). In *Introduction to Quantitative Genetics Chapitre 10 : Heritability 160-177 Longman,*.
- Fan, Y., Menon, R.K., Cohen, P., Hwang, D., Clemens, T., DiGirolamo, D.J., Kopchick, J.J., Le Roith, D., Trucco, M., and Sperling, M.A. (2009). Liver-specific deletion of the growth hormone receptor reveals essential role of growth hormone signaling in hepatic lipid metabolism. *J. Biol. Chem.* **284**, 19937–19944.
- Fang, P., Schwartz, I.D., Johnson, B.D., Derr, M.A., Roberts, C.T., Hwa, V., and Rosenfeld, R.G. (2009). Familial short stature caused by haploinsufficiency of the insulin-like growth factor I receptor due to nonsense-mediated messenger ribonucleic acid decay. *J. Clin. Endocrinol. Metab.* **94**, 1740–1747.

- Farrall, M. (2004). Quantitative genetic variation: a post-modern view. *Hum. Mol. Genet. 13 Spec No 1*, R1–R7.
- Feinberg, A.P. (2004). The epigenetics of cancer etiology. *Semin. Cancer Biol. 14*, 427–432.
- Feinberg, A.P. (2007). Phenotypic plasticity and the epigenetics of human disease. *Nature 447*, 433–440.
- Feinberg, A.P., and Irizarry, R.A. (2010). Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc. Natl. Acad. Sci. 107*, 1757–1764.
- Feinberg, A.P., Irizarry, R.A., Fradin, D., Aryee, M.J., Murakami, P., Aspelund, T., Eiriksdottir, G., Harris, T.B., Launer, L., Gudnason, V., et al. (2010). Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci. Transl. Med. 2*, 49ra67.
- Filipowicz, W., Jaskiewicz, L., Kolb, F.A., and Pillai, R.S. (2005). Post-transcriptional gene silencing by siRNAs and miRNAs. *Curr. Opin. Struct. Biol. 15*, 331–341.
- Finer, S., Holland, M.L., Nanty, L., and Rakyán, V.K. (2011). The hunt for the epiallele. *Environ. Mol. Mutagen. 52*, 1–11.
- Fisher, R.A. (1919). XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Earth Environ. Sci. Trans. R. Soc. Edinb. 52*, 399–433.
- Fitzpatrick, G.V., Soloway, P.D., and Higgins, M.J. (2002). Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nat. Genet. 32*, 426–431.
- Fofanova-Gambetti, O.V., Hwa, V., Kirsch, S., Pihoker, C., Chiu, H.K., Höglér, W., Cohen, L.E., Jacobsen, C., Derr, M.A., and Rosenfeld, R.G. (2009). Three novel IGFALS gene mutations resulting in total ALS and severe circulating IGF-I/IGFBP-3 deficiency in children of different ethnic origins. *Horm. Res. 71*, 100–110.
- Ford, E., and Thanos, D. (2009). Time's up: Bursting out of Transcription. *Cell 138*, 430–432.
- Foret, S., Kucharski, R., Pellegrini, M., Feng, S., Jacobsen, S.E., Robinson, G.E., and Maleszka, R. (2012). DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proc. Natl. Acad. Sci. 109*, 4968–4973.
- Foyt, H.L., LeRoith, D., and Roberts, C.T. (1991). Differential association of insulin-like growth factor I mRNA variants with polysomes in vivo. *J. Biol. Chem. 266*, 7300–7305.
- Fradin, D., Le Fur, S., Mille, C., Naoui, N., Groves, C., Zelenika, D., McCarthy, M.I., Lathrop, M., and Bougnères, P. (2012). Association of the CpG methylation pattern of the proximal insulin gene promoter with type 1 diabetes. *PLoS One 7*, e36278.
- Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suñer, D., Cigudosa, J.C., Urioste, M., Benitez, J., et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. U. S. A. 102*, 10604–10609.
- Galton, F. (1876). On the Height and Weight of Boys Aged 14, in Town and Country Public Schools. *J. Anthropol. Inst. G. B. Irel. 5*, 174.

- Galton, F. (1886). Regression Towards Mediocrity in Hereditary Stature. *J. Anthropol. Inst. G. B. Irel.* *15*, 246–263.
- Galton, F., and Dickson, J.D.H. (1886). Family Likeness in Stature. *Proc. R. Soc. Lond.* *40*, 42–73.
- Gauguin, L., Delaine, C., Alvino, C.L., McNeil, K.A., Wallace, J.C., Forbes, B.E., and De Meyts, P. (2008). Alanine scanning of a putative receptor binding surface of insulin-like growth factor-I. *J. Biol. Chem.* *283*, 20821–20829.
- Gavalli-Sforza L.L, and Menozzi P (1994). PYGMIES in Chapter 3: Africa. In *The History and Geography of Human Genes*, (Princeton University Press), pp. 177–180.
- Gibbs, J.R., van der Brug, M.P., Hernandez, D.G., Traynor, B.J., Nalls, M.A., Lai, S.-L., Arepalli, S., Dillman, A., Rafferty, I.P., Troncoso, J., et al. (2010). Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* *6*, e1000952.
- Goddard, A.D., Dowd, P., Chernausek, S., Geffner, M., Gertner, J., Hintz, R., Hopwood, N., Kaplan, S., Plotnick, L., Rogol, A., et al. (1997). Partial growth-hormone insensitivity: the role of growth-hormone receptor mutations in idiopathic short stature. *J. Pediatr.* *131*, S51–S55.
- Goldberg, A.D., Allis, C.D., and Bernstein, E. (2007). Epigenetics: A Landscape Takes Shape. *Cell* *128*, 635–638.
- Greenhalgh, C.J., Rico-Bautista, E., Lorentzon, M., Thaus, A.L., Morgan, P.O., Willson, T.A., Zervoudakis, P., Metcalf, D., Street, I., Nicola, N.A., et al. (2005). SOCS2 negatively regulates growth hormone action in vitro and in vivo. *J. Clin. Invest.* *115*, 397–406.
- Griffiths A.J.F, Miller J.H, Suzuki D.T, Lewontin R.C, and Gelbart W.M (1999). *An Introduction to Genetic Analysis*.
- Gudbjartsson, D.F., Walters, G.B., Thorleifsson, G., Stefansson, H., Halldorsson, B.V., Zusmanovich, P., Sulem, P., Thorlacius, S., Gylfason, A., Steinberg, S., et al. (2008). Many sequence variants affecting diversity of adult human height. *Nat. Genet.* *40*, 609–615.
- Guibert, S., and Weber, M. (2013). Functions of DNA methylation and hydroxymethylation in mammalian development. *Curr. Top. Dev. Biol.* *104*, 47–83.
- Guibert, S., Forné, T., and Weber, M. (2009). Dynamic regulation of DNA methylation during mammalian development. *Epigenomics* *1*, 81–98.
- Haberland, M., Montgomery, R.L., and Olson, E.N. (2009). The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat. Rev. Genet.* *10*, 32–42.
- Haisa, M. (2013). The type 1 insulin-like growth factor receptor signalling system and targeted tyrosine kinase inhibition in cancer. *J. Int. Med. Res.* *41*, 253–264.
- Hall, L.J., Kajimoto, Y., Bichell, D., Kim, S.W., James, P.L., Counts, D., Nixon, L.J., Tobin, G., and Rotwein, P. (1992). Functional analysis of the rat insulin-like growth factor I gene and identification of an IGF-I gene promoter. *DNA Cell Biol.* *11*, 301–313.



- Hallgrímsson, B., Willmore, K., and Hall, B.K. (2002). Canalization, developmental stability, and morphological integration in primate limbs. *Am. J. Phys. Anthropol. Suppl* 35, 131–158.
- Hao, S., and Baltimore, D. (2009). The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. *Nat. Immunol.* 10, 281–288.
- Harris, R.A., Nagy-Szakal, D., and Kellermayer, R. (2013). Human metastable epiallele candidates link to common disorders. *Epigenetics Off. J. DNA Methylation Soc.* 8, 157–163.
- He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333, 1303–1307.
- Hemberger, M., Dean, W., and Reik, W. (2009). Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington’s canal. *Nat. Rev. Mol. Cell Biol.* 10, 526–537.
- Hepler, J.E., Van Wyk, J.J., and Lund, P.K. (1990). Different half-lives of insulin-like growth factor I mRNAs that differ in length of 3’ untranslated sequence. *Endocrinology* 127, 1550–1552.
- Hermanussen, M., and Cole, J. (2003). The calculation of target height reconsidered. *Horm. Res.* 59, 180–183.
- Holliday, R. (1990). Mechanisms for the control of gene activity during development. *Biol. Rev. Camb. Philos. Soc.* 65, 431–471.
- Holliday, R. (1994). Epigenetics: an overview. *Dev. Genet.* 15, 453–457.
- ’t Hooft, G., Susskind, L., Witten, E., Fukugita, M., Randall, L., Smolin, L., Stachel, J., Rovelli, C., Ellis, G., Weinberg, S., et al. (2005). A theory of everything? *Nature* 433, 257–259.
- Houseman, E.A., Kelsey, K.T., Wiencke, J.K., and Marsit, C.J. (2015). Cell-composition effects in the analysis of DNA methylation array data: a mathematical perspective. *BMC Bioinformatics* 16, 95.
- Huang, J., Wang, Y., Guo, Y., and Sun, S. (2010). Down-regulated microRNA-152 induces aberrant DNA methylation in hepatitis B virus-related hepatocellular carcinoma by targeting DNA methyltransferase 1. *Hepatol. Baltim. Md* 52, 60–70.
- Iguchi-Ariga, S.M., and Ariga, H. (1989). Concerted mechanism of DNA replication and transcription. *Cell Struct. Funct.* 14, 649–651.
- Illingworth, R., Kerr, A., DeSousa, D., Jørgensen, H., Ellis, P., Stalker, J., Jackson, D., Clee, C., Plumb, R., Rogers, J., et al. (2008). A Novel CpG Island Set Identifies Tissue-Specific Methylation at Developmental Gene Loci. *PLoS Biol* 6, e22.
- Inagaki, K., Tiulpakov, A., Rubtsov, P., Sverdlova, P., Peterkova, V., Yakar, S., Terekhov, S., and LeRoith, D. (2007). A familial insulin-like growth factor-I receptor mutant leads to short stature: clinical and biochemical characterization. *J. Clin. Endocrinol. Metab.* 92, 1542–1548.
- Isaksson, O.G., Edén, S., and Jansson, J.O. (1985). Mode of action of pituitary growth hormone on target cells. *Annu. Rev. Physiol.* 47, 483–499.

Isaksson, O.G., Lindahl, A., Nilsson, A., and Isgaard, J. (1987). Mechanism of the stimulatory effect of growth hormone on longitudinal bone growth. *Endocr. Rev.* *8*, 426–438.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* *333*, 1300–1303.

Jablonka, E., and Lamb, M.J. (2002). The changing concept of epigenetics. *Ann. N. Y. Acad. Sci.* *981*, 82–96.

Jacques Monod (1970). *Le Hasard et la Nécessité*.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* *33 Suppl*, 245–254.

Jansen, E., Steenbergh, P.H., LeRoith, D., Roberts, C.T., and Sussenbach, J.S. (1991). Identification of multiple transcription start sites in the human insulin-like growth factor-I gene. *Mol. Cell. Endocrinol.* *78*, 115–125.

Jansen, E., Steenbergh, P.H., van Schaik, F.M., and Sussenbach, J.S. (1992). The human IGF-I gene contains two cell type-specifically regulated promoters. *Biochem. Biophys. Res. Commun.* *187*, 1219–1226.

Jeffries, M.A., Dozmorov, M., Tang, Y., Merrill, J.T., Wren, J.D., and Sawalha, A.H. (2011). Genome-wide DNA methylation patterns in CD4<sup>+</sup> T cells from patients with systemic lupus erythematosus. *Epigenetics Off. J. DNA Methylation Soc.* *6*, 593–601.

Jelinic, P., Stehle, J.-C., and Shaw, P. (2006). The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in H19 imprinting control region methylation. *PLoS Biol.* *4*, e355.

Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A., and Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* *449*, 248–251.

Jin, S.-G., Kadam, S., and Pfeifer, G.P. (2010). Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res.* *38*, e125.

Jirtle, R.L., and Skinner, M.K. (2007). Environmental epigenomics and disease susceptibility. *Nat. Rev. Genet.* *8*, 253–262.

Jobling M.A, Hurles M.E, and Tyler-Smith C *Human Evolutionary Genetics : Origins, Peoples and Disease*.

Johannes, F., Colot, V., and Jansen, R.C. (2008). Epigenome dynamics: a quantitative genetics perspective. *Nat. Rev. Genet.* *9*, 883–890.

Jones, J.I., and Clemmons, D.R. (1995). Insulin-like growth factors and their binding proteins: biological actions. *Endocr. Rev.* *16*, 3–34.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to

repress transcription. *Nat. Genet.* *19*, 187–191.

Juul, A., Bang, P., Hertel, N.T., Main, K., Dalgaard, P., Jørgensen, K., Müller, J., Hall, K., and Skakkebaek, N.E. (1994). Serum insulin-like growth factor-I in 1030 healthy children, adolescents, and adults: relation to age, sex, stage of puberty, testicular size, and body mass index. *J. Clin. Endocrinol. Metab.* *78*, 744–752.

Kakutani, T., Munakata, K., Richards, E.J., and Hirochika, H. (1999). Meiotically and Mitotically Stable Inheritance of DNA Hypomethylation Induced by *ddm1* Mutation of *Arabidopsis thaliana*. *Genetics* *151*, 831–838.

Kaneda, M., Sado, T., Hata, K., Okano, M., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Role of De Novo DNA Methyltransferases in Initiation of Genomic Imprinting and X-Chromosome Inactivation. *Cold Spring Harb. Symp. Quant. Biol.* *69*, 125–130.

Kangaspeska, S., Stride, B., Métivier, R., Polycarpou-Schwarz, M., Ibberson, D., Carmouche, R.P., Benes, V., Gannon, F., and Reid, G. (2008). Transient cyclical methylation of promoter DNA. *Nature* *452*, 112–115.

Kenneth M Weiss (1993). Genes and phenotypes in populations. In *Genetic Variation and Human Disease Principles and Evolutionary Approaches*, (CAMBRIDGE UNIVERSITY PRESS), pp. 52–62.

Kerkel, K., Spadola, A., Yuan, E., Kosek, J., Jiang, L., Hod, E., Li, K., Murty, V.V., Schupf, N., Vilain, E., et al. (2008). Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. *Nat. Genet.* *40*, 904–908.

Kimura M (1983). *The neutral Theory of Molecular Evolution* (University Press, Cambridge).

Kofoed, E.M., Hwa, V., Little, B., Woods, K.A., Buckway, C.K., Tsubaki, J., Pratt, K.L., Bezrodnik, L., Jasper, H., Tepper, A., et al. (2003). Growth hormone insensitivity associated with a *STAT5b* mutation. *N. Engl. J. Med.* *349*, 1139–1147.

Kokura, K., Kaul, S.C., Wadhwa, R., Nomura, T., Khan, M.M., Shinagawa, T., Yasukawa, T., Colmenares, C., and Ishii, S. (2001). The Ski protein family is required for MeCP2-mediated transcriptional repression. *J. Biol. Chem.* *276*, 34115–34121.

Koutsilieris, M., Mitsiades, C., Lembessis, P., and Sourla, A. (2000). Cancer and bone repair mechanism: clinical applications for hormone refractory prostate cancer. *J. Musculoskelet. Neuronal Interact.* *1*, 15–17.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell* *128*, 693–705.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base, 5-hydroxymethylcytosine is present in brain and enriched in Purkinje neurons. *Science* *324*, 929–930.

Kuström, B., Lundberg, E., Jonsson, B., Albertsson-Wikland, K., and study group (2014). IGF-1 and growth response to adult height in a randomized GH treatment trial in short non-GH-deficient children. *J. Clin. Endocrinol. Metab.* *99*, 2917–2924.

Kruis, T., Klammt, J., Galli-Tsinopoulou, A., Wallborn, T., Schlicke, M., Müller, E., Kratzsch, J., Körner, A., Odeh, R., Kiess, W., et al. (2010). Heterozygous mutation within a kinase-

conserved motif of the insulin-like growth factor I receptor causes intrauterine and postnatal growth retardation. *J. Clin. Endocrinol. Metab.* *95*, 1137–1142.

Kuroda, A., Rauch, T.A., Todorov, I., Ku, H.T., Al-Abdullah, I.H., Kandeel, F., Mullen, Y., Pfeifer, G.P., and Ferreri, K. (2009). Insulin Gene Expression Is Regulated by DNA Methylation. *PLoS ONE* *4*, e6953.

Lange, K. (1978). Central limit theorems of pedigrees. *J. Math. Biol.* *6*, 59–66.

Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S., et al. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* *467*, 832–838.

Laron, Z. (2001). Insulin-like growth factor 1 (IGF-1): a growth hormone. *Mol. Pathol. MP* *54*, 311–316.

Laron, Z. (2004). Laron syndrome (primary growth hormone resistance or insensitivity): the personal experience 1958-2003. *J. Clin. Endocrinol. Metab.* *89*, 1031–1044.

Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* *11*, 204–220.

Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* *71*, 865–873.

LeRoith, D., and Roberts, C.T. (1991). Insulin-like growth factor I (IGF-I): a molecular basis for endocrine versus local action? *Mol. Cell. Endocrinol.* *77*, C57–C61.

Lettre, G. (2011). Recent progress in the study of the genetics of height. *Hum. Genet.* *129*, 465–472.

Lettre, G., Butler, J.L., Ardlie, K.G., and Hirschhorn, J.N. (2007). Common genetic variation in eight genes of the GH/IGF1 axis does not contribute to adult height variation. *Hum. Genet.* *122*, 129–139.

Lettre, G., Jackson, A.U., Gieger, C., Schumacher, F.R., Berndt, S.I., Sanna, S., Eyheramendy, S., Voight, B.F., Butler, J.L., Guiducci, C., et al. (2008). Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat. Genet.* *40*, 584–591.

List, E.O., Berryman, D.E., Funk, K., Jara, A., Kelder, B., Wang, F., Stout, M.B., Zhi, X., Sun, L., White, T.A., et al. (2014). Liver-specific GH receptor gene-disrupted (LiGHRKO) mice have decreased endocrine IGF-I, increased local IGF-I, and altered body size, body composition, and adipokine profiles. *Endocrinology* *155*, 1793–1805.

Lister, R., Pelizzola, M., Downen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* *462*, 315–322.

Liu, J.-L., and LeRoith, D. (1999). Insulin-Like Growth Factor I Is Essential for Postnatal Growth in Response to Growth Hormone. *Endocrinology* *140*, 5178–5184.

Lowe, W.L., Roberts, C.T., Lasky, S.R., and LeRoith, D. (1987). Differential expression of

alternative 5' untranslated regions in mRNAs encoding rat insulin-like growth factor I. *Proc. Natl. Acad. Sci. U. S. A.* *84*, 8946–8950.

Lucifero, D., Mann, M.R.W., Bartolomei, M.S., and Trasler, J.M. (2004). Gene-specific timing and epigenetic memory in oocyte imprinting. *Hum. Mol. Genet.* *13*, 839–849.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* *389*, 251–260.

Lund, P.K., Hoyt, E.C., and Van Wyk, J.J. (1989). The size heterogeneity of rat insulin-like growth factor-I mRNAs is due primarily to differences in the length of 3'-untranslated sequence. *Mol. Endocrinol. Baltim. Md* *3*, 2054–2061.

Luo, Z.C., Albertsson-Wikland, K., and Karlberg, J. (1998). Target height as predicted by parental heights in a population-based study. *Pediatr. Res.* *44*, 563–571.

Magee, B.A., Shooter, G.K., Wallace, J.C., and Francis, G.L. (1999). Insulin-like growth factor I and its binding proteins: a study of the binding interface using B-domain analogues. *Biochemistry (Mosc.)* *38*, 15863–15870.

Mahmoudabady, M., Mathieu, M., Touihri, K., Hadad, I., Da Costa, A.M., Naeije, R., and Mc Entee, K. (2009). Cardiac insulin-like growth factor-1 and cyclins gene expression in canine models of ischemic or overpacing cardiomyopathy. *BMC Cardiovasc. Disord.* *9*, 49.

Maiti, A., and Drohat, A.C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J. Biol. Chem.* *286*, 35334–35338.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex diseases. *Nature* *461*, 747–753.

Martin, L. (1958). Private medical care; advantages we enjoy under our present system. *J. Kans. Med. Soc.* *59*, 339–340.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Demethylation of the zygotic paternal genome. *Nature* *403*, 501–502.

Le May, N., Fradin, D., Iltis, I., Bougnères, P., and Egly, J.-M. (2012). XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes. *Mol. Cell* *47*, 622–632.

Meaburn, E.L., Schalkwyk, L.C., and Mill, J. (2010). Allele-specific methylation in the human genome: implications for genetic studies of complex disease. *Epigenetics Off. J. DNA Methylation Soc.* *5*, 578–582.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* *454*, 766–770.

Melas, P.A., Wei, Y., Wong, C.C.Y., Sjöholm, L.K., Åberg, E., Mill, J., Schalling, M., Forsell, Y., and Lavebratt, C. (2013). Genetic and epigenetic associations of MAOA and NR3C1 with

depression and childhood adversities. *Int. J. Neuropsychopharmacol. Off. Sci. J. Coll. Int. Neuropsychopharmacol. CINP* *16*, 1513–1528.

Metherell, L.A., Akker, S.A., Munroe, P.B., Rose, S.J., Caulfield, M., Savage, M.O., Chew, S.L., and Clark, A.J. (2001). Pseudoexon activation as a novel mechanism for disease resulting in atypical growth-hormone insensitivity. *Am. J. Hum. Genet.* *69*, 641–646.

Métivier, R., Gallais, R., Tiffoche, C., Le Péron, C., Jurkowska, R.Z., Carmouche, R.P., Ibberson, D., Barath, P., Demay, F., Reid, G., et al. (2008). Cyclical DNA methylation of a transcriptionally active promoter. *Nature* *452*, 45–50.

Michels, K.B., Binder, A.M., Dedeurwaerder, S., Epstein, C.B., Grealley, J.M., Gut, I., Houseman, E.A., Izzi, B., Kelsey, K.T., Meissner, A., et al. (2013). Recommendations for the design and analysis of epigenome-wide association studies. *Nat. Methods* *10*, 949–955.

Mill, J., Dempster, E., Caspi, A., Williams, B., Moffitt, T., and Craig, I. (2006). Evidence for monozygotic twin (MZ) discordance in methylation level at two CpG sites in the promoter region of the catechol-O-methyltransferase (COMT) gene. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet. Off. Publ. Int. Soc. Psychiatr. Genet.* *141B*, 421–425.

Mill, J., Tang, T., Kaminsky, Z., Khare, T., Yazdanpanah, S., Bouchard, L., Jia, P., Assadzadeh, A., Flanagan, J., Schumacher, A., et al. (2008). Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am. J. Hum. Genet.* *82*, 696–711.

Mohandas, T., Sparkes, R.S., and Shapiro, L.J. (1981). Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science* *211*, 393–396.

Murgatroyd, C., Patchev, A.V., Wu, Y., Micale, V., Bockmühl, Y., Fischer, D., Holsboer, F., Wotjak, C.T., Almeida, O.F.X., and Spengler, D. (2009). Dynamic DNA methylation programs persistent adverse effects of early-life stress. *Nat. Neurosci.* *12*, 1559–1566.

Myer, V.E., Fan, X.C., and Steitz, J.A. (1997). Identification of HuR as a protein implicated in AUUUA-mediated mRNA decay. *EMBO J.* *16*, 2130–2139.

Nan, X., Cross, S., and Bird, A. (1998). Gene silencing by methyl-CpG-binding proteins. *Novartis Found. Symp.* *214*, 6–16; discussion 16–21, 46–50.

Netchine, I., Azzi, S., Houang, M., Seurin, D., Perin, L., Ricort, J.-M., Daubas, C., Legay, C., Mester, J., Herich, R., et al. (2009). Partial primary deficiency of insulin-like growth factor (IGF)-I activity associated with IGF1 mutation demonstrates its critical role in growth and brain development. *J. Clin. Endocrinol. Metab.* *94*, 3913–3921.

Nicholas Wright Gilham (2001). *A life of sir Francis GALTON From African Exploration to the Birth of Eugenics.*

Nolten, L.A., van Schaik, F.M., Steenbergh, P.H., and Sussenbach, J.S. (1994). Expression of the insulin-like growth factor I gene is stimulated by the liver-enriched transcription factors C/EBP alpha and LAP. *Mol. Endocrinol. Baltim. Md* *8*, 1636–1645.

Nolten, L.A., Steenbergh, P.H., and Sussenbach, J.S. (1996). The hepatocyte nuclear factor 3beta stimulates the transcription of the human insulin-like growth factor I gene in a direct

and indirect manner. *J. Biol. Chem.* **271**, 31846–31854.

Oberbauer, A.M. (2013). The Regulation of IGF-1 Gene Transcription and Splicing during Development and Aging. *Front. Endocrinol.* **4**, 39.

Oda, M., Yamagiwa, A., Yamamoto, S., Nakayama, T., Tsumura, A., Sasaki, H., Nakao, K., Li, E., and Okano, M. (2006). DNA methylation regulates long-range gene silencing of an X-linked homeobox gene cluster in a lineage-specific manner. *Genes Dev.* **20**, 3382–3394.

Okada, Y., Kamatani, Y., Takahashi, A., Matsuda, K., Hosono, N., Ohmiya, H., Daigo, Y., Yamamoto, K., Kubo, M., Nakamura, Y., et al. (2010). A genome-wide association study in 19 633 Japanese subjects identified LHX3-QSOX2 and IGF1 as adult height loci. *Hum. Mol. Genet.* **19**, 2303–2312.

Okano, M., Xie, S., and Li, E. (1998). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res.* **26**, 2536–2540.

Orr, H.A. (1998). The Population Genetics of Adaptation: The Distribution of Factors Fixed during Adaptive Evolution. *Evolution* **52**, 935–949.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol. CB* **10**, 475–478.

Palles, C., Johnson, N., Coupland, B., Taylor, C., Carvajal, J., Holly, J., Fentiman, I.S., Silva, I. dos S., Ashworth, A., Peto, J., et al. (2008). Identification of genetic variants that influence circulating IGF1 levels: a targeted search strategy. *Hum. Mol. Genet.* **17**, 1457–1464.

Pearson, K. (1895). Contributions to the Mathematical Theory of Evolution. III. Regression, Heredity, and Panmixia. **59**.

Petronis, A. (2010). Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature* **465**, 721–727.

Piazza, A. (1992). L’histoire génétique des populations : est-ce-que les races existent ? *Mots* **33**, 151–159.

Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. *Nat. Biotechnol.* **28**, 1057–1068.

Posadas, D.M., and Carthew, R.W. (2014). MicroRNAs and their roles in developmental canalization. *Curr. Opin. Genet. Dev.* **27**, 1–6.

Rakyan, V.K., Blewitt, M.E., Druker, R., Preis, J.I., and Whitelaw, E. (2002). Metastable epialleles in mammals. *Trends Genet. TIG* **18**, 348–351.

Rakyan, V.K., Hildmann, T., Novik, K.L., Lewin, J., Tost, J., Cox, A.V., Andrews, T.D., Howe, K.L., Otto, T., Olek, A., et al. (2004). DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. *PLoS Biol.* **2**, e405.

Rakyan, V.K., Down, T.A., Thorne, N.P., Flicek, P., Kulesha, E., Gräf, S., Tomazou, E.M.,

Bäckdahl, L., Johnson, N., Herberth, M., et al. (2008). An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res.* *18*, 1518–1529.

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* *447*, 425–432.

Reinius, L.E., Acevedo, N., Joerink, M., Pershagen, G., Dahlén, S.-E., Greco, D., Söderhäll, C., Scheynius, A., and Kere, J. (2012). Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One* *7*, e41361.

Richards, E.J. (2006). Inherited epigenetic variation--revisiting soft inheritance. *Nat. Rev. Genet.* *7*, 395–401.

Riggs, A.D. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* *14*, 9–25.

Rinderknecht, E., and Humbel, R.E. (1978). The amino acid sequence of human insulin-like growth factor I and its structural homology with proinsulin. *J. Biol. Chem.* *253*, 2769–2776.

Robertson A (1967). the nature of quantitative variation. In Brink, A. *Heritage from Mendel* (University of Wisconsin, Madison).

Rodgers, J.L., and Nicewander, W.A. (1988). Thirteen Ways to Look at the Correlation Coefficient. *Am. Stat.* *42*, 59–66.

Rogers, I., Metcalfe, C., Gunnell, D., Emmett, P., Dunger, D., Holly, J., and Avon Longitudinal Study of Parents Children Study Team (2006a). Insulin-like growth factor-I and growth in height, leg length, and trunk length between ages 5 and 10 years. *J. Clin. Endocrinol. Metab.* *91*, 2514–2519.

Rogers, I., Metcalfe, C., Gunnell, D., Emmett, P., Dunger, D., Holly, J., and Avon Longitudinal Study of Parents Children Study Team (2006b). Insulin-like growth factor-I and growth in height, leg length, and trunk length between ages 5 and 10 years. *J. Clin. Endocrinol. Metab.* *91*, 2514–2519.

Rosenbloom, A.L. (1999). IGF-I deficiency due to GH receptor deficiency. *Horm. Metab. Res. Horm. Stoffwechselforschung Horm. Métabolisme* *31*, 161–171.

Rosenfeld, R.G., and Bakker, B. (2008). Compliance and persistence in pediatric and adult patients receiving growth hormone therapy. *Endocr. Pract. Off. J. Am. Coll. Endocrinol. Am. Assoc. Clin. Endocrinol.* *14*, 143–154.

Rotwein, P. (2012). Mapping the growth hormone--Stat5b--IGF-I transcriptional circuit. *Trends Endocrinol. Metab. TEM* *23*, 186–193.

Rotwein, P., Naylor, S.L., and Chirgwin, J.M. (1986). Human insulin-related DNA sequences map to chromosomes 2 and 11. *Somat. Cell Mol. Genet.* *12*, 625–631.

Salmon, W.D., and Daughaday, W.H. (1957). A hormonally controlled serum factor which stimulates sulfate incorporation by cartilage in vitro. *J. Lab. Clin. Med.* *49*, 825–836.



- Salmon, W.D., and Daughaday, W.H. (1990). A hormonally controlled serum factor which stimulates sulfate incorporation by cartilage in vitro. 1956. *J. Lab. Clin. Med.* *116*, 408–419.
- Dos Santos, C., Essioux, L., Teinturier, C., Tauber, M., Goffin, V., and Bougnères, P. (2004). A common polymorphism of the growth hormone receptor is associated with increased responsiveness to growth hormone. *Nat. Genet.* *36*, 720–724.
- Schalkwyk, L.C., Meaburn, E.L., Smith, R., Dempster, E.L., Jeffries, A.R., Davies, M.N., Plomin, R., and Mill, J. (2010). Allelic skewing of DNA methylation is widespread across the genome. *Am. J. Hum. Genet.* *86*, 196–212.
- Schmidt, I.M., Jørgensen, M.H., and Michaelsen, K.F. (1995). Height of conscripts in Europe: is postneonatal mortality a predictor? *Ann. Hum. Biol.* *22*, 57–67.
- Schomacher, L. (2013). Mammalian DNA demethylation: multiple faces and upstream regulation. *Epigenetics Off. J. DNA Methylation Soc.* *8*, 679–684.
- Seisenberger, S., Peat, J.R., Hore, T.A., Santos, F., Dean, W., and Reik, W. (2013). Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* *368*, 20110330.
- Shabalin, A.A., Aberg, K.A., and van den Oord, E.J.C.G. (2015). Candidate gene methylation studies are at high risk of erroneous conclusions. *Epigenomics* *7*, 13–15.
- Shahbazian, M.D., and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. *Annu. Rev. Biochem.* *76*, 75–100.
- Shemer, J., Adamo, M.L., Roberts, C.T., and LeRoith, D. (1992). Tissue-specific transcription start site usage in the leader exons of the rat insulin-like growth factor-I gene: evidence for differential regulation in the developing kidney. *Endocrinology* *131*, 2793–2799.
- Shen, H., Qiu, C., Li, J., Tian, Q., and Deng, H.-W. (2013). Characterization of the DNA methylome and its interindividual variation in human peripheral blood monocytes. *Epigenomics* *5*, 255–269.
- Shimatsu, A., and Rotwein, P. (1987). Mosaic evolution of the insulin-like growth factors. Organization, sequence, and expression of the rat insulin-like growth factor I gene. *J. Biol. Chem.* *262*, 7894–7900.
- Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R., and Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* *479*, 74–79.
- Siddle, K., Ursø, B., Niesler, C.A., Cope, D.L., Molina, L., Surinya, K.H., and Soos, M.A. (2001). Specificity in ligand binding and intracellular signalling by insulin and insulin-like growth factor receptors. *Biochem. Soc. Trans.* *29*, 513–525.
- Silventoinen, K., Sammalisto, S., Perola, M., Boomsma, D.I., Cornes, B.K., Davis, C., Dunkel, L., De Lange, M., Harris, J.R., Hjelmborg, J.V.B., et al. (2003). Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Res. Off. J. Int. Soc. Twin Stud.* *6*, 399–408.

Siracusa, L.D., Washburn, L.L., Swing, D.A., Argeson, A.C., Jenkins, N.A., and Copeland, N.G. (1995). Hypervariable yellow (Ahvy), a new murine agouti mutation: Ahvy displays the largest variation in coat color phenotypes of all known agouti alleles. *J. Hered.* *86*, 121–128.

Sjögren, K., Liu, J.L., Blad, K., Skrtic, S., Vidal, O., Wallenius, V., LeRoith, D., Törnell, J., Isaksson, O.G., Jansson, J.O., et al. (1999). Liver-derived insulin-like growth factor I (IGF-I) is the principal source of IGF-I in blood but is not required for postnatal body growth in mice. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 7088–7092.

Sjögren, K., Jansson, J.-O., Isaksson, O.G.P., and Ohlsson, C. (2002). A transgenic model to determine the physiological role of liver-derived insulin-like growth factor I. *Minerva Endocrinol.* *27*, 299–311.

Smith, Z.D., Chan, M.M., Humm, K.C., Karnik, R., Mekhoubad, S., Regev, A., Eggan, K., and Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo. *Nature* *511*, 611–615.

Song, F., Smith, J.F., Kimura, M.T., Morrow, A.D., Matsuyama, T., Nagase, H., and Held, W.A. (2005). Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 3336–3341.

Song, F., Mahmood, S., Ghosh, S., Liang, P., Smiraglia, D.J., Nagase, H., and Held, W.A. (2009). Tissue specific differentially methylated regions (TDMR): Changes in DNA methylation during development. *Genomics* *93*, 130–139.

Stein, R., Razin, A., and Cedar, H. (1982). In vitro methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells. *Proc. Natl. Acad. Sci. U. S. A.* *79*, 3418–3422.

Stewart, D.M., Tian, L., Notarangelo, L.D., and Nelson, D.L. (2008). X-linked hypogammaglobulinemia and isolated growth hormone deficiency: an update. *Immunol. Res.* *40*, 262–270.

Stigler, S.M. (1989). Francis Galton’s Account of the Invention of Correlation. *Stat. Sci.* *4*, 73–79.

Straughen, J.K., Sipahi, L., Uddin, M., Misra, D.P., and Misra, V.K. (2015). Racial differences in IGF1 methylation and birth weight. *Clin. Epigenetics* *7*, 47.

Strous, G.J., dos Santos, C.A., Gent, J., Govers, R., Sachse, M., Schantl, J., and van Kerkhof, P. (2004). Ubiquitin system-dependent regulation of growth hormone receptor signal transduction. *Curr. Top. Microbiol. Immunol.* *286*, 81–118.

Sunami, T., Byrne, N., Diehl, R.E., Funabashi, K., Hall, D.L., Ikuta, M., Patel, S.B., Shipman, J.M., Smith, R.F., Takahashi, I., et al. (2010). Structural basis of human p70 ribosomal S6 kinase-1 regulation by activation loop phosphorylation. *J. Biol. Chem.* *285*, 4587–4594.

Sutherland, J.E., and Costa, M. (2003). Epigenetics and the environment. *Ann. N. Y. Acad. Sci.* *983*, 151–160.

Sutter, N.B., Bustamante, C.D., Chase, K., Gray, M.M., Zhao, K., Zhu, L., Padhukasahasram, B.,

- Karlins, E., Davis, S., Jones, P.G., et al. (2007). A single IGF1 allele is a major determinant of small size in dogs. *Science* *316*, 112–115.
- Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* *9*, 465–476.
- Svedružić, Ž.M. (2011). Dnmt1 structure and function. *Prog. Mol. Biol. Transl. Sci.* *101*, 221–254.
- Szulwach, K.E., Li, X., Li, Y., Song, C.-X., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levey, A.I., et al. (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat. Neurosci.* *14*, 1607–1616.
- Szyf, M. (2014). Lamarck revisited: epigenetic inheritance of ancestral odor fear conditioning. *Nat. Neurosci.* *17*, 2–4.
- Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* *324*, 930–935.
- Talens, R.P., Boomsma, D.I., Tobi, E.W., Kremer, D., Jukema, J.W., Willemsen, G., Putter, H., Slagboom, P.E., and Heijmans, B.T. (2010). Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *24*, 3135–3144.
- Tannenbaum, G.S., and Ling, N. (1984). The interrelationship of growth hormone (GH)-releasing factor and somatostatin in generation of the ultradian rhythm of GH secretion. *Endocrinology* *115*, 1952–1957.
- Teglund, S., McKay, C., Schuetz, E., van Deursen, J.M., Stravopodis, D., Wang, D., Brown, M., Bodner, S., Grosveld, G., and Ihle, J.N. (1998). Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. *Cell* *93*, 841–850.
- Teh, A.L., Pan, H., Chen, L., Ong, M.-L., Dogra, S., Wong, J., Maclsaac, J.L., Mah, S.M., McEwen, L.M., Saw, S.-M., et al. (2014). The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Res.* *24*, 1064–1074.
- Thambirajah, A.A., Dryhurst, D., Ishibashi, T., Li, A., Maffey, A.H., and Ausió, J. (2006). H2A.Z stabilizes chromatin in a way that is dependent on core histone acetylation. *J. Biol. Chem.* *281*, 20036–20044.
- Tijsterman, M., Okihara, K.L., Thijssen, K., and Plasterk, R.H.A. (2002). PPW-1, a PAZ/PIWI protein required for efficient germline RNAi, is defective in a natural isolate of *C. elegans*. *Curr. Biol. CB* *12*, 1535–1540.
- Variot, G. (1904). Gouttes de Lait Et Consultations de Nourrissons. *Br. Med. J.* *1*, 1125–1126.
- Visscher, P.M. (2008). Sizing up human height variation. *Nat. Genet.* *40*, 489–490.
- Visscher, P.M., Medland, S.E., Ferreira, M.A.R., Morley, K.I., Zhu, G., Cornes, B.K., Montgomery, G.W., and Martin, N.G. (2006). Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet.* *2*, e41.

Visscher, P.M., McEVOY, B., and Yang, J. (2010). From Galton to GWAS: quantitative genetics of human height. *Genet. Res.* *92*, 371–379.

Waddington, C.H. (1940). *Organisers and genes*. x + 160 pp.

Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characters. *Nature* *150*, 563–565.

Waddington, C.H. (1957). *The strategy of the genes; a discussion of some aspects of theoretical biology*. (London: Allen & Unwin).

Waddington, C.H. (1974). *EPIGENETICS: A TREATISE ON THEORETICAL BIOLOGY*. By Søren Løvtrup. John Wiley and Sons, London, 1974. Pp. ix+547. £15.00. *Q. J. Exp. Physiol. Cogn. Med. Sci.* *59*, 261–262.

Waddington, C.H. (2012). The epigenotype. 1942. *Int. J. Epidemiol.* *41*, 10–13.

Walenkamp, M.J.E., Losekoot, M., and Wit, J.M. (2013). Molecular IGF-1 and IGF-1 receptor defects: from genetics to clinical management. *Endocr. Dev.* *24*, 128–137.

Wallenius, K., Sjögren, K., Peng, X.D., Park, S., Wallenius, V., Liu, J.L., Umaerus, M., Wennbo, H., Isaksson, O., Frohman, L., et al. (2001). Liver-derived IGF-I regulates GH secretion at the pituitary level in mice. *Endocrinology* *142*, 4762–4770.

Wang, L., Wang, X., and Adamo, M.L. (2000). Two putative GATA motifs in the proximal exon 1 promoter of the rat insulin-like growth factor I gene regulate basal promoter activity. *Endocrinology* *141*, 1118–1126.

Wang, Y., Price, S.E., and Jiang, H. (2003). Cloning and characterization of the bovine class 1 and class 2 insulin-like growth factor-I mRNAs. *Domest. Anim. Endocrinol.* *25*, 315–328.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* *40*, 897–903.

Waterland, R.A. (2003). Do maternal methyl supplements in mice affect DNA methylation of offspring? *J. Nutr.* *133*, 238; author reply 239.

Waterland, R.A. (2014). Epigenetic mechanisms affecting regulation of energy balance: many questions, few answers. *Annu. Rev. Nutr.* *34*, 337–355.

Waterland, R.A., and Michels, K.B. (2007). Epigenetic epidemiology of the developmental origins hypothesis. *Annu. Rev. Nutr.* *27*, 363–388.

Waterland, R.A., Dolinoy, D.C., Lin, J.-R., Smith, C.A., Shi, X., and Tahiliani, K.G. (2006). Maternal methyl supplements increase offspring DNA methylation at Axin Fused. *Genes. N. Y. N* *2000* *44*, 401–406.

Wayne, R.K., and Ostrander, E.A. (2007). Lessons learned from the dog genome. *Trends Genet. TIG* *23*, 557–567.

Wayne, R.K., and vonHoldt, B.M. (2012). Evolutionary genomics of dog domestication.

Mamm. Genome Off. J. Int. Mamm. Genome Soc. 23, 3–18.

Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., and Schübeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* 37, 853–862.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* 39, 457–466.

Weedon, M.N., Lettre, G., Freathy, R.M., Lindgren, C.M., Voight, B.F., Perry, J.R.B., Elliott, K.S., Hackett, R., Guiducci, C., Shields, B., et al. (2007). A common variant of HMGA2 is associated with adult and childhood height in the general population. *Nat. Genet.* 39, 1245–1250.

Williamson, C.M., Turner, M.D., Ball, S.T., Nottingham, W.T., Glenister, P., Fray, M., Tymowska-Lalanne, Z., Plagge, A., Powles-Glover, N., Kelsey, G., et al. (2006). Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster. *Nat. Genet.* 38, 350–355.

Woelfle, J., Chia, D.J., and Rotwein, P. (2003). Mechanisms of growth hormone (GH) action. Identification of conserved Stat5 binding sites that mediate GH-induced insulin-like growth factor-I gene activation. *J. Biol. Chem.* 278, 51261–51266.

Wong, C.C.Y., Caspi, A., Williams, B., Craig, I.W., Houts, R., Ambler, A., Moffitt, T.E., and Mill, J. (2010). A longitudinal study of epigenetic variation in twins. *Epigenetics Off. J. DNA Methylation Soc.* 5, 516–526.

Wood, A.R., Esko, T., Yang, J., Vedantam, S., Pers, T.H., Gustafsson, S., Chu, A.Y., Estrada, K., Luan, J., Kutalik, Z., et al. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.*

Woods, K.A., Camacho-Hübner, C., Savage, M.O., and Clark, A.J. (1996). Intrauterine growth retardation and postnatal growth failure associated with deletion of the insulin-like growth factor I gene. *N. Engl. J. Med.* 335, 1363–1367.

Woods, K.A., Camacho-Hübner, C., Barter, D., Clark, A.J., and Savage, M.O. (1997). Insulin-like growth factor I gene deletion causing intrauterine growth retardation and severe short stature. *Acta Paediatr. Oslo Nor.* 1992 Suppl. 423, 39–45.

Wu Ct, null, and Morris, J.R. (2001). Genes, genetics, and epigenetics: a correspondence. *Science* 293, 1103–1105.

Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., and Kellis, M. (2005). Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* 434, 338–345.

Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J., and Viegas-Péquignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* 402, 187–191.

Yakar, S., Liu, J.L., Stannard, B., Butler, A., Accili, D., Sauer, B., and LeRoith, D. (1999). Normal

growth and development in the absence of hepatic insulin-like growth factor I. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 7324–7329.

Yakar, S., Pennisi, P., Wu, Y., Zhao, H., and LeRoith, D. (2005). Clinical Relevance of Systemic and Local IGF-I. In *Endocrine Development*, S. Cianfarani, D.R. Clemmons, and M.O. Savage, eds. (Basel: KARGER), pp. 11–16.

Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* *42*, 565–569.

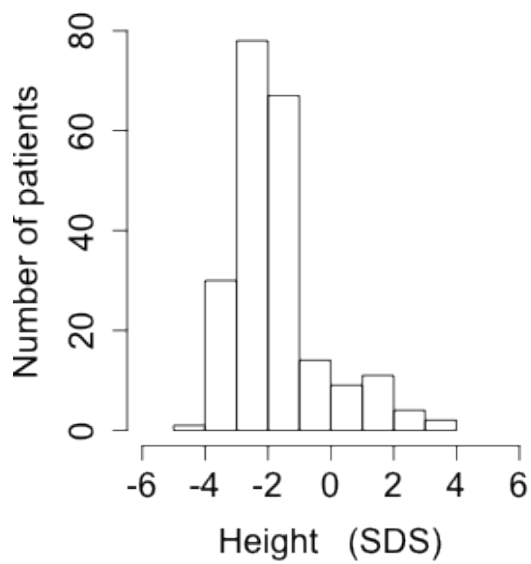
Zhao, S., Wang, Y., Liang, Y., Zhao, M., Long, H., Ding, S., Yin, H., and Lu, Q. (2011). MicroRNA-126 regulates DNA methylation in CD4+ T cells and contributes to systemic lupus erythematosus by targeting DNA methyltransferase 1. *Arthritis Rheum.* *63*, 1376–1386.

Zhou, Y., Lu, Y., and Tian, W. (2012). Epigenetic features are significantly associated with alternative splicing. *BMC Genomics* *13*, 123.

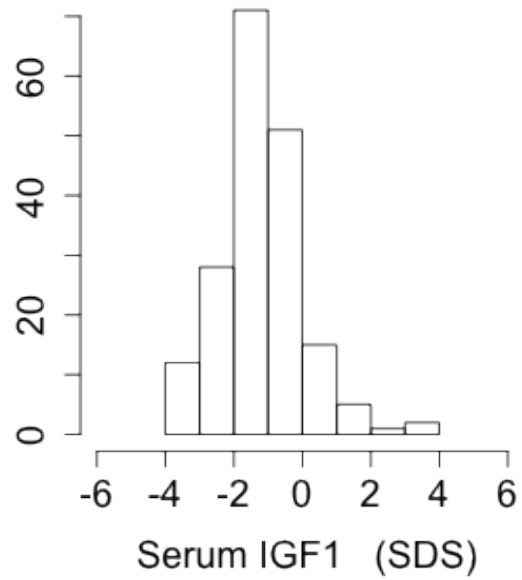
Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat. Genet.* *39*, 61–69.

## ANNEXES

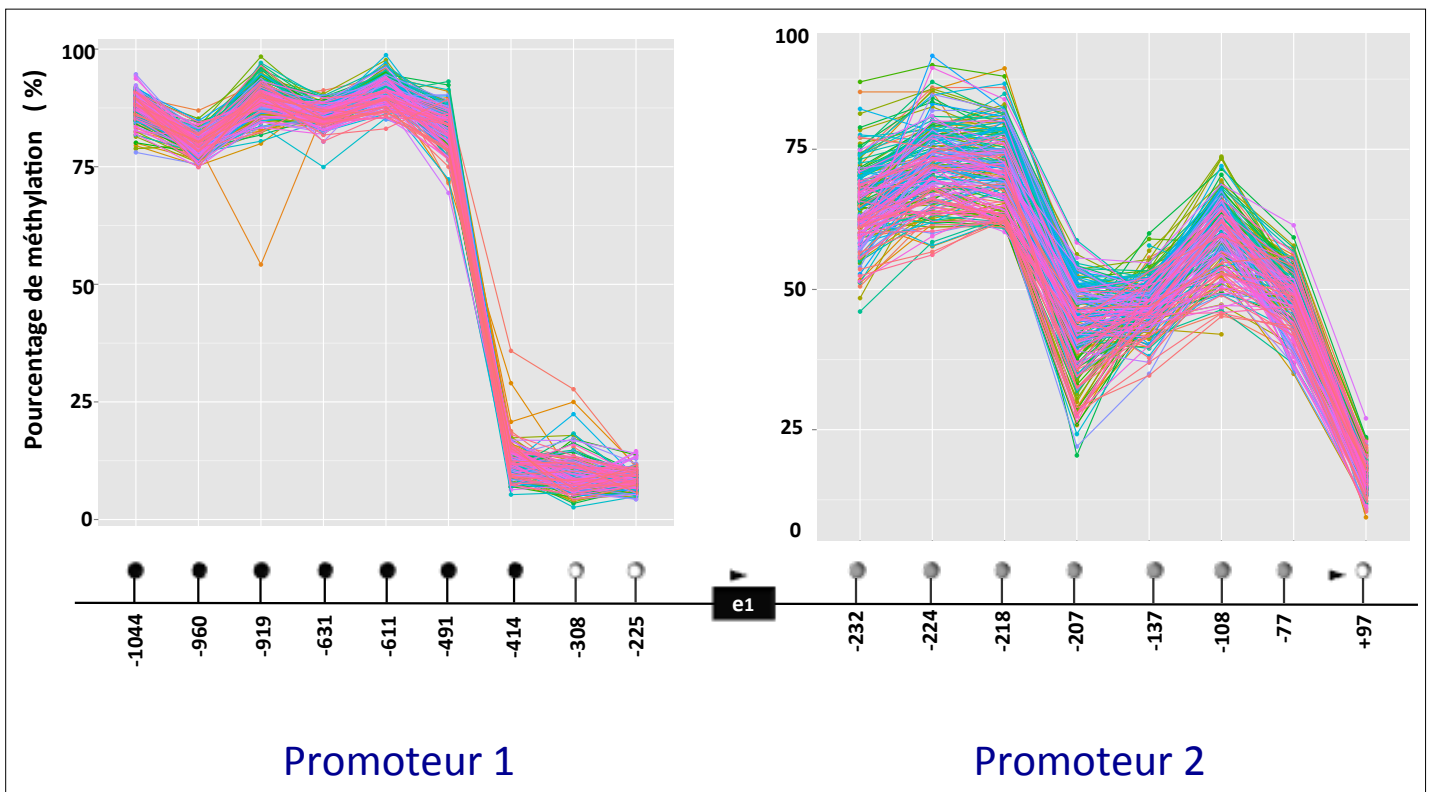
Distribution de la taille



Distribution de l'IGF1 circulant



### Profil individuel de méthylation des promoteurs du gène *IGF1* des patients





## Summary

At the interface of genetics and environment, epigenetics contributes to phenotypic diversity. Quantifying the impact of epigenetic variation on quantitative traits (QT), an emerging challenge in humans. Growth provides a handset of quantitative traits to epigenetic studies. We studied the variability of several inter-related QTs: clinical QTs (height, height response to growth hormone and biological QT (serum IGF1 and serum IGF1 response to GH).

Since insulin-like growth factor 1 (IGF1) controls postnatal growth in mammals including human, we tested whether the CG methylation of the two promoters (P1 and P2) of the *IGF1* gene could be an epigenetic contributor to the individual variation i) in circulating IGF1 and stature in growing children. ii) on response of these parameters to treatment with (GH).

### i) Child height and circulating IGF1

To explore the relation between *IGF1* promoter methylation and height, we studied two cohorts of pediatric endocrinology department, totalling 216 prepubertal children with various statures.

The methylation of a cluster of six CGs located within the proximal part of the *IGF1* P2 promoter showed a strong negative association with serum IGF1 and growth. These correlations were observed in two cohorts of growing children. Tall children show lower levels of methylation in several CGs in P2 and P1 promoters of *IGF1* gene than short children with idiopathic short stature. CG methylation contributed 13% to the variance of height and 10% to the variance of serum IGF1.

To test if the found association reflected biological causality, we tested if methylation at the P2 promoter affects the transcriptional activity of the *IGF1* gene.

The transcriptional activity of the P2 promoter was inversely correlated with the CG methylation in mononuclear blood cells.

We established that high levels of CG methylation at the two promoters of *IGF1* contributed to the many molecular mechanisms responsible for “idiopathic” short stature.

### ii) Response to treatment with (GH).

Short children using growth hormone (GH) to accelerate their growth respond to this treatment with a variable efficacy. The causes of this individual variability are partially understood and could involve epigenetics. In this aim, we investigated the contribution of DNA methylation to the response to GH at two levels: direct effect of GH on transcription of *IGF1* gene, on circulating IGF1 and on the growth response to GH. Following a GH injection, we found a variable increase in *IGF1* transcripts across the studied children. The increase in P2-driven transcripts showed a strong inverse correlation with 4/8 of P2 CGs. Among the CGs of P1 promoter, only CG-611 showed an inverse correlation with P1-driven transcripts. Variability of DNA methylation in these CGs contributes with 27% to 67% of increase in transcripts.

In 136 children with idiopathic short stature, we showed that DNA methylation of the P2 promoter is associated with growth response to GH during the first year of GH administration, for both increment in growth rate and circulating IGF1. CG-137 methylation of P2 promoter contributes 25% to variance of growth response to GH. The link between DNA methylation and the response to a treatment in humans illustrating the role of epigenetic marks as potent contributors to conclusion « pharmacoeugenetics ».

Our work can find application in growth physiology and therapeutics, as well as for studies in aging, longevity or cancer where IGF1 has a prominent role.