

KARINE TREMBLAY

**COMPARAISON DE DIFFÉRENTES APPROCHES  
MÉTHODOLOGIQUES DANS L'IDENTIFICATION DE  
DÉTERMINANTS GÉNÉTIQUES DE SUSCEPTIBILITÉ  
OU DE PROTECTION POUR L'ASTHME**

Thèse présentée  
à la Faculté des études supérieures de l'Université Laval  
dans le cadre du programme de doctorat en médecine expérimentale  
pour l'obtention du grade de Philosophiae Doctor (Ph.D.)

FACULTÉ DE MÉDECINE  
UNIVERSITÉ LAVAL  
QUÉBEC

2009

## Résumé

L'asthme est une maladie respiratoire chronique qui affecte autant les enfants que les adultes et qui se présente sous diverses manifestations cliniques. Cette maladie est également reconnue comme un trait complexe puisque plusieurs facteurs environnementaux et génétiques sont impliqués dans son développement. Les déterminants génétiques de susceptibilité ou de protection à l'asthme interagissent entre eux de même qu'avec l'environnement dans son développement. Depuis les dernières décennies, les études d'association ont permis d'identifier plus de 170 gènes associés à l'asthme ou à ses manifestations cliniques. Malgré cet avancement des connaissances en génétique de l'asthme, la confirmation des associations déjà observées et l'identification de nouveaux déterminants génétiques fait l'objet d'une recherche très active. C'est dans un tel contexte que le principal objectif de la présente thèse s'inscrit, à savoir l'identification de déterminants génétiques de susceptibilité ou de protection pour l'asthme. Pour ce faire, de nouvelles approches méthodologiques ainsi que des approches plus classiques ont été employées pour mener à terme les six études d'association présentées. Grâce à méthodologies novatrices, quatre nouvelles associations génétiques positives ont été mises en évidence avec les gènes *CX3CR1*, *ALOX15*, *PLAU* et *PTPRE*. De tels résultats ont permis, *a priori*, de démontrer l'utilité des nouvelles approches employées dans les études d'association génétique et de démontrer leur efficacité. L'acquisition de ces nouvelles connaissances a également mené à l'élaboration de nouvelles hypothèses de recherche et les projets en cours devraient améliorer la compréhension des mécanismes biologiques en cause dans l'asthme. Finalement, l'étude plus approfondie de ces gènes permettra, éventuellement, de développer de nouvelles thérapies et de surcroît, améliorer les conditions de vie des personnes asthmatiques.

## **Avant-propos**

La présente thèse est consacrée à l'asthme, l'une des maladies respiratoires des plus communes qui affecte des centaines de millions de personnes dans le monde et dont la thérapie actuelle ne peut soulager tous les individus qui en sont atteints. C'est ainsi que l'asthme mobilise plusieurs équipes de chercheurs à l'échelle internationale, ce qui a pour effet d'augmenter les connaissances et de mieux comprendre son histoire naturelle. À sa définition de plus en plus spécifique, présentée dans le Chapitre 1, il est généralement admis que l'asthme est un trait complexe induit par l'interaction d'une multitude de gènes agissant les uns sur les autres et étant modulés par l'environnement. En fait, plus de 170 gènes sont associés à l'asthme ou à ses manifestations cliniques et parmi eux, certains contribueraient à la susceptibilité à la maladie tandis que d'autres auraient un effet protecteur. Le résumé de l'état des connaissances sur la génétique de l'asthme est présenté dans le Chapitre 2. L'identification des gènes de l'asthme s'effectue grâce aux études d'épidémiologie génétique, qui sont précisément le sujet des travaux présentés dans cette thèse et dont les étapes nécessaires à leur réalisation sont décrites dans le Chapitre 3.

La recherche en génétique de l'asthme repose sur des méthodologies scientifiques rigoureuses. À cet effet, le Chapitre 4 met en évidence les hypothèses et objectifs de recherche qui ont motivé les études d'association présentées dans cette thèse et dont tous les résultats se traduisent en cinq articles scientifiques publiés, un sous presse et un en préparation (Chapitres 5 à 8, incluant la mise en contexte de chacun des articles et la contribution des auteurs). Il est à noter que les quatre premiers articles font partie de la même approche méthodologique, c'est pourquoi ils sont regroupés dans le Chapitre 5. Suivant ces résultats, une synthèse de l'information génétique recueillie suivie d'une discussion axée sur les méthodologies employées, leurs forces et leurs faiblesses ainsi que des constats qu'elles animent font l'objet du Chapitre 9. Enfin, le Chapitre 10 propose un modèle de perspectives de recherche suivant les études d'association génétique en lien avec les résultats observés. Pour ce qui est de la conclusion, elle fait l'objet du Chapitre 11. Quant aux références citées dans l'ensemble des sections de cette thèse, elles sont listées dans le Chapitre 12. Toutefois, cette liste exclut les citations des articles puisque chacun possède sa propre liste de références.

Avant d'entrer dans le vif du sujet, j'aimerais prendre quelques lignes pour souligner et remercier les personnes qui ont contribué à la réalisation de mon doctorat et par conséquent, à la rédaction de la présente thèse. Sans leur présence, leur aide et leur support, rien de tout ceci n'aurait été possible. Ces personnes m'ont beaucoup apporté tant sur les plans professionnel que personnel et aujourd'hui, je leur dois reconnaissance et gratitude.

Tout d'abord, à mon mentor Catherine Laprise. En plus de m'avoir transmis sa passion pour la recherche, elle m'a accordé sa confiance, sa disponibilité ainsi que l'environnement et les encouragements nécessaires pour réaliser mon doctorat. Je suis honorée d'avoir été l'étudiante d'une scientifique aussi intègre et authentique. Avec tout mon respect, merci Catherine pour ton exemple si inspirant.

À mon codirecteur Michel Laviolette, un grand merci de m'avoir accueillie au centre de recherche de l'Hôpital Laval, d'avoir suivi la progression de mes travaux et d'avoir partagé toutes ces discussions si enrichissantes.

Aux participant(e)s, pour qui nous sommes tous dévoués. Merci pour votre généreuse contribution et surtout, de croire en nos efforts. C'est la source de notre motivation.

À tous les étudiant(e)s, professionnel(le)s de recherche et chercheur(e)s que j'ai rencontrés dans tous les centres de recherche que j'ai visité et congrès auxquels j'ai participé. Ce fut un réel plaisir de faire votre connaissance et de passer tous ces moments si agréables avec vous.

Aux coauteur(e)s de chacun des articles. La publication est un processus lent et laborieux. Grâce à votre aide, le travail en fut allégé. Merci pour toutes ces précieuses collaborations.

Aux divers organismes subventionnaires qui m'ont attribué des prix et des bourses d'étude en plus de m'avoir permis de voyager. Merci beaucoup, ce fut plus qu'apprécié.

À mes parents, mes sœurs et leur compagnon de vie, ma belle-famille et mes amis. Votre présence est essentielle à mon bonheur. Soyez assurés que chacune des pensées et

paroles d'encouragement ou de réconfort que vous m'avez prodiguées tout au cours de ces années passées au doctorat ont toujours touché leur cible. Avec tout mon cœur, merci d'avoir cru en moi et pour votre support inconditionnel dans cette autre étape de ma vie.

Finalement, et non le moindre, à mon conjoint. Pendant toutes ces années, tu étais là, supportant mes bons moments autant que les plus difficiles. L'aboutissement de ce doctorat n'est pas étranger à ta présence à mes côtés et à défaut de trouver les mots justes pour décrire toute l'importance que tu représentes pour moi, je te dédie ma thèse, avec tout mon amour.

*À Martin*

## Table des matières

<b>Résumé</b>	<b>ii</b>
<b>Avant-propos</b>	<b>iii</b>
<b>Table des matières</b>	<b>vii</b>
<b>Liste des figures</b>	<b>xiii</b>
<b>Liste des tableaux</b>	<b>xiv</b>
<b>Liste des abréviations</b>	<b>xv</b>
<b>Chapitre 1 : L'asthme</b>	<b>1</b>
1.1 Définition	2
1.2 Étiologie	2
1.3 Types d'asthme	3
1.4 Physiopathologie	4
1.4.1 Hyperréactivité bronchique	5
1.4.2 Inflammation	5
1.4.3 Remodelage bronchique	7
1.4.4 Asthme allergique	11
<b>Chapitre 2 : Génétique de l'asthme</b>	<b>16</b>
2.1 Agrégation familiale	16
2.2 Étude de jumeaux	17
2.3 Analyse de ségrégation	17
2.4 L'asthme : un trait complexe	18
<b>Chapitre 3 : Épidémiologie génétique de l'asthme</b>	<b>20</b>
3.1 Étude de liaison	20
3.2 Étude d'association	23
3.2.1 Approche génomique	24
3.2.2 Approche par gènes candidats	28
3.2.2.1 Sélection des gènes candidats	29
3.2.2.1.1 Clonage positionnel	29
3.2.2.1.2 Fonction biologique	30
3.2.2.1.3 Profil d'expression génique	31
3.2.2.1.4 Outil bioinformatique Candidate Genes to Inherited Diseases (G2D)	32
3.2.2.1.5 Association génétique connue avec le trait complexe étudié	33
3.2.2.1.6 Association génétique connue un autre trait complexe qui partage une composante physiopathologique avec le trait complexe étudié	34
3.2.2.2 Sélection des SNPs	35

3.2.2.2.1 Association directe – SNPs fonctionnels	35
3.2.2.2.2 Association indirecte – TagSNPs	36
3.2.2.3 Échantillonnage	38
3.2.2.3.1 Cas-témoins	38
3.2.2.3.2 Trios familiaux	41
3.2.2.4 Évaluation phénotypique	43
3.2.2.5 Évaluation génotypique	45
3.2.2.6 Analyses statistiques	47
<b>Chapitre 4 : Hypothèses et objectifs du projet de recherche</b>	<b>52</b>
4.1 Hypothèse générale	52
4.2 Hypothèses spécifiques	52
4.3 Objectif général	53
4.4 Objectifs spécifiques	53
<b>Chapitre 5 : Les gènes ciblés par une étude du profil d'expression de millier de gènes dans les biopsies bronchiques de sujets asthmatiques comparés aux sujets témoins</b>	<b>56</b>
5.1 Premier article : Association entre le gène <i>CX3CR1</i> et l'asthme	60
5.1.1 TITLE PAGE	61
5.1.2 ABSTRACT	62
5.1.3 INTRODUCTION	62
5.1.4 RESULTS	64
5.1.5 DISCUSSION	65
5.1.6 SUBJECTS AND METHODS	68
5.1.7 ACKNOWLEDGMENTS	70
5.1.8 REFERENCES	71
5.1.9 LEGEND TO FIGURES	75
5.1.10 Table 1	77
5.1.11 Table 2	78
5.1.12 Table 3	79
5.1.13 Table 4	80
5.1.14 Table 5	81
5.1.15 Figure 1	82
5.1.16 Figure 2	83
5.1.18 SUPPLEMENTARY MATERIAL	84
5.2 Deuxième article : Variations génétiques dans les gènes des molécules de signalisation immune différemment exprimés dans les biopsies bronchiques de sujets asthmatiques	92
5.2.1 TITLE PAGE	93
5.2.2 ABSTRACT	94



5.2.3 RELEVANCE	94
5.2.4 ABBREVIATIONS	95
5.2.5 INTRODUCTION	96
5.2.6 METHODS	97
5.2.7 RESULTS	99
5.2.8 DISCUSSION	100
5.2.9 REFERENCES	103
5.2.10 Table 1	108
5.2.11 Table 2	111
5.2.12 Table 3	112
5.2.13 Table 4	113
5.2.14 Table 5	114
5.2.15 Table 6	116
5.2.16 ONLINE REPOSITORY	117
5.3 Troisième article : Étude d'association entre le gène de l'activateur du plasminogène de type urokinase ( <i>PLAU</i> ), l'asthme et l'atopie	144
5.3.1 TITLE PAGE	145
5.3.2 ABSTRACT	146
5.3.3 INTRODUCTION	147
5.3.4 MATERIALS AND METHODS	148
5.3.5 RESULTS	150
5.3.6 DISCUSSION	154
5.3.7 ACKNOWLEDGMENTS	156
5.3.8 REFERENCES	156
5.3.9 FIGURE LEGENDS	161
5.3.10 Table 1	163
5.3.11 Table 2	164
5.3.12 Table 3	165
5.3.13 Table 4	166
5.3.14 Table 5	167
5.3.15 Table 6	168
5.3.16 Table 7	169
5.3.17 Table 8	170
5.3.18 Table 9	171
5.3.19 Figure 1	172
5.3.20 Figure 2	173
5.3.21 SUPPLEMENTARY MATERIAL	174
5.4 Quatrième article : Étude comparative de deux expérimentations de puces à ADN pour définir le patron d'expression génique de l'asthme allergique	193
5.4.1 TITLE PAGE	194

5.4.2 ABSTRACT	195
5.4.3 BACKGROUND	195
5.4.3 METHODS	196
5.4.5 RESULTS	198
5.4.6 DISCUSSION	199
5.4.7 CONCLUSION	201
5.4.8 ACKNOWLEDGMENTS	202
5.4.9 REFERENCES	202
5.4.10 Table 1	206
5.4.11 Table 2	207
5.4.12 Table 3	208
5.4.13 Additional Table 4	209

## **Chapitre 6 : Les gènes ciblés à l'aide de l'outil bioinformatique *Genes to Diseases (G2D)* 215**

6.1 Cinquième article : Application de l'outil bioinformatique G2D dans l'identification de gènes candidats pour l'asthme	218
6.1.1 TITLE PAGE	219
6.1.2 ABSTRACT	220
6.1.3 INTRODUCTION	220
6.1.4 METHODS	222
6.1.5 RESULTS	225
6.1.6 DISCUSSION	229
6.1.7 ACKNOWLEDGMENTS	231
6.1.8 REFERENCES	232
6.1.9 LEGEND TO FIGURES	236
6.1.10 Table 1	237
6.1.11 Table 2	238
6.1.12 Table 3	240
6.1.13 Table 4	241
6.1.14 Figure 1	242
6.1.15 Figure 2	243
6.1.16 SUPPORTING INFORMATION FILES	244

## **Chapitre 7 : Gène ciblé en raison de son association génétique connue avec l'asthme 254**

7.1 Sixième article : Réplication d'une association entre les SNPs du 17q21 et l'asthme dans un échantillon familial Canadien-Français	255
7.1.1 TITLE PAGE	256
7.1.2 ABSTRACT	256
7.1.3 INTRODUCTION	257
7.1.4 METHODS	257

7.1.5 RESULTS AND DISCUSSION	257
7.1.6 ACKNOWLEDGMENTS	258
7.1.7 REFERENCES	258
7.1.8 Table 1	260
7.1.9 ONLINE SUPPLEMENT	261

## **Chapitre 8 : Gène ciblé en raison de son association génétique connue avec un trait complexe partageant une composante physiopathologique avec l'asthme** **265**

8.1 Septième article : Association des variants du gène filaggrine ( <i>FLG</i> ) avec l'asthme et l'atopie dans l'échantillon familial du Saguenay-Lac-St-Jean	266
8.1.1 TITLE PAGE	267
8.1.2 ABSTRACT	268
8.1.3 INTRODUCTION	268
8.1.4 SUBJECTS AND METHODS	270
8.1.5 RESULTS	272
8.1.6 DISCUSSION	274
8.1.7 ACKNOWLEDGMENTS	276
8.1.8 WEB RESOURCES	277
8.1.9 REFERENCES	277
8.1.10 FIGURE LEGENDS	281
8.1.11 Table 1	282
8.1.12 Table 2	283
8.1.13 Table 3	284
8.1.14 Table 4	285
8.1.16 Figure 1	286
8.1.17 ONLINE SUPPLEMENT	287

## **Chapitre 9 : Discussion** **300**

9.1 Synthèse de l'information génétique recueillie	300
9.2 Comparaison des approches méthodologiques employées pour la sélection des gènes candidats étudiés	301
9.2.1 Approches classiques	302
9.2.2 Approches novatrices	303
9.2.2.1 Profil d'expression génique	304
9.2.2.2 Outil G2D	306
9.3 Difficultés rencontrées dans les études d'association pour l'asthme	310
9.3.1 La réplification stricte	312
9.3.2 La réplification partielle	313
9.3.3 La non-réplification	314

9.4 Facteurs influençant la distribution des associations génétiques parmi les études indépendantes	315
9.4.1 Conception de l'échantillon	316
9.4.2 Erreur de géotypage	318
9.4.3 Puissance statistique des échantillons	319
9.4.4 Corrections statistiques	321
9.4.5 Biais de publication	323
9.4.6 Structure génétique des populations	324
9.4.7 Interactions gène-gène et gène-environnement	326
9.4.8 Hétérogénéité phénotypique	328
<b>Chapitre 10 : Perspectives de recherche</b>	<b>330</b>
10.1 Exemples de modèle d'intégration biologique	330
10.2 Nouvelles questions et hypothèses de recherche	333
<b>Chapitre 11 : Conclusion</b>	<b>339</b>
<b>Chapitre 12 : Références</b>	<b>341</b>

## Liste des figures

- Figure 1 : Représentation schématique des principaux mécanismes de remodelage se produisant dans les voies respiratoires des sujets asthmatiques ainsi que des principales cellules structurales et inflammatoires impliquées \_\_\_\_\_ 9
- Figure 2 : Représentation schématique d'une boucle de régulation inflammatoire présente dans les voies respiratoires de sujets asthmatiques allergiques \_\_\_\_\_ 14
- Figure 3 : Représentation schématique de la molécule d'ADN double brin et des deux principaux types de marqueurs génétiques utilisés en épidémiologie génétique \_\_\_\_\_ 21
- Figure 4 : Diagramme conceptuel représentant les deux principales approches et les diverses étapes nécessaires à la réalisation d'une étude d'association génétique \_\_\_\_\_ 24
- Figure 5 : Représentation schématique des étapes menant à la réalisation d'une GWAS \_\_\_\_\_ 25
- Figure 6 : Représentation schématique des combinaisons haplotypiques possibles à l'intérieur des blocs de déséquilibre de liaison formés par des groupes de SNPs transmis ensemble \_\_\_\_\_ 27
- Figure 7 : Représentation schématique d'un gène et de ses différentes régions \_\_\_\_\_ 36
- Figure 8 : Critères utilisés dans l'établissement des phénotypes relatifs à l'asthme pour les sujets de l'échantillon familial du SLSJ \_\_\_\_\_ 44
- Figure 9 : Représentation schématique et hypothétique de la voie biologique du *CX3CR1* et de son ligand dans les voies respiratoires chez l'humain \_\_\_\_\_ 331
- Figure 10 : Représentation schématique des voies respiratoires asthmatiques montrant l'expression du gène *ALOX15* ( $ARN_m$ ) et de son enzyme, la 15-lipoxygénase (15-LO-1) \_\_\_\_\_ 332
- Figure 11 : Diagramme conceptuel des étapes logiques suivant une étude d'association génétique \_\_\_\_\_ 334

## Liste des tableaux

Tableau I : Gènes de susceptibilité associés à l'asthme ou à ses phénotypes dans au moins cinq études d'association indépendantes et basées sur l'approche par gènes candidats \_\_\_\_\_ 50

Tableau II : Différences dans l'expression des gènes des molécules immunes entre les sujets témoins et asthmatiques \_\_\_\_\_ 57

Tableau III : Sélection des gènes G2D selon le nombre d'apparitions dans les analyses, la fonction biologique ainsi que les études génétiques et fonctionnelles publiées \_\_\_\_ 216

Tableau IV : Matrice de réplication utilisée pour évaluer le degré de réplication d'une étude d'association effectuée pour répliquer des observations initiales \_\_\_\_\_ 311

Tableau V : Résumé des associations et de leur réplication au sein des divers échantillons étudiés \_\_\_\_\_ 311

Tableau VI : Questions de recherche soulevées par les résultats d'une étude d'association \_\_\_\_\_ 333

Tableau VII : Gènes impliqués dans la voie biologique du *CX3CR1* \_\_\_\_\_ 335

## Liste des abréviations

\*Seules les abréviations fréquemment employées ou au cœur des éléments traités dans cette thèse paraissent dans cette liste. Quelques abréviations étant plus spécifiques à certaines figures, tableaux ou articles, leur description y est incluse.

*ADAM8* : *ADAM metallopeptidase domain 8*

ADN : Acide désoxyribonucléique

*ALOX15* : *Arachidonate 15-lipoxygenase (15-LO)*

ARN<sub>m</sub> : Acide ribonucléique messenger

ATS : *American Thoracic Society*

CAMP : *Childhood Asthma Management Program Study*

CAPPS : *Canadian Asthma Primary Prevention Study*

CP<sub>20</sub> : Concentration de métacholine qui provoque une chute de 20 % du VEMS

*CD14* : *CD14 molecule (antigen)*

*CD27* : *CD27 molecule (antigen)*

*CX3CR1* : *Chemokine (C-X3-C motif) receptor 1*

*CXCL12* : *Chemokine (C-X-C motif) ligand 12*

*DOCK1* : *Dedicator of cytokinesis 1*

EHW : Équilibre de Hardy-Weinberg

*FANK1* : *Fibronectin type III and ankyrin repeat domains 1*

*FLG* : *Filaggrin*

G2D : *Candidate Genes to Inherited Diseases (Genes to diseases)*

GWAS : *Genome-Wide Association Study*

*GPR123* : *G protein-coupled receptor 123*

HRB : Hyperréactivité bronchique

IgE : Immunoglobuline E

*IL2RB* : *Interleukin 2 receptor, beta*

*IL7R* : *Interleukin 7 receptor*

*LPA* : *Lipoprotein, Lp(a)*

NCBI : *National Center for Biotechnology Information*

*NOS2A* : *Nitric oxide synthase 2, inducible*

*NOX3* : *NADPH oxidase 3*

OMIM : *Online Mendelian Inheritance in Man*

*ORMDL3* : *ORM1-like 3 (S. cerevisiae)*

PCR : *Polymerase Chain Reaction*

PLAU : *Plasminogen activator, urokinase*

PTPRE : *Protein tyrosine phosphatase, receptor type, E*

RFLP : *Restriction Fragment Length Polymorphism*

SAGE : *Study of Asthma Genes and the Environment Cohort*

SERPINB2 : *Serpin peptidase inhibitor, clade B (ovalbumin), member 2*

SFRP1 : *Secreted frizzled-related protein 1*

SLSJ : *Saguenay–Lac-St-Jean*

SNP : *Single Nucleotide Polymorphism*

SNX9 : *Sorting nexin 9*

SPT : *Skin Prick Test*

TagSNP : *Tagging single nucleotide polymorphism*

VEMS : *Volume expiratoire maximal en une seconde*

VIL2 : *Villin 2 (ezrin)*

VIP : *Vasoactive intestinal peptide*



# Chapitre 1

## L'asthme

L'asthme est une maladie respiratoire chronique des plus communes qui affecte approximativement 300 millions de personnes dans le monde (Beasley, 2004; Masoli et al., 2004). Au Canada, ce nombre est estimé à trois millions, incluant environ 700 000 Québécois (dont 300 000 enfants) (Association pulmonaire du Québec, 2007). Selon sa sévérité et son degré de contrôle, cette maladie peut parfois être fatale, ayant entraîné le décès de 287 Canadiens en 2003 (Association pulmonaire du Canada, 2008). Les divers efforts nationaux pour lutter contre l'asthme tendent toutefois à faire chuter les taux de mortalité, ce qui fait que la mortalité associée à l'asthme est devenue un événement qui se raréfie depuis les années 1980. Néanmoins, la morbidité associée à l'asthme est toujours présente en 2008. Ce sont particulièrement les cas les plus sévères d'asthme qui permettent de témoigner de l'importance de son impact socio-économique (mesurable en coûts directs et indirects) (Marion et al., 1985) et du degré auquel cette maladie affecte la qualité de vie des personnes atteintes (Bousquet et al., 1994). Les coûts directs incluent les soins médicaux, les médicaments et les séjours à l'hôpital, ce qui représente des milliards de dollars par année que les familles et les sociétés doivent assumer (Smith et al., 1997). Au Québec par exemple, l'asthme occasionne chaque année plus de 760 000 visites chez le médecin, environ 100 000 visites à l'urgence, 56 000 jours d'hospitalisation et 4 000 appels pour des services ambulanciers (Association pulmonaire du Québec, 2007). Quant aux coûts indirects, ils incluent l'absentéisme scolaire et professionnel (perte de productivité) et sont liés au degré d'invalidité que cause la maladie (Smith, 1988; Hill et al., 1989). D'ailleurs, en 2004, l'asthme se classait au 25<sup>e</sup> rang mondial des maladies invalidantes, derrière les conditions périnatales, les infections respiratoires, le sida, les troubles bipolaires, l'alcoolisme, la schizophrénie et le diabète pour ne nommer que ces exemples (Masoli et al., 2004).

Suivant ce survol de l'épidémiologie relative à l'asthme, le présent chapitre est consacré à l'état des connaissances de la maladie, portant une attention particulière à sa définition usuelle suivie des principales causes et facteurs de risque reconnus pour son développement (étiologie) pour finir avec les principaux mécanismes physiologiques qui régulent sa pathologie (regroupés sous le terme physiopathologie).

## 1.1 Définition

L'asthme se définit comme une affection pulmonaire chronique caractérisée par une inflammation de la paroi bronchique médiée par les différentes cellules structurales et inflammatoires présentes ou recrutées dans les poumons. Cette inflammation chronique cause de multiples changements structuraux et physiologiques qui sont responsables, individuellement ou collectivement, du rétrécissement des voies aériennes et de l'apparition des deux composantes physiologiques majeures observées dans l'asthme, soit une hyperréactivité des bronches à divers stimuli (endogènes ou exogènes) ainsi qu'une obstruction variable et réversible (spontanément ou suivant un traitement) du passage de l'air. Ces changements physiopathologiques mènent aux symptômes aigus ou persistants observés dans l'asthme comme l'essoufflement, l'oppression thoracique, la respiration sifflante, la production de mucus et la toux (Boulet et al., 1999; Lemanske et Busse, 2003; Global initiative for asthma, 2007). L'hyperréactivité et l'obstruction bronchiques sont variables d'un individu asthmatique à l'autre, ce qui contribue à la diversité des manifestations cliniques observées (appelée hétérogénéité clinique ou phénotypique), incluant les divers degrés de sévérité et les réponses thérapeutiques variées aux diverses médications (Lemanske et Busse, 2003).

## 1.2 Étiologie

L'hétérogénéité clinique observée chez les personnes asthmatiques est le résultat de l'interaction complexe de plusieurs facteurs inducteurs (responsables du développement de la maladie), de facteurs déclencheurs (responsables de l'expression et de la sévérité de la maladie) ainsi que de facteurs qui sont à la fois inducteurs et déclencheurs (Busse et Lemanske, 2001).

Les facteurs de risque inducteurs incluent en premier lieu les prédispositions génétiques à l'asthme, à l'hyperréactivité bronchique ou à l'atopie. À ce jour, plus de 170 gènes ont été associés à l'asthme ou à ses manifestations cliniques (détails dans le Chapitre 3). Malgré toutes ces connaissances, le profil génétique d'un individu asthmatique, c'est-à-dire quels sont les gènes impliqués et quelles sont les interactions entre ceux-ci et les diverses expositions environnementales, reste toujours inconnu et demeure à documenter (Tattersfield et al., 2002). Le genre est un autre facteur inducteur important

dans l'expression clinique de la maladie. Bien qu'il soit observé que le sexe masculin soit un facteur de risque à développer l'asthme dans l'enfance et que cette tendance se renverse à l'âge adulte pour les femmes, les mécanismes ne sont pas encore bien caractérisés (Postma, 2007). Les autres facteurs inducteurs de l'asthme sont liés aux expositions environnementales. Parmi ces derniers, on retrouve les infections respiratoires, les allergènes domestiques (exemple : animaux, moisissures ou acariens) ou extérieurs (exemple : pollens), le tabagisme (actif ou passif), la pollution atmosphérique, le style de vie (rural ou urbain), la diète, l'utilisation d'antibiotiques, la manifestation d'autres maladies (exemple : dermatite atopique ou obésité), l'exposition aux diverses substances irritantes des milieux de travail, etc (Maddox et Schwartz, 2002; Lemanske et Busse, 2006). Une fois que l'asthme est établi chez un individu, de multiples facteurs sont responsables de déclencher et/ou d'exacerber les symptômes. Dans cette catégorie de facteurs déclencheurs se retrouvent, en plus des expositions environnementales précédemment énumérées, la période de la journée, l'exercice, l'air froid et les facteurs psychosociaux tel le stress (Lemanske et Busse, 2006).

### **1.3 Types d'asthme**

La diversité des facteurs de risque de l'asthme fait en sorte que la maladie se manifeste différemment parmi les personnes atteintes, ce qui rend possible l'établissement de différents types d'asthme, classifiés selon les facteurs étiologiques en cause, l'âge du début de la maladie, le type d'inflammation impliqué ou le degré de sévérité.

Une des plus grandes distinctions concerne l'âge de l'apparition de la maladie, soit à l'enfance ou à l'âge adulte. La plupart des personnes à haut risque de développer la maladie présentent généralement les premiers signes cliniques avant l'âge de cinq ans et la maladie persiste à l'âge adulte pour 30 à 40 % d'entre eux (Kiley et al., 2007). L'origine du développement de l'asthme à l'enfance repose sur une interaction complexe entre les susceptibilités génétiques (incluant l'hyperréactivité bronchique, les allergies ou l'histoire familiale d'asthme et d'allergie), la réponse aux expositions environnementales (incluant les allergènes, les virus, la fumée secondaire ou la pollution) ainsi que le développement des systèmes pulmonaires et immunitaires (Lemanske et Busse, 2006; Kiley et al., 2007). Quant à l'asthme qui se développe à l'âge adulte, les facteurs de risque ne sont pas bien définis (Panettieri et al., 2008). On

reconnait toutefois les mêmes prédispositions génétiques et expositions environnementales que chez l'enfant en plus des expositions irritantes liées au milieu de travail (asthme professionnel), le tabagisme, le genre ou la présence d'autres maladies respiratoires comme la rhinite (Kiley et al., 2007; Panettieri et al., 2008).

L'atopie, plus particulièrement les allergies, est reconnue comme le plus grand facteur de risque lié au développement de l'asthme (Lemanske et Busse, 2003; Guilbert et al., 2004; Holgate et al., 2007). Le type d'asthme le plus prévalent est donc l'asthme allergique qui se développe chez les personnes sensibilisées à un ou plusieurs allergènes (Holt et al., 1999). Bien que la majorité des personnes asthmatiques soit allergique (Holgate, 2008), à l'inverse, seulement 10 % des personnes allergiques développent de l'asthme (Holgate et al., 2000). Les individus qui présentent un asthme allergique ont généralement manifesté leurs premiers symptômes à l'enfance car l'interaction entre les gènes et les expositions environnementales influencent le développement du système immunitaire inné et acquis (Maddox et Schwartz, 2002). Par conséquent, les individus développent une sensibilisation progressive aux allergènes de leur environnement, ce qui fait que les expositions subséquentes déclenchent la réponse inflammatoire et l'apparition des symptômes (Kiley et al., 2007). La majorité des données documentées relatives à la pathogénèse de l'asthme est concentrée sur l'asthme allergique, tout comme le sont les travaux de la présente thèse. C'est pourquoi une attention particulière sera portée à la physiopathologie de l'asthme allergique (sous-section 1.4.4).

Parmi les autres types d'asthme, on reconnaît l'asthme qui se manifeste chez les personnes allergiques ou intolérantes à l'aspirine et aux anti-inflammatoires non stéroïdiens (Szczeklik et Sanak, 2006), l'asthme résistant à la corticothérapie anti-inflammatoire (Ito et al., 2006) ainsi que les divers degrés de sévérité de l'asthme (Lemanske et Busse, 2003). Puisque ces autres manifestations de l'asthme ne font pas l'objet de la présente thèse, elles ne sont pas traitées plus en détail.

## **1.4 Physiopathologie**

Peu importe le type d'asthme ou ses facteurs déclencheurs, les événements physiopathologiques qui en résultent reposent sur des mécanismes inflammatoires et de remodelage responsables du développement de l'hyperréactivité et de l'obstruction

bronchiques, ainsi que des symptômes relatifs à l'asthme. Bien que ces caractéristiques soient variables d'une personne asthmatique à une autre, il est tout de même possible de dresser un portrait général des mécanismes qui les engendrent afin d'élucider les bases biologiques nécessaires à la compréhension du choix des gènes étudiés dans la présente thèse et qui feront l'objet des Chapitres 5 à 8.

#### **1.4.1 Hyperréactivité bronchique**

L'hyperréactivité des bronches à divers stimuli, comme l'inhalation de diverses substances irritantes, l'air froid, l'exercice ou l'hyperventilation (Cockcroft, 2001), est le résultat d'une contraction exagérée du muscle lisse qui rétrécit le passage de l'air (en combinaison avec les événements inflammatoires, sous-section 1.4.2), ce qui mène à l'apparition des symptômes de l'asthme (Berend et al., 2008). Cette bronchoconstriction est régulée par plusieurs médiateurs (histamine, leucotriènes, prostaglandines, facteur activateur des plaquettes) libérés par les cellules résidentes (mastocytes, macrophages alvéolaires, cellules épithéliales, cellules endothéliales) ou infiltrées (éosinophiles, lymphocytes, neutrophiles, basophiles et plaquettes) dans les voies respiratoires asthmatiques (Lemanske et Busse, 2003). Bien que l'hyperréactivité bronchique soit une caractéristique fondamentale du développement de l'asthme (Boulet, 2003; Berend et al., 2008), elle n'est pas spécifique à l'asthme puisqu'elle peut être asymptomatique (Woolcock et al., 1987) ou se présenter chez des personnes atteintes d'autres maladies respiratoires comme la fibrose kystique ou la maladie obstructive chronique (Rijcken et al., 1987; van Haren et al., 1995). Des individus ne présentant aucune de ces maladies peuvent également présenter une hyperréactivité bronchique quelque temps après une infection respiratoire virale (Lemanske et Busse, 2003). La prévalence d'hyperréactivité bronchique sans symptôme d'asthme dans la population générale varie selon les études, une proportion inférieure à 15 % étant la plus souvent rapportée (Jansen et al., 1997). Par contraste, dans une population de sujets hyperréacteurs, cette proportion de sujets asymptomatiques est généralement supérieure à 30 % (Jansen et al., 1997).

#### **1.4.2 Inflammation**

Le processus inflammatoire et ses effets sur la structure des voies aériennes semblent être le principal mécanisme qui provoque l'apparition et la persistance de l'asthme

(Boulet et al., 1999; Busse et Lemanske, 2001). Parmi ces mécanismes, certains sont communs à toutes les réactions inflammatoires comme la congestion vasculaire et l'exsudation des protéines ainsi que la migration des cellules inflammatoires hors des vaisseaux sanguins vers l'interstitium tissulaire (Busse et al., 1999). Dans l'asthme, les mécanismes inflammatoires impliquent l'ensemble des cellules structurales (cellules épithéliales, cellules endothéliales, fibroblastes et cellules du muscle lisse) et immunitaires de l'arbre bronchique en particulier les éosinophiles, les lymphocytes T (essentiellement de type Th2), les lymphocytes B, les macrophages, les cellules dendritiques, les neutrophiles et les mastocytes (Bousquet et al., 2000). Ces cellules, une fois activées, libèrent les cytokines qui initient et perpétuent la réponse inflammatoire grâce à leur habileté à activer les divers facteurs de transcription qui régulent l'expression des gènes d'un éventail d'autres médiateurs tels les cytokines pro-inflammatoires, les chimiokines, les molécules d'adhésion et les facteurs de survie (Busse et Lemanske, 2001).

Une des composantes majeures de l'inflammation dans l'asthme est une dense infiltration de cellules immunitaires dans la paroi bronchique, particulièrement les éosinophiles, les neutrophiles, les lymphocytes et les mastocytes (Lemanske et Busse, 2003). Ces cellules sont recrutées par diverses chimiokines, comme l'éotaxine, l'interleukine-5, l'interleukine-8 ou les RANTES (pour *regulated upon activation, normal T cell expressed and secreted*), secrétées par les cellules structurales et inflammatoires activées (Tattersfield et al., 2002). Les cellules recrutées se lient par la suite aux molécules d'adhésion endothéliales, principalement la ICAM-1 (pour *intercellular adhesion molecule 1*) et la VCAM-1 (pour *vascular-cell adhesion molecule 1*), ce qui initie leur transmigration de la circulation sanguine vers le foyer d'inflammation dans la paroi bronchique (Schleimer et Bochner, 1998; Busse et Lemanske, 2001; Lemanske et Busse, 2003). Sous l'effet d'autres médiateurs inflammatoires tels les facteurs de survie, ces cellules infiltrées présentent une apoptose réduite, ce qui fait qu'elles persistent au foyer d'inflammation tout en libérant de nombreux autres médiateurs inflammatoires comme les leucotriènes, l'histamine, les facteurs nécrosant des tumeurs ou les cytokines de type Th2 (Bousquet et al., 2000), ce qui résulte en la persistance de l'inflammation. La recherche des divers médiateurs inflammatoires impliqués dans l'asthme est un domaine qui progresse continuellement. Certains articles de revue proposent une liste exhaustive, mais non définitive, des

cytokines et chimiokines impliquées dans l'asthme ainsi que de leurs effets (Chung et Barnes, 1999; Busse et Lemanske, 2001).

La progression de la réponse inflammatoire dans les voies respiratoires entraîne le développement de l'hyperréactivité bronchique et aggrave l'obstruction bronchique, ce qui mène à l'apparition des symptômes de l'asthme, à sa sévérité clinique, à sa chronicité, à ses exacerbations ainsi qu'à sa résistance à la thérapie anti-inflammatoire (Busse et al., 1999; Lemanske et Busse, 2003). Lorsque l'inflammation devient chronique, des changements structuraux se produisent dans l'arbre bronchique, lesquels constituent l'un des premiers signes que le remodelage s'installe (Folli et al., 2008). Le remodelage bronchique, accompagné de ses conséquences physiologiques néfastes et irréversibles dans les poumons, résulte en une diminution progressive de la fonction respiratoire, le développement d'une obstruction bronchique fixe et possiblement d'autres anomalies de la fonction pulmonaire qui peuvent résulter en une hyperréactivité bronchique permanente et ce, malgré un traitement anti-inflammatoire approprié (Busse et al., 1999; Lemanske et Busse, 2003; Pascual et Peters, 2005).

#### **1.4.3 Remodelage bronchique**

Bien que le remodelage de la paroi bronchique touche tous les sujets asthmatiques et soit variable d'un sujet à l'autre (Bousquet et al., 1999), les mécanismes responsables de sa progression ne sont pas encore bien compris (Broide, 2008). De plus, il n'est pas encore certain si le remodelage bronchique se produit suite à une inflammation chronique (comme cause secondaire) ou *a priori* dans les voies aériennes (comme cause primaire) (Pascual et Peters, 2005; Moore et Peters, 2007; Sumi et Hamid, 2007). En effet, de récentes observations obtenues de biopsies bronchiques d'enfants asthmatiques montrent la présence de certaines composantes de remodelage (Barbato et al., 2006; Kicic et al., 2006), laissant supposer que le remodelage serait une caractéristique majeure dans la pathogénèse de l'asthme et non le résultat de la chronicité ou de la sévérité (Moore et Peters, 2007; Sumi et Hamid, 2007). Néanmoins, la littérature s'accorde généralement sur un concept de remodelage bronchique qui résulterait de lésions répétées survenant dans les bronches suite aux événements inflammatoires persistants, accompagnées d'un processus mal contrôlé de réparation tissulaire qui permet normalement l'élimination ou le remplacement des cellules structurales altérées

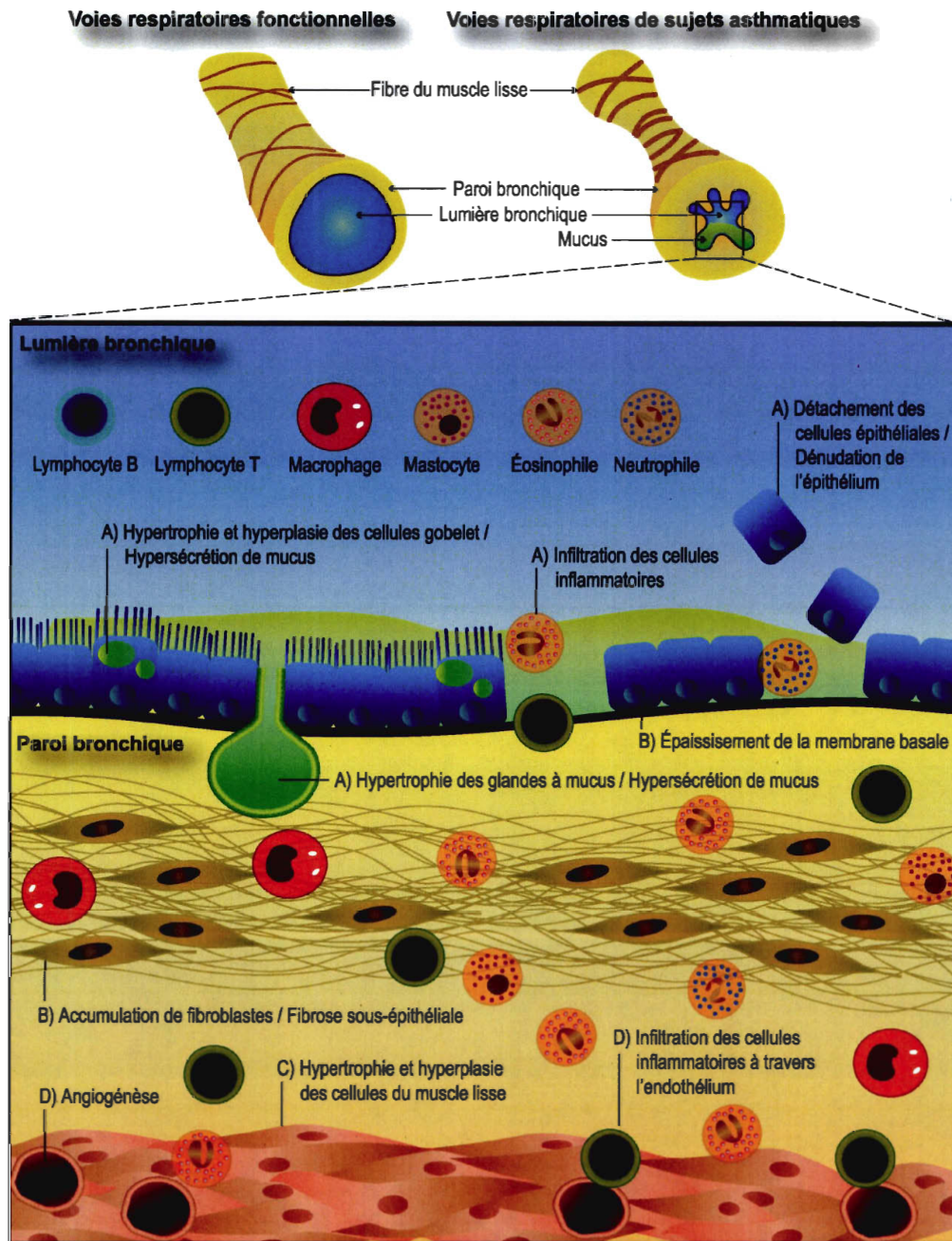
ou inflammatoires infiltrées (Bousquet et al., 2000; Broide, 2008). Un déséquilibre dans les processus qui contrôlent la survie, la prolifération et/ou l'apoptose de ces cellules pourrait également favoriser l'inflammation chronique et par conséquent, le remodelage bronchique (Benayoun et Pretolani, 2003). Par exemple, un défaut dans l'élimination des éosinophiles par apoptose pourrait favoriser leur accumulation incontrôlée et leur persistance dans les voies aériennes, ce qui contribuerait au maintien de l'inflammation chronique et des lésions tissulaires. En revanche, une apoptose accrue, associée à un déficit de prolifération des cellules épithéliales bronchiques, pourrait être à l'origine d'un défaut de réparation de l'épithélium, ce qui perpétuerait le contact entre les allergènes et les structures sous-épithéliales dans le maintien de l'inflammation et des lésions tissulaires (Benayoun et Pretolani, 2003). Il y a également plusieurs autres observations que le remodelage bronchique peut faire suite à une inflammation bronchique (Laprise et al., 1999; Chu et al., 2006; Holgate, 2008).

Sur le plan histopathologique, le remodelage exerce ses effets sur l'épithélium bronchique, les structures sous-épithéliales, le muscle lisse ainsi que sur les vaisseaux sanguins bronchiques et implique une variété de médiateurs inflammatoires, de cytokines et de facteurs de croissance sécrétés par les cellules structurales et infiltrées (Tagaya et Tamaoki, 2007). Le remodelage se caractérise donc par différentes composantes, présentées de façon schématique dans la Figure 1 et brièvement décrites dans les prochains paragraphes.

Au niveau de l'épithélium, l'activation des cellules épithéliales, consécutive aux événements inflammatoires, provoque la libération d'une variété de cytokines inflammatoires et de facteurs de croissance, le principal étant le TGF- $\beta$  (pour *Transforming Growth Factor beta*), qui induisent des changements néfastes sur les composantes tissulaires (Tagaya et Tamaoki, 2007), comme un détachement de cellules épithéliales de la membrane basale (Sumi et Hamid, 2007). L'augmentation de l'espace ainsi créé entre les cellules épithéliales basales résulte en une dénudation de l'épithélium (Tagaya et Tamaoki, 2007), agrandissant la surface de contact avec les allergènes inhalés et favorisant la transmigration des cellules inflammatoires à travers la membrane. Parmi les autres changements observés au niveau de l'épithélium se retrouvent l'hypertrophie (augmentation de la taille) des glandes à mucus ainsi que l'hypertrophie et l'hyperplasie (augmentation du nombre) des cellules épithéliales



glandulaires (cellules à gobelet), phénomènes responsables d'une hypersécrétion de mucus (Davies et al., 2003; Lemanske et Busse, 2003).



**Figure 1 : Représentation schématique des principaux mécanismes de remodelage se produisant dans les voies respiratoires des sujets asthmatiques ainsi que des principales cellules structurales et inflammatoires impliquées.** Suivant les événements inflammatoires (section 1.4.2), la paroi bronchique subit des modifications au niveau A) de l'épithélium, B) de la zone sous-épithéliale, C) du muscle lisse et D) des vaisseaux sanguins. La description complète, incluant les principaux médiateurs impliqués et les références, est disponible dans le texte de la section 1.4.3.

Au niveau de la zone sous-épithéliale, le remodelage se caractérise par un épaissement apparent de la membrane basale causé par l'accumulation des fibroblastes activés par le TGF- $\beta$  ou l'interleukine (IL)-11 produits par les cellules épithéliales et les éosinophiles activés (Sumi et Hamid, 2007; Warner et Knight, 2008). Ces événements fibrotiques sont le résultat d'une augmentation du dépôt des protéines de la matrice extracellulaire tels le collagène, la ténascine, la fibronectine, la laminine et les protéoglycans (Sumi et Hamid, 2007; Tagaya et Tamaoki, 2007). La déposition de ces protéines dans les tissus est contrôlée par une balance entre l'expression des métalloprotéases, qui dégradent ces protéines, et de leurs inhibiteurs (TIMPs pour *tissue inhibitors of metalloproteases*) (Bousquet et al., 2000). La surexpression des TIMPs causerait le dépôt des protéines de la matrice extracellulaire par inhibition de leur dégradation (Tagaya et Tamaoki, 2007). Finalement, la perte d'élasticité observée dans les voies aériennes des sujets asthmatiques serait le résultat de la lyse des fibres élastiques présentes dans la zone sous-épithéliale, ce qui contribuerait à rigidifier la paroi bronchique (Bousquet et al., 2000).

Au niveau du muscle lisse bronchique, le remodelage résulte en une augmentation de la masse totale due à une hypertrophie et à une hyperplasie des cellules musculaires (Davies et al., 2003; Lemanske et Busse, 2003; Bosse et al., 2008). Ces cellules participent également au remodelage par leur sécrétion de cytokines, de facteurs de croissance, de protéines de la matrice extracellulaire, ainsi que par l'expression de molécules d'adhésion cellulaire et de facteurs de co-stimulation (pour une liste complète, consulter (Sumi et Hamid, 2007; Tagaya et Tamaoki, 2007)). Ainsi, en plus d'avoir un rôle important dans la contraction exagérée observée dans l'hyperréactivité bronchique, le muscle lisse contribue à l'orchestration et à la perpétuation de l'inflammation et du remodelage dans les voies aériennes asthmatiques (Busse et al., 1999). L'hypertrophie pourrait être le résultat de bronchospasmes répétés, d'une activité myogénique incontrôlée ou être le reflet d'une synthèse protéique accrue (Bousquet et al., 2000) et son effet pourrait se faire sentir sur la contractilité des fibres (Sumi et Hamid, 2007). Quant à l'hyperplasie, elle est la conséquence d'une prolifération excessive des cellules musculaires lisses résultant d'une stimulation de la mitose combinée à une suppression de l'apoptose (Sumi et Hamid, 2007) induite par différents systèmes de récepteurs comme les récepteurs des tyrosine kinases, les récepteurs

couplés aux protéines G, les récepteurs des cytokines ainsi que leurs ligands (Tagaya et Tamaoki, 2007; Bosse et al., 2008).

Pour terminer, le remodelage au niveau des vaisseaux sanguins des voies aériennes s'observe par une angiogénèse (croissance de nouveaux vaisseaux sanguins à partir de vaisseaux existants) (Tagaya et Tamaoki, 2007). La vascularité ainsi accrue s'accompagne d'une augmentation de l'expression de facteurs de croissance angiogéniques et de perméabilité, le principal étant le VEGF (pour *vascular endothelium growth factor*) (Warner et Knight, 2008), ce qui a pour conséquence l'induction d'un œdème tissulaire et l'exsudation des protéines plasmatiques (Sumi et Hamid, 2007). En combinaison avec la sécrétion des molécules d'adhésion ICAM-I et VCAM-I, cette prolifération des capillaires sanguins contribue à une augmentation du recrutement des cellules inflammatoires circulantes (et à leur transmigration à travers l'endothélium), ce qui perpétue l'inflammation et le remodelage déjà entamés (Busse et Lemanske, 2001; Lemanske et Busse, 2003).

#### **1.4.4 Asthme allergique**

L'allergie se définit comme un état clinique d'hypersensibilité d'origine héréditaire, prédisposant l'individu à l'apparition de symptômes allergiques lorsqu'il est exposé à certaines substances, les allergènes, qui ne déclencheraient aucune réaction chez des personnes non sensibilisées (Galli et al., 2008). En fait, la réaction allergique est considérée comme une manifestation exagérée ou inappropriée du mécanisme de défense immunitaire de l'organisme contre ces allergènes (Dubé et Boulet, 1996). Tout comme dans l'asthme non-allergique, les mécanismes physiopathologiques de l'asthme allergique impliquent l'inflammation et du remodelage des voies aériennes tels que décrits précédemment. La distinction vient de la spécificité des facteurs inducteurs et déclencheurs en cause, c'est-à-dire les allergènes.

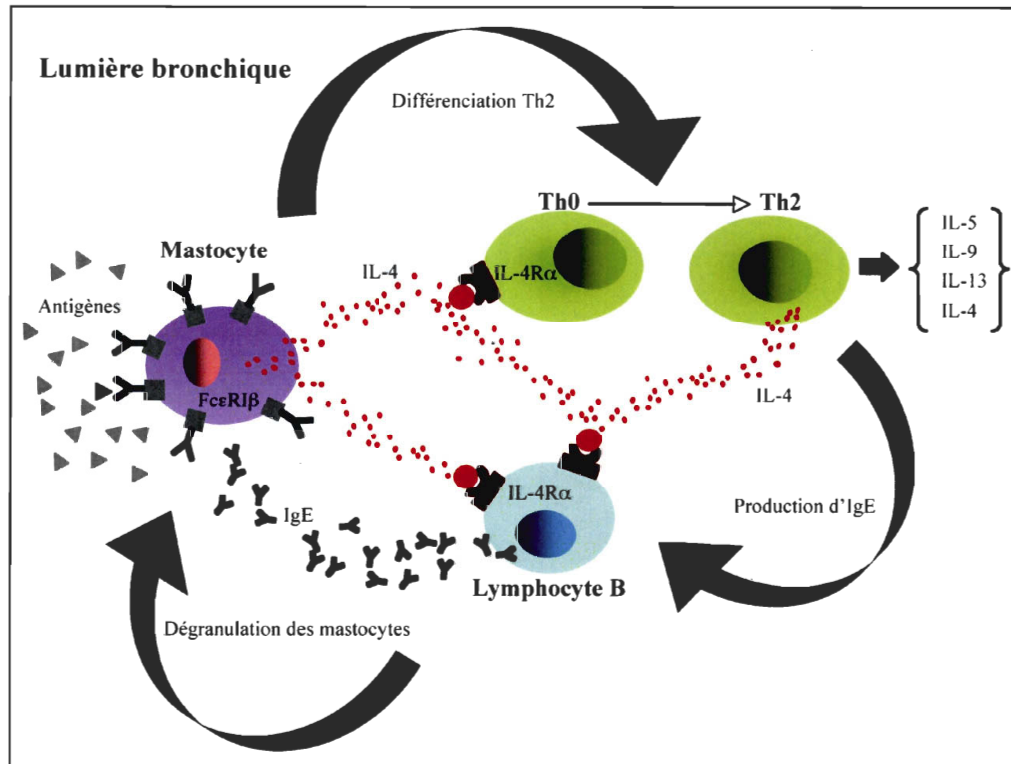
Le développement de l'asthme allergique débute généralement au cours de l'enfance et se traduit par l'élaboration d'une mémoire immunologique spécifique (immunité acquise) contre un allergène (ou un antigène) inhalé (Holgate, 1999; Holt et al., 1999). Ce processus implique en premier lieu les cellules dendritiques qui, une fois mises en contact avec ces allergènes, migrent jusqu'aux ganglions lymphatiques où elles agissent

en tant que cellules présentatrices d'antigène aux molécules du complexe majeur d'histocompatibilité situées sur les lymphocytes T précurseurs (Th0) (Busse et Lemanske, 2001; Galli et al., 2008). Les lymphocytes Th0 ainsi activés se différencient en lymphocytes Th1 (si exposés simultanément à l'IL-12) ou en lymphocytes Th2 (si exposés simultanément à l'IL-4) (Maddox et Schwartz, 2002). Les lymphocytes Th1 ou Th2 jouent un rôle important, et distinct selon le type, dans le développement, l'initiation et la régulation de la réponse immunitaire qui sous-tend la pathogénèse de l'asthme allergique (Maddox et Schwartz, 2002). Les lymphocytes Th1 activés secrètent principalement l'IL-2, l'IL-3, l'interféron- $\gamma$ , le GM-CSF (pour *granulocyte-macrophage colony-stimulating factor*) et le facteur nécrosant des tumeurs (TNF)- $\alpha$ , et sont impliqués dans l'hypersensibilité retardée et la réponse immunitaire cytotoxique (Maddox et Schwartz, 2002; Packard et Khan, 2003). Quant aux lymphocytes Th2, ils secrètent principalement l'IL-3, l'IL-4, l'IL-5, l'IL-6, l'IL-9, l'IL-10, l'IL-13 ainsi que le GM-CSF, et régulent la réaction allergique par l'activation des lymphocytes B (Packard et Khan, 2003) ainsi que par le recrutement et l'activation des mastocytes (par l'IL-4 et l'IL-9), des éosinophiles (par l'IL-3, l'IL-5 et le GM-CSF) et des basophiles (par l'IL-3) (Holgate et al., 2000; Holgate, 2008). Pour ce qui est des lymphocytes B, une fois activés, ils produisent une quantité anormalement élevée d'IgE en réponse aux allergènes, ce qui contribue à créer la mémoire immunologique contre cet allergène, c'est-à-dire la sensibilisation (Galli et al., 2008). L'inhalation subséquente de ces allergènes chez les individus sensibilisés implique d'autres cellules présentatrices d'antigènes comme les mastocytes et les basophiles qui, grâce à leurs récepteurs de haute affinité aux IgE (Fc $\epsilon$ RI $\beta$ ), retiennent à leur surface membranaire ces IgE spécifiques, ce qui permet de reconnaître et d'intercepter les allergènes immédiatement suite à leur inhalation ainsi que d'enclencher la réponse inflammatoire (Holt et al., 1999; Bradding et al., 2006; Holgate, 2008). Par conséquent, l'activation de la cascade de signalisation membranaire et intracellulaire initie la dégranulation des mastocytes, libérant une variété de médiateurs tels les cytokines (IL-4, IL-5, IL-13 ou TGF- $\beta$ ), l'histamine, la prostaglandine (PG) D<sub>2</sub> ou le leucotriène (LT) C<sub>4</sub> (Busse et Lemanske, 2001; Packard et Khan, 2003). Ces médiateurs participent aux divers phénomènes physiologiques et aux symptômes aigus de la réponse asthmatique immédiate (bronchospasme, œdème, obstruction du passage de l'air) (Bradding et al., 2006) ainsi qu'à ceux de la réponse tardive (inflammation, obstruction et hyperréactivité des voies

aériennes) (Busse et Lemanske, 2001). C'est au cours de cette réponse tardive que les éosinophiles activés deviennent les cellules majoritaires et principaux acteurs dans la mise en place et le maintien de l'inflammation chronique et de ses conséquences fonctionnelles (Maddox et Schwartz, 2002), l'éosinophilie étant reconnue comme la principale caractéristique de l'inflammation des voies respiratoires asthmatiques allergiques (Kay, 2005). L'IL-5 est la cytokine essentielle permettant la différenciation, la maturation, le recrutement et l'activation des éosinophiles dans les sites inflammatoires (Holgate, 2008). Une fois activés, les éosinophiles libèrent à leur tour une série de médiateurs inflammatoires incluant des cytokines de types Th2, du LTC<sub>4</sub>, des facteurs de croissance et des protéines basiques, leur conférant un rôle prépondérant dans le remodelage des voies aériennes, en particulier dans la dénudation de l'épithélium, dans la production de mucus, dans la fibrose sous-épithéliale et dans la bronchoconstriction (Pascual et Peters, 2005; Broide, 2008; Galli et al., 2008; Holgate, 2008), intensifiant l'hyperréactivité bronchique et augmentant la sévérité clinique de l'asthme (Busse et Lemanske, 2001). La Figure 2 résume une boucle de régulation immunitaire se produisant chez un individu asthmatique sensibilisé à un allergène mettant en scène un mastocyte activé, les lymphocytes T et B ainsi que de l'IL-4, une des principales cytokines pro-inflammatoires reconnue dans l'asthme allergique avec l'IL-13 et l'IL-5 (James, 2005).

Le profil traditionnel de la pathogénèse de l'asthme allergique est basé sur l'hypothèse d'une augmentation de la réponse immunitaire de type Th2 combinée à une diminution de la réponse immunitaire de type Th1 (Walter et al., 2002; Busse et Rosenwasser, 2003). L'allergie résulterait donc d'un déséquilibre du système immunitaire acquis en faveur de l'immunité humorale de type Th2 pro-allergique au détriment de l'immunité de type Th1 anti-allergique (Packard et Khan, 2003; Nauta et al., 2008). D'ailleurs, ce déséquilibre s'observe par un excès de produits Th2 et une diminution de produits Th1 présents dans les bronches des sujets asthmatiques (Walter et al., 2002; Busse et Rosenwasser, 2003). Bien que l'origine de ce déséquilibre dans le développement du système immunitaire chez les personnes asthmatiques allergiques ne soit pas caractérisée, l'hypothèse de l'hygiène est l'une des propositions globalement acceptée pour expliquer ce phénomène (Bresciani et al., 2005; Nauta et al., 2008). Cette hypothèse stipule qu'une exposition réduite aux infections (ou à tout autre facteur favorisant le développement de l'immunité protectrice de type Th1) durant l'enfance,

combinée à une exposition aux facteurs favorisant le phénotype Th2 pro-allergique (style de vie urbain, environnement stérilisé, utilisation d'antibiotiques, etc.), est associée à un plus grand risque de développer des maladies allergiques plus tard dans la vie de l'individu (Busse et Lemanske, 2001; Bresciani et al., 2005).



**Figure 2 : Représentation schématique d'une boucle de régulation inflammatoire présente dans les voies respiratoires de sujets asthmatiques allergiques.** La liaison de l'antigène à son anticorps (IgE), situé sur le récepteur de haute affinité pour les IgE (FcεRIβ) des mastocytes, stimule la dégranulation de ces derniers afin de libérer l'IL-4. L'IL-4 induit par la suite la différenciation des lymphocytes Th0 en lymphocytes Th2 sécréteurs de médiateurs inflammatoires tels l'IL-4 lui-même, l'IL-13, l'IL-5 ou l'IL-9. L'IL-4 exerce également son action sur les lymphocytes B où il induit la commutation et la production d'IgE, lesquels iront ensuite s'attacher sur leurs récepteurs FcεRIβ des mastocytes, ce qui a pour effet d'augmenter la sensibilité aux antigènes. Les références utilisées pour construire ce schéma se retrouvent dans le texte de la section 1.4.4.

Suivant cette revue de littérature générale sur la physiopathologie de l'asthme, il est possible de constater l'implication d'une grande variété de cellules immunitaires et structurales communiquant entre elles et avec leur milieu au moyen des médiateurs qu'elles sécrètent. Chacun de ces médiateurs, et leur récepteur, est une molécule spécifique traduite d'un gène spécifique dont l'expression est régulée par l'activation d'un réseau de signalisation intracellulaire et de facteurs de transcription. Ces cascades

de signalisation résultent donc en l'expression de plusieurs gènes qui agissent de concert, sous l'influence de l'environnement, dans l'établissement d'une réponse inflammatoire et immunitaire déséquilibrée responsable du développement de l'asthme. Bien que les facteurs environnementaux et leurs interactions avec les gènes sont d'une importance fondamentale dans l'étiologie et la physiopathologie de l'asthme, la présente thèse est consacrée à l'identification des gènes de l'asthme. Par conséquent, la mise en contexte des hypothèses et objectifs de recherche nécessite une revue de littérature sur la génétique et l'épidémiologie génétique relatives à l'asthme, ce qui fera l'objet des deux prochains chapitres.

## Chapitre 2

### Génétique de l'asthme

Plusieurs évidences indiquent que l'asthme est une maladie sous l'influence de facteurs génétiques (Bleecker et al., 1997; Cookson, 1999). D'ailleurs, le fait que la génétique joue un rôle dans la pathogénèse de l'asthme est reconnu depuis longtemps, les premières observations d'une agrégation familiale remontant au début du XX<sup>e</sup> siècle (Cooke et van der Veer, 1916). Depuis, de nombreuses études familiales, incluant des études de jumeaux et des analyses de ségrégation, ont été menées afin de mesurer la contribution relative des gènes dans la susceptibilité à l'asthme et de démontrer, hors de tout doute, l'implication de déterminants génétiques dans l'expression clinique de cette affection.

#### 2.1 Agrégation familiale

Le degré de concentration familiale d'une maladie peut être évalué en comparant sa prévalence observée parmi les fratries (sujets apparentés) des sujets atteints à la prévalence observée chez les apparentés des sujets non-atteints (sujets témoins provenant de la population générale) (Sibbald et al., 1980). En d'autres termes, une agrégation familiale pour l'asthme sera démontrée si le risque de faire de l'asthme est significativement plus grand chez les apparentés d'un sujet asthmatique que chez les apparentés d'un sujet témoin (Steinke et al., 2008). Plus ce risque relatif (risque des apparentés aux asthmatiques divisé par le risque des apparentés aux témoins) sera grand, plus grande sera la contribution génétique (Sandford et al., 1996). Pour ce qui est de l'asthme, un risque relatif d'environ 5-6 a été calculé, signifiant que les personnes apparentées à un sujet asthmatique ont 5 à 6 fois plus de chances de faire de l'asthme, ce qui suggère une forte contribution génétique pour la maladie (Sandford et al., 1996). Le risque peut également être évalué en fonction du statut asthmatique des parents. Ainsi, la prévalence de l'asthme et des phénotypes liés à l'asthme (atopie, hyperréactivité bronchique, taux d'IgE, etc.) est plus importante chez les enfants de parents atteints comparativement aux enfants de parents non-atteints (Sibbald et al., 1980; Laprise et Boulet, 1996). Il a même été estimé que le risque qu'un enfant



développe de l'asthme est trois fois plus élevé si l'un de ses parents est asthmatique et ce risque double si les deux parents sont atteints (Illig et Wjst, 2002).

## **2.2 Étude de jumeaux**

Les études de jumeaux sont une autre approche utilisée pour mesurer la contribution relative des facteurs génétiques dans la susceptibilité à l'asthme (Sandford et al., 1996; Borish, 1999). Dans ces études, le taux de concordance de la maladie entre les jumeaux identiques (monozygotes, partageant 100 % de leurs caractéristiques génétiques) est comparé avec celui des jumeaux non-identiques (dizygotes, partageant approximativement 50 % de leurs caractéristiques génétiques) (Los et al., 2001). Si la maladie étudiée est influencée par des facteurs génétiques, le taux de concordance entre les monozygotes devrait être plus grand que chez les dizygotes (Sandford et al., 1996). Ces taux de concordance permettent également d'évaluer l'héritabilité de la maladie (paramètre calculé selon diverses méthodes statistiques) afin d'estimer la contribution relative provenant des facteurs génétiques par opposition aux facteurs environnementaux : 0 % signifiant aucune contribution génétique (donc la maladie serait d'origine environnementale) et 100 % indiquant une prédisposition génétique complète (comme dans le cas des maladies monogéniques) (Borish, 1999). Plusieurs études de jumeaux ont déjà démontré que le taux de concordance pour l'asthme est significativement plus élevé chez les jumeaux monozygotes (de 30 % à 80 %) que chez les jumeaux dizygotes (de 4 % à 45 %) (Falliers et al., 1971; Hopp et al., 1984; Duffy et al., 1990; Hanson et al., 1991; Nieminen et al., 1991). Quant à son héritabilité, elle a été estimée entre 50 % à 75 % (Duffy et al., 1990; Laitinen et al., 1998; Skadhauge et al., 1999). Comme l'on suppose que les jumeaux partagent le même environnement, ces taux de concordance et l'héritabilité suggèrent aussi une forte contribution génétique pour l'asthme (Duffy et al., 1990; Cookson, 1999; Huss et Huss, 2000).

## **2.3 Analyse de ségrégation**

Les analyses de ségrégation permettent de déterminer le mode de transmission génétique, le plus simple, qui explique la ségrégation (ou la distribution) d'un phénotype au sein d'une famille (Sandford et al., 1996). Les analyses de ségrégation sont une méthode analytique qui utilise les structures familiales (généalogies) pour

établir une relation statistique entre la transmission du phénotype observée dans la descendance et celle qui serait attendue selon les modèles de transmission génétique théoriques (par exemple autosomal récessif, autosomal dominant) ou complexes (par exemple polygénique, multifactoriel) (Sandford et al., 1996; Steinke et al., 2008). Dans l'asthme, les analyses de ségrégation confirment l'importance de la contribution génétique, mais ne sont pas concluantes quant au mode de transmission (Los et al., 2001). En fait, depuis Cooke et Van der Veer (1916), qui ont suggéré un mode de transmission autosomal dominant pour l'asthme, à peu près tous les modes de transmission lui ont été attribués (Sandford et al., 1996). Encore aujourd'hui, le mode de transmission génétique de l'asthme demeure toujours inconnu (Ober et al., 2001). Donc, pour définir la génétique de l'asthme, il est consensuellement admis par la communauté scientifique que l'asthme est une maladie polygénique et multifactorielle (Palmer et Cookson, 2000; Los et al., 2001; Illig et Wjst, 2002), c'est-à-dire sous l'influence de plusieurs facteurs génétiques (dont le nombre et l'identité ne sont pas encore définitifs) et de plusieurs facteurs environnementaux qui agissent les uns avec les autres dans des réseaux d'interaction complexes qui demeurent également à documenter. L'asthme est donc un « trait complexe », tout comme le sont la majorité des autres maladies communes à haute prévalence comme les cancers, les maladies cardiovasculaires, la schizophrénie ou l'arthrite rhumatoïde, pour ne nommer que ces exemples.

#### **2.4 L'asthme : un trait complexe**

L'asthme ne présente donc pas un mode de transmission mendélien classique récessif ou dominant normalement attribué aux maladies monogéniques (c'est-à-dire lorsqu'un seul gène est responsable) comme la fibrose kystique (Sandford et al., 1996). Ainsi, l'ensemble de la littérature considère l'asthme comme un trait complexe puisque son expression clinique est dictée par l'effet combiné de plusieurs gènes dans des conditions environnementales particulières (Elston, 1995; Sandford et al., 1996; Silverman et Palmer, 2000; Cookson, 2002). En général, la complexité vient de l'absence de correspondance entre le phénotype (manifestations détectables d'un gène, par exemple l'asthme ou chacune de ses conditions cliniques associées) et le génotype (ensemble des allèles – formes possibles – d'un gène porté par un individu). Ainsi, il est souvent impossible de trouver un gène qui montre une parfaite coségrégation (transmission)

avec un trait complexe, soit parce que le génotype résulte en différents phénotypes, soit parce que différents génotypes résultent en un seul phénotype (Lander et Schork, 1994). En plus de la complexité des interactions avec d'autres gènes et/ou l'environnement, plusieurs raisons peuvent expliquer ce phénomène : 1) la pénétrance incomplète du gène, c'est-à-dire qu'un individu porteur d'un génotype donné ne présentera pas le phénotype attendu au moment où l'évaluation clinique est faite ; 2) les phénocopies, c'est-à-dire qu'un individu non porteur du génotype présentera le phénotype ; 3) l'hétérogénéité génétique, c'est-à-dire que plusieurs gènes différents ainsi que leurs différentes combinaisons peuvent résulter au même phénotype ; 4) la transmission polygénique, c'est-à-dire l'effet additif de gènes qui doivent être transmis simultanément pour voir apparaître le phénotype ; et finalement 5) la pléiotropie, c'est-à-dire qu'un gène ou une combinaison de plusieurs gènes influencent différents phénotypes (Lander et Schork, 1994; Sandford et al., 1996; Los et al., 2001; Hoffjan et Ober, 2002).

En résumé, les traits complexes sont causés par l'interaction d'une multitude de gènes et de facteurs environnementaux qui forment un gradient de susceptibilité ou de protection à développer et présenter la maladie (Weeks et Lathrop, 1995). Pour toutes les raisons mentionnées ci-haut, sans oublier la nature complexe de l'asthme *per se* et de ses nombreux phénotypes intermédiaires (qui peuvent résulter de causes génétiques communes ou spécifiques), il est difficile d'identifier les déterminants génétiques de l'asthme et leurs interactions. C'est par les études d'épidémiologie génétique, avec ses différentes approches méthodologiques toujours en évolution, qu'il est possible d'accroître les connaissances sur les facteurs génétiques de l'asthme, lesquelles ont augmenté substantiellement depuis les dernières décennies. Ainsi, il reste encore beaucoup de chemin à faire afin de connaître tous les gènes impliqués dans la physiopathologie de l'asthme et c'est dans ce contexte que s'inscrit le présent projet de recherche. Le prochain chapitre aborde en détail les approches employées en épidémiologie génétique et permettra l'introduction et une meilleure compréhension des articles inclus dans la présente thèse (Chapitres 5 à 8).

## Chapitre 3

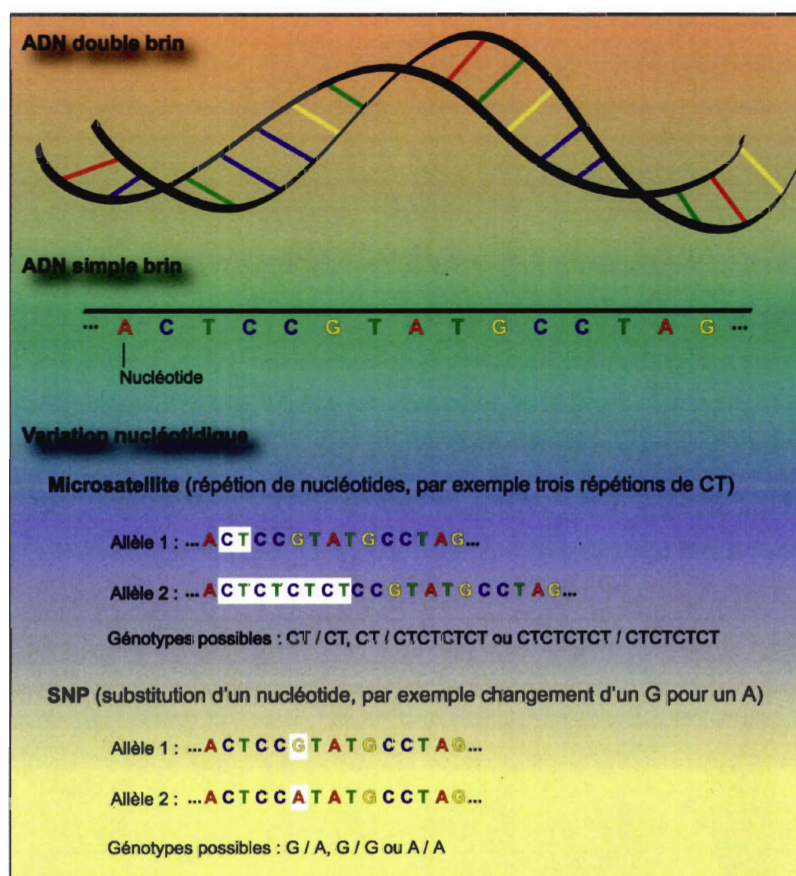
### Épidémiologie génétique de l'asthme

Par convention, l'épidémiologie génétique est une discipline scientifique créée par la fusion des champs de recherche de l'épidémiologie traditionnelle (étude de la distribution, des causes environnementales et des facteurs de risque associés à une maladie) ainsi que de la génétique (étude de l'hérédité d'une maladie et de ses facteurs génétiques) et de ses sous-disciplines comme la génétique moléculaire (étude des marqueurs génétiques qui seront définis dans ce chapitre et de leur forme allélique) et la génétique des populations (étude des marqueurs génétiques dans différentes populations) (Permutt et al., 2005). Dans l'étude d'un trait complexe comme l'asthme, l'épidémiologie génétique a pour buts d'identifier les déterminants génétiques qui prédisposent à la maladie et d'expliquer comment ces déterminants confèrent leur effet et dans quels contextes environnementaux (Burton et al., 2005). Pour ce faire, l'épidémiologie génétique repose sur deux types d'études essentielles, à savoir les études de liaison et les études d'association (Eley et Rijsdijk, 2005), qui représentent les deux principales sections de ce chapitre.

#### 3.1 Étude de liaison

L'étude de liaison (aussi appelée analyse de liaison ou analyse de linkage) sert à identifier des régions chromosomiques transmises dans les familles en relation avec un phénotype d'intérêt. Ainsi, à partir de l'acide désoxyribonucléique (ADN) de chacun des membres des familles, il est possible d'obtenir le génotype (selon divers procédés de génotypage en laboratoire) de marqueurs génétiques spécifiques (Vercelli, 2008). Chacun de ces marqueurs génétiques représente un locus (plu. loci - endroit spécifique sur un chromosome) polymorphique, c'est-à-dire un site où l'ADN présente au moins deux allèles (Sandford et al., 1996). Les marqueurs génétiques les plus couramment employés dans les analyses de liaison sont les marqueurs microsatellites (Figure 3) dont les formes alléliques sont représentées par les différentes longueurs des nucléotides répétés (Sandford et al., 1996). Grâce à la complétion de la séquence du génome humain (International Human Genome Sequencing Consortium, 2004), la localisation de ces marqueurs moléculaires est connue, ce qui permet d'en sélectionner un nombre

déterminé espacé à intervalles réguliers pour couvrir l'ensemble de tous les chromosomes et par conséquent, le génome en entier (Vercelli, 2008). Par un tel criblage génomique (ou *genome-wide scan*), l'analyse de liaison permet l'étude de la transmission de ces marqueurs en fonction du statut d'atteint dans les familles. Si un des allèles d'un des marqueurs est significativement plus souvent retrouvé parmi les atteints en comparaison aux non-atteints, ce marqueur sera « lié » à la maladie, donc co-transmis avec la maladie (Eley et Rijdsdijk, 2005). Comme la séquence d'ADN couverte par un tel locus peut s'étendre jusqu'à 30 millions de paires de bases nucléotidiques et contenir des centaines de gènes, ce locus n'est pas nécessairement causal pour la maladie (Vercelli, 2008). Il est plutôt un indicateur de la présence d'un ou de plusieurs gènes d'importance dans cette région chromosomique qui ne pourront être identifiés qu'avec l'utilisation de marqueurs encore plus rapprochés (Eley et Rijdsdijk, 2005).



**Figure 3 : Représentation schématique de la molécule d'ADN double brin et des deux principaux types de marqueurs génétiques utilisés en épidémiologie génétique.** SNP = Polymorphisme d'un seul nucléotide, A = Adénine, C = Cytosine, G = Guanine et T = Thymine.

Depuis le premier criblage génomique réalisé pour l'asthme et l'atopie (Daniels et al., 1996), au moins 18 autres ont été complétés, ayant fait l'objet de plusieurs articles de revue (Hoffjan et Ober, 2002; Blumenthal, 2005; Steinke et al., 2008). De nombreuses régions chromosomiques liées à l'asthme ou à ses phénotypes intermédiaires ont ainsi été identifiées, lesquelles sont variables d'une publication à l'autre dû aux différences dans la conception des études, les populations étudiées et l'hétérogénéité clinique de l'asthme (Hoffjan et Ober, 2002). Les régions chromosomiques les plus rapportées, en plus d'avoir été répliquées dans plus d'une étude, sont situées sur les chromosomes 1, 2, 3, 4, 5, 6, 7, 11, 12, 13, 14, 15, 19, 20 et 21. Plus précisément, pour l'asthme il s'agit des régions 1p, 2q, 4q, 5q, 6p, 12q, 13q, 14q, 19q, 20p et 21q ; pour l'atopie, les régions 3q, 4q, 6p, 11q, 13q, 17q et 20p ; pour le taux d'IgE total, les régions 2q, 3q, 5q, 6p, 7q, 13q et 12q ; et finalement, pour le taux d'éosinophile sanguin élevé, la région 15q (consulter (Hoffjan et Ober, 2002; Blumenthal, 2005) pour les références à chacune des études originales). Ces régions contiennent des gènes impliqués dans les processus de régulation immunologique, inflammatoire et de remodelage qui ont potentiellement un rôle dans la pathogénèse de l'asthme (Hakonarson et Halapi, 2002).

L'analyse de liaison a été et est toujours utile dans l'identification des gènes responsables des maladies héréditaires, surtout celles à transmission mendélienne classique comme la fibrose kystique qui implique un seul gène (Rommens et al., 1989). L'application des analyses de liaison aux maladies complexes n'est pas aussi efficace en raison des nombreux paramètres de complexité déjà mentionnés (Chapitre 2, section 2.4) (Lander et Schork, 1994; Palmer et Cookson, 2000) et aussi dû au fait que les traits complexes impliquent plusieurs gènes à effet individuel faible ou modeste plus difficile à détecter (Rao, 2008). Néanmoins, le principal avantage lié à l'utilisation de l'analyse de liaison dans l'asthme est la découverte de régions chromosomiques contenant des gènes et des voies biologiques encore non soupçonnés dans sa physiopathologie (Ober et Hoffjan, 2006). Par conséquent, ces gènes, dû à leur localisation dans une région chromosomique liée à l'asthme, deviennent des « gènes candidats » pour les études d'association subséquentes.

### 3.2 Étude d'association

L'étude d'association a pour principal objectif de détecter l'association génétique entre le phénotype et un ou des gènes candidats particuliers. Ainsi, l'étude d'association entre le phénotype et le génotype de ces gènes est la méthode statistique la plus puissante pour identifier les déterminants génétiques de susceptibilité ou de protection qui sous-tendent les traits complexes (Risch et Merikangas, 1996).

Par contraste avec les analyses de liaison qui emploient les marqueurs microsatellites, les marqueurs génétiques les plus employés dans les études d'association sont les polymorphismes d'un seul nucléotide (SNP pour *Single Nucleotide Polymorphism*) (Figure 3). Parmi les 3,3 milliards (approximativement) de nucléotides qui composent la totalité de l'ADN retrouvée dans le génome humain, il a été estimé à près de 7 millions le nombre de SNPs uniformément distribués et présentant une fréquence de l'allèle mineur plus grande que 5 % (Rao, 2008). Ce nombre est évalué à 10 millions si l'on considère une fréquence de l'allèle mineur de 1 % (Reich et al., 2003; Thorisson et al., 2005). Donc, tout comme dans l'analyse de liaison, les études d'association peuvent être effectuées à l'échelle génomique. On parle alors d'études d'association génomiques (GWAS pour *Genome-Wide Association Study*) qui analysent plusieurs milliers de SNPs à la fois et par conséquent, la presque totalité des gènes retrouvés dans le génome. À l'inverse, les études d'association classiques reposent sur l'approche par gènes candidats, c'est-à-dire où les SNPs analysés sont limités à un ou quelques gènes, préalablement sélectionnés dépendant d'une hypothèse de départ. Peu importe que l'approche soit génomique ou par sélection de gènes candidats, les SNPs sélectionnés sont génotypés dans des échantillons populationnels de sujets asthmatiques pour permettre la vérification de l'association au moyen de diverses analyses statistiques. Au bout du compte, si une association statistique positive est trouvée avec un ou plusieurs SNPs d'un gène donné, le gène sera associé à la maladie et le statut de gène de susceptibilité ou de protection lui sera conféré. Toutefois, l'association devra être reproduite dans au moins une autre étude d'association indépendante (appelée dans ce cas étude de répliation). Les études d'association sont donc composées par diverses approches méthodologiques réunies dans le diagramme conceptuel de la Figure 4 et qui feront l'objet des prochaines sous-sections.

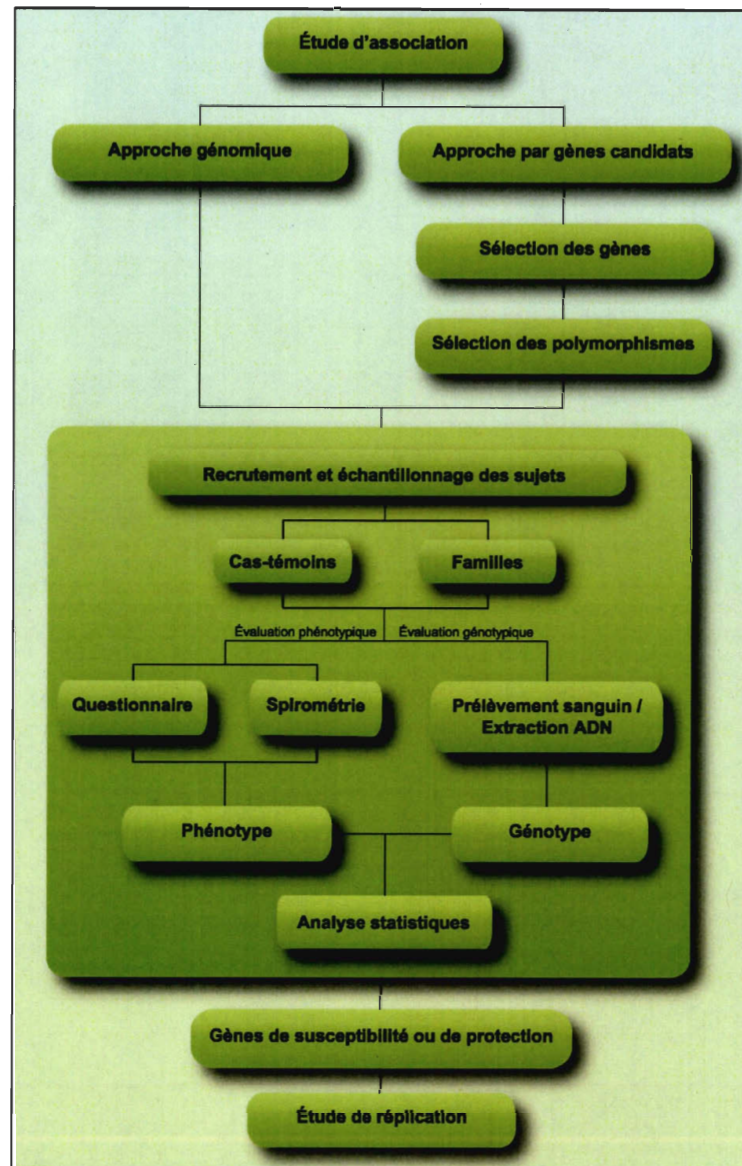


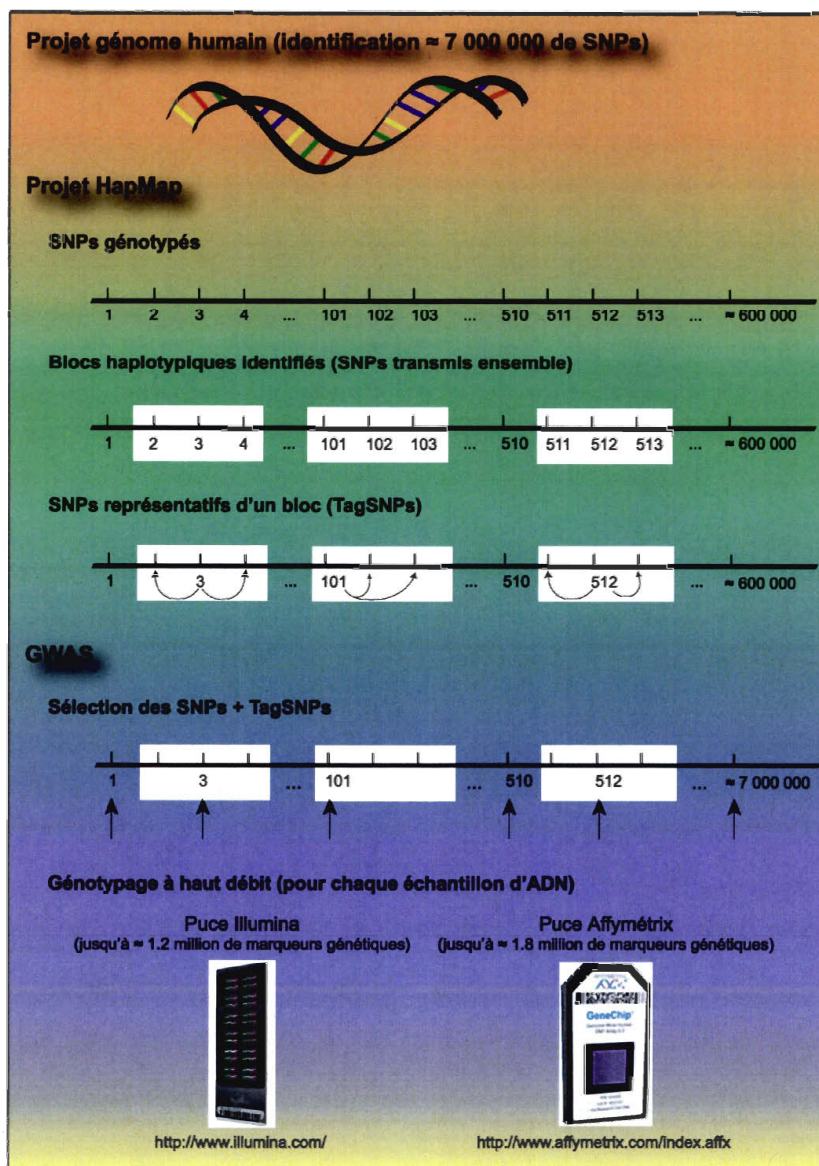
Figure 4 : Diagramme conceptuel représentant les deux principales approches et les diverses étapes nécessaires à la réalisation d'une étude d'association génétique.

### 3.2.1 Approche génomique

Les GWAS sont une révolution en épidémiologie génétique et ont été rendues possible grâce à la réalisation de deux projets d'envergure internationale, le projet Génome Humain (International Human Genome Sequencing Consortium, 2004) et le projet HapMap (International HapMap Consortium, 2003; International HapMap Consortium, 2005; Thorisson et al., 2005), ainsi que par les avancées biotechnologiques, notamment celles au niveau du génotypage à débit très élevé (Grant et Hakonarson, 2008). Ces



progrès scientifiques sont mis en relation dans la Figure 5 et décrits dans les prochains paragraphes.

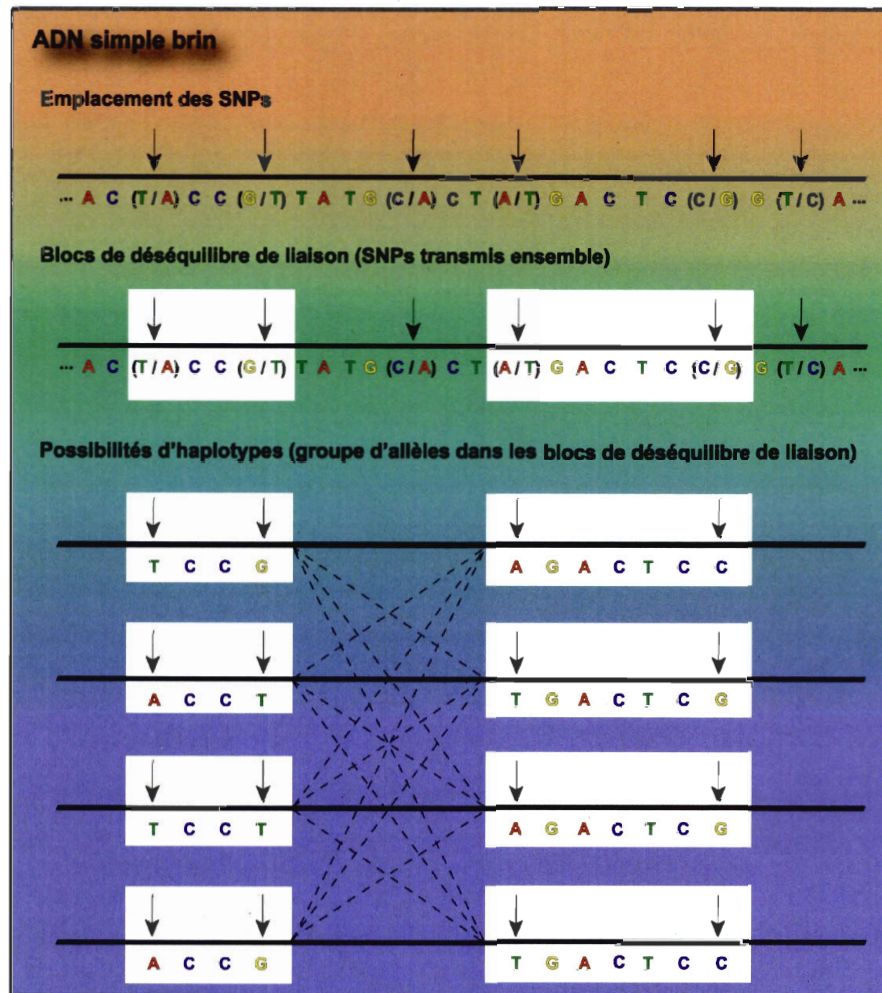


**Figure 5 : Représentation schématique des étapes menant à la réalisation d'une GWAS.** Les boîtes blanches représentent des blocs de déséquilibre de liaison à l'intérieur desquels se retrouvent les haplotypes. GWAS = *Genome-Wide Association Study* et SNP = Polymorphisme d'un seul nucléotide.

En ce qui concerne le projet génome humain, il a permis de déterminer la séquence complète du génome et de la rendre accessible aux chercheurs, via les bases de données publiques disponibles en ligne (par exemple celles du *National Center for*

*Biotechnology Information* [NCBI], *UCSC Genome Browser* ou *Ensembl Genome Browser*) en plus de localiser les variations génétiques communes comme les SNPs (International Human Genome Sequencing Consortium, 2004). Quant au projet HapMap, il a permis le génotypage, dans plusieurs populations, de plus de 600 000 SNPs espacés uniformément sur tout le génome (International HapMap Consortium, 2003; International HapMap Consortium, 2005; Thorisson et al., 2005), ce qui a permis de rendre compte de la diversité génétique au sein des populations étudiées (Grant et Hakonarson, 2008), en plus de mettre en évidence comment ces SNPs sont organisés en blocs de déséquilibre de liaison (aussi appelés blocs haplotypiques) (Rao, 2008). Ce concept de déséquilibre de liaison stipule que les allèles de SNPs voisins dans une région chromosomique donnée ont tendance à demeurer regroupés lors de la transmission, ce que l'on appelle des haplotypes (Ardlie et al., 2002) (Figure 6). Bien qu'une région chromosomique puisse contenir de nombreux SNPs, la localisation de quelques SNPs-marqueurs dans ces blocs haplotypiques, appelés « TagSNPs » (pour *Tagging* SNPs - Figure 5), suffit pour connaître la nature de la variation génétique contenue dans une région donnée. Considérant le fait qu'il serait financièrement impensable de génotyper les 10 millions de SNPs du génome pour chaque sujet dans la recherche des gènes d'une maladie, la connaissance d'une quantité définie de TagSNPs qui contiennent l'information génétique du génome est une avancée essentielle pour la génétique des traits complexes apportée par le projet HapMap (Thorisson et al., 2005). En parallèle avec la progression de ces connaissances génomiques, l'avènement des micropuces à ADN s'est avéré d'une nécessité pour la réalisation des GWAS puisque c'est grâce à cette technologie que le génotypage simultané des milliers de SNPs pour un même échantillon d'ADN a été rendu possible (Steemers et al., 2006) (Figure 5).

Les GWAS analysent donc la relation entre le phénotype d'une maladie à trait complexe et des milliers de SNPs et de TagSNPs. Le site internet de l'Institut de recherche national sur le génome humain (<http://www.genome.gov/26525384>) tient une liste exhaustive des GWAS publiées depuis 2005 (McCarthy et al., 2008). Ce fureteur permet un accès rapide à ces études et à leurs informations essentielles à savoir les phénotypes et échantillons étudiés, les régions chromosomiques et gènes identifiés en plus des références complètes.



**Figure 6 : Représentation schématique des combinaisons haplotypiques possibles à l'intérieur des blocs de déséquilibre de liaison formés par des groupes de SNPs transmis ensemble.** SNP = Polymorphisme d'un seul nucléotide, A = Adénine, C = Cytosine, G = Guanine et T = Thymine.

En ce qui concerne l'asthme, deux GWAS ont été réalisées à ce jour. La première de Moffatt et collègues (2007), a permis d'associer des SNPs du gène *ORMDL3*, une protéine transmembranaire du réticulum endoplasmique, à une susceptibilité génétique à l'asthme chez les enfants (Moffatt et al., 2007). La deuxième, de Ober et collaborateurs (2008), a permis d'associer un SNP du gène *CHI3L1*, codant pour une protéine chitinase YKL-40 impliquée dans l'inflammation et le remodelage des tissus, comme un facteur de susceptibilité pour l'asthme, l'hyperréactivité bronchique et le déclin de la fonction pulmonaire (Ober et al., 2008). Comme l'efficacité des GWAS a été démontrée dans plusieurs traits complexes (Grant et Hakonarson, 2008), il est fort probable que cette approche fera l'objet de nouvelles études génétiques pour l'asthme et ses

phénotypes intermédiaires au cours des prochaines années (Holloway et Koppelman, 2007). Il a même été prédit que les GWAS deviendront l'approche standard dans la recherche de gènes des traits complexes au cours de la prochaine décennie (Ober et Hoffjan, 2006).

Les étapes pour réaliser une GWAS reposent sensiblement sur les mêmes que celles des études d'association basées sur l'approche par gènes candidats (prochaine sous-section) à la différence que, dans la première, l'ensemble des gènes du génome sont étudiés (grâce au génotypage à haut débit), ce qui complexifie les analyses statistiques, notamment pour les corrections statistiques à appliquer dû aux tests multiples effectués et pour l'interprétation des résultats obtenus (Rao, 2008). Outre cette spécificité, les deux types d'approche se concentrent à identifier des facteurs de susceptibilité ou de protection à la maladie dans un échantillon de sujets dont l'évaluation clinique, le prélèvement d'ADN ou la définition du phénotype repose sur des critères standardisés (Figure 4). En plus, dans les deux cas, lorsqu'une association positive est trouvée, elle doit nécessairement faire l'objet d'une étude de réplication. Le présent projet de doctorat ne contient pas d'étude d'association basée sur l'approche génomique, c'est pourquoi l'accent n'est pas porté sur ses particularités méthodologiques (voir (McCarthy et al., 2008) pour détails). Néanmoins, les concepts qui ont été amenés dans cette sous-section sont nécessaires à la compréhension de certains critères de sélection des gènes et de leurs marqueurs génétiques pour les études d'association basées sur l'approche par gènes candidats.

### **3.2.2 Approche par gènes candidats**

L'approche par gènes candidats (Figure 4) est la méthode classique pour réaliser les études d'association pour un trait complexe (Palmer et Cookson, 2001) et elle est employée pour toutes les études génétiques présentées dans cette thèse. En fait, dans ces études, la première étape consiste à sélectionner un ou des gènes potentiellement impliqués dans la pathogénèse de la maladie étudiée selon plusieurs critères. Suivant cette priorisation de gènes candidats, l'étape qui suit nécessite la sélection de SNPs localisés à l'intérieur de ces gènes. Ces SNPs seront soit fonctionnels (c'est-à-dire qui ont un effet sur l'expression du gène ou sur la structure de la protéine produite), soit des TagSNPs, ou soit les deux. La dernière étape consiste à réaliser l'étude d'association en

tant que telle. Celle-ci s'effectue dans un échantillon de sujets présentant le phénotype d'intérêt et pour lesquels les génotypes des SNPs sélectionnés auront été obtenus via l'un des procédés disponibles pour le génotypage. Par des analyses statistiques appropriées au type d'échantillons utilisés (cas-témoins ou trios familiaux), il est ainsi possible d'évaluer s'il existe une association phénotype-génotype significative.

### ***3.2.2.1 Sélection des gènes candidats***

Le génome humain contient approximativement de 20 000 à 25 000 gènes codant pour les protéines nécessaires au développement et au fonctionnement de l'organisme humain (International Human Genome Sequencing Consortium, 2004). Ce nombre de gènes étant relativement très élevé, les chercheurs, se basant sur leurs connaissances de la littérature scientifique, doivent choisir le ou les meilleurs gènes candidats possibles pour réaliser leur étude d'association. Ces connaissances regroupent diverses informations comme la localisation des gènes dans des régions chromosomiques liées à la maladie étudiée (clonage positionnel), leur rôle biologique dans la pathophysiologie relative à la maladie, leur expression différentielle dans les tissus de la pathologie concernée de sujets atteints comparé à des sujets témoins ou leur association génétique connue avec la maladie. Certains outils bioinformatiques spécialisés dans le recoupage de ces informations ont également été développés afin de faciliter la tâche de priorisation. Bien que l'un ou l'autre de ces critères de sélection des gènes candidats soit suffisant pour mener une étude d'association, la plupart des chercheurs reposent leur choix sur une combinaison des informations recueillies pour poser leurs hypothèses de départ.

#### ***3.2.2.1.1 Clonage positionnel***

Le terme « clonage positionnel » provient de la méthode originale de cette approche qui consiste à cloner les régions chromosomiques autour d'un marqueur génétique lié à une maladie (suivant une étude de liaison) pour identifier le gène et son rôle dans la physiopathologie relative à cette maladie (Rosenwasser, 1996). Suivant ces études fonctionnelles, l'association génétique du gène est ensuite vérifiée dans des échantillons de sujets atteints (Rosenwasser, 1996). Depuis, le clonage positionnel est devenu une approche expérimentale standard pour cibler des gènes candidats pour une étude

d'association, à la différence que les étapes de clonage et d'études fonctionnelles ne sont plus nécessaires. Il suffit de connaître une région chromosomique liée à la maladie et les gènes qui y sont localisés représentent tous des gènes candidats potentiels pour une étude d'association génétique (Moffatt et Cookson, 1998). Cette sélection, basée sur la position chromosomique, augmente les chances de trouver des associations positives car le potentiel d'implication des gènes dans la maladie, même si leur fonction n'est pas encore caractérisée, est non négligeable (Hattersley et McCarthy, 2005). En contrepartie, les études de liaison ont une puissance statistique réduite en ce qui concerne l'identification des régions chromosomiques contenant des gènes qui ont un effet modeste dans la susceptibilité à une maladie (Hattersley et McCarthy, 2005). Par conséquent, ces derniers ne pourront pas être sélectionnés comme gènes candidats par clonage positionnel pour une étude d'association.

#### *3.2.2.1.2 Fonction biologique*

Une alternative au clonage positionnel est la sélection de gènes candidats pour leur fonction biologique connue ou potentiellement importante dans la physiopathologie et la pharmacologie relatives à la maladie (Sandford et al., 1996; Hoffjan et Ober, 2002; Hattersley et McCarthy, 2005). Par exemple, des gènes impliqués dans l'inflammation (activation et recrutement des cellules inflammatoires), des gènes impliqués dans la reconnaissance et la régulation immunitaire, des gènes qui régulent et entretiennent les composantes structurales du tissu pulmonaire, des gènes qui codent pour des protéines impliquées dans les mécanismes d'action des agents thérapeutiques, des gènes qui régulent la production des médiateurs lipidiques ou cytokiniques et leurs récepteurs ainsi que des gènes de facteurs de transcription (Anderson et Cookson, 1999; Illig et Wjst, 2002; Hoffjan et al., 2003) seraient de bons candidats pour l'asthme et ses phénotypes. Cette approche présente un très grand avantage lorsque la physiopathologie de la maladie est connue (Hattersley et McCarthy, 2005), puisque ces gènes, issus des études fonctionnelles, ont un rôle certain dans les mécanismes et par conséquent, représentent une bonne hypothèse de départ. Toutefois, lorsque la physiopathologie de la maladie n'est pas entièrement caractérisée, comme c'est le cas pour l'asthme, des gènes impliqués dans des cascades biologiques encore non caractérisées ne seront pas sélectionnés comme gènes candidats. D'où la nécessité de combiner d'autres approches.

### 3.2.2.1.3 Profil d'expression génique

Les études du profil d'expression génique analysent simultanément les transcripts (acide ribonucléique messager - ARN<sub>m</sub>) de milliers de gènes d'un tissu donné dont la détection s'effectue à l'aide de sondes spécifiques disposées sur un support de verre. Cette technologie des micropuces à ADN est la plus employée actuellement pour réaliser ces études en raison de sa fiabilité, de sa spécificité et de sa flexibilité démontrées. Dans une maladie comme l'asthme, la comparaison du profil d'expression du tissu pulmonaire de sujets asthmatiques comparé à celui de sujets témoins permet d'étudier plusieurs composantes de la maladie, à savoir les événements moléculaires qui sous-tendent son hétérogénéité, sa réponse aux thérapies ou même, sa réponse à une provocation allergénique (Rolph et al., 2006). Par exemple, l'étude de Laprise et collègues (2004) a comparé le profil d'expression des biopsies bronchiques de sujets asthmatiques avant et après traitement à la corticothérapie en plus de le comparer au profil d'expression des sujets témoins. Les gènes différemment exprimés entre les sujets asthmatiques et témoins dénotent une implication dans les mécanismes physiopathologiques comme l'inflammation ou le remodelage, et les gènes différemment exprimés avant et après traitement pourraient être impliqués dans la régulation de la réponse au traitement (Laprise et al., 2004).

Dans un contexte d'étude d'association, le profil d'expression différentiel des gènes dans les biopsies de sujets atteints comparé à celui de sujets témoins permet de sélectionner des gènes candidats, même si ces gènes ont une fonction inconnue dans la physiopathologie de la maladie (Holloway et Koppelman, 2007). En fait, cette caractéristique de l'utilisation des études d'expression présente l'avantage de cibler des gènes qui ne l'auraient pas été autrement. La sélection de tels gènes, ou des gènes impliqués dans les mêmes cascades biologiques, repose sur l'hypothèse que la différence d'expression observée est le résultat d'un ou de polymorphismes présents à l'intérieur des gènes et que ceux-ci seraient responsables de moduler le niveau d'expression du gène ou d'affecter l'intégrité des protéines produites. Cette approche, novatrice, en est à ses balbutiements en épidémiologie génétique de l'asthme et son efficacité reste à être démontrée. Toutefois, elle représente un potentiel d'identification de gènes et de voies biologiques encore non soupçonnés dans la physiopathologie relative à l'asthme ainsi qu'à ses déterminants génétiques.

#### 3.2.2.1.4 Outil bioinformatique *Candidate Genes to Inherited Diseases (G2D)*

Le développement d'outils bioinformatiques capables de prédire l'importance d'un gène dans une maladie en fonction des informations entreposées dans les bases de données publiques est un domaine de recherche en pleine expansion (Yu, 2009). Actuellement, il existe au moins 14 de ces outils qui possèdent chacun leurs particularités. Parmi eux, l'outil bioinformatique G2D (pour *Candidate Genes to Inherited Diseases*), un algorithme informatique publique (disponible à l'adresse suivante : [http://www.ogic.ca/projects/g2d\\_2/](http://www.ogic.ca/projects/g2d_2/)), développé pour permettre le criblage des gènes localisés dans une région chromosomique liée à une maladie génétique afin de prioriser des gènes candidats pour des études d'association génétique et dont la mise à jour a été effectuée spécifiquement en fonction de l'asthme (Perez-Iratxeta et al., 2002; Perez-Iratxeta et al., 2005). En fait, comme ces régions chromosomiques peuvent contenir des centaines de gènes, il est important de prioriser ceux qui auront plus de chances d'être associés à la maladie, ce qui permet de cibler les efforts de génotypage à un nombre plus restreint de gènes (Perez-Iratxeta et al., 2002). Le principe de l'outil G2D est de relier ces gènes avec la maladie étudiée en établissant une correspondance entre les données stockées dans les bases de données publiques, toutes accessibles via NCBI, comme Medline (pour la littérature biomédicale sur le gène et le phénotype, incluant les associations génétiques publiées), OMIM (pour les informations sur la maladie héréditaire) ou RefSeq (pour la séquence annotée du gène ainsi que la correspondance à sa fonction biologique) (Perez-Iratxeta et al., 2005). L'utilisation de l'algorithme nécessite deux types d'entrées : un phénotype et une région chromosomique (Perez-Iratxeta et al., 2005). Pour le phénotype, l'utilisateur doit fournir l'identifiant OMIM et pour la région chromosomique, l'utilisateur a le choix de fournir les deux marqueurs microsatellites, bandes cytogénétiques ou positions nucléotidiques qui entourent la région étudiée (Perez-Iratxeta et al., 2005). L'algorithme G2D est basé sur l'hypothèse que pour la maladie étudiée dont un gène X encore non associé est recherché et une maladie présentant un phénotype similaire dont un gène Y est associé, des fonctions entre les gènes X et Y peuvent être reliées et par conséquent, être potentiellement importantes pour les deux phénotypes (Perez-Iratxeta et al., 2005). La sortie de l'algorithme est une liste de gènes, connus ou prédits pour la maladie, ordonnés en fonction de la probabilité des gènes d'être impliqués dans le développement du phénotype d'intérêt (Perez-Iratxeta et al., 2007). Pour les présents travaux, plusieurs identifiants OMIM, représentant les différents phénotypes de l'asthme, ont été utilisés



pour produire la liste de gènes candidats initiale dans les deux régions chromosomiques liées à l'asthme et à l'atopie (6q26 et 10q26.3). Ainsi, la liste de gènes produite avec ces différents phénotypes de l'asthme inclut des gènes candidats qui, en plus de présenter une probabilité non négligeable d'être impliqués dans la maladie compte tenu de la littérature scientifique, présentent des similarités biologiques avec ceux de fonction connue dans des mécanismes physiopathologiques communs à ceux de l'asthme (Perez-Iratxeta et al., 2002; Perez-Iratxeta et al., 2005). Bref, l'utilisation de l'outil G2D est un exemple de combinaison des méthodes pour identifier des gènes candidats puisque, à partir de la liste de gènes produite, c'est au chercheur de décider s'il analysera tous les gènes pour une association génétique ou s'il introduira un autre niveau de priorisation dans sa sélection de gènes candidats, comme leur rôle biologique.

Tout comme les études de profil d'expression génique, l'utilisation de l'outil G2D n'est pas une approche standard en épidémiologie génétique de l'asthme (et des traits complexes en général) et son efficacité dans les études d'association reste également à être démontrée. Toutefois, les gènes ciblés par G2D, qui ne sont pas connus dans la physiopathologie relative à la maladie, présentent une probabilité non nulle de l'être compte tenu du recoupage de la littérature effectué. Par conséquent, l'outil G2D pourrait permettre d'identifier de nouveaux gènes dans les prédispositions à l'asthme.

#### 3.2.2.1.5 Association génétique connue avec le trait complexe étudié

Le dernier type de sélection des gènes candidats repose sur la connaissance *a priori* d'une association génétique du gène avec le trait complexe étudié (Hattersley et McCarthy, 2005). Ce type d'approche est classique et largement employé en épidémiologie génétique des traits complexes, comme en témoignent les nombreux exemples répertoriés dans la littérature scientifique pour l'asthme (Hoffjan et al., 2003). En fait, lorsqu'une étude rapporte une première association génétique d'un gène avec un trait complexe, un chercheur pourra sélectionner ce gène comme gène candidat pour une étude d'association, qui sera alors appelée une étude de répliation. L'étude de répliation est une méthodologie standard en recherche génétique : chaque association génétique trouvée doit être confirmée dans un échantillon indépendant à l'étude originale pour démontrer l'importance du gène dans la maladie (Nature Genetics Editorial, 1999; Ioannidis et al., 2001; Hirschhorn et Altshuler, 2002; Ober et Hoffjan,

2006; Vercelli, 2008). Un gène répliqué dans plus d'une étude présente un effet non négligeable sur les prédispositions génétiques au trait complexe étudié, ce qui fait que les associations génétiques positives précédentes sont un indicateur de l'implication potentielle d'un gène dans l'étiologie d'une maladie (Hattersley et McCarthy, 2005). Par conséquent, une telle connaissance préalable renforce la sélection d'un gène pour une étude d'association (Hattersley et McCarthy, 2005). Toutefois, il est bon de mentionner que la majorité des associations génétiques rapportées pour les traits complexes ne sont pas systématiquement répliquées à travers les études (Ioannidis et al., 2001; Hirschhorn et al., 2002). En fait, plusieurs facteurs peuvent moduler la distribution des associations génétiques, comme les différentes structures génétiques des populations étudiées, les différentes conceptions des études, les différents phénotypes ou environnements et leurs interactions (Ioannidis et al., 2001; Hirschhorn et al., 2002; Guerra et Martinez, 2008). Cet aspect de la non-réplication possible des associations génétiques dans les traits complexes sera abordé en détails dans la discussion de cette thèse (Chapitre 9).

#### *3.2.2.1.6 Association génétique connue un autre trait complexe qui partage une composante physiopathologique avec le trait complexe étudié*

La sélection d'un gène candidat basée sur son association génétique connue avec le trait complexe étudié peut être étendue à une sélection d'un gène préalablement associé à un autre trait complexe qui partage une ou plusieurs composantes physiopathologiques. Il s'agit d'un principe de sélection basé sur l'implication moléculaire potentielle du gène dans la maladie. Par exemple, les études génétiques dans les maladies inflammatoires de l'intestin, dans l'arthrite rhumatoïde ou dans le diabète de type I ont permis de constater que des gènes ou des familles de gènes sont communs entre ces maladies inflammatoires et immunitaires et que, par conséquent, ces gènes pourraient également avoir un rôle dans l'asthme ou l'allergie (Cookson, 2002). D'autres traits complexes partagent une composante physiopathologique et génétique avec l'asthme, c'est le cas pour l'obésité (Beuther et al., 2006) et la dermatite atopique (Cookson et Moffatt, 2002; Cookson, 2004; Kiyohara et al., 2008) pour ne nommer que ces exemples. Ainsi, la connaissance d'une implication génétique et biologique d'un gène dans une maladie qui partage des composantes physiopathologiques avec l'asthme telles l'inflammation ou l'atopie permet de cibler des gènes candidats de choix pour des études d'association.

Cette approche offre également la possibilité d'identifier de nouveaux gènes qui régulent les mécanismes physiopathologiques de l'asthme.

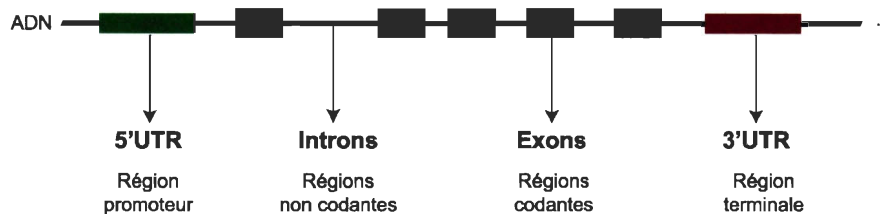
### **3.2.2.2 Sélection des SNPs**

Une fois que les gènes candidats sont choisis, le ou les SNPs localisés à l'intérieur de ces gènes sont sélectionnés et génotypés selon les diverses technologies disponibles (voir évaluation génotypique, section 3.2.2.5) dans les échantillons de sujets asthmatiques étudiés (Gray et al., 2000). La plupart des technologies de génotypage sont concentrées sur les SNPs étant donné (1) leur très grand nombre et (2) la facilité méthodologique que représente leur détection sur l'ADN (Schork et al., 2008). Les 10 millions de SNPs présentant une fréquence de l'allèle mineur supérieure à 1 % (Reich et al., 2003) représentent 90 % des variations génétiques du génome humain (Palmer et Cardon, 2005). Parmi les autres variations génétiques, apparaissent les répétitions nucléotidiques, les insertions, délétions ainsi que les duplications ou inversions de fragments d'ADN (Redon et al., 2006). Comme ces autres variations génétiques ne sont détectables qu'avec le séquençage et que les coûts du séquençage à grande échelle ne peuvent actuellement pas concurrencer avec ceux du génotypage à grande échelle (Schork et al., 2008), les études d'association pour les traits complexes emploient majoritairement les SNPs comme marqueurs génétiques (Schork et al., 2008).

Étant donné le nombre important de SNPs connus dans le génome, un choix ciblé aux gènes candidats étudiés doit être effectué. Ainsi, grâce aux nombreuses bases de données publiques disponibles en ligne (consulter (Palmer et Cardon, 2005) pour un résumé), il est possible de choisir des SNPs fonctionnels ou des TagSNPs (Figure 5), selon le type d'association désirée, directe ou indirecte, respectivement.

#### **3.2.2.2.1 Association directe – SNPs fonctionnels**

Les SNPs fonctionnels sont des variations génétiques qui se situent dans des régions précises d'un gène et qui affectent ses produits, c'est-à-dire l'ARN<sub>m</sub> et les protéines. Ces régions préférentielles, représentées de façon schématique dans la Figure 7, sont la région du promoteur (5'UTR pour *Untranslated Region*), la région terminale (3'UTR) ou les régions codantes (exons) (Murthy et al., 1996; Rosenwasser, 1996).



**Figure 7 : Représentation schématique d'un gène et de ses différentes régions.**

Cette sélection est faite sur la base qu'un SNP situé au niveau du 5'UTR (qui contient des sites d'initiation à la transcription) ou du 3'UTR (qui contient des sites de polyadénylation – stabilité de l'ARN<sub>m</sub>) d'un gène joue un rôle dans son niveau d'expression (Dreyfuss et al., 1996; Murthy et al., 1996; Zhang et al., 2008). Par ailleurs, un SNP situé dans les exons d'un gène affecte la conformation de la protéine synthétisée, laquelle sera plus ou moins efficace, voire même complètement inefficace (Goldstein et al., 2003). En ce qui concerne les SNPs localisés dans les régions introniques du gène, il ne sont pas considérés comme des SNPs fonctionnels même si ces derniers peuvent affecter l'épissage (mécanisme qui enlève les introns et réunit les exons) de l'ARN<sub>m</sub> lors de la transcription du gène, ce qui pourrait donner lieu à différents ARN<sub>m</sub> (phénomène appelé épissage alternatif) (Ast, 2004; Zhang et al., 2008).

Lorsqu'une association positive est trouvée avec un SNP fonctionnel, il y a de fortes chances pour que ce dernier joue un rôle dans le développement de la maladie étudiée (Cordell et Clayton, 2005), surtout si ce rôle a déjà été démontré par des études fonctionnelles précédentes. Toutefois, même si le SNP est situé dans l'une ou l'autre des régions préférentielles du gène, il n'est pas certain qu'il soit causal car les connaissances sur les effets de tous les SNPs fonctionnels sont limitées et incomplètes (Neale et Sham, 2004). C'est pourquoi la plupart des études d'association génétique sont suivies d'études fonctionnelles des SNPs associés afin de caractériser leur rôle dans la physiopathologie de la maladie (Steinke et al., 2008; Zhang et al., 2008).

#### 3.2.2.2.2 Association indirecte – TagSNPs

Les TagSNPs (Figure 5) sont des SNPs-marqueurs représentatifs d'un bloc de déséquilibre de liaison (Rao, 2008). Brièvement, ces blocs contiennent des SNPs dont les allèles sont transmis ensemble (sous forme d'haplotype – Figure 6) étant donné leur

proximité physique. En fait, cette proximité réduit les évènements de recombinaison possible entre eux, ce qui fait que la transmission des allèles ne se fait plus aléatoirement, elle est déséquilibrée en faveur des haplotypes (Ardlie et al., 2002). Le déséquilibre de liaison est un concept-clé en génétique des populations. En fait, la longueur d'un bloc de déséquilibre de liaison et sa distribution sont variables d'une population à l'autre dû à la structure génétique de chaque population, soumise à différents facteurs démographiques au cours de son établissement. À titre indicatif, ces facteurs regroupent la dérive génétique (changement dans les fréquences alléliques et haplotypiques de génération en génération – responsable des fréquences alléliques variables d'une population à l'autre), le taux de croissance des populations, les évènements migratoires et de mélange entre les populations, la sélection naturelle ou les différents taux de mutation et de recombinaison (Ardlie et al., 2002).

Comme les TagSNPs représentent les variations génétiques localisées dans un bloc haplotypique, il est possible d'en déterminer un nombre minimal pour couvrir l'ensemble de l'information génétique contenue dans un gène (Grant et Hakonarson, 2008). Les méthodes pour identifier ces TagSNPs sont décrites dans les articles concernés des Chapitres 5 à 8. Bien que ces TagSNPs soient parfois localisés dans les régions préférentielles du gène, ils sont en général localisés dans les régions introniques ou dans les séquences en aval et en amont du gène (*upstream* et *downstream*).

L'utilisation des TagSNPs est une approche efficace, tel que démontré avec les GWAS, et couramment employée dans les études d'association basées sur l'approche par gènes candidats (Neale et Sham, 2004). Une association génétique positive avec un TagSNP est considérée comme indirecte puisque ce dernier a été choisi comme un SNP représentatif et non comme un SNP fonctionnel, même si l'un n'empêche pas l'autre. Ainsi, le TagSNP associé pourra (1) être lui-même fonctionnel si localisé dans les régions préférentielles du gène, (2) être en déséquilibre de liaison avec un SNP fonctionnel fréquent ou rare, ou (3) être en déséquilibre de liaison avec d'autres variations génétiques causales fréquentes ou rares (par exemple les insertions, les délétions, les duplications, les inversions) (Neale et Sham, 2004; Redon et al., 2006; Schork et al., 2008). Les conclusions pour une association génétique avec un TagSNP sont donc conservatrices car elles nécessitent, au minimum, le séquençage de la région du gène qui entoure le TagSNP associé. C'est seulement ainsi que la ou les variations

généétiques responsables pourront être identifiées. Leur rôle dans la physiopathologie de la maladie sera ensuite confirmé par des études fonctionnelles (Zhang et al., 2008).

La sélection des SNPs est une étape cruciale dans la réalisation des études d'association et repose en bonne partie sur les connaissances acquises par les projets Génome Humain et HapMap à l'aide de l'ADN d'individus ne présentant pas de problèmes de santé et faisant partie de populations ou de groupes ethniques précis (International HapMap Consortium, 2003; International Human Genome Sequencing Consortium, 2004; International HapMap Consortium, 2005; Thorisson et al., 2005). Choisir un SNP sur cette base nécessite donc que le groupe de sujets (échantillon) sélectionné pour réaliser l'étude d'association soit comparable en termes de population ou d'ethnie au groupe de sujets employés par ces deux projets de référence. De cette manière, les fréquences alléliques (ou haplotypiques) des SNPs sélectionnées seront présumées représentatives de celles qui devraient être retrouvées dans l'échantillon étudié. Une fréquence des allèles mineurs supérieure à 5 % est généralement employée comme valeur seuil pour la sélection d'un SNP (Bodmer et Bonilla, 2008). L'obtention de SNPs non polymorphes (c'est-à-dire avec une fréquence allélique de 0 %) est ainsi limitée puisqu'ils ne sont d'aucune utilité pour les analyses statistiques en plus d'engager des frais de génotypage inutiles. De plus, les SNPs présentant des fréquences de l'allèle mineur inférieures à 5 % sont considérés comme des variants rares et sont généralement spécifiques aux populations, particulièrement en raison des phénomènes de dérive génétique ou d'effet fondateur qui pourraient avoir eu lieu (Bodmer et Bonilla, 2008).

### **3.2.2.3 Échantillonnage**

Il existe plusieurs conceptions d'études (ou *study design*) communément employées pour réaliser les études d'association dans les traits complexes (Cordell et Clayton, 2005), les deux plus populaires étant l'utilisation des échantillons cas-témoins et des échantillons de trios familiaux (Rao, 2008).

#### **3.2.2.3.1 Cas-témoins**

Les études de type cas-témoins représentent la conception standard pour les études d'association et sont les plus fréquemment employées (Cordell et Clayton, 2005;

Howson et al., 2005; Laird et Lange, 2008; Zhu et al., 2008). L'initiation de telles études nécessite le recrutement d'un nombre de sujets « cas », c'est-à-dire atteints de la maladie étudiée, non apparentés entre eux et sélectionnés aléatoirement à partir de différents registres médicaux, et d'un nombre de sujets « témoins », c'est-à-dire non-atteints de la maladie, non apparentés entre eux et choisis aléatoirement dans les registres de la population choisie pour l'échantillonnage (Cordell et Clayton, 2005; Newton-Cheh et Hirschhorn, 2005). De plus, le modèle idéal tentera d'apparier les cas et les témoins pour l'âge et le genre.

Dans ce type d'étude, le groupe de sujets cas est présumé avoir une plus grande fréquence des allèles de susceptibilité que le groupe de sujets témoins (McCarthy et al., 2008), c'est pourquoi les fréquences alléliques (ou haplotypiques ou génotypiques) sont comparées entre ces deux groupes de sujets (Hoffjan et al., 2003; Vercelli, 2008). Ainsi, un allèle (ou un haplotype ou un génotype) d'un marqueur d'un gène donné sera dit associé à la maladie si sa fréquence est significativement plus grande dans le groupe de sujets cas que dans le groupe de sujets témoins (Eley et Rijdsdijk, 2005). D'autre part, un allèle moins fréquent dans le groupe de sujets cas sera dit de protection.

Le plus grand avantage à l'utilisation d'une conception d'étude cas-témoins est sans doute lié à la facilité relative que représente le recrutement des sujets (Newton-Cheh et Hirschhorn, 2005). Comparativement à un échantillon de familles, le recrutement d'un échantillon cas-témoins ne nécessite aucun suivi des sujets, réduisant les coûts liés au maintien de l'échantillon, et nécessite un nombre total de sujets moins élevé pour obtenir les mêmes puissances statistiques (Cordell et Clayton, 2005), ce qui réduit considérablement les coûts liés au génotypage. Un autre avantage réside dans la simplicité des analyses statistiques qui permettent d'évaluer plusieurs paramètres d'exposition en plus du risque génétique (Cordell et Clayton, 2005).

Toutefois, les échantillons cas-témoins sont sensibles aux populations étudiées et sujets à des facteurs confondants (biais) si l'échantillon n'est pas rigoureusement collecté. Idéalement, les cas et les témoins devraient être sélectionnés à partir de populations présentant la même histoire démographique et ancestrale pour éviter les biais liés à la stratification (sous-ensembles) de la population (Lander et Schork, 1994; Howson et al., 2005; McCarthy et al., 2008). En fait, les échantillons de sujets cas, autant que ceux des

sujets témoins, peuvent être composés par divers sous-ensembles de la population distincts de par leur structure génétique et exposés à divers facteurs de risque environnementaux (Zhu et al., 2008). Si les cas et les témoins diffèrent par ces caractéristiques, il est possible que les fréquences alléliques des marqueurs génétiques étudiés soient variables au sein de ces sous-ensembles, ce qui affecterait les résultats de l'association génétique (McCarthy et al., 2008). Par conséquent, un allèle d'un marqueur génétique donné plus fréquemment observé chez les cas comparativement aux témoins ne sera pas nécessairement associé à la maladie, il pourrait en effet être lié à la stratification de la population présente dans le groupe de sujets cas. Une telle association est considérée comme étant confondante, c'est-à-dire comme un faux-positif (erreur de type I) (Cordell et Clayton, 2005; Zhu et al., 2008). Ce biais, lié à la stratification des populations, peut être contourné dans les analyses statistiques en appariant les données selon les régions géographiques ou les origines ethniques des sujets ainsi que par l'emploi de contrôles génomiques (Pritchard et Rosenberg, 1999; Devlin et al., 2001; Cordell et Clayton, 2005).

La conception d'étude cas-témoins est également très sensible au biais de sélection, surtout en ce qui concerne les sujets témoins. La sélection de ces sujets doit être effectuée de façon très rigoureuse et leur phénotype de non-atteint doit absolument être bien caractérisé (Cordell et Clayton, 2005) et ce groupe de sujets doit être le portrait de la population générale, c'est-à-dire du groupe d'âge ou d'autres facteurs démographiques, qu'ils sont supposés représenter (McCarthy et al., 2008).

Parmi les études d'association incluses dans la présente thèse, deux échantillons de type cas-témoins ont été utilisés. Le premier, utilisé dans le premier article (Chapitre 5, section 5.1), a été composé à partir des participants d'une cohorte de sujets construite pour évaluer un programme d'intervention et de prévention de l'asthme et d'autres maladies allergiques, et dont le recrutement a été effectué à partir de deux centres de recherche canadiens (Vancouver et Winnipeg) (Chan-Yeung et al., 2000). Pour ce qui est du deuxième échantillon cas-témoins, utilisé dans le deuxième article (Chapitre 5, section 5.2), il s'agit de celui collecté dans la population de Busselton dans le cadre d'une étude australienne de santé publique (Hill et al., 1995; Daniels et al., 1996; Moffatt et al., 2001; Palmer et al., 2001; James et al., 2005). Pour plus de détails sur la conception de ces deux



échantillons cas-témoins ou sur la caractérisation des sujets, consulter la description résumée dans les articles concernés ou les articles de référence cités.

#### 3.2.2.3.2 *Trios familiaux*

Les études familiales, essentiellement les trios familiaux formés d'au moins un enfant atteint et de ses deux parents, représentent une conception d'étude d'association également très employée (Cordell et Clayton, 2005). Étant donné que les études familiales sont également utilisées dans les autres types d'études génétiques (d'agrégation, de ségrégation ou de liaison), l'emploi des trios familiaux pour des études d'association nécessite des adaptations afin que les résultats obtenus soient bien le fruit d'une association et non l'effet d'une liaison génétique (Laird et Lange, 2008). Pour ce faire, les analyses statistiques sont adaptées pour tenir compte de cet aspect.

Dans ce type d'étude, les allèles, les génotypes ou les haplotypes transmis aux enfants atteints sont comparés à ceux non-transmis (Spielman et al., 1993; Wang et Zhao, 2003). Ainsi, des données de transmission des allèles de susceptibilité sont appariées à des données de transmission témoins pour chaque enfant atteint (Cordell et Clayton, 2005), une sorte de contrôle interne qui élimine l'effet de stratification potentiel lié à la structure de la population étudiée (Howson et al., 2005). En d'autres termes, une association positive est indiquée lorsqu'un allèle présente une transmission préférentielle des parents aux enfants atteints comparativement aux allèles non-transmis. Comme tous les allèles proviennent d'une même famille, il n'y a aucune possibilité d'association faux-positif liée à la stratification de la population (Eley et Rijdsdijk, 2005), ce qui représente le plus grand avantage à l'utilisation des trios familiaux (Cordell et al., 2004; Newton-Cheh et Hirschhorn, 2005). Un autre avantage à leur utilisation est la possibilité d'estimer l'origine parentale de la transmission des allèles de susceptibilité à la maladie (Cordell et al., 2004; Cordell et Clayton, 2005).

Malgré ces avantages, le recrutement des échantillons familiaux est plus long et plus coûteux que celui des échantillons cas-témoins, à moins que les familles aient déjà été collectées pour des études de liaison préalables (Cordell et Clayton, 2005; Eley et Rijdsdijk, 2005; Howson et al., 2005; Ober et Hoffjan, 2006). Ainsi, il est plus difficile d'obtenir une taille d'échantillons comparable à celle des échantillons cas-témoins, ce

qui fait que la puissance statistique d'un échantillon de trios familiaux est généralement moindre que celle des cas-témoins (Cordell et Clayton, 2005).

Dans le cadre des études d'association présentées dans cette thèse, quatre échantillons de trios familiaux, composés d'au moins un enfant asthmatique accompagné de ses deux parents, ont été employés pour une ou l'autre des études. Le premier, utilisé dans toutes les études d'association de la présente thèse, est l'échantillon familial de sujets asthmatiques originaires de la région du Saguenay–Lac-Saint-Jean (ou SLSJ) (Labuda et al., 1996; Heyer et al., 1997; Laitinen et al., 2001; Scriver, 2001; Lemire et al., 2004). Le deuxième, employé dans le deuxième et le troisième article (Chapitre 5, sections 5.2 et 5.3), est un échantillon de trios familiaux collectés à partir des participants d'une cohorte de naissance canadienne à haut risque pour l'asthme nommée « *The Canadian Asthma Primary Prevention Study* » (ou CAPPS) (Chan-Yeung et al., 2000; Becker et al., 2004; Chan-Yeung et al., 2005). Le troisième, également employé dans le deuxième et le troisième article (Chapitre 5, sections 5.2 et 5.3), est un échantillon collecté à partir de la cohorte de naissance populationnelle canadienne nommée « *The Study of Asthma Genes and the Environment Cohort* » (ou SAGE). Ces trois premiers échantillons de trios familiaux sont entièrement décrits dans le troisième article (Chapitre 5, section 5.3). Le dernier échantillon de trios familiaux a été employé dans le cinquième article (Chapitre 6, section 6.1) et provient d'une cohorte de naissance américaine nommée « *The Childhood Asthma Management Program Study* » (ou CAMP) (CAMP, 1999; The Childhood Asthma Management Program Research Group, 2000). Pour plus de détails sur la conception de ces échantillons ou sur la caractérisation des sujets, consulter la description résumée dans les articles concernés ou les articles de référence cités.

Peu importe la conception d'étude favorisée pour mener une étude d'association, la contrainte de la taille des échantillons doit être respectée afin que le nombre de sujets inclus dans l'étude soit assez grand pour atteindre une puissance statistique suffisante pour détecter les associations génétiques (Rao, 2008). Cet aspect de la puissance statistique est d'une importance capitale dans les études d'association génétique et sera abordé en détails dans la discussion de cette thèse (Chapitre 9). Une fois le nombre de participants requis ainsi déterminé, ces derniers sont recrutés selon diverses méthodes (cliniques médicales, centres de recherche, publicités dans les médias, etc.) ainsi que selon divers critères d'inclusion, et leur participation se fait sur une base volontaire.

Ensuite, une évaluation clinique, incluant la réponse à un questionnaire de santé respiratoire, des tests de fonction respiratoire (spirométrie) et un prélèvement sanguin pour extraction de l'ADN, leur est prodiguée par un professionnel de la santé suivi par le diagnostic d'un pneumologue. Ces données cliniques obtenues, ainsi que les échantillons d'ADN prélevés, permettent respectivement de déterminer le(s) phénotype(s) et le(s) génotype(s) de chaque individu afin de pouvoir établir la relation statistique phénotype-génotype qui existe entre le trait complexe étudié et les marqueurs génétiques des gènes candidats sélectionnés.

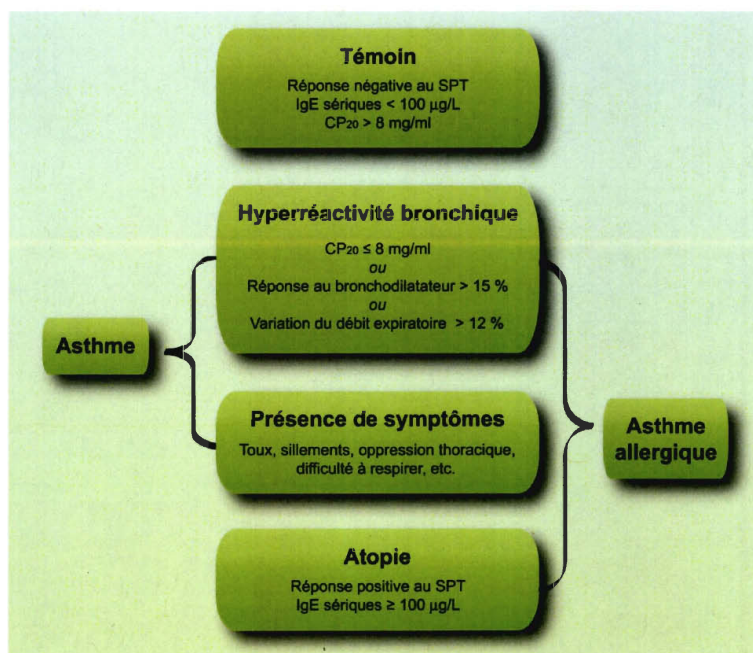
#### ***3.2.2.4 Évaluation phénotypique***

L'évaluation phénotypique consiste à utiliser une définition standardisée des phénotypes de la maladie étudiée, basée sur des critères cliniques, et de caractériser chacun des sujets en fonction de ces critères. Peu importe le type d'échantillon étudié (cas-témoins ou trios familiaux), cette étape est également d'une importance capitale pour réaliser une étude d'association génétique. En fait, ne pas bien classifier les sujets, selon qu'ils soient atteints ou non atteints de la maladie étudiée, peut causer des biais de classification (McCarthy et al., 2008). Ainsi, des phénotypes mal définis augmentent le risque d'association génétique confondante (Rousseau et Laflamme, 2003).

Pour l'asthme, les critères recommandés par l'*American Thoracic Society* (ou ATS) ont été employés pour définir les phénotypes de toutes les études d'association de la présente thèse, avec certaines particularités selon les échantillons. Pour ce qui est de l'échantillon familial du SLSJ, les critères de l'évaluation phénotypique sont résumés dans le prochain paragraphe et disponibles en entier dans le deuxième et le troisième article (Chapitre 5, sections 5.2 et 5.3). Pour les autres échantillons utilisés, certaines différences sont possibles (consulter les articles concernés pour les particularités).

Donc, suivant les réponses au questionnaire de santé respiratoire standardisé (ATS, 1987), les tests de fonction respiratoire, incluant le volume expiratoire maximal en une seconde (VEMS), la capacité vitale forcée, un test de provocation bronchique à la méthacholine (CP<sub>20</sub>), la réponse au bronchodilatateur ainsi que la variation du débit expiratoire de pointe (ATS, 1995), sont effectués pour chacun des sujets suivi d'un test d'allergie cutané à la piqûre (SPT pour *Skin Prick Test*). À partir de ces informations

cliniques, un sujet est considéré comme asthmatique s'il présente les symptômes respiratoires associés à l'asthme en plus d'une hyperréactivité bronchique, définie ici comme une  $CP_{20} < 8$  mg/ml (ATS, 1995). Si la  $CP_{20}$  n'est pas mesurable, comme c'est parfois le cas chez de jeunes enfants, l'hyperréactivité bronchique est définie comme une augmentation supérieure ou égale à 15 % du VEMS après l'inhalation d'un bronchodilatateur ou comme une variation biquotidienne du débit expiratoire de pointe plus grand ou égal à 12 % de la valeur de base sur une période de deux semaines (ATS, 1995). Un sujet est considéré comme allergique s'il présente un taux anormalement élevé d'IgE et que son SPT est positif pour au moins un allergène (c'est-à-dire  $\geq 3$  mm d'induration après 10 minutes) (Sub-Committee on Skin Tests of the European Academy of Allergology and Clinical Immunology, 1989). Finalement, un sujet sera considéré comme asthmatique allergique s'il présente les symptômes respiratoires, l'hyperréactivité bronchique et un SPT positif. Parallèlement, les sujets témoins ne souffrent pas d'asthme ou de maladies associées à l'asthme, ne présentent aucune réaction positive au SPT, aucune hyperréactivité bronchique ( $CP_{20} > 8$  mg/ml) et leur quantité d'IgE sériques est dite normale (ATS, 1995). La Figure 8 présente un schéma récapitulatif de ces critères de classification des différents phénotypes.



**Figure 8 : Critères utilisés dans l'établissement des phénotypes relatifs à l'asthme pour les sujets de l'échantillon familial du SLSJ.** SPT = Test d'allergie cutanée à la piqûre, IgE = Immunoglobuline E,  $CP_{20}$  = Concentration de méthacholine qui provoque une chute de 20 % du VEMS.

### 3.2.2.5 Évaluation génotypique

L'évaluation génotypique (appelée génotypage) consiste en premier lieu à extraire l'ADN, dans le cas présent à partir des globules blancs présents dans les prélèvements sanguins, et de déterminer les génotypes des SNPs étudiés pour chaque sujet à l'étude. Les différentes technologies de génotypage disponibles reposent toutes sur le principe de l'amplification PCR (pour *Polymerase Chain Reaction*) qui, avec l'utilisation d'amorces spécifiques, permet d'amplifier la section d'ADN précise contenant le SNP d'intérêt. Ce fragment d'ADN amplifié permet ensuite la détection du SNP au moyen de techniques variées. La plus classique est la technique RFLP (pour *Restriction Fragments Length Polymorphisms*) qui repose sur l'utilisation d'une enzyme de restriction dont la fonction est de parcourir chacun des brins d'ADN amplifiés et de « couper » si son motif de reconnaissance est présent dans le fragment. L'action de cette enzyme permet de scinder l'ADN en plusieurs fragments de restriction de longueurs différentes (en nombre de nucléotides), ce qui permet leur reconnaissance sur un gel d'agarose après une migration électrophorétique. Selon que l'ADN ait été fragmenté ou non, il est possible d'établir le génotype de chacun des individus porteurs de la mutation (hétérozygote), non porteurs de la mutation (homozygote normal) et porteurs de deux allèles mutés (homozygote muté). D'autres techniques reposent sur un système de génotypage automatisé à haut rendement qui consistent en une réaction PCR initiale en multiplex (c'est-à-dire plusieurs SNPs dans une seule réaction PCR) suivie d'une réaction d'extension d'un nucléotide unique au moyen de didésoxynucléotide triphosphate (ddNTP) colorés avec des marqueurs fluorescents (spécifiques à chaque technologie). La fluorescence émise par les colorants est ensuite détectée par un appareil branché à un ordinateur contenant les logiciels d'interprétation et de détermination des génotypes développés par les différentes compagnies. Dans les études d'association incluses dans la présente thèse, plusieurs technologies à haut rendement ont été employées : *SNPstream*<sup>®</sup> *UHT* (pour *Ultra-High Throughput SNP genotyping*) de la compagnie Orchid Biosciences (États-Unis), *HEFP*<sup>™</sup> (pour *High Efficiency Fluorescence Polarization detection technology*) de la compagnie Molecular Devices (Etats-Unis), *GoldenGate Genotyping BeadChip Assay* de la compagnie Illumina (États-Unis), *TaqMan*<sup>™</sup> *SNP genotyping assay* de la compagnie Applied Biosystems (États-Unis) et *Sequenom*<sup>®</sup> *matrix-assisted laser mass array* de la compagnie Sequenom (États-Unis). Par souci de brièveté, le détail des particularités méthodologiques de chacune de ces technologies, incluant les conditions PCR, les appareils et logiciels

utilisés ainsi que les références, sont disponibles dans la section méthode des articles de chacune des études d'association de la présente thèse (Chapitres 5 à 8). Le choix de l'une ou l'autre des technologies se fait généralement selon la disponibilité des appareils dans les laboratoires concernés, selon le débit (peu ou beaucoup de SNPs à vérifier à la fois) et bien sûr, selon le coût de la technologie prise en considération.

Pour que les données génotypiques puissent être utilisées à des fins d'analyse, la qualité du génotypage doit être évaluée. En effet, des erreurs de génotypage peuvent modifier de façon notable les fréquences alléliques observées et les puissances statistiques de l'échantillon de sujets, ce qui peut avoir un impact sur le résultat des analyses statistiques et les conclusions qui en seront tirées (Rousseau et Laflamme, 2003). Une première vérification concerne le taux de réussite du génotypage pour chaque SNP étudié, c'est-à-dire le pourcentage de génotypes obtenus sur le total d'échantillons d'ADN évalués. Ce taux de réussite doit être supérieur à 95 % pour limiter les biais provoqués par les erreurs de type I. Une deuxième vérification, spécifique aux échantillons de trios familiaux, est le décompte des erreurs mendéliennes qui se définissent comme les impossibilités de transmission des allèles des parents aux enfants. Par exemple, deux parents de génotype homozygote pour un des allèles ne pourront transmettre l'allèle complémentaire à l'enfant atteint. Lorsque des erreurs mendéliennes sont trouvées, les génotypes des familles présentant ces erreurs ne sont pas utilisés pour les analyses. De plus, si un SNP présente un total de plus de deux erreurs mendéliennes, il ne sera pas non plus considéré pour les analyses. Finalement, les erreurs de génotypage peuvent modifier la fréquence allélique des SNPs et causer une perte de l'équilibre théorique de Hardy-Weinberg (EHW) (Rousseau et Laflamme, 2003). Cette vérification est particulièrement nécessaire dans les études cas-témoins où les individus ne sont pas apparentés et où les erreurs mendéliennes ne peuvent pas être vérifiées (Hoffjan et al., 2003). La vérification de l'EHW commence avec la détermination des fréquences alléliques observées « p » (pour l'allèle majeur ou le plus fréquent) et « q » (pour l'allèle mineur ou le moins fréquent) à partir des fréquences génotypiques observées dans l'échantillon. Avec les fréquences alléliques, il est possible d'établir les proportions génotypiques attendues dans la population sous l'EHW en appliquant la formule  $[p^2 + 2pq + q^2 = 1]$  où les fréquences génotypiques sont représentées par  $p^2$  (homozygote normal),  $2pq$  (hétérozygote) et  $q^2$  (homozygote muté). Ainsi, les proportions génotypiques observées sont comparées à celles attendues par un test

d'hypothèse du chi-deux ( $\chi^2$ ) pour accepter ou non l'hypothèse de départ stipulant que les proportions génotypiques sont à l'EHW (Hartl et Clark, 1997). Une déviation de cet équilibre populationnel est signe que des erreurs de génotypage sont survenues ou qu'un des génotypes est favorisé en faveur de la maladie (Xu et al., 2002), faisant augmenter sa fréquence par rapport à celle attendue et ainsi, contribuer à déséquilibrer les fréquences génotypiques attendues.

### **3.2.2.6 Analyses statistiques**

Les analyses statistiques dépendent de la conception des échantillons pour l'étude d'association. Pour une étude cas-témoins, des méthodes comme la régression logistique, le test d'indépendance du chi-carré, les rapports des cotes d'exposition (odds ratio) et le risque relatif (ratio de la présence d'un allèle chez les cas sur celle chez les témoins) sont employés comme tests statistiques (Sandford et al., 1996; Cordell et Clayton, 2005; Newton-Cheh et Hirschhorn, 2005). Ces tests peuvent être réalisés avec des logiciels d'analyses statistiques standards et permettent d'estimer le risque que représente le facteur génétique étudié sur le développement de la maladie (Cordell et Clayton, 2005).

En ce qui concerne les échantillons de trios familiaux, le test statistique utilisé pour évaluer l'association entre l'asthme et les déterminants génétiques est le test de déséquilibre de transmission (Spielman et al., 1993). Ce test vise à évaluer le nombre de fois où un allèle de susceptibilité est transmis ou non chez l'enfant atteint. Comme la transmission des allèles attendue selon les lois de Mendel est de 50 %, tout déséquilibre significatif, c'est-à-dire avec une valeur-p du test inférieure ou égale à 0,05, indiquera une association génétique de l'allèle avec la maladie (Spielman et Ewens, 1996; Peltonen et al., 2000). Un allèle transmis significativement à plus de 50 % des parents hétérozygotes aux enfants atteints signifie que l'allèle est de susceptibilité pour la maladie. Parallèlement, un allèle transmis à moins de 50 % signifie qu'il est de protection pour la maladie. Le test de déséquilibre de transmission est effectué à l'aide d'un logiciel spécialisé nommé FBAT (pour *Family Based Tests of Association*) développé par Horvath et collègues en 2001 et largement utilisé dans les études d'association employant les trios familiaux (Horvath et al., 2001).

Étant donné que plusieurs phénotypes et plusieurs SNPs sont testés pour une association dans une même étude, des erreurs de type I peuvent survenir en raison des tests multiples qui sont effectués simultanément (Newton-Cheh et Hirschhorn, 2005). En effet, plus de tests sont effectués à la fois, plus les probabilités de voir apparaître une valeur-p significative par hasard augmentent, d'où la possibilité d'apparition de résultats faux-positifs. Pour minimiser ces faux-positifs liés aux tests multiples, des corrections statistiques doivent être appliquées. Bien que plusieurs approches soient développées dans ce domaine (Cordell et Clayton, 2005; Newton-Cheh et Hirschhorn, 2005), la correction statistique employée dans toutes les études d'association de la présente thèse est une correction Bonferroni traditionnelle (Rao, 2008). Ce type de correction statistique est le plus conservateur et consiste, brièvement, à multiplier la valeur-p du test statistique par le nombre total de tests effectués (Newton-Cheh et Hirschhorn, 2005). Par exemple, si une étude d'association est effectuée pour quatre phénotypes de l'asthme et 90 SNPs, les valeurs-p obtenues pour chaque test devront être multipliées par le total des tests effectués, c'est-à-dire 360 tests (4 phénotypes X 90 SNPs). Donc, si la valeur-p d'un test est de 0,001, après correction elle sera de 0,36. Ceci départage les associations significatives des faux-positifs potentiels. La description des corrections statistiques est disponible dans la section des analyses statistiques dans chacun des articles présentant une étude d'association.

Les analyses statistiques représentent la dernière étape nécessaire à la réalisation d'une étude d'association (Figure 4) ainsi qu'à l'établissement des conclusions sur la présence d'une association entre l'asthme et un marqueur génétique donné. La suite logique est les études de réplification pour confirmer ces associations. Ainsi, une association statistiquement significative entre un SNP d'un gène candidat et le phénotype d'une maladie sera le résultat d'une de ces trois explications : 1) l'allèle a des conséquences fonctionnelles et affecte l'expression du phénotype; 2) l'allèle est en déséquilibre de liaison avec un autre locus lequel affecte l'expression du phénotype ; et 3) l'association est un faux-positif qui résulte d'une erreur de type I (erreur de génotypage, de phénotypage ou d'échantillonnage) (Silverman et Palmer, 2000; Hoffjan et al., 2003).

Plus d'une centaine d'études d'association basées sur l'approche par gènes candidats pour l'asthme et ses phénotypes ont été réalisées à ce jour dans différentes populations et ont fait l'objet de plusieurs articles de revue (Hoffjan et al., 2003; Cookson, 2004;



Wills-Karp et Ewart, 2004; Ober et Hoffjan, 2006; Bosse et Hudson, 2007; Guerra et Martinez, 2008; Vercelli, 2008; Zhang et al., 2008). À elle seule, la sélection des gènes candidats par clonage positionnel a permis d'identifier six gènes associés à l'asthme ou à ses phénotypes et répliqués dans plus d'une étude. Il s'agit des gènes *ADAM33*, situé dans la région 20p13 liée à l'asthme et à l'hyperréactivité bronchique, un membre de la famille des métalloprotéases exerçant diverses fonctions, incluant la dégradation des protéines de surface cellulaire comme les cytokines et leurs récepteurs (Van Eerdewegh et al., 2002) ; *DPP10*, localisé dans la région 2q14 liée à l'asthme, un homologue des dipeptidyl peptidases dont le rôle est de cliver les dipeptides terminaux des cytokines et des chimiokines (Allen et al., 2003) ; *PHF11*, situé dans la région 13q14 liée à l'atopie et à un taux élevé d'IgE sérique, une protéine intracellulaire à motifs doigts de zinc impliquée dans la régulation de la transcription (Zhang et al., 2003) ; *GPR4*, un récepteur couplé à une protéine G, situé dans la région 7p liée à la susceptibilité à l'asthme (Laitinen et al., 2004) ; *HLA-G*, un antigène des leucocytes humains, situé dans la région 6p21 liée à l'asthme et ses phénotypes (Nicolae et al., 2005) ; et *CYFIP2*, situé dans la région 5q33 liée à la susceptibilité à l'asthme, une protéine d'interaction cytoplasmique impliquée dans l'adhésion des lymphocytes T et dans l'induction de l'apoptose (Noguchi et al., 2005).

Pour ce qui est des autres approches de sélection des gènes candidats, elles ont permis d'identifier un total de 173 gènes associés avec l'asthme ou à ses phénotypes dans au moins une étude (Zhang et al., 2008). La revue de littérature a tout d'abord été effectuée par Hoffjan et collaborateurs en 2003, où 64 gènes étaient rapportés comme associés à l'asthme ou à ses phénotypes dans au moins une étude (Hoffjan et al., 2003). Trois ans plus tard, Ober et Hoffjan ont mis cette liste à jour et le nombre de gènes associés passa à 120 (Ober et Hoffjan, 2006). La dernière mise à jour de cette liste date de 2008 et inclut 53 nouvelles associations identifiées en 2006-2007 (Zhang et al., 2008). Ces articles de revue sont cités par l'ensemble de la communauté scientifique lorsqu'il est question de référer au nombre de gènes associés à l'asthme ou à ses phénotypes. En plus, la liste de tous les gènes connus pour être associés sont disponibles, ainsi que les références aux études qui les ont identifiés et répliqués. Par souci de brièveté, cette liste complète n'est pas incluse dans la présente thèse, mais est facilement accessible via les sites Internet des articles de revue en question.

À partir de ces informations, un regroupement des gènes associés selon leur rôle biologique et selon leur nombre d'associations dans des études indépendantes a été proposé par Vercelli (2008). Selon ce travail de synthèse, 33 gènes de susceptibilité pour l'asthme ou ses phénotypes ont été rapportés dans au moins cinq études d'association indépendantes (Vercelli, 2008). Le Tableau I présente ces 33 gènes, leur localisation chromosomique et une brève description de leur rôle biologique. Pour les références à chacune des études d'association, consulter (Vercelli, 2008).

**Tableau I**

Gènes de susceptibilité associés à l'asthme ou à ses phénotypes dans au moins cinq études d'association indépendantes et basées sur l'approche par gènes candidats. Traduit et modifié de (Vercelli, 2008).

Gène	Locus	Nom officiel <sup>a</sup>	Fonction et cascade biologique
<i>GSTM1</i>	1p13.3	<i>Glutathione S-transferase mu 1</i>	Interface environnementale, stress oxydatif et détoxification
<i>FLG</i>	1q21.3	<i>Filaggrin</i>	Intégrité de la barrière épithéliale
<i>IL10</i>	1q31-q32	<i>Interleukin 10</i>	Immunorégulation
<i>CTLA4</i>	2q33	<i>Cytotoxic T-lymphocyte-associated protein 4</i>	Inhibition et immunorégulation de la réponse des cellules T
<i>IL13</i>	5q31	<i>Interleukin 13</i>	Effecteur dans la différenciation des Th2
<i>IL4</i>	5q31.1	<i>Interleukin 4</i>	Effecteur dans la différenciation des lymphocytes Th2 et dans l'induction de IgE
<i>CD14</i>	5q31.1	<i>CD14 molecule</i>	Immunité innée, reconnaissance microbiale
<i>SPINK5</i>	5q32	<i>Serine peptidase inhibitor, Kazal type 5</i>	Inhibiteur des protéases épithéliales
<i>ADRB2</i>	5q31-q32	<i>Adrenergic receptor, beta 2</i>	Relaxation du muscle lisse
<i>HAVCR1</i>	5q33.2	<i>Hepatitis A virus cellular receptor 1</i>	Régulation de la réponse des cellules T
<i>LTC4S</i>	5q35	<i>Leucotriene C4 synthase</i>	Synthèse des leukotriènes cystéinyl, inflammation
<i>LTA</i>	6p21.3	<i>Lymphotoxin alpha (TNF superfamily, member 1)</i>	Inflammation
<i>TNF</i>	6p21.3	<i>Tumor necrosis factor (TNF superfamily, member 2)</i>	Inflammation
<i>HLA-DRB1</i>	6p21	<i>Major histocompatibility complex, class II, DR beta 1</i>	Présentation de l'antigène
<i>HLA-DQB1</i>	6p21	<i>Major histocompatibility complex, class II, DQ beta 1</i>	Présentation de l'antigène
<i>HLA-DPB1</i>	6p21	<i>Major histocompatibility complex, class II, DP beta 1</i>	Présentation de l'antigène
<i>GPRA</i>	7p14.3	<i>G protein-coupled receptor 154</i>	Régulation de la croissance cellulaire et des mécanismes neuronaux
<i>NAT2</i>	8p22	<i>N-acetyltransferase 2</i>	Détoxification des drogues et des carcinogènes
<i>MS4A2 (FCER1B)</i>	11q13	<i>Membrane-spanning 4-domains, subfamily A, member 2 (Fc fragment of IgE, high affinity 1, receptor for; beta polypeptide)</i>	Effecteur dans la réponse Th2
<i>CC16</i>	11q12.3-q13.1	<i>Clara-cell specific 10-kD protein</i>	Protéines anti-inflammatoire dérivée de l'épithélium
<i>GSTP1</i>	11q13	<i>Glutathione S-transferase pi 1</i>	Interface environnementale, stress oxydatif et détoxification
<i>IL18</i>	11q22.2-q22.3	<i>Interleukin 18</i>	Induit la production de l'interféron- $\gamma$ et du facteur nécrosant des tumeurs
<i>STAT6</i>	12q13	<i>Signal transducer and activator of transcription 6, interleukin-4 induced</i>	Signalisation des interleukines -4 et -13
<i>NOS1</i>	12q24.2-q24.31	<i>Nitric oxide synthase 1 (neuronal)</i>	Synthèse de l'oxide nitrique et communication cellulaire
<i>CMA1</i>	14q11.2	<i>Chymase 1, mast cell</i>	Sérine protéase chymotryptique des mastocytes
<i>IL4R</i>	16p12.1-p12.2	<i>Interleukin 4 receptor alpha chain</i>	Signalisation des interleukines -4 et -13
<i>CCL11</i>	17q21.1-q21.2	<i>Chemokine (C-C motif) ligand 11 (eotaxin)</i>	Chimioattractant des éosinophiles, dérivé de l'épithélium

<i>CCL5</i>	17q11.2-q12	<i>Chemokine (C-C motif) ligand 5 (RANTES)</i>	Chimioattractant pour les monocytes, les cellules T et les éosinophiles
<i>ACE</i>	17q23.3	<i>Angiotensin I converting enzyme (peptidyl-dipeptidase A)</i>	Inactivation des médiateurs inflammatoires
<i>TBXA2R</i>	19p13.3	<i>Thromboxane A2 receptor</i>	Contraction du muscle lisse, inflammation
<i>TGFBI</i>	19q13.1	<i>Transforming growth factor, beta 1</i>	Immunorégulation, prolifération cellulaire
<i>ADAM33</i>	20p13	<i>ADAM metallopeptidase domain 33</i>	Interactions cellule-cellule et cellule-tissu
<i>GSTT1</i>	22q11.23	<i>Glutathione S-transferase theta 1</i>	Interface environnementale, stress oxydatif et détoxification

<sup>a</sup> Source : NCBI (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene>)

Ces gènes peuvent être regroupés selon quatre rôles biologiques distincts impliqués dans la physiopathologie de l'asthme (Vercelli, 2008) : 1) les gènes associés avec l'immunité innée et l'immunorégulation (par exemple *CD14*, *IL10* ou les *HLA*) ; 2) les gènes associés avec la différenciation des lymphocytes Th2 et la fonction d'effecteur (par exemple *IL13*, *IL4*, *FCER1B* ou *STAT6*) ; 3) les gènes associés à l'épithélium, c'est-à-dire à l'interface de l'immunité innée et acquise (par exemple *CCL11*, *CCL5* ou *SPINK5*) ; et 4) les gènes associés à la fonction pulmonaire, le remodelage des voies aériennes ou la sévérité de l'asthme (par exemple *ADRB2*, *TNF*, *TGFBI* ou *LTA*). Bien que tous les 33 gènes listés dans le Tableau I représentent des facteurs de susceptibilité ou de protection potentiels pour l'asthme, huit d'entre eux sont identifiés comme étant les facteurs génétiques les plus importants dans le développement de l'asthme puisqu'ils ont été répertoriés dans plus de 20 études indépendantes (Vercelli, 2008). Il s'agit des gènes *IL13*, *IL4*, *CD14*, *ADRB2*, *TNF*, *HLA-DRB1*, *FCER1B* et *IL4R* (Hoffjan et al., 2003; Vercelli, 2008).

Malgré les nombreux gènes déjà associés à l'asthme, le portrait génétique de l'asthme n'est pas encore complet. D'autres études d'association sont nécessaires pour premièrement, confirmer les associations déjà observées, mais également pour en trouver de nouvelles, ce qui représente précisément le but ultime de la présente thèse et introduit le prochain chapitre consacré aux objectifs et hypothèses qui ont motivé les présents travaux de recherche.

## Chapitre 4

### Hypothèses et objectifs du projet de recherche

Le but principal du présent projet de recherche est de comparer différentes approches méthodologiques dans la sélection des gènes candidats pour réaliser des études d'association avec l'asthme et les phénotypes liés à l'asthme (hyperréactivité bronchique et atopie). Pour y arriver, différentes hypothèses et objectifs de recherche sont nécessaires, tels que présentés dans les prochaines sections où chacune des hypothèses est directement en lien avec son objectif et avec l'article et/ou le chapitre concerné.

#### 4.1 Hypothèse générale

Des déterminants génétiques de susceptibilité ou de protection sont impliqués dans la pathogénèse de l'asthme.

#### 4.2 Hypothèses spécifiques

- (1) Les gènes des molécules immunes, ciblés en raison de leur différence d'expression dans les biopsies bronchiques de sujets asthmatiques comparée à celle des témoins, sont des gènes candidats pour l'asthme et les phénotypes liés à l'asthme (Chapitre 5, deux premiers articles).
- (2) Le gène *PLAU*, ciblé en raison de son implication dans la cascade biologique du gène *SERPINB2* qui présente une différence d'expression dans les biopsies bronchiques de sujets asthmatiques comparée à celle des témoins, est un gène candidat pour l'asthme et les phénotypes liés à l'asthme (Chapitre 5, troisième article).
- (3) Les expérimentations de puces à ADN sont efficaces dans l'identification de biomarqueurs pour l'asthme (Chapitre 5, quatrième article).

- (4) Les gènes localisés dans des régions chromosomiques liées à l'asthme et à l'atopie, et ciblés par l'outil bioinformatique *Genes to Diseases* (G2D), sont des gènes candidats pour l'asthme et les phénotypes liés à l'asthme (Chapitre 6, cinquième article).
- (5) Les SNPs du locus 17q21, associés à l'asthme dans une GWAS, sont des polymorphismes candidats pour une étude de réplication avec l'asthme (Chapitre 7, sixième article).
- (6) Le gène filaggrin (*FLG*), identifié grâce à son association génétique avec la dermatite atopique qui partage une composante allergique avec l'asthme, est un gène candidat pour l'asthme et les phénotypes liés à l'asthme (Chapitre 8, septième article).
- (7) Les quatre approches méthodologiques employées (puces à ADN, G2D, association génétique connue avec l'asthme et association génétique connue avec un autre trait complexe partageant une composante physiopathologique commune à l'asthme) sont efficaces dans l'identification de gènes candidats destinés aux études d'association génétique pour l'asthme (Chapitre 9, Discussion).
- (8) Les gènes associés selon l'une ou l'autre des approches employées ont un rôle dans la susceptibilité et/ou la protection pour l'asthme et ont une implication biologique dans sa physiopathologie (Chapitre 10, Perspectives de recherche).

### **4.3 Objectif général**

Identifier des déterminants génétiques de susceptibilité ou de protection pour l'asthme et ses phénotypes associés à l'aide de quatre différentes approches méthodologiques de sélection des gènes candidats.

### **4.4 Objectifs spécifiques**

Afin d'alléger le texte de cette sous-section, les symboles officiels des gènes ont été utilisés. Pour une description rapide de ces symboles, consulter la liste des abréviations

à la page vx. Pour une information plus complète, incluant la fonction générale de ces gènes, se référer aux articles concernés indiqués entre parenthèses.

- (1) Effectuer une étude d'association entre les phénotypes de l'asthme et les SNPs (et/ou TagSNPs) de neuf gènes de molécules de signalisation immunes différemment exprimés dans les biopsies bronchiques de sujets asthmatiques (comparativement à des sujets témoins) dans l'échantillon de trios familiaux du SLSJ. Par la suite, reproduire les résultats dans des échantillons indépendants comme ceux des trios familiaux canadiens (CAPPS et SAGE) ou de l'échantillon cas-témoins australien (Busselton).  
Plus précisément, avec des SNPs (et/ou TagSNPs)
  - (1a) du gène *CX3CR1* (Chapitre 5, premier article) ;
  - (1b) des gènes *ALOX15*, *CD14*, *CD27*, *CXCL12*, *IL2RB*, *IL7R*, *NOS2A* et *SFRP1* (Chapitre 5, deuxième article).
  
- (2) Effectuer une étude d'association entre les phénotypes de l'asthme et les SNPs du gène *PLAU* impliqué dans la cascade biologique du gène *SERPINB2* (codant pour une enzyme protéolytique) différemment exprimé dans les biopsies bronchiques de sujets asthmatiques comparativement aux sujets témoins dans l'échantillon de trios familiaux du SLSJ mais qui n'est pas associé à l'asthme dans cet échantillon. Par la suite, reproduire les résultats dans des échantillons indépendants des trios familiaux canadiens CAPPS et SAGE (Chapitre 5, troisième article).
  
- (3) Démontrer l'efficacité des expérimentations de puces à ADN dans l'identification de biomarqueurs différemment exprimés dans les biopsies bronchiques de sujets asthmatiques comparés à celles des sujets témoins (Chapitre 5, quatrième article).
  
- (4) Effectuer une étude d'association entre les phénotypes de l'asthme et des SNPs (et TagSNPs) de dix gènes (*LPA*, *NOX3*, *SNX9*, *VIL2*, *VIP*, *ADAM8*, *DOCK1*, *FANK1*, *GPR123*, *PTPRE*), localisés dans deux régions chromosomiques liées à l'asthme et/ou à l'atopie au SLSJ (6q26 et 10q26.3) et ciblés par l'outil informatique G2D, dans l'échantillon de trios familiaux du SLSJ. Par la suite, répliquer les associations positives dans l'échantillon indépendant de trios

familiaux provenant de la cohorte de naissance américaine CAMP (Chapitre 6, cinquième article).

- (5) Effectuer une étude de réplication de l'association génétique des dix SNPs du chromosome 17q21, associés à l'asthme dans une GWAS effectuée sur une cohorte de sujets asthmatiques provenant de l'Europe, dans l'échantillon de trios familiaux du SLSJ (Chapitre 7, sixième article).
- (6) Effectuer une étude d'association entre l'asthme et les phénotypes liés à l'asthme et les mutations connues ainsi que des TagSNPs du gène *FLG* dans l'échantillon de trios familiaux du SLSJ. Par la suite, séquencer le gène *FLG* afin d'identifier les mutations causales (Chapitre 8, septième article).
- (7) Synthétiser l'information génétique recueillie et comparer l'efficacité des quatre approches méthodologiques dans l'identification de gènes candidats pour des études d'association génétique pour l'asthme (Chapitre 9, Discussion).
- (8) Résumer l'implication biologique potentielle des gènes associés dans la physiopathologie de l'asthme afin de proposer de nouvelles hypothèses et perspectives de recherche (Chapitre 10, Perspectives de recherche).

## Chapitre 5

### **Les gènes ciblés par une étude du profil d'expression de millier de gènes dans les biopsies bronchiques de sujets asthmatiques comparés aux sujets témoins**

L'analyse du transcriptome de sujets asthmatiques effectuée par les travaux conjoints des Drs Laprise et Laviolette a permis d'identifier plusieurs biomarqueurs génétiques, répartis dans une douzaine de catégories biologiques (incluant les molécules immunes et les enzymes protéolytiques), différemment exprimés dans les biopsies bronchiques de sujets asthmatiques comparés à des témoins (Laprise et al., 2004). C'est à partir de ces résultats innovateurs que le présent projet de doctorat a vu le jour. Ainsi, basé sur l'hypothèse que cette différence d'expression s'explique à un niveau génétique, neuf gènes de molécules immunes (Tableau II) et un gène impliqué dans la voie biologique d'une enzyme protéolytique furent sélectionnés comme gènes candidats pour des études d'association avec l'asthme et ses phénotypes. Ce projet de recherche a mené à la publication de trois études d'association et d'une étude comparative, faisant chacune l'objet d'une section du présent chapitre.

La section 5.1 présente le tout premier article du présent projet doctoral et concerne l'étude d'association génétique entre le gène *CX3CR1* (Tableau II) et l'asthme. À cette époque, en 2004, ce gène immunitaire représentait le meilleur potentiel pour une étude d'association et c'est pourquoi il a été génotypé avant les huit autres. L'association fortement significative et sa réplique dans une étude indépendante ont motivé la publication dans le journal *Genes and Immunity*. J'ai effectué toutes les analyses statistiques, l'interprétation des résultats et la rédaction du manuscrit. Le Dr Mathieu Lemire a validé les analyses statistiques. Tomi Pastinen et Yannick Renaud ont effectué le génotypage de l'échantillon du SLSJ et l'équipe du Dr Andrew Sandford a effectué le génotypage pour l'échantillon de Vancouver. Le Dr Laviolette a contribué à l'interprétation grâce à ses connaissances de la physiopathologie de l'asthme et a révisé le manuscrit. La Dre Laprise a conçu l'étude, dirigé toutes les étapes et a révisé le manuscrit. Tous les coauteurs ont révisé le manuscrit.



**Tableau II**

Différences dans l'expression des gènes des molécules immunes entre les sujets témoins et asthmatiques (adapté de Laprise et al. 2004).

Gène	NCBI GeneID	Locus	Nom officiel <sup>a</sup>	Fonction <sup>b</sup>	Témoins (T)		Asthme (A)		A vs T <sup>c</sup>	
					Moyenne <sup>c</sup>	SD	Moyenne <sup>c</sup>	SD	MFC <sup>d</sup>	Valeur-p <sup>f</sup>
<i>ALOX15</i>	246	17p13.3	<i>Arachidonate 15-lipoxygénase</i>	Impliquée dans les processus anti-inflammatoires, de remodelage membranaire et de développement du cancer/métastases.	3,59	0,08	4,21	0,09	4,190	0
<i>CD14</i>	929	5q31.1	<i>CD14 molécule</i>	Augmente la régulation des molécules de surface incluant les molécules d'adhésion. Active le NF-kappa-B et la sécrétion de cytokines dans la réponse inflammatoire.	3,38	0,08	2,94	0,10	-2,764	4,41E-10
<i>CD27</i>	939	12p13	<i>CD27 molécule</i>	Rôle dans l'activation des lymphocytes B et T (réduit la production d'IgE).	3,19	0,08	2,63	0,13	-3,606	8,13E-13
<i>CX3CR1</i>	1524	3p21.3	<i>Chemokine (C-X3-C motif) receptor 1</i>	Rôle dans le recrutement, l'adhésion et la migration des lymphocytes et des monocytes.	2,90	0,08	2,55	0,10	-2,210	1,51E-07
<i>CXCL12</i>	6387	10q11.1	<i>Chemokine (C-X-C motif) ligand 12</i>	Chémoattractant actif dans les lymphocytes T et les monocytes, mais pas dans les neutrophiles.	3,09	0,08	2,71	0,12	-2,392	8,07E-07
<i>IL2RB</i>	3560	22q13.1	<i>Interleukin 2 receptor, beta</i>	Impliqué dans l'endocytose et la transduction des signaux mitogénique de l'IL-2.	3,08	0,08	2,60	0,14	-2,997	1,42E-08
<i>IL7R</i>	3575	5p13	<i>Interleukin 7 receptor</i>	Rôle important dans la prolifération et la maturation des cellules B.	3,32	0,08	2,94	0,10	-2,408	3,22E-09
<i>NOS2A</i>	4843	17q11.2-q1	<i>Nitric oxide synthase 2, inducible</i>	Vasodilatateur, neurotransmetteur et médiateur inflammatoire dans les voies respiratoires. Il est produit en grande quantité dans l'asthme.	2,72	0,10	3,58	0,09	7,251	0
<i>SFRP1</i>	6422	8p12-p11.1	<i>Secreted frizzled-related protein 1</i>	Récepteurs transmembranaires. Rôle dans la croissance et la différenciation cellulaires.	3,17	0,08	2,84	0,11	-2,155	8,87E-07

**Abréviations utilisées :** NCBI = Centre national d'information biotechnologique, SD = Déviation standard, A = Asthmatique, T = Témoin.

<sup>a</sup> Source : NCBI (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene>)

<sup>b</sup> Tirées de bases de données publiques disponibles en ligne : *Swiss-Prot* (<http://ca.expasy.org/sprot/sprot-top.html>), NCBI (<http://www.ncbi.nlm.nih.gov/LocusLink/list.cgi>) et/ou *Genecards* (<http://bioinfo1.weizmann.ac.il/genecards/>).

<sup>c</sup> Moyenne de la mesure d'expression en valeur logarithmique.

<sup>d</sup> MFC = *Mean fold change*, désigne la différence d'expression moyenne entre les sujets A et T. Une valeur positive indique une sur-expression chez les sujets asthmatiques.

<sup>e</sup> Asthmatique et témoins sélectionnés pour exclure les gènes avec une différence d'expression plus petite que deux.

<sup>f</sup> Valeurs-p identifiées par 0 sont très petites ( $p < 1 \times 10^{-13}$ ).

Le succès obtenu avec l'approche employée pour sélectionner le gène *CX3CR1* a motivé l'étude d'association des huit autres gènes des molécules de signalisation immune (Tableau II), présentée dans l'article de la section 5.2. Le génotypage des huit gènes a été effectué dans le contexte d'AllerGen Inc. dans quatre échantillons indépendants de sujets asthmatiques (les échantillons de trios familiaux du SLSJ, CAPPS et SAGE et l'échantillon cas-témoins de Busselton) sous la responsabilité d'Alexandre Montpetit. Les résultats significatifs et résistants aux corrections statistiques obtenus pour le gène *ALOX15* dans deux des échantillons ont motivé leur publication dans le *Journal of Allergy and Clinical Immunology*. L'article fut également sélectionné par l'éditeur sous la rubrique « *Editor's Choice* » de cette revue médicale. Pour cette étude, j'ai effectué les analyses statistiques, l'interprétation des résultats et la rédaction du manuscrit. J'ai également contribué à la coordination entre les coauteurs. Les Dr Sandford, Chan-Yeung et Paré représentent l'échantillon CAPPS, les Drs Becker et Kozyrskyj représentent l'échantillon SAGE, et les Drs James et Musk représentent l'échantillon de Busselton. Les Drs Daley et Lemire ont validé les analyses statistiques et contribué à la rédaction du manuscrit. La Dre Laprise a conçu l'étude, dirigé toutes les étapes et révisé le manuscrit. Tous les coauteurs ont contribué à la révision finale du manuscrit.

Le troisième article en lien avec l'hypothèse des puces à ADN, présenté dans la section 5.3, est une étude d'association entre le gène *PLAU* et les phénotypes de l'asthme. Il s'agit du projet de maîtrise de Philippe Bégin pour lequel j'ai collaboré. Le gène *PLAU*, localisé dans la région chromosomique 10q24 liée à l'asthme dans l'échantillon familial du SLSJ, code pour l'activateur du plasminogène de type urokinase (u-PA) et est impliqué dans la réponse immunitaire et le remodelage tissulaire. L'activité de l'u-PA serait inhibée par l'inhibiteur des activateurs de plasminogène (PAI-2), enzyme protéolytique codée par le gène *SERPINB2*. L'expression de *SERPINB2* a d'ailleurs été montrée comme 11 fois plus élevée dans les biopsies bronchiques de sujets asthmatiques comparée aux témoins (Laprise et al. 2004), mais ce gène n'a montré aucune association avec l'asthme au SLSJ (résultats non présentés). Le gène *PLAU* a donc été sélectionné comme gène candidat sur la base de sa localisation chromosomique et de son rôle dans la voie biologique de l'enzyme protéolytique PAI-2. L'article, publié dans l'*American Journal of Respiratory Critical Care Medicine*, rapporte l'association positive du gène *PLAU* dans l'échantillon de trios familiaux du SLSJ et sa répllication

dans deux échantillons de trios familiaux indépendants (CAPPS et SAGE). Pour cette étude, Philippe Bégin a effectué l'ensemble des travaux incluant la sélection des SNPs, le génotypage, les analyses ainsi que rédigé le manuscrit. Pour ma part, j'ai participé à l'ensemble des travaux de laboratoire relatifs au génotypage de l'échantillon des trios familiaux du SLSJ et j'ai contribué à la rédaction manuscrit, à la validation des résultats et à toutes les révisions demandées. Sabine Kacel (étudiante à la maîtrise) a effectué le séquençage aidé de Sébastien Claveau (professionnel de recherche). Charleen Salesse (B.Sc.) et Alexandre Montpetit (Ph.D.) ont participé au génotypage. Les Drs Chang-Yeung, Becker et Kozyrskyj sont les principaux investigateurs pour les échantillons CAPPS et SAGE. La Dre Denise Daley a participé aux analyses statistiques. La Dre Laprise a conçu l'étude, dirigé toutes les étapes et révisé le manuscrit. Tous les coauteurs ont contribué à la révision finale du manuscrit.

Le dernier article de ce chapitre, section 5.4, présente une étude comparative entre deux expérimentations indépendantes de puces à ADN. Cette étude s'inscrit dans les travaux de doctorat d'Annie Chamberland, qui a réalisé les analyses et rédigé le manuscrit qui est présentement en révision à la revue *Experimental Lung Research*. Anne-Marie Madore, étudiante au doctorat, a contribué aux analyses et à la revision du manuscrit. Pour ma part, j'ai participé à la rédaction et à la révision du manuscrit en plus de contribuer à la présentation et à l'interprétation des résultats. Le Dr Laviolette a révisé le manuscrit. La Dre Laprise a conçu l'étude, dirigé toutes les étapes et révisé le manuscrit. Étant donné que les résultats contenus dans cet article contribuent à l'avancement et à l'interprétation des résultats présentés dans les études d'association des sections 5.1 et 5.2, les éléments apportés par cette étude ont été jugés pertinents au sujet de cette thèse.

## 5.1 Premier article

### Association entre le gène *CX3CR1* et l'asthme

Karine Tremblay, Mathieu Lemire, Véronique Provost, Tomi Pastinen, Yannick Renaud, Andrew J Sandford, Michel Laviolette, Thomas J Hudson et Catherine Laprise

*Genes & Immunity* (2006) 7: 632-639

#### Résumé

Le gène *CX3CR1* code pour le récepteur de la fractalkine afin de permettre l'adhésion et la migration de divers leucocytes dans les sites d'inflammation. Puisque *CX3CR1* est moins exprimé dans les tissus bronchiques de sujets asthmatiques, nous avons émis l'hypothèse que des variations génétiques pourraient être associées à la susceptibilité à l'asthme. L'objectif de cette étude est donc d'effectuer une étude d'association génétique ainsi que les analyses haplotypiques de certains polymorphismes (SNPs) du *CX3CR1* dans un échantillon familial de sujets asthmatiques originaires d'une population à effet fondateur. Les analyses génétiques effectuées à l'aide du logiciel FBAT ont montré que cinq SNPs du *CX3CR1* (rs938203, rs2669849, rs1050592, T280M and V249I) présentent une association significative entre leur allèle fréquent et l'asthme sous un modèle dominant ( $p < 0,004$ ). Un haplotype formé avec les allèles communs des SNPs rs1050592, T280M and V249I est aussi plus transmis chez les probands asthmatiques sous un modèle dominant ( $p = 0,005$ ). Les associations des SNPs V249I et rs2669849 ont été validées dans un échantillon cas-témoins indépendant. Pour le V249I, les odds ratio sont de 2,16 (homozygotes fréquents) et de 2,11 (hétérozygotes) sous le modèle dominant ( $p = 0,031$ ). Pour le rs2669849, les odds ratio sont de 2,75 (homozygotes fréquents) et de 1,86 (hétérozygotes) sous un modèle additif ( $p = 0,007$ ) et dominant ( $p = 0,059$ ). Ces résultats suggèrent un effet protecteur des allèles mineurs chez les porteurs sains. Des études fonctionnelles du *CX3CR1* et des SNPs identifiés sont nécessaires afin de documenter son rôle dans la pathophysiologie relative à l'asthme.

### 5.1.1 TITLE PAGE

#### **Association study between the CX3CR1 gene and asthma**

Karine Tremblay <sup>1,2</sup>, Mathieu Lemire <sup>3</sup>, Véronique Provost <sup>2</sup>, Tomi Pastinen <sup>3</sup>, Yannick Renaud <sup>3</sup>, Andrew J Sandford <sup>4</sup>, Michel Laviolette <sup>2</sup>, Thomas J Hudson <sup>3</sup> and Catherine Laprise <sup>1,5</sup>

<sup>1</sup> University of Montreal Community Genomic Medicine Centre, Chicoutimi University Hospital, Saguenay (QC), Canada; <sup>2</sup> Hôpital Laval, Institut universitaire de cardiologie et de pneumologie de l'Université Laval, Québec (QC), Canada; <sup>3</sup> McGill University and Genome Québec Innovation Centre, Montreal (QC), Canada; <sup>4</sup> James Hogg iCapture Centre, University of British Columbia, Vancouver (BC), Canada; <sup>5</sup> Université du Québec à Chicoutimi, Saguenay (QC), Canada.

#### **Corresponding Author**

Catherine Laprise, Ph.D.  
Université du Québec à Chicoutimi  
555, boulevard de l'Université,  
Chicoutimi (Québec), Canada, G7H 2B1  
Telephone :(418) 541-1234, 3070  
Fax :(418) 541-1116  
Email: claprise@saglac.qc.ca

#### **Running title**

Association between CX3CR1 and asthma

#### **Keywords**

Association study, asthma, CX3CR1, polymorphisms, haplotypic analyses

This article has an online data supplement, which is accessible from the issue's table of content online at <http://www.nature.com/gene/index.html>.

### 5.1.2 ABSTRACT

CX3CR1, a fractalkine receptor, mediates cell adhesive and migratory functions in inflammation. Based on CX3CR1 expression observed in bronchial tissues of asthmatic subjects, we hypothesized that genetic variation at this locus may affect susceptibility to asthma. OBJECTIVE : To carry out an association study and haplotypic analyses with selected polymorphisms of the CX3CR1 in a asthmatic families sample from a founder population. RESULTS : Genetic analyses performed by FBAT showed five CX3CR1 SNPs (rs938203, rs2669849, rs1050592, T280M and V249I) with significant associations between their common alleles and asthma ( $p < 0.004$ ) under a dominant model. A haplotype formed with common alleles of rs1050592, T280M and V249I is also overtransmitted in asthmatic subjects ( $p = 0.005$ ) under a dominant model. The associations of V249I and rs2669849 have been validated in a independent case-control sample. For V249I, OR are 2.16 (common homozygous) and 2.11 (heterozygous) under dominant model ( $p = 0.031$ ). For rs2669849, OR are 2.75 (common homozygous) and 1.86 (heterozygous) under additive model ( $p = 0.007$ ) and dominant model ( $p = 0.059$ ). These results suggest an asthma protective effect of the minor alleles for healthy control carriers. Further functional studies of CX3CR1 are needed to document its role in asthma pathophysiology.

### 5.1.3 INTRODUCTION

Asthma, a multifactorial respiratory disease, involves genetic and environmental factors in its development, chronicity and severity (1, 2). The pathophysiology of asthma is characterized by airway inflammation and remodeling processes modulated by resident and inflammatory cells notably epithelial cells, eosinophils and T lymphocytes, particularly of the Th2 subset (3, 4). Although some of the underlying mechanisms have been elucidated in recent years, more work is needed to gain a clearer understanding of the genetic and environmental factors that contribute to the development of different forms of asthma.

Numerous genetic linkage and association studies were performed within the last decade - see Ober & Hoffjan (5) for a review – as well as the considerable advances in laboratory and statistical methodologies. However, more works are needed to characterize genetic risk factors for asthma and its sub-phenotypes. New technological

advances, including large single nucleotide polymorphism (SNP) databases, high-resolution haplotype structure analysis, and global gene expression profiling from normal and affected tissues may be useful in the screening for new genes and pathways relevant to asthma. In this context, we applied high-density microarray technology to bronchial biopsies and identified 79 genes that show a significant differential expression in asthmatics compared to controls, including numerous genes previously implicated in asthma as well as new potential candidates, such as CX3CR1 (6).

The CX3CR1 gene, localized in the 3p21.3 chromosomal region, encodes for a high affinity seven-transmembrane G-protein coupled receptor specific for the fractalkine (FKN) chemokine (7). CX3CR1 has been shown to be expressed on inflammatory leukocytes such as natural killer cells, monocytes and T lymphocytes (8-11). FKN expression can be observed on IL-1 $\beta$ , IFN $\gamma$  and TNF $\alpha$  activated primary endothelial cells (12, 13), on bronchial epithelial cells (14) and on the cells of many other tissues such as lung, brain and intestines (12). This chemokine exists in two forms; as a membrane-anchored mucin-like stalk proadhesive protein that mediates strong adherence of cytotoxic leukocytes expressing CX3CR1 and leads to their migration into the endothelium to promote tissue damage (8, 9, 12) or as a soluble chemotactic peptide that recruits CX3CR1 inflammatory expressing cells to induce their cellular adhesion into the inflammation sites (9). A role for CX3CR1-FKN mediated inflammation has been suggested in various inflammatory diseases including vascular injury (15), atopic dermatitis (16) and allergic airway diseases (17, 18). To our knowledge, no genetic association study has been reported between CX3CR1 and asthma. However, two coding SNPs (T280M and V249I) have been shown to affect risk to other diseases involving immune and inflammatory diseases (19-23).

Based on our observations that the CX3CR1 gene is less expressed in bronchial biopsies of asthmatic subjects (6), on its biological importance in the recruitment of leukocytes into inflammatory sites (8, 9, 12) and on its potential role in allergic asthma (17, 18), we hypothesize that CX3CR1 is a candidate gene in asthma pathogenesis. The aims of this study were to carry out an association study and an haplotypic analysis with selected SNPs in CX3CR1 gene in a founder effect familial asthmatic sample from Northeastern Quebec.

#### 5.1.4 RESULTS

##### **Family-based association analysis of CX3CR1 SNPs with asthma and related phenotypes**

One thousand one hundred and thirty nine individuals, aged 3 to 88 years (Table 1), were included in the study (570 asthmatics, of which 223 were probands and 347 were family members). The median age of asthma onset for cases and their affected siblings was 5 years (2 – 46 years). The sex ratio (male:female) was 1:1.2 for probands, 1:1.4 for affected family members and 1:1.2 for unaffected family members. The eleven SNPs were tested individually for association with asthma, airway hyperresponsiveness (AHR), atopy and elevated IgE level under an additive and a dominant genetic model. Given that the analyses showed non-significant results for the AHR, atopy and elevated IgE level phenotypes, the results presented in this section and their interpretation are restricted to the asthma phenotype. Five mendelian errors have been detected in the entire sample for all the SNPs tested: the genotypes of these families were considered as missing data in the analysis. None of the SNPs gave a significant deviation from equilibrium when considering the unrelated founders of the sample. Calculation of effective number of independent tests was done using SNPSdP (24) (see online supplement for detail on simulation parameters) and allowed to find that the eleven correlated SNPs correspond to eight effective independent ones. The simulation study indicates that the four phenotypes studies correspond to between 2.85 and 3.34 effective independent ones, depending on the minor allele frequency of the SNP simulated, with a median of 3.09. Hence, an estimate of the effective number of tests that we performed at this stage is roughly 42 (7 SNPs, 3 phenotypes, 2 genetic models). The lowest p values observed (0.00058, rounded up to 0.001 in Table 2) is thus still significant even after allowing for the total number of tests that we performed with a Bonferroni correction with the effective number of independent tests (corrected p value:  $0.00058 \times 42 = 0.0243$ ).

Under an additive model, no significant associations were found (data not shown). Under a dominant model, results are shown in Table 2. Five common alleles (rs1050592, T280M, V249I, rs938203 and rs2669849) were significantly overtransmitted to asthmatic offspring (p values < 0.004). The association observed with V249I and rs2669849 has been independently replicated in a case-control sample from Vancouver/Winnipeg. The Armitage test for trend (Table 3) supports the



susceptibility to the same alleles. In Table 3 , the odds ratios (OR) are expressed under the arbitrary choice to describe the common allele as being a susceptibility allele. Equivalently, the rare allele could be considered to be a protective allele: this would not have an effect on the association results. For V249I, the OR are approximately the same for both common homozygous and heterozygous (2.16 and 2.11 respectively) which supports a dominant model and the association with the common allele (p value = 0.031). For rs2669849, the dominant model no longer best describes the data (the risks are 2.75 for the common homozygous and 1.86 for the heterozygous); the association is nevertheless also with the common allele (p value = 0.007 under an additive model; p=0.059 under a dominant model).

### **CX3CR1 linkage disequilibrium and haplotype specific association analyses**

Figure 1A shows the linkage disequilibrium (LD) between the SNPs measured by D'. The eleven SNPs separated the region into four blocks of tightly associated SNPs. We used Gabriel's criteria implemented in Haploview 3.2 for block definition (see methods). Block 1 is located towards the 3'end and exon 2 and consists of SNPs rs1050592, T280M and V249I. Block 2 is located in the intronic region and consists of SNPs rs2853712 and rs2669841. Block 3 is located towards intronic and 5'end regions and consists of SNPs rs2853711, rs871610 and rs871144. Block 4 is located towards the 5'end and consists of SNPs rs938203, rs3020453 and rs2669849. Common haplotypes were observed in each block : block 1, haplotypes TGC (0.69), CAT (0.15) and CGT (0.14); block 2, GC (0.49), AT (0.23) and AC (0.28); block 3, ACA (0.53), CCG (0.23), ATA (0.23) and ACG (0.02); and block 4, AAG (0.25), GAA (0.16), AAA (0.39) and AGA (0.20) (Figure 1B). To characterize common CX3CR1 haplotype transmissions, previously inferred using Haploview, we assessed for non-random transmission using FBAT (v1.4) software (25, 26) under a dominant genetic model. For asthma, only one haplotype (TGC) of the first block showed non-random transmission (overtransmitted, p = 0.005) (Table 4). The other haplotypes tested (including all blocks combinations) were non significant (data not presented).

### **5.1.5 DISCUSSION**

We report an association between five genetic variants and one haplotype of the CX3CR1 gene and asthma in a French-Canadian sample. The selection of CX3CR1 was

based on our previous report of its differential expression in asthmatic bronchial biopsies compared to control subjects using microarrays (6), on its biological functions in inflammatory processes such in allergic asthma (17) and on its genetic association with other inflammatory complex traits (19-21). This study demonstrates how microarray studies can be used to motivate original areas of investigation and to provide new research hypotheses to highlight candidate genes which have not been identified by classical genome screens (27).

To our knowledge, this genetic association study is the first to investigate CX3CR1 with asthma. However, the CX3CR1 gene has been tested with other complex traits where the less common alleles of two SNPs in strong linkage disequilibrium, I249 and M280 (22), were associated with both increased risk of HIV infection and accelerated HIV disease progression (22, 23), with reduced prevalence of coronary events and atherosclerosis (19-21) but not with peripheral arterial diseases (28) or cerebrovascular diseases (29). The associations with HIV infection (22, 23) have not been confirmed by an independent study of McDermott and colleagues (30). We tested these SNPs in the present sample and found positive associations with asthma. According to our results, the common alleles of these two SNPs (V249 and T280) are overtransmitted in asthmatic probands. These single marker association results were further supported by the pairwise disequilibrium pattern analysis that showed strong linkage disequilibrium (0.98) between these two SNPs and with the rs1050592 in the block 1 haplotype. Haplotype TGC spanning in this block was also overtransmitted in asthmatic probands, suggesting a significant non-random distribution of this haplotype for asthma. As the minor alleles were undertransmitted to asthmatic probands, this positive association suggests that the less common I249/M280 haplotype could represent a protective factor for asthma. Concerning the association with the three other SNPs involving variants in the promoter region (rs938203, rs3020453, rs2669849), these genetic markers could be in linkage disequilibrium with the CX3CR1 V249I/T280M studied variants (LD = 0.88, 0.68, 0.63, respectively) or could be associated with regulatory variants affecting expression. The latter is also suggested by the reported imbalance between expressed alleles of CX3CR1 in heterozygous lymphoblastoid cell lines (31, 32).

Family-based tests of association are not robust in the presence of Mendelian-consistent genotyping errors: these undetected errors give the false impression of overtransmission of the common alleles from parents to offspring (33). Overtransmission of common alleles is precisely what we observed. It is thus reassuring to see a validation in a case-control study: even though undetected genotyping errors can change the outcome of association results, they do not inflate the type I error of the test (34). It could be suggested that these common alleles are susceptibility alleles in asthma, that moreover seem to act in a dominant fashion, indicating a role of this CX3CR1-FKN pathway in the disease. The association of V249I and rs2669849 has been validated in the Vancouver/Winnipeg case-control sample in our preliminary works, and additional association studies are in process using three Canadian asthma studies with 3600 individuals (AllerGen Network of Excellence).

Because CX3CR1 is preferentially expressed in Th1 cells compared to Th2 cells in response to FKN, cell infiltration via FKN-CX3CR1 interaction promotes Th1 response and may contribute to the pathogenesis of numerous vascular and tissue injury (13, 15, 35, 36). FKN expressed on inflamed endothelium may play a role as a vascular gateway for cytotoxic effector cells expressing CX3CR1 by rapidly capturing them from the blood and by promoting their migration into the tissue, where inflammation and Th1 polarization may occur through IFN $\gamma$  production (15). CX3CR1 I249/M280 haplotype has been associated with a reduced risk of coronary events (19-21) that could be explained by a decreased ability of circulating monocytes harboring CX3CR1 I249/M280 haplotype to adhere to vascular endothelium as a result of a decreased binding affinity of CX3CR1 for its ligand (15). This decrease in monocyte infiltration results in a reduced inflammation and Th1 response responsible for this protective effect in cardiovascular disease (21). On the other hand, HIV infected individuals carrying this I249/M280 haplotype also presented a reduced binding affinity between CX3CR1 and FKN, resulting in a decreased Th1 response and host cellular defense that would lead to the more rapid development of AIDS (22). Based on these genetic and functional studies and because FKN is also expressed on human bronchial epithelial cells (14), we hypothesize that common CX3CR1 V249/T280 haplotype may be a genetic risk factor for asthma possibly by supporting inflammatory leukocytes entry in the bronchial airway walls and by enhancing the inflammatory response. Then, the individuals carrying the I249/M280 haplotype are protected against the development of

asthma by this decrease in their inflammatory response, as observed in cardiovascular diseases. Delineation of the real contribution of the genetic risks associated to CX3CR1 in asthma pathogenesis awaits further characterization of the molecular structure of the gene and of the functional consequences of the identified variants on lung physiology.

In this regard, we plan to compare CX3CR1 level of protein expression in bronchial biopsies of asthmatic and control subjects and to characterize the cellular source of their mRNA expression. As FKN is constitutively expressed in pulmonary endothelial and epithelial cells contributing to the recruitment and the persistence of inflammatory cells in the airways of allergic individuals (17), we believe that further functional works on the CX3CR1 and the FKN will help delineate the impact of this immune-inflammatory pathway in the asthma pathophysiology.

In conclusion, microarrays allowed us to identify CX3CR1 as a new candidate gene for asthma and our association studies showed a potential susceptibility risk if carrying CX3CR1 common alleles of five SNPs, suggesting an asthma protective effect of the minor alleles for healthy control carriers. Based on previous genetic and functional studies, we hypothesize that CX3CR1 could be implicated in the asthma pathophysiology, but we actually pursue the functional studies of either associated SNPs in asthma pathogenesis to elucidate the real biological relevance of CX3CR1 in asthma.

#### **5.1.6 SUBJECTS AND METHODS**

##### **Samples**

The study sample, described in a recent report (37), comprises asthmatic probands residing in Saguenay–Lac-St-Jean (SLSJ), a region of Northeastern Quebec (Canada). In this region resides a founder population (38-41) of approximately 287,000 individuals who descend predominantly from approximately 2500 founders originating from France that settled in Quebec in the 17<sup>th</sup> century (42). All information concerning the recruitment and the clinical evaluation of the subjects are available in this journal's online supplement as well as the information concerning the Vancouver/Winnipeg case-control sample (43) used to validate associations.

### **Phenotyping**

Following the American Thoracic Society's criteria (44), we defined participants as asthmatics if 1) they had a reported history of asthma (validated by a physician), or if 2) they showed a positive methacholine provocation test (PC<sub>20</sub>) and asthma related symptoms. Airway hyperresponsiveness (AHR) was defined as a PC<sub>20</sub> methacholine  $\leq$  8 mg/ml. If PC<sub>20</sub> was unavailable, AHR was defined as either a rise in the post-bronchodilator forced expiratory volume in one second (FEV<sub>1</sub>) of  $\geq$  15% from baseline or a variation of the morning-evening peak expiratory flow rate (PEFR) of  $\geq$  12% over a two-week period (45). Atopy was defined as at least one positive response (wheal diameter  $\geq$  3 mm at 10 minutes) on skin prick tests and elevated serum IgE level was fixed at a concentration of 100 mg/l following the sample distribution.

### **Genotyping**

Blood samples were drawn from all participants and DNA was extracted from whole blood leukocytes using Genomic-tip 100/G kit (Qiagen Inc.) according to manufacturer's instructions. A panel of twelve SNPs was selected from public databases (NCBI and The SNP Consortium) according to the presence of allelic imbalance in a SLSJ sample previously noted by our group (data not shown) and tested for association with each asthma subphenotype. One SNP, rs2669850, was excluded due to poor performance of the genotyping assay. SNPs described in this report are cited using their reference SNP identifier (rs#) from the NCBI database. See Table 5 for CX3CR1 SNP characteristics and Figure 2 for CX3CR1 genomic organization. Among the eleven working SNPs, there were two nonsynonymous coding variants in the transmembrane domain of the receptor, T280M (a G-to-A substitution changed Thr<sup>280</sup> to Met) and V249I (a C-to-T substitution changed Val<sup>249</sup> to Ile), five 5'UTR variants, one 3'UTR variant and three non-coding variant. SNPs genotyping was performed using SNPstream® UHT (Orchid Biosciences) (46) and HEFP™ (Molecular Devices) (47). All protocols, primers and conditions used are available in this journal's online repository.

### **Statistical analysis**

For the SLSJ familial sample, allele distribution patterns and Mendelian errors were assessed by FBAT software (v 1.4) (25, 26, 48) under an additive and a dominant genetic model for each CX3CR1 SNP. To account for the inclusion of multiple affected

family members, an empirical estimate for the variance was used (48). Hardy-Weinberg equilibrium was tested for each SNP with a chi-square test using one degree of freedom. The calculation of effective number of independent tests is described in online supplement. Strength of linkage disequilibrium (LD) between pairs of SNPs was measured as D prime (D') (49), using Haploview software (v 3.2) (50). Regions of strongly associated markers (LD blocks) were inferred using the definition proposed by Gabriel et al. (51) and implemented in Haploview. Haplotype-specific associations were investigated under a dominant genetic model using the 'haplo' command implemented in FBAT (25, 26) and an empirical variance estimator was used (48). For the Vancouver/Winnipeg case-control sample, the Armitage test for trend (52) was used to test for association between asthma and the rs2669849 and V249I SNPs. Odds ratio and confidence intervals were estimated from 2x2 contingency tables (baseline is homozygous rare). The R statistical software (<http://www.R-project.org>) was used to evaluate these tests.

### **5.1.7 ACKNOWLEDGMENTS**

We thank all families of the Saguenay–Lac-Saint-Jean population for their enthusiastic participation in this study. We also want to thank the following individuals: Paul Bégin, Jean-Pierre Leblanc, Christian Allard, Marie Laforte, Hélène Archibald and Diane Gagné for their invaluable participation in the ascertainment of subjects, Nathalie Pagé from the Laval Hospital for her manuscript revision. K. Tremblay and V. Provost are supported by the Respiratory Health Network (RHN) of the Fonds de la recherche en santé du Québec (FRSQ). K.Tremblay is also supported by a grant from AllerGen (to T.J. Hudson, A. Sandford and C. Laprise). T.J. Hudson is recipient of an Investigator Award from the Canadian Institutes of Health Research (CIHR) and a Clinician-scientist Award in Translational Research from the Burroughs Wellcome Fund. A. Sandford is the chairholder of the Canadian research chair on genetic susceptibility to inflammatory airway diseases. C. Laprise is the chairholder of the Canada Research Chair on genetic determinants in asthma ([www.chairs.gc.ca](http://www.chairs.gc.ca)) and the director of of the Genetics & Genomics Thematic Unit of the RHN of the FRSQ. This work was supported by the Canadian Institutes of Health Research (CIHR) and the Respiratory Health Network (RHN) of the Fonds de la recherche en santé du Québec (FRSQ).

### 5.1.8 REFERENCES

1. Maddox L, Schwartz DA. The pathophysiology of asthma. *Annu Rev Med* 2002; **53**: 477-498.
2. Cookson WO, Moffatt MF. Genetics of asthma and allergic disease. *Hum Mol Genet* 2000; **9**(16): 2359-2364.
3. Howarth PH, Wilson J, Djukanovic R, Wilson S, Britten K, Walls A *et al.* Airway inflammation and atopic asthma: a comparative bronchoscopic investigation. *Int Arch Allergy Appl Immunol* 1991; **94**(1-4): 266-269.
4. Bousquet J, Jeffery PK, Busse WW, Johnson M, Vignola AM. Asthma. From bronchoconstriction to airways inflammation and remodeling. *Am J Respir Crit Care Med* 2000; **161**(5): 1720-1745.
5. Ober C, Hoffjan S. Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 2006; **7**(2): 95-100.
6. Laprise C, Sladek R, Ponton A, Bernier MC, Hudson TJ, Laviolette M. Functional classes of bronchial mucosa genes that are differentially expressed in asthma. *BMC Genomics* 2004; **5**(1): 21-31.
7. Maho A, Bensimon A, Vassart G, Parmentier M. Mapping of the CCXCR1, CX3CR1, CCBP2 and CCR9 genes to the CCR cluster within the 3p21.3 region of the human genome. *Cytogenet Cell Genet* 1999; **87**(3-4): 265-268.
8. Fong AM, Robinson LA, Steeber DA, Tedder TF, Yoshie O, Imai T *et al.* Fractalkine and CX3CR1 mediate a novel mechanism of leukocyte capture, firm adhesion, and activation under physiologic flow. *J Exp Med* 1998; **188**(8): 1413-1419.
9. Imai T, Hieshima K, Haskell C, Baba M, Nagira M, Nishimura M *et al.* Identification and molecular characterization of fractalkine receptor CX3CR1, which mediates both leukocyte migration and adhesion. *Cell* 1997; **91**(4): 521-530.
10. Combadiere C, Salzwedel K, Smith ED, Tiffany HL, Berger EA, Murphy PM. Identification of CX3CR1. A chemotactic receptor for the human CX3C chemokine fractalkine and a fusion coreceptor for HIV-1. *J Biol Chem* 1998; **273**(37): 23799-23804.

11. Foussat A, Coulomb-L'Hermine A, Gosling J, Krzysiek R, Durand-Gasselien I, Schall T *et al.* Fractalkine receptor expression by T lymphocyte subpopulations and in vivo production of fractalkine in human. *Eur J Immunol* 2000; **30**(1): 87-97.
12. Bazan JF, Bacon KB, Hardiman G, Wang W, Soo K, Rossi D *et al.* A new class of membrane-bound chemokine with a CX3C motif. *Nature* 1997; **385**(6617): 640-644.
13. Fraticelli P, Sironi M, Bianchi G, D'Ambrosio D, Albanesi C, Stoppacciaro A *et al.* Fractalkine (CX3CL1) as an amplification circuit of polarized Th1 responses. *J Clin Invest* 2001; **107**(9): 1173-1181.
14. Fujimoto K, Imaizumi T, Yoshida H, Takanashi S, Okumura K, Satoh K. Interferon-gamma stimulates fractalkine expression in human bronchial epithelial cells and regulates mononuclear cell adherence. *Am J Respir Cell Mol Biol* 2001; **25**(2): 233-238.
15. Umehara H, Bloom ET, Okazaki T, Nagano Y, Yoshie O, Imai T. Fractalkine in vascular biology: from basic research to clinical disease. *Arterioscler Thromb Vasc Biol* 2004; **24**(1): 34-40.
16. Echigo T, Hasegawa M, Shimada Y, Takehara K, Sato S. Expression of fractalkine and its receptor, CX3CR1, in atopic dermatitis: possible contribution to skin inflammation. *J Allergy Clin Immunol* 2004; **113**(5): 940-948.
17. Rimaniol AC, Till SJ, Garcia G, Capel F, Godot V, Balabanian K *et al.* The CX3C chemokine fractalkine in allergic asthma and rhinitis. *J Allergy Clin Immunol* 2003; **112**(6): 1139-1146.
18. Sukkar MB, Issa R, Xie S, Oltmanns U, Newton R, Chung KF. Fractalkine/CX3CL1 production by human airway smooth muscle cells: induction by IFN-gamma and TNF-alpha and regulation by TGF-beta and corticosteroids. *Am J Physiol Lung Cell Mol Physiol* 2004; **287**(6): L1230-L1240.
19. Moatti D, Faure S, Fumeron F, Amara Mel W, Seknadji P, McDermott DH *et al.* Polymorphism in the fractalkine receptor CX3CR1 as a genetic risk factor for coronary artery disease. *Blood* 2001; **97**(7): 1925-1928.
20. McDermott DH, Halcox JP, Schenke WH, Waclawiw MA, Merrell MN, Epstein N *et al.* Association between polymorphism in the chemokine receptor CX3CR1 and coronary vascular endothelial dysfunction and atherosclerosis. *Circ Res* 2001; **89**(5): 401-407.



21. McDermott DH, Fong AM, Yang Q, Sechler JM, Cupples LA, Merrell MN *et al.* Chemokine receptor mutant CX3CR1-M280 has impaired adhesive function and correlates with protection from cardiovascular disease in humans. *J Clin Invest* 2003; **111**(8): 1241-1250.
22. Faure S, Meyer L, Costagliola D, Vaneensberghe C, Genin E, Autran B *et al.* Rapid progression to AIDS in HIV+ individuals with a structural variant of the chemokine receptor CX3CR1. *Science* 2000; **287**(5461): 2274-2277.
23. Faure S, Meyer L, Genin E, Pellet P, Debre P, Theodorou I *et al.* Deleterious genetic influence of CX3CR1 genotypes on HIV-1 disease progression. *J Acquir Immune Defic Syndr* 2003; **32**(3): 335-337.
24. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 2004; **74**(4): 765-769.
25. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM. Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet Epidemiol* 2004; **26**(1): 61-69.
26. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. *Genet Epidemiol* 2000; **19**(Suppl 1): S36-S42.
27. Malerba G, Pignatti PF. A review of asthma genetics: gene expression studies and recent candidates. *J Appl Genet* 2005; **46**(1): 93-104.
28. Gugl A, Renner W, Seinost G, Brodmann M, Pabst E, Wascher TC *et al.* Two polymorphisms in the fractalkine receptor CX3CR1 are not associated with peripheral arterial disease. *Atherosclerosis* 2003; **166**(2): 339-343.
29. Hattori H, Ito D, Tanahashi N, Murata M, Saito I, Watanabe K *et al.* T280M and V249I polymorphisms of fractalkine receptor CX3CR1 and ischemic cerebrovascular disease. *Neurosci Lett* 2005; **374**(2): 132-135.
30. McDermott DH, Colla JS, Kleeberger CA, Plankey M, Rosenberg PS, Smith ED *et al.* Genetic polymorphism in CX3CR1 and risk of HIV disease. *Science* 2000; **290**(5499): 2031.
31. Pastinen T, Sladek R, Gurd S, Sammak A, Ge B, Lepage P *et al.* A survey of genetic and epigenetic variation affecting human gene expression. *Physiol Genomics* 2004; **16**(2): 184-193.

32. Pastinen T, Ge B, Gurd S, Gaudin T, Dore C, Lemire M *et al.* Mapping common regulatory variants to human haplotypes. *Hum Mol Genet* 2005; **14**(24): 3963-3971.
33. Mitchell AA, Cutler DJ, Chakravarti A. Undetected genotyping errors cause apparent overtransmission of common alleles in the transmission/disequilibrium test. *Am J Hum Genet* 2003; **72**(3): 598-610.
34. Gordon D, Ott J. Assessment and management of single nucleotide polymorphism genotype errors in genetic association analysis. *Pac Symp Biocomput* 2001; 18-29.
35. Umehara H, Bloom E, Okazaki T, Domae N, Imai T. Fractalkine and vascular injury. *Trends Immunol* 2001; **22**(11): 602-607.
36. Nishimura M, Umehara H, Nakayama T, Yoneda O, Hieshima K, Kakizaki M *et al.* Dual functions of fractalkine/CX3C ligand 1 in trafficking of perforin<sup>+</sup>/granzyme B<sup>+</sup> cytotoxic effector lymphocytes that are defined by CX3CR1 expression. *J Immunol* 2002; **168**(12): 6173-6180.
37. Poon AH, Laprise C, Lemire M, Montpetit A, Sinnett D, Schurr E *et al.* Association of vitamin D receptor genetic variants with susceptibility to asthma and atopy. *Am J Respir Crit Care Med* 2004; **170**(9): 967-973.
38. Laitinen T, Daly MJ, Rioux JD, Kauppi P, Laprise C, Petays T *et al.* A susceptibility locus for asthma-related traits on chromosome 7 revealed by genome-wide scan in a founder population. *Nat Genet* 2001; **28**(1): 87-91.
39. Heyer ET, M. Desjardins, B. Seventeenth-century European origins of hereditary diseases in the Saguenay population (Quebec, Canada). *Hum Biol* 1997; **69**(2): 209-225.
40. Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, Weissenbach J *et al.* Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am J Hum Genet* 1996; **59**(3): 633-643.
41. Sriver CR. Human genetics: lessons from Quebec populations. *Annu Rev Genomics Hum Genet* 2001; **2**: 69-101.
42. Heyer E, Tremblay M. Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 1995; **56**(4): 970-978.

43. Chan-Yeung M, Manfreda J, Dimich-Ward H, Ferguson A, Watson W, Becker A. A randomized controlled study on the effectiveness of a multifaceted intervention program in the primary prevention of asthma in high-risk infants. *Arch Pediatr Adolesc Med* 2000; **154**(7): 657-663.
44. American Thoracic Society. ATS statement--Snowbird workshop on standardization of spirometry. *Am Rev Respir Dis* 1979; **119**: 831-838.
45. American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. *Am Rev Respir Dis* 1987; **136**(1): 225-244.
46. Bell PA, Chaturvedi S, Gelfand CA, Huang CY, Kochersperger M, Kopla R *et al*. SNPstream UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery. *Biotechniques*. 2002; **Suppl:70-2**: 74, 76-77.
47. Chen X, Levine L, Kwok PY. Fluorescence polarization in homogeneous nucleic acid analysis. *Genome Res* 1999; **9**(5): 492-498.
48. Lake SL, Blacker D, Laird NM. Family-based tests of association in the presence of linkage. *Am J Hum Genet* 2000; **67**(6): 1515-1525.
49. Lewontin RC. On measures of gametic disequilibrium. *Genetics* 1988; **120**(3): 849-852.
50. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**(2): 263-265.
51. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B *et al*. The structure of haplotype blocks in the human genome. *Science* 2002; **296**(5576): 2225-2229.
52. Sasiemi PD. From genotypes to genes: doubling the sample size. *Biometrics* 1997; **53**(4): 1253-1261.

### 5.1.9 LEGEND TO FIGURES

**Figure 1:** Pairwise linkage disequilibrium (LD) pattern of CX3CR1 measured by D' and the common haplotypes of the LD regions

A) The location of each tested SNP along the chromosome is indicated on top. The number in each diamond indicates the magnitude of LD in percent between respective pairs of SNPs. For example, the pairwise D' for SNPs rs2853712 and rs2669841 is 0.97. Squares without values represent perfect LD (D' = 1.0). Strength of LD is depicted by

progression of color, for all  $D'$  with  $LOD > 2$ , the color moves from red to light blue as  $D'$  runs from 1 to 0;  $D'$  with  $LOD < 2$  are represented in white. B) Common haplotypes of the four blocks are listed with their frequencies in parentheses. Thick lines joining haplotypes from each block represent combined haplotypes with frequency  $> 0.1$  and thin lines for frequency  $< 0.1$ .

**Figure 2:** Genomic organization of CX3CR1 (shown on reverse complement strand)  
Positions and names of the eleven SNPs analyzed are presented by arrows below the gene structure. Source: November 2005 Genome Browser [NM\_001337] (<http://genome.ucsc.edu/>) and NCBI [GeneID:1524] (<http://www.ncbi.nlm.gov/>) databases.

**5.1.10 Table 1**

Clinical baseline characteristics of the Saguenay–Lac-Saint-Jean sample subjects

	Families		
	Probands (n=223)	Affected relatives (n=347)	Unaffected relatives (n=569)
Male : Female ratio	1 : 1.2	1 : 1.4	1 : 1.2
Mean age in years (range)	18 (3-46)	40 (2-83)	48 (3-96)
Median age in years	16	41	48
Smoking status			
Never	186 (84%)	176 (51%)	243 (43%)
Ex-smoker	12 (5%)	105 (30%)	199 (35%)
Smoker	25 (11%)	66 (19%)	127 (22%)
FEV <sub>1</sub> as % predicted (SD) <sup>1</sup>	92.2 (16.3)	88.9 (22.7)	98.9 (17.0)
PC <sub>20</sub> in mg/ml (SD) <sup>2</sup>	2.66 (3.33)	3.36 (4.53)	26.91 (3.04)
Serum IgE in µg/L (SD) <sup>2</sup>	229.1 (4.6)	157.4 (4.6)	80.9 (3.7)
Number of persons with subphenotypes			
Asthma	223 (100%)	347 (100%) <sup>3</sup>	0 (0%)
Atopy <sup>4</sup>	182 (82%)	237 (68%)	218 (38%)
AHR <sup>5</sup>	169 (90%)	217 (82%)	64 (11%)
IgE >100 µg/L	141 (63%)	185 (53%)	155 (27%)
IgE >280 µg/L	95 (43%)	106 (31%)	70 (12%)

<sup>1</sup> FEV<sub>1</sub> = Forced expiratory volume in one second<sup>2</sup> PC<sub>20</sub> = concentration of methacholine inducing a 20 % fall in FEV<sub>1</sub>; geometric mean<sup>3</sup> Present asthma or past documented clinical history of asthma<sup>4</sup> Defined as at least one positive response (wheal diameter ≥ 3 mm at 10 minutes) on skin prick tests<sup>5</sup> AHR = Airway hyperresponsiveness as defined in method section (evaluated for 188 probands, 264 affected family members and 486 unaffected family members)

**5.1.11 Table 2**

FBAT results for association of CX3CR1 and asthma

SNP	Allele	Allele frequency	S <sup>1</sup>	E(S) <sup>2</sup>	Var(S)	Z <sup>3</sup>	p value
rs1050592	C	0.314	119.000	114.417	39.326	0.731	0.465
	T	0.686	93.000	83.250	11.688	2.852	0.004
T280M	A	0.162	65.000	68.917	30.674	-0.707	0.479
	G	0.838	18.000	14.250	1.188	3.441	0.001
V249I	C	0.704	73.000	63.472	9.751	3.051	0.002
	T	0.296	103.000	100.056	38.147	0.477	0.634
rs2853712	G	0.481	105.000	108.583	27.312	-0.686	NS <sup>4</sup>
	A	0.519	100.000	100.500	25.250	-0.100	NS
rs2669841	C	0.754	54.000	51.500	6.250	1.000	NS
	T	0.246	77.000	76.750	26.562	0.049	NS
rs2853711	C	0.245	97.000	94.833	35.236	0.365	NS
	A	0.755	63.000	61.750	7.812	0.447	NS
rs871610	C	0.770	69.000	69.664	10.238	-0.207	NS
	T	0.230	104.000	101.414	33.628	0.446	NS
rs871144	G	0.261	101.000	97.574	30.674	0.619	NS
	A	0.739	62.000	61.491	8.500	0.175	NS
rs938203	G	0.156	74.000	78.488	26.887	-0.866	0.387
	A	0.844	52.000	45.155	3.994	3.425	0.001
rs3020453	G	0.198	77.000	68.139	27.779	1.681	NS
	A	0.802	42.000	38.222	4.605	1.760	NS
rs2669849	G	0.257	94.000	93.917	36.882	0.014	0.989
	A	0.743	72.000	63.750	7.188	3.077	0.002

<sup>1</sup> S = FBAT statistic<sup>2</sup> E(S) = Expected FBAT statistic<sup>3</sup> Z = Z-score (negative score shows a lower transmission frequency in probands)<sup>4</sup> NS = not significant (p value > 0.05)

Note au lecteur :

La valeur S désigne la transmission des allèles observée chez les probands comparativement à celle attendue (E(S)) en fonction du génotype des parents. La Variance (Var(S)) est l'écart observé entre S et E(S) et transformée en valeur Z, désigne le résultat statistique du test avec la valeur p qui lui est associée.

**5.1.12 Table 3**

Statistical analysis of the two SNPs tested in the case-control Vancouver/Winnipeg sample

SNP	Genotype	Case	Control	OR <sup>1</sup>	Additive <sup>2</sup>		Dominant <sup>2</sup>	
					Z	p value	Z	p value
rs2669849	AA	119	151	2.75				
	AG	55	103	1.86	2.44	0.0073	1.56	0.059
	GG	4	14					
V249I	CC	93	134	2.16				
	CT	73	108	2.11	1.31	0.095	1.86	0.031
	TT	8	25					

<sup>1</sup> OR : Estimated odds ratio

<sup>2</sup> Armitage test for trend of association with the common allele (52) under an additive or dominant model

Note au lecteur :

Parmi les cinq SNPs associés, seulement rs2669849 et V249I ont été demandés pour une réplication étant donné que le génotypage de ces deux SNPs a été obtenu avant les autres. À la lumière des résultats finaux et compte tenu du fait que ces deux SNPs sont localisés dans les deux blocs de déséquilibre de liaison qui incluent la totalité des SNPs associés, les auteurs de l'article n'ont pas jugé essentiel de demander la réplication pour les trois autres SNPs associés. Et ce, sur la base que rs2669849 et V249I sont représentatifs de la région chromosomique associée.

**5.1.13 Table 4**

FBAT results for Block 1 haplotype transmission patterns related to asthma

rs1050592	Block 1		S <sup>1</sup>	E(S) <sup>2</sup>	Z <sup>3</sup>	p value
	T280M	V249I				
T	G	C	45.968	37.191	2.823	0.005
C	A	T	45.969	42.939	0.616	NS <sup>4</sup>
C	G	T	40.021	42.532	-0.479	NS

<sup>1</sup> S = FBAT statistic<sup>2</sup> E(S) = Expected FBAT statistic<sup>3</sup> Z = Z-score<sup>4</sup> NS = not significant (p value > 0.05)



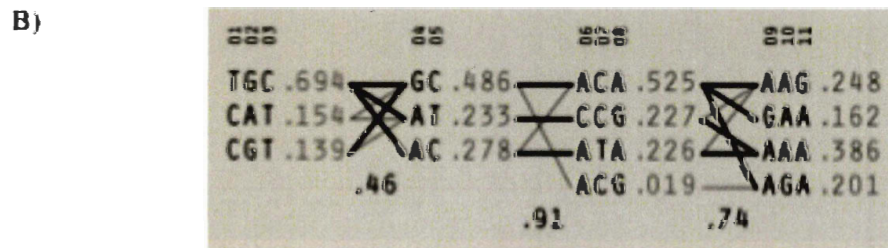
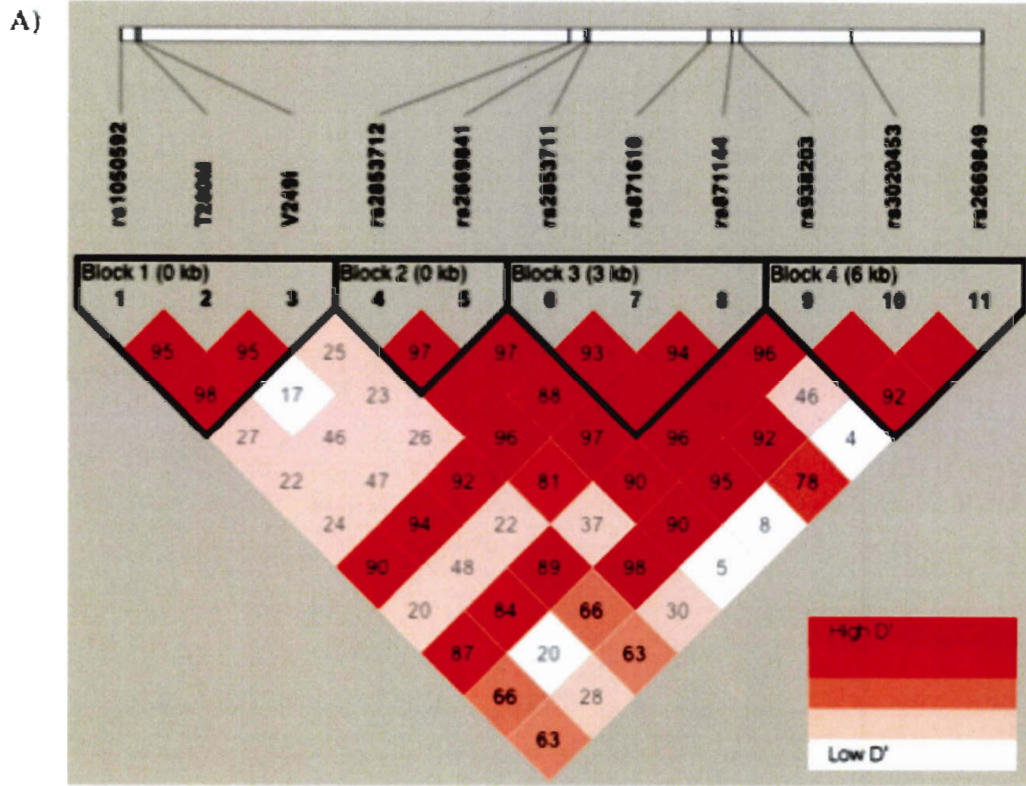
**5.1.14 Table 5**  
CX3CR1 polymorphism characteristics

SNP	Base change	Gene map locus	MAF <sup>1</sup>	Chromosome location <sup>2</sup>	Genomic features
rs1050592	T/C	3p21	0.314	39281787	3'UTR
rs3732378	G/A	3p21	0.162	39282165	Nonsynonymous T280M
rs3732379	C/T	3p21	0.296	39282259	Nonsynonymous V249I
rs2853712	A/G	3p21	0.481	39293291	intron
rs2669841	C/T	3p21	0.246	39293707	intron
rs2853711	A/C	3p21	0.245	39293800	intron
rs871610	C/T	3p21	0.230	39296870	5'UTR
rs871144	A/G	3p21	0.261	39297469	5'UTR
rs938203	A/G	3p21	0.156	39297668	5'UTR
rs3020453	A/G	3p21	0.198	39300526	5'UTR
rs2669849	A/G	3p21	0.257	39303877	5'UTR

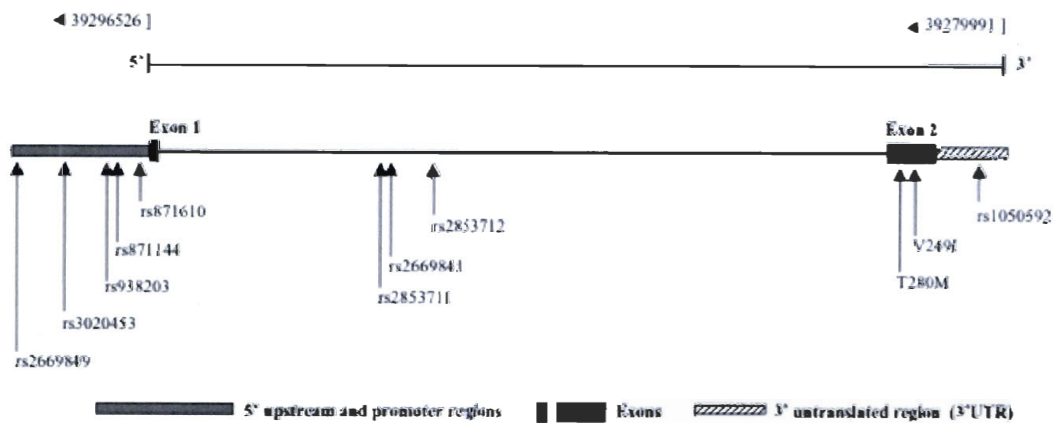
<sup>1</sup> MAF= minor allele frequency

<sup>2</sup> Chromosome location for each SNP was obtained from the November 2005 freeze of the University of California Santa Cruz Genome Browser database using [NM\_001337] (<http://genome.ucsc.edu/>)

5.1.15 Figure 1



5.1.16 Figure 2



## **5.1.18 SUPPLEMENTARY MATERIAL**

### **Association study between the CX3CR1 gene and asthma**

Karine Tremblay, Mathieu Lemire, Véronique Provost, Tomi Pastinen, Yannick Renaud, Andrew J Sandford, Michel Laviolette, Thomas J Hudson and Catherine Laprise

### **SUBJECTS AND METHODS**

#### **Studied population and sample for the association study**

The Saguenay–Lac-Saint-Jean (SLSJ) region of Northeastern Quebec, Canada, is inhabited by 287,000 individuals who descend predominantly from approximately 2500 founders originating from France that settled in Quebec in the 17<sup>th</sup> century (E1). The population grew at a high rate with little admixture over 12 to 14 generations, and is an example of a young founder population (E2). Recruitment occurred through media advertisement (newspaper and radio) and from the Chicoutimi Hospital's specialized clinics (pediatric and respiratory). In order to be included, probands were required to fulfill at least two of the following criteria: 1) a minimum of three clinic visits for acute asthma within one year; 2) two or more asthma-related hospital admissions within one year; or 3) steroid dependency, as defined by either six months of oral, or one year of inhaled corticosteroid use. Families were included in the study if parents (mother and father) were available for phenotypic assessment, if at least one parent was unaffected, and if all four grandparents were of French-Canadian origin. Family participation rate was about 60%. A total of 223 independent families with family size ranging from 3 to 17 and the number of affected family members ranging from 1 to 10 were analyzed.

#### **Clinical evaluation of the SLSJ sample subjects**

The affection status of all study participants was determined by clinical evaluation and by the completion of a standard respiratory questionnaire that was modified to include questions about asthma and atopy severity, family history of asthma and/or atopy, age-of-onset and the presence of other respiratory diseases (E3). In 41 cases, the age-of-onset described by parents was below two years; because of the uncertainty of this information, we used a default class of two years. Measurements of expiratory flow

were done with a Morgan spirometer (Morgan Spiro 232, P.K. Morgan Ltd) according to the American Thoracic Society's recommendations (E4). Bronchodilator response (BD) was measured as the increase in FEV<sub>1</sub> at 15 minutes after a 200 mg dose of salbutamol was inhaled. Peak expiratory flow rates (PEFRs) were measured in the morning and evening over a period of two weeks using a mini-Wright peak-flow meter (Armstrong Medical). The best of three repeated measurements was recorded on a diary card. Skin-prick tests were performed for 26 inhalant allergens. Serum IgE was measured with enzyme immunofluorometry (QuantiCLONE total IgE kit, Kallestad Diagnostics). Methacholine based bronchoprovocation tests were performed according to the method described by Juniper and colleagues (E5). The dose of methacholine that resulted in a 20% fall in FEV<sub>1</sub> was recorded as the PC<sub>20</sub>. The Chicoutimi Hospital (SLSJ) local ethics committee approved the study and all subjects gave their informed consent.

### **Genotyping**

For the UHT Orchid platform genotyping, SNP flanking sequences were tested for the presence of repeats or duplicated regions using the BLAT program (<http://www.genome.ucsc.edu>). PCR and single-base extension (SBE) primers were designed using the Autoprimer program (<http://www.autoprimer.com>). The program selects PCR primers that will generate products ranging between 90 and 180 basepairs (bp) and an optimized single bp extension primer 5' to the SNP site (E6). See Tables E1 and E2 for oligonucleotides used for genotyping in this study. The program also assembles SNPs into groups of twelve by SNP extension type (e.g. T/C, A/T, etc.) and appends a unique DNA tag to the 5' ends of the extension primers. Twelve-plex PCR reactions were performed in 384-well plates (MJS BioLynx) in a 5  $\mu$ L volume using 6 ng of DNA, 75  $\mu$ M dNTPs, 0.5 Unit of AmpliTaq Gold (Perkin-Elmer), and the 24 PCR primers at a concentration of 50 nM each in 1 X PCR buffer. Thermal cycling was performed in GeneAmp PCR system 9700 thermal cyclers (Applied Biosystems) using the following program: initial denaturation at 95°C for 5 min followed by 40 cycles of 95°C for 30 sec, 50-55°C for 55 sec, 72°C for 30 sec. After the last cycle, the reaction was held at 72°C for 7 min. Following PCR, plates were centrifuged briefly and 3  $\mu$ L of a mixture containing 0.67 Unit Exonuclease I (Amersham Pharmacia) and 0.33 Unit Shrimp Alkaline Phosphatase (Amersham Pharmacia) were added to each well. The

plates were sealed and incubated for 30 min at 37°C and at 95°C for 10 min. Extension reactions and hybridizations to the Orchid UHT microarray plates were carried out as described in reference (E6), using the SNPware UHT reagent kit and the appropriate extension mix kit containing two dideoxynucleotides labeled with either Bodipy-Fluorescein or TAMRA dye (Beckman Coulter). Finally, the plates were read using the SNPstream Array Imager (Beckman Coulter) and fluorescence intensities were measured with the UHTImage software (Beckman Coulter). Intensities were plotted and genotypes were called by the UHTGetGenos software (Beckman Coulter). After visual inspection of the clusters, manual adjustments were made for some of the assays.

For the HEFP platform, PCR primers were designed using the Primer3 software ([http://www-genome.wi.mit.edu/cgi-bin/primer/primer3\\_www.cgi](http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi)). PCR reactions were carried out as follows: 6.0 ng of genomic DNA were added to an 8.0 µL reaction mixture containing 2.5 mM of MgCl<sub>2</sub>, 25mM of dNTPs, 0.2 Unit of HotstartTaq DNA polymerase (Qiagen) and 100 nM of each primer. PCR was initiated by denaturing the samples at 94°C for 15 min followed by 45 cycles of denaturation at 94°C for 30 sec, annealing at a temperature between 56°C-57°C (primers specific) for 30 sec and extension at 72°C for 30 sec. Final extension was done at 72°C for 6 min. PCR products were treated with Exonuclease I and Shrimp Alkaline Phosphatase as recommended by the manufacturer (AcycloPrime-FP SNP Detection Kit, Perkin Elmer). SBE detection primers of minimum 16 bp length and melting temperature above 55°C were designed in one or both orientations for each SNP using Primer3. See Tables E1 and E2 for oligonucleotides used in the HEFP platform. FP-SBE reactions were performed in one or both orientations as suggested by the manufacturer (AcycloPrime-FP SNP Detection Kit, Perkin Elmer). If necessary, we optimised the FP-SBE reaction by varying the annealing temperature and MgCl<sub>2</sub> concentration. After addition of reading buffer, we read the plates using an Analyst HT reader (Molecular Devices) as described previously (E7, E8).

### **Vancouver/Winnipeg sample**

This sample was drawn from the participants of a study designed to assess the effectiveness of a multifaceted intervention program in the primary prevention of asthma and other atopic disorders (E9). The individuals were the parents of children at high risk of developing asthma and were recruited from two centers (Vancouver and

Winnipeg, Canada) and asthma was diagnosed by questionnaire (E9). The CX3CR1 rs2669849 and V249I SNPs were genotyped utilizing standard restriction fragment length polymorphism assays. The region surrounding rs2669849 was amplified with the following primers : 5'-CCA AGT GTA TTG ACC GAG TGA-3' and 5'-TGG CTG TAG AAG GTA TTT ATT TTG C-3'. The PCR contained 1X PCR buffer (Qiagen), 0.2 mM dNTPs, 0.5  $\mu$ M of each primer, 0.5 units HotStar Taq DNA polymerase (Qiagen) and 100 ng of genomic DNA. PCR was carried out in a PCR Express Thermal Cycler (Hybaid Ltd) with the following cycles : 95°C 15 min, 35 cycles of 94°C 30 sec, 55°C 30 sec, 72°C 30 sec and a final cycle of 72°C 5 min. The PCR product was digested with 10 units of BsiE1 (New England Biolabs) for 5 hours at 60°C. Following electrophoresis on a 2.5% agarose gel the A allele appeared as a 152 bp band and the G allele as 74 and 78 bp bands. The region surrounding V249I was amplified with the following primers : 5'-GGT GGT CAT CGT GTT TTT CC-3' and 5'-AGG CAT TTC CCA TAC AGG TG-3'. The PCR conditions and cycles were as described for the rs2669849 polymorphism. The PCR product was digested with 10 units of AcII (New England Biolabs) for 5 hours at 37°C. Following electrophoresis on a 2.5% agarose the T allele appeared as a 231 bp band and the G allele as 192 and 39 bp bands.

#### **Statistical analysis : Effective number of independent tests**

Since the 11 SNPs are correlated, as are the four phenotypes studied, accounting for the number of tests using a simple Bonferroni correction is expected to be conservative. An effective number of independent SNPs can be calculated by the method proposed by Li and Ji (E10) implemented in SNPSdP (E11) : this number takes into account the observed  $r^2$  measure of pairwise linkage disequilibrium between all SNPs. Out of the four phenotypes investigated, an effective number of independent phenotypes was calculated using 10000 simulated replicates of the SLSJ sample of families: genotypes for 10 SNPs with minor allele frequencies varying from 5% to 50% were simulated under the null hypothesis of no linkage and no association, for all members of the sample who have DNA. FastSLINK (E12-E14) and SUP (E15) were used to create the replicates. Each replicates of genotypes was then analysed for all four phenotypes studied, as they have been observed in the sample, and the minimum p value observed over the four phenotypes was stored for each of the SNPs. The effective number of independent phenotypes was then taken to be the ratio between 0.05 and the quantile of order 5% of these minimum p values.

## REFERENCES

- E1. Heyer E, Tremblay M. Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 1995; **56**(4): 970-978.
- E2. Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, Weissenbach J *et al.* Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am J Hum Genet* 1996; **59**(3): 633-643.
- E3. American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. *Am Rev Respir Dis* 1987; **136**(1): 225-244.
- E4. American Thoracic Society. ATS statement--Snowbird workshop on standardization of spirometry. *Am Rev Respir Dis* 1979; **119**: 831-838.
- E5. American Thoracic Society. Standardization of Spirometry, 1994 Update. *Am J Respir Crit Care Med* 1995; **152**: 1107-1136.
- E6. Bell PA, Chaturvedi S, Gelfand CA, Huang CY, Kochersperger M, Kopla R *et al.* SNPstream UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery. *Biotechniques* 2002; **Jun;Suppl:70-2, 74**: 76-77.
- E7. Hsu TM, Law SM, Duan S, Neri BP, Kwok PY. Genotyping single-nucleotide polymorphisms by the invader assay with dual-color fluorescence polarization detection. *Clin Chem* 2001; **47**(8): 1373-1377.
- E8. Hsu TM, Chen X, Duan S, Miller RD, Kwok PY. Universal SNP genotyping assay with fluorescence polarization detection. *Biotechniques* 2001; **31**(3): 560, 562, 564-8, *passim*.
- E9. Chan-Yeung M, Manfreda J, Dimich-Ward H, Ferguson A, Watson W, Becker A. A randomized controlled study on the effectiveness of a multifaceted intervention program in the primary prevention of asthma in high-risk infants. *Arch Pediatr Adolesc Med* 2000; **154**(7): 657-663.
- E10. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 2005; **95**(3): 221-227.
- E11. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 2004; **74**(4): 765-769.



- E12. Ott J. Computer-simulation methods in human linkage analysis. *Proc Natl Acad Sci U S A* 1989; **86**(11): 4175-4178.
- E13. Weeks DE, Lathrop GM, Ott J. Multipoint mapping under genetic interference. *Hum Hered* 1993; **43**(2): 86-97.
- E14. Cottingham RW, Jr., Idury RM, Schaffer AA. Faster sequential genetic linkage computations. *Am J Hum Genet* 1993; **53**(1): 252-263.
- E15. Lemire M. SUP: An extension to SLINK to allow a larger number of marker loci to be simulated in pedigrees conditional on trait values. *BMC Genet* 2006; **7**(1): 40.

**Table E1**  
Oligonucleotides used in genotyping assays

SNP	Technology	Type <sup>1</sup>	Sequence	Cycling <sup>2</sup>
rs1050592	SNPstream (Orchid)	P	GACCTGGGTGTCGATACCTAAACAACCCTA	a
		F	GAGTGTGTTGAGAA	
		R	GATGGACCCAATGCACAC	
rs2853711	SNPstream (Orchid)	P	CATTCAAAGAGTTCAATTTGTTCA	a
		F	CGACTGTAGGTGCGTAACTCGGGAAAGTTG	
		R	GGGCTTCTAGCTGAA	
rs871144	SNPstream (Orchid)	P	ACTCTGCCCAGAAGAGCC	a
		F	AGCCAGCAGATTCTGAGC	
		R	GACCTGGGTGTCGATACCTAGCCAGAGGGG	
rs938203	SNPstream (Orchid)	P	CCTTGAAGGGGTCAC	a
		F	TGGGATGTGGGCAGCAAG	
		R	TATATAATTCTGATCCATCCCTAACTAG	
rs2669841	HEFP	P	GTGATTCTGTACGTGTCGCCGTGCTGCCGGC	a
		F	TGGTGGCCTGCTC	
		R	AGAGCTGGGGGCTGCAGT	
rs2669849	HEFP	P	CCAAGAACAAGAAAAGAAAAGT	b
		F	CTGCTTACAGGTGGGATTCT	
		R	CCAAACAAGATGAGGGGGT	
rs2853712	HEFP	P	TGGCCACAGCTGATGCTTC	b
		F	TTTAAGGGGAAAACAAAGGTGGT	
		R	CATCACCAAGTGATTGACCGAG	
rs3020453	HEFP	P	GCAGGTGGTAGGTCCTCAAG	b
		F	CCCTCCTGGCCAAACGGTC	
		R	TTGCTGTTACTGTCGTGCAGAT	
T280M	HEFP	P	CTCTCCCAAAAAGAACCAAGT	b
		F	CCAAATATGAAGATCAAAAATAGG	
		R	TCCCTAAACAATATAGCTTAGCTTT	
V249I	HEFP	P	GCCAAACCAAAAAGTATGAATCC	b
		F	ACTCCCCAACAAGTCCTTTTAGC	
		R	CCCATTTCCTCTGATGTGATTAT	
rs871610	HEFP	P	TTGATACAGGCATGTAACGTGT	b
		F	ATACCCCATTTCTCTGATGTGA	
		R	TTCTTTCCAGTTGTGACATGAG	
rs871610	HEFP	P	CAACAATGGCTAAATGCAACC	b
		F	GGCCCTCAGTGTGACTGAGA	
		R	TTTCCATACAGGTGGTAAAGGT	
rs871610	HEFP	P	GTTTTCTCTTCTGGACACCCT	b
		F	AGCGTCTCCAGGAAAATCATAA	
		R	CCTCTTCTGGACACCCTACAAC	
rs871610	HEFP	P	ATTCAGGCAACAATGGCTAAATG	b
		F	CTCTCATTAATGGGGAGCCTT	
		R	AGTACTCTGGGCTCCCTCTC	
rs871610	HEFP	P	GCCCTCAGCTTCCCTGCCCA	b
		F	TTTGAGGAGAAGAGCTGTGTGTG	
		R	TTTGAGGAGAAGAGCTGTGTGTG	

<sup>1</sup> P = Probe; F = Forward; R = Reverse; Ps = Probe Sense; Pas = Probe AntiSense

<sup>2</sup> See Table E2 for cycling conditions

**Table E2**  
Genotyping cycling conditions

<b>Cycling</b>	<b>Procedure</b>	<b>Step</b>	<b>Description</b>	
a	PCR	Step1	95°C for 5:00 min.	
		Step2	95°C for 0:30 min.	
		Step3	50°C for 0:55 min.	
		Step4	72°C for 0:30 min.	
		Step5	Go to step 2 for 5 times	
		Step6	95°C for 0:30 min.	
		Step7	50°C for 0:55 min. increase by 0.2°C with each cycle	
		Step8	72°C for 0:30 min.	
		Step9	Go to step 6 for 25 times	
		Step10	95°C for 0:30 min.	
		Step11	55°C for 0:55 min.	
		Step12	72°C for 0:30 min.	
		Step13	Go to step 9 for 15 times	
		Step14	72°C for 7:00 min.	
		Step15	4°C	
	Exo/SAP	Step1	37°C for 30:00 min.	
		Step2	100°C for 10:00 min.	
		Step3	4°C	
	SBE	Step1	96°C for 3:00 min.	
		Step2	94°C for 0:20 min.	
		Step3	40°C for 0:11 min.	
		Step4	Go to step 2 for 45 times	
		Step5	4°C	
b	PCR	Step1	95°C for 10:00 min.	
		Step2	95°C for 0:30 min.	
		Step3	60°C for 0:30 min.	
		Step4	72°C for 0:30 min.	
		Step5	Go to step 2 for 40 times	
		Step6	72°C for 7:00 min.	
		Exo/SAP	Step1	37°C for 45:00 min.
			Step2	95°C for 25:00 min.
		SBE	Step1	95°C for 2:00 min.
			Step2	95°C for 0:10 min.
			Step3	54°C for 0:30 min.
			Step4	Go to step 2 for 30 times

## 5.2 Deuxième article

### **Variations génétiques dans les gènes des molécules de signalisation immune différemment exprimés dans les biopsies bronchiques de sujets asthmatiques \***

Karine Tremblay, Denise Daley, Annie Chamberland, Mathieu Lemire, Alexandre Montpetit, Michel Laviolette, Arthur W Musk, Alan L James, Moira Chan-Yeung, Allan Becker, Anita L Kozyrskyj, Andrew J Sandford, Thomas J Hudson, Peter D Paré et Catherine Laprise

*Journal of Allergy and Clinical Immunology* (2008) Sep;122(3):529-36.e17

\* Sélectionné dans la rubrique *The Editors' Choice* de l'issue de septembre 2008

#### **Résumé**

Huit gènes des molécules de signalisation immune différemment exprimés dans les biopsies bronchiques de sujets asthmatiques ont été sélectionnés comme gènes candidats pour l'asthme. L'objectif de la présente étude est d'effectuer une étude d'association entre ces huit gènes et l'asthme et trois phénotypes d'asthme (atopie, hyperréactivité bronchique (HRB) et asthme allergique) dans trois échantillons de trios familiaux canadiens et un échantillon cas-témoins australien (AllerGen Inc.). Des TagSNPs (pour *tagging single nucleotide polymorphisms*) ont été sélectionnés à l'aide de la base de données publique HapMap ( $r^2 > 0,8$  et fréquence de l'allèle mineur  $> 0,10$ ) et génotypés sur une plate-forme Illumina. Les tests d'association ont été effectués pour les quatre phénotypes dans chaque échantillon et des corrections pour tests multiples ont été appliquées. Avant les corrections statistiques, six gènes ont été associés à au moins un phénotype dans un ou plusieurs échantillons étudiés ( $0,001 < P < 0,05$ ). Après les corrections, l'association entre *ALOX15* et l'HRB et l'asthme allergique demeure significative dans deux des échantillons canadiens ( $P$  corrigée = 0,022 et 0,049, respectivement), et l'association entre *CD14* et l'asthme demeure significative dans un échantillon canadien ( $P$  corrigée = 0,042). Dans tous les cas, un effet protecteur des allèles mineurs a été observé. Ces résultats démontrent l'efficacité des puces à ADN à cibler des gènes candidats pour l'asthme puisqu'elles ont permis de trouver une nouvelle association pour *ALOX15* dans deux échantillons indépendants. Le potentiel anti-inflammatoire de ce gène dans la pathophysiologie relative à l'asthme n'est pas encore connu et reste à être déterminé par des études fonctionnelles.

## 5.2.1 TITLE PAGE

### **Genetic variation in immune signaling genes differentially expressed in asthmatic lung tissues**

Karine Tremblay, MSc<sup>1,2</sup>, Denise Daley, PhD<sup>3</sup>, Annie Chamberland, MSc<sup>2</sup>, Mathieu Lemire, PhD<sup>4</sup>, Alexandre Montpetit, PhD<sup>5</sup>, Michel Laviolette, MD<sup>1</sup>, Arthur (Bill) W Musk, MD<sup>6</sup>, Alan L James, MD<sup>7</sup>, Moira Chan-Yeung, PhD<sup>8</sup>, Allan Becker, MD<sup>9</sup>, Anita L Kozyrskyj, BScPhm, PhD<sup>9,10</sup>, Andrew J Sandford, PhD<sup>3</sup>, Thomas J Hudson, MD<sup>4</sup>, Peter D Paré, MD<sup>3</sup> and Catherine Laprise, PhD<sup>2</sup>

<sup>1</sup> Hôpital Laval, Université Laval, Québec, QC, G1V 4G5, Canada; <sup>2</sup> Département des sciences fondamentales, Université du Québec à Chicoutimi, Saguenay, QC, G7H 2B1, Canada; <sup>3</sup> James Hogg iCAPTURE Centre for Cardiovascular and Pulmonary Research, University of British Columbia, Vancouver, BC, V6Z 1Y6 Canada; <sup>4</sup> Ontario Institute for Cancer Research, Toronto, ON, M5G 0A3, Canada; <sup>5</sup> McGill University and Genome Quebec Innovation Centre, Montreal, QC, H3A 1A4, Canada; <sup>6</sup> Department of Respiratory Medicine, Sir Charles Gairdner Hospital, Nedlands, WA, 6009 Australia; <sup>7</sup> Western Australian Sleep Disorders Research Institute, Sir Charles Gairdner Hospital, Nedlands, WA, 6009, Australia; <sup>8</sup> Occupational and Environmental Lung Disease Unit, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada; <sup>9</sup> Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, MB, R3A 1S1, Canada; <sup>10</sup> Faculty of Pharmacy; Department of Community Health Sciences, Manitoba Centre for Health Policy, Faculty of Medicine, University of Manitoba, Winnipeg, MB, R3E 3P5, Canada; <sup>11</sup> University of Montreal Community Genomic Medicine Centre, Chicoutimi University Hospital, Saguenay, QC, G7H 5H6, Canada.

#### **Corresponding Author**

Catherine Laprise, Ph.D.  
Université du Québec à Chicoutimi  
555, boulevard de l'Université,  
Saguenay, QC, G7H 2B1, Canada  
Telephone: (418) 545-5011, 5659  
Fax: (418) 615-1203  
Email: catherine\_laprise@uqac.ca

Funded by a grant from the AllerGen NCE Inc. (<http://www.allergen-nce.ca/>), the Canadian Institutes of Health Research, the Respiratory Health Network of the Fonds de la recherche en santé du Québec and the Western Australia Healthway.

### 5.2.2 ABSTRACT

**Background:** Eight genes in immune signaling pathway, shown to be differentially expressed in asthmatic lung biopsies in a previous microarray experiment, were selected as candidate genes for asthma susceptibility. **Objective:** To perform an association study with these genes and asthma-related phenotypes in three independent Canadian familial asthma collections and one Australian asthma case-control study. **Methods:** Tagging single nucleotide polymorphisms were selected using the HapMap public database ( $r^2 > 0.8$ ; minor allele frequency  $> 0.10$ ) and genotyped using the Illumina platform. Family-based association and trend tests for asthma, atopy, airway hyperresponsiveness (AHR) and allergic asthma phenotypes were done in each sample correcting for multiple testing. **Results:** Uncorrected associations with polymorphisms within seven genes were detected with one or more of the phenotypes in one or more of the four populations ( $0.001 < P < 0.05$ ). After correction, the 15-lipoxygenase associations with AHR and allergic asthma remained significant in two Canadian samples (corrected  $P = 0.022$  and  $0.049$ , respectively), and the association of the CD14 antigen with asthma remained significant in one Canadian sample (corrected  $P = 0.042$ ). In both cases, a protective effect of the minor alleles was observed. **Conclusion:** Expression profiling studies are a useful way to identify candidate genes for asthma as this approach has led to the first report of an association with 15-lipoxygenase in two independent populations. Since 15-lipoxygenase is involved in anti-inflammatory processes, further functional and clinical investigation of the role of this biological pathway in asthma is warranted.

### 5.2.3 RELEVANCE

#### Key messages

- This study demonstrates the efficiency of the microarrays in the identification of asthma candidate genes.
- The findings suggest a novel asthma protective role for the 15-lipoxygenase gene.

- Since 15-lipoxygenase is involved in anti-inflammatory processes, functional and clinical investigation of its role in asthma is warranted.

### **Key words**

Asthma, microarrays expression profiling, candidate genes, genetic association study, tagging single nucleotide polymorphism, 15-lipoxygenase

### **Capsule summary**

This study demonstrates the usefulness of microarray gene profiling on bronchial biopsies as a method to identify asthma candidate genes and allowed the identification of a novel genetic association between the 15-lipoxygenase gene and asthma related phenotypes.

### **5.2.4 ABBREVIATIONS**

15(S)-HETE = 15(S)-hydroxyeicosatetraenoic acid

15-LO = 15-lipoxygenase

AHR = Airway hyperresponsiveness

ALOX15 = Arachidonate 15-lipoxygenase gene

CAPPS = Canadian Asthma Primary Prevention Study

CD14 = CD14 molecule gene

CD27 = CD27 molecule gene

CXCL12 = Chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1) gene

FBAT = Family-Based Association Test

IL2RB = Interleukin 2 receptor beta gene

IL7R = Interleukin 7 receptor gene

LD = Linkage Disequilibrium

mRNA = Messenger RiboNucleic Acid

NCBI = National Center for Biotechnology Information

NOS2A = Nitric oxide synthase 2A gene

SAGE = Study of Asthma Genetics and Environment Cohort

SLSJ = Saguenay–Lac-Saint-Jean

SNP = Single Nucleotide Polymorphism

SNPSpD = Single Nucleotide Polymorphism Spectral Decomposition

SFRP1 = Secreted frizzled-related protein 1 gene

TagSNP = Tagging Single Nucleotide Polymorphism

### 5.2.5 INTRODUCTION

Asthma is a chronic inflammatory disease of the airways associated with variable airflow obstruction and airway hyperresponsiveness.<sup>1</sup> Asthma is also recognized as a complex heritable disease, implying the interaction of numerous genetic and environmental factors.<sup>2</sup> Identification of the genes that influence asthma susceptibility and severity has been the subject of intense research over the last decade.<sup>3, 4</sup> This knowledge has significantly contributed to our understanding of the pathophysiology of asthma but much work is needed to replicate reported associations, to characterize the functional effects of the genes already identified<sup>5</sup>, and to identify new candidate genes for asthma.

Expression profiling studies in which transcript levels in tissues of affected and control subjects are compared allow the identification of genes related to a disease process.<sup>6</sup> These genes, which may or may not have known function, can be selected as candidate genes in subsequent genetic association studies on the basis of their differential expression in affected tissues.<sup>5</sup> We used this approach in the selection of the *CX3CR1* gene, the fractalkine receptor, for a previous asthma association study<sup>7</sup> since *CX3CR1* was underexpressed in bronchial biopsies of asthmatic subjects compared with controls.<sup>8</sup>

*CX3CR1* is related to immune signaling and eight other genes which we reported as differentially expressed in asthmatic bronchial biopsies are also part of this biological cluster.<sup>8</sup> Three of these have been previously suggested as asthma genes (*CD14*, *CXCL12* and *NOS2A*)<sup>4</sup> but five have not been recognized as asthma genes or clearly implicated in asthma pathophysiology (*ALOX15*, *CD27*, *IL2RB*, *IL7R* and *SFRP1*) (see Table 1 for a description of the genes). Since the approach used for *CX3CR1* allowed us to find a positive association which has been replicated in an independent cohort<sup>7</sup>, we have applied a similar strategy to these eight additional differentially expressed genes. We postulated that these genes represent asthma biomarkers, and if genetic variation explains the difference in gene expression, then we should observe an association



between variants of these candidate genes and asthma and related phenotypes in a collection of asthmatic subjects. Specifically, we hypothesized that, based on their differential expression in asthmatic bronchial biopsies<sup>8</sup> and their biological role in immunity, *ALOX15*, *CD14*, *CD27*, *CXCL12*, *IL2RB*, *IL7R*, *NOS2A* and *SFRP1* represent candidate genes for asthma susceptibility.

The objectives of the present study are to report tests for association between asthma-related phenotypes and the eight candidate genes in four independent asthmatic collections (different by their recruitment, ascertainment of the subjects as well as by their genetic background) included in the Canadian AllerGen network (AllerGen NCE Inc.). Three of these are Canadian family studies and the fourth one is an Australian asthma case-control study. Our results demonstrated the usefulness of the expression profiling studies in the identification of asthma candidate genes since they allow us to report the first asthma genetic association for *ALOX15* in two independent Canadian population samples.

### **5.2.6 METHODS**

#### *Association study samples*

In the context of the AllerGen (Allergy, Genes and Environment) NCE Inc. Canadian network, four well-characterized samples of individuals affected with asthma and their family members as well as case-control data were combined as a resource to investigate the genetics of asthma and related phenotypes (see the complete description in the journal's online repository). Briefly, these include three Canadian independent studies: 1) the Saguenay–Lac-Saint-Jean (SLSJ) founder population family collection<sup>9-14</sup>; 2) the Canadian Asthma Primary Prevention Study (CAPPS) high-risk birth cohort<sup>15-17</sup>; and 3) the Study of Asthma Genes and the Environment Cohort (SAGE), a nested case-control sample from a population-based birth cohort. These three samples have been fully described in a recent paper by Bégin and coworkers.<sup>18</sup> The fourth asthma sample is the Australian Busselton Health Study, which is a nested case-control sample selected from the population based cohort which has a longitudinal design and a standardized phenotype classification.<sup>19-23</sup> Informed consents were approved by each center's Research Ethics Board and collected for each study participant.

### *Phenotype definitions*

Four dichotomous asthma related-phenotypes were analyzed: (1) Asthma, defined as doctor-diagnosed in SLSJ, CAPPs and SAGE, and as a positive response to the question “has a doctor ever told you that you had asthma, or bronchial asthma?” at either survey (1981 and 1994) in the Busselton cohort; (2) Atopy, defined as at least one positive response (wheal diameter  $\geq$  3 mm at 10 minutes) on skin prick testing for any allergen tested; (3) Airway hyperresponsiveness (AHR), assessed by methacholine-challenge testing. A positive response was defined as a 20% decrease in forced expiratory volume in one second ( $PC_{20}$  or  $PD_{20}$ ). For the SLSJ, CAPPs and SAGE samples, AHR was defined as  $PC_{20} < 8$  mg/ml and for Busselton,  $PD_{20} \leq 3.9$  micromoles; and (4) Allergic asthma, defined as individuals diagnosed with both asthma and atopy.

### *Genotyping*

DNA samples were available for each studied sample. AllerGen tagSNP genotyping was performed using the Illumina BeadLab (<http://www.illumina.com>, Illumina Inc., San Diego, CA) platform, as described previously.<sup>24</sup> TagSNPs were selected from the International HapMap Project (<http://www.hapmap.org>) release 16c, from among SNPs located in the vicinity of the candidate genes (if they were not located in neighboring genes) and included the flanking regions 10 kb upstream and 10 kb downstream of the first and the last exon, respectively. TagSNPs were selected in a manner similar to LD Select<sup>25</sup> using an  $r^2$  cutoff of 0.8. If equivalent tagSNPs were identified, those with the highest Illumina design score were kept. Only SNPs with a design score above 0.4 were retained for the panel. Samples with call rates below 95% were excluded from the final dataset. Assays with poor cluster separation and with call rates below 90% were removed from the final dataset. All primers and conditions used are available on request to the corresponding author. For the eight studied candidate genes, a panel of 77 tagSNPs was tested for association with asthma, atopy, AHR and allergic asthma phenotypes in each sample. SNPs described in this report are cited using their NCBI reference sequence number (rs#).

### *Statistical analyses*

Allele distribution patterns and Mendelian errors for the three familial samples (SLSJ, CAPPs and SAGE) were assessed by FBAT software (version 1.7.2)<sup>26-28</sup> under an additive genetic model. An empirical estimate of the variance has been used for the

SLSJ sample.<sup>26</sup> FBAT results are reported for the minor alleles. For the Busselton sample, association testing was performed using a general allelic likelihood ratio test, Chi-square (one degree of freedom) as implemented in UNPHASED.<sup>29</sup> This approach uses a standard retrospective case-control likelihood for case-control designs. Odds ratios were estimated using the most common allele as the referent and are reported for each minor allele. For each gene and for each SNPs within that gene, we applied a correction to the significance level that accounts for both the effective number of independent SNPs tested in that gene and the effective number of phenotypes tested. A Sidak correction was applied with respect to that number (independent SNPs X independent phenotypes). Since the tagSNPs are correlated, the effective number of independent SNPs was calculated by the method proposed by Li and Ji<sup>30</sup> implemented in the Single Nucleotide Polymorphism Spectral Decomposition (SNPSpD) web resource (<http://gump.qimr.edu.au/general/daledN/SNPSpD/>).<sup>31</sup> We used a similar procedure, the Matrix Spectral Decomposition approach, to estimate the number of independent phenotypes ( $n = 3$ ) in each sample. Hardy-Weinberg equilibrium p-values was calculated using the Haploview software (version 3.32)<sup>32</sup> for the SLSJ, CAPPs and SAGE samples and using a Chi-square test for the Busselton cohort, considering one degree of freedom and a significance threshold of 0.001. The strength of linkage disequilibrium (LD) between pairs of SNPs was evaluated using the Haploview software (version 3.32).<sup>32</sup> LD plots for each gene in each population are presented in the journal's online repository.

## 5.2.7 RESULTS

For the SLSJ sample, the probands had a male:female ratio of 1:1.2, a mean age of 18 years and a mean PC<sub>20</sub> of 2.66 mg/ml. For the CAPPs cohort, the children were 7 years old, the sex ratio was 1.05:1 and they had a mean PC<sub>20</sub> of 2.99 mg/ml. For the SAGE cohort, the children were aged between 8 and 10 years, they had a sex ratio of 1.26:1 and a mean PC<sub>20</sub> of 2.72 mg/ml. Finally, for the Busselton cohort, the age range was from 17 to 91 years, the sex ratio was 1:1.35 and they had a mean PD<sub>20</sub> of 1.14  $\mu$ /mol. The number of individuals genotyped and the phenotypic distribution in each studied sample are presented in Table 2. Genotyping was achieved with a mean success rate of 99.9 % for the SLSJ, 99.6 % for the CAPPs, 99.7 % for the SAGE and 99.7 % for the Busselton samples. A total of 17 Mendelian errors were observed. Families that showed

Mendelian errors were reset to zero for FBAT analyses. All SNPs were in Hardy-Weinberg equilibrium and none of them showed more than two Mendelian errors.

The family-based association analyses for the SLSJ, CAPPs and SAGE samples are presented in the Tables 3 to 5, respectively. Odds ratios for the Busselton sample are presented in Table 6. For the sake of brevity, only results showing a p-value less than the 0.05 threshold for one or more phenotypes are presented with their respective corrected p-values indicated in parentheses. The significance values for each population for all tagSNPs that generated significant associations are available in online Tables E1 to E4. After correction for multiple testing, only three associations remained significant. In the CAPPs sample, *ALOX15* rs7217186 showed a positive association for the AHR phenotype, suggesting a protective effect of the minor allele since it is undertransmitted to the hyperresponsive probands (corrP = 0.022). In the SAGE sample, *ALOX15* showed a positive association for rs2664593 and allergic asthma. The undertransmission of the minor allele to the allergic asthmatic probands suggests a protective effect (corrP = 0.049). In this sample, a protective association was also found for the *CD14* gene (corrP = 0.042). For all other tagSNPs tested, no association remained significant after correcting for multiple testing, regardless of the studied phenotype in a given sample.

### 5.2.8 DISCUSSION

This study reports association analyses of eight candidate genes, selected based on our previous microarray results<sup>7</sup>, in four independent asthma samples. Before correction for multiple testing, this strategy identified seven associations with asthma or related phenotypes in one or more of the samples studied. The genes which showed positive associations included *CD14*, *CXCL12* and *NOSA*, which are well known in asthma genetics<sup>4</sup> and novel ones, such as *ALOX15*, *CD27*, *IL2RB* and *SFRP1*. However we applied a correction of our p-values in order to reduce the possibility of false positive associations. Thus, without minimizing the potential impact of the five genes that did not survive correction, only the *ALOX15* and the *CD14* associations remained significant. Taken together, these results demonstrate the usefulness as well as the efficiency of the microarray approach to identify genes relevant to asthma.

The *ALOX15* association remained significant in two of the four studied samples and is the first report of its association with asthma phenotypes, thus accentuating our interest in this gene and its biological pathway. In the CAPPS sample, the association found was for rs7217186 and AHR and in the SAGE sample, it was for the rs2664593 and allergic asthma. In both cases, a protective effect of the minor allele was suggested. The *ALOX15* gene, located on chromosome 17p13.3, encodes arachidonate 15-lipoxygenase (15-LO). This enzyme is expressed in several lung cell types, including human airway epithelial cells<sup>33</sup>, alveolar macrophages<sup>34, 35</sup> as well as neutrophils<sup>36</sup>, and converts arachidonic acid into 15(S)-hydroxyeicosatetraenoic acid (15(S)-HETE), a major metabolite of the 15-LO pathway. Alterations in the 15-LO pathway have been suggested as potential contributors to the pathogenesis of asthma<sup>34, 37</sup>, but a protective anti-inflammatory role has also been proposed. Firstly, 15-LO could act via the uptake of soluble 15(S)-HETE into the cellular membrane phospholipids, contributing to a decrease in the pericellular concentration of this proinflammatory mediator.<sup>33, 34</sup> Secondly, 15-LO could have a protective effect via lipoxin A4 biosynthesis (from 15(S)-HETE) which exerts diverse inhibitory mechanisms to protect against AHR and asthmatic inflammation.<sup>35, 38</sup> Finally, 15-LO could act via the inhibition of the chemotactic mediator leukotriene B4, which modulates neutrophilic trafficking and infiltration.<sup>36, 39</sup> Thus, the protective association we report for *ALOX15* rs2664593 and rs7217186 with respect to allergic asthma and AHR could be explained by these anti-inflammatory roles of the 15-LO pathway. Obviously, this hypothesis remains speculative as further functional investigations of the *ALOX15* gene in inflammation are needed to delineate its role in the pathophysiology of asthma.

The variability in association patterns among the four studies may reflect the differences in their respective design<sup>18</sup> as well as genetic and environmental heterogeneity. Indeed, the recruitment of the subjects differed between the studies: the SLSJ probands were recruited at a mean age of 18; the CAPPS subjects were recruited prenatally based on a family history of asthma and/or allergy; the SAGE subjects were ascertained as part of a population-based sample; and the Busselton cohort is a study of adult asthmatics. Since the natural history of asthma shows considerable variability implying complex pathological mechanisms<sup>40, 41</sup>, the variable recruitment strategies could have resulted in subjects who are at different disease stages or actual phenotypic sub-sets (e.g. childhood

transient versus persistent asthma). Since different genes may predispose to these different phenotypes, this could be a cause of heterogeneous results. Another potential source of different association patterns is variable linkage disequilibrium (LD) patterns for the studied SNPs across the populations.<sup>42</sup> The SLSJ sample is recognized as a founder population<sup>11, 12, 43-46</sup> while for the CAPPs, SAGE and Busselton studies, individuals come from admixed Canadian or Australian populations. These populations present variable *ALOX15* LD patterns, such that a polymorphism might be in LD with a nearby disease allele in the CAPPs or SAGE samples but not in the SLSJ or Busselton studies. A last possible factor that could explain, in part, the different associations found for *ALOX15* is population specific gene-gene and gene-environment interactions.<sup>42</sup> This possibility is also likely to explain the unique *CD14* association found in the SAGE sample. In fact, literature review reports *CD14* as a classical example of gene-environment interactions for asthma and atopy phenotypes, and to date, numerous positive and negative associations have been reported across diverse populations.<sup>47</sup>

In conclusion, we believe that hypothesis-driven candidate gene selection from microarray experiments is a methodological approach that improves the likelihood of finding positive as well as novel asthma associations within a limited number of tests performed. We also think that studying numerous independent asthma samples has the advantage of being able to document the biological relevance of candidate genes across diverse populations living in different environments. Based on the results presented here, we believe that *ALOX15* represents an interesting potential candidate for further studies.

We are thankful to the participants in each asthma study presented in this report. K. Tremblay is an AllerGen PhD trainee and is supported by the Fondation de l'Université Laval studentship. D. Daley is supported by grants from the Canadian Institutes of Health Research (CIHR), Institutes of Gender and Health, Genetics, Population and Public Health, and a CIHR STIHR IMPACT fellowship. D. Daley is also supported by the Lung Association of British Columbia. A. Chamberland is supported by the Fondation de l'Université du Québec à Chicoutimi studentship. A.L. Kozyrskyj is the recipient of a CIHR New Investigator award. A. Sandford is the chairholder of the Canada Research Chair ([www.chairs.gc.ca](http://www.chairs.gc.ca)) on genetic susceptibility to inflammatory

disease. P.D. Paré is a Michael Smith Foundation and Jacob Churg Scholar. T.J. Hudson is recipient of an Investigator Award from the CIHR and a Clinician-scientist Award in Translational Research from the Burroughs Wellcome Fund. C. Laprise is the chairholder of the Canada Research Chair on genetic determinants in asthma and the director of the Genetics platform of the Respiratory Health Network (RHN) of the Fonds de la recherche en santé du Québec (FRSQ).

### 5.2.9 REFERENCES

1. Bousquet J, Jeffery PK, Busse WW, Johnson M, Vignola AM. Asthma. From bronchoconstriction to airways inflammation and remodeling. *Am J Respir Crit Care Med* 2000; 161:1720-1745.
2. Sandford A, Weir T, Pare P. The genetics of asthma. *Am J Respir Crit Care Med* 1996; 153:1749-1765.
3. Hoffjan S, Nicolae D, Ober C. Association studies for asthma and atopic diseases: a comprehensive review of the literature. *Respir Res* 2003; 4:14.
4. Ober C, Hoffjan S. Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 2006; 7:95-100.
5. Holloway JW, Koppelman GH. Identifying novel genes contributing to asthma pathogenesis. *Curr Opin Allergy Clin Immunol* 2007; 7:69-74.
6. Schmidt-Weber CB. Gene expression profiling in allergy and asthma. *Chem Immunol Allergy* 2006; 91:188-194.
7. Tremblay K, Lemire M, Provost V, Pastinen P, Renaud Y, Sandford AJ, Laviolette M, Hudson TJ, Laprise C. Association study between the CX3CR1 gene and asthma. *Genes Immun* 2006; 7:632-639.
8. Laprise C, Sladek R, Ponton A, Bernier MC, Hudson TJ, Laviolette M. Functional classes of bronchial mucosa genes that are differentially expressed in asthma. *BMC Genomics* 2004; 5:21-31.
9. Laitinen T, Daly MJ, Rioux JD, Kauppi P, Laprise C, Petays T, Green T, Cargill M, Haahtela T, Lander ES, Laitinen LA, Hudson TJ, Kere J. A susceptibility locus for asthma-related traits on chromosome 7 revealed by genome-wide scan in a founder population. *Nat Genet* 2001; 28:87-91.

10. Heyer E, Tremblay M, Desjardins B. Seventeenth-century European origins of hereditary diseases in the Saguenay population (Quebec, Canada). *Hum Biol* 1997; 69:209-225.
11. Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, Weissenbach J, Popowska E, Pronicka E, Root AW, Glorieux FH. Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am J Hum Genet* 1996; 59:633-643.
12. Scriver CR. Human genetics: lessons from Quebec populations. *Annu Rev Genomics Hum Genet* 2001; 2:69-101.
13. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, November 1986. *Am Rev Respir Dis* 1987; 136:225-244.
14. Lemire M, Roslin NM, Laprise C, Hudson TJ, Morgan K. Transmission-ratio distortion and allele sharing in affected sib pairs: a new linkage statistic with reduced bias, with application to chromosome 6q25.3. *Am J Hum Genet* 2004; 75:571-586.
15. Chan-Yeung M, Manfreda J, Dimich-Ward H, Ferguson A, Watson W, Becker A. A randomized controlled study on the effectiveness of a multifaceted intervention program in the primary prevention of asthma in high-risk infants. *Arch Pediatr Adolesc Med* 2000; 154:657-663.
16. Chan-Yeung M, Ferguson A, Watson W, Dimich-Ward H, Rousseau R, Lilley M, Dybuncio A, Becker A. The Canadian Childhood Asthma Primary Prevention Study: outcomes at 7 years of age. *J Allergy Clin Immunol* 2005; 116:49-55.
17. Becker A, Watson W, Ferguson A, Dimich-Ward H, Chan-Yeung M. The Canadian asthma primary prevention study: outcomes at 2 years of age. *J Allergy Clin Immunol* 2004; 113:650-656.
18. Bégin P, Tremblay K, Daley D, Lemire M, Claveau S, Salesse C, Kacel S, Montpetit A, Becker A, Kozyrskyj A, Hudson T, Laprise C. Association of Urokinase-type Plasminogen Activator with Asthma and Atopy. *A J Resp Crit Care Med* 2007; 175:1109-1116.
19. James AL, Palmer LJ, Kicic E, Maxwell PS, Lagan SE, Ryan GF, Musk AW. Decline in lung function in the Busselton Health Study: the effects of asthma and cigarette smoking. *Am J Respir Crit Care Med* 2005; 171:109-114.



20. Palmer LJ, Knuiiman MW, Divitini ML, Burton PR, James AL, Bartholomew HC, Ryan G, Musk AW. Familial aggregation and heritability of adult lung function: results from the Busselton Health Study. *Eur Respir J* 2001; 17:696-702.
21. Moffatt MF, Schou C, Faux JA, Abecasis GR, James A, Musk AW, Cookson WO. Association between quantitative traits underlying asthma and the HLA-DRB1 locus in a family-based population sample. *Eur J Hum Genet* 2001; 9:341-346.
22. Hill MR, James AL, Faux JA, Ryan G, Hopkin JM, le Souef P, Musk AW, Cookson WO. Fc epsilon RI-beta polymorphism and risk of atopy in a general population sample. *BMJ* 1995; 311:776-779.
23. Daniels SE, Bhattacharrya S, James A, Leaves NI, Young A, Hill MR, Faux JA, Ryan GF, le Souef PN, Lathrop GM, Musk AW, Cookson WO. A genome-wide search for quantitative trait loci underlying asthma. *Nature* 1996; 383:247-250.
24. Lincoln MR, Montpetit A, Cader MZ, Saarela J, Dyment DA, Tiislar M, Ferretti V, Tienari PJ, Sadovnick AD, Peltonen L, Ebers GC, Hudson TJ. A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. *Nat Genet* 2005; 37:1108-1112.
25. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 2004; 74:106-120.
26. Lake SL, Blacker D, Laird NM. Family-based tests of association in the presence of linkage. *Am J Hum Genet* 2000; 67:1515-1525.
27. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM. Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet Epidemiol* 2004; 26:61-69.
28. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. *Genet Epidemiol* 2000; 19 (Suppl 1):S36-S42.
29. Dudbridge F. Pedigree disequilibrium tests for multilocus haplotypes. *Genet Epidemiol* 2003; 25:115-121.
30. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 2005; 95:221-227.
31. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 2004; 74:765-769.

32. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; 21:263-265.
33. Profita M, Vignola AM, Sala A, Mirabella A, Siena L, Pace E, Folco G, Bonsignore G. Interleukin-4 enhances 15-lipoxygenase activity and incorporation of 15(S)-HETE into cellular phospholipids in cultured pulmonary epithelial cells. *Am J Respir Cell Mol Biol* 1999; 20:61-68.
34. Profita M, Sala A, Riccobono L, Paterno A, Mirabella A, Bonanno A, Guerrera D, Pace E, Bonsignore G, Bousquet J, Vignola AM. 15-Lipoxygenase expression and 15(S)-hydroxyeicosatetraenoic acid release and reincorporation in induced sputum of asthmatic subjects. *J Allergy Clin Immunol* 2000; 105:711-716.
35. Levy BD, Romano M, Chapman HA, Reilly JJ, Drazen J, Serhan CN. Human alveolar macrophages have 15-lipoxygenase and generate 15(S)-hydroxy-5,8,11-cis-13-trans-eicosatetraenoic acid and lipoxins. *J Clin Invest* 1993; 92:1572-1579.
36. Profita M, Sala A, Riccobono L, Pace E, Paterno A, Zarini S, Siena L, Mirabella A, Bonsignore G, Vignola AM. 15(S)-HETE modulates LTB(4) production and neutrophil chemotaxis in chronic bronchitis. *Am J Physiol Cell Physiol* 2000; 279:C1249-C1258.
37. Chu HW, Balzar S, Westcott JY, Trudeau JB, Sun Y, Conrad DJ, Wenzel SE. Expression and activation of 15-lipoxygenase pathway in severe asthma: relationship to eosinophilic phenotype and collagen deposition. *Clin Exp Allergy* 2002; 32:1558-1565.
38. Levy BD, De Sanctis GT, Devchand PR, Kim E, Ackerman K, Schmidt BA, Szczeklik W, Drazen JM, Serhan CN. Multi-pronged inhibition of airway hyper-responsiveness and inflammation by lipoxin A(4). *Nat Med* 2002; 8:1018-1023.
39. Takata S, Matsubara M, Allen PG, Janmey PA, Serhan CN, Brady HR. Remodeling of neutrophil phospholipids with 15(S)-hydroxyeicosatetraenoic acid inhibits leukotriene B4-induced neutrophil migration across endothelium. *J Clin Invest* 1994; 93:499-508.
40. Reed CE. The natural history of asthma. *J Allergy Clin Immunol* 2006; 118:543-548; quiz 549-550.
41. Koh MS, Irving LB. The natural history of asthma from childhood to adulthood. *Int J Clin Pract* 2007; 61:1371-1374.
42. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002; 4:45-61.

43. Heyer E, Tremblay M. Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 1995; 56:970-978.
44. Engert JC, Berube P, Mercier J, Dore C, Lepage P, Ge B, Bouchard JP, Mathieu J, Melancon SB, Schalling M, Lander ES, Morgan K, Hudson TJ, Richter A. ARSACS, a spastic ataxia common in northeastern Quebec, is caused by mutations in a new gene encoding an 11.5-kb ORF. *Nat Genet* 2000; 24:120-125.
45. Richter A, Rioux JD, Bouchard JP, Mercier J, Mathieu J, Ge B, Poirier J, Julien D, Gyapay G, Weissenbach J, Hudson TJ, Melancon SB, Morgan K. Location score and haplotype analyses of the locus for autosomal recessive spastic ataxia of Charlevoix-Saguenay, in chromosome region 13q11. *Am J Hum Genet* 1999; 64:768-775.
46. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994; 265:2037-2048.
47. Simpson A, John SL, Jury F, Niven R, Woodcock A, Ollier WE, Custovic A. Endotoxin exposure, CD14, and allergic disease: an interaction between genes and the environment. *Am J Respir Crit Care Med* 2006; 174:386-392.

### 5.2.10 Table 1

Gene and TagSNP characteristics

Gene location *	Microarray †	Description ‡	Function ‡	TagSNP	Genomic Feature §
<i>ALOX15</i> 17p13.3 GeneID:246	4.19	Arachidonate 15-lipoxygenase	Converts arachidonic acid to 15S-hydroperoxyeicosatetraenoic acid. Also acts on C-12 of arachidonate as well as on linoleic acid.	rs1076039	Downstream 3'UTR
				rs4790210	Downstream 3'UTR
				rs916055	3'UTR
				rs743646	Exon 11 (T485T)
				rs7217186	Intron 7
				rs2664593	Upstream 5'UTR
<i>CD14</i> 5q31.1 GeneID:929	-2.76	CD14 molecule	Mediates the innate immune response to bacterial lipopolysaccharides. Leads the NF-kappa-B activation, cytokine secretion and the inflammatory response. Up-regulates cell surface adhesion molecules.	rs748694	Upstream 5'UTR
				rs778583	Downstream 3'UTR
				rs778584	Downstream 3'UTR
				rs4914	Exon 2 (L367L)
				rs2569190	Upstream 5'UTR
				rs5744455	Upstream 5'UTR
<i>CD27</i> 12p13 GeneID:939	-3.61	CD27 molecule	Member of the TNF-receptor superfamily required for generation and long-term maintenance of T cell immunity. Regulates B-cell activation and immunoglobulin synthesis. This receptor transduces signals that lead to the activation of NF-kappaB and MAPK8/JNK.	rs2569193	Upstream 5'UTR
				rs2364493	Upstream 5'UTR
				rs7970260	Upstream 5'UTR
				rs2250246	Upstream 5'UTR
				rs3136551	5'UTR
				rs25680	Exon 2 (T59A)
<i>CXCL12</i> 10q11.1 GeneID:6387	-2.39	Chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)	Small cytokines that belongs to the intercrine family, members of which activate leukocytes and are often induced by proinflammatory stimuli such as lipopolysaccharide, TNF or IL-1.	rs2267966	Intron 2
				rs266105	Downstream 3'UTR
				rs185545	Downstream 3'UTR
				rs2839696	3'UTR
				rs197452	Intron 3
				rs266087	Intron 3
				rs2297630	Intron 3
				rs3780891	Intron 1
				rs1413519	Upstream 5'UTR
				rs2861442	Upstream 5'UTR
rs6593412	Upstream 5'UTR				

<b><i>IL2RB</i></b> 22q13.1 GeneID:3560	-3.00	Interleukin 2 receptor, beta	Receptor for interleukin 2. This beta subunit is involved in receptor mediated endocytosis and transduces the mitogenic signals of IL-2.	rs6000570	Upstream 3'UTR
				rs228934	Upstream 3'UTR
				rs228935	Upstream 3'UTR
				rs228937	Upstream 3'UTR
				rs3218339	Upstream 3'UTR
				rs228942	Exon 10 (E391D)
				rs84459	Intron 9
				rs228947	Intron 9
				rs3218322	Intron 9
				rs3218316	Intron 8
				rs3218315	Intron 8
				rs3218312	Intron 8
				rs228954	Intron 6
				rs3218292	Intron 6
				rs2281094	Intron 4
				rs1003694	Intron 3
				rs2235330	Intron 2
rs3218258	Intron 1				
rs228979	Intron 1				
rs2016771	Upstream 5'UTR				
rs743776	Upstream 5'UTR				
<b><i>IL7R</i></b> 5p13 GeneID:3575	-2.41	Interleukin 7 receptor	Receptor for interleukin 7. Requires the interleukin 2 receptor, gamma chain, which is a common chain shared by the receptors of various cytokines, including interleukin 2, 4, 7, 9, and 15.	rs6890853	Upstream 5'UTR
				rs1353251	Intron 1
				rs1389832	Intron 1
				rs987106	Intron 6
				rs3194051	Exon 8 (V356I)
				rs9292618	Downstream 3'UTR
<b><i>NOS2A</i></b> 17q11.2-q12 GeneID:4843	7.25	Nitric oxide synthase 2A	Produces nitric oxide (NO) which is a messenger molecule with diverse functions throughout the body. In macrophages, NO mediates tumoricidal and bactericidal actions.	rs4796017	Downstream 3'UTR
				rs8081248	Downstream 3'UTR
				rs2255929	Intron 23
				rs2297518	Exon 16 (L608S)
				rs1137933	Exon 10 (D385D)
				rs3794764	Intron 5

				rs8072199	Intron 2
				rs3794766	Intron 2
				rs2779248	5'UTR
				rs2779252	Upstream 5'UTR
				rs2531860	Upstream 5'UTR
				rs3242	3'UTR
				rs1127379	3'UTR
				rs7013368	Intron 2
				rs11774662	Intron 2
				rs10098898	Intron 2
				rs10958672	Intron 2
				rs747417	Intron 2
				rs968427	Intron 1
				rs921142	Upstream 5'UTR
				rs4736965	Upstream 5'UTR
<b><i>SFRP1</i></b>					
8p12-p11.1	-2.16	Secreted frizzled-	SFRPs function as modulators of Wnt signaling through		
GeneID:6422		related protein 1	direct interaction with Wnts. They have a role in regulating		
			cell growth and differentiation in specific cell types.		
			SFRP1 decreases intracellular beta-catenin levels.		

Abbreviations used: TagSNP = Tagging single nucleotide polymorphism; UTR = Untranslated region

\* Gene location obtained from NCBI Gene (<http://www.ncbi.nlm.nih.gov>) (Freeze on August 25 2007).

† Data with p value < 0.000001. From U95A microarray study performed by Laprise *et al.* (2004)<sup>8</sup> using RNA from bronchial biopsies of control subjects (without asthma or atopy) (n = 4), allergic asthmatic subjects (n = 4); Mean fold change in logarithmic values.

‡ From public databases: NCBI (Gene, OMIM, PubMed); Genecards (<http://bioinformatics.weizmann.ac.il/cards/>); Swiss-Prot (<http://ca.expasy.org/sprot/sproc t-top.html>).

§ Genomic features as described in HapMap and NCBI SNP.

**5.2.11 Table 2**

Distribution of asthma, atopy, airway hyperresponsiveness (AHR) and allergic asthma phenotypes among the four studied samples

Phenotypes	SLSJ (n = 1186) *		CAPPS (n = 1253)		SAGE (n = 1247)		Busselton (n = 1524)	
	Probands (n = 209) †	%	Children (n = 509) †	%	Children (n = 684) †	%	Cases ‡	%
	Asthma	209/209	100	71/375	18.9	228/683	33.4	653/1422
Atopy	168/208	80.8	162/362	44.8	298/680	43.8	655/1264	51.8
AHR	158/178	88.8	254/343	74.1	342/624	54.8	222/1083	20.5
Allergic asthma	168/208	80.8	52/374	13.9	146/683	21.4	388/795	48.8

\* Number of available genotypes for the study on each entire sample (probands + family members for SLSJ, CAPPS and SAGE, and cases + controls for Busselton).

† Number of probands or children presenting the phenotype on the total number of available phenotype.

‡ Number of subjects presenting the phenotype on the total number of available phenotype.  
Individuals in each control groups may include cases from the other phenotype groups.

### 5.2.12 Table 3

Significant FBAT results for tagSNPs of the studied candidate genes and asthma, atopy, airway hyperresponsiveness and allergic asthma phenotypes under an additive genetic model in the SLSJ familial sample

Gene	TagSNP	Base change	MAF	Asthma			Atopy			AHR			Allergic asthma		
				N	Z	P (corrP) *	N	Z	P (corrP)	N	Z	P (corrP)	N	Z	P (corrP)
<i>IL2RB</i>	rs2016771	G > T	0.34	128	2.26	0.02 (0.99)	-	-	-	-	-	-	-	-	-
	rs228979	C > A	0.27	- †	-	-	92	-2.10	0.04 (1.00)	-	-	-	94	-2.29	0.02 (0.92)
<i>NOS2A</i>	rs2779248	A > G	0.45	-	-	-	-	-	-	-	-	-	111	-2.34	0.02 (0.47)
	rs3794766	G > A	0.26	-	-	-	-	-	-	-	-	-	109	-2.13	0.03 (0.80)

Abbreviations used: TagSNP = Tagging single nucleotide polymorphism, MAF = Minor allele frequency, AHR = Airway hyperresponsiveness, N = Number of families contributing the statistic, Z = Z score, P = p-value, corrP = Corrected p-value.

\* Number of independent SNPs is estimated to be 14 for *IL2RB* and 8 for *NOS2A*. Number of independent phenotypes is estimated to be 3.

† The dash means that the test is not significant (p-value > 0.05).



### 5.2.13 Table 4

Significant FBAT results for tagSNPs of the studied candidate genes and asthma, atopy, airway hyperresponsiveness and allergic asthma phenotypes under an additive genetic model in the CAPPS familial sample

Gene	TagSNP	Base change	MAF	Asthma			Atopy			AHR			Allergic asthma		
				N	Z	P (corrP) *	N	Z	P (corrP)	N	Z	P (corrP)	N	Z	P (corrP)
<i>ALOX15</i>	rs4790210	T > C	0.21	29	-2.61	0.009 (0.162)	61	-2.85	0.004 (0.079)	101	-2.06	0.04 (0.71)	22	-2.75	0.006 (0.109)
	rs7217186	A > G	0.47	43	-2.21	0.03 (0.48)	-	-	-	154	-3.24	0.001 <b>(0.022)</b>	-	-	-
	rs916055	T > C	0.34	-†	-	-	-	-	-	129	-2.59	0.001 (0.172)	-	-	-
	rs743646	T > C	0.10	-	-	-	-	-	-	66	2.15	0.03 (0.57)	-	-	-
	rs2664593	G > C	0.21	-	-	-	-	-	-	107	1.98	0.05 (0.87)	-	-	-
<i>CD14</i>	rs4914	G > C	0.11	-	-	-	-	-	-	65	2.02	0.04 (0.52)	-	-	-
<i>IL2RB</i>	rs228979	C > A	0.29	-	-	-	-	-	-	-	-	-	28	-2.27	0.02 (1.00)
<i>NOS2A</i>	rs2297518	G > A	0.19	30	2.14	0.03 (0.78)	-	-	-	-	-	-	21	2.20	0.03 (0.67)
	rs2531860	C > T	0.15	-	-	-	-	-	-	-	-	-	18	-2.29	0.02 (0.52)

P-values highlighted in bold and underlined indicate that they remain significant after correction for multiple testing.

Abbreviations used: TagSNP = Tagging single nucleotide polymorphism, MAF = Minor allele frequency, AHR = Airway hyperresponsiveness, N = Number of families contributing the statistic, Z = Z score, P = p-value, corrP = Corrected p-value.

\* Number of independent SNPs is estimated to be 6 for *ALOX15*, 4 for *CD14*, 21 for *IL2RB* and 8 for *NOS2A*. Number of independent phenotypes is estimated to be 3.

† The dash means that the test is not significant (p-value > 0.05).

**5.2.14 Table 5**

Significant FBAT results for tagSNPs of the studied candidate genes and asthma, atopy, airway hyperresponsiveness and allergic asthma phenotypes under an additive genetic model in the SAGE familial sample

Gene	TagSNP	Base change	MAF	Asthma			Atopy			AHR			Allergic asthma		
				N	Z	P (corrP)*	N	Z	P (corrP)	N	Z	P (corrP)	N	Z	P (corrP)
<i>ALOX15</i>	rs4790210	T > C	0.20	63	2.19	0.03 (0.51)	-	-	-	74	2.16	0.03 (0.56)	43	2.38	0.02 (0.31)
	rs2664593	G > C	0.23	71	-2.09	0.04 (0.67)	109	-2.23	0.03 (0.46)	84	-2.05	0.04 (0.73)	47	-2.99	0.003 <b>(0.049)</b>
<i>CD14</i>	rs4914	G > C	0.11	36	-2.92	0.004 <b>(0.042)</b>	-	-	-	-	-	-	-	-	-
<i>CXCL12</i>	rs2861442	A > G	0.26	- †	-	-	104	2.03	0.04 (1.00)	-	-	-	-	-	-
<i>IL2RB</i>	rs6000570	G > A	0.47	-	-	-	139	1.96	0.05 (1.00)	-	-	-	-	-	-
<i>NOS2A</i>	rs4796017	T > C	0.46	82	2.35	0.02 (0.45)	-	-	-	-	-	-	53	2.33	0.02 (0.48)
	rs2779252	G > T	0.06	-	-	-	-	-	-	27	-1.98	0.05 (1.00)	-	-	-
<i>SFRP1</i>	rs11774662	T > C	0.32	90	-2.03	0.04 (0.77)	-	-	-	-	-	-	-	-	-
	rs968427	T > C	0.48	-	-	-	-	-	-	108	-2.06	0.04 (0.71)	-	-	-
	rs921142	T > C	0.42	-	-	-	-	-	-	106	-2.23	0.03 (0.46)	-	-	-
	rs1127379	T > C	0.44	-	-	-	127	-2.11	0.03 (0.63)	-	-	-	-	-	-
	rs7013368	C > T	0.34	-	-	-	119	-2.21	0.03 (0.49)	-	-	-	-	-	-

P-values highlighted in bold and underlined indicate that they remain significant after correction for multiple testing.

Abbreviations used: TagSNP = Tagging single nucleotide polymorphism, MAF = Minor allele frequency, AHR = Airway hyperresponsiveness, N = Number of families contributing the statistic, Z = Z score, P = p-value, corrP = Corrected p-value.

\* Number of independent SNPs is estimated to be 6 for *ALOX15*, 4 for *CD14*, 8 for *CXCL12*, 21 for *IL2RB*, 8 for *NOS2A* and 6 for *SFRP1*. Number of independent phenotypes is estimated to be 3.

† The dash means that the test is not significant (p-value > 0.05).

### 5.2.15 Table 6

Significant odds ratio for tagSNPs of the studied candidate genes and asthma, atopy, airway hyperresponsiveness and allergic asthma phenotypes in the Busselton case-control cohort

Gene	TagSNP	Base change	MAF	Asthma		Atopy		AHR		Allergic asthma	
				OR	P (corrP)	OR	P (corrP)	OR	P (corrP)	OR	P (corrP)
<i>ALOX15</i>	rs7217186	A > G	0.46	-	-	-	-	-	-	0.82	0.05 (0.73)
<i>CD14</i>	rs778583	C > T	0.28	-	-	-	-	0.72	0.008 (0.071)	-	-
<i>CD27</i>	rs3136551	A > G	0.04	0.59	0.008 (0.113)	-	-	-	-	0.45	0.004 (0.056)
	rs7970260	A > C	0.38	-	-	-	-	1.30	0.02 (0.27)	-	-
<i>IL2RB</i>	rs228954	A > G	0.44	-	-	-	-	1.32	0.01 (0.46)	-	-
	rs2281094	C > T	0.27	-	-	-	-	1.31	0.03 (1.00)	-	-
	rs3218312	A > G	0.23	-	-	-	-	0.76	0.04 (1.00)	-	-
<i>NOS2A</i>	rs2531860	C > T	0.15	-	-	-	-	-	-	1.33	0.05 (0.95)

Abbreviations used: TagSNP = Tagging single nucleotide polymorphism, MAF = Minor allele frequency, AHR = Airway hyperresponsiveness, OR = 95% confidence interval odd ratio, P = p-value, corrP = Corrected p-value.

\* Number of independent SNPs is estimated to be 5 for *ALOX15*, 3 for *CD14*, 5 for *CD27*, 14 for *IL2RB* and 7 for *NOS2A*. Number of independent phenotypes is estimated to be 3.

† The dash means that the test is not significant (p-value > 0.05).

## 5.2.16 ONLINE REPOSITORY

### **Genetic variation in immune signaling genes differentially expressed in asthmatic lung tissues**

Karine Tremblay, Denise Daley, Annie Chamberland, Mathieu Lemire, Alexandre Montpetit, Michel Laviolette, Arthur (Bill) W Musk, Alan L James, Moira Chan-Yeung, Allan Becker, Anita L Kozyrskyj, Andrew J Sandford, Thomas J Hudson, Peter D Paré and Catherine Laprise

#### **METHODS**

##### **The Saguenay–Lac-Saint-Jean (SLSJ) familial sample**

The Saguenay–Lac-Saint-Jean (SLSJ) region of Northeastern Quebec, Canada, is inhabited by 287,000 individuals who descend predominantly from approximately 2500 founders originating from France that settled in Quebec in the 17<sup>th</sup> century.<sup>E1</sup> The population grew at a high rate with little admixture over 12 to 14 generations, and is an example of a young founder population.<sup>E2</sup> Recruitment occurred through media advertisement (newspaper and radio) and from the Chicoutimi Hospital's specialized clinics (pediatric and respiratory). In order to be included, probands were required to fulfill at least two of the following criteria: 1) a minimum of three clinic visits for acute asthma within one year; 2) two or more asthma-related hospital admissions within one year; or 3) steroid dependency, as defined by either six months of oral, or one year of inhaled corticosteroid use. Families were included in the study if parents were available for phenotypic assessment, if at least one parent was unaffected, and if all four grandparents were of French-Canadian origin. Family participation rate was about 60%. A total of 223 independent families with family size ranging from 3 to 17 and the number of affected family members ranging from 1 to 10 were analyzed.

The affection status of the SLSJ subjects was determined by clinical evaluation and by the completion of a standard respiratory questionnaire that was modified to include questions about asthma and atopy severity, family history of asthma and/or atopy, age-of-onset and the presence of other respiratory diseases.<sup>E3</sup> In 41 cases, the age-of-onset described by parents was below two years; because of the uncertainty of this

information, we used a default class of two years. Measurements of expiratory flow were done with a Morgan spirometer (Morgan Spiro 232, P.K. Morgan Ltd) according to the American Thoracic Society's recommendations.<sup>E4</sup> Bronchodilator response was measured as the increase in FEV<sub>1</sub> at 15 minutes after a 200 mg dose of salbutamol was inhaled. Peak expiratory flow rates (PEFRs) were measured in the morning and evening over a period of two weeks using a mini-Wright peak-flow meter (Armstrong Medical). The best of three repeated measurements was recorded on a diary card. Skin-prick tests were performed for 24 common airborne allergens, which were divided into six main categories: 1) cat, dog, horse, cow and feathers; 2) dust; 3) *D.farinae* and *D.Pterionisus*; 4) grass, weed, Ambrosia, Tymotheny, ryegrass; 5) tree mix, birch, maple, oak, elm; 6) *Chladosporium*, *Hormodendrum*, *Alternaria Alternata*, *Alternaria Tenuis*, *Aspergillus* and *Penicillium*. Serum IgE was measured with enzyme immunoassay (QuantiCLONE total IgE kit, Kallestad Diagnostics). Methacholine based bronchoprovocation tests were performed for participants 12 years and older according to the method described by Juniper et al.<sup>E5</sup> The dose of methacholine that resulted in a 20% fall in FEV<sub>1</sub> was recorded as the PC<sub>20</sub>. All the tests were performed by the investigators for every participant (except for bronchoprovocation) at the University of Montreal Community Genomic Medicine Centre in Chicoutimi, Canada. Medical charts were also reviewed for any previous function tests and medication. Airway hyperresponsiveness (AHR) was defined as a PC<sub>20</sub> < 8 mg/ml. If PC<sub>20</sub> was unavailable, AHR was defined as either a rise in the post-bronchodilator FEV<sub>1</sub> of > 15% from baseline or a variation of the morning-evening PEFr of > 12% over a two-week period. Participants were defined as asthmatic if 1) they had a reported history of asthma (validated with physician), or if 2) they showed asthma related symptoms and a positive PC<sub>20</sub>. Subjects were deemed atopic if they had at least one positive response (wheal diameter ≥ 3 mm at 10 minutes) on skin prick testing. The Chicoutimi University Hospital ethics committee approved the study and subjects gave informed consent.

### **The Canadian Asthma Primary Prevention Study cohort (CAPPS)**

The Canadian Asthma Primary Prevention Study was initiated in 1995 to assess the effectiveness of a multifaceted intervention program applied in the first twelve months of life on the primary prevention of asthma and other atopic disorders in high-risk infants.<sup>E6</sup> Infants at high risk for developing asthma and other allergic disorders were

identified and their mothers were recruited during their third trimester of pregnancy. Infants at high risk were defined as those who had a parent with asthma or two first-degree relatives with allergies or atopic disorders. The study had two recruitment centers in Canada (Vancouver and Winnipeg). The diagnosis of asthma at the seven year time point<sup>E7</sup> was made by a pediatric allergist in each center without knowledge of the group allocation status of the children and who did not provide health care services to the families. The physician conducted a structured interview with parents by using a standardized form to record symptoms and physical findings. Spirometry and methacholine challenge testing were performed after obtaining parental consent. The diagnoses of asthma and other atopic disorders were clinical decisions made by the pediatric allergists without knowledge of the results of the questionnaire, allergy skin tests, spirometry, or methacholine challenge tests. Atopy was defined as at least one positive response (wheal diameter  $\geq 3$  mm at 10 minutes) on skin prick tests. Allergy skin tests were performed with the epicutaneous method by using a prick lanceter (Hollister-Stier, Omega Laboratories LTD, Montreal, Canada) with the following allergens (Hollister-Stier, Omega Laboratories LTD): house dust mite (*Dermatophagoides pteronyssinus* and *Dermatophagoides farinae*), cat, dog, cockroach, *Alternaria* species, *Cladosporium* species, tree, grass and weed (ragweed) pollens, cow's milk, egg white, wheat, soy, and peanut. Histamine (1 mg/mL) was used as the positive control, and saline was used as the negative control. The largest wheal diameter and its perpendicular were measured at 15 minutes after testing. A mean wheal diameter of 3 mm or greater than that elicited by the negative control was considered a positive reaction. Atopy was defined as a positive skin test reaction to one or more of these common allergens.

Methacholine (Methapharm Inc, Brantford, Ontario, Canada) challenge testing was carried out according to the protocol of Cockcroft et al.<sup>E8</sup> Two children who had FEV<sub>1</sub> values of less than 70% of predicted value were excluded from methacholine testing, but both of them had a postbronchodilator change in FEV<sub>1</sub> of 12% or more, and they were considered to have airway hyperresponsiveness (AHR). The provocative concentration of methacholine that induced a 20% decrease in FEV<sub>1</sub> from postsaline value (PC<sub>20</sub>) was determined. The Ethics Committees of the University of British Columbia and the University of Manitoba approved the study and parents provided written informed consent for participation.

### **The Study of Asthma Genes and the Environment Cohort (SAGE)**

The Study of Asthma Genes and the Environment (SAGE), is derived from a population-based cohort of 16,320 children born in province of Manitoba, Canada, between January 1, 1995 through December 31, 1995. A survey on child health and home environment exposures was mailed to the parents of this birth cohort in 2002. Following parent response to the mailed survey and approval to be contacted, a subset of children was invited to join the study at age 8-10 years. This included all children with parent-declared asthma and a comparable number of children without asthma, randomly selected following stratification of the received surveys by urban / rural location and family income. All recruited children underwent clinical assessment for asthma by a pediatric allergist (AB) on the basis of the Canadian Asthma Consensus Guidelines. To aid in the diagnosis, a standardized history was used including questions on cough, wheeze, shortness of breath, response to current medications (i.e. bronchodilators, corticosteroids) and the presence of other allergic conditions (e.g. allergic rhinitis, atopic dermatitis, and food allergies) were elicited. Physical examination included examination for chest symptoms (hyperinflation, wheeze, prolonged expiration and decreased breath sounds). Hospitalization and medical visits for breathing difficulty in the past year were also noted. All findings were recorded and a diagnosis of asthma in the last 12 months was made. In total 723 study trios (child and both parents) were recruited, 247 of which have a child affected with asthma. A child affected with asthma was defined as follows: pediatric allergist diagnosis of asthma based upon history and the physical exam. Airway hyperresponsiveness (AHR) was assessed in all children following a methacholine challenge test (Methapharm Inc, Brantford, Ontario Canada), carried out according to the protocol of Cockcroft et al.<sup>E8</sup> The children with  $FEV_1 < 70\%$  had reversibility testing with an aerosol bronchodilator. Reversibility was defined as an improvement in  $FEV_1$  by  $> 12\%$  following inhalation of a beta2-adrenergic agonist. Bronchoprovocation was not attempted for these children. Atopic status was determined for all children; skin prick tests to common relevant allergens were performed (Tree pollen mix, Weed pollen mix, Ragweed (separate), Grass pollen mix, Alternaria, Cladosporium, Penicillium, House dust mites (*D. pteronyssinus*, *D. farinae*), Cockroach, Cat, Dog, Feathers and Peanut) and atopy was defined as at least one positive response (wheal diameter  $\geq 3$  mm at 10 minutes). For children in Winnipeg and immediate surroundings the assessment was carried out in the investigator's laboratories in the Manitoba Institute of Child Health. For children in



remote and northern communities a team travelled to the communities to assess children in and surrounding those communities. The study was approved by the Ethics Committee at the University of Manitoba and parents provided written informed consent for participation.

#### **The Busselton Health Study Population (Busselton cohort)**

Residents of the town of Busselton in the southwest of Western Australia have been involved in a series of health surveys since 1966. The population is predominantly of European origin. In the present nested case-control study, all subjects ( $n = 1,599$ , of which 1,549 were sent for genotyping (679 cases and 870 controls)) who attended any of 6 cross-sectional surveys from 1996 to 1981 and the follow-up survey in 1994 and who had available DNA were included. Subjects were considered to have asthma if they answered yes to the question "Has your doctor ever told you that you had asthma/bronchial asthma?" at any survey. All other subjects were considered as controls. The panel includes a sample of allergic and non-allergic individuals as determined by skin-prick tests to common allergens. Methacholine challenge tests using a modification of the Yan method<sup>E9</sup> were performed on all participants of the 1994 follow-up survey. The early dose-response slope with provocative dose producing a 20% fall in FEV1 (PD20) was determined.

#### **Pairwise linkage disequilibrium (LD) pattern**

The strength of linkage disequilibrium (LD) between pairs of SNPs in each gene has been measured as  $D'$ <sup>E10</sup>, using Haploview software (3.32 version).<sup>E11</sup> Figures E1 to E8 present these LD patterns for each studied gene in each studied sample.

#### **REFERENCES**

- E1. Heyer E, Tremblay M. Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 1995; 56:970-978.
- E2. Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, Weissenbach J, Popowska E, Pronicka E, Root AW, Glorieux FH. Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am J Hum Genet* 1996; 59:633-643.

- E3. American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. *Am Rev Respir Dis* 1987; 136:225-44.
- E4. American Thoracic Society. ATS statement--Snowbird workshop on standardization of spirometry. *Am Rev Respir Dis* 1979; 119:831-838.
- E5. American Thoracic Society. Standardization of Spirometry, 1994 Update. *Am J Respir Crit Care Med* 1995; 152: 1107-1136.
- E6. Chan-Yeung M, Manfreda J, Dimich-Ward H, Ferguson A, Watson W, Becker A. A randomized controlled study on the effectiveness of a multifaceted intervention program in the primary prevention of asthma in high-risk infants. *Arch Pediatr Adolesc Med* 2000; 154:657-663.
- E7. Chan-Yeung M, Ferguson A, Watson W, Dimich-Ward H, Rousseau R, Lilley M, Dybuncio A, Becker A. The Canadian Childhood Asthma Primary Prevention Study: outcomes at 7 years of age. *J Allergy Clin Immunol* 2005; 116:49-55.
- E8. Crockcroft DW, Murdock KY, Berscheid BA. Relationship between atopy and bronchial responsiveness to histamine in a random population. *Ann Allergy* 1984; 53:26-29.
- E9. Bremner PR, de Klerk NH, Ryan GF, James AL, Musk M, Murray C, Le Souef PN, Young S, Spargo R, Musk AW. Respiratory symptoms and lung function in aborigines from tropical Western Australia. *Am J Respir Crit Care Med* 1998; 158:1724-1729.
- E10. Lewontin RC. On measures of gametic disequilibrium. *Genetics* 1988; 120:849-852.
- E11. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; 21:263-265.

## LEGEND TO FIGURES

### Figure E1: LD pattern for *ALOX15*

A) *ALOX15* LD pattern in the SLSJ familial sample; B) *ALOX15* LD pattern in the CAPPS cohort; C) *ALOX15* LD pattern in the SAGE cohort; D) *ALOX15* LD pattern in the Busselton cohort. The location of each tested tagSNP along the chromosome is indicated in the upper part of the figure. The number in each diamond indicates the magnitude of LD ( $D'$ ) in percent between respective pairs of SNPs. Strength of LD is

depicted by progression of color: the color moves from red to white as  $D'$  runs from 1 to 0.

**Figure E2:** LD pattern for *CD14*

A) *CD14* LD pattern in the SLSJ familial sample; B) *CD14* LD pattern in the CAPPS cohort; C) *CD14* LD pattern in the SAGE cohort; D) *CD14* LD pattern in the Busselton cohort. The location of each tested tagSNP along the chromosome is indicated in the upper part of the figure. The number in each diamond indicates the magnitude of LD ( $D'$ ) in percent between respective pairs of SNPs. Strength of LD is depicted by progression of color: the color moves from red to white as  $D'$  runs from 1 to 0.

**Figure E3:** LD pattern for *CD27*

A) *CD27* LD pattern in the SLSJ familial sample; B) *CD27* LD pattern in the CAPPS cohort; C) *CD27* LD pattern in the SAGE cohort; D) *CD27* LD pattern in the Busselton cohort. The location of each tested tagSNP along the chromosome is indicated in the upper part of the figure. The number in each diamond indicates the magnitude of LD ( $D'$ ) in percent between respective pairs of SNPs. Strength of LD is depicted by progression of color: the color moves from red to white as  $D'$  runs from 1 to 0.

**Figure E4:** LD pattern for *CXCL12*

A) *CXCL12* LD pattern in the SLSJ familial sample; B) *CXCL12* LD pattern in the CAPPS cohort; C) *CXCL12* LD pattern in the SAGE cohort; D) *CXCL12* LD pattern in the Busselton cohort. The location of each tested tagSNP along the chromosome is indicated in the upper part of the figure. The number in each diamond indicates the magnitude of LD ( $D'$ ) in percent between respective pairs of SNPs. Strength of LD is depicted by progression of color: the color moves from red to white as  $D'$  runs from 1 to 0.

**Figure E5:** LD pattern for *IL2RB*

A) *IL2RB* LD pattern in the SLSJ familial sample; B) *IL2RB* LD pattern in the CAPPS cohort; C) *IL2RB* LD pattern in the SAGE cohort; D) *IL2RB* LD pattern in the Busselton cohort. The location of each tested tagSNP along the chromosome is indicated in the upper part of the figure. The number in each diamond indicates the magnitude of LD ( $D'$ ) in percent between respective pairs of SNPs. Strength of LD is

depicted by progression of color: the color moves from red to white as  $D'$  runs from 1 to 0.

**Figure E6:** LD pattern for *IL7R*

A) *IL7R* LD pattern in the SLSJ familial sample; B) *IL7R* LD pattern in the CAPPS cohort; C) *IL7R* LD pattern in the SAGE cohort; D) *IL7R* LD pattern in the Busselton cohort. The location of each tested tagSNP along the chromosome is indicated in the upper part of the figure. The number in each diamond indicates the magnitude of LD ( $D'$ ) in percent between respective pairs of SNPs. Strength of LD is depicted by progression of color: the color moves from red to white as  $D'$  runs from 1 to 0.

**Figure E7:** LD pattern for *NOS2A*

A) *NOS2A* LD pattern in the SLSJ familial sample; B) *NOS2A* LD pattern in the CAPPS cohort; C) *NOS2A* LD pattern in the SAGE cohort; D) *NOS2A* LD pattern in the Busselton cohort. The location of each tested tagSNP along the chromosome is indicated in the upper part of the figure. The number in each diamond indicates the magnitude of LD ( $D'$ ) in percent between respective pairs of SNPs. Strength of LD is depicted by progression of color: the color moves from red to white as  $D'$  runs from 1 to 0.

**Figure E8:** LD pattern for *SFRP1*

A) *SFRP1* LD pattern in the SLSJ familial sample; B) *SFRP1* LD pattern in the CAPPS cohort; C) *SFRP1* LD pattern in the SAGE cohort; D) *SFRP1* LD pattern in the Busselton cohort. The location of each tested tagSNP along the chromosome is indicated in the upper part of the figure. The number in each diamond indicates the magnitude of LD ( $D'$ ) in percent between respective pairs of SNPs. Strength of LD is depicted by progression of color: the color moves from red to white as  $D'$  runs from 1 to 0.

**Table E1**

FBAT values for all tagSNPs of the SLSJ sample that generated significant associations in any one population for one or more of the studied phenotypes

Gene	TagSNP	Base change	MAF	Asthma			Atopy			AHR			Allergic asthma		
				Fam	Z	P	Fam	Z	P	Fam	Z	P	Fam	Z	P
<i>ALOX15</i>	rs4790210	T > C	0.18	102	0.96	0.34	86	0.53	0.60	88	0.33	0.74	90	0.53	0.60
	rs916055	T > C	0.35	136	-0.55	0.58	115	-0.40	0.69	116	-0.68	0.50	115	-0.70	0.49
	rs743646	T > C	0.10	59	-0.12	0.90	45	0.00	1.00	48	0.26	0.80	48	0.20	0.84
	rs7217186	A > G	0.47	130	0.68	0.49	115	1.13	0.26	113	0.44	0.66	114	0.64	0.52
	rs2664593	G > C	0.19	95	0.31	0.75	77	0.33	0.74	79	1.12	0.26	82	0.45	0.65
<i>CD14</i>	rs778583	C > T	0.26	113	0.76	0.45	91	0.39	0.70	93	0.07	0.94	95	0.85	0.39
	rs4914	G > C	0.11	64	-0.88	0.38	54	-0.60	0.55	53	-1.15	0.25	54	-0.41	0.68
<i>CD27</i>	rs7970260	A > C	0.45	117	0.40	0.69	93	0.38	0.71	97	0.97	0.33	94	0.34	0.74
	rs3136551	A > G	0.04	29	0.43	0.67	24	0.57	0.57	26	0.65	0.52	24	0.00	1.00
<i>CXCL12</i>	rs2861442	A > G	0.25	113	-0.47	0.64	96	-0.78	0.44	92	-0.85	0.40	98	-0.74	0.46
<i>IL2RB</i>	rs6000570	G > A	0.42	117	0.85	0.39	102	1.34	0.18	99	0.94	0.35	103	1.38	0.17
	rs3218312	A > G	0.19	83	-1.54	0.12	68	-1.17	0.24	65	-1.49	0.14	70	-0.97	0.33
	rs228954	A > G	0.42	118	-0.10	0.92	105	0.66	0.51	106	-0.25	0.80	105	0.78	0.43
	rs2281094	C > T	0.25	109	-0.40	0.69	97	0.46	0.65	95	-0.44	0.66	101	0.14	0.89
	rs228979*	C > A	0.27	109	-1.59	0.11	92	-2.10	<b>0.0358</b> ( > 1.0)	92	-1.20	0.23	94	-2.29	<b>0.0220</b> (0.9240)
	rs2016771	G > T	0.34	128	2.26	<b>0.0236</b> (0.9912)†	109	1.46	0.14	106	1.36	0.17	113	1.61	0.11
<i>NOS2A</i>	rs4796017	T > C	0.41	123	-1.60	0.11	111	-1.29	0.20	108	-1.20	0.23	112	-1.44	0.15
	rs2297518	G > A	0.22	107	-0.70	0.48	96	-0.90	0.37	94	-0.07	0.94	94	-1.13	0.26

	rs3794766	G > A	0.26	127	-1.56	0.12	111	-1.38	0.17	105	-0.87	0.39	109	-2.13	<b>0.0334</b> (0.8016)
	rs2779248	A > G	0.45	132	-1.26	0.21	112	-1.92	0.06	102	-0.51	0.61	111	-2.34	<b>0.0194</b> (0.4656)
	rs2779252	G > T	0.05	36	0.44	0.66	33	0.47	0.64	30	0.72	0.47	34	0.33	0.74
	rs2531860	C > T	0.24	107	-0.40	0.69	93	-0.07	0.95	92	-0.15	0.88	92	-0.10	0.92
	rs1127379	T > C	0.47	132	-0.43	0.67	111	0.04	0.97	110	-0.59	0.55	116	-0.49	0.62
	rs7013368	C > T	0.33	124	0.64	0.52	101	0.98	0.33	105	0.22	0.82	108	0.57	0.57
<i>SFRP1</i>	rs11774662	T > C	0.34	125	1.19	0.23	106	1.45	0.15	105	1.13	0.26	112	1.39	0.16
	rs968427	T > C	0.48	130	-0.07	0.95	114	-0.59	0.55	118	0.25	0.80	119	-0.43	0.67
	rs921142	T > C	0.38	125	-0.18	0.86	110	-0.19	0.85	106	0.20	0.84	115	-0.43	0.66

P-values  $\leq 0.05$  are marked in bold. Grey shadow indicates that the tagSNP has at least one p-value under 0.05.

Abbreviations used: TagSNP = Tagging single nucleotide polymorphism, MAF = Minor allele frequency, AHR = Airway hyperresponsiveness, N = Number of families contributing the statistic, Z = Z score, P = p-value (corrected p-value are indicated in parentheses).

\* This tagSNP has been tested in SLSJ and CAPPS only.

† Number of independent SNPs is estimated to be 14 for *IL2RB* and 8 for *NOS2A*. Number of independent phenotypes is estimated to be 3.

**Table E2**

FBAT values for all tagSNPs of the CAPPS sample that generated significant associations in any one population for one or more of the studied phenotypes

Gene	TagSNP	Base change	MAF	Asthma			Atopy			AHR			Allergic asthma		
				Fam	Z	P	Fam	Z	P	Fam	Z	P	Fam	Z	P
<i>ALOX15</i>	rs4790210	T > C	0.21	29	-2.61	<b>0.0090</b> (0.1620) <sup>†</sup>	61	-2.85	<b>0.0044</b> (0.0792)	101	-2.06	<b>0.0397</b> (0.7146)	22	-2.75	<b>0.0060</b> (0.1080)
	rs916055	T > C	0.34	29	-1.33	0.18	78	-1.69	0.09	129	-2.59	<b>0.0095</b> (0.1710)	21	-1.57	0.12
	rs743646	T > C	0.10	26	0.58	0.56	44	0.44	0.66	66	2.15	<b>0.0314</b> (0.5652)	21	0.43	0.67
	rs7217186	A > G	0.47	43	-2.21	<b>0.0269</b> (0.4842)	99	-1.31	0.19	154	-3.24	<b>0.0012</b> <b>(0.0216)</b>	32	-1.94	0.05
	rs2664593	G > C	0.21	34	1.27	0.21	69	0.66	0.51	107	1.98	<b>0.0482</b> (0.8676)	26	1.06	0.29
<i>CD14</i>	rs778583	C > T	0.26	34	1.12	0.26	72	-0.21	0.83	109	0.88	0.38	25	0.56	0.58
	rs4914	G > C	0.11	17	0.47	0.64	47	-0.71	0.48	65	2.02	<b>0.0436</b> (0.5232)	12	-0.28	0.78
<i>CD27</i>	rs7970260	A > C	0.40	40	0.67	0.50	89	0.29	0.78	142	-0.73	0.47	32	1.07	0.29
	rs3136551	A > G	0.05	6	NA	NA	22	0.00	1.00	32	-0.69	0.49	4	NA	NA
<i>CXCL12</i>	rs2861442	A > G	0.27	38	0.00	1.00	79	-0.40	0.69	123	0.98	0.33	29	-0.33	0.74
<i>IL2RB</i>	rs6000570	A > G	0.48	40	-0.14	0.89	95	-0.64	0.52	143	0.58	0.56	31	-0.16	0.87
	rs3218312	A > G	0.22	32	0.48	0.63	67	1.13	0.26	112	0.00	1.00	23	1.35	0.18
	rs228954	A > G	0.43	42	0.53	0.60	84	-1.17	0.24	142	1.03	0.30	30	-0.16	0.87
	rs2281094	C > T	0.25	40	0.83	0.41	79	-1.11	0.27	119	1.07	0.28	29	0.65	0.52

	rs228979*	C > A	0.29	41	-1.24	0.22	79	-1.58	0.11	133	0.08	0.94	28	-2.27	<b>0.0231</b> ( > 1.0)
	rs2016771	G > T	0.40	43	0.79	0.43	95	0.36	0.72	140	0.74	0.46	32	1.07	0.29
	rs4796017	T > C	0.43	44	0.94	0.35	98	0.62	0.54	150	-0.63	0.53	32	0.95	0.34
	rs2297518	G > A	0.19	30	2.14	<b>0.0326</b> (0.7824)	68	1.33	0.19	110	0.00	1.00	21	2.20	<b>0.0278</b> (0.6672)
<i>NOS2A</i>	rs3794766	G > A	0.23	35	0.95	0.34	76	0.52	0.60	124	-0.83	0.40	25	1.46	0.14
	rs2779248	A > G	0.35	45	-0.39	0.70	87	0.20	0.84	138	-0.83	0.41	34	0.15	0.88
	rs2779252	G > T	0.08	12	1.16	0.25	38	-0.32	0.75	51	0.69	0.49	9	NA	NA
	rs2531860	C > T	0.15	27	-1.92	0.06	43	-1.54	0.12	80	-1.56	0.12	18	-2.29	<b>0.0218</b> (0.5232)
	rs1127379	T > C	0.47	41	0.40	0.69	100	-1.69	0.09	154	-1.34	0.18	30	0.62	0.54
	rs7013368	C > T	0.36	40	-0.41	0.69	89	-1.64	0.10	149	-0.91	0.36	30	-0.31	0.76
<i>SFRP1</i>	rs11774662	T > C	0.36	40	-1.07	0.29	90	-1.81	0.07	150	-1.48	0.14	28	-1.23	0.22
	rs968427	T > C	0.43	44	-0.52	0.61	102	0.69	0.49	156	-1.10	0.27	34	-0.30	0.77
	rs921142	T > C	0.38	42	-0.53	0.60	103	0.85	0.39	150	-1.33	0.18	31	-0.46	0.65

P-values  $\leq 0.05$  are marked in bold and underlined if they remain significant after correction for multiple testing. Grey shadow indicates that the tagSNP has at least one p-value under 0.05.

Abbreviations used: TagSNP = Tagging single nucleotide polymorphism, MAF = Minor allele frequency, AHR = Airway hyperresponsiveness, N = Number of families contributing the statistic, Z = Z score, P = p-value (corrected p-value are indicated in parentheses), NA = Not available because the number of families is too low (< 10) to contribute to the test.

\* This tagSNP has been tested in SLSJ and CAPPS only.

† Number of independent SNPs is estimated to be 6 for *ALOX15*, 4 for *CD14*, 21 for *IL2RB* and 8 for *NOS2A*. Number of independent phenotypes is estimated to be 3.



**Table E3**

FBAT values for all tagSNPs of the SAGE sample that generated significant associations in any one population for one or more of the studied phenotypes

Gene	TagSNP	Base change	MAF	Asthma			Atopy			AHR			Allergic asthma		
				Fam	Z	P	Fam	Z	P	Fam	Z	P	Fam	Z	P
<i>ALOX15</i>	rs4790210	T > C	0.20	63	2.19	<b>0.0282</b> (0.5076)†	93	1.73	0.08	74	2.16	<b>0.0310</b> (0.5580)	43	2.38	<b>0.0173</b> (0.3114)
	rs916055	T > C	0.35	92	0.83	0.41	132	0.31	0.76	103	0.69	0.49	61	0.78	0.44
	rs743646	T > C	0.10	44	-1.13	0.26	54	-0.38	0.71	45	-0.94	0.35	24	-1.35	0.18
	rs7217186	A > G	0.48	101	0.61	0.54	145	0.58	0.56	114	0.16	0.87	67	1.39	0.16
	rs2664593	G > C	0.23	71	-2.09	<b>0.0370</b> (0.6660)	109	-2.23	<b>0.0258</b> (0.4644)	84	-2.05	<b>0.0404</b> (0.7272)	47	-2.99	<b>0.002755</b> (0.04959)
<i>CD14</i>	rs778583	C > T	0.27	75	0.31	0.76	106	-0.61	0.54	84	-0.19	0.85	53	0.87	0.39
	rs4914	G > C	0.11	36	-2.92	<b>0.0035</b> (0.0420)	51	-1.60	0.11	41	-1.68	0.09	25	-1.57	0.12
<i>CD27</i>	rs7970260	A > C	0.41	90	0.09	0.93	127	1.49	0.14	100	0.71	0.48	58	0.69	0.49
	rs3136551	A > G	0.04	14	0.00	1.00	23	0.21	0.83	20	0.45	0.65	8	NA	NA
<i>CXCL12</i>	rs2861442	A > G	0.26	74	1.92	0.06	104	2.03	<b>0.0429</b> (> 1.0)	83	1.85	0.06	49	1.34	0.18
	rs6000570	G > A	0.47	105	0.43	0.66	139	1.96	<b>0.0500</b> (> 1.0)	116	1.24	0.22	71	0.42	0.67
<i>IL2RB</i>	rs3218312	A > G	0.20	63	-0.47	0.64	99	-0.37	0.72	74	0.11	0.92	45	-0.54	0.59
	rs228954	A > G	0.45	93	-0.71	0.48	125	-1.55	0.12	106	-1.03	0.30	62	-0.66	0.51
	rs2281094	C > T	0.31	83	0.78	0.44	107	-0.78	0.44	87	-0.29	0.78	50	1.12	0.26
	rs3218266*	C > G	0.38	90	2.44	<b>0.0149</b> (0.9387)	129	-0.40	0.69	103	0.09	0.93	62	1.73	0.08
	rs2016771	G > T	0.44	90	1.46	0.14	126	-0.39	0.70	104	0.42	0.67	56	0.93	0.35

<i>NOS2A</i>	rs4796017	T > C	0.46	82	2.35	<b>0.0187</b> (0.4488)	124	1.64	0.10	100	1.29	0.20	53	2.33	<b>0.0201</b> (0.4824)
	rs2297518	G > A	0.18	68	1.25	0.21	88	1.12	0.26	69	1.59	0.11	42	1.16	0.25
	rs3794766	G > A	0.24	75	0.43	0.67	105	0.27	0.79	90	1.17	0.24	47	0.14	0.89
	rs2779248	A > G	0.37	92	-1.11	0.27	128	-0.62	0.54	109	-0.59	0.55	58	-0.58	0.56
	rs2779252	G > T	0.06	25	-1.13	0.26	37	-0.32	0.75	27	-1.98	<b>0.0482</b> (> 1.0)	17	-1.61	0.11
	rs2531860	C > T	0.15	52	-1.03	0.30	69	-1.44	0.15	58	-1.18	0.24	30	-0.51	0.61
<i>SFRP1</i>	rs1127379	T > C	0.44	95	-0.79	0.43	127	-2.11	<b>0.0348</b> (0.6264)	111	-1.10	0.27	62	-0.22	0.83
	rs7013368	C > T	0.34	93	-1.38	0.17	119	-2.21	<b>0.0270</b> (0.4860)	100	-1.38	0.17	63	-0.56	0.57
	rs11774662	T > C	0.32	90	-2.03	<b>0.0428</b> (0.7704)	111	-1.58	0.11	92	-1.30	0.19	58	-0.92	0.36
	rs968427	T > C	0.48	99	-0.43	0.66	139	1.08	0.28	108	-2.06	<b>0.0392</b> (0.7056)	71	0.31	0.76
	rs921142	T > C	0.42	96	-0.72	0.47	135	0.68	0.50	106	-2.23	<b>0.0258</b> (0.4644)	68	0.32	0.75

P-values  $\leq 0.05$  are marked in bold and underlined if they remain significant after correction for multiple testing. Grey shadow indicates that the tagSNP has at least one p-value under 0.05.

Abbreviations used: TagSNP = Tagging single nucleotide polymorphism, MAF = Minor allele frequency, AHR = Airway hyperresponsiveness, N = Number of families contributing the statistic, Z = Z score, P = p-value (corrected p-value are indicated in parentheses), NA = Not available because the number of families is too low (< 10) to contribute to the test.

\* This tagSNP has been tested in SAGE and Busselton only.

† Number of independent SNPs is estimated to be 6 for *ALOX15*, 4 for *CD14*, 8 for *CXCL12*, 21 for *IL2RB*, 8 for *NOS2A* and 6 for *SFRP1*. Number of independent phenotypes is estimated to be 3.

**Table E4**

Odds ratio values for all tagSNPs of the Busselton sample that generated significant associations in any one population for one or more of the studied phenotypes

Gene	TagSNP	Base change	MAF	Asthma		Atopy		AHR		Allergic asthma	
				OR	P	OR	P	OR	P	OR	P
<i>ALOX15</i>	rs4790210	T > C	0.18	0.99	0.89	0.89	0.28	0.83	0.21	0.91	0.47
	rs916055	T > C	0.31	0.99	0.92	0.85	0.07	0.90	0.37	0.86	0.15
	rs743646	T > C	0.12	1.00	0.98	1.07	0.61	1.13	0.48	1.04	0.81
	rs7217186	A > G	0.46	0.95	0.47	0.83	<b>0.026</b> (0.39)	0.90	0.35	0.82	<b>0.0488</b> (0.7320)
	rs2664593	G > C	0.22	1.10	0.28	1.25	<b>0.028</b> (0.42)	1.18	0.19	1.26	0.06
<i>CD14</i>	rs778583	C > T	0.28	1.03	0.71	0.91	0.32	0.72	<b>0.0079</b> (0.0711)	0.91	0.43
	rs4914	G > C	0.12	1.14	0.27	1.03	0.80	0.90	0.5410	1.09	0.57
<i>CD27</i>	rs7970260	A > C	0.38	0.99	0.94	1.02	0.81	1.30	<b>0.0182</b> (0.2730)	1.00	1.00
	rs3136551	A > G	0.04	0.59	<b>0.0075</b> (0.1125)†	0.71	0.10	1.14	0.63	0.45	<b>0.0037</b> (0.0555)
<i>CXCL12</i>	rs2861442	A > G	0.31	0.95	0.51	0.94	0.49	0.98	0.86	0.91	0.3902
<i>IL2RB</i>	rs6000570	G > A	0.49	1.04	0.65	1.08	0.38	1.09	0.43	0.90	0.28
	rs3218312	A > G	0.23	1.07	0.47	1.02	0.88	0.76	<b>0.0393</b> (> 1.0)	1.04	0.74
	rs228954	A > G	0.44	1.10	0.20	1.06	0.47	1.32	<b>0.0108</b> (0.4536)	1.14	0.19
	rs2281094	C > T	0.27	1.10	0.27	1.19	0.06	1.31	<b>0.0255</b> (> 1.0)	1.25	0.0513

	rs3218266*	C > G	0.33	1.03	0.70	1.14	0.15	1.20	0.11	1.17	0.14
	rs2016771	G > T	0.41	1.01	0.95	1.03	0.75	1.00	0.98	1.03	0.79
<i>NOS2A</i>	rs4796017	T > C	0.43	1.03	0.75	1.02	0.84	0.96	0.69	1.02	0.85
	rs2297518	G > A	0.19	1.06	0.58	0.98	0.84	0.85	0.27	1.05	0.71
	rs3794766	G > A	0.23	1.15	0.12	1.07	0.48	0.86	0.27	1.19	0.15
	rs2779248	A > G	0.37	1.01	0.90	1.13	0.14	0.97	0.80	1.14	0.20
	rs2779252	G > T	0.06	1.26	0.16	0.94	0.72	0.72	0.20	1.13	0.58
	rs2531860	C > T	0.15	1.15	0.18	1.18	0.16	1.15	0.35	1.33	<b>0.0454</b> (0.9534)
<i>SFRP1</i>	rs1127379	T > C	0.42	0.92	0.26	1.01	0.95	1.19	0.11	0.94	0.57
	rs7013368	C > T	0.37	0.90	0.17	1.02	0.82	1.01	0.93	0.93	0.50
	rs11774662	T > C	0.36	0.89	0.16	0.98	0.81	0.96	0.75	0.89	0.28
	rs968427	T > C	0.44	1.06	0.45	1.04	0.66	1.01	0.93	1.05	0.62
	rs921142	T > C	0.39	1.04	0.63	1.03	0.73	0.94	0.56	1.04	0.71

P-values  $\leq 0.05$  are marked in bold. Grey shadow indicates that the tagSNP has at least one p-value under 0.05.

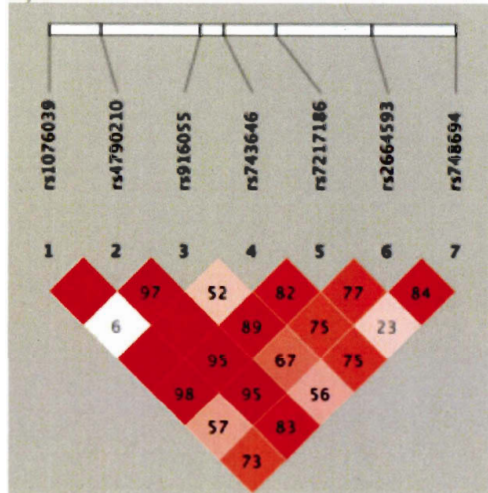
Abbreviations used: TagSNP = Tagging single nucleotide polymorphism, MAF = Minor allele frequency, AHR = Airway hyperresponsiveness, OR = 95% confidence interval odd ratio, P = p-value (corrected p-value are indicated in parentheses).

\* This tagSNP has been tested in SAGE and Busselton only.

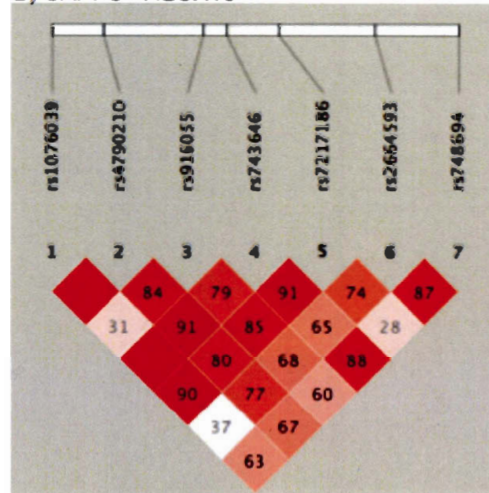
† Number of independent SNPs is estimated to be 5 for *ALOX15*, 3 for *CD14*, 5 for *CD27*, 14 for *IL2RB* and 7 for *NOS2A*. Number of independent phenotypes is estimated to be 3.

Figure E1

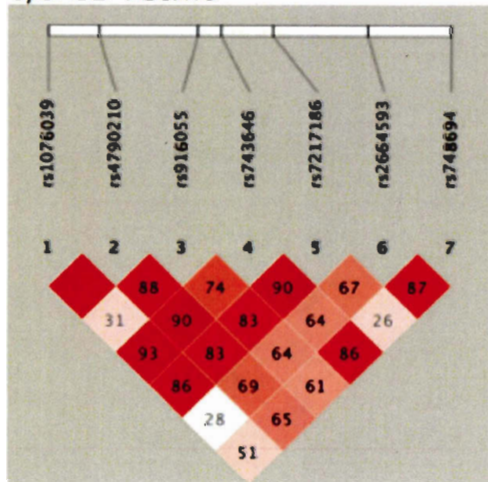
A) SLSJ - ALOX15



B) CAPPs - ALOX15



C) SAGE - ALOX15



D) Busselton - ALOX15

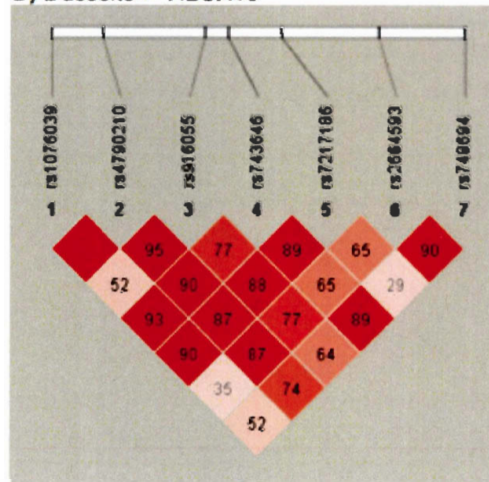
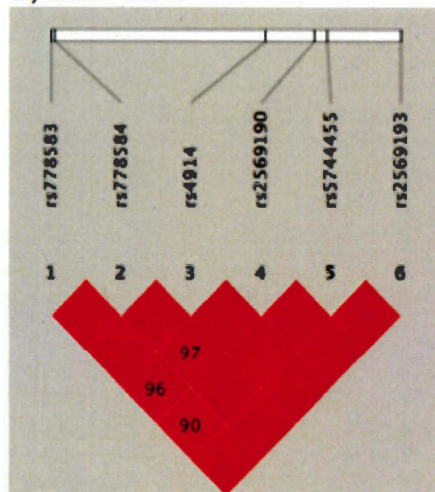
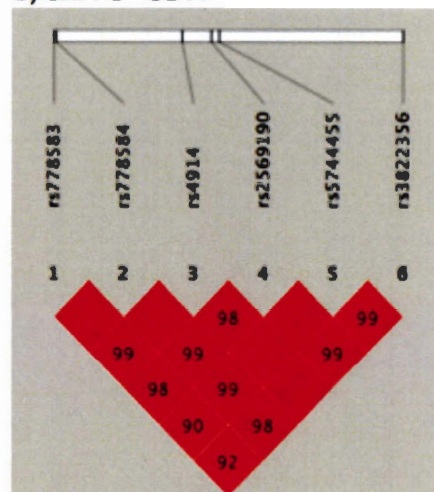


Figure E2

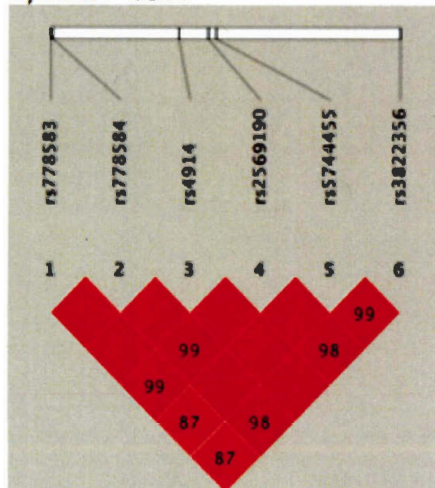
A) SLSJ - CD14



B) CAPPS - CD14



C) SAGE - CD14



D) Busselton - CD14

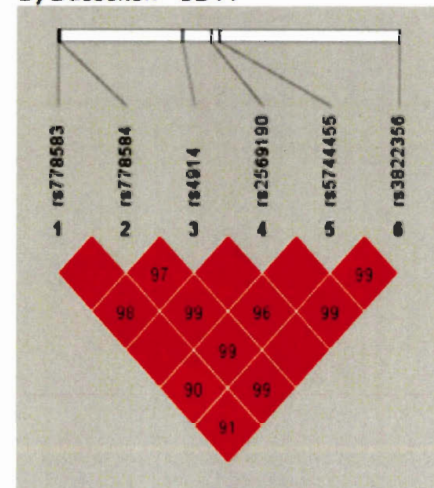
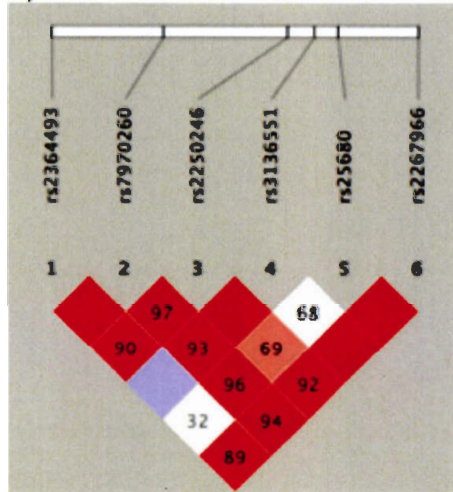
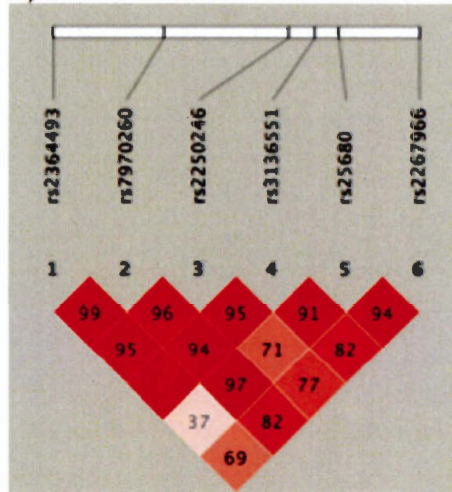


Figure E3

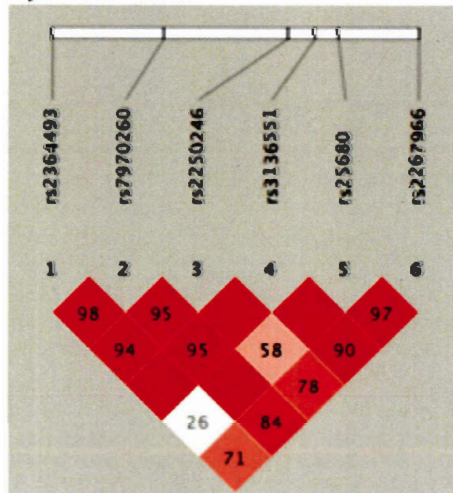
A) SLSJ - CD27



B) CAPPS - CD27



C) SAGE - CD27



D) Busselton - CD27

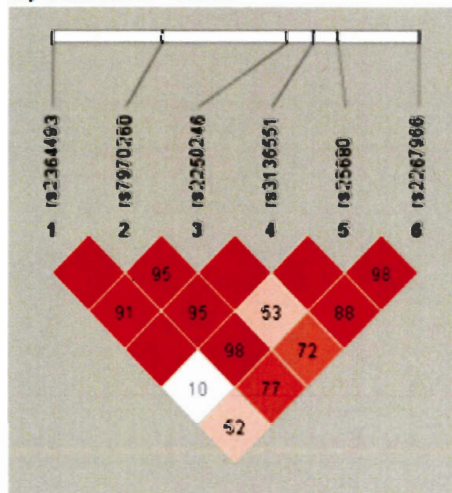










Figure E5 (continued)

C) SAGE - IL2RB

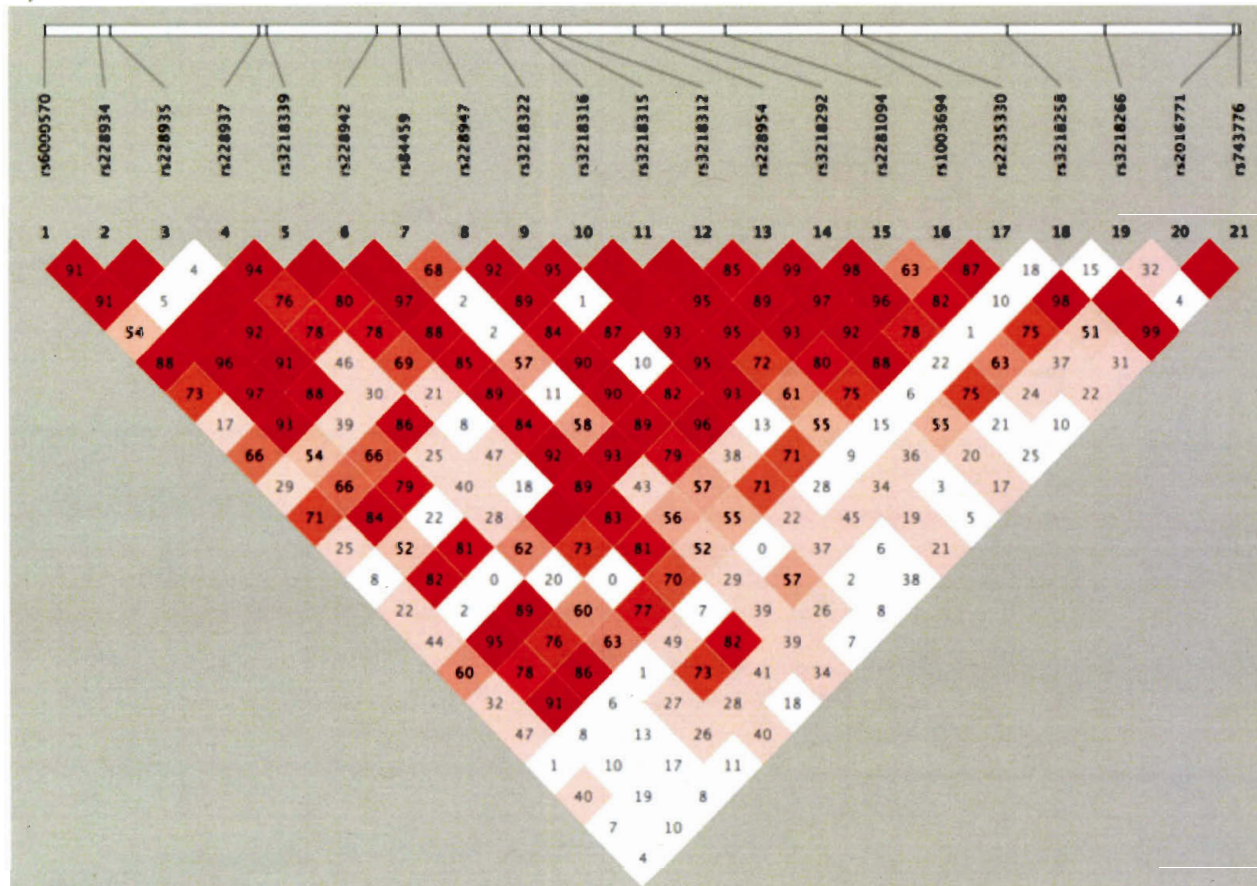
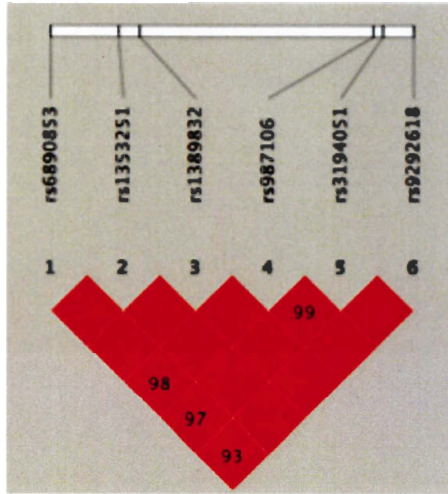
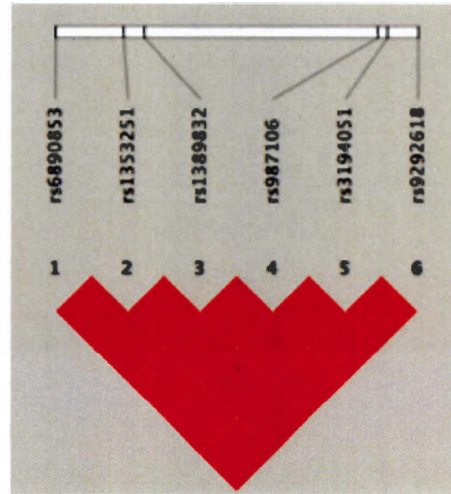
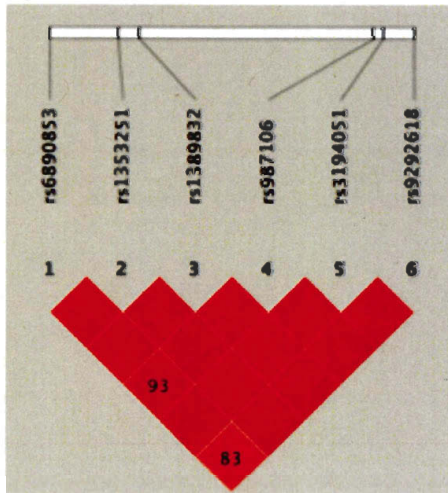
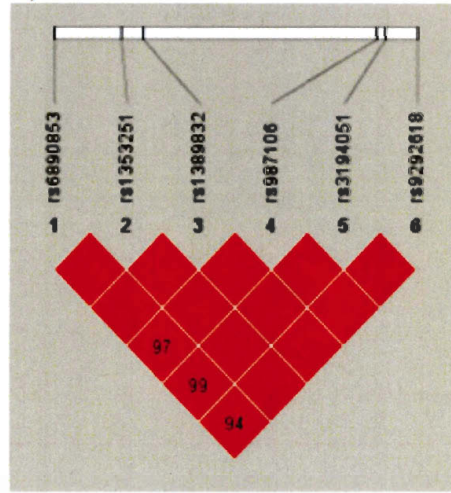




Figure E6

A) SLSJ - *IL7R*B) CAPPS - *IL7R*C) SAGE - *IL7R*D) Busselton - *IL7R*





### 5.3 Troisième article

#### **Étude d'association entre le gène de l'activateur du plasminogène de type urokinase (*PLAU*), l'asthme et l'atopie**

Philippe Bégin, Karine Tremblay, Denise Daley, Mathieu Lemire, Sébastien Claveau, Charleen Salesse, Sabine Kacel, Alexandre Montpetit, Allan Becker, Anita L. Kozyrskyj, Thomas J. Hudson et Catherine Laprise

*American Journal of Respiratory Critical Care Medicine* (2007) 175(11): 1109-1116

#### **Résumé**

L'activateur du plasminogène de type urokinase interagit avec son récepteur sur les cellules inflammatoires et migratrices pour moduler la protéolyse extracellulaire, l'adhésion et l'activation cellulaires. Il est essentiel au développement d'une réponse immunitaire appropriée et est impliqué dans le remodelage tissulaire. Le gène *plasminogen activator, urokinase (PLAU)*, codant pour cette enzyme, est situé en chromosome 10q24, une région chromosomique liée à l'asthme dans un échantillon familial du Saguenay–Lac-Saint-Jean. Nous posons ici l'hypothèse que le gène *PLAU* agirait comme un modulateur de la susceptibilité à l'asthme. Le gène *PLAU* a été séquencé et ses variants analysés pour associations génétiques avec des traits liés à l'asthme dans l'échantillon familial du Saguenay–Lac-Saint-Jean (231 familles, 1139 sujets). Deux cohortes additionnelles ont été étudiées pour valider les observations (CAPPS, 238 trios, et SAGE, 237 trios). Dans l'échantillon original, selon un modèle dominant, les allèles rs2227564C (P141) et rs2227566T ont été associés avec l'asthme ( $p = 0,011$ ;  $p = 0,045$ ) et l'hyperréactivité bronchique (HRB) ( $p = 0,026$  ;  $p = 0,038$ ). L'haplotype fréquent était également associé à l'asthme, l'atopie et l'HRB ( $p = 0,031$  ;  $p = 0,043$  ;  $p = 0,006$ ). Bien qu'aucune association significative n'ait été détectée pour les polymorphismes individuels dans la cohorte CAPPS, une association négative a été observée dans la cohorte SAGE entre l'allèle rs4065C et l'atopie selon des modèles additif ( $p = 0,005$ ) et dominant ( $p = 0,0001$ ). Ces résultats suggèrent un rôle de l'activateur du plasminogène urokinase dans la pathogenèse de l'asthme.



### 5.3.1 TITLE PAGE

#### **Association of Urokinase-type Plasminogen Activator with Asthma and Atopy**

Philippe Bégin<sup>1,2</sup>, Karine Tremblay<sup>1,3</sup>, Denise Daley<sup>4</sup>, Mathieu Lemire<sup>5</sup>, Sébastien Claveau<sup>1,6</sup>, Charleen Salesse<sup>1,6</sup>, Sabine Kacel<sup>1,6</sup>, Alexandre Montpetit<sup>5</sup>, Allan Becker<sup>7</sup>, Anita L. Kozyrskyj<sup>7,8</sup>, Thomas J. Hudson<sup>9</sup> and Catherine Laprise<sup>1,6</sup>

<sup>1</sup> University of Montreal Community Genomic Medicine Centre, Chicoutimi University Hospital, Saguenay (Qc), Canada; <sup>2</sup> Department of Medicine, University of Montreal, Montreal (Qc), Canada; <sup>3</sup> Department of Medicine, Université Laval, Québec (Qc), Canada; <sup>4</sup> James Hogg iCapture Centre, University of British Columbia, Vancouver (BC), Canada; <sup>5</sup> McGill University and Genome Quebec Innovation Centre, Montréal (Qc), Canada; <sup>6</sup> Department of fundamental sciences, Université du Québec à Chicoutimi, Saguenay (Qc), Canada; <sup>7</sup> Department of Pediatrics and Child Health, Faculty of Medicine, University of Manitoba, Winnipeg (Man), Canada; <sup>8</sup> Faculty of Pharmacy and department of Community Health Sciences, Manitoba Centre for Health Policy, Winnipeg (Man) Canada; <sup>9</sup> Ontario Institute for Cancer Research, Toronto, (ON), Canada

#### **Corresponding author**

Catherine Laprise, Ph.D.  
Université du Québec à Chicoutimi  
555, boulevard de l'Université,  
Chicoutimi (Québec), Canada, G7H 2B1  
Telephone: (418) 541-1234, 3070; Fax: (418) 541-1116  
Email: claprise@saglac.qc.ca

#### **Funding sources**

This work was supported by the Canadian Institutes of Health Research and the Respiratory Health Network of the Fonds de la recherche en santé du Québec (FRSQ). AllerGen NCE Inc paid for the genotyping.

#### **Running title**

Association between PLAU and asthma

This article has an online data supplement, which is accessible from this issue's online table of contents at [www.atsjournals.org](http://www.atsjournals.org).

### 5.3.2 ABSTRACT

Urokinase plasminogen activator interacts with its receptor on inflammatory and migrating cells and regulates extracellular matrix degradation, cell adhesion and inflammatory cell activation. It is necessary for the development of an appropriate immune response and is involved in tissue remodeling. The PLAU gene codes for this enzyme and is located on 10q24. This region has demonstrated evidence for linkage in a genome scan for asthma in a sample from northeastern Quebec. Here, we hypothesized that PLAU may function as a regulator of asthma susceptibility. We sequenced PLAU and tested for genetic association between identified variants and asthma related traits in a French-Canadian cohort (231 families, 1139 subjects). Under dominant model, the common alleles rs2227564C (P141) and rs2227566T were associated with asthma ( $p = 0.011$ ;  $p = 0.045$ ) and with airway hyperresponsiveness (AHR) ( $p = 0.026$ ;  $p = 0.038$ ). Analysis of the linkage disequilibrium pattern also revealed association of the common haplotype for asthma, atopy and AHR ( $p = 0.031$ ;  $p = 0.043$ ;  $p = 0.006$ ). Additional studies were carried out in two other family-based Canadian cohorts. Whereas no significant association was detected for single PLAU SNPs in the CAPPS cohort (238 trios), association was observed in the SAGE cohort (237 trios) between the rs4065C allele and protection from atopy in additive ( $p = 0.005$ ) and dominant ( $p = 0.0001$ ) genetic models. This suggests a role for the urokinase plasminogen activation pathway in the pathogenesis of the disease.

Word count: 214

Key words: Asthma, Airway Hyperresponsiveness, Atopy, Association study, Haplotypes, Urokinase-type Plasminogen Activator (PLAU gene).

### 5.3.3 INTRODUCTION

Asthma is recognized as a complex trait and the genetic risk factors for its development, its progression and its subphenotypes remain largely unknown despite several genetic studies performed in the past decade. In a genome-wide scan conducted in 2004 to identify susceptibility loci for asthma and related phenotypes in a French-Canadian population, we identified eight chromosomal regions with suggestive evidence for linkage (1). Two regions were suggestive for asthma and atopy phenotypes (6q26 and 16p12.1, asthma) and (10q24 and 11q13, atopy), while the remaining four were associated with IgE levels (5q23.3, 6q27, 7q22.2 and 9q21.32). This was the first time that the 10q24 region was identified in a linkage study for asthma (LOD = 2.32,  $p = 0.00016$ ). This novel region was furthermore found to be associated with severe asthma in a more recent genome-wide association study conducted in a Puerto-Rican population ( $p < 10^{-4}$ ) (2).

The 10q24 region harbors the PLAU gene, which is expressed in most cell types and encodes for the urokinase type plasminogen activator, a secreted proteolytic enzyme involved in many biological processes, including inflammation and tissue remodeling (3). This enzyme is recruited to the cell membrane by the urokinase plasminogen activator receptor, where it interacts with both extracellular and membrane components (4) to activate pericellular plasmin-mediated proteolysis (3) and to activate several intracellular signaling pathways. It has notably been implicated in the activation and infiltration of lung tissue by T cells and eosinophils (5, 6). Urokinase plasminogen activator and its receptor have been shown to be upregulated in a human cell culture model of asthma after application of compressive stress mimicking bronchospasm as well as in post-mortem lung tissue from patients that died from status asthmaticus (7). Its two main inhibitors, SERPINE1 and SERPINB2, have also been shown to be overexpressed in asthmatic bronchial tissue (8) and a variant of SERPINE1 has been associated with an increased risk for asthma (9).

To the best of our knowledge, the PLAU gene itself has never been studied for association with asthma. It has however been found to be associated with other diseases involving inflammation or extracellular proteolysis. Common PLAU gene variant P141 (rs2227564C) has been associated with increased susceptibility for late-onset

Alzheimer's (10, 11) and colorectal cancer invasion (12) while common 3' UTR variant rs4065C has been associated with a decreased risk for rheumatoid arthritis (13), mitral valve prolapse (14), urolithiasis (15) and Alzheimer's (16).

Combining positional cloning and candidate gene approaches, and based on reported genetic associations with other diseases, we hypothesized that PLAU may function as a regulator of asthma susceptibility. To verify this, we sequenced PLAU and conducted an association study in a French-Canadian family-based sample. We then investigated two additional Canadian trio-based samples for replication purpose. Some of the results of this study have been previously reported in the form of an abstract (*Proc Am Thorac Soc*, 2005; 2:A620) (17).

### 5.3.4 MATERIALS AND METHODS

#### Samples

The original study sample comprises 231 asthmatic families from Saguenay-Lac-St-Jean (SLSJ) (1139 subjects), a region of Northeastern Quebec (Canada) characterized with founder effect (18-22). This sample was described in a recent report (23) and all information concerning recruitment and clinical evaluation of subjects is available in the online supplement. Briefly, asthma phenotype was characterized by the investigators for all participants using a respiratory health questionnaire and function tests following ATS standards (22). Airway hyperresponsiveness (AHR) was defined as a positive challenge to methacholine ( $PC_{20} < 8$  mg/ml) at time of recruitment. If  $PC_{20}$  was not measurable, a 15% augmentation in forced expiratory volume in one second ( $FEV_1$ ) after inhalation of a bronchodilator or a variation in peak expiratory flow (PEF) of at least 12% within a two-week period were also considered diagnostic of AHR (22). Participants were defined as asthmatic if 1) they had a reported history of asthma (validated with physician), or if 2) they showed asthma related symptoms and a positive methacholine provocation test ( $PC_{20}$ ) at the time of recruitment. Subjects were deemed atopic if they had at least one positive response (wheal diameter  $\geq 3$  mm at 10 minutes) on skin prick test (24). Subject with  $PC_{20} > 8$ mg/ml; without history of physician diagnosed asthma and without both symptoms of asthma and  $PC_{20} > 8$ mg/ml; and with no positive response on skin prick test were considered unaffected for AHR, asthma and atopy respectively.

The Canadian Asthma Primary Prevention Study cohort (CAPPS cohort) (25-27) and the Study of Asthma Genes and the Environment birth cohort (SAGE cohort) were used as replication cohorts. The two cohorts are fully described in the online supplement. Briefly, children affected with asthma or related phenotypes were selected along with their parents from these two population asthma cohorts to form study trios for replication analysis. Phenotypes were characterized for all children in each recruiting center, either at time of recruitment (for the SAGE cohort) or at 7 years of age (for the CAPPS cohort), and asthma was diagnosed by a pediatric allergist based on clinical evaluation. Challenge to methacholine was performed for all children in both cohorts. AHR was defined as  $PC_{20} < 6.6$  mg/mL (28, 29). Atopy was defined as at least one positive response on skin prick test (24).

### **PLAU Sequencing**

In order to identify single nucleotide polymorphisms (SNPs) for genetic analyses, fifty unrelated asthmatic probands (25 males and 25 females) were selected from the SLSJ sample (validated using BALSAC population register (19)). PLAU sequence information was obtained from Ensembl v33 (<http://www.ensembl.org/index.html>). The sequenced region spanned 7.6 kb starting from 850bp upstream of the mRNA transcription start site (30) to 430bp downstream of the 3' UTR. The gene was divided in ten regions; the amplification primers, conditions and material are presented in online supplement. SNPs described in this report are cited using their reference sequence number (rs#) from the NCBI database if documented.

### **Genotyping**

Blood samples were drawn from all participants and DNA was extracted from whole blood leukocytes using Genomic-tip 100/G kit (Qiagen Inc.) according to manufacturer's instructions. Genotyping was performed for eight SNPs (Table 3) in the SLSJ sample using either Fluorescence Polarization (Analyst<sup>TM</sup> HT) (31), TaqMan<sup>TM</sup> or Illumina technologies, according to the lab's availability. Four of these SNPs (rs2688607, rs2227562, rs2227564, rs4065) were genotyped in the two additional samples using Illumina technologies. Consult this journal's online supplement for protocols, primers and PCR conditions.

### **Statistical analysis**

Allele distribution patterns and Mendelian errors in the original sample were assessed by FBAT software (1.7 version) (32-34) under additive and dominant genetic models using the empirical variance estimator “-e” for each PLAU SNP and tested phenotypes (32). Hardy-Weinberg equilibrium (HWE) for each SNP and strength of linkage disequilibrium between pairs of SNPs, measured as D prime (D’) (35), were done using Haploview software (3.31 version) (36). Regions of strongly associated markers were inferred using the definition proposed by Gabriel et al. (37) implemented in Haploview. Based on the linkage disequilibrium block of variants, haplotypes were inferred and assessed for nonrandom transmission using the “hbat” command of the FBAT software (38) under additive and dominant models using the empirical variance estimator “-e” (32).

A similar approach was taken for the replication analyses in the CAPPS and SAGE cohorts. Indeed, the theoretical expression of the variance of the FBAT statistic was used (default parameters), under additive and dominant models. Haplotypes were inferred and assessed for both cohorts using the “hbat” command of the FBAT software, under additive and dominant models.

## **5.3.5 RESULTS**

### **Characteristics of the studied samples**

Clinical characteristics of the three studied samples have been reported previously and are described in Tables 1 and 2 (23, 25-27).

### **PLAU sequencing**

All exonic regions were sequenced except for short gaps in exon 4 and in the 3’UTR; most of the intronic regions were also sequenced (Figure 1). As no documented SNPs are known in these gaps and as we judged we had a sufficient SNP coverage, we did not complete the sequence of these gaps. Sequence analysis allowed the identification of 13 polymorphisms. The SNP selection criteria for the association study were based on the allele frequency that had to be greater than 5% as determined by sequencing and on the location within the PLAU gene with priority put on coding regions, regulatory regions (i.e. 5’ and 3’ UTR) and intron / exon junctions.

Seven SNPs were initially selected for PLAU characterization and association testing (presented in Figure 1 in the 5'→3' order): rs1916341, rs2227562, rs2227564, rs2227566, rs2227568, rs2227571 and rs4065. Among the SNPs identified, an A/T substitution located upstream of the PLAU gene was not documented in the dbSNP, this SNP had an allele frequency less than 5% and therefore was not selected for the association study. Of the remaining five discarded SNPs, four were uninformative, with minor allele frequencies <0.025, and one SNP, rs2227560, had a failed genotyping assay. Among the seven selected SNPs were the two previously studied variants rs2227564 (P141L) (10, 12) and rs4065 (13-15). In addition to this initial selection, SNP rs2688607 was genotyped in all three cohorts through the AllerGen gene panel (refer to the online supplement for details on its selection). It is located 7kb upstream of exon 1 (SNP characteristics presented in Table 3).

### **Genotyping**

In the SLSJ sample, genotyping completion rate was >98% for the eight selected SNPs. Mendelian error was present in only two families for three and four SNPs each, suggesting non-paternity or sample error. These families were excluded from further analyses. Hardy-Weinberg equilibrium (HWE) was tested for each SNP with a chi-square test using one degree of freedom. Of the eight SNPs, only rs2227564 gave a significant deviation from HWE ( $p = 0.036$ ); given the number of SNPs tested, this deviation could have occurred by chance alone or as a result of association.

In both replication cohorts, genotyping completion rate was >95% for the four SNPs and there were no Mendelian errors in either the CAPPs or SAGE cohorts. We examined the evidence for departure from HWE separately for the parents and offspring. Significant deviation from Hardy-Weinberg equilibrium is observed in the parental generation of the SAGE cohort for rs4065 ( $p = 0.0031$ ), and in the parental generation of the CAPPs cohort for rs2688607 ( $p = 0.006$ ), and rs2227564 ( $p = 0.027$ ). There is no evidence for departure from HWE in the offspring generation of either cohort. We note that one potential explanation for this finding may be admixture in the parental generation, as both the CAPPs and SAGE cohorts include minority populations. To examine this hypothesis we stratified the cohorts by ethnicity and examined allele frequencies and HWE. We found that for SNPs rs2688607 and rs4065

the major and minor allele depends upon population ethnicity as the most frequent allele in Caucasians is the minor allele in the Asian population of these samples (see Table E3 in online supplement). When stratified by ethnicity the parental generation of the CAPPS cohort is in HWE for the Asian population and there is only minor evidence for departure for rs2227564 ( $p = 0.04$ ) in the Caucasian population. For the SAGE cohort when testing for HWE stratified by ethnicity there is no evidence for departure from HWE in any ethnicity. This suggests that the departure from HWE observed in the combined analyses of these cohorts is due to population stratification. One of the benefits of the trio design is protection against type 1 error due to population stratification. However, our data indicates that caution should be exercised when examining this locus in case control designs.

#### **Family-based association analyses of PLAU SNPs in the SLSJ sample**

The eight PLAU SNPs were tested individually for association with asthma, atopy and AHR under additive and dominant genetic models (Tables 4 and 5). The FBAT software calculates the number of time a given allele is transmitted to affected offspring (Statistic S) and compares it to the number of time it should be transmitted under the null hypothesis of transmission equilibrium (Expected Statistic E(S)). The Z-score is the measure of the deviation of S from E(S) and thus a measure of transmission disequilibrium (32). Under the additive model, statistical analysis results (Table 4) show one suggestive association for rs2227566T with atopy ( $p = 0.041$ ). Under the dominant model (Table 5), the common alleles rs2688607C, rs2227564C (P141) and rs2227566T are overtransmitted to asthmatic probands ( $p = 0.035$ ,  $0.011$  and  $0.045$ , respectively) and to probands presenting AHR ( $p = 0.062$ ,  $0.026$  and  $0.038$ , respectively), but not significantly to atopic probands.

#### **PLAU linkage disequilibrium and haplotype specific association analyses**

Figure 2 shows Haploview linkage disequilibrium plot forming a unique block indicating that SNPs are likely to be transmitted together. Within this region of high linkage disequilibrium, three haplotypes (H8 (1), H8 (2) and H8 (3)) with frequencies higher than 5% were observed (Table 6) and assessed for nonrandom transmission in the SLSJ familial sample (Table 7). Frequent haplotype H8 (1) was shown to be overtransmitted to probands with asthma, atopy and AHR under additive model ( $p =$



0.024, 0.056 and 0.015, respectively) and with asthma, atopy and AHR in dominant genetic model ( $p = 0.031$ ,  $p = 0.043$  and  $0.006$ , respectively).

#### **Family-based association analysis of PLAU SNPs in the CAPPS and SAGE cohorts**

Four PLAU SNPs were tested individually for association with asthma, atopy and AHR under additive and dominant genetic models. No significant findings were found in the CAPPS cohort under both genetic models (see Tables E4 and E5 in online supplement). Analyses were however suggestive of a negative association between the rs2227562A and rs4065C alleles and AHR ( $p$  values between 0.06 and 0.12). In the SAGE cohort, a significant association between atopy and the rs4065C allele has been found. Under the additive model (Table 8) the C allele was undertransmitted to children with atopy ( $p$  value = 0.005), and a similar pattern has been observed for the dominant genetic model with the C allele being undertransmitted ( $p$  value = 0.0001) (Table 9). No association with asthma or airway hyperresponsiveness has been observed in the SAGE cohort.

Haplotype analyses were also carried out with the four PLAU SNPs (Haploview linkage disequilibrium plot presented in the online supplement Figure E1). In the CAPPS cohort, the common H4 (1) haplotype was overtransmitted to children with AHR ( $p = 0.061$ , additive model) while the minor H4 (3) haplotype was significantly undertransmitted to the same children ( $p = 0.051$  and  $0.045$ , additive and dominant) (Table E6 in online supplement). In the SAGE cohort, the common H4 (1) haplotype was overtransmitted to children with atopy ( $p = 0.010$ , additive model) (see Table E7 in online supplement). These findings are consistent with the H8 haplotype analysis from the original SLSJ familial sample (Table 7). When using the same four SNPs for comparison however, common haplotype H4 (1) fails to reach significance in the SLSJ sample (Table E8 in online supplement). This suggests that the additional SNPs are necessary to better delineate the functional variant in the original SLSJ sample.

Given their similar structures, the CAPPS and SAGE samples were combined for further analyses (Tables E9 and E10 in online supplement). The results showed a negative association between allele rs2227562A and asthma ( $p = 0.046$ , additive; and  $p = 0.028$ , dominant), as well as between allele rs4065 and both asthma and atopy ( $p = 0.058$  and  $p = 0.0002$ , dominant). Haplotype analysis showed common haplotype H4 (1) to be overtransmitted to probands affected with atopy ( $p = 0.006$ , additive), while

minor haplotype H4 (3) was under transmitted to probands with asthma and AHR ( $p = 0.049$  and  $p = 0.050$ ; dominant).

### 5.3.6 DISCUSSION

Using family-based samples, we observed associations between allelic variants of novel candidate gene PLAU and asthma phenotypes in three independent Canadian population samples. In the original French-Canadian familial sample (231 families, for a total of 1139 subjects), the common alleles of three SNPs, rs2688607, rs2227564C (P141) and rs2227566T, were associated with asthma and AHR. Additional studies were carried out in two other Canadian cohorts (CAPPS, 238 trios; SAGE, 237 trios), where common allele rs4065C exhibited a highly significant association with protection from atopy. These single marker association results were further supported by haplotype association analysis that identified significant non-random distribution of marker haplotypes covering 12 kb the PLAU gene.

These results are in accordance with current literature suggesting a role for urokinase plasminogen activator in asthma pathogenesis. At the cellular level, it is thought to increase pericellular proteolysis which in turn releases chemotactic mediators and favors adhesion and migration of eosinophils, fibroblasts and smooth muscle cells through the extracellular matrix in lung tissue (3-5, 39, 40). By activating its membrane receptor, it also triggers several intracellular pathways (39, 41), which induce rearrangement of cytoskeleton and cell movement (40, 41), as well as activation of inflammatory cells (42) and T cell priming (43). Knock-out mice for PLAU fail to generate appropriate response to *Cryptococcus neoformans* infection and to schistosomal antigen challenge, suggesting that PLAU is essential for the development of both type 1 and type 2 immune responses (44, 45).

Previous studies have also shown the rs2227564C (P141) and rs4065C alleles to be respectively associated with susceptibility and protection for other diseases involving inflammation and tissue remodeling (10-15). However, only the rs2227564 (P141L) variant has been studied for its functional impact. This C/T mutation codes for the nonsynonymous change of a proline (P) for a leucine (L) at amino acid position 141, located in the kringle domain of the urokinase protein (46). The kringle structure has

the affinity for heparin, plasminogen and extracellular matrix components, such as proteoglycans, and is necessary for binding to the urokinase receptor (12, 47). The L141 allele has been shown to decrease the affinity for fibrin and possibly for other extracellular matrix components, by enhancing the hydrophobicity of the kringle structure (48).

Not much documentation is available for the other associated variants. As SNPs rs2688607 and rs2227566 fall in the same haplotype bloc as rs2227564 (P141L) ( $D' = 0.92$  and  $0.94$ ), it is likely that these SNPs act as markers of P141L rather than as causal mutations. The rs4065 variant is located in the 3' untranslated region (UTR) of PLAU mRNA, which contains multiple instability-determining regions (49). Although its functional impact has not been studied, it could be hypothesized that the rs4065 variant would modulate PLAU expression at a post-transcriptional level either by directly modifying the stability of the 3'UTR or by altering its affinity for mRNA-stabilizing factors (50). It is also possible that rs4065 would be linked to another functional SNP located in a portion of 3'UTR not covered by gene sequencing. More functional work is needed to delineate the real impact of these SNPs in asthma pathogenesis.

Populations with founder effect, such as Saguenay–Lac-Saint-Jean, are important in multigenic disease mapping because they exhibit a lower number of risk genes for a given disease and thus allow the identification of susceptibility loci that would have otherwise been concealed by other high risk loci (51). Neale and Sham have suggested that the gene, rather than specific variants or haplotypes, should be considered as the unit of replication (52). In a recent review, Ober and Hoffjan support this definition of replication because there are now many examples of established associations with different functional variants within the same gene in different populations (53). Therefore, although association results were not directly replicated in the SAGE cohort, the association of the rs4065C variant with protection from atopy does support the association of PLAU with asthma phenotypes and provides evidence that it is not a unique feature of Saguenay–Lac-Saint-Jean. More work will however be needed to better judge the association's importance amongst, and interaction with, other genetic and environmental factors.

In conclusion, we found that common alleles rs2227564C (P141) and rs2227566C were associated with asthma and AHR in a Northeastern Quebec sample. We also found a common PLAU haplotype to be associated with susceptibility to asthma, AHR and atopy. We showed these findings were consistent with data from two other Canadian cohorts, where allele rs4065C could be associated with protection from atopy. This further supports the potential role of the plasminogen activating pathway in the pathogenesis of asthma.

### 5.3.7 ACKNOWLEDGMENTS

We thank all families for their enthusiastic participation in this study. The authors would like to thank Diane Gagné, Muriel Grenon and Dr Paul Bégin for their invaluable participation in the ascertainment of the subjects. P. Bégin and K. Tremblay are supported by the Allergy, Genes and Environment Network (AllerGen). T.J. Hudson is recipient of an Investigator Award from the Canadian Institutes of Health Research and a Clinician-scientist Award in Translational Research from the Burroughs Wellcome Fund. C. Laprise is the chairholder of the Canada Research Chair on genetic determinants in asthma and the director of the Genetics platform of the Respiratory Health Network (RHN) of the Fonds de la recherche en santé du Québec (FRSQ), which financially supports the SLSJ sample. D. Daley is supported by grants from the Canadian Institutes of Health Research (CIHR) Institutes of Gender and Health, Genetics, Population and Public Health, and the IMPACT fellowship. D. Daley is also supported by the Lung Association of British Columbia. The AllerGen Network financially supported genotyping.

### 5.3.8 REFERENCES

1. Laprise C, Lemire M, Sammak A, Verner A, Hudson TJ. Asthma Genome Scan in a French-Canadian Founder Population [abstract]. *Am J Respir Crit Care Med* 2003;169:A628.
2. Choudhry S, Mei R, Tsai HJ, Matsuzaki H, Tang H, Loi H, Matallana H, Dong S, Ung N, Narzarrio S, Casal J, Marcus G, Walsh S, Shriver M, Jones K, Rodriguez W, Ziv E, Burchard EG. Genome-Wide Association Study of Severe Asthma among Puerto Ricans [abstract]. *Proc Am Thorac Soc* 2005;2:A31.

3. Irigoyen JP, Munoz-Canoves P, Montero L, Koziczak M, Nagamine Y. The plasminogen activator system: biology and regulation. *Cell Mol Life Sci* 1999;56(1-2):104–132.
4. Waltz DA, Fujita RM, Yang X, Natkin L, Zhuo S, Gerard CJ, Rosenberg S, Chapman HA. Nonproteolytic role for the urokinase receptor in cellular migration in vivo. *Am J Respir Cell Mol Biol* 2000;22(3):316-322.
5. Brooks AM, Bates ME, Vrtis RF, Jarjour NN, Bertics PJ, Sedgwick JB. Urokinase-type Plasminogen Activator Modulates Airway Eosinophil Adhesion in Asthma. *Am J Respir Cell Mol Biol* 2006;35:503-511.
6. Gyetko MR, Sud S, Sonstein J, Polak T, Sud A, Curtis JL. Antigen-driven lymphocyte recruitment to the lung is diminished in the absence of urokinase-type plasminogen activator (uPA) receptor, but is independent of uPA. *J Immunol* 2001;167(10):5539–5542.
7. Chu EK, Foley JD, Cheng J, Drazen JM, Tschumperlin DJ. Mechanical Regulation of the Urokinase System as a Potential Mediator of Remodeling in a Human Cell Culture of Asthma [abstract]. *Proc Am Thorac Soc* 2005;2:A517.
8. Laprise C, Sladek R, Ponton A, Bernier MC, Hudson TJ, Laviolette M. Functional classes of bronchial mucosa genes that are differentially expressed in asthma. *BMC Genomics* 2004;5(1):21-30.
9. Buckova D, Izakovicova Holla L, Vacha J. Polymorphism 4G/5G in the plasminogen activator inhibitor-1 (PAI-1) gene is associated with IgE-mediated allergic diseases and asthma in the Czech population. *Allergy* 2002;57(5):446–448.
10. Finckh U, van Hadeln K, Muller-Thomsen T, Alberici A, Binetti G, Hock C, Nitsch RM, Stoppe G, Reiss J, Gal A. Association of late-onset Alzheimer disease with a genotype of PLAU, the gene encoding urokinase-type plasminogen activator on chromosome 10q22.2. *Neurogenetics* 2003;4(4):213–217.
11. Riemenschneider M, Konta L, Friedrich P, Schwarz S, Taddei K, Neff F, Padovani A, Kolsch H, Laws SM, Klopp N, Bickeboller H, Wagenpfeil S, Mueller JC, Rosenberger A, Diehl-Schmid J, Archetti S, Lautenschlager N, Borroni B, Muller U, Illig T, Heun R, Egensperger R, Schlegel J, Forstl H, Martins RN, Kurz A. A functional polymorphism within plasminogen activator urokinase (PLAU) is associated with Alzheimer's disease. *Hum Mol Genet* 2006;15(16):2446-2456.

12. Przybylowska K, Smolarczyk K, Kulig A, Romanowicz-Makowska H, Dziki A, Ulanska J, Pander B, Blasiak J. Antigen levels of the urokinase-type plasminogen activator and its gene polymorphisms in colorectal cancer. *Cancer Lett* 2002;181(1):23-30.
13. Huang CM, Chen CL, Tsai JJ, Tsai CH, Tsai FJ. Association between urokinase gene 3'-UTR T/C polymorphism and Chinese patients with rheumatoid arthritis in Taiwan. *Clin Exp Rheumatol* 2004;22(2):219-222.
14. Chou HT, Chen YT, Wu JY, Tsai FJ. Association between urokinase-plasminogen activator gene T4065C polymorphism and risk of mitral valve prolapse. *Int J Cardiol* 2004;96(2):165-170.
15. Tsai FJ, Lin CC, Lu HF, Chen HY, Chen WC. Urokinase gene 3'-UTR T/C polymorphism is associated with urolithiasis. *Urology* 2002;59(3):458-461.
16. Ozturk A, Minster RL, DeKosky ST, Kamboh MI. Association of tagSNPs in the urokinase-plasminogen activator (PLAU) gene with Alzheimer's disease and associated quantitative traits. *Am J Med Genet B Neuropsychiatr Genet* 2007;144(1):79-82.
17. Begin P, Provost V, Page N, Hudson TJ, Laviolette M, Laprise C. Association between the PLAU Gene and Asthma in a French Canadian Population [abstract]. *Proc Am Thorac Soc* 2005;2:A620.
18. Laitinen T, Daly MJ, Rioux JD, Kauppi P, Laprise C, Petays T, Green T, Cargill M, Haahtela T, Lander ES, Laitinen LA, Hudson TJ, Kere J. A susceptibility locus for asthma-related traits on chromosome 7 revealed by genome-wide scan in a founder population. *Nat Genet* 2001;28(1):87-91.
19. Heyer E, Tremblay M. Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 1995;56(4):970-978.
20. Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, Weissenbach J, Popowska E, Pronicka E, Root AW, Glorieux FH. Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am J Hum Genet* 1996;59(3):633-643.
21. Sriver CR. Human genetics: lessons from Quebec populations. *Annu Rev Genomics Hum Genet* 2001;2:69-101.

22. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, November 1986. *Am Rev Respir Dis* 1987;136(1):225-244.
23. Poon AH, Laprise C, Lemire M, Montpetit A, Sinnott D, Schurr E, Hudson TJ. Association of vitamin D receptor genetic variants with susceptibility to asthma and atopy. *Am J Respir Crit Care Med* 2004;170(9):967-973.
24. Skin tests used in type I allergy testing Position paper. Sub-Committee on Skin Tests of the European Academy of Allergology and Clinical Immunology. *Allergy* 1989;44 Suppl 10:1-59.
25. Chan-Yeung M, Manfreda J, Dimich-Ward H, Ferguson A, Watson W, Becker A. A randomized controlled study on the effectiveness of a multifaceted intervention program in the primary prevention of asthma in high-risk infants. *Arch Pediatr Adolesc Med* 2000;154(7):657-663.
26. Chan-Yeung M, Ferguson A, Watson W, Dimich-Ward H, Rousseau R, Lilley M, Dybuncio A, Becker A. The Canadian Childhood Asthma Primary Prevention Study: outcomes at 7 years of age. *J Allergy Clin Immunol* 2005;116(1):49-55.
27. Becker A, Watson W, Ferguson A, Dimich-Ward H, Chan-Yeung M. The Canadian asthma primary prevention study: outcomes at 2 years of age. *J Allergy Clin Immunol* 2004;113(4):650-656.
28. Godfrey S. Bronchial hyper-responsiveness in children. *Paediatr Respir Rev* 2000;1(2):148-155.
29. Godfrey S, Springer C, Bar-Yishay E, Avital A. Cut-off points defining normal and asthmatic bronchial reactivity to exercise and inhalation challenges in children and young adults. *Eur Respir J* 1999;14(3):659-668.
30. Riccio A, Grimaldi G, Verde P, Sebastio G, Boast S, Blasi F. The human urokinase-plasminogen activator gene and its promoter. *Nucleic Acids Res* 1985;13(8):2759-2771.
31. Kwok PY. SNP genotyping with fluorescence polarization detection. *Hum Mutat* 2002;Apr;19(4):315-323.
32. Lake SL, Blacker D, Laird NM. Family-based tests of association in the presence of linkage. *Am J Hum Genet* 2000;67(6):1515-1525.

33. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM. Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet Epidemiol* 2004;26(1):61-69.
34. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. *Genet Epidemiol* 2000;19 Suppl 1:S36-S42.
35. Lewontin RC. On measures of gametic disequilibrium. *Genetics* 1988;120(3):849-852.
36. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21(2):263-265.
37. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. *Science* 2002;296(5576):2225-2229.
38. Laird N, Horvath S, Xu X. Implementing a unified approach to family based tests of association. *Genetic Epi* 2000;19(Suppl 1):S36-S42.
39. Dumler I, Weis A, Mayboroda OA, Maasch C, Jerke U, Haller H, Gulba DC. The Jak/Stat pathway and urokinase receptor signaling in human aortic vascular smooth muscle cells. *J Biol Chem* 1998;273(1):315-321.
40. Bernstein AM, Greenberg RS, Taliana L, Masur SK. Urokinase Anchors uPAR to the Actin Cytoskeleton. *Invest Ophthalmol Vis Sci* 2004;45(9):2967-2977.
41. Aguirre Ghiso JA, Kovalski K, Ossowski L. Tumor dormancy induced by downregulation of urokinase receptor in human carcinoma involves integrin and MAPK signaling. *J Cell Biol* 1999;147(1):89-104.
42. Abraham E, Gyetko MR, Kuhn K, Arcaroli J, Strassheim D, Park JS, Shetty S, Idell S. Urokinase-type plasminogen activator potentiates lipopolysaccharide-induced neutrophil activation. *J Immunol* 2003;170(11):5644-5651.
43. Gyetko MR, Libre EA, Fuller JA, Chen GH, Toews G. Urokinase is required for T lymphocyte proliferation and activation in vitro. *J Lab Clin Med* 1999;133(3):274-288.
44. Gyetko MR, Sud S, Chen GH, Fuller JA, Chensue SW, Toews GB. Urokinase-type plasminogen activator is required for the generation of a type 1 immune response to pulmonary *Cryptococcus neoformans* infection. *J Immunol* 2002;168(2):801-809.



45. Gyetko MR, Sud S, Chensue SW. Urokinase-deficient mice fail to generate a type 2 immune response following schistosomal antigen challenge. *Infect Immun* 2004;72(1):461-467.
46. Hansen AP, Petros AM, Meadows RP, Nettesheim DG, Mazar AP, Olejniczak ET, Xu RX, Pederson TM, Henkin J, Fesik SW. Solution structure of the amino-terminal fragment of urokinase-type plasminogen activator. *Biochemistry* 1994;33(16):4847-4864.
47. Bdeir K, Kuo A, Sachais BS, Rux AH, Bdeir Y, Mazar A, Higazi AA, Cines DB. The kringle stabilizes urokinase binding to the urokinase receptor. *Blood* 2003;102(10):3600-3608.
48. Yoshimoto M, Ushiyama Y, Sakai M, Tamaki S, Hara H, Takahashi K, Sawasaki Y, Hanada K. Characterization of single chain urokinase-type plasminogen activator with a novel amino-acid substitution in the kringle structure. *Biochim Biophys Acta* 1996;1293(1):83-89.
49. Nanbu R, Menoud PA, Nagamine Y. Multiple instability-regulating sites in the 3' untranslated region of the urokinase-type plasminogen activator mRNA. *Mol Cell Biol* 1994;14(7):4920-4928.
50. Tran H, Maurer F, Nagamine Y. Stabilization of urokinase and urokinase receptor mRNAs by HuR is linked to its cytoplasmic accumulation induced by activated mitogen-activated protein kinase-activated protein kinase 2. *Mol Cell Biol* 2003;23(20):7177-7188.
51. Malerba G, Pignatti PF. A review of asthma genetics: gene expression studies and recent candidates. *J Appl Genet* 2005;46(1):93-104.
52. Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. *Am J Hum Genet* 2004;75(3):353-362.
53. Ober C, Hoffjan S. Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 2006;7(2):95-100.

### 5.3.9 FIGURE LEGENDS

**Figure 1:** PLAU gene sequenced region and identified SNPs scaled location

Schematic presentation of the PLAU genomic organization and localization of identified SNPs. Exons are depicted as black boxes with the corresponding exon numbers on top. Introns and non-coding regions are depicted as lines between exon

boxes. Horizontal bold lines indicate the gaps that were not sequenced. Black diamonds and triangles indicate the genotyping technology used. An “x” indicates that the SNP has not been genotyped (see text). Image source: Hapmap ([www.hapmap.org](http://www.hapmap.org)) Jul 06 freeze; modified according to study design.

**Figure 2:** Pairwise linkage disequilibrium pattern of PLAU

The location of each tested SNP along the chromosome is indicated on top. The number in each diamond indicates the magnitude of linkage disequilibrium ( $D'$ ) between respective pairs of SNPs. For example, the pairwise magnitude of linkage disequilibrium for variants rs2227571C and rs4065C is 0.98. Squares without numbers written on them represent perfect linkage disequilibrium ( $D' = 1.0$ ).

**5.3.10 Table 1**

Clinical characteristics of the SLSJ familial sample subjects

	Families		
	Probands (n = 223)	Affected relatives (n = 355)	Unaffected relatives (n = 581)
Male : Female ratio	1 : 1.2	1 : 1.4	1 : 1.2
Mean age (years (range))	18 (3-46)	40 (2-83)	48 (3-96)
Median age (years)	16	41	48
Smoking status (number (%))			
Never	186 (84)	176 (51)	243 (43)
Ex-smoker	12 (5)	105 (30)	199 (35)
Smoker	25 (11)	66 (19)	127 (22)
FEV <sub>1</sub> % of predicted value (SD) *	92.2 (16.3)	88.9 (22.7)	98.9 (17.0)
PC <sub>20</sub> in mg/ml (SD) †	2.66 (3.33)	3.36 (4.53)	26.91 (3.04)
Serum IgE in µg/L (SD) †	229.1 (4.6)	157.4 (4.6)	80.9 (3.7)
Number of persons with subphenotypes (%)			
Asthma ‡	223 (100)	347 (100)	0 (0)
Atopy §	182 (82)	237 (68)	218 (38)
AHR **	188/188 (100)	217/264 (82)	64/486 (13)

\* FEV<sub>1</sub> = Forced expiratory volume in one second† PC<sub>20</sub> = concentration of methacholine inducing a 20 % fall in FEV<sub>1</sub>, IgE = Immunoglobulin E; geometric mean

‡ Present asthma or past documented clinical history of asthma

§ Defined as at least one positive response (wheal diameter ≥ 3 mm at 10 minutes) on skin prick tests

\*\* AHR = Airway hyperresponsiveness at time of recruitment as defined in method section

**5.3.11 Table 2**

Clinical characteristics of the CAPPS and SAGE cohorts subjects

	CAPPS cohort		SAGE cohort	
	Probands * (n = 238)	All Children † (n = 380)	Probands * (n = 237)	All Children (n = 723)
Male : Female ratio	1.31 : 1	1.05 : 1	1.66 : 1	1.26 : 1
Ethnicity (number (%))				
Caucasian	191 (80)	301 (79)	185 (78)	525 (73)
Noncaucasian	47 (20)	79 (21)	52 (22)	198 (27)
PC <sub>20</sub> in mg/ml ‡	1.899	2.987	1.660	2.718
Number of children with subphenotypes (%)				
Asthma §	55 (23)	71 (19)	126 (53)	247 (34)
Atopy **	130 (55)	147/365 (40)	177 (75)	326/719 (45)
AHR ††	205 (86)	257/348 (74)	144 (61)	401/461 (87)

\* Children Affected with Asthma, Atopy, or AHR tested for PLAU genotypes

† Of the 549 children born between 1994 and 1997, median year 1995, 380 were evaluated at 7 yr follow-up

‡ PC<sub>20</sub> = concentration of methacholine inducing a 20 % fall in FEV1

§ Doctor diagnosed asthma.

\*\* Defined as at least one positive response (wheal diameter ≥ 3 mm at 10 minutes) on skin prick tests

†† AHR = Airway hyperresponsiveness; Defined as a PC<sub>20</sub> methacholine ≤ 6.6 mg/ml

**5.3.12 Table 3**  
Genotyped SNPs characteristics

SNP	Base change *	MAF †	Chromosomal location ‡	Genomic features
rs2688607	C/T	0.22	75333742	7kb upstream of Exon 1
rs1916341	C/A	0.39	75341168	Intron
rs2227562	G/A	0.18	75342967	Intron
rs2227564	C/T	0.20	75343107	Exon 5 (nonsynonymous P141L)
rs2227566	T/C	0.41	75343737	Intron
rs2227568	C/T	0.18	75343885	Exon 7 (synonymous N274N)
rs2227571	T/C	0.41	75344746	Intron
rs4065	T/C	0.41	75346470	3'UTR

\* Major/minor (with forward orientation)

† MAF = minor allele frequency as observed in the original SLSJ familial sample

‡ According to NCBI public database build 36.1 ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov))

**5.3.13 Table 4**

Family-based association test results for asthma and related phenotypes under the additive model for the SLSJ familial sample

Phenotype	SNP	Allele	S *	E(S) †	Z ‡	P value
Asthma	rs2688607	C	234	228	0.797	NS §
		T	88	94		
	rs1916341	C	210	209	0.152	NS
		A	156	157		
	rs227562	G	217	216	0.174	NS
		A	75	76		
	rs2227564 (P141L)	C	231	221	1.45	NS
		T	83	93		
	rs2227566	T	216	201	1.9	0.057
		C	118	132		
	rs2227568	C	222	217	0.711	NS
		T	76	81		
rs2227571	T	208	197	1.44	NS	
	C	132	143			
rs4065	T	222	219	0.362	NS	
	C	158	161			
Atopy	rs2688607	C	234	230	0.438	NS
		T	92	96		
	rs1916341	C	210	208	0.194	NS
		A	160	161		
	rs227562	G	179	175	0.566	NS
		A	63	66		
	rs2227564 (P141L)	C	220	214	0.704	NS
		T	84	90		
	rs2227566	T	214	198	2.043	<b>0.041</b>
		C	118	134		
	rs2227568	C	188	180	1.134	NS
		T	64	71		
rs2227571	T	188	177	1.481	NS	
	C	122	132			
rs4065	T	193	191	0.257	NS	
	C	143	145			
AHR	rs2688607	C	230	225	0.651	NS
		T	92	97		
	rs1916341	C	193	195	0.205	NS
		A	155	153		
	rs227562	G	195	195	0.027	NS
		A	73	73		
	rs2227564 (P141L)	C	223	215	1.102	NS
		T	87	95		
	rs2227566	T	190	178	1.773	0.076
		C	112	124		
	rs2227568	C	192	188	0.606	NS
		T	70	74		
rs2227571	T	195	187	1.185	NS	
	C	131	139			
rs4065	T	199	198	0.114	NS	
	C	151	152			

\* S = Association test statistic

† E(S) = Expected statistic

‡ Z = Z-score

§ NS = not significant (p value > 0.05)

**5.3.14 Table 5**

Family-based association test results for asthma and related phenotypes under the dominant model for the SLSJ familial sample

Phenotype	SNP	Allele	S *	E(S) †	Z ‡	P value
Asthma	rs2688607	C	55	49	2.105	<b>0.035</b>
	rs1916341	C	100	95	0.924	NS §
	rs227562	A	72	74	-0.411	NS
	rs2227564 (P141L)	C	56	49	2.54	<b>0.011</b>
	rs2227566	T	112	102	2.01	<b>0.045</b>
	rs2227568	C	41	38	1.117	NS
	rs2227571	T	86	82	0.815	NS
	rs4065	C	106	110	-0.330	NS
Atopy	rs2688607	C	50	47	0.871	NS
	rs1916341	C	87	82	0.802	NS
	rs227562	A	67	71	-0.632	NS
	rs2227564 (P141L)	C	45	42	0.856	NS
	rs2227566	T	95	86	1.901	0.057
	rs2227568	C	43	39	1.245	NS
	rs2227571	T	72	69	0.661	NS
	rs4065	C	91	92	-0.208	NS
AHR	rs2688607	C	49	44	1.863	0.062
	rs1916341	C	88	83	1.023	NS
	rs227562	A	68	70	-0.349	NS
	rs2227564 (P141L)	C	49	43	2.592	<b>0.026</b>
	rs2227566	T	99	90	2.068	<b>0.038</b>
	rs2227568	C	42	41	0.483	NS
	rs2227571	T	77	73	1.011	NS
	rs4065	C	98	98	0.076	NS

\* S = Association test statistic

† E(S) = Expected statistic

‡ Z = Z-score

§ NS = not significant (p value > 0.05)

**5.3.15 Table 6**

Common haplotypes found in the linkage disequilibrium block listed with their frequency

<b>Common 8-SNP Haplotypes and Frequencies for the original SLSJ familial sample</b>									
Haplotype	rs2688607	rs1916341	rs2227562	rs2227564	rs2227566	rs2227568	rs2227571	rs4065	Frequency
H8 (1)	C	C	G	C	T	C	T	T	0.553
H8 (2)	T	A	G	T	C	C	C	C	0.185
H8 (3)	C	A	A	C	C	T	C	C	0.155

<b>Common 4-SNP Haplotypes and Frequencies for all the cohorts</b>								
Haplotype	rs2688607	rs2227562	rs2227564	rs4065	Frequency			
					SLSJ	CAPPS	SAGE	
H4 (1)	C	G	C	T	0.595	0.514	0.541	
H4 (2)	T	G	T	C	0.192	0.227	0.200	
H4 (2)	C	A	C	C	0.173	0.178	0.179	



**5.3.16 Table 7**

Family-based association test results for the PLAU haplotype transmission pattern under the additive and dominant models for the SLSJ familial sample

Phenotype	Haplotype *	Additive				Dominant			
		S †	E(S) ‡	Z §	P value	S	E(S)	Z	P value
Asthma	H8 (1)	181	167	2.254	<b>0.024</b>	132	124	2.151	<b>0.031</b>
	H8 (2)	56	58	-0.342	NS **	53	54	-0.219	NS
	H8 (3)	56	58	-0.536	NS	53	54	-0.299	NS
Atopy	H8 (1)	160	148	1.907	0.056	115	107	2.022	<b>0.043</b>
	H8 (2)	42	45	-0.628	NS	40	42	-0.472	NS
	H8 (3)	51	53	-0.356	NS	48	48	0.024	NS
AHR	H8 (1)	166	152	2.430	<b>0.015</b>	124	115	2.702	<b>0.006</b>
	H8 (2)	53	55	-0.505	NS	50	51	-0.259	NS
	H8 (3)	49	50	-0.262	NS	46	46	0.014	NS

\* Refer to Table 6 for haplotype description

† S = Association test statistic

‡ E(S) = Expected statistic

§ Z = Z-score

\*\* NS = non significant (p value > 0.05)

**5.3.17 Table 8**

Family-based association test results for asthma and related phenotypes under the additive model for the SAGE cohort

Phenotype	SNP	Allele	S *	E(S) †	Z ‡	P value
Asthma	rs2688607	C	94	99.5	1.153	NS §
		T	62	56.5		
	rs2227562	G	90	83	1.673	0.094
		A	32	39		
	rs2227564 (P141L)	C	95	100	1.054	NS
		T	59	54		
rs4065	C	78	82.5	0.862	NS	
	T	94	89.5			
Atopy	rs2688607	C	130	123	1.278	NS
		T	64	71		
	rs2227562	G	130	125.5	0.87	NS
		A	56	60.5		
	rs2227564 (P141L)	C	134	127.5	1.172	NS
		T	64	70.5		
rs4065	C	93	110	2.758	<b>0.0058</b>	
	T	147	130			
AHR	rs2688607	C	103	105	0.202	NS
		T	79	82.5		
	rs2227562	G	107	105.5	0.325	NS
		A	45	46.5		
	rs2227564 (P141L)	C	106	109	0.606	NS
		T	58	55		
rs4065	C	79	82.5	0.642	NS	
	T	111	107.5			

\* S = Association test statistic

† E(S) = Expected statistic

‡ Z = Z-score

§ NS = not significant (p value > 0.05)

**5.3.18 Table 9**

Family-based association test results for asthma and related phenotypes under the dominant model for the SAGE cohort

Phenotype	SNP	Allele	S *	E(S) †	Z ‡	P value
Asthma	rs2688607	C	15	15.25	-0.11	NS §
	rs2227562	A	24	30.75	-1.824	0.068
	rs2227564	C	12	14.25	-1.039	NS
	rs4065	C	29	34.75	-1.591	NS
Atopy	rs2688607	C	25	22.75	0.847	NS
	rs2227562	A	39	46.5	-1.651	0.098
	rs2227564	C	24	22.5	0.577	NS
	rs4065	C	34	51	-3.85	<b>0.0011</b>
AHR	rs2688607	C	19	17.25	0.750	NS
	rs2227562	A	37	38.25	-0.299	NS
	rs2227564	C	16	15	0.471	NS
	rs4065	C	41	42	-0.246	NS

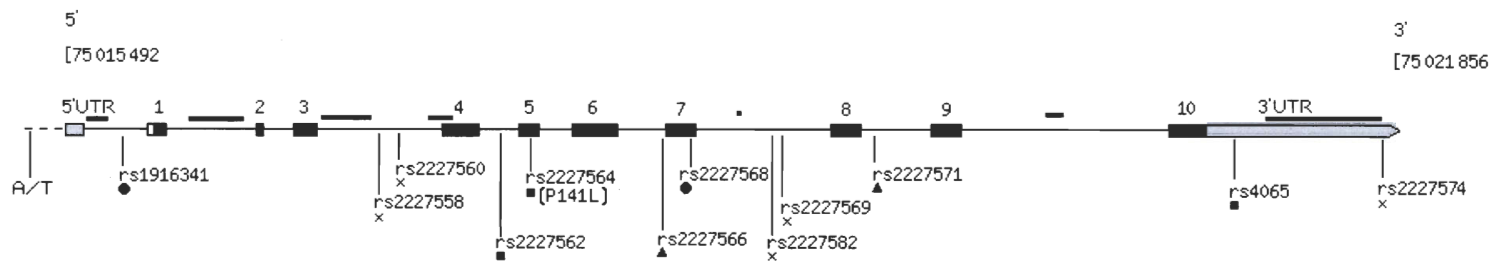
\* S = Association test statistic

† E(S) = Expected statistic

‡ Z = Z-score

§ NS = not significant (p value > 0.05)

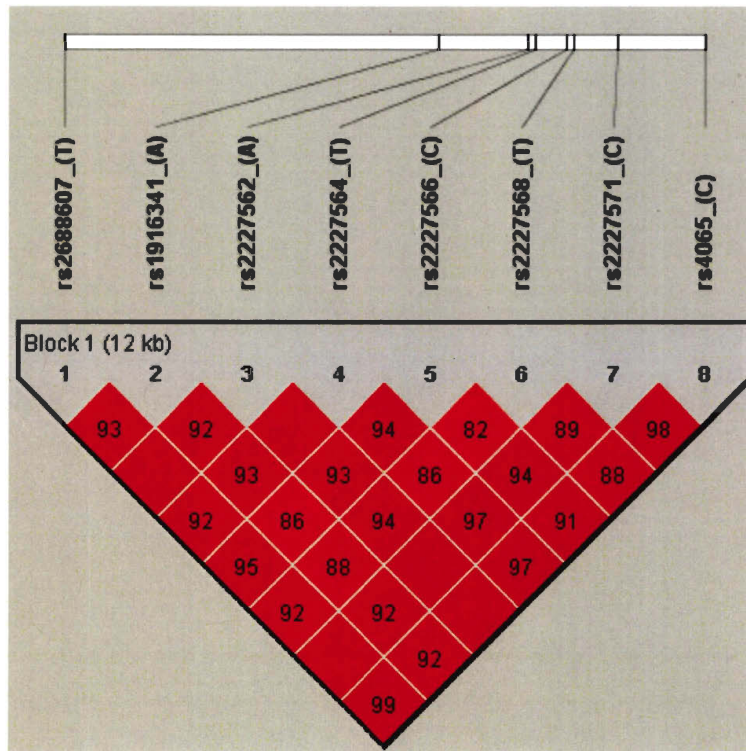
5.3.19 Figure 1



Legend:

- TaqMan genotyping, ▲ FP genotyping, × not genotyped, ■ exon, □ UTR, — intron, — not sequenced, --- 5' upstream region
- Illumina genotyping

5.3.20 Figure 2



### 5.3.21 SUPPLEMENTARY MATERIAL

#### **Association of Urokinase-type plasminogen activator with asthma**

Philippe Bégin, Karine Tremblay, Denise Daley, Mathieu Lemire, Sébastien Claveau, Charleen Salesse, Sabine Kacel, Alexandre Montpetit, Allan Becker, Anita L Kozyrskyj, Thomas J Hudson and Catherine Laprise

#### **SUBJECTS AND METHODS**

##### **The Saguenay–Lac-Saint-Jean (SLSJ) familial sample**

The Saguenay–Lac-Saint-Jean (SLSJ) region of Northeastern Quebec, Canada, is inhabited by 287,000 individuals who descend predominantly from approximately 2500 founders originating from France that settled in Quebec in the 17<sup>th</sup> century (E1). The population grew at a high rate with little admixture over 12 to 14 generations, and is an example of a young founder population (E2). Recruitment occurred through media advertisement (newspaper and radio) and from the Chicoutimi Hospital's specialized clinics (pediatric and respiratory). In order to be included, probands were required to fulfill at least two of the following criteria: 1) a minimum of three clinic visits for acute asthma within one year; 2) two or more asthma-related hospital admissions within one year; or 3) steroid dependency, as defined by either six months of oral, or one year of inhaled corticosteroid use. Families were included in the study if parents (mother and father) were available for phenotypic assessment, if at least one parent was unaffected, and if all four grandparents were of French-Canadian origin. Family participation rate was about 60%. A total of 223 independent families with family size ranging from 3 to 17 and the number of affected family members ranging from 1 to 10 were analyzed.

The affection status of all study participants was determined by clinical evaluation and by the completion of a standard respiratory questionnaire that was modified to include questions about asthma and atopy severity, family history of asthma and/or atopy, age-of-onset and the presence of other respiratory diseases (E3). In 41 cases, the age-of-onset described by parents was below two years; because of the uncertainty of this information, we used a default class of two years. Measurements of expiratory flow were done with a Morgan spirometer (Morgan Spiro 232, P.K. Morgan Ltd) according

to the American Thoracic Society's recommendations (E4). Bronchodilator response (BD) was measured as the increase in FEV<sub>1</sub> at 15 minutes after a 200 mg dose of salbutamol was inhaled. Peak expiratory flow rates (PEFRs) were measured in the morning and evening over a period of two weeks using a mini-Wright peak-flow meter (Armstrong Medical). The best of three repeated measurements was recorded on a diary card. Skin-prick tests were performed for 24 common airborne allergens, which were divided into six main categories: 1) cat, dog, horse, cow and feathers; 2) dust; 3) *D.farinae* and *D.Pterionisus*; 4) grass, weed, Ambrosia, Tymothen, ryegrass; 5) tree mix, birch, maple, oak, elm; 6) *Chladosporium*, *Hormodendrum*, *Alternaria Alternata*, *Alternaria Tenuis*, *Aspergillus* and *Penicillium*. Serum IgE was measured with enzyme immunofluorometry (QuantiCLONE total IgE kit, Kallestad Diagnostics). Methacholine based bronchoprovocation tests were performed for participants 12 years and older according to the method described by Juniper and colleagues (E5). The dose of methacholine that resulted in a 20% fall in FEV<sub>1</sub> was recorded as the PC<sub>20</sub>. All the tests were performed by the investigators for every participant (except for bronchoprovocation) at the University of Montreal Community Genomic Medicine Centre in Saguenay (Qc), Canada. Medical charts were also reviewed for any previous function tests and medication. Airway hyperresponsiveness (AHR) was defined as a PC<sub>20</sub> < 8 mg/ml. If PC<sub>20</sub> was unavailable, AHR was defined as either a rise in the post-bronchodilator FEV<sub>1</sub> of > 15% from baseline or a variation of the morning-evening peak expiratory flow rate (PEFR) of > 12% over a two-week period. Participants were defined as asthmatic if 1) they had a reported history of asthma (validated with physician), or if 2) they showed asthma related symptoms and a positive methacholine provocation test (PC<sub>20</sub>). Subjects were deemed atopic if they had at least one positive response (wheal diameter ≥ 3 mm at 10 minutes) on skin prick test. The Chicoutimi University Hospital local ethics committee approved the study and all subjects gave their informed consent.

#### **The Canadian Asthma Primary Prevention Study cohort (CAPPS cohort)**

The Canadian Asthma Primary Prevention Study was initiated in 1995 to assess the effectiveness of a multifaceted intervention program applied in the first twelve months of life on the primary prevention of asthma and other atopic disorders in high-risk infants (E6). Infants at high risk for developing asthma and other allergic disorders

were identified and their mothers were recruited during their third trimester of pregnancy. Infants at high risk were defined as those who had a parent with asthma or two first-degree relatives with allergies or atopic disorders. The study had two recruitment centers in Canada (Vancouver and Winnipeg). The diagnosis of asthma at the seven year time point (E7) was made by a pediatric allergist in each center without knowledge of the group allocation status of the children and who did not provide health care services to the families. The physician conducted a structured interview with parents by using a standardized form to record symptoms and physical findings. Spirometry and methacholine challenge testing were performed after obtaining parental consent. The diagnoses of asthma and other atopic disorders were clinical decisions made by the pediatric allergists without knowledge of the results of the questionnaire, allergy skin tests, spirometry, or methacholine challenge tests. Atopy was defined as at least one positive response (wheal diameter  $\geq 3$  mm at 10 minutes) on skin prick tests (refer to main SLSJ familial sample description for complete list of allergens). Allergy skin tests were performed with the epicutaneous method by using a prick lanceter (Hollister-Stier, Omega Laboratories LTD, Montreal, Quebec, Canada) with the following allergens (Hollister-Stier, Omega Laboratories LTD): house dust mite (*Dermatophagoides pteronyssinus* and *Dermatophagoides farinae*), cat, dog, cockroach, *Alternaria* species, *Cladosporium* species, tree, grass and weed (ragweed) pollens, cow's milk, egg white, wheat, soy, and peanut. Histamine (1 mg/mL) was used as the positive control, and saline was used as the negative control. The largest wheal diameter and its perpendicular were measured at 15 minutes after testing. A mean wheal diameter of 3 mm or greater than that elicited by the negative control was considered a positive reaction. Atopy was defined as a positive skin test reaction to one or more of these common allergens.

Methacholine (Methapharm Inc, Brantford, Ontario, Canada) challenge testing was carried out according to the protocol of Cockcroft et al. (E8). Two children who had FEV<sub>1</sub> values of less than 70% of predicted value were excluded from methacholine testing, but both of them had a postbronchodilator change in FEV<sub>1</sub> of 12% or more, and they were considered to have airway hyperresponsiveness (AHR). The provocative concentration of methacholine that induced a 20% decrease in FEV<sub>1</sub> from postsaline value (PC<sub>20</sub>) was determined. AHR was defined as a PC<sub>20</sub> of less than 6.6 mg/mL methacholine (E9, E10). The percentages of children whose FEV<sub>1</sub> values decreased by



20% at 2, 4, 6, and 7.8 mg/mL methacholine were also determined. The Ethics Committees of the University of British Columbia and the University of Manitoba approved the study and parents provided written informed consent for participation.

### **The Study of Asthma Genes and the Environment Cohort (SAGE)**

The Study of Asthma Genes and the Environment (SAGE), is derived from a population-based cohort of 16,320 children born in province of Manitoba, Canada, between January 1, 1995 through December 31, 1995. A survey on child health and home environment exposures was mailed to the parents of this birth cohort in 2002. Following parent response to the mailed survey and approval to be contacted, a subset of children was invited to join the study at age 8-10 years. This included all children with parent-declared asthma and a comparable number of children without asthma, randomly selected following stratification of the received surveys by urban / rural location and family income. All recruited children underwent clinical assessment for asthma by a pediatric allergist (AB or JL) on the basis of the Canadian Asthma Consensus Guidelines. To aid in the diagnosis, a standardized history was used including questions on cough, wheeze, shortness of breath, response to current medications (i.e. bronchodilators, corticosteroids) and the presence of other allergic conditions (e.g. allergic rhinitis, atopic dermatitis, and food allergies) were elicited. Physical examination included examination for chest symptoms (hyperinflation, wheeze, prolonged expiration and decreased breath sounds). Hospitalization and medical visits for breathing difficulty in the past year were also noted. All findings were recorded and a diagnosis of asthma in the last 12 months was made. In total 723 study trios (child and both parents) were recruited, 247 of which have a child affected with asthma. A child affected with asthma was defined as follows: pediatric allergist diagnosis of asthma based upon history and the physical exam. Airway hyperresponsiveness (AHR) was assessed in all children following a methacholine challenge test (Methapharm Inc, Brantford, Ontario Canada), carried out according to the protocol of Crockcroft et al. (E8). AHR was defined as a PC<sub>20</sub> (dose of methacholine that resulted in a 20% fall in FEV<sub>1</sub>) of less than 6.6 mg/ml (E9, E10). This definition is based on recent literature supporting the use of this more specific cut-off value when testing AHR in children (E9, E10). The children with FEV<sub>1</sub> < 70% had reversibility testing with an aerosol bronchodilator. Reversibility was defined as an improvement in FEV<sub>1</sub> by > 12% following inhalation of a beta2-adrenergic agonist.

Bronchoprovocation was not attempted for these children. Atopic status was determined for all children; skin prick tests to common relevant allergens were performed (Tree pollen mix, Weed pollen mix, Ragweed (separate), Grass pollen mix, Alternaria, Cladosporium, Penicillium, House dust mites (*D. pteronyssinus*, *D. farinae*), Cockroach, Cat, Dog, Feathers and Peanut) and atopy was defined as at least one positive response (wheal diameter  $\geq 3$  mm at 10 minutes). For children in Winnipeg and immediate surroundings the assessment was carried out in the investigator's laboratories in the Manitoba Institute of Child Health. For children in remote and northern communities a team travelled to the communities to assess children in and surrounding those communities. The study was approved by the Ethics Committee at the University of Manitoba and parents provided written informed consent for participation.

### **PLAU Sequencing**

The gene was divided in ten regions amplified with the following conditions: 95°C 15 min followed by 35 cycles of 94°C 30 sec, 59-66°C 40 sec and 72°C 1 min, then 72°C 5 min. Oligonucleotides and annealing temperatures are listed in Table E1. PCR were done in 25 $\mu$ L volume containing of 0.6 unit of Taq DNA polymerase (New England Biolabs), 50 $\mu$ M of each dNTPs, 20 mM Tris-HCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM KCl, 2 mM MgSO<sub>4</sub>, 0.1 % Triton, pH 8.8, 0.4 $\mu$ M of each primer and 10 ng of human genomic DNA. Amplification products were purified with multiscreen PCR plates (Millipore Corporation), sequenced with BigDye terminator v3.1 chemistry following instructions of the manufacturer and analyzed on a 3100 Genetic analyzer (Applied Biosystems). Sequence analysis was performed with Codoncode Aligner software (CodonCode Corporation).

### **Genotyping**

#### *Template-directed dye-terminator incorporation assay with fluorescence polarization detection (FP-TDI)*

Single nucleotide polymorphisms (SNPs) rs2227566 and rs2227571 were genotyped with the FP-TDI technique, which distinguishes the polymorphic base of a SNP by the template-directed incorporation of a dye-labeled dideoxynucleotide onto an oligonucleotide primer that anneals just 5' to the polymorphic base (E11). The FP-TDI assays are accomplished in four steps according to the method described by Kwok (E11). Briefly, the first step involves PCR amplification of the genomic DNA regions

containing the SNPs. The forward and reverse primers were designed according to the human PLAU gene sequence (NM\_002658) using Primer 3 software ([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)) (see Table E2 for oligonucleotides used designed as “Primer OUT”). In the second step, the PCR products are cleaned by digesting the excess primers and deoxynucleotides enzymatically with exonuclease I and shrimp alkaline phosphatase. The third step was primer extension containing DNA polymerase, SNP-specific primer, and dye-terminators. The sequences of forward and reverse TDI probes were also designed using Primer 3 software (see Table E2 for oligonucleotides used designed as “Primer IN”). Finally, the SNP analysis was done using the Analyst HT (LJL Biosystems, Sunnyvale, CA). An automatically created genotype cluster chart and output worksheet was obtained by plotting the TAMRA polarization vs. R110 polarization in dimensionless units (mP) as previously described (E11, E12). Full technical details are available on request.

#### *TaqMan® SNP Genotyping Assays*

SNPs rs1916341 and rs2227568 (part number C\_26165619\_10 and C\_11458608\_10, respectively) were done with TaqMan® SNP Genotyping Assays multiplex platform (Applied Biosystems) using the Rotor Gene (Corbett Research). PCR was carried out using 15 ng of genomic DNA, 0.2 units of AmpliTaq Gold (Applied Biosystems), 0.5 µl of 10x PCR Gold Buffer (Applied Biosystems), 3.5 mM of MgCl<sub>2</sub>, 0.25 mM of dNTPs, 0.15 µl of 20X TaqMan® SNP Genotyping Assay mix and water for a final volume of 5 µl with the following cycling conditions: 95°C 10 min, 45 cycles of 92°C 15 sec and 58°C 1 min. Genotypes were attributed by the Rotor Gene software using the scatter graph analysis option.

#### *AllerGen SNPs selection and Illumina genotyping assays*

Genotype information for SNPs rs2688607, rs2227562, rs2227564 and rs4065 was provided by the Allergen gene panel (AllerGen Network of Excellence, Theme I: Genes & Early life determinants (<http://www.genapha.ca>) for the original and both replication cohorts. SNP genotyping was performed using the Illumina BeadLab platform as described previously (E13) with oligonucleotides sequences available on the AllerGen website. These SNPs were initially selected from the International HapMap Project release 16c using 10 kb upstream and downstream of the gene. Tags were selected using

the program Tagger using a  $r^2$  cutoff of 0.8. If equivalent tags were identified, those with the highest Illumina design score were kept. Only SNPs with a design score above 0.4 were retained for the panel. Samples with call rate below 95% were excluded from the final dataset. Also, SNPs with poor cluster separation, with a call rate below 90% and with more than two Mendelian errors were removed from the final dataset.

## REFERENCES

- E1. Heyer E, Tremblay M. Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 1995;56(4):970-978.
- E2. Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, Weissenbach J, Popowska E, Pronicka E, Root AW, Glorieux FH. Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am J Hum Genet* 1996;59(3):633-643.
- E3. American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. *Am Rev Respir Dis* 1987;136 (1): 225-244.
- E4. American Thoracic Society. ATS statement--Snowbird workshop on standardization of spirometry. *Am Rev Respir Dis* 1979;119:831-838.
- E5. American Thoracic Society. Standardization of Spirometry, 1994 Update. *Am J Respir Crit Care Med* 1995;152: 1107-1136.
- E6. Chan-Yeung M, Manfreda J, Dimich-Ward H, Ferguson A, Watson W, Becker A. A randomized controlled study on the effectiveness of a multifaceted intervention program in the primary prevention of asthma in high-risk infants. *Arch Pediatr Adolesc Med* 2000;154(7):657-663.
- E7. Chan-Yeung M, Ferguson A, Watson W, Dimich-Ward H, Rousseau R, Lilley M, Dybuncio A, Becker A. The Canadian Childhood Asthma Primary Prevention Study: outcomes at 7 years of age. *J Allergy Clin Immunol* 2005;116(1):49-55.
- E8. Crockcroft DW, Murdock KY, Berscheid BA. Relationship between atopy and bronchial responsiveness to histamine in a random population. *Ann Allergy* 1984;53:26-29.

- E9. Godfrey S. Bronchial hyper-responsiveness in children. *Paediatr Respir Rev* 2000;1(2):148-155.
- E10. Godfrey S, Springer C, Bar-Yishay E, Avital A. Cut-off points defining normal and asthmatic bronchial reactivity to exercise and inhalation challenges in children and young adults. *Eur Respir J* 1999;14(3):659-668.
- E11. Kwok PY. SNP genotyping with fluorescence polarization detection. *Hum Mutat* 2002;19(4):315-323.
- E12. PerkinElmer Life Sciences. 2000. AcycloPrime™-FP SNP Detection Kit. PerkinElmer, Inc., Boston.
- E13. Lincoln MR, Montpetit A, Cader MZ, Saarela J, Dymont DA, Tiislar M, et al. A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. *Nat Genet* 2005;37(10):1108-1112.

#### FIGURE LEGENDS

**Figure E1:** Pairwise linkage disequilibrium pattern of PLAU.

The location of each tested SNP along the chromosome is indicated on top. The number in each diamond indicates the magnitude of linkage disequilibrium ( $D'$ ) between respective pairs of SNPs. Squares without numbers written on them represent perfect linkage disequilibrium ( $D' = 1.0$ ). A) Haplotype blocks for CAPPS cohort. B) Haplotype blocks for SAGE cohort.

**Table E1**

Oligonucleotides and annealing temperatures used for amplification of sequenced fragments

Oligo name	Covered region	Product size	Annealing temperature	Sequence
PLAU-1F	5' upstream	837	59°C	GTCCTCCAGTCTCCCAATTC
PLAU-1R				CGCAACGCTCACAAAGAT
PLAU-2FB	5'UTR, exon 1	1023	62,5°C	ACAAGCCTCTCGATTCTCA
PLAU-2RB				ACTCTCCCTCTCCCTCTCG
PLAU-3FB	Exons 2, 3	973	60°C	AGCTTGCTTGGGTCAATCC
PLAU-3RB				AGTTGGAGTGCAGTGGTGTG
PLAU-4F	Exons 4, 5	998	66°C	GACCCTGTCCAGTCTTCTGG
PLAU-4R				GGAGGCAGGTAGGAGAAAGC
PLAU-5F	Exons 5, 6, 7	1046	62,5°C	TTGTCCAAGAGTGCATGGTG
PLAU-5R				CCAAGCCATATACCCACAGG
PLAU-6F	Exons 8, 9	1037	66°C	GGAGTGGCAGGGAAGAGTTC
PLAU-6R				GGTGGGAAATCAGCTTCACA
PLAU-7FB	Exon 9	693	66°C	GGCCTTCCCTGGTAGAGATAC
PLAU-7RB				ATATGCCAGCCTCCTTCTCA
PLAU-8F	Exon 10, 3'UTR	1100	66°C	GCAGAGGGAAGAAAGGAAGG
PLAU-8R				CACCCATGCACAGGAGATG
PLAU-9F	Exon 10, 3'UTR	994	66°C	GCAGTCACACCAAGGAAGAGA
PLAU-9R				CCATACAAGCACCAGGGAGA
PLAU-10FB	3'UTR, 3' downstream	668	62,5°C	GTTTCATCCAATCCTCACTGG
PLAU-10RB				TGCTACCACACCTGGCTAAT

**Table E2**  
Oligonucleotides used in FP-TDI genotyping assays

SNP	Primer OUT		Primer IN	
	Sequence 5' - 3'	T (°C)	Sequence 5' - 3'	T (°C)
rs2227566	F – GGGAGGGAAGGAAGAAGTGG R – CCCTGCGTGTGGAGTTAAG	52	F – GCTGTTGATGGGTATCTTCT R – CTTGGGTAATCACTGTGG	54
rs2227571	F – TGTGAGATCACTGGCTTTGG R – CTCTTCATCCATCCCTGGAC	55	F – CCTGAAAATGAGCCCAG R – GCAGTCTTCCCTTGATCAC	54

F = Forward, R = Reverse, T = Annealing temperature

**Table E3**  
Allele frequencies stratified by ethnicity in the CAPPS and SAGE cohorts

SNP	Caucasian *		Asian †		Other ‡	
	Minor	Major	Minor	Major	Minor	Major
<b>CAPPS Cohort</b>						
rs2227562	A (0.17)	G (0.83)	G (0.25)	A (0.75)	G (0.18)	A (0.82)
rs2227564	T (0.23)	C (0.77)	T (0.31)	C (0.67)	T (0.23)	C (0.74)
rs2688607	T (0.25)	C (0.75)	C (0.40)	T (0.60)	T (0.25)	C (0.75)
rs4065	C (0.43)	T (0.57)	T (0.16)	C (0.84)	C (0.25)	T (0.75)
<b>SAGE Cohort</b>						
rs2227562	A (0.17)	G (0.83)	A (0.37)	G (0.63)	A (0.33)	G (0.67)
rs2227564	T (0.24)	C (0.76)	T (0.28)	C (0.72)	T (0.16)	C (0.84)
rs2688607	T (0.24)	C (0.76)	T (0.38)	C (0.62)	T (0.32)	C (0.68)
rs4065	C (0.43)	T (0.57)	T (0.19)	C (0.81)	T (0.30)	C (0.70)

\* Caucasians-child and both parents race were reported as Caucasian

† Asian- child and both parents were reported as Asian

‡ Other-Child has two parents with different ethnicities

Note: Rows highlighted above are SNPs where the major and minor allele depends upon the ethnicity of the population being studied.



**Table E4**

Family-based association test results for asthma and related phenotypes under the additive model for the CAPPS cohort

Phenotype	SNP	Allele	S *	E(S) †	Z ‡	P value
Asthma	rs268807	C	39	37.5	0.48	NS §
		T	25	26.5		
	rs2227562	G	52	48.5	1.093	NS
		A	18	21.5		
	rs2227564 (P141L)	C	36	34.5	0.539	NS
		T	18	19.5		
rs4065	C	39	43.5	1.236	NS	
	T	47	42.5			
Atopy	rs2688607	C	97	93	0.800	NS
		T	63	67		
	rs2227562	G	109	106	0.632	NS
		A	45	48		
	rs2227564 (P141L)	C	93	89	0.883	NS
		T	45	49		
rs4065	C	90	96	1.163	NS	
	T	103	97			
AHR	rs2688607	C	151	147	0.658	NS
		T	85	89		
	rs2227562	G	173	163	1.859	0.06
		A	63	74		
	rs2227564 (P141L)	C	153	150.50	0.427	NS
		T	77	79.50		
rs4065	C	133	144	1.596	0.11	
	T	165	154			

\* S = Association test statistic

† E(S) = Expected statistic

‡ Z = Z-score

§ NS = not significant (p value > 0.05)

**Table E5**

Family-based association test results for asthma and related phenotypes under the dominant model for the CAPPS cohort

Phenotype	SNP	Allele	S *	E(S) †	Z ‡	P value
Asthma	rs268807	C	7	8.75	-1.000	NS §
	rs2227562	A	15	18.5	-1.228	NS
	rs2227564	C	4	5	-0.756	NS
	rs4065	C	13	15.5	-1.031	NS
Atopy	rs2688607	C	22	22.75	-0.268	NS
	rs2227562	A	38.00	40.25	-0.535	NS
	rs2227564	C	12	13.75	-0.831	NS
	rs4065	C	34	38.50	-1.203	NS
AHR	rs2688607	C	30	30	0	NS
	rs2227562	A	54	62.50	-1.632	0.10
	rs2227564	C	20	22	-0.763	NS
	rs4065	C	53	30.25	-1.531	0.12

\* S = Association test statistic

† E(S) = Expected statistic

‡ Z = Z-score

§ NS = not significant (p value > 0.05)

**Table E6**

Family-based association test results for the PLAU haplotype transmission pattern under the additive and dominant models in the CAPPS cohort

Phenotype	Haplotype *	Additive				Dominant			
		S †	E(S) ‡	Z §	P value	S	E(S)	Z	P value
Asthma	H4 (1)	48	43.5	1.236	NS	35	33.5	0.626	NS **
	H4 (2)	21	23	-0.707	NS	17	20	-1.279	NS
	H4 (3)	18	21.5	-1.093	NS	16	19.5	-1.257	NS
Atopy	H4 (1)	112	105	1.257	NS	81	79	0.563	NS
	H4 (2)	52	56	-0.863	NS	43	49	-1.576	NS
	H4 (3)	45	48	-0.632	NS	40	42.5	-0.602	NS
AHR	H4 (1)	184	171	1.867	0.061	138	129	1.434	NS
	H4 (2)	84	87.5	-0.598	NS	68	73.50	-1.153	NS
	H4 (3)	60	71	-1.954	<b>0.051</b>	55	65	-2.00	<b>0.045</b>

\* Refer to Table 6 for haplotype description

† S = Association test statistic

‡ E(S) = Expected statistic

§ Z = Z-score

\*\* NS = non significant (p value > 0.05)

**Table E7**

Family-based association test results for the PLAU haplotype transmission pattern under the additive and dominant models in the SAGE cohort

Phenotype	Haplotype *	Additive				Dominant			
		S †	E(S) ‡	Z §	P value	S	E(S)	Z	P value
Asthma	H4 (1)	123	120	0.543	NS	85	85.25	-0.77	NS **
	H4 (2)	62	57	0.990	NS	54	47.25	1.691	NS
	H4 (3)	38	44.5	-1.501	NS	32	37.5	-1.511	NS
Atopy	H4 (1)	174	158	2.546	<b>0.010</b>	112	111.25	0.213	NS
	H4 (2)	58	63.5	-1.044	NS	51	54.25	-0.747	NS
	H4 (3)	56	61.5	-1.026	NS	46	52.50	-1.463	NS
AHR	H4 (1)	136	134.5	0.275	NS	97	95.25	0.598	NS
	H4 (2)	56	53	0.600	NS	50	44.75	1.285	NS
	H4 (3)	43	45	-0.436	NS	38	39.5	-0.372	NS

\* Refer to Table 6 for haplotype description

† S = Association test statistic

‡ E(S) = Expected statistic

§ Z = Z-score

\*\* NS = non significant (p value > 0.05)

**Table E8**

Family-based association test results for the PLAU 4-SNP haplotype transmission pattern under the additive and dominant models in the SLSJ familial sample

Phenotype	Haplotype *	Additive				Dominant			
		S †	E(S) ‡	Z §	P value	S	E(S)	Z	P value
Asthma	H4 (1)	270	266.5	0.425	NS	198	198.5	-0.107	NS **
	H4 (2)	107	112	-0.716	NS	99	99.7	-0.138	NS
	H4 (3)	95	96	-0.174	NS	84	86.3	-0.389	NS
Atopy	H4 (1)	258	255.5	0.324	NS	190	191	-0.378	NS
	H4 (2)	104	105	-0.136	NS	93	93	-0.012	NS
	H4 (3)	95	98.5	-0.566	NS	84	87.2	-0.599	NS
AHR	H4 (1)	212	207	0.724	NS	159	155.1	0.963	NS
	H4 (2)	82	89.6	-1.243	NS	76	78.7	-0.529	NS
	H4 (3)	69	70.1	-0.216	NS	62	63.5	-0.323	NS

\* Refer to Table 6 for haplotype description

† S = Association test statistic

‡ E(S) = Expected statistic

§ Z = Z-score

\*\* NS = non significant (p value > 0.05)

**Table E9**

Family-based association test results under the additive and dominant models for the combined CAPPS and SAGE cohorts

Phenotype	SNP	Allele	Additive		Dominant	
			Combined Z-score	P value	Combined Z-score	P value
Asthma	rs268807	C	-0.7016	NS	-0.693	NS
	rs2227562	A	-1.993	<b>0.0462</b>	-2.194	<b>0.0281</b>
	rs2227564 (P141L)	C	0.6364	NS	-1.280	NS
	rs4065	C	-1.236	NS	-1.895	<b>0.0579</b>
Atopy	rs2688607	C	-0.695	NS	-1.495	NS
	rs2227562	A	-1.3018	NS	-1.1656	NS
	rs2227564 (P141L)	C	0.4221	NS	-1.313	NS
	rs4065	C	-1.163	NS	-3.714	<b>0.0002</b>
AHR	rs2688607	C	0.3825	NS	0.4546	NS
	rs2227562	A	-1.666	0.096	-1.4605	NS
	rs2227564 (P141L)	C	-0.6523	NS	-0.2965	NS
	rs4065	C	-1.6497	0.099	-1.3221	NS

**Table E10**

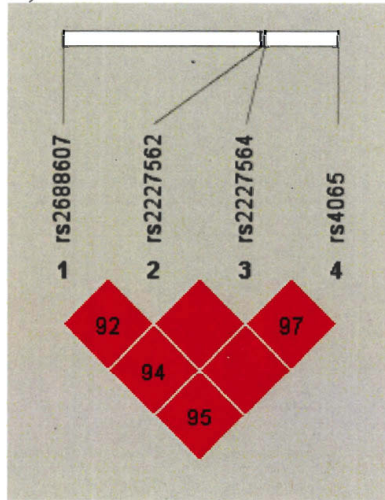
Family-based association test results for the PLAU haplotype transmission pattern under the additive and dominant models in the combined CAPPS and SAGE cohorts

Phenotype	Haplotype *	Additive		Dominant	
		Combined Z-score	P value	Combined Z-score	P value
Asthma	H4 (1)	1.133	NS	0.31068	NS
	H4 (2)	0.518	NS	0.8099	NS
	H4 (3)	-1.856	0.06	-1.9639	<b>0.049</b>
Atopy	H4 (1)	2.739	<b>0.006</b>	0.5563	NS
	H4 (2)	-1.3536	NS	-1.5996	NS
	H4 (3)	-1.1873	NS	-1.4795	NS
AHR	H4 (1)	1.6391	NS	0.3106	NS
	H4 (2)	-0.0649	NS	0.8099	NS
	H4 (3)	-1.7856	0.074	-1.9639	<b>0.0495</b>

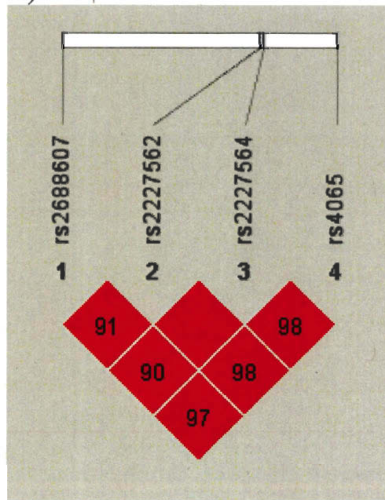
\* Refer to Table 6 for haplotype description

Figure E1

A)



B)





## 5.4 Quatrième article

### **Étude comparative de deux expérimentations de puces à ADN pour définir le patron d'expression génique de l'asthme allergique**

Annie Chamberland, Anne-Marie Madore, Karine Tremblay, Michel Laviolette et Catherine Laprise

*Experimental Lung Research* (octobre 2008) [sous presse]

#### **Résumé**

L'asthme allergique est reconnu comme une maladie à trait complexe. Plusieurs approches ont été employées au cours de la dernière décennie afin d'identifier les biomarqueurs génétiques impliqués dans sa pathogénèse ou sa physiopathologie. La présente étude a pour but de démontrer l'efficacité et la validité des expérimentations de puces à ADN dans la caractérisation des profils d'expression génique de l'asthme. Ainsi, le niveau d'expression des transcrits de biopsies bronchiques a été obtenu à l'aide de deux expérimentations de puces à ADN indépendantes. Ces deux expérimentations ont été réalisées à deux ans d'intervalle avec deux puces différentes (*Affymetrix GeneChip® Human Genome U95Av2* et *U133A*). Toutes deux ont comparé les niveaux d'expression entre un groupe de quatre sujets asthmatiques allergiques et un groupe de quatre sujets témoins, sans asthme ni allergie, qui sont également des sujets différents entre les deux expérimentations. Les puces *U95Av2* et *U133A* ont détecté respectivement 89 et 40 gènes différemment exprimés dans les biopsies des deux groupes de sujets. Cinquante-cinq pourcent des gènes identifiés avec la puce *U133A* l'avaient été dans la précédente puce *U95Av2*. Les gènes des molécules immunes et des enzymes protéolytiques sont les deux voies biologiques les plus conservées entre les deux expérimentations : 75 % des gènes de chacune de ces catégories identifiés avec la puce *U133A* l'avaient été dans la puce *U95Av2*. Ces résultats démontrent l'efficacité des expérimentations de puces à ADN utilisant les biopsies bronchiques de sujets asthmatiques allergiques. La comparaison des résultats de ces deux études de puces suggère que cette technologie est valable pour cibler de nouveaux biomarqueurs et de nouvelles voies biologiques dans l'asthme allergique.

### 5.4.1 TITLE PAGE

#### **A comparison of two sets of microarray experiments to define allergic asthma expression pattern**

Annie Chamberland<sup>1</sup>, Anne-Marie Madore<sup>1,2</sup>, Karine Tremblay<sup>1,2</sup>, Michel Laviolette<sup>2,3</sup>,  
Catherine Laprise<sup>1</sup>§

<sup>1</sup>Université du Québec à Chicoutimi, Département des sciences fondamentales, Saguenay, Canada <sup>2</sup>Université Laval, Département de médecine, Québec, Canada and <sup>3</sup>Centre de recherche de l'Hôpital Laval, Institut universitaire de cardiologie et pneumologie de l'Université Laval, Québec (Qc), Canada

Running head title: Comparison of two sets of microarray experiments

§ Corresponding author: Catherine Laprise  
Université du Québec à Chicoutimi  
555, boulevard de l'Université  
Chicoutimi (Québec), Canada, G7H 2B1  
Telephone: (418) 545-5011, 5659  
Fax: (418) 545-5012

Email addresses:

AC: annie.chamberland1@uqac.ca  
AMM: anne-marie.madore.1@ulaval.ca  
KT: karine.tremblay.19@ulaval.ca  
ML: michel.laviolette@med.ulaval.ca  
CL: catherine\_laprise@uqac.ca

### 5.4.2 ABSTRACT

**Background:** Allergic asthma is a complex trait. Several approaches have been used to identify biomarkers involved in this disease. This study aimed at demonstrating the relevance and the validity of microarrays in the definition of allergic asthma expression pattern. **Methods and Results:** We compared the transcript expressions of bronchial biopsy of two different microarray experiments done 2 years apart both including non-allergic healthy and allergic asthmatic subjects (n = 4 in each experiment). U95Av2 and U133A GeneChips detected respectively 89 and 40 differentially expressed genes. Fifty-five percent of the U133A genes were previously identified with the U95Av2 arrays. The immune signalling molecules and the proteolytic enzymes were the most preserved categories between the two experiments, since 3/4 of the genes identified by the U133A were also significant in the U95Av2 study for both categories. **Conclusion:** These results demonstrate the relevance of microarray experiments using bronchial tissues in allergic asthma. The comparison of these GeneChip studies suggests that earlier microarray results are as relevant as actual ones to target new genes of interest, particularly in function categories linked to the studied disease. Moreover, it demonstrates that microarrays are a valuable technology to target novel allergic asthma pathways as well as biomarkers.

### 5.4.3 BACKGROUND

Allergic asthma is a multifactorial bronchial disease that involves genetic and environmental factors in its development, chronicity and severity [1, 2]. It is characterized by airway hyperresponsiveness, inflammation and remodeling processes [3]. Gene expression profiling studies are increasingly used to identify pathways and novel biomarkers involved in allergic asthma pathogenesis and Hansel and Diette recently published a review regrouping them according to tissues or cell types used [4]. Only two studies reported the use of whole bronchial biopsy specimens [5, 6]. The great diversity of allergic asthma phenotypes, the heterogeneity of bronchial tissue samples, the possible variation of the pathology distribution in the lung add complexity to microarray data interpretation and make comparison between studies difficult [4, 7].

The genes identified by microarrays should not be viewed as an end by themselves. Indeed the results must be validated using low-throughput techniques such as quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR) or northern blots [8]. It can also be validated by a positive association between the selected gene and the studied disease [9-11]. Since these techniques allow validation of a limited number of differentially expressed genes, it is important to know if microarray study results are reproducible and if previous studies are as efficient as recent ones to target genes of interest. A possible way to validate microarray results may be to compare two independent microarray experiments [12]. Based on these assumptions, the present study proposes to compare our previous microarray results [6] with a second microarray experiment performed with different subjects, GeneChips and algorithms. This comparative study attempts to demonstrate the relevance of microarray assays in the definition of allergic asthma expression patterns and the potential and validity of this approach to identify biomarkers involved in the pathophysiology of this disease.

### **5.4.3 METHODS**

#### **Subjects' selection and characteristics**

The two gene expression studies were compared. For each of them, four healthy controls without history of allergy and asthma and four allergic asthmatic subjects according to the ATS criteria were selected [13], the healthy controls and the allergic asthmatic subjects not being the same ones in both experiments. The selection of the subjects and the procedure for bronchoscopy and bronchial biopsies were performed in the second study as they were in the first about two years later [6]. Briefly, the baseline evaluation included a questionnaire on health status and respiratory conditions, a physical examination, skin prick tests using 26 common aeroallergens and the measurements of the FEV<sub>1</sub> values and of the dose of methacholine that resulted in a 20% fall in FEV<sub>1</sub> (PC<sub>20</sub>). For each subject bronchoscopy, eight biopsies evaluated as acceptable by the research assistant were taken from bronchi of the right lung using alligator forceps (Olympus FB-15C-1). Biopsies were obtained from the sub-segmental and segmental bronchial carinae of the lower (3–5 biopsies) and upper (1–2 biopsies) lobes, as well as the lobar carinae of the middle (2 biopsies) and upper (1–2 biopsies) lobes to provide adequate samples of the large airways (3<sup>rd</sup> to 5<sup>th</sup> bronchial generation) [6]. The samples were immediately placed in RNAlater (Ambion) to stabilize

intracellular RNA during transport. The samples were mechanically homogenized (Rotor/Stator homogeniser, PowerGen) and total RNA was extracted using the RNEasy kit (Qiagen) according to the manufacturer's instructions. RNA integrity and yield was assessed using Agilent Bioanalyser RNA LabChips<sup>®</sup>. Approval for the study was obtained from ethics committee at Laval Hospital. All subjects provided informed consent.

The clinical features of subjects included in the two studies were compared with an independent sample t-test (including the Levene test) performed using SPSS 11.5 (SPSS Inc. Chicago, IL). The significant threshold was fixed at 0.05.

### **Microarray studies and analysis**

The previously reported microarray expression study was performed using Affymetrix U95Av2 and the actual study used the U133A GeneChips (Affymetrix, Santa Clara, CA). The U95Av2 contains 12,000 probe sets, which represent approximately 10,000 full-length genes whereas the U133A contains 22,283 probe sets targeting 18,400 transcripts for the study of 14,500 genes. Ninety-seven percent of the U95Av2 probe sets have corresponding matches in U133A GeneChips. A single expression array was hybridized for each RNA sample. Following washing, the specifically bound probe was detected by incubating the arrays with SAPE (streptavidin phycoerythrin, Molecular Probes) and scanning the chips using a Gene Array Scanner (Agilent). Detailed probe synthesis hybridization and washing protocols have been previously described [14].

To calculate gene expression two different algorithms were used, the MAS5 [15] for the U95Av2 data and the RMA [16] for the U133A data. To identify significant differences in gene expression, a pairwise comparison between controls and allergic asthmatic subjects was performed with the microarray results, in the first case with the ArrayStat microarray analysis package, as described in our previous study [6], and in the second case with the "limma" package available in Bioconductor [17]. The comparison of the two assays was made between the new expression differences in U133A and the previously observed ones in U95Av2 [6].

As described in the study of Laprise et al. [6], the U95Av2 data were corrected for multiple testing using the Bonferroni correction (p values < 6.6E-6). However, no

correction was applied when using the U133A GeneChips since the RMA method is more stringent [18, 19]. However, Smyth's moderated t-test was used instead of standard t-test ( $p$  value  $< 0.05$ ). This statistic allows for more stable inference particularly when using small numbers of arrays (between 3 and 5) [20].

#### 5.4.5 RESULTS

Clinical features of the healthy control and allergic asthmatic subjects recruited for both sets of experiments are presented in table 1. Statistics for each studied clinical feature revealed a significant difference between U95Av2 and U133A experiments for the forced expiratory volume in one second ( $FEV_1$ ):  $95.7 \pm 8.8$  and  $102.3 \pm 1.5$  respectively ( $p = 0.004$ ). However, no significant difference was observed for allergic asthmatic subjects. Table 2 presents detailed clinical features of the allergic asthmatic subjects in both studies. All of these subjects were allergic to common indoor and outdoor allergens. Table 3 presents the comparison of methodology used for the novel U133A GeneChips experiments and the U95Av2 experiments [6]. Given the difference of standardization that appeared over time, two different algorithms, the Robust Multi-array Analysis (RMA) [16] and the Microarray Suite 5.0 (MAS5) [15], were respectively used for the two microarray analyses. The RMA algorithm being more stringent, a 1.8-fold-change was used with Smyth's moderated t-test for the U133A microarray experiments instead of a two-fold change with Bonferroni correction for the U95Av2 arrays.

The U95Av2 and U133A GeneChips detected respectively 89 and 40 genes differentially expressed between healthy control and allergic asthmatic subjects. Even if 55% less genes were identified using the U133A RMA method, the U133A GeneChips contained 4500 more genes than the U95Av2 GeneChips. All the results of the U95Av2 arrays are available online [6] and table 4 presents the 40 genes identified by the U133A arrays (see Additional file 1). To help comparison, the table 4 also contains the data of the first study for the genes implicated in present comparative analysis. Nine of these 40 genes obtained in U133A were not available with the U95Av2 microarrays [21]. Of these 31 genes, 17 genes (55%) revealed by the U133A array were also identified by the U95Av2 arrays [6]. All the fold changes in expression values for these genes were in the same direction in both studies. We classified these genes based on the twelve

categories established in our U95Av2 previous study when it was possible. Compared to U95Av2, the U133A GeneChips identified two transcripts in a new category: structural proteins (group 13).

Interestingly, two important categories - due to their implication in the pathophysiology of allergic asthma, the immune signalling molecules and the proteolytic enzymes - showed the most conservative differential expression. For both categories, 3/4 genes identified by the novel microarray experiment were also detected by the previous microarray experiment [6]. Thus, the immune molecules arachidonate 15-lipoxygenase (*ALOX15*), chemokine (C-X3-C motif) receptor 1 (*CX3CR1*) and nitric oxide synthase 2A (*NOS2A*) as well as the proteolytic enzymes cathepsin C (*CTSC*), serpin peptidase inhibitor clade B member 2 (*SERPINB2*) and member 4 (*SERPINB4*) were identified by both arrays (Table 4) [6].

#### 5.4.6 DISCUSSION

The aim of the present report was to demonstrate the reproducibility of microarrays in the definition of allergic asthma expression patterns by comparing two independent experiments using different GeneChips and different subjects selected according to the same clinical criteria [22]. The more stringent U133A microarrays analysis identified less genes differentially expressed in allergic asthmatic subjects compared to control subjects than the U95Av2 GeneChips. But from those genes identified by the U133A, 55% were also detected in the U95Av2 experiment.

Diverse variability factors may have influenced the concordance rate between the two studies. First, an inescapable heterogeneity between the two groups of allergic asthmatic and control subjects could explain some variability [7]. In order to limit the variation, the measured clinical features of the two groups of allergic asthmatics and controls were similar. The difference in FEV<sub>1</sub> observed in the control groups was not clinically significant. Moreover, the difference in gender between the allergic asthmatics of both studies has been considered in the analyses and we eliminated all biomarkers that were related to the gender such as the ones on the X and Y chromosomes. Secondly, the use of bronchial biopsies is a source of variation by itself due to cell heterogeneity of the tissue samples and to the possible difference in the patterns of inflammation and

remodelling present in the tissue bronchial sections [6, 7]. This variation may also come from differences in cellular composition due to differing proportions of epithelial cells, fibroblasts, and smooth muscle cells that may occur in asthmatic and control tissue as a result of biopsy techniques. However, this was minimized, in the two studies, by using a pool of eight biopsies taken from similar subcarina generations using the same biopsy technique. Thirdly, the transcriptome variability may also play an important role depending on the GeneChips used. Together, all these factors influence the material obtained from the biopsies even if the experiment is carefully executed [14]. Consequently, for each experiment, meticulous precautions such as biopsy sampling by the same physician in the same period of the day, tissue processing by the same person, use of the same RNA quantity obtained from the same number of biopsies and microarray hybridizations performed at the same moment, were taken in order to limit the maximum of variability factors. Finally, since the standardization of analyses changed over the years, the two experiments used different algorithms for their normalization [18]. Being more stringent, the RMA algorithm identified less differentially expressed genes than the MAS5.0 algorithm [16, 23], representing an inherent source of variation between compared studies.

Despite a great heterogeneity between both array studies, two allergic asthma relevant functional categories, the immune signalling molecules and proteolytic enzymes, were highly reproduced. Since the two experiments were performed using different subjects, different GeneChips and data analysis, these results strongly suggest that the genes for which the differential expression was detected in both arrays are relevant biomarkers in the pathophysiology of allergic asthma. Among these, some are known to be involved in allergic asthma pathogenesis such as mucin 2 and mucin 5AC, oligomeric mucus/gel-forming (*MUC2* and *MUC5AC* respectively) [24], *NOS2A* [25-28], NAD(P)H dehydrogenase, quinone 1 (*NQO1*) [29], *SERPINB2* [30] and *SERPINB4* [31, 32]. Other molecules such as *ALOX15*, *CTSC*, *CX3CR1*, *IGHG1* and periostin (osteoblast specific factor) (*POSTN*) were identified as newly ones in our previous work [6]. Finally, the last six genes that were differentially expressed in both arrays (*CEACAM5*, *C7*, *IGKC*, *LGALS7*, *S100P* and *TCF21*) do not seem to have any obvious function in allergic asthma pathogenesis. However, the fact that they were detected using the U95Av2 and U133A GeneChips suggests a significant role of their pathway in allergic asthma pathophysiology. For example, *CX3CR1* was identified as a candidate gene by a



published association study between the *CX3CR1* gene and asthma [11] as well as by a recent genetic association study using the immune signaling molecule biomarkers showed up in the first microarray study that allowed finding seven positive associations with asthma and/or related phenotypes like *ALOX15* and *CX3CR1* (JACI, in press).

#### **5.4.7 CONCLUSION**

In conclusion, performing two independent microarray experiments, using the standard methods respective to the time they were carried out, allowed demonstrating that strong biomarkers appear in both studies, the recent study with up-to-date statistical tools and in the previous one. It also demonstrates that the use of p value, fold change cut-off and also functional categories allowed selecting possible genes of interest in earlier microarray studies. Such results reveal the potential and the validity to use microarrays and bronchial biopsies to increase our knowledge on asthma pathogenesis. Thus, the discovery of allergic asthma susceptibility and/or protection genes by both arrays is a convincing demonstration that the microarrays are likely to accelerate the definition of allergic asthma expression pattern, and to improve our understanding of the molecular pathophysiology of this disease.

***Declaration of interest:*** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

#### 5.4.8 ACKNOWLEDGMENTS

The authors thank Luce Trépanier for her very valuable help in recruiting and evaluating the subjects for the study and Marie-Claude Bernier for the RNA extraction from bronchial biopsies. The authors wish to acknowledge Dr André Ponton and his team of the Functional Genomics platform of the McGill University and Génome Québec Innovation Centre for their technical assistance. The authors would like to thank Dr Fernando Martinez for his critical revision of the manuscript. A.C. is supported by the Fondation de l'Université du Québec à Chicoutimi (FUQAC). AMM received a studentship from Canadian Institutes of Health Research (CIHR). K.T. is supported by the Fondation de l'Université Laval. C.L. is the chairholder of the Canada Research Chair on genetic determinants in asthma ([www.chairs.gc.ca](http://www.chairs.gc.ca)).

#### 5.4.9 REFERENCES

1. Cookson WO and Moffatt MF: Genetics of asthma and allergic disease. *Hum Mol Genet.* 2000; 9: 2359-64.
2. Maddox L and Schwartz DA: The pathophysiology of asthma. *Annu Rev Med.* 2002; 53: 477-98.
3. National Institute of Health, Guidelines for the diagnosis and management of asthma. 1997, National Institutes of Health National Heart, Lung, and Blood Institutes: Washington. p. 153.
4. Hansel NN and Diette GB: Gene expression profiling in human asthma. *Proc Am Thorac Soc.* 2007; 4: 32-6.
5. Dolganov GM, Woodruff PG, Novikov AA, Zhang Y, Ferrando RE, Szubin R, and Fahy JV: A novel method of gene transcript profiling in airway biopsy homogenates reveals increased expression of a Na<sup>+</sup>-K<sup>+</sup>-Cl<sup>-</sup> cotransporter (NKCC1) in asthmatic subjects. *Genome Res.* 2001; 11: 1473-83.
6. Laprise C, Sladek R, Ponton A, Bernier MC, Hudson TJ and Laviolette M: Functional classes of bronchial mucosa genes that are differentially expressed in asthma. *BMC Genomics.* 2004; 5: 21.
7. Rolph MS, Sisavanh M, Liu SM and Mackay CR: Clues to asthma pathogenesis from microarray expression studies. *Pharmacol Ther.* 2005.

8. Chuaqui RF, Bonner RF, Best CJ, Gillespie JW, Flaig MJ, Hewitt SM, Phillips JL, Krizman DB, Tangrea MA, Ahram M, Linehan WM, Knezevic V, and Emmert-Buck MR: Post-analysis follow-up and validation of microarray experiments. *Nat Genet.* 2002; 32 Suppl: 509-14.
9. Begin P, Tremblay K, Daley D, Lemire M, Claveau S, Salesse C, Kacel S, Montpetit A, Becker A, Chan-Yeung M, Kozyrskyj AL, Hudson TJ, and Laprise C: Association of urokinase-type plasminogen activator with asthma and atopy. *Am J Respir Crit Care Med.* 2007; 175: 1109-16.
10. Tremblay K, Daley D, Chamberland A, Lemire M, Montpetit A, Laviolette M, Musk AW, James AL, Chan-Yeung M, Becker A, Kozyrskyj AL, Sandford AJ, Hudson TJ, Pare PD, and Laprise C: Genetic variation in immune signaling genes differentially expressed in asthmatic lung tissues. *J Allergy Clin Immunol.* 2008; 122: 529-36 e17.
11. Tremblay K, Lemire M, Provost V, Pastinen T, Renaud Y, Sandford AJ, Laviolette M, Hudson TJ, and Laprise C: Association study between the CX3CR1 gene and asthma. *Genes Immun.* 2006; 7: 632-9.
12. Moreau Y, Aerts S, De Moor B, De Strooper B and Dabrowski M: Comparison and meta-analysis of microarray data: from the bench to the computer desk. *Trends Genet.* 2003; 19: 570-7.
13. American Thoracic Society: Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, November 1986. *Am Rev Respir Dis.* 1987; 136: 225-44.
14. Novak JP, Sladek R and Hudson TJ: Characterization of variability in large-scale gene expression data: implications for study design. *Genomics.* 2002; 79: 104-13.
15. Affymetrix. Microarray Suite User Guide, Version 5. <http://www.affymetrix.com/support/technical/manuals.affx>.
16. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, and Speed TP: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics.* 2003; 4: 249-64.
17. Smyth GK, Limma: linear models for microarray data. In: Springer, Eds. *Bioinformatics and Computational Biology Solutions using R and Bioconductor.* New York, 2005:397-420.

18. Millenaar FF, Okyere J, May ST, van Zanten M, Voeselek LA and Peeters AJ: How to decide? Different methods of calculating gene expression from short oligonucleotide array data will give different results. *BMC Bioinformatics*. 2006; 7: 137.
19. Barash Y, Dehan E, Krupsky M, Franklin W, Geraci M, Friedman N, and Kaminski N: Comparative analysis of algorithms for signal quantitation from oligonucleotide microarrays. *Bioinformatics*. 2004; 20: 839-46.
20. Smyth GK: Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004; 3: Article3.
21. Affymetrix technical note. Current NetAffx Annotation Files. (2007) <https://www.affymetrix.com/support/technical/annotationfilesmain.affx>.
22. Begin P, Tremblay K, Daley D, Lemire M, Claveau S, Salesse C, Kacel S, Montpetit A, Becker A, Kozyrskyj AL, Hudson TJ, and Laprise C: Association of Urokinase-type Plasminogen Activator with Asthma and Atopy. *Am J Respir Crit Care Med*. 2007.
23. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B and Speed TP: Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res*. 2003; 31: e15.
24. Morcillo EJ and Cortijo J: Mucus and MUC in asthma. *Curr Opin Pulm Med*. 2006; 12: 1-6.
25. Batra J, Pratap Singh T, Mabalirajan U, Sinha A, Prasad R and Ghosh B: Association of inducible nitric oxide synthase with asthma severity, total serum immunoglobulin E and blood eosinophil levels. *Thorax*. 2007; 62: 16-22.
26. Hoffjan S, Ostrovnaja I, Nicolae D, Newman DL, Nicolae R, Gangnon R, Steiner L, Walker K, Reynolds R, Greene D, Mirel D, Gern JE, Lemanske RF, Jr., and Ober C: Genetic variation in immunoregulatory pathways and atopic phenotypes in infancy. *J Allergy Clin Immunol*. 2004; 113: 511-8.
27. Eynott PR, Paavolainen N, Groneberg DA, Noble A, Salmon M, Nath P, Leung SY, and Chung KF: Role of nitric oxide in chronic allergen-induced airway cell proliferation and inflammation. *J Pharmacol Exp Ther*. 2003; 304: 22-9.

28. Gao PS, Kawada H, Kasamatsu T, Mao XQ, Roberts MH, Miyamoto Y, Yoshimura M, Saitoh Y, Yasue H, Nakao K, Adra CN, Kun JF, Moro-oka S, Inoko H, Ho LP, Shirakawa T, and Hopkin JM: Variants of NOS1, NOS2, and NOS3 genes in asthmatics. *Biochem Biophys Res Commun.* 2000; 267: 761-3.
29. Yang IA, Fong KM, Zimmerman PV, Holgate ST and Holloway JW: Genetic susceptibility to the respiratory effects of air pollution. *Thorax.* 2008; 63: 555-63.
30. Woodruff PG, Boushey HA, Dolganov GM, Barker CS, Yang YH, Donnelly S, Ellwanger A, Sidhu SS, Dao-Pick TP, Pantoja C, Erle DJ, Yamamoto KR, and Fahy JV: Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci U S A.* 2007; 104: 15858-63.
31. Izuhara K: The role of interleukin-4 and interleukin-13 in the non-immunologic aspects of asthma pathogenesis. *Clin Chem Lab Med.* 2003; 41: 860-4.
32. Yuyama N, Davies DE, Akaiwa M, Matsui K, Hamasaki Y, Suminami Y, Yoshida NL, Maeda M, Pandit A, Lordan JL, Kamogawa Y, Arima K, Nagumo F, Sugimachi M, Berger A, Richards I, Roberds SL, Yamashita T, Kishi F, Kato H, Arai K, Ohshima K, Tadano J, Hamasaki N, Miyatake S, Sugita Y, Holgate ST, and Izuhara K: Analysis of novel disease-related genes in bronchial asthma. *Cytokine.* 2002; 19: 287-96.

**5.4.10 Table 1**

Clinical features of the control and allergic asthmatic subjects

	Control subjects			Allergic asthmatic subjects		
	U95Av2 (n = 4)*	U133A (n = 4)	p value	U95Av2 (n = 4)*	U133A (n = 4)	p value
Age, mean $\pm$ SD, y	22.5 $\pm$ 4.3	26.7 $\pm$ 7.5	0.491	25.7 $\pm$ 8.9	24.2 $\pm$ 3.9	0.205
Sex, (F/M)	2/2	2/2	–	4/0	0/4	–
FVC, mean $\pm$ SD, % <sup>†</sup>	95.5 $\pm$ 11.3	100.7 $\pm$ 3.4	0.201	100.5 $\pm$ 8.7	103.7 $\pm$ 7.5	0.971
FEV <sub>1</sub> , mean $\pm$ SD, % <sup>†</sup>	95.7 $\pm$ 8.8	102.3 $\pm$ 1.5	0.004	98.0 $\pm$ 8.5	87.0 $\pm$ 14.8	0.396
PC <sub>20</sub> , mean $\pm$ SD, mg/ml <sup>‡</sup>	73.7 $\pm$ 1.5	83.2 $\pm$ 2.0	0.282	2.4 $\pm$ 2.1	0.25 $\pm$ 3.9	0.325
Allergy	No	No	–	Yes	Yes	–

\* Data from our previous study [6]

<sup>†</sup> Forced vital capacity (FVC) and forced expiratory volume in one second (FEV<sub>1</sub>) expressed in % of predicted<sup>‡</sup> Concentration of methacholine inducing 20% fall in FEV<sub>1</sub>

**5.4.11 Table 2**

Detailed clinical features of the allergic asthmatic subjects in both studies

<b>Allergic asthmatic subjects in U95Av2*</b>					
Age	Sex	FVC† 1 (%)	FEV <sub>1</sub> † 1 (%)	PC <sub>20</sub> mg/ml‡	Allergy§
(1) 21	F	3.60 (91)	3.25 (96)	6.12	C, D, DPt, Du, G, Rw, T
(2) 20	F	3.50 (99)	2.80 (90)	2.53	C, D, DPt, G, Du, Rw, T
(3) 23	F	3.84 (100)	3.07 (96)	1.01	C, D, G, DPt, T
(4) 39	F	3.96 (112)	3.21 (110)	2.23	C, Du, G, Rw, T
<b>Allergic asthmatic subjects in U133A</b>					
(1) 21	M	4.83 (100)	2.72 (66)	0.22	C, D, Du, DPt, DFa, G, F, M
(2) 21	M	5.88 (115)	4.43 (100)	0.04	G, Rw, T
(3) 29	M	5.10 (100)	3.79 (88)	1.00	Du, DPt, DFa, F
(4) 26	M	4.82 (100)	3.72 (94)	0.43	C, D, Du, G, Rw, H

\* Data from our previous study [6].

† Forced vital capacity (FVC) and forced expiratory volume in one second (FEV<sub>1</sub>) expressed in liter and % of predicted.‡ Concentration of methacholine inducing 20% fall in FEV<sub>1</sub>.§ Skin response read at 10 min if mean diameter  $\geq$  3mm; C: Cat, D: Dog, Du: Dust, DPt: D. Pteronyssinus, DFa: D. Farinae, G: Grass pollens, Rw: Ragweed, T: tree pollens, F: Feather, H: Horse, M: Mold. A wheal response to histamine  $\geq$  3 mm served as a positive control.

**5.4.12 Table 3**  
Comparison between GeneChips experiments

	U95Av2*	U133A
Algorithm	MAS 5.0	RMA
Fold change threshold	2	1.8
Comparison statistic	t-test with Bonferroni correction	Smyth's moderated t-test
Probes sets (n)	12000 †	22283
Genes (n)	10000	14500
Differentially expressed genes (n)	89	40
Biological cluster categories (n)	12	13

\* According to our previous study [6].

† With Bonferonni correction.

‡ Ninety-seven percent of the HG-U95Av2 probe sets have corresponding matches in HG-U133A GeneChips.



### 5.4.13 Additional Table 4

Gene expression differences between control and asthmatic subjects using the U133A GeneChips.

Significant genes from Smyth's moderated t-test implemented in the "limma" package in Bioconductor.

Probe Set U133A / U95Av2	Accession	Gene *†	Cytoband †	Description	Control (C)		Asthma (A)		A vs C		A vs C <sup>  </sup>	
					Mean ‡	SD	Mean ‡	SD	MFC §	p value	MFC §	p value ¶
<b>Group 1: Immune signalling molecules</b>												
207328_at / 34636_at	NM_001140	<b>ALOX15</b>	17p13.3	Arachidonate 15-lipoxygenase	6,22	0,43	8,02	1,09	3,47	0,007	4,17	0,00E+00
205898_at / 40646_at	U20350	<b>CX3CR1</b>	3p21.3	Chemokine (CX3-C motif) receptor 1	7,72	0,18	6,59	0,51	-2,20	0,001	-2,21	1,51E-07
205403_at / 998_s_at	NM_004633	IL1R2	2q12-q22	interleukin 1 receptor, type II	5,59	0,13	6,86	0,98	2,42	0,016	NA	NA
211372_s_at / 998_s_at	U64094	IL1R2	2q12-q22	interleukin 1 receptor, type II	5,22	0,11	6,45	1,07	2,34	0,028	NA	NA
210037_s_at / 1948_f_at	L24553	<b>NOS2A</b>	17q11.2-q12	Nitric oxide synthase 2A (inducible, hepatocytes)	5,14	0,20	6,26	0,63	2,17	0,004	7,25	0,00E+00
<b>Group 2: Extracellular proteins</b>												
204673_at / 896_at	NM_002457	<b>MUC2</b>	11p15.5	Mucin 2, oligomeric mucus/gel-forming	5,85	0,10	6,89	1,02	2,06	0,043	50,98	0,00E+00
217110_s_at / 35912_at	AJ242547	MUC4	3q29	Mucin 4, cell surface associated	6,16	0,34	7,39	0,88	2,36	0,015	NS	NS
204895_x_at / 35912_at	NM_004532	MUC4	3q29	Mucin 4, cell surface associated	7,52	0,29	8,61	0,66	2,13	0,007	NS	NS

217109_at / 35912_at	AJ242547	MUC4	3q29	Mucin 4, cell surface associated	7,58	0,39	8,52	0,50	1,91	0,009	NS	NS
214385_s_at / 41383_at	AI521646	<b>MUC5AC</b>	11p15.5	Mucin 5AC, oligomeric mucus/gel-forming	9,05	0,56	10,49	1,04	2,72	0,020	5,89	0,00E+00
214303_x_at / 298_at	AW192795	<b>MUC5AC</b>	11p15.5	Mucin 5AC, oligomeric mucus/gel-forming	9,24	0,52	10,50	1,04	2,40	0,034	5,89	0,00E+00
204051_s_at / 41405_at	AW089415	SFRP4	7p14.1	Secreted frizzled-related protein 4	8,47	0,13	7,49	0,93	-1,98	0,038	NS	NS

**Group 3: Immune response**

201884_at / 1582_at	NM_004363	<b>CEACAM5</b>	19q13.1- q13.2	Carcinoembryonic antigen-related cell adhesion molecule 5	7,36	0,54	9,98	1,16	6,15	0,001	5,88	0,00E+00
219759_at / -	NM_022350	ERAP2	5q15	Endoplasmic reticulum aminopeptidase 2	6,10	1,07	4,66	0,33	-2,71	0,016	-	-
215118_s_at / 35530_f_at	AW519168	<b>IGHG1</b>	14q32.33	Immunoglobulin heavy constant gamma 1 (G1m marker)	5,32	0,38	4,34	0,24	-1,98	0,001	-2,13	1,53E-08
217022_s_at / 35530_f_at	S55735	<b>IGHG1</b>	14q32.33	Immunoglobulin heavy constant gamma 1 (G1m marker)	12,96	0,11	12,06	0,45	-1,87	0,002	-2,13	1,53E-08
212592_at / 37006_at	AV733266	IGJ	4q21	Immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides	10,88	0,44	9,82	0,66	-2,07	0,014	NS	NS

211645_x_at / 31586_f_at	M85256	<b>IGKC</b>	2p12	Immunoglobulin kappa constant	9,66	0,37	8,39	1,00	-2,42	0,023	-2,13	4,42E-07
217179_x_at / 31347_f_at	X79782	<b>IGL@</b>	22q11.1- q11.2	Immunoglobulin lambda locus	8,25	0,15	7,27	0,79	-1,96	0,022	NS	NS
217227_x_at / 31347_f_at	X93006	<b>IGL@</b>	22q11.1- q11.2	Immunoglobulin lambda locus	7,07	0,23	6,10	0,83	-1,96	0,029	NS	NS
216207_x_at / -	AW408194	<b>IGKV1D-13</b>	2p12	Immunoglobulin kappa variable 1D-13	10,51	0,28	9,57	0,72	-1,92	0,021	-	-

#### Group 4: Intracellular signalling

219250_s_at / -	NM_013281	<b>FLRT3</b>	20p11	fibronectin leucine rich transmembrane protein 3	7,56	0,34	6,69	0,63	-1,83	0,021	-	-
-----------------	-----------	--------------	-------	---	------	------	------	------	-------	-------	---	---

#### Group 5: Proteolytic enzymes

201487_at / 133_at	NM_001814	<b>CTSC</b>	11q14.1-q1	Cathepsin C	9,16	0,22	10,09	0,28	1,90	0,000	2,32	1,03E-08
204614_at / 37185_at	NM_002575	<b>SERPINB2</b>	18q21.3	Serpin peptidase inhibitor, clade B (ovalbumin), member 2	5,26	0,43	7,88	1,84	6,13	0,011	11,37	0,00E+00
211906_s_at / 1549_s_at	AB046400	<b>SERPINB4</b>	18q21.3	Serpin peptidase inhibitor, clade B (ovalbumin), member 4	6,18	0,51	8,53	0,96	5,12	0,001	2,05	3,27E-06
210413_x_at / 1549_s_at	U19557	<b>SERPINB4</b>	18q21.3	Serpin peptidase inhibitor, clade B (ovalbumin), member 4	9,77	0,08	10,65	0,31	1,85	0,000	2,05	3,27E-06
205185_at / 40315_at	NM_006846	<b>SPINK5</b>	5q32	Serine peptidase inhibitor, Kazal type 5	6,90	0,18	7,84	0,72	1,91	0,018	NS	NS

#### Group 6: Transmembrane proteins

209869_at / 264_at	AF284095	<b>ADRA2A</b>	10q24-q26	adrenergic, alpha- 2A-, receptor	6,43	0,31	7,30	0,75	1,83	0,036	NS	NS
--------------------	----------	---------------	-----------	-------------------------------------	------	------	------	------	------	-------	----	----

206165_s_at/ 40356_at	NM_006536	CLCA2	1p31-p22	CLCA family member 2, chloride channel regulator	8,19	0,30	9,16	0,50	1,97	0,005	NS	NS
205328_at/ 39579_at	NM_006984	CLDN10	13q31-q34	Claudin 10	7,21	0,35	8,07	0,30	1,81	0,002	NS	NS
206400_at/ 38608_at	NM_002307	<b>LGALS7</b>	19q13.2	Lectin, galactoside-binding, soluble, 7	6,10	0,46	7,34	0,68	2,35	0,007	2,36	4,62E-08
219564_at/-	NM_018658	KCNJ16	17q23.1-q24.2	Potassium inwardly-rectifying channel, subfamily J, member 16	6,12	0,53	7,58	1,08	2,75	0,021	-	-

---

**Group 7: Free radical metabolism**

---

201467_s_at/ 38066_at	AI039874	<b>NQO1</b>	16q22.1	NAD(P)H dehydrogenase, quinone 1	5,44	0,28	6,32	0,74	1,84	0,031	1,93	6,58E-06
--------------------------	----------	-------------	---------	----------------------------------	------	------	------	------	------	-------	------	----------

---

**Group 8: Gene transcription**

---

204931_at/ 37247_at	NM_003206	<b>TCF21</b>	6pter-qter	Transcription factor 21	6,04	0,41	4,96	0,33	-2,12	0,001	-2,25	2,42E-06
------------------------	-----------	--------------	------------	-------------------------	------	------	------	------	-------	-------	-------	----------

---

**Group 9: Cell growth and proliferation**

---

218730_s_at/-	NM_014057	OGN	9q22.3	Osteoglycin	5,39	0,39	4,43	0,50	-1,95	0,008	-	-
218677_at/-	NM_020672	S100A14	1q21.1	S100 calcium binding protein A14	7,67	0,24	8,68	0,76	2,01	0,018	-	-
204351_at/ 34319_at	NM_005980	<b>S100P</b>	4p16	S100 calcium-binding protein P	9,46	0,36	10,82	0,78	2,57	0,006	3,08	1,07E-12
209260_at/ 33322_i_at	BC000329	SFN	1p36.11	Stratifin	7,05	0,31	8,04	0,81	1,99	0,028	NS	NS
33322_i_at/ 33322_i_at	X57348	SFN	1p36.11	Stratifin	9,94	0,21	10,80	0,69	1,82	0,023	NS	NS

33323_r_at / 33323_r_at	X57348	SFN	1p36.11	Stratifin	9,30	0,22	10,16	0,73	1,81	0,029	NS	NS
<b>Group 10: Cell adhesion molecules</b>												
210809_s_at / 1451_s_at	D13665	POSTN	13q13.3	Periostin, osteoblast specific factor	8,89	0,33	10,85	1,07	3,88	0,003	2,79	2,73E-09
<b>Group 11: Complement components</b>												
202992_at / 37394_at	NM_000587	C7	5p13	Complement component 7	5,90	0,35	4,93	0,24	-1,96	0,001	-2,21	2,74E-06
205382_s_at / 40282_s_at	NM_001928	CFD	19p13.3	Complement factor D (adipsin)	8,44	0,72	7,28	0,38	-2,23	0,010	NS	NS
<b>Group 12: Metabolic enzymes</b>												
219429_at / -	NM_024306	FA2H	16q23	Fatty acid 2- hydroxylase	5,32	0,31	6,19	0,64	1,84	0,020	-	-
209613_s_at / 35730_at	M21692	ADH1B	4q21-q23	Alcohol dehydrogenase 1B (class I), beta polypeptide	8,73	0,29	6,72	0,87	-4,05	0,001	NS	NS
209612_s_at / 35730_at	M24317	ADH1B	4q21-q23	Alcohol dehydrogenase 1B (class I), beta polypeptide	9,51	0,08	8,00	0,93	-2,85	0,005	NS	NS
212240_s_at /1269_at	A1679268	PIK3R1	5q12-q13	Phosphoinositide -3-kinase, regulatory subunit 1 (alpha)	8,93	0,26	7,90	0,53	-2,05	0,003	NS	NS
<b>Group 13: Structural proteins</b>												
218963_s_at / -	NM_015515	KRT23	17q21.2	Keratin 23 (histone deacetylase inducible)	6,68	0,14	7,65	0,50	1,97	0,002	-	-
201645_at / 32818_at	NM_002160	TNC	9q33	Tenascin C	7,74	0,51	8,81	0,50	2,11	0,007	NS	NS

Abbreviations used: MFC = Mean fold change; SD = standard deviation

\* Genes in bold are differently expressed in both platforms (U95Av2 and U133A)

† Gene symbol and location obtained from National Center for Biotechnology Information public database build 128 (<http://www.ncbi.nlm.nih.gov>).

‡ Logarithmic values.

§ Mean fold change are indicated for each probe set significantly 1.8-fold (U133A) or 2-fold (U95Av2) differentially expressed by allergic asthmatic subjects compared with control subjects. Positive data indicate genes which are over-expressed and negative data indicate genes which are under-expressed by allergic asthmatic subjects. Mean fold change are in logarithmic values.

|| For comparison, data from the study of Laprise et al. [6].

¶ According to Bonferroni correction, p values  $<6.6E-6$  are considered significant. p values identified by 0 are very low (  $p <1E-13$  ).

## Chapitre 6

### Les gènes ciblés à l'aide de l'outil bioinformatique *Genes to Diseases (G2D)*

Cette étude est le fruit de la collaboration entre les Drs Catherine Laprise et Miguel Andrade-Navarro afin de valider l'outil G2D pour la première fois afin de prioriser des gènes candidats pour une étude d'association dans le trait complexe qu'est l'asthme. Donc, à partir de deux loci liés à l'asthme (6q26) et à l'atopie (10q26.3) dans des familles du SLSJ, l'outil G2D a été utilisé pour produire une liste initiale de gènes contenus dans ces régions. Après revue de la littérature de la biologie de ces gènes, ceux représentant le meilleur potentiel (totalisant 10 gènes –Tableau III) ont été sélectionnés comme gènes candidats pour une étude d'association avec les phénotypes de l'asthme dans l'échantillon de trios familiaux du SLSJ.

Ce cinquième article (section 6.1) s'inscrit dans le projet de maîtrise de Camille Potvin auquel j'ai significativement collaboré. L'article a été publié dans le journal *PLoS ONE* et rapporte une association génétique pour le gène *PTPRE* avec l'asthme allergique dans l'échantillon du SLSJ. Pour cette étude, j'ai participé à la sélection des gènes candidats et des TagSNPs, effectué les analyses statistiques et l'interprétation des résultats. J'ai également rédigé le manuscrit (sauf les sections G2D) et coordonné la correspondance entre les collaborateurs. Mathieu Lemire a participé aux analyses de l'étude de liaison, validé les analyses d'association et contribué à la rédaction de la section des analyses statistiques du manuscrit. Camille Potvin a contribué à la sélection TagSNPs, au génotypage, au séquençage et à la rédaction du manuscrit. Alexandre Tremblay a participé au génotypage des ADN de l'échantillon du SLSJ. Gary Hunninghake a réalisé les analyses pour l'échantillon CAMP. Benjamin Raby est le chercheur principal de l'échantillon CAMP et a révisé le manuscrit. Thomas Hudson a supervisé l'étude de liaison et révisé le manuscrit. Carolina Perez-Ixatra a participé au design de l'étude, effectué les analyses G2D et contribué à la rédaction des sections G2D du manuscrit. Miguel Andrade-Navarro a participé au design de l'étude, révisé le manuscrit et approuvé sa version finale. Catherine Laprise a conçu l'étude, réalisé l'étude de liaison, dirigé l'ensemble des étapes, révisé le manuscrit et approuvé sa version finale.

**Tableau III**

Sélection des gènes G2D selon le nombre d'apparitions dans les analyses, la fonction biologique ainsi que les études génétiques et fonctionnelles publiées.

Gène	Analyse <sup>a</sup>	Nom officiel <sup>b</sup>	Fonction <sup>c</sup>	Études d'association et fonctionnelles <sup>d</sup>
<i>LPA</i> 6q26	Non	<i>Lipoprotein, Lp(a)</i>	Possède une activité sérine protéinase et est capable d'autoprotéolyse. Inhibe l'activateur du plasminogène de type 1.	Propriétés athérogéniques (Sotiriou, 2006). Cofacteur dans l'hypertension (Antonicelli, 2001). Contrôle génétique potentiel des niveaux de Lp(a) dans le diabète de type 2 (Tian, 2003). Un niveau élevé de Lp(a) est un facteur de risque pour le lupus érythémateux systémique (Frostedgard, 2002). Associée à la maladie d'Alzheimer (Carter, 2007).
<i>NOX3</i> 6q26	Oui	<i>NADPH oxidase 3</i>	Enzyme plasmatique membranaire trouvée dans plusieurs types cellulaires. Catalyse la production des superoxydes.	La sur-régulation de cette enzyme dans les poumons et les cellules endothéliales contribue à la pathologie de l'emphysème (Zhang, 2006).
<i>SNX9</i> 6q26	Non	<i>Sorting nexin 9</i>	Les membres de cette famille de protéines contiennent un domaine phox (PX), un domaine de liaison phosphoinositide, impliqué dans plusieurs étapes de la signalisation intracellulaire.	Rôle dans la co-signalisation des cellules T (Badour et al., 2007).
<i>VIL2</i> 6q26	Non	<i>Villin 2 (ezrin)</i>	Protéine cytoplasmique située près de la membrane et qui peut agir comme substrat pour les tyrosines kinases dans les microvilli. Sert d'intermédiaire entre la membrane plasmique et le cytosquelette d'actine. Rôle clé dans l'adhésion, la migration et l'organisation cellulaires.	Rôle dans le développement de complications dans le diabète (McRobert, 2003). L'IL-13 agit sur cette protéine pour induire des dommages et l'obstruction des voies aériennes (Laoukili, 2001).
<i>VIP</i> 6q26	Non	<i>Vasoactive intestinal peptide</i>	Protéine sécrétée de la famille des glucagons. Stimule la contraction du myocarde, cause la vasodilatation, augmente la glycogénolyse, diminue la pression artérielle et relaxe le muscle lisse de la trachée, de l'estomac et de la vessie.	Niveau élevé dans le sérum de sujets atteints de dermatite atopique (Umemoto, 2003). Dilatateur du muscle lisse bronchique (Ichinose, 1999). Peut moduler la réponse immunitaire dans l'arthrite rhumatoïde (Jorgensen, 1994). Peut être une importante composante des mécanismes anti-asthme (Szema, 2006). Augmente les propriétés de la fractalkine à attirer les mastocytes dans l'asthme (El-Shazly, 2006).



<i>ADAM8</i> 10q26.3	Non	<i>ADAM metalloproteinase domain 8</i>	Protéine membranaire impliquée dans une variété de processus biologiques, notamment les interactions cellule-cellule ou cellule-tissu. Peut être impliquée dans l'adhésion cellulaire durant la neurodégénération.	Rôles dans la progression du cancer du poumon (Ishikawa, 2004) et dans le système immunitaire (Yamamoto, 1999). Peut induire la perte des molécules d'adhésion vasculaires (VCAM-1) (spécifiques pour l'adhésion des éosinophiles), ce qui peut être important dans la suppression de l'asthme dans un modèle expérimental (Matsuno, 2005). Responsable de la perte des ectodomains des protéines membranaires impliquées dans l'inflammation et la neurodégénération (Naus, 2006). Régule l'infiltration des leucocytes, directement ou indirectement (Higuchi, 2002). Exprimée dans les lymphocytes B sanguins périphériques (Richens et al., 2007). Exprimée par les cellules musculaires des myopathies inflammatoires (Dehmel et al., 2007). Associée l'éosinophilie et à l'inflammation pulmonaire chez l'humain (Matsuno et al., 2007).
<i>DOCK1</i> 10q26.3	Oui	<i>Dedicator of cytokinesis 1</i>	Protéine impliquée dans les réarrangements du cytosquelette requis pour la phagocytose des cellules apoptotiques et pour la mobilité cellulaire.	-
<i>FANK1</i> 10q26.3	Non	<i>Fibronectin type III and ankyrin repeat domains 1</i>	-	-
<i>GPR123</i> 10q26.3	Non	<i>G protein-coupled receptor 123</i>	Récepteur orphelin.	Distribution dans plusieurs tissus (Bjarnadottir, 2004).
<i>PTPRE</i> 10q26.3	Oui	<i>Protein tyrosine phosphatase, receptor type, E</i>	Les protéines tyrosine phosphatases sont des molécules de signalisation qui régulent une variété de processus cellulaires comme la croissance et la différenciation cellulaires et la transformation oncogénique.	Rôle dans le développement des maladies inflammatoires chroniques du cerveau (Schumman, 1998). Rôle dans le contrôle des fonctions des macrophages (Sully, 2001).

- = pas d'information disponible.

<sup>a</sup> Oui = Apparition dans au moins deux analyses G2D.

<sup>b</sup> Source : NCBI (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene>).

<sup>c</sup> Information tirée des bases de données publiques NCBI (<http://www.ncbi.nlm.nih.gov/>) et GeneCards (<http://www.genecards.org/index.shtml>).

<sup>d</sup> Références disponibles dans le Chapitre 12.

## 6.1 Cinquième article

### Application de l'outil bioinformatique G2D dans l'identification de gènes candidats pour l'asthme

Karine Tremblay, Mathieu Lemire, Camille Potvin, Alexandre Tremblay, Gary M Hunninghake, Benjamin A Raby, Thomas J Hudson, Carolina Perez-Iratxeta, Miguel A Andrade-Navarro et Catherine Laprise

*PLoS ONE* (2008) Aug 6;3(8):e2907

#### Résumé

Un criblage génomique réalisé dans un échantillon de sujets asthmatiques provenant du Saguenay–Lac-St-Jean (SLSJ) a permis d'identifier une région chromosomique liée à l'asthme (6q26) et une liée à l'atopie (10q26.3). L'outil informatique *Genes to Diseases* (G2D) a été employé pour prioriser dix gènes (*LPA*, *NOX3*, *SNX9*, *VIL2*, *VIP*, *ADAM8*, *DOCK1*, *FANK1*, *GPR123*, *PTPRE*) localisés dans ces deux régions afin de démontrer l'utilité de G2D dans l'identification de gènes candidats pour l'asthme. Une étude d'association pour l'asthme et l'atopie a été conduite dans l'échantillon familial du SLSJ (n = 1167). Le génotypage des polymorphismes sélectionnés dans les huit gènes a été effectué avec les technologies Sequenom® et Taqman® et un test d'association, incluant une correction statistique pour tests multiples, a été appliqué. Le séquençage du gène *PTPRE* a été fait avec les ADN de 40 individus asthmatiques allergiques provenant du SLSJ. La réplication des associations positives a été effectuée dans l'échantillon indépendant CAMP (*Childhood Asthma Management Program*). Les résultats ont montré une association positive du *PTPRE* rs7081735 pour l'asthme allergique au SLSJ (p = 0,0006; p corrigée = 0,01). Cette association n'a pas été trouvée dans CAMP. Le séquençage des régions entourant le *PTPRE* rs7081735 a permis d'identifier sept SNPs, dont quatre qui n'ont pas été ciblés pour les tests d'association préliminaires. Leur génotypage et leur analyse génétique n'a révélé aucune association positive. Ces résultats démontrent l'utilité de l'outil G2D dans la sélection de gènes candidats pour l'asthme, nous ayant permis de documenter une nouvelle association de protection pour l'asthme allergique au SLSJ avec un SNP du gène *PTPRE*.

### 6.1.1 TITLE PAGE

#### **Genes to Diseases (G2D) Computational Method to Identify Asthma Candidate Genes**

Karine Tremblay<sup>1,2</sup>, Mathieu Lemire<sup>3</sup>, Camille Potvin<sup>1,2</sup>, Alexandre Tremblay<sup>2</sup>, Gary M Hunninghake<sup>4</sup>, Benjamin A Raby<sup>4</sup>, Thomas J Hudson<sup>3,5</sup>, Carolina Perez-Iratxeta<sup>6</sup>, Miguel A Andrade-Navarro<sup>6,7,8</sup> and Catherine Laprise<sup>2,9</sup>

<sup>1</sup> Department of Medicine, Laval University, Québec, QC, G1V 0A6, Canada; <sup>2</sup> University of Montreal Community Genomic Centre, Chicoutimi Hospital, Saguenay, QC, G7H 5H6, Canada; <sup>3</sup> Ontario Institute for Cancer Research, Toronto, ON, M5G 0A3, Canada; <sup>4</sup> Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, 02115 – 5804, United-States; <sup>5</sup> McGill University and Genome Quebec Innovation Centre, Montreal, QC, H3A 1A4, Canada; <sup>6</sup> Molecular Medicine, Ottawa Health Research Institute, Ottawa, ON, K1H 8L6, Canada; <sup>7</sup> Faculty of Medicine, University of Ottawa, Ottawa, ON, K1H 8M5, Canada; <sup>8</sup> Max Delbrück Center for Molecular Medicine, Berlin, 13125, Germany; <sup>9</sup> Département des Sciences fondamentales, Université du Québec à Chicoutimi, Saguenay, QC, G7H 2B1, Canada.

#### **Contact**

Catherine Laprise, Ph.D.  
Université du Québec à Chicoutimi  
555, boulevard de l'Université, Chicoutimi (QC), Canada, G7H 2B1  
Telephone: (418) 545-5011, #5659; Fax: (418) 541-1116  
Email: catherine\_laprise@uqac.ca

#### **Funding sources**

The Canadian Institutes of Health Research and the Respiratory Health Network of the Fonds de la recherche en santé du Québec financially supported this work. The Childhood Asthma Management Program Genetics Ancillary Study is supported by U01 HL075419, U01 HL65899, P01 HL083069, R01 HL 086601, and T32 HL07427 from the National Heart, Lung and Blood Institute, National Institutes of Health.

### 6.1.2 ABSTRACT

Asthma is a complex trait for which different strategies have been used to identify its environmental and genetic predisposing factors. Here, we describe a novel methodological approach to select candidate genes for asthma genetic association studies. In this regard, the Genes to Diseases (G2D) computational tool has been used in combination with a genome-wide scan performed in a sub-sample of the Saguenay–Lac-St-Jean (SLSJ) asthmatic familial collection ( $n = 609$ ) to identify candidate genes located in two suggestive loci shown to be linked with asthma (6q26) and atopy (10q26.3), and presenting differential parent-of-origin effects. This approach combined gene selection based on the G2D data mining analysis of the bibliographic and protein public databases, or according to the genes already known to be associated with the same or a similar phenotype. Ten genes (*LPA*, *NOX3*, *SNX9*, *VIL2*, *VIP*, *ADAM8*, *DOCK1*, *FANK1*, *GPR123* and *PTPRE*) were selected for a subsequent association study performed in a large SLSJ sample ( $n = 1167$ ) of individuals tested for asthma and atopy related phenotypes. Single nucleotide polymorphisms ( $n = 91$ ) within the candidate genes were genotyped and analysed using a family-based association test. The results suggest a protective association to allergic asthma for *PTPRE* rs7081735 in the SLSJ sample ( $p = 0.000463$ ; corrected  $p = 0.0478$ ). This association has not been replicated in the Childhood Asthma Management Program (CAMP) cohort. Sequencing of the regions around rs7081735 revealed additional polymorphisms, but additional genotyping did not yield new associations. These results demonstrate that the G2D tool can be useful in the selection of candidate genes located in chromosomal regions linked to a complex trait.

### 6.1.3 INTRODUCTION

Asthma involves genetic and environmental factors in its development, chronicity and severity [1, 2]. Although some of its underlying mechanisms have been elucidated in recent years, more work is needed to gain a clearer understanding of genetic determinants. The mapping of asthma has been one of the most important areas of human genetics in the last two decades. According to an overview by Blumenthal (2005), twelve complete and two incomplete genome scans for asthma have been published, identifying a total of twenty chromosomal linked regions to asthma [3].

Discrepancies often appeared between linkage studies [3-5], leading to the use of standard phenotype definitions and founder populations as a way to decrease phenotypic and genetic heterogeneity [6-8].

To date, common strategies have been employed to identify genes involved in asthma predisposition. Linkage studies followed by positional cloning identified six genes while association studies identified over a hundred genes, the majority of these having quite small effects on asthma susceptibility (see [9, 10] for a review). Here, we describe a novel methodological approach that combines classical genetic approaches with a computational data-mining tool. In this regard, a genome-wide scan for asthma and atopy in families originating from the Saguenay–Lac-St-Jean (SLSJ) founder population (Northeastern Quebec, Canada) [11-16] has been combined with the Genes to Diseases (G2D) new computational tool in the prioritization of asthma candidate genes [17, 18]. G2D performs the selection of the candidates in the chromosomal regions genetically linked to a disease by highlighting genes whose functions are related to the phenotype of the disease according to a data mining analysis of the bibliographic and protein public databases, or according to the genes already known to be associated with the same or a similar phenotype.

Single nucleotide polymorphisms (SNPs) within the candidate genes prioritized by the G2D tool have been subsequently used to conduct an association study with asthma, atopy and allergic asthma phenotypes in the SLSJ asthma familial collection. The chromosomal regions around associated SNPs have then been sequenced in order to find causal mutations, and a replication of the positive association findings has been assessed in the Childhood Asthma Management Program (CAMP) independent cohort. The goal was to apply the G2D tool in the search of genetic determinants associated with a complex trait, using asthma as a model. This approach led to the hypothesis-driven identification of ten genes that may not have been selected otherwise (*LPA*, *NOX3*, *SNX9*, *VIL2*, *VIP*, *ADAM8*, *DOCK1*, *FANK1*, *GPR123* and *PTPRE*). Of these, one positive genetic association, resisting to corrections for multiple testing, has been found between the protein tyrosine phosphatase receptor type E gene (*PTPRE*) and allergic asthma in the SLSJ sample.

## 6.1.4 METHODS

### Subjects

Clinical evaluation and phenotyping criteria of the Saguenay-Lac-St-Jean (SLSJ) subjects have been described in recent reports [19-21] and summarized in Table 1. This familial sample is predominantly composed of probands that reported an onset age of asthma below 12 years old (81.6% of the probands). The mean age of onset for the probands is 7 years and the mean age of onset for the asthmatic family members is 22 years. The entire sample has been used for the association study while for the genome scan, the first 79 recruited families, that were available at the time when the genome scan was performed have been used ( $n = 609$  individuals – see supplementary Table S1 for subjects characteristics and studied phenotypes). The Chicoutimi Hospital local ethics committee approved the study and all subjects provided informed consent.

### Genome Scan

DNA was extracted for all SLSJ participants from whole blood by using the QIAGEN genomic purification procedure (QIAGEN Inc., Valencia, CA). Genotyping was completed on 367 autosomal and 21 X-chromosome microsatellite markers evenly spaced throughout the genome (average spacing of 9.2 cM). The marker set is a modification of the Cooperative Human Linkage Centre Screening Set (<http://gai.nci.nih.gov/CHLC/>, version 6.0), showing an average heterozygosity of 0.72 in our data set. Each primer was amplified separately and then pooled into panels of eight markers and products were interrogated using ABI 3700 sequencers (Applied Biosystems, Foster City, CA) with a size standard ladder. Duplicate of two CEPH control DNAs and one water were included in each genotyping plate. In our data set of 195939 autosomal genotypes, 91.4% of the alleles were called and the proportion of observed Mendelian error was 1.2%. Linkage of chromosomal regions with putative genetic risk factors for a given trait was assessed by evaluating the extent of excess sharing of alleles identical by descent in affected relatives within families. Test statistics are reported on the LOD scale. Briefly, a multipoint, one-parameter likelihood ratio test that is robust against incompleteness of marker data (when the descent of alleles in a pedigree is not fully known) was used [22, 23]. We moreover evaluated the specific contribution of mothers and fathers to the test of linkage, to look for parent-of-origin effects. The above tests of linkage, linkage through mothers and linkage through fathers are described in details in [24].

### **Genes to Diseases (G2D)**

The G2D tool has been applied in the two best genome scan susceptibility regions: 6q26 between markers D6S476 and D6S305, and 10q26.3 between markers D10S1223 and D10S1248. We used two of the approaches considered in G2D [18] to pre-select gene lists in these regions. The first approach uses a description of the phenotype to point to genes in a region. This method works with automatically derived relationships between the disease symptoms (as MeSH C terms) and gene features (as Gene Ontology or GO terms [25]) that are obtained from the literature and Entrez gene database (<http://www.ncbi.nlm.nih.gov/>). We call this procedure the "PHENOTYPE" method. The second approach consists on automatically finding genes in a region that are similar to other genes previously associated to asthma. To do this, G2D measures the semantic distance between the annotations of the "known genes" and the annotations of genes in the problem region assigned by homology searches. We call this procedure the "KNOWN GENES" method. For details about the algorithm see the G2D web site at [http://www.ogic.ca/projects/g2d\\_2/](http://www.ogic.ca/projects/g2d_2/) and [17, 18].

### **Association Study**

*SNP selection:* The HapMap database (<http://www.hapmap.org/>) has been used to identify SNPs assumed to be polymorphic in the SLSJ population. TagSNPs were then selected with the tagger program implemented in the Haploview software (version 3.32) [26] using an  $r^2$  cutoff of 0.8 and a minor allele frequency (MAF) over 0.10 to cover each whole gene. SNPs were also prioritized on their localization (coding or untranslated regions - see supplementary Table S2). All SNPs are referred using their reference sequence number (rs#).

*SNP genotyping:* Eighty SNPs have been genotyped by the Sequenom® matrix-assisted laser desorption/ionization time-of-flight mass array spectrometer (Sequenom Inc., San Diego, CA) (Table S2). Sequenom primers were designed using the Sequenom SNP Assay Design software version 3.0 for iPLEX reactions. A total of 74 assays were designed for a single multiplex reaction. The assay group file containing the PCR primers and the iPLEX extension probes can be supplied on request to the corresponding author. The protocol and reaction conditions are in accordance with the manufacturer [27]. The genotypes were viewed and analyzed using the MassARRAY Typer software version 3.4 (Sequenom Inc., San Diego, CA). The ten remaining SNPs

(Table S2) have been genotyped by the TaqMan® SNP Genotyping Assays (Applied Biosystems, Foster City, CA) using the Rotor-Gene™ real-time PCR (Corbett Research Ltd, Sydney, Australia). Protocol and method were supplied by the manufacturer and PCR conditions were optimized to get a good cluster separation between different genotypes (see supplementary Table S3 for PCR conditions). Genotypes were attributed by the Rotor Gene software using the scatter graph analysis option.

*Statistical analysis:* Family-based association testing has been performed with the FBAT software (version 1.7) using an empirical estimate of the variance [28-30] to correctly account for linkage. A Sidak correction for multiple testing has been applied on the p-values accounting for the effective number of independent phenotypes and SNPs, according to the definition of Li and Ji [31], as implemented in the SNPSpD program [32] (see the online supporting text file for supplementary details). Parent-specific transmission disequilibrium tests were performed using sib\_tdt from the ASPEX package (<http://aspex.sourceforge.net/>). Mendelian errors have been assessed by FBAT and Hardy-Weinberg equilibrium has been assessed with Haploview software (version 3.32)[26].

### ***PTPRE* Sequencing**

Forty unrelated SLSJ subjects (validated with the BALSAC database [33]) presenting full-fit allergic asthmatic criteria and that have contributed to the *PTPRE* association were selected. *PTPRE* sequence information was obtained from Ensembl database (<http://www.ensembl.org/index.html>, release 46). The sequencing was divided in six regions that spanned 2.7 kb, starting from the exon 2 to the exon 4. Oligonucleotides and PCR conditions are listed in Table S4. Amplification products were purified with multiscreen PCR plates (Millipore Corporation, Billerica, MA), sequenced with BigDye terminator v3.1 chemistry following instructions of the manufacturer and analyzed on a 3100 Genetic analyzer (Applied Biosystems, Foster City, CA). Sequence analysis was performed with Codoncode Aligner software (CodonCode Corporation, <http://www.codoncode.com/>). The identified SNPs presenting MAF over 0.05 have been genotyped in the SLSJ sample using a Sequenom panel and analyzed with FBAT, as described above. Newly described SNPs have been submitted to NCBI (<http://www.ncbi.nlm.nih.gov/>) SNP database.



### Replication Study

The *PTPRE* rs7081735 association has been assessed in the Childhood Asthma Management Program (CAMP) study [34, 35]. This analysis includes the 497 non-Hispanic white children and their parents for whom adequate DNA was available. The genotyping of the *PTPRE* rs7081735 was performed using a TaqMan® SNP Genotyping Assays (Applied Biosystems, Foster City, CA) (see Table S3 for PCR conditions). Plates were scanned using the 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA) and genotypes were assigned by the SDS 2.2 software using the scatter graph analysis option. The Institutional Review Board of the Brigham and Women's Hospital (BWH), as well as those of the other CAMP study centers, approved this study. Informed assent and consent were obtained from the study participants and their parents to collect DNA for genetic studies.

## 6.1.5 RESULTS

### Genome Scan

The genome-wide linkage scan analysis revealed at least two regions showing suggestive evidence for linkage, as well as differential maternal and paternal contribution (Figure 1). The two regions are 6q26 for asthma (LOD = 1.54,  $p = 0.0038$ ) and 10q26.3 for atopy (LOD = 2.82,  $p = 0.00016$ ). In these two regions, affected sibs tend to share more alleles inherited from the mothers than from the fathers. The linkage tests through the mothers reach a LOD of 2.19 ( $p = 0.00074$ ) in 6q26 for asthma and a LOD of 2.96 ( $p = 0.00011$ ) in 10q26.3 for atopy. The respective LODs obtained through the fathers at the same loci are only 0.01 ( $p = 0.40$ ) and 0.58 ( $p = 0.05$ ). Moreover, even though the 10q region does not show a great strength of linkage with asthma (LOD = 0.57,  $p = 0.051$ ), the asthmatic sibs tend to share these alleles when received from their mothers (LOD = 2.82,  $p = 0.00016$ ). Each of the two chromosomal regions show LOD values (either for tests of linkage, or parent-specific LODs, or both) that are, in order of magnitude, consistent with what has been defined as « suggestive » for linkage by Lander and Kruglyak [36], which are expected to occur once per whole genome scan on average. Because of their great hypothesis generating potential, the 6q26 and 10q26 regions as well as asthma and atopy phenotypes were selected for the following G2D and association studies.

## G2D

We applied the two algorithms PHENOTYPE and KNOWN GENES to the 6q26 and 10q26.3 regions. For the PHENOTYPE method, we used the OMIM [37] record 600807 as input, particularly the MeSH C terms from the MEDLINE references in OMIM entry 600807 which deals with susceptibility to asthma and asthma related traits. The most prevailing MeSH C terms are "Asthma" and "Bronchial Hyperreactivity", but the list also includes "Hypersensitivity", "Respiratory Hypersensitivity" and "Eosinophilia". We thus considered that the terms that refer to asthma were "Asthma" and "Bronchial Hyperreactivity" and that the terms that refer to atopy were "Hypersensitivity", "Respiratory Hypersensitivity" and "Eosinophilia". The highest scoring GO terms, associated to these MeSH C terms, describe a variety of molecular functions and processes that include leukotriene and interleukin signaling, glutathione metabolism, etc (see header of supplementary Table S5). We applied this method to the 6q26 region between D6S476 and D6S305, the two markers directly flanking the linkage peak seen in the region, which corresponds to the 10.48 MB band between positions 151,685,574 and 162,165,587 of chromosome 6. After discarding candidates that did not overlap with any known or hypothetical Entrez Gene sequence, 16 genes remained (see supplementary Table S5-A). A similar analysis was carried between D10S1223 and D10S1248, the two markers directly flanking the linkage peak in 10q26.3, between positions 129,150,822 and 130,982,363 in chromosome 10. In that case, 12 candidates were obtained (see supplementary Table S5-B). For the KNOWN GENES method, we compiled a list of genes reported to be associated with asthma and atopy from the literature [38] and from the Genetic Association Database GAD [39] (supplementary Table S6). We then extracted the GO annotation of those genes in Entrez Gene [37]. We derived a scoring system for the candidates according to the minimal semantic distance between their GO annotation and the ones from the compiled known-gene list. For example, genes annotated with GO terms such as "dipeptidyl-peptidase IV activity" or "chemokine receptor binding" would score high as candidates. The method takes into account the hierarchical structure of GO as well as the specificity of GO terms. In that sense, similarity with more infrequent terms receive higher scores. For the complete list of GO terms see the header of supplementary Table S7. We applied the KNOWN GENES method to both regions 6q26 and 10q26.3, obtaining 15 and 10 genes, respectively, after filtering out those candidates that did not overlap with either known or hypothetical genes (see supplementary Table S7).

*Use of a complementary method based on genomic sequence:* We applied the G2D complementary Disease Gene Prediction (DGP) tool [40] to both genetically linked regions in order to predict the involvement of genes in inherited disease by their sequence features. This method analyses the probability of a gene to be associated with any disease phenotype. Genes with no associated phenotype and a probability greater than 0.7 were retained. With this criterion the DGP method identified one candidate in chromosome 10 (MMP21) and three candidates in the 6q26 region: *TFB1M* [MIM:607033], *RGS17* [MIM:607191] and *VIP* [MIM:192320].

*Final list of candidate genes:* Genes that received the higher scores in the pre-selected lists, and those that were pointed by the two different analyses were preferred, allowing the construction of a list of 17 candidates (displayed in the Table 2). We then applied a candidate gene approach to select a final list of ten genes with the best biological potential related to asthma pathophysiology for genotyping (five in each chromosomal region): *LPA*, *NOX3*, *SNX9*, *VIL2*, *VIP*, *ADAM8*, *DOCK1*, *FANK1*, *GPR123* and *PTPRE* (marked in bold in Table 2).

### **Association Study**

A final panel of 91 SNPs (75 tagSNPs and 16 non-tagSNPs) was selected among the ten candidate genes (supplementary Table S2). Of these, four were non-polymorphic, six failed Sequenom genotyping assays, none presented deviation from Hardy-Weinberg equilibrium (p-values > 0.001) and one presented more than two Mendelian errors (Table S2). For the remaining 80 SNPs, genotypes of individuals with Mendelian errors were considered as missing data in the FBAT analyses. Genotyping presented a mean success rate of 99.0% for the Sequenom assays and a mean success rate of 98.3% for the TaqMan assays. Accounting for the residual correlation between the tagSNPs, it is estimated that the 80 partially correlated SNPs correspond to an effective number of 59 independent ones [31]. As for the effective number of phenotypes, simulation indicates that the three studied phenotypes (asthma, atopy and allergic asthma) correspond to 1.75 effective independent ones (see supplementary text file). Accordingly, the estimated total effective number of independent tests is 103.25 (59 effective independent SNPs x 1.75 effective independent phenotypes). Thus, applying Sidak correction, the p-value threshold of significance is estimated to be 0.000483.

The FBAT single marker analyses were performed under an additive genetic model for each SNP and the three studied phenotypes. For the sake of brevity, only results showing a p-value under 0.05 before correction for multiple testing for one or more phenotypes are presented in Table 3. For the asthma phenotype, minor alleles LPA\_rs12175867C, GPR123\_rs11101913T, GPR123\_rs11101932T and GPR123\_rs12257731A were overtransmitted to the asthmatic probands, suggesting a susceptibility effect of these alleles for asthma ( $0.0085 < p < 0.047$ ). Inversely, minor alleles DOCK1\_rs1051039G, PTPRE\_rs4369314A and PTPRE\_rs7081735G were undertransmitted to the asthmatic probands, suggesting a protective effect of these alleles for asthma ( $0.010 < p < 0.049$ ). For the atopy phenotype, only ADAM8\_rs11101672G and PTPRE\_rs7081735G minor alleles have been undertransmitted to the atopic probands, suggesting a protective effect ( $p = 0.039$  and  $0.037$ , respectively). Finally, for the allergic asthma phenotype, minor alleles LPA\_rs12175867C, GPR123\_rs11101913T, GPR123\_rs11101932T and GPR123\_rs12257731A were overtransmitted to the allergic asthmatic probands, suggesting a susceptibility effect of these alleles for allergic asthma ( $0.035 < p < 0.041$ ). Inversely, minor alleles ADAM8\_rs11101672G, GPR123\_rs11101916A, GPR123\_rs761777G, PTPRE\_rs11016002A, PTPRE\_rs4002572C and PTPRE\_rs7081735G were undertransmitted to the allergic asthmatic probands, suggesting a protective effect of these alleles for allergic asthma ( $0.000463 < p < 0.037$ ). None of the SNPs reported above showed a significantly greater extent of transmission distortion from mothers than from fathers, thus not providing insights to the observed parental distortions seen in the linkage results (not shown).

### ***PTPRE* Sequencing**

*PTPRE* rs7081735 shows the strongest association to allergic asthma ( $p = 0.000463$ ) and is the only SNP shown to be significant after multiple testing correction (corrected  $p = 0.0478$ ). We thus sequenced strategic genomic regions around the rs7081735, including coding regions in order to identify causal mutation. Figure 2 shows the *PTPRE* sequenced regions, the rs7081735 localization and the nine identified variants (diamonds), including four novel ones (c.86172A>G, c.140901G>A, c.140903G>A and c.141102C>T) (Table 4). It is worth to note that any of these identified polymorphisms may affect either *PTPRE* isoforms. A family-based association analysis has been performed between the three studied phenotypes and variants presenting a MAF over

0.05 that were not included in the first genotyping panel, which was the case for four of the nine identified variants (rs7911506, c.86172A>G, rs7895103 and c.140901G>A). Only rs7895103 showed a modest association with asthma, at a level that does not provide additional insights ( $p = 0.013$ , compared to  $p = 0.000463$  for rs7081735).

### Replication Study

Considering that the *PTPRE* rs7081735 association to allergic asthma is the strongest in the SLSJ sample, we evaluated it in an independent familial cohort, the Childhood Asthma Management Program (CAMP) [34]. The genotyping completion rate was 96% and no discordance was observed upon repeat genotyping of two random plates. The minor allele frequency was 0.33 and was in Hardy-Weinberg equilibrium ( $p = 0.90$ ). Family-based association showed no significant association for asthma or atopy phenotypes (data not shown). However, to ensure that the selection of the CAMP study for the replication is appropriate and to demonstrate that the association found could result from a childhood subset of asthma instead of an adulthood one, we stratified the SLSJ association analyses considering only the probands that reported an asthma age of onset below 12 years old. Thus, the positive association found for *PTPRE* rs7081735 and allergic asthma ( $p = 0.000463$ , Table 3) remains significant when considering only those probands (Family number = 74;  $Z$  for the minor allele = -3.166;  $p = 0.001546$ ). Even with the loss of statistical power due to the stratified analysis, this comparison allowed to assume that the association found for *PTPRE* is probably more related to a childhood asthma, comforting the choice of the CAMP cohort as a replication study.

### 6.1.6 DISCUSSION

This study proposes a novel approach for the selection of candidate genes for asthma association studies using the computational G2D tool to find genetic determinants for this disease. Based on a genome-wide scan performed in an asthmatic familial sample from the SLSJ founder population, we selected the two best-linked regions (6q26 and 10q26.3) and applied the G2D data mining approach [17, 18] to identify ten candidate genes for an association in the SLSJ sample. Among these, five (*LPA*, *ADAM8*, *DOCK1*, *GPR123* and *PTPRE*) presented modest associations with asthma, atopy or allergic asthma. After corrections for multiple testing, only the *PTPRE* rs7081735

association to allergic asthma remained significant. These findings demonstrate that the G2D tool can be useful in the selection of candidate genes for asthma genetic studies.

Because the *PTPRE* association to allergic asthma remained significant after correction for multiple testing, we sequenced strategic regions around the rs7081735, aiming to find the causal mutation. Sequencing allowed the identification of five known and four novel variants. Testing four SNPs with MAF > 0.05 did not identify additional associated SNPs with asthma or atopy related phenotypes in the SLSJ sample. Thus, the rs7081735, located in the 5' untranslated region, could be the causal variant, or be in linkage disequilibrium with an unknown causal mutation.

*PTPRE* is a member of the protein tyrosine phosphatase (PTPs) family, which includes genes that are important regulators of signal transduction pathways involved in various cellular processes such as control of metabolic pathways, cellular adhesion, cell cycle progression and immune response [41, 42]. *PTPRE* encodes two different isoforms, cytoplasmic and transmembrane [43], based on its different promoters [44]. PTPs receptors participate in transmembrane signaling and cellular adhesion processes, whereas intracellular PTPs take part in signal transduction within the cell [45]. For example in mice, PTP $\epsilon$ -deficient macrophages present abnormalities in the regulation of the respiratory burst and the production of cytokines in response to bacterial lipopolysaccharide, suggesting a role of the PTP $\epsilon$  isoform in inflammation as well as in host defense [46]. However, *PTPRE* has been shown to be highly expressed in peripheral human monocytes and granulocytes, and antigen-receptor stimulation induces the expression of *PTPRE* in activated lymphocytes [47]. These observations and our finding suggest a protective effect of *PTPRE* in allergic asthma, and lead us to hypothesize that this potential protective role could involve the leukocyte cellular processes in the limitation of lung inflammation following an allergen sensitization. Further work is needed to define the *PTPRE* possible role in asthma pathophysiology.

The *PTPRE* rs7081735 association has been evaluated in the independent CAMP cohort. Results showed no positive association for asthma and atopic phenotypes. Taking into account that the SLSJ and the CAMP studies are both family-based designed, well powered [48, 49] and presenting a childhood onset asthma, the *PTPRE* association lack of replication in the CAMP study can be explained by other reasons:

natural variability of asthma history [50, 51], differences of proband mean age between samples (SLSJ mean age of 18, and CAMP mean age of 8, respectively), differences in the genetic background of the two populations [52] (SLSJ individuals descend predominantly from French European founders [11-16] whereas CAMP individuals come from a white North-American admixed population), or population specific gene-gene and gene-environment interactions [52]. Based on these observations, we conclude that the *PTPRE* rs7081735 association to allergic asthma is more penetrant in the SLSJ population, possibly resulting from an interaction with others genes and/or environmental factors that are more common in this founder population.

In summary, this study demonstrates that the G2D tool can be useful in the prioritization of candidate genes for a complex disease as it allowed us to find a novel asthma genetic association with the *PTPRE* gene in the SLSJ familial asthma sample. This association represents a potential protective factor for asthma pathogenesis, as it is more likely related to a childhood onset asthma. The present genetic study is an example of how the combination of different methodological approaches can be relevant to target asthma genetic determinants and to motivate further genetic and functional investigations.

#### **6.1.7 ACKNOWLEDGMENTS**

We thank families of the Saguenay–Lac-St-Jean for their participation in this study. We also thank Janet Murphy for her participation in the genome scan as well as Alexandre Belisle, Pierre Lepage and Charleen Salesse for the genotyping assays. K. Tremblay is an AllerGen PhD trainee and is supported by the Fondation de l'Université Laval studentship. T.J. Hudson received an Investigator Award from the Canadian Institutes of Health Research (CIHR) and a Clinician-scientist Award in Translational Research from the Burroughs Wellcome Fund. M.A. Andrade-Navarro holds a Canada Research Chair in Bioinformatics. C. Laprise is the chairholder of the Canada Research Chair (<http://www.chairs.gc.ca>) on genetic determinants in asthma and the director of the Genetics platform of the Respiratory Health Network of the Fonds de la recherche en santé du Québec. We also thank all subjects of the CAMP study for their ongoing participation. We acknowledge the CAMP investigators and research team, supported by NHLBI, for collection of CAMP Genetic Ancillary Study data. All work on data

collected from the CAMP Genetic Ancillary Study was conducted at the Channing Laboratory of the Brigham and Women's Hospital under appropriate CAMP policies and human subject's protections.

#### 6.1.8 REFERENCES

1. Maddox L, Schwartz DA (2002) The pathophysiology of asthma. *Annu Rev Med* 53: 477-498.
2. Cookson WO, Moffatt MF (2000) Genetics of asthma and allergic disease. *Hum Mol Genet* 9: 2359-2364.
3. Blumenthal MN (2005) The role of genetics in the development of asthma and atopy. *Curr Opin Allergy Clin Immunol* 5: 141-145.
4. Sandford A, Weir T, Pare P (1996) The genetics of asthma. *Am J Respir Crit Care Med* 153: 1749-1765.
5. Hoffjan S, Ober C (2002) Present status on the genetic studies of asthma. *Curr Opin Immunol* 14: 709-717.
6. Rannala B (2001) Finding genes influencing susceptibility to complex diseases in the post-genome era. *Am J Pharmacogenomics* 1: 203-221.
7. Ober C, Tsalenko A, Parry R, Cox NJ (2000) A second-generation genomewide screen for asthma-susceptibility alleles in a founder population. *Am J Hum Genet* 67: 1154-1162.
8. Laitinen T, Daly MJ, Rioux JD, Kauppi P, Laprise C, et al. (2001) A susceptibility locus for asthma-related traits on chromosome 7 revealed by genome-wide scan in a founder population. *Nat Genet* 28: 87-91.
9. Hoffjan S, Nicolae D, Ober C (2003) Association studies for asthma and atopic diseases: a comprehensive review of the literature. *Respir Res* 4: 14.
10. Ober C, Hoffjan S (2006) Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 7: 95-100.
11. Heyer E, Tremblay M (1995) Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 56: 970-978.
12. Sriver CR (2001) Human genetics: lessons from Quebec populations. *Annu Rev Genomics Hum Genet* 2: 69-101.



13. Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, et al. (1996) Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am J Hum Genet* 59: 633-643.
14. Engert JC, Berube P, Mercier J, Dore C, Lepage P, et al. (2000) ARSACS, a spastic ataxia common in northeastern Quebec, is caused by mutations in a new gene encoding an 11.5-kb ORF. *Nat Genet* 24: 120-125.
15. Richter A, Rioux JD, Bouchard JP, Mercier J, Mathieu J, et al. (1999) Location score and haplotype analyses of the locus for autosomal recessive spastic ataxia of Charlevoix-Saguenay, in chromosome region 13q11. *Am J Hum Genet* 64: 768-775.
16. Lander ES, Schork NJ (1994) Genetic dissection of complex traits. *Science* 265: 2037-2048.
17. Perez-Iratxeta C, Wjst M, Bork P, Andrade MA (2005) G2D: a tool for mining genes associated with disease. *BMC Genet* 6: 45.
18. Perez-Iratxeta C, Bork P, Andrade-Navarro MA (2007) Update of the G2D tool for prioritization of gene candidates to inherited diseases. *Nucleic Acids Res* Jul 1: W212-W216.
19. Poon AH, Laprise C, Lemire M, Montpetit A, Sinnett D, et al. (2004) Association of vitamin D receptor genetic variants with susceptibility to asthma and atopy. *Am J Respir Crit Care Med* 170: 967-973.
20. Tremblay K, Lemire M, Provost V, Pastinen P, Renaud Y, et al. (2006) Association study between the CX3CR1 gene and asthma. *Genes Immun* 7: 632-639.
21. Begin P, Tremblay K, Daley D, Lemire M, Claveau S, et al. (2007) Association of urokinase-type plasminogen activator with asthma and atopy. *Am J Respir Crit Care Med* 175: 1109-1116.
22. Kong A, Cox NJ (1997) Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 61: 1179-1188.
23. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58: 1347-1363.

24. Lemire M (2005) A simple nonparametric multipoint procedure to test for linkage through mothers or fathers as well as imprinting effects in the presence of linkage. *BMC Genet* 6 Suppl 1: S159.
25. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25: 25-29.
26. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263-265.
27. Oeth P, Beaulieu M, Park C, Kosman D, del Mistro G, et al. (2005) iPLEX assay: Increased plexing efficiency and flexibility for MassArray system through single base primer extension with mass-modified terminators. *SEQUENOM Application Note*.
28. Lake SL, Blacker D, Laird NM (2000) Family-based tests of association in the presence of linkage. *Am J Hum Genet* 67: 1515-1525.
29. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, et al. (2004) Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet Epidemiol* 26: 61-69.
30. Laird NM, Horvath S, Xu X (2000) Implementing a unified approach to family-based tests of association. *Genet Epidemiol* 19 Suppl 1: S36-S42.
31. Li J, Ji L (2005) Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 95: 221-227.
32. Nyholt DR (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 74: 765-769.
33. Bouchard G, Roy R, Casgrain B, Hubert M (1989) [Population files and database management: the BALSAC database and the INGRES/INGRID system]. *Hist Mes* 4: 39-57.
34. The Childhood Asthma Management Program Research Group (1999) The Childhood Asthma Management Program (CAMP): design, rationale, and methods. *Control Clin Trials* 20: 91-120.
35. The Childhood Asthma Management Program Research Group (2000) Long-term effects of budesonide or nedocromil in children with asthma. *N Engl J Med* 343: 1054-1063.

36. Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11: 241-247.
37. Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, et al. (2007) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 35: D5-D12.
38. Wills-Karp M, Ewart SL (2004) Time to draw breath: asthma-susceptibility genes are identified. *Nat Rev Genet* 5: 376-387.
39. Becker KG, Barnes KC, Bright TJ, Wang SA (2004) The genetic association database. *Nat Genet* 36: 431-432.
40. Lopez-Bigas N, Ouzounis CA (2004) Genome-wide identification of genes likely to be involved in human genetic disease. *Nucleic Acids Res* 32: 3108-3114.
41. Alonso A, Sasin J, Bottini N, Friedberg I, Friedberg I, et al. (2004) Protein tyrosine phosphatases in the human genome. *Cell* 117: 699-711.
42. Li L, Dixon JE (2000) Form, function, and regulation of protein tyrosine phosphatases and their involvement in human diseases. *Semin Immunol* 12: 75-84.
43. Nakamura K, Mizuno Y, Kikuchi K (1996) Molecular cloning of a novel cytoplasmic protein tyrosine phosphatase PTP epsilon. *Biochem Biophys Res Commun* 218: 726-732.
44. Tanuma N, Nakamura K, Kikuchi K (1999) Distinct promoters control transmembrane and cytosolic protein tyrosine phosphatase epsilon expression during macrophage differentiation. *Eur J Biochem* 259: 46-54.
45. Schumann G, Fiebich BL, Menzel D, Hüll M, Butcher R, et al. (1998) Cytokine-induced transcription of protein-tyrosine-phosphatases in human astrocytoma cells. *Brain Res Mol Brain Res* 62: 56-64.
46. Sully V, Pownall S, Vincan E, Bassal S, Borowski AH, et al. (2001) Functional abnormalities in protein tyrosine phosphatase epsilon-deficient macrophages. *Biochem Biophys Res Commun* 286: 184-188.
47. Wabakken T, Hauge H, Finne EF, Wiedlocha A, Aasheim H (2002) Expression of human protein tyrosine phosphatase epsilon in leucocytes: a potential ERK pathway-regulating phosphatase. *Scand J Immunol* 56: 195-203.

48. Lemire M, Roslin NM, Laprise C, Hudson TJ, Morgan K (2004) Transmission-ratio distortion and allele sharing in affected sib pairs: a new linkage statistic with reduced bias, with application to chromosome 6q25.3. *Am J Hum Genet* 75: 571-586.
49. Hersh CP, Raby BA, Soto-Quiros ME, Murphy AJ, Avila L, et al. (2007) Comprehensive Testing of Positionally Cloned Asthma Genes in Two Populations. *Am J Respir Crit Care Med* 176: 849-857.
50. Reed CE (2006) The natural history of asthma. *J Allergy Clin Immunol* 118: 543-548; quiz 549-550.
51. Koh MS, Irving LB (2007) The natural history of asthma from childhood to adulthood. *Int J Clin Pract* 61: 1371-1374.
52. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K (2002) A comprehensive review of genetic association studies. *Genet Med* 4: 45-61.

### 6.1.9 LEGEND TO FIGURES

#### **Figure 1:** Genome scan results summary

Results from the tests of linkage with atopy (top three panels) and asthma (bottom three panels) reported on the LOD scale. For each phenotype, the top panel shows the results of the test of linkage (excess allele sharing), the middle panel shows results from the tests of linkage through mothers (excess allele sharing transmitted from mothers) and the bottom panel shows results from the tests of linkage through fathers (excess allele sharing transmitted from fathers).

#### **Figure 2:** *PTPRE* gene sequenced regions and identified SNPs scaled location

The black thick boxes above the gene define its sequenced parts, which are identified by the same numbers used in the Table 4, in which exact chromosomal positions are available. Studied SNPs are represented below the gene. The TagSNPs correspond to an asterisk (\*) and the SNPs identified by sequencing correspond to a diamond (◊). All *PTPRE* numbers for the discovered SNPs are based on the mRNA sequence NM\_006504 (variant 1, receptor form) and from the NCBI (<http://www.ncbi.nlm.nih.gov/>) SNP database (build 127). Image source: HapMap ([www.hapmap.org/](http://www.hapmap.org/)) October 2007 (Genome Browser, <http://genome.ucsc.edu/>, version 1.69); modified according to our study design.

**6.1.10 Table 1**

Clinical and phenotypic characteristics of the Saguenay–Lac-St-Jean association study sample subjects

	<b>Probands</b> (n = 226)	<b>Family members</b> (n = 941)
Male: Female ratio	1: 1.2	1: 1.1
Mean age in years (range)	18 (3-46)	44 (2-96)
Smoking status (n (%))		
Never	185 (83.3)	424 (45.8)
Ex-smoker	12 (5.4)	307 (33.2)
Smoker	25 (11.3)	195 (21.0)
FEV <sup>1</sup> % of predicted value (SD) *	92.3 (16.3)	94.4 (20.5)
PC <sub>20</sub> in mg/ml (SD) †	2.68 (3.59)	10.87 (5.23)
Serum IgE in µg/l (SD) ‡	233.2 (4.6)	108.0 (4.2)
Number of persons with subphenotypes (%)		
Asthma §	226/226 (100)	353/935 (37.8)
Atopy ¶	185/224 (82.6)	468/920 (50.9)
Allergic asthma (Asthma + Atopy)	185/224 (82.6)	239/344 (69.5)

\* FEV<sub>1</sub> = Geometric mean of the force expiratory volume in one second evaluated for 208 probands and for 683 family members.

† PC<sub>20</sub> = Concentration of methacholine inducing a 20% fall in FEV<sub>1</sub>. Geometric mean and SD were obtained from the log transformed PC<sub>20</sub> values. Evaluated for 186 probands and for 614 family members.

‡ IgE = Immunoglobulin E serum concentration. Geometric mean and SD were obtained from the log transformed IgE values. Evaluated for 211 probands and 704 family members.

§ Present asthma or past documented clinical history of asthma. The reported mean age of onset is 7 years among the asthmatic probands and 22 years among the asthmatic family members.

¶ Defined as at least one positive response on skin prick testing (wheal diameter > 3 mm at 10 minutes).

**6.1.11 Table 2**

Genes identified by G2D data mining analysis and those selected for the association study based on their number of appearance in G2D analyses and on their biological function

<b>Gene *</b>	<b>GeneID / [MIM]</b>	<b>Analysis †</b>	<b>Gene definition ‡</b>	<b>Gene function ‡</b>
<i>ARID1B</i> 6q26	57492 / [MIM:NA]	No	AT rich interactive domain 1B (SWI1-like)	Involved in transcriptional activation and repression of select genes by chromatin remodeling (alteration of DNA-nucleosome topology). Binds DNA non-specifically.
<i>ESR1</i> 6q26	2099 / [MIM:133430]	Yes	Estrogen receptor 1	Nuclear hormone receptor. The steroid hormones and their receptors are involved in the regulation of eukaryotic gene expression and affect cellular proliferation and differentiation in target tissues.
<i>IGF2R</i> 6q26	3482 / [MIM:147280]	Yes	Insulin-like growth factor 2 receptor	Transmembrane protein with a short cytoplasmic tail containing an internalization signal. This receptor binds insulin-like growth factor IGF2.
<i>LPA</i> 6q26	4018 / [MIM:152200]	No	Lipoprotein, Lp(a)	Apo(a) is the main constituent of lipoprotein(a). It has a serine proteinase activity and is able of autoproteolysis. Inhibits tissue-type plasminogen activator 1.
<i>NOX3</i> 6q26	50508 [MIM:607105]	Yes	NADPH oxidase 3	NADPH oxidases, such as NOX3, are plasma membrane-associated enzymes found in many cell types. They catalyze the production of superoxide by a 1-electron reduction of oxygen, using NADPH as the electron donor.
<i>SLC22A2</i> (OCT2) 6q26	6582 / [MIM:602608]	No	Solute carrier family 22 (organic cation transporter), member 2	Polyspecific transporter of organic cations, mainly expressed in kidney (luminal membrane of distal tube).
<i>SNX9</i> 6q26	51429 / [MIM:605952]	No	Sorting nexin 9	This gene encodes a member of the sorting nexin family. Members of this family contain a phox (PX) domain, which is a phosphoinositide binding domain, and are involved in several stages of intracellular trafficking.
<i>TFB1M</i> 6q26	51106 / [MIM:607033]	No	Transcription factor B1, mitochondrial	The transcription of genes from mitochondrial DNA requires a mitochondrial RNA polymerase and a DNA-binding transcription factor. Transcription factor B1 (TFB1M) is a part of this transcription complex and is implied in the rRNA processing.
<i>VIL2</i> 6q26	7430 / [MIM:123900]	No	Villin 2 (ezrin)	The cytoplasmic peripheral membrane protein that functions as a protein-tyrosine kinase substrate in microvilli. Serves as an intermediate between the plasma membrane and the actin cytoskeleton. Plays a key role in cell surface structure adhesion, migration, and organization.

<i>VIP</i> 6q26	7432 / [MIM:192320]	No	Vasoactive intestinal peptide	Secreted protein belongs to the glucagon family. It stimulates myocardial contractility, causes vasodilatation, increases glycogenolysis, lowers arterial blood pressure and relaxes the smooth muscle of trachea, stomach and gall bladder.
<i>ADAM8</i> 10q26.3	101 / [MIM:602267]	No	ADAM metallopeptidase domain 8	Membrane-anchored proteins structurally related to snake venom disintegrins expressed in granulocyte, monocyte, and macrophage. Implicated in a variety of biological processes involving cell-cell and cell-matrix interactions. May be involved in cell adhesion during neurodegeneration.
<i>CPXM2</i> 10q26.3	119587 / [MIM:NA]	Yes	Carboxypeptidase X, member 2	Implicated in cell adhesion and proteolysis.
<i>DOCK1</i> 10q26.3	1793 / [MIM:601403]	Yes	Dedicator of cytokinesis 1	Role in signaling from focal adhesions. This gene product binds to the SH3 domain of CRK protein. It may regulate cell surface extension and may have a role in the cell surface extension of an engulfing cell around a dying cell during apoptosis.
<i>FANK1</i> 10q26.3	92565 / [MIM:611640]	No	Fibronectin type III and ankyrin repeat domains 1	NA
<i>GPR123</i> 10q26.3	84435 / [MIM:NA]	No	G protein- coupled receptor 123	Orphan receptor : multi-pass membrane protein.
<i>MKI67</i> 10q26.3	4288 / [MIM:176741]	Yes	Antigen identified by monoclonal antibody Ki-67	Predominantly localized in the G1 phase in the perinucleolar region and in the nuclear matrix. In mitosis, present on all chromosomes. Thought to be required for maintaining cell proliferation. Regulation of progression through cell cycle.
<i>PTPRE</i> 10q26.3	5791 / [MIM:600926]	Yes	Protein tyrosine phosphatase, receptor type E	PTPs are known to be signaling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation.

Genes marked in bold indicate that they have been selected for the association study. NA = Not available.

\* Gene symbol, chromosomal location and GeneID number (Entrez Gene identifier in the NCBI database (<http://www.ncbi.nlm.nih.gov/>)).

† Appears in at least two G2D analyses.

‡ From NCBI (<http://www.ncbi.nlm.nih.gov/>) and GeneCards (<http://www.genecards.org/>) databases.

**6.1.12 Table 3**

Significant Family-Based Association Test (FBAT) results between the ten G2D candidate genes studied SNPs and asthma, atopy and allergic asthma phenotypes under an additive genetic model

Gene	SNP	Base change	MAF	Asthma			Atopy			Allergic Asthma		
				N	Z	P	N	Z	P	N	Z	P
<i>LPA</i>	rs12175867	T > C	0.16	81	1.98	0.0474	68	1.94	0.0530	71	2.04	0.0412
<i>ADAM8</i>	rs11101672	C > G	0.23	95	-1.76	0.0784	79	-2.06	0.0390	82	-2.12	0.0340
<i>DOCK1</i>	rs1051039	C > G	0.41	117	-1.97	0.0486	101	-0.72	0.4739	103	-0.48	0.6336
<i>GPR123</i>	rs11101913	C > T	0.29	112	2.63	0.0085	94	1.81	0.0707	97	2.04	0.0412
	rs11101916	G > A	0.18	95	-1.79	0.0731	73	-1.31	0.1916	80	-2.16	0.0308
	rs11101932	C > T	0.15	80	2.26	0.0240	60	1.74	0.0816	66	2.26	0.0397
	rs761777	A > G	0.29	113	-1.30	0.1948	92	-1.00	0.3152	99	-2.31	0.0209
	rs12257731	G > A	0.12	77	2.15	0.0314	60	1.65	0.0995	65	2.11	0.0353
<i>PTPRE</i>	rs11016002	A > T	0.35	113	-1.04	0.2983	89	-1.33	0.1836	93	-2.47	0.0136
	rs4369314	G > A	0.21	96	-2.57	0.0101	81	-1.65	0.0999	83	-1.67	0.0958
	rs4002572	C > T	0.45	113	-1.61	0.1067	86	-1.03	0.3047	93	-2.09	0.0367
	rs7081735	A > G	0.32	111	-2.23	0.0256	88	-2.09	0.0371	91	-3.50	<b>0.000463</b>

Abbreviations used: SNP = Single nucleotide polymorphism, MAF = Minor allele frequency, N = Number of families contributing to the statistic, Z = Z score, P = p-value (significance threshold of 0.000483; p-values under are marked in bold).



**6.1.13 Table 4***PTPRE* sequenced regions and characteristics of identified SNPs

Sequenced regions [bp] *	Identified SNP †	Base change	MAF (SLSJ) ‡	MAF (CEU) §	SNP location *
1 [75542 – 75977]	<i>No SNPs found</i>				
2 [85761 – 86205]	<b>rs7911506</b>	C > T	0.49	0.296	Intron 2
	<b>c.86172 A&gt;G</b>	A > G	0.31	NA	Intron 2
3 [88011 – 88457]	rs4369314	G > A	0.17	0.192	Intron 2
4 [92004 – 92507]	rs7081735	A > G	0.47	0.342	Intron 2
	<b>rs7895103</b>	T > C	0.05	0.217	Intron 2
5 [134156 – 134591]	<i>No SNPs found</i>				
6 [140684 – 141152]	rs4262642	C > A	0.19	0.25	Intron 3
	<b>c.140901 G&gt;A</b>	G > A	0.40	NA	Intron 3
	c.140903 G>A	G > A	0.03	NA	Intron 3
	c.141102 C>T	C > T	0.03	NA	Intron 4

SNPs marked in bold have been genotyped and analyzed.

Abbreviations used: bp = base pairs, SNP = single nucleotide polymorphism, MAF = Minor allele frequency, SLSJ = Saguenay-Lac-St-Jean, CEU = CEPH Utah residents with ancestry from northern and western Europe, NA = Not available.

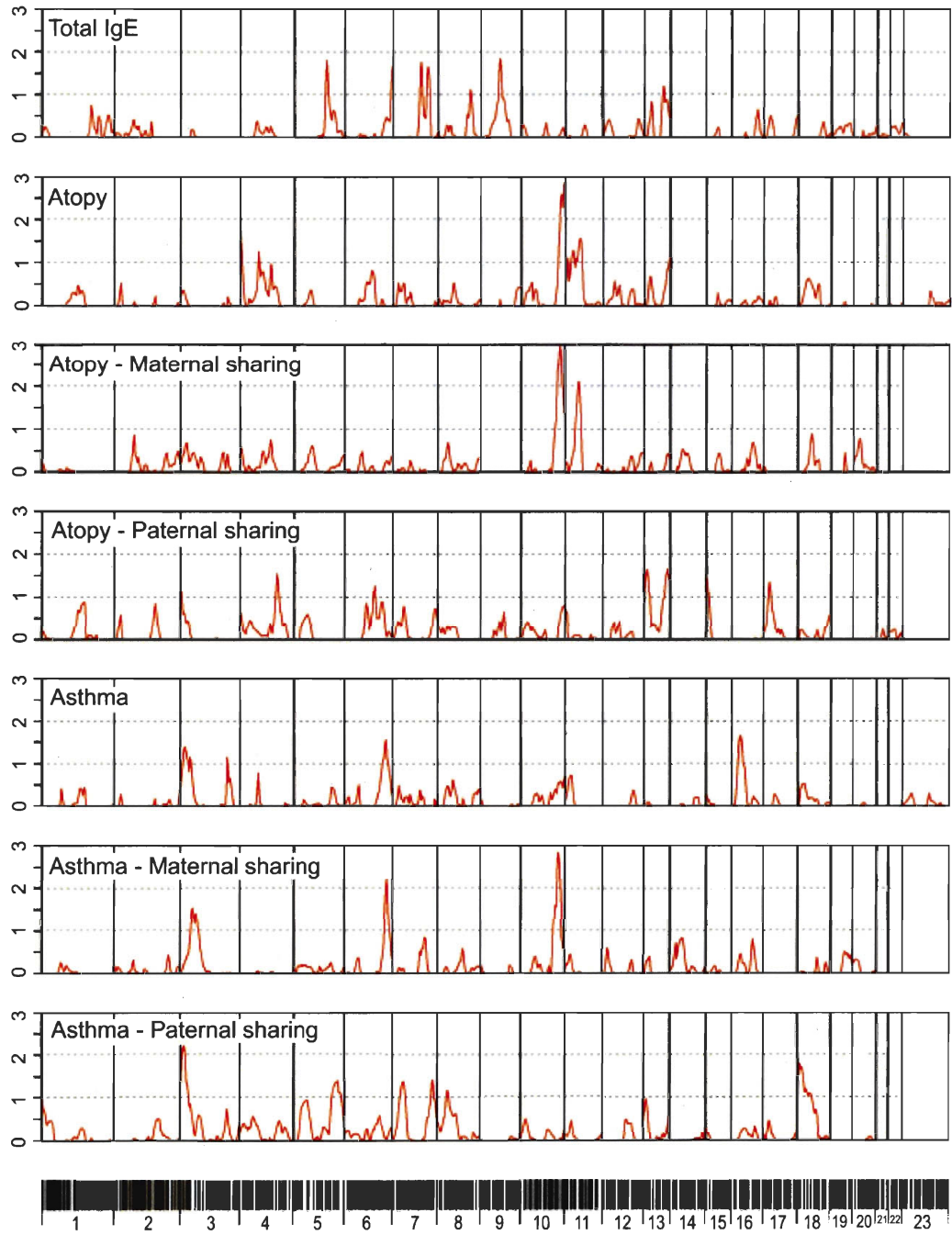
\* Base pair numbering according to NCBI (<http://www.ncbi.nlm.nih.gov/>) genome assembly (Build 36.2).

† If known from NCBI SNP and Ensembl (<http://www.ensembl.org/index.html>) databases, reported by the NCBI rs#.

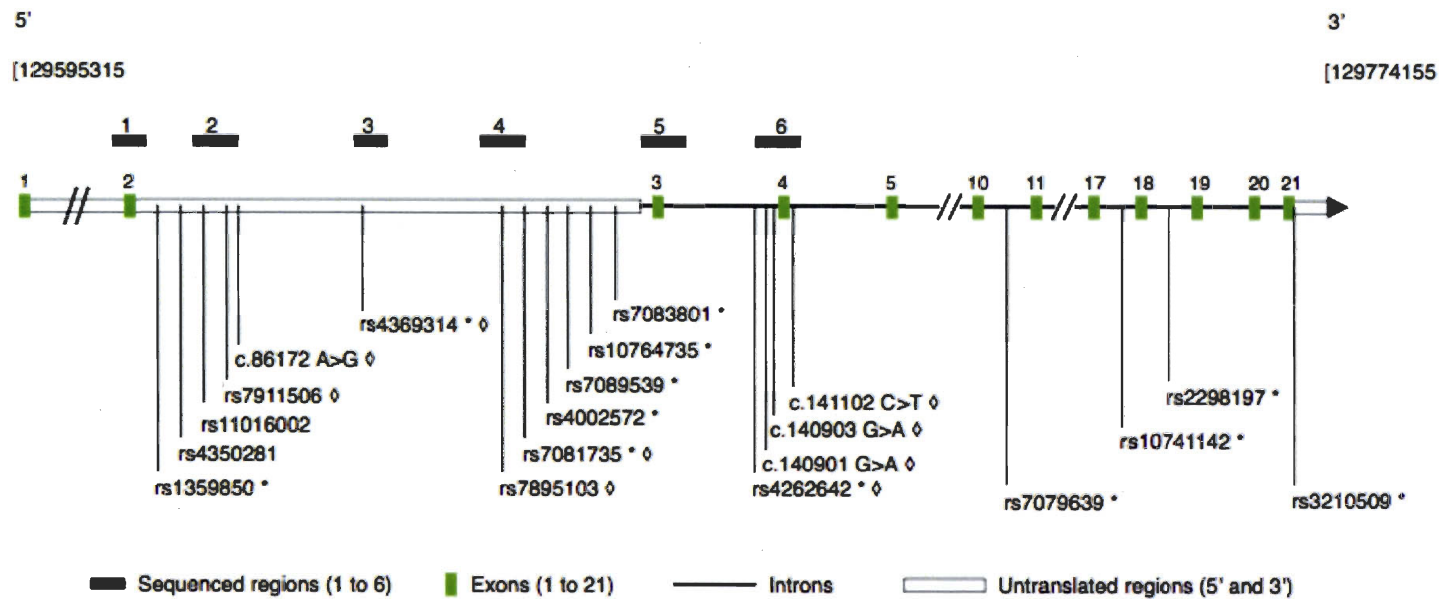
‡ Calculated on 39 individuals originating from the SLSJ population.

§ Obtained from the HapMap international project database (<http://www.hapmap.org/>).

6.1.14 Figure 1



6.1.15 Figure 2



## 6.1.16 SUPPORTING INFORMATION FILES

### **Genes to Diseases (G2D) Computational Method to Identify Asthma Candidate Genes**

Karine Tremblay, Mathieu Lemire, Camille Potvin, Alexandre Tremblay, Gary M Hunninghake, Benjamin A Raby, Thomas J Hudson, Carolina Perez-Iratxeta, Miguel A Andrade-Navarro and Catherine Laprise

#### **METHODS**

##### **Association study: Effective number of independent tests**

Since the SNPs are correlated, as are the three phenotypes studied, accounting for the number of tests using a simple Bonferroni correction is expected to be conservative. An effective number of independent SNPs can be calculated by the method proposed by Li and Ji [S1] implemented in SNPSdP [S2]: this number takes into account the observed  $r^2$  measure of pairwise linkage disequilibrium between all SNPs. Out of the three phenotypes investigated, an effective number of independent phenotypes was calculated using 10000 simulated replicates of the SLSJ sample of families: genotypes for ten SNPs with minor allele frequencies varying from 5% to 50% were simulated under the null hypothesis of no linkage and no association, for all members of the sample who have DNA. FastSLINK [S3-S5] and SUP [S6] were used to create the replicates. Each replicate of genotypes was then analysed for all three studied phenotypes, as they have been observed in the sample, and the minimum p-value observed over the four phenotypes was stored for each of the SNPs. The effective number of independent phenotypes was then taken to be the ratio between 0.05 and the quantile of order 5% of these minimum p-values.

#### **REFERENCES**

- S1. Li J, Ji L (2005) Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 95: 221-227.
- S2. Nyholt DR (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 74: 765-769.

- S3. Ott J (1989) Computer-simulation methods in human linkage analysis. *Proc Natl Acad Sci U S A* 86: 4175-4178.
- S4. Weeks DE, Lathrop GM, Ott J (1993) Multipoint mapping under genetic interference. *Hum Hered* 43: 86-97.
- S5. Cottingham RW, Jr., Idury RM, Schaffer AA (1993) Faster sequential genetic linkage computations. *Am J Hum Genet* 53: 252-263.
- S6. Lemire M (2006) SUP: An extension to SLINK to allow a larger number of marker loci to be simulated in pedigrees conditional on trait values. *BMC Genet* 7: 40.

**Table S1**

Genome-wide scan Saguenay–Lac-St-Jean subjects clinical characteristics and studied phenotypes

	Families	
	Probands (n = 79)	Family Members (n = 530)
Male: Female ratio	1: 1.2	1: 1.2
Mean age in years (range)	17.8 (4-62)	40.7 (4-88)
Smoking status (n (%))		
Never	68 (86)	242 (52)
Ex-smoker	6 (8)	129 (28)
Smoker	5 (6)	91 (19)
FEV <sub>1</sub> as % predicted (SD) *	96.5 (13.1)	97.6 (20.9)
PC <sub>20</sub> in mg/ml (SD) †	2.6 (3.8)	9.4 (5.1)
Serum IgE in µg/L (SD) †	228.0 (3.9)	122.7 (4.4)
Number of Persons with Subphenotypes (%)		
Asthma ‡	79 (100)	189 (40)
Atopy §	68 (86)	249 (54)
AHR ¶	64 (81)	165 (52)
IgE ** > 100 µg/L	58 (73)	190 (53)
IgE > 280 µg/L	38 (48)	107 (30)

\* FEV<sub>1</sub> = Forced expiratory volume in one second.† Geometric mean; PC<sub>20</sub> = provocative concentration of methacholine inducing a 20% fall in FEV<sub>1</sub>.

‡ 466 subjects with clinical status documented for asthma.

§ 462 with skin prick test performed.

¶ AHR = Airway hyperreactivity according to American Thoracic Society criteria (&lt; 8 mg/ml) Methacholine test was performed for 316 subjects (age ≥ 12 years).

\*\* IgE level was measured for 356 subjects.

**Table S2**  
Characteristics of the 91 selected SNPs

Gene *	SNP	Chromosome position †	Genomic location †	Techno ‡	TagSNP Selection §	HWE	ME
<i>LPA</i> NM_005577.2	rs7770628	6:160938164	Intron 20	S	Yes	0.109	0
	rs12175867	6:160939128	Intron 20	S	Yes	0.584	1
	<b>rs9355817</b>	6:160953578	Intron 16	S	Yes	Failed	NA
	rs12526465	6:160988597	Intron 1	S	Yes	0.531	0
	rs1652507	6:161002451	Intron 0	S	Yes	1.0	0
	rs1367211	6:161002685	Intron 0	S	Yes	0.596	0
	rs1321195	6:161004146	Intron 0	S	Yes	0.718	0
	<b>rs9346833</b>	6:161004632	Intron 0	S	Yes	Failed	NA
<i>NOX3</i> NM_015718.1	rs12665231	6:155758289	Exon14	S	No	0.440	0
	rs231951	6:155791096	Intron 9	S	Yes	0.755	0
	rs231959	6:155797666	Intron 7	S	Yes	0.525	0
	rs3749930	6:155802938	Exon 6 - K171T	S	No	1.0	0
	rs2235673	6:155812124	Intron 4	S	Yes	1.0	0
	rs7754577	6:155812942	Intron 4	S	Yes	1.0	0
	rs1016259	6:155816650	Intron 3	S	Yes	0.900	0
	rs9480140	6:155818406	Intron 1	S	Yes	0.689	0
<i>SNX9</i> NM_016224.3	rs2490565	6:158246186	Intron 6	S	Yes	0.881	0
	rs3805780	6:158249963	Intron 7	S	Yes	0.494	0
	rs9456870	6:158263794	Intron 10	S	Yes	0.090	0
	rs9365587	6:158264825	Intron 10	S	Yes	0.286	0
	rs11759002	6:158265875	Intron 10	S	Yes	0.368	0
	rs11755474	6:158275006	Intron 13	S	Yes	0.635	0
	rs6455919	6:158279463	Intron 15	S	Yes	0.329	0
	rs912879	6:158281887	Intron 16	S	Yes	0.388	0
<i>VIL2</i> NM_003379.3	rs3102974	6:159112907	Intron 7	S	Yes	0.082	0
	rs744893	6:159122228	Intron 6	S	Yes	0.749	0
	<b>rs11550707</b>	6:159126572	Exon 4 - L75P	S	No	NP	NA
	rs756144	6:159127578	Intron 3	T	Yes	0.933	1
	rs3123116	6:159130114	Intron 2	S	Yes	0.742	1
	rs9295086	6:159134798	Intron 1	T	Yes	0.448	0
	rs7754951	6:159151086	Intron 1	T	Yes	0.607	0
	rs1994350	6:159158992	Intron 1	T	Yes	0.041	1
<i>VIP</i> NM_003381.2	rs1282449	6:153104299	Upstream 5' UTR	T	No	0.225	0
	rs2756117	6:153108601	Upstream 5' UTR	T	No	0.219	0
	rs7764067	6:153115892	Intron 2	T	No	0.476	0
	rs3823082	6:153116015	Intron 2	T	No	0.056	0
	rs688136	6:153121754	Exon 7	T	No	0.689	1
	rs671330	6:153126927	Downstream 3' UTR	T	No	0.053	1
<i>ADAM8</i> NM_001109.3	rs2230576	10:134926587	Exon 23 - A846A	S	Yes	0.234	0
	rs11101672	10:134931744	Intron 19	S	Yes	0.274	0
	rs2275720	10:134932336	Exon 19 - L657F	S	No	0.090	0
	rs1131718	10:134935744	Exon 10 - T300T	S	Yes	0.092	0
	<b>rs3810960</b>	10:134937295	Exon 5 - L115P	S	No	NP	NA
	rs2275725	10:134939025	Exon 2 - R35W	S	Yes	0.108	0

<i>DOCK1</i> NM_001380.3	rs7092254	10:128582862	Upstream 5' UTR	S	No	0.035	1
	<b>rs2720994</b>	10:128979119	Intron 27	S	Yes	Failed	NA
	rs4751240	10:129026399	Intron 27	S	Yes	1.0	0
	rs2296635	10:129050448	Exon 32	S	Yes	1.0	0
	rs7899059	10:129075564	Intron 36	S	Yes	0.107	0
	rs2229599	10:129097419	Exon 40	S	Yes	1.0	0
	<b>rs2296626</b>	10:129106789	Exon 44	S	Yes	NP	NA
	rs869801	10:129135674	Exon 50	S	Yes	1.0	0
	rs2229603	10:129139652	Exon 51	S	Yes	1.0	0
rs1051039	10:129140593	3'UTR	S	Yes	0.155	0	
<i>FANK1</i> NM_145235.3	<b>rs4010025</b>	10:127575163	Exon1	S	No	Failed	NA
	<b>rs4962499</b>	10:127623699	Intron 1	S	Yes	Failed	NA
	rs10751537	10:127637438	Intron 1	S	Yes	0.836	0
	<b>rs7894453</b>	10:127639645	Intron 1	S	Yes	0.555	4
	rs10901491	10:127650767	Intron 1	S	Yes	0.584	0
	rs7918092	10:127652925	Intron 1	S	Yes	0.920	0
	rs11244753	10:127672950	Intron 3	S	Yes	0.961	0
	rs3736477	10:127675081	Intron 4	S	Yes	0.720	0
	rs4962316	10:127677666	Intron 6	S	Yes	0.697	0
	<b>rs17153976</b>	10:127687987	Exon 11 - F343C	S	No	NP	NA
<i>GPR123</i> NM_001083909.1	rs7912475	10:134739856	Intron 2	S	Yes	0.302	0
	rs4838694	10:134748584	Intron 7	S	Yes	0.395	0
	rs11101913	10:134757229	Intron 9	S	Yes	1.0	0
	rs11101916	10:134762304	Intron 12	S	Yes	0.443	0
	rs11101925	10:134763348	Intron 12	S	Yes	0.255	0
	rs9419004	10:134764512	Intron 12	S	Yes	0.279	0
	rs9419103	10:134765822	Intron 12	S	Yes	0.901	0
	rs736017	10:134767287	Intron 13	S	Yes	0.331	0
	rs11101932	10:134768842	Intron 13	S	Yes	1.0	0
	rs4838797	10:134769155	Intron 13	S	Yes	0.239	0
	rs761777	10:134788065	Intron 13	S	Yes	0.729	0
	rs12257731	10:134791129	Intron 14	S	Yes	0.520	0
	<i>PTPRE</i> NM_006504.3	rs1359850	10:129675326	Intron 1	S	Yes	0.709
rs4350281		10:129679923	Intron 1	S	No	0.344	0
rs11016002		10:129680794	Intron 2	S	No	0.331	0
rs4369314		10:129683048	Intron 2	S	Yes	0.385	0
rs7081735		10:129687046	Intron 2	S	Yes	0.195	0
rs4002572 (or rs4281406)		10:129690845	Intron 2	S	Yes	0.714	0
<b>rs4073016</b>		10:129691719	Intron 2	S	Yes	Failed	NA
rs7089539		10:129706236	Intron 2	S	Yes	0.304	0
rs10764735		10:129715444	Intron 2	S	Yes	0.211	0
rs7083801		10:129725918	Intron 2	S	Yes	0.574	0
rs4262642		10:129735537	Intron 3	S	Yes	0.122	0
rs7079639		10:129751782	Intron 10	S	Yes	0.126	0
rs10741142		10:129763083	Intron 17	S	Yes	0.401	0
rs2298197		10:129765704	Intron 18	S	Yes	0.105	0
rs3210509		10:129771768	Exon 21	S	Yes	0.401	0

SNPs marked in bold have not been analyzed for the association study.

Abbreviations used: SNP = Single nucleotide polymorphism; HWE = Hardy-Weinberg Equilibrium (Haploview P-values); ME = Mendelian errors reported in FBAT analyses (number); NA = Not available.

\* Gene symbol and the transcript variant used (NM number).

† From NCBI (<http://www.ncbi.nlm.nih.gov/>) SNP database (build 127).

‡ Technology used : S = Sequenom, T = Taqman.

§ Yes = tagSNP selection, No ≠ tagSNP selection.



**Table S3**  
Taqman PCR conditions

Gene	SNPs	Assay ID *	PCR	
			Mix †	Cycling conditions
VIL2 (SLSJ)	rs756144	AssayID C_8305215_10	10 ng of genomic DNA, 0.2 units of AmpliTaq Gold™ DNA Polymerase, 0.5 µl of 10x PCR Gold Buffer, concentration of MgCl <sub>2</sub> ranging from 1.25 to 4 mM depending on the SNP, 0.25 mM of dNTPs, 0.15 µl of 20X TaqMan® SNP Genotyping Assay mix, and water for a final volume of 5 µl.	95°C 10 min, 45 to 50 cycles of 92°C 15 sec and 58°C 1 min.
	rs9295086	AssayID C_29990632_10		
	rs7754951	AssayID C_11551998_10		
	rs1994350	AssayID C_12033615_10		
VIP (SLSJ)	rs1282449	AssayID C_462198_10	6.25 ng of genomic DNA, 2.5 µl of Applied Biosystems 2x Universal Master Mix, 0.125 µl of 40X TaqMan® SNP Genotyping Assay mix, and water for a final volume of 5 µl.	50°C 2 min, 95°C 10 min, 50 cycles of 95°C 15 sec and 60°C 1 min.
	rs2756117	AssayID C_16286919_10		
	rs7764067	AssayID C_29430252_10		
	rs3823082	AssayID C_27491244_10		
	rs688136	AssayID C_3250639_10		
PTPRE (CAMP)	rs7081735	AssayID C_29345951_10		

\* Refers to the Taqman kit identification number supplied by the company.

† From Applied Biosystems Inc.

**Table S4**  
Oligonucleotides used for *PTPRE* sequencing

Oligo name *	Covered region	Product size	Sequence (5' – 3') †
PTPRE_Exon2_L	Exon 2	436	CACAGCAGAGGCAGCTGAC
PTPRE_Exon2_R			CCTGCTCAAGCAAGGTAAGG
PTPRE_rs11016002_L	Around rs11016002	445	GGGAAAACAACATGACCAG
PTPRE_rs11016002_R			GTGTCTGGCTTCTGATGCA
PTPRE_rs4369314_L	Around rs4369314	447	AGAACGCAGGCCCATAG
PTPRE_rs4369314_R			TTGGATTGGTGGGTACATCC
PTPRE_rs7081735_L	Around rs7081735	504	GGCTTCCATTTCTGTCTGG
PTPRE_rs7081735_R			CCCCTTCTACAGCACTTCCTAA
PTPRE_Exon3_R	Exon 3	436	TGGGTTCCAGTCCCTGATC
PTPRE_Exon3_L			AGCCCCTTTATCACCCCTCTG
PTPRE_Exon4_L	Exon 4	468	GGGAACTCACCCCTGTCAT
PTPRE_Exon4_R			CCTGGCCTCTGTGTGTCTC

\* Annealing temperatures for each oligo pair is 64°C.

† PCR were done in 25µL volume containing of 0.6 unit of Taq DNA polymerase (New England Biolabs), 50µM of each dNTPs, 20 mM Tris-HCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM KCl, 2 mM MgSO<sub>4</sub>, 0.1 % Triton, pH 8.8, 0.4µM of each primer and 10 ng of human genomic DNA amplified with the following conditions: 95°C 15 min followed by 35 cycles of 94°C 30 sec, 59-66°C 40 sec and 72°C 1 min, then 72°C 5 min.

**Table S5**

Genes selected by G2D « PHENOTYPE » analysis

The complete list of GO scores associated by G2D to asthma phenotype can be found at [http://www.ogic.ca/projects/g2d\\_2/asthma\\_go\\_phenotype.html](http://www.ogic.ca/projects/g2d_2/asthma_go_phenotype.html).

Table S5-A List of 16 genes proposed by G2D in region 6q26. Table S5-B List of 12 genes proposed by G2D in region 10q26.3.

## Table S5-A

OPRM1  
MAS1  
NOX3  
SOD2  
SLC22A3  
TIAM2  
ESR1  
IGF2R  
SYTL3  
TFB1M  
RGS17  
AKAP12  
TULP4  
SNX9  
MYCT1  
ARID1B

## Table S5-B

DOCK1  
PTPRE  
EBF3  
C10orf39  
GPR123  
KNDC1  
ADAM12  
CPXM2  
FANK1  
OAT  
INPP5A  
MKI67

**Table S6**

Genes known or suspected to be associated with asthma that were considered for the G2D «KNOWN GENES» analysis

Gene ID *	Symbol †	Description	GO annotation ‡
387129	GPRA	G protein-coupled receptor 154	No
404744	AAA1	Asthma associated protein (MMystery protein)	No
3596	IL13	Interleukin 13	Yes
154	ADRB2	adrenergic, beta-2-, receptor	Yes
3593	IL12B	interleukin 12B (natural killer cell stimulatory factor 2, cytotoxic lymphocyte maturation factor 2, p40)	Yes
7356	SCGB1A1	secretoglobin, family 1A, member 1	Yes
7124	TNF	tumor necrosis factor (TNF superfamily, member 2)	Yes
51131	PHF11	PHD finger protein 11	No
117156	SCGB3A2	secretoglobin, family 3A, member 2	Yes
2206	MS4A2	membrane-spanning 4-domains, subfamily A, member 2	Yes
5729	PTGDR	prostaglandin D2 receptor (DP)	Yes
4056	LTC4S	leukotriene C4 synthase	Yes
3567	IL5	interleukin 5 (colony-stimulating factor, eosinophil)	Yes
2944	GSTM1	glutathione S-transferase M1	Yes
80332	ADAM33	ADAM metallopeptidase domain 33	Yes
3565	IL4	interleukin 4	Yes
3566	IL4R	interleukin 4 receptor	Yes
57628	DPP10	dipeptidyl-peptidase 10	Yes
727	C5	complement component	Yes

\* Entrez Gene identifier in the NCBI database (<http://www.ncbi.nlm.nih.gov/>).

† Gene symbol.

‡ Genes not annotated with GO terms could not be used for this analysis.

**Table S7**

Genes selected by G2D « KNOWN GENES » analysis

The complete list of GO scores associated by G2D to asthma known compiled genes can be found at [http://www.ogic.ca/projects/g2d\\_2/asthma\\_go\\_known.html](http://www.ogic.ca/projects/g2d_2/asthma_go_known.html).

Table S7-A List of 15 genes proposed by G2D in region 6q26. Table S7-B List of 10 genes proposed by G2D in region 10q26.3.

## Table S7-A

RGS17  
MAP3K4  
OPRM1  
NOX3  
VIL2  
MAS1  
LPA  
TULP4  
SOD2  
MYCT1  
TIAM2  
ESR1  
IGF2R  
PARK2  
SLC22A2

## Table S7-B

DOK1  
PPP2R2D  
ADAM12  
ADAM8  
CPXM2  
INPP5A  
CTBP2  
MKI67  
PTPRE  
DPYSL4

## Chapitre 7

### **Gène ciblé en raison de son association génétique connue avec l'asthme**

L'étude d'association présentée dans ce chapitre consiste en une étude de réplication des SNPs du gène *ORMDL3* associés à une susceptibilité à l'asthme chez les enfants dans une GWAS (Moffatt et al., 2007). Il s'agit donc de la sélection d'un gène candidat à partir d'une association connue *a priori* dans l'asthme. Dans le cadre de cette étude publiée dans le journal *Human Genetics*, Anne-Marie Madore, étudiante au doctorat, a effectué le génotypage des SNPs, les analyses statistiques et la rédaction du manuscrit. Pour ma part, j'ai validé les analyses statistiques, contribué à la rédaction et à la révision du manuscrit. Le Dr Hudson a révisé le manuscrit. La Dre Laprise a conçu l'étude, dirigé chacune des étapes et révisé le manuscrit.

## 7.1 Sixième article

### **Réplication d'une association entre les SNPs du 17q21 et l'asthme dans un échantillon familial Canadien-Français**

Anne-Marie Madore, Karine Tremblay, Thomas J Hudson et Catherine Laprise

*Human Genetics* (2008) 123(1): 93-5

#### **Résumé**

L'asthme est un trait complexe sous l'influence des facteurs génétiques et environnementaux. Dans une récente GWAS, réalisée avec des sujets asthmatiques d'origine européenne, de nouvelles associations génétiques ont été rapportées entre les polymorphismes d'un seul nucléotide (SNPs) du chromosome 17q21 et l'asthme chez l'enfant. Afin de répliquer ces associations, nous avons effectué une analyse génétique avec les dix meilleurs SNPs (qui présentaient les plus fortes associations statistiques) dans l'échantillon familial du SLSJ. Nos analyses statistiques ont révélé des associations positives pour huit SNPs ( $0,005 < \text{valeurs-p} < 0,017$ ) et pour deux haplotypes ( $\text{valeurs-p} = 0,0004$  et  $0,002$ ). Ces résultats confirment le 17q21 comme un locus d'asthme.

### 7.1.1 TITLE PAGE

#### **Replication of an association between 17q21 SNPs and asthma in a French-Canadian familial collection**

Anne-Marie Madore<sup>1,2,3</sup>, Karine Tremblay<sup>1,2,3</sup>, Thomas J. Hudson<sup>4</sup>, Catherine Laprise<sup>2,3</sup>

<sup>1</sup> Laval University, Department of Medicine, Québec (Québec), Canada; <sup>2</sup> University of Montreal Community Genomics Medicine Centre, Chicoutimi Hospital, Saguenay (Québec), Canada; <sup>3</sup> Université du Québec à Chicoutimi, Department of Fundamental Sciences, Saguenay (Québec), Canada; <sup>4</sup> Ontario Institute for Cancer Research, Toronto (Ontario), Canada.

#### **Corresponding author**

Catherine Laprise, PhD,  
Université du Québec à Chicoutimi,  
555 blv de l'Université, Chicoutimi (Québec), Canada, G7H 2B1  
Telephone: (418) 545-5011, 5659  
Fax: (418) 545-5012  
Email: catherine\_laprise@uqac.ca

#### **Key words**

Asthma, SNP, 17q21, ORMDL3, family-based association

### 7.1.2 ABSTRACT

Asthma is a complex trait which is influenced by environmental and genetic factors. In a recent genome-wide association study for asthma in European subjects, novel associations were reported between 17q21 single nucleotide polymorphisms (SNPs) and childhood asthma. We performed an association study with the ten SNPs that showed the strongest association in a French-Canadian asthmatic familial collection. Family-based association tests revealed significant associations for eight SNPs ( $0.005 < p < 0.017$ ) and for two haplotypes ( $p = 0.0004$  and  $0.002$ ), confirming the 17q21 as an asthma locus.



### 7.1.3 INTRODUCTION

Several association studies have been performed to characterize the genetic etiology of asthma (for a review see (Ober et al. 2006)). A genome-wide association study for childhood asthma reported strong associations for 63 single nucleotide polymorphisms (SNPs) at the 17q21 locus, and an association between these SNPs and *ORMDL3* transcript abundance, a gene encoding a transmembrane protein anchored in the endoplasmic reticulum (Hjelmqvist et al. 2002; Moffatt et al. 2007). Based on these observations, we aimed to replicate the association of the ten SNPs showing the strongest associations with asthma in a French-Canadian asthmatic familial collection and found significant associations for eight SNPs and two haplotypes.

### 7.1.4 METHODS

**Samples:** The French-Canadian familial collection comprises 253 asthmatic probands for a total of 1,275 subjects. A proportion of 36.6% of relatives are affected by asthma. For phenotypic characterization of the subjects, see the online supplement.

**Genotyping:** Blood samples were drawn from all participants and DNA was extracted to be genotyped with the TaqMan® SNP Genotyping Assays (Applied Biosystems, Foster City, USA) or the Sequenom® technology (Sequenom Inc., San Diego, USA). See the online supplement for further details.

**Statistical analysis:** Mendelian errors and allele distribution patterns were assessed by FBAT software (version 1.7) under an additive model using the empirical variance estimator “-e” (Lake et al. 2000). Hardy-Weinberg equilibrium (HWE) for each SNP and strength of linkage disequilibrium (LD) between pairs of SNPs, measured as D prime (D’), were done using Haploview software (3.31 version) (Barrett et al. 2005).

### 7.1.5 RESULTS AND DISCUSSION

One SNP failed genotyping (rs8069176). The genotyping mean success rate for the remaining SNPs was 99.8%. Six mendelian errors were found across the entire dataset, leading to the exclusion of five families in the analyses. No significant deviation from HWE was detected. As observed in the study by Moffatt and coworkers, LD has been

found between rs2290400 and rs7219923 but not for rs11557467. These SNPs were also in LD with rs2305479, rs9303281 and rs7216389 ( $D' > 0.97$ ) (see the online Figure E1). We evaluated the number of effective independent SNPs and considering the use of one phenotype, one genetic model and three effective SNPs, we fixed the significance threshold at 0.017 to keep the type I error rate at a 5% level (Nyholt 2004).

Table 1 presents the gene symbol, position and chromosomal location for each SNP and FBAT association for asthma and the nine SNPs genotyped. The association observed by Moffatt and coworkers was replicated for eight of them. In all cases, the reference allele was overtransmitted to the affected offspring ( $0.005 < p < 0.017$ ), suggesting a susceptibility effect for asthma. Two haplotypes were also associated ( $p = 0.0004$  and  $0.002$ ). These results were observed despite the differences in sample origin (English and Germanic versus French-Canadian) as well as ascertainment (childhood versus persistent asthma). This replication study highlights the importance of the 17q21 locus in asthma.

#### 7.1.6 ACKNOWLEDGMENTS

We express our gratitude to all probands and their family for participating in the study. We thank S. Claveau for his help in genotyping. A.-M. Madore is supported by the Canadian Institutes of Health Research (CIHR). K. Tremblay is supported by a grant from the Laval University Foundation. T.J. Hudson was supported by a Clinician-Scientist Award in Translational Research from the Burroughs Wellcome Fund. C. Laprise is the chairholder of the Canada Research Chair for genetic determinants in asthma ([www.chairs.gc.ca](http://www.chairs.gc.ca)) and the director of the Genetics & Genomics Thematic Unit of the Respiratory Health Network of the *Fonds de la Recherche en Santé du Québec* (FRSQ).

#### 7.1.7 REFERENCES

Barrett JC, Fry B, Maller J and Daly MJ (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21(2): 263-5.

- Hjelmqvist L, Tuson M, Marfany G, Herrero E, Balcells S and Gonzalez-Duarte R (2002). ORMDL proteins are a conserved new family of endoplasmic reticulum membrane proteins. *Genome Biol* 3(6): RESEARCH0027.
- Lake SL, Blacker D and Laird NM (2000). Family-based tests of association in the presence of linkage. *Am J Hum Genet* 67(6): 1515-25.
- Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, Depner M, von Berg A, Bufe A, Rietschel E, Heinzmann A, Simma B, Frischer T, Willis-Owen SA, Wong KC, Illig T, Vogelberg C, Weiland SK, von Mutius E, Abecasis GR, Farrall M, Gut IG, Lathrop GM and Cookson WO (2007). Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 448(7152): 470-3.
- Nyholt DR (2004). A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 74(4): 765-9.
- Ober C and Hoffjan S (2006). Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 7(2): 95-100.

### 7.1.8 Table 1

FBAT results between the nine studied SNPs, the two common haplotypes and asthma in the French-Canadian familial collection

SNPs / Haplotypes*	Gene symbol / position	Chromosomal location (bases) <sup>†</sup>	Allele <sup>‡</sup>	Allele frequency	No. family	S	E(S)	Z score	P-value <sup>£</sup>
SNPs									
rs11557467	ZBPB2 / exon 4	35,282,160	<b>G</b>	0.607	152	307	280	2.568	0.010
			T	0.393		185	212	-2.568	
rs2305479	GSDML / exon 6	35,315,743	<b>A</b>	0.385	116	136	160	-2.817	0.005
			<b>G</b>	0.615		244	220	2.817	
rs2290400	GSDML / intron	35,319,766	<b>G</b>	0.382	150	176	201	-2.607	0.009
			<b>A</b>	0.618		312	287	2.607	
rs7216389	GSDML / intron	35,323,475	<b>C</b>	0.422	154	194	218	-2.448	0.014
			<b>T</b>	0.578		300	276	2.448	
rs9303281		35,327,572	<b>A</b>	0.590	120	239	215	2.795	0.005
			<b>G</b>	0.410		149	173	-2.795	
rs7219923		35,328,044	<b>C</b>	0.410	121	152	176	-2.827	0.005
			<b>T</b>	0.590		240	216	2.827	
rs4378650	ORMDL3 / intron	35,334,391	<b>C</b>	0.602	116	233	213	2.392	0.017
			<b>T</b>	0.398		145	165	-2.392	
rs8076131	ORMDL3 / intron	35,334,438	<b>A</b>	0.633	117	246	225	2.631	0.009
			<b>G</b>	0.367		128	149	-2.631	
rs6503525		35,348,700	<b>C</b>	0.530	110	183	170	1.603	NS
			<b>G</b>	0.470		163	176	-1.603	
Haplotypes									
H1	[rs2305479G; rs2290400T; rs7216389T; rs9303281A; rs7219923T]			0.579	137	249	219	3.548	0.0004
H2	[rs2305479A; rs2290400C; rs7216389C; rs9303281G; rs7219923C]			0.362	127	131	157	-3.110	0.002

Abbreviations used: SNP = Single nucleotide polymorphism, S = Association test statistic, E(S) = Expected statistic, ZBPB2 = zona pellucida binding protein 2, GSDML = gasdermin-like, ORMDL3 = orosomucoid I (ORM1)-like 3, NS = non significant.

\* Common haplotypes (H1 and H2) as defined by the LD block (see supplement Figure E1)

<sup>†</sup> According to National Center for Biotechnology Information public database build 128 (<http://www.ncbi.nlm.nih.gov>).

<sup>‡</sup> Alleles underlined in bold indicate the major allele

<sup>£</sup> P-value significance threshold of 0.017

## 7.1.9 ONLINE SUPPLEMENT

### **Replication of an association between 17q21 SNPs and asthma in a French-Canadian familial collection**

Anne-Marie Madore, Karine Tremblay, Thomas J. Hudson, Catherine Laprise

#### **METHODS**

##### **Samples**

The sample comprises 253 probands for a total of 1,275 subjects. This sample and all information concerning recruitment and clinical evaluation of the subjects were described in a recent report (Begin et al. 2007). Briefly, asthma phenotype was characterized by the investigators for all participants using a respiratory health questionnaire and function tests following American Thoracic Society standards (American Thoracic Society, 1987). Participants were defined as having asthma if (1) they had a reported history of asthma (validated by a physician), or (2) they showed asthma-related symptoms and a positive PC<sub>20</sub> (provocative concentration causing a 20% drop in FEV<sub>1</sub> [PC<sub>20</sub>] < 8 mg/ml (American Thoracic Society, 1987)) at the time of recruitment. Subjects with a PC<sub>20</sub> greater than 8 mg/ml; without history of physician-diagnosed asthma, and without both symptoms of asthma, and a PC<sub>20</sub> greater than 8 mg/ml, were considered unaffected. See Table E1 for the clinical characteristics of the subjects.

##### **Genotyping**

Blood samples were drawn from all participants and DNA was extracted from blood leukocytes using Genomic-tip 100/G kit (Qiagen, Inc., Valencia, CA) according to manufacturer's instructions. Genotyping of the rs11557467, rs2290400, and rs7216389 SNPs was performed with TaqMan® SNP Genotyping Assays (Applied Biosystems, Foster City, CA) using the Rotor-Gene 3000 (Corbett Research, Sydney, Australia) with 10 ng of genomic DNA, 2.15 µl of 2X TaqMan® Genotyping Master Mix and 0.11 µl of 20X TaqMan® SNP Genotyping Assay (C\_9272244\_10 , C\_15884536\_10 and C\_29062108\_10 respectively (Applied Biosystems, Foster City, CA)) for a final volume of 4.3 µl with the following cycling conditions: 95°C 10 min, 40 cycles of 92°C 15 sec

and 58°C 1 min. The Rotor-Gene software (version 6.0) was used to determine genotypes using the scatter graph analysis option.

Genotyping for the seven remaining SNPs (rs2305479, rs4378650, rs6503525, rs7219923, rs8069176, rs8076131, rs9303281) was done using the Sequenom® matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass array spectrometer (Sequenom Inc., San Diego, CA, USA). Primers were designed using the Sequenom SNP Assay Design software version 3.0 for iPLEX reactions. The protocol and reaction conditions are in accordance with the manufacturer (Oeth P et al. 2005). The genotypes were viewed and analyzed using the MassARRAY Typer software version 3.4 (Sequenom).

## REFERENCES

- American Thoracic Society (1987). Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. *Am Rev Respir Dis* 136(1): 225-44.
- Begin P, Tremblay K, Daley D, Lemire M, Claveau S, Salesse C, Kacel S, Montpetit A, Becker A, Chan-Yeung M, Kozyrskyj AL, Hudson TJ and Laprise C (2007). Association of urokinase-type plasminogen activator with asthma and atopy. *Am J Respir Crit Care Med* 175(11): 1109-16.
- Oeth P, Beaulieu M, Park C, Kosman D, del Mistro G, van den Boom D and Jurinke C. (2005). iPLEX assay: Increased plexing efficiency and flexibility for MassArray system through single base primer extension with mass-modified terminators. SEQUENOM Application Note.

## LEGENDS TO FIGURE

**Figure E1:** Pairwise linkage disequilibrium pattern of the nine 17q21 locus single nucleotide polymorphisms (SNPs).

The location of each tested SNP along the chromosome is indicated on top. The number in each diamond indicates the magnitude of linkage disequilibrium ( $D'$ ) between respective pairs of SNPs (e.g., the pairwise magnitude of linkage disequilibrium for variants rs9303281 and rs7219923 is 0.99). Diamonds without numbers represent perfect linkage disequilibrium ( $D'=1.0$ ).

**Table E1**

Clinical characteristics of the French-Canadian familial collection subjects

	Families		
	Probands n = 253	Affected relatives n = 379	Unaffected relatives n = 643
Male : Female ratio	1 : 1.12	1 : 1.27	1 : 1.10
Mean age, yr (range)	17 (2-50)	39 (2-83)	46 (3-93)
Median age, yr	15	40	46
Smoking status, n (%)			
Never	212 (85)	192 (51)	272 (43)
Ex-smoker	12 (5)	112 (30)	234 (37)
Smoker	25 (10)	71 (19)	128 (20)
FEV <sub>1</sub> *, % predicted (SD)	92.7 (16.0)	88.9 (22.4)	99.1 (17.2)
PC <sub>20</sub> <sup>†</sup> , mg/ml (SD)	2.50 (3.63)	3.35 (4.43)	21.36 (3.215)

Abbreviation used: FEV<sub>1</sub> = Forced expiratory volume in one second, PC<sub>20</sub> = Provocative concentration of methacholine that induces a 20% fall in FEV<sub>1</sub>, SD = Standard deviation.

\* Geometric mean calculated for 218 probands, 325 affected relatives and 440 unaffected relatives

† Geometric mean of the log-transformed PC<sub>20</sub> calculated for 193 probands, 285 affected relatives and 418 unaffected relatives





## Chapitre 8

### **Gène ciblé en raison de son association génétique connue avec un trait complexe partageant une composante physiopathologique avec l'asthme**

Pour ce qui est du septième article inclus dans ce chapitre, il s'agit d'une étude d'association entre le gène filaggrine (*FLG*) et les phénotypes liés à l'asthme dans l'échantillon de trios familiaux du SLSJ. La sélection de ce gène candidat s'est effectuée sur la base de son association génétique connue dans la dermatite atopique et de son rôle dans les mécanismes allergiques. Le manuscrit, dans sa forme actuelle, comprend l'étude d'association entre les TagSNPs du gène *FLG*, les phénotypes de l'asthme ainsi que les sous-phénotypes de dermatite atopique. Suite à l'association d'un des TagSNPs, le séquençage complet du gène a été réalisé par Cécilia Hernández et David Roquis pour identifier la ou les mutations causales. Pour cette étude, j'ai réalisé l'analyse d'association et l'analyse de la séquence, préparé les résultats et leur interprétation, et rédigé le manuscrit. Yohan Bossé a effectué le génotypage des TagSNPs et révisé le manuscrit. Les Drs Hudson et Laprise ont conçu l'étude et révisé le manuscrit dans sa forme actuelle. La Dre Laprise a dirigé toutes les étapes. Comme le séquençage a permis d'identifier 54 nouvelles variations génétiques, leur génotypage et de nouvelles analyses d'association sont requises. L'article sera soumis lorsque ces travaux seront terminés et ajoutés au manuscrit.

## 8.1 Septième article

### Association des variants du gène filaggrine (*FLG*) avec l'asthme et l'atopie dans l'échantillon familial du Saguenay-Lac-St-Jean

Karine Tremblay, Yohan Bossé, Cecilia Hernández, David Roquis, Thomas J. Hudson et Catherine Laprise

Article en préparation.

#### Résumé

Le gène *FLG* code pour une protéine structurale, la filaggrine, essentielle à l'agrégation des filaments de kératine dans l'épiderme et à l'intégrité de la barrière épidermique. Une production réduite de filaggrine est reconnue comme un facteur de prédisposition au développement de la dermatite atopique. Étant donné son implication dans la sensibilisation aux allergènes et considérant le fait que l'asthme et la dermatite atopique partagent une composante atopique, *FLG* a été sélectionné comme gène candidat pour une étude d'association avec l'asthme allergique. Les objectifs sont donc (1) d'effectuer une étude d'association entre des TagSNPs du gène *FLG* et l'asthme, l'atopie, l'asthme allergique ainsi que chacun de ces phénotypes apparaissant simultanément avec la dermatite atopique dans l'échantillon de trios familiaux du SLSJ; et (2) de séquencer le gène pour trouver les mutations causales. Cinq TagSNPs ont été génotypés dans 249 trios à l'aide d'un panel Sequenom® et les tests de déséquilibre de transmission ont été effectués. Le séquençage des exons 2 et 3 du gène *FLG* a été réalisé avec l'ADN de 48 sujets asthmatiques allergiques provenant du SLSJ. Des associations positives ont été identifiées pour rs2184951 et rs12730241 avec l'asthme ( $p = 0,014$  et  $0,011$ , respectivement), l'atopie ( $p = 0,025$  et  $0,020$ ), l'atopie apparaissant avec la dermatite atopique ( $p = 0,0097$  et  $0,0097$ ) et l'asthme allergique apparaissant avec la dermatite atopique ( $p = 0,036$  et  $0,036$ ). Dans tous les cas, les allèles mineurs sont moins transmis aux enfants atteints, suggérant un effet de protection. Le séquençage a permis d'identifier 54 nouvelles variations génétiques. Ces résultats, en plus de documenter de nouvelles variations génétiques du gène *FLG* et de montrer leur diversité au sein des populations, suggèrent une implication du gène *FLG* dans la pathogénèse de l'asthme, plus spécifiquement en lien avec sa composante atopique.

### 8.1.1 TITLE PAGE

#### **Sequencing the Filaggrin Gene in Atopic Asthmatic Subjects From the Saguenay–Lac-Saint-Jean Founder Population**

Karine Tremblay<sup>1,2,3</sup>, Yohan Bossé<sup>1,4</sup>, Cecilia Hernández<sup>4</sup>, David Roquis<sup>4</sup>, Thomas J. Hudson<sup>4,5</sup> and Catherine Laprise<sup>2,3</sup>

<sup>1</sup> Hôpital Laval, Université Laval, Québec, QC, G1V 4G5, Canada; <sup>2</sup> Département des sciences fondamentales, Université du Québec à Chicoutimi, Saguenay, QC, G7H 2B1, Canada; <sup>3</sup> University of Montreal Community Genomic Medicine Centre, Chicoutimi University Hospital, Saguenay, QC, G7H 5H6, Canada; <sup>4</sup> McGill University and Genome Quebec Innovation Centre, Montreal, QC, H3A 1A4, Canada; <sup>5</sup> Ontario Institute for Cancer Research, Toronto, ON, M5G 0A3, Canada

#### **Corresponding author**

Catherine Laprise, Ph.D.  
Université du Québec à Chicoutimi  
555, boulevard de l'Université,  
Chicoutimi (Qc), Canada, G7H 2B1  
Telephone: (418) 545-5011, 5659  
Fax: (418) 615-1203  
Email: catherine\_laprise@uqac.ca

#### **Running title**

Filaggrin Gene and Atopic Asthma

### 8.1.2 ABSTRACT

**Background:** The filaggrin (*FLG*) gene encodes a structural protein essential to aggregate keratin filaments into the epidermis. Reduced filaggrin production results in an impaired epidermal barrier formation, which is a strong predisposing factor for atopic dermatitis (AD). Due to its implication in allergic sensitization and considering that asthma shares an atopic component with AD, we selected *FLG* as a candidate gene for asthma disease, especially regarding atopic component of the disease. **Objectives:** To perform an association study between *FLG* tagging single nucleotide polymorphisms (tagSNPs) and asthma, atopy, atopic asthma as well as both phenotypes co-occurring with AD in a familial asthma sample from the Saguenay–Lac-Saint-Jean (SLSJ) founder population, and to sequence the gene to find causal mutations. **Methods:** Five tagSNPs (rs2184951, rs12730241, rs11582620, rs1933063 and rs1933064) covering the whole *FLG* gene have been genotyped in 249 trios using a Sequenom panel, and family-based association tests have been performed. Direct sequencing of the *FLG* exons 2 and 3 has been done on 48 SLSJ atopic asthmatic subjects using ten designed PCR fragments. **Results:** Positive associations were found for *FLG* rs2184951 and rs12730241 with asthma ( $p = 0.014$  and  $0.011$ , respectively), atopy ( $p = 0.025$  and  $0.020$ ), atopy co-occurring with AD ( $p = 0.0097$  and  $0.0097$ ) and atopic asthma co-occurring with AD ( $p = 0.036$  and  $0.036$ ). In all cases, minor alleles were undertransmitted to the affected probands, suggesting a genetic protective effect. The *FLG* sequencing allowed to the identification of 104 genetic variants, including 54 new synonymous and non-synonymous nucleotide substitutions specific to the SLSJ population and/or to the atopic asthma disease. **Conclusions:** These results, in addition to document new *FLG* genetic variants and to demonstrate their variability among diverse populations, may suggest an implication of *FLG* in asthma pathogenesis, specifically related to its atopic component.

### 8.1.3 INTRODUCTION

Asthma is a complex chronic respiratory disease that involves both environmental and genetic factors in its development, and that is characterized by inflammation and remodeling of the airways modulated by resident and inflammatory cells [1, 2]. Asthma also presents a strong atopic component characterized by the production of high levels

of immunoglobulin E in response to common allergens [3]. To date, asthma and atopy genetic studies identified several linked chromosomal loci and associated genes [4-6]. Nevertheless, finding of novel genetic predisposing factors to better understand the asthma natural history is an active research area.

The human filaggrin gene (*FLG*), originally described in ichthyosis vulgaris (IV) [7], encodes a skin structural protein essential to aggregate keratin filaments in the epidermis [8]. The common R501X and 2282del4 null *FLG* mutations have been shown to result in a complete loss of filaggrin production, which is responsible for the impaired epidermal barrier formation observed in IV disease [7]. A strong predisposing effect of these two *FLG* mutations has been reported for atopic dermatitis (AD), a complex skin disease sharing clinical phenotypes and genetic risk factors with IV [9]. Subsequent genetic studies in AD have shown *FLG* as the gene with the most widely replicated associations [10-19]. In asthma, the *FLG* genetic risk has been especially reported as co-occurring with AD, as a secondary trait [9, 11, 17, 18, 20-22].

The *FLG* gene is located on chromosome 1q21 within a cluster of genes of the epidermal differentiation complex and consists of the exon 1 (15 bp), the exon 2 (159 bp: containing the initiation codon) and the unusually large, highly repetitive exon 3 (12,753 bp: 10-12 ~1kb repeats) that encodes the bulk of the profilaggrin protein (> 435 kDa) [23]. As sequencing analysis of the exon 3 is problematic [24], PCR protocols allowing routine *FLG* sequencing have been developed [14] and frequently used in IV and AD. However, no complete sequencing of the *FLG* gene in asthma has been reported.

Based on these observations and in order to delineate the genetic effect of *FLG* in asthma disease, we performed an association study between five *FLG* tagging single nucleotide polymorphisms (tagSNPs) and asthma, atopy, atopic asthma and these phenotypes co-occurring with AD in a familial collection of asthmatic subjects originating from the Saguenay–Lac-Saint-Jean (SLSJ) founder population [25-29]. We then performed a direct sequencing of the *FLG* exons 2 and 3 in 48 SLSJ atopic asthmatic subjects in order to find causal mutations. Thus, we found positive associations for two tagSNPs located downstream the *FLG* 3' untranslated region and

the sequencing assays allowed to detect 104 genetic variants, including 54 new synonymous or non-synonymous nucleotide substitutions. These results document a genetic association of *FLG* with asthma and related atopic component as well as new genetic variants likely specific to the SLSJ population and/or to the atopic asthma disease.

#### 8.1.4 SUBJECTS AND METHODS

##### *Association Study Subjects*

Clinical characteristics and phenotyping of the subjects have been described in recent reports [30, 31] and summarized in Table 1. Briefly, asthma phenotype was characterized by the investigators for all participants using a respiratory health questionnaire and function tests following ATS standards [32]. Subjects were considered asthmatic if they exhibited symptoms of asthma and if function tests were compatible with airway hyperresponsiveness (AHR) either at time of recruitment or as previously recorded in the medical chart. AHR was defined as a provocative concentration of methacholine (PC<sub>20</sub>) causing a 20% fall in forced expiratory volume (PC<sub>20</sub> < 8 mg/ml). Atopy was defined as at least one positive response (wheal diameter ≥ 3 mm at 10 minutes) on a 26 inhalant allergens skin prick testing [33]. Serum IgE was measured with enzyme immunoassay (QuantiCLONE total IgE kit, Kallestad Diagnostics). Atopic asthma was defined as individuals diagnosed with both asthma and atopy. Personal history of AD was self-reported in the respiratory questionnaire. Families were included in the study if at least one parent was available for phenotypic assessment, at least one parent was unaffected, and all four grandparents were of French Canadian origin (validated using the BALSAC file [34]). The family participation rate was approximately 60%. The Chicoutimi Hospital local ethics committee approved the study and all subjects (or their parents) provided their informed consent.

##### *TagSNPs Selection*

The CEU (CEPH Utah residents with ancestry from northern and western Europe) genotyping data from the International HapMap project [35] was used to select SNPs covering ten kilobase up- and downstream of the *FLG* gene. A maximally informative set of SNPs (tagSNPs) were selected using a pairwise tagging algorithm implemented in the Haploview software (version 3.32)[36]. The  $r^2$  and the minor allele frequency

cutoffs were set to 0.8 and 0.05, respectively. TagSNPs described in this report are cited using their NCBI reference sequence number (rs#) and their characteristics are presented in Table 2.

#### *TagSNPs Genotyping*

Blood samples were obtained from all participants and DNA was extracted from whole blood leukocytes using Genomic-tip 100/G kit (Qiagen Inc., Valencia, CA, USA) according to the manufacturer's instructions. All tagSNPs were genotyped using the Sequenom® matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass array spectrometer (Sequenom Inc., San Diego, CA, USA). Primers were designed using the Sequenom SNP Assay Design software version 3.0 for iPLEX reactions. The protocol and reaction conditions are in accordance with the manufacturer [37] and the assay group file containing the PCR primers and the iPLEX extension probes can be supplied on request to the corresponding author. The genotypes were viewed and analyzed using the MassARRAY Typer software version 3.4 (Sequenom).

#### *FLG Sequencing*

Forty eight unrelated SLSJ subjects (validated with the BALSAC database [34]) presenting full-fit atopic asthmatic criteria (as described above), heterozygous for *FLG* rs2184951 and rs12730241 associated tagSNPs, and that have contributed to their associations were selected in order to increase the likelihood to find a causal mutation. This strategy is based on the assumption that genetic variants that will present important heterozygosity should be highly in LD with the associated tagSNP. *FLG* sequence information was obtained from UCSC Genome Browser. The PCR was divided in ten regions that spanned 13 kb, starting from the exon 2 and covering all exon 3. The amplification has been performed with the following conditions: 95°C 15 min followed by 33-40 cycles of 95°C 30 sec, 60-67°C 30 sec and 72°C 60-240 sec, then a final extension of 72°C 10 min. PCR primers specific annealing temperatures, extension time and number of cycles are listed in the supplementary Table E1. PCR were done in 25 µL volume containing 0-1 mM of Magnesium Chloride, 1X of Qiagen HotStart Taq PCR Buffer, 0.2-0.4 mM of dNTPs, 0.02-0.04 U/ml of HotStart Taq DNA polymerase (Qiagen, Valencia, Ca), 0.8 µM of each primer and 20 ng human genomic DNA. Amplification products were purified with multiscreen PCR plates (Millipore Corporation, Bedford, Ma), sequenced with BigDye terminator v3.1 chemistry

following instructions of the manufacturer and analyzed on a 3730XL Genetic analyzer (Applied Biosystems, Foster City, Ca). Sequencing reactions were performed with the following conditions: 25 cycles of 96°C 10 sec, 50°C 5 sec and 60°C 4 min sec and analyzed on a 3730XL Genetic analyzer (Applied Biosystems, Foster City, Ca). Primers used for sequencing are also listed in the supplementary Table E1. Sequence analysis was performed with phredPhrap (version 5.04).

#### *Statistical Analysis*

Allele distribution patterns and Mendelian errors were assessed using the Family-Based Association Test (FBAT) software (version 1.7) [38-40]. All tagSNPs analyses were performed under an additive genetic model using an empirical estimate of the variance [38]. Hardy-Weinberg equilibrium (HWE) p-values, considering a significance threshold of 0.001, and linkage disequilibrium (LD)  $r^2$  values between tagSNPs were calculated using Haploview software (version 4.0) [36]. Considering that the *FLG* has been selected on a precise biological plausibility and that genetic associations have already been reported in AD and asthma, we did not apply statistical corrections for the total number of tests performed.

### **8.1.5 RESULTS**

#### *Association of FLG With Asthma and Related Phenotypes*

None of the tagSNPs gave a significant deviation from the HWE (Table 2) and no Mendelian error has been found in the entire sample for all tested tagSNPs. Minor allele frequencies of the tagSNPs range from 0.09 to 0.45 (Table 3).

The association results for asthma, atopy and atopic asthma phenotypes are presented in Table 3. Of the five tested tagSNPs, rs2184951 and rs12730241 showed positive associations for asthma (p-values of 0.014 and 0.011, respectively) and for atopy (p-values of 0.025 and 0.020, respectively). For both associated tagSNP, the minor allele was undertransmitted to the affected probands, suggesting a protective effect. Perfect LD ( $r^2 = 1$ ) is observed between the two associated tagSNPs (supplementary Figure E1).

To verify the influence of AD on these associations, we combined these three previous phenotypes with the presence of a personal history of AD (supplementary Table E2).



When selecting all asthmatic subjects and discriminating for the presence of AD (Asthma + AD), the positive associations of the tagSNPs were lost ( $p$ -values  $> 0.05$ ). In contrast, atopic subjects discriminated for the presence of AD (Atopy + AD) showed stronger associations for the same rs2184951 and rs12730241 tagSNPs ( $p$ -values of 0.0097 and 0.0097, respectively). These associations were also observed in atopic asthmatic subjects that reported a personal history of AD (Atopic asthma + AD) ( $p$ -values of 0.0357 and 0.0357, respectively). In all cases, minor alleles were undertransmitted to the affected probands, also suggesting a protective genetic effect.

*FLG Sequencing: Identification of a New Termination Nucleotide Substitution*

As these two associated tagSNPs are located downstream the *FLG* 3' untranslated region and as the common R501X and 2282del4 mutations are located in the *FLG* repeat 1 (nearest from the rs11582620), we found intriguing that the association was not observed with this tagSNP. Thus, we sequenced the *FLG* coding region in atopic asthmatic subjects in order to find causal mutations. Figure 1 shows the *FLG* sequenced region (13Kb), the five tagSNPs location and the 104 variants identified, numbered 1 to 104 and presented in details in the supplementary Table E3. The delimitation of the eleven *FLG* repeats used to localize genetic variants in the exon 3 was based on the sequence previously reported by Sandilands and coworkers [14]. Among the 104 identified variants, 50 were reported in the NCBI SNP database (including the tagSNP rs11582620) and 54 did not have a NCBI reference sequence number. The Table 4 presents the distribution of the known and new identified variants in regard of their HWE deviation, their resulting amino acid substitution as well as their minor allele frequency. Among the known null or frameshift mutations reported [14], only the R501X and R2447X have been detected (refer as #8 and #47, respectively in Figure 1 and supplementary Table E3). The remaining variants were non-synonymous ( $n = 34$ ) or synonymous nucleotide substitutions ( $n = 14$ ) reaching the HWE or not. Among the new identified genetic variants, the c.10534C > T substitution (referred as #84 in Figure 1 and Table E3) is located in the *FLG* repeat 10 and is predicted to lead to the substitution of a glutamine (Q) for a termination codon (X) at position 3512 of the profilaggrin protein. In addition to this new p.Gln3512Term substitution (quoted Q3512X), 35 non-synonymous and 18 synonymous nucleotide substitutions, in HWE or not, have been identified. For each kind of variants presented in Table 4, the minor allele frequencies (MAF) have been regrouped in three representative classes in order to

highlight the rare variants ( $0 < \text{MAF} \leq 0.05$ ), the common variants ( $0.05 < \text{MAF} \leq 0.15$ ) as well as the variants with the highest heterozygosity ( $0.42 < \text{MAF} \leq 0.50$ ).

In order to define the *FLG* LD structure in the SLSJ sequenced sample,  $r^2$  between each pair of identified variants ( $n = 104$ ) has been calculated and are presented as online supplementary data (Figure E2).

### 8.1.6 DISCUSSION

The present study reports positive associations for two *FLG* tagSNPs (rs2184951 and rs12730241) with asthma and atopy, as well as with atopy and atopic asthma co-occurring with AD in a familial asthma sample originating from the Canadian SLSJ founder population. In addition, this study documents the first direct sequencing of the *FLG* exons 2 and 3 in atopic asthmatic SLSJ subjects.

#### *Association Study*

The *FLG* rs2184951 and rs12730241 associations showed that the minor alleles were undertransmitted to the affected probands, suggesting a protective effect for asthma and atopy related phenotypes. Palmer and coworkers, studying *FLG* R510X and 2282del4, already observed a trend for a protective effect in asthma (odd ratio = 0.8), but this association was not significant [9]. The mechanisms by which *FLG* mutations are strong predisposing factors in IV and AD diseases [9, 11, 12, 17, 18], and are suggested to be protective factors in asthma and its atopic component are difficult to explain at this stage of knowledge. If these associated tagSNPs were in LD with a gain-of-function mutation that would improve the keratin filaments aggregation in the epithelial barriers, it would result in a reduced surface contact with allergens, protecting against the development of atopy and asthma. Such hypothesis remained to be verify and raised important questions about the implication of the epithelial barrier in the asthma genetics [41, 42] and the relevance to consider this aspect in addition to the immune and inflammatory responses in candidate genes selection for asthma association studies [41].

Results obtained from asthma, atopy and atopic asthma co-occurring with AD phenotypes allow to suggest an influence of AD on the p-values observed between atopy as well as atopic asthma and rs2184951 and rs12730241 tagSNPs. As this influence have not been observed for asthma, it may reflect the importance of the *FLG* in atopy development, prior to asthma, which is in accordance with the atopic disease march [43]. On the other hand, it is interesting to note that *FLG* could be implicated in the progression of asthma as indicated by a recent study suggesting *FLG* as a marker of asthma severity independent of AD status [44]. This observation was not replicated in an independent study [16]. The present SLSJ familial sample being composed predominantly of mild asthmatic subjects, we did not investigate the association between variants and asthma severity phenotype in the present work.

#### *FLG Sequencing*

In order to characterize the *FLG* molecular structure in a SLSJ atopic asthma sub-sample, a direct sequencing of the gene has been done. An unexpected number of new missense nucleotide substitutions, including the Q3512X, has been detected and only two of the ones reported by Sandilands and coworkers [14] were observed (R501X and R2447X). In addition, none of the deletions reported in the *FLG* literature have been detected by our sequencing assays and some of the most difficult parts of the gene remained unfortunately unsequenced, as example by the Gap 5 that spans *FLG* repeats 8-9 and that contains the new S2899X and S3296X mutations described in Japanese subjects [45, 46]. The *FLG* common R501X and 2282del4 mutation frequencies have been examined in the CEPH samples and both mutations were absent in non-European populations such as those of African and Asian origin [9], showing a different distribution of these variants depending on the studied population. As the SLSJ population is recognized as a young founder population settled by French ancestors in the 17<sup>th</sup> century [25-29], it is likely that the 2282del4 and other European mutations have been modulated by the genetic drift and that their alleles have not been introduced in the SLSJ population.

The studied tagSNPs have been selected using the CEU Hapmap genotyping data and assuming the similar genetic background between CEU and SLSJ populations. Thus, the perfect LD observed between the two associated tagSNPs (rs2184951 and

rs12730241) is unexpected and highlights the difference in genetic LD structure between these populations. Such difference may indicate that even if the SLSJ individuals descend from European ancestors, the founder effect has modulated the genetic variants presence or absence as well as their frequencies (CEU: 0.07 and 0.04, respectively; SLSJ: 0.18 and 0.18, respectively). It is then not surprising that mutations specific in other populations such as the 3702delG in Irish IV families [47] or the 3321delA and S2554X in Japanese AD subjects [48] are not found in the SLSJ sample. In counterpart, all newly identified nucleotide substitutions that are not reported in NCBI database nor in other diseases, are likely to be specific to the SLSJ population and/or to the atopic asthma disease itself. Whether they are known or newly identified, the genetic variants responsible of the *FLG* observed genetic effect with asthma related phenotypes in the SLSJ remained to be elucidated by further investigations.

To conclude, the tagSNP *FLG* associations with asthma and atopy related phenotypes underline a role of one or numerous genetic variants identified by our sequencing assay. Delineation of the real genetic risks and the role of this *FLG* biological pathway in atopic asthma pathogenesis and progression needs additional genetic as well as functional studies.

### **8.1.7 ACKNOWLEDGMENTS**

We thank families of the Saguenay–Lac-Saint-Jean for their participation in this study. The authors wish to acknowledge Dr Pierre Lepage and Joëlle Fontaine of the Functional Genomics platform of the McGill University and Génome Québec Innovation Centre for their technical assistance for the sequencing assay. K. Tremblay was supported by AllerGen NCE Inc. and Laval University Foundation studentships. Y. Bossé is recipient of a fellowship award from the Canadian Institutes of Health and Research (CIHR). T.J. Hudson is recipient of an Investigator Award from the CIHR and a Clinician-scientist Award in Translational Research from the Burroughs Wellcome Fund. C. Laprise is the chairholder of the Canada Research Chair on genetic determinants in asthma and the director of the Genetics platform of the Respiratory Health Network (RHN) of the Fonds de la recherche en santé du Québec (FRSQ),

which financially supports the Saguenay–Lac-St-Jean family collection. This work was supported by the CIHR and the RHN of the FRSQ.

### 8.1.8 WEB RESOURCES

Canada Research Chair, [www.chairs.gc.ca](http://www.chairs.gc.ca)

dbSNP, <http://www.ncbi.nlm.nih.gov/projects/SNP/>

Ensembl Genome Browser database, <http://www.ensembl.org/index.html>

International HapMap Project, <http://www.hapmap.org/>

NCBI, <http://www.ncbi.nlm.nih.gov/>

Tagger program, <http://broad.mit.edu/mpg/tagger/>

### 8.1.9 REFERENCES

1. Howarth PH, Wilson J, Djukanovic R, Wilson S, Britten K, et al. (1991) Airway inflammation and atopic asthma: a comparative bronchoscopic investigation. *Int Arch Allergy Appl Immunol* 94: 266-269.
2. Bousquet J, Jeffery PK, Busse WW, Johnson M, Vignola AM (2000) Asthma. From bronchoconstriction to airways inflammation and remodeling. *Am J Respir Crit Care Med* 161:1720-45.
3. Cookson W (2004) The immunogenetics of asthma and eczema: a new focus on the epithelium. *Nat Rev Immunol* 4: 978-988.
4. Blumenthal MN (2005) The role of genetics in the development of asthma and atopy. *Curr Opin Allergy Clin Immunol* 5: 141-145.
5. Hoffjan S, Nicolae D, Ober C (2003) Association studies for asthma and atopic diseases: a comprehensive review of the literature. *Respir Res* 4: 14.
6. Ober C, Hoffjan S (2006) Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 7: 95-100.
7. Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, et al. (2006) Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat Genet* 38: 337-342.
8. Candi E, Schmidt R, Melino G (2005) The cornified envelope: a model of cell death in the skin. *Nat Rev Mol Cell Biol* 6: 328-340.

9. Palmer CN, Irvine AD, Terron-Kwiatkowski A, Zhao Y, Liao H, et al. (2006) Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat Genet* 38: 441-446.
10. Morar N, Cookson WO, Harper JI, Moffatt MF (2007) Filaggrin mutations in children with severe atopic dermatitis. *J Invest Dermatol* 127: 1667-1672.
11. Marenholz I, Nickel R, Ruschendorf F, Schulz F, Esparza-Gordillo J, et al. (2006) Filaggrin loss-of-function mutations predispose to phenotypes involved in the atopic march. *J Allergy Clin Immunol* 118: 866-871.
12. Barker JN, Palmer CN, Zhao Y, Liao H, Hull PR, et al. (2007) Null Mutations in the Filaggrin Gene (FLG) Determine Major Susceptibility to Early-Onset Atopic Dermatitis that Persists into Adulthood. *J Invest Dermatol* 127: 564-7.
13. Ruether A, Stoll M, Schwarz T, Schreiber S, Folster-Holst R (2006) Filaggrin loss-of-function variant contributes to atopic dermatitis risk in the population of Northern Germany. *Br J Dermatol* 155: 1093-1094.
14. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, et al. (2007) Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nat Genet* 39: 650-654.
15. Stemmler S, Parwez Q, Petrasch-Parwez E, Epplen JT, Hoffjan S (2007) Two common loss-of-function mutations within the filaggrin gene predispose for early onset of atopic dermatitis. *J Invest Dermatol* 127: 722-724.
16. Rogers AJ, Celedon JC, Lasky-Su JA, Weiss ST, Raby BA (2007) Filaggrin mutations confer susceptibility to atopic dermatitis but not to asthma. *J Allergy Clin Immunol* 120: 1332-1337.
17. Weidinger S, Illig T, Baurecht H, Irvine AD, Rodriguez E, et al. (2006) Loss-of-function variations within the filaggrin gene predispose for atopic dermatitis with allergic sensitizations. *J Allergy Clin Immunol* 118: 214-219.
18. Weidinger S, Rodriguez E, Stahl C, Wagenpfeil S, Klopp N, et al. (2007) Filaggrin Mutations Strongly Predispose to Early-Onset and Extrinsic Atopic Dermatitis. *J Invest Dermatol* 127: 724-726.
19. Baurecht H, Irvine AD, Novak N, Illig T, Buhler B, et al. (2007) Toward a major risk factor for atopic eczema: meta-analysis of filaggrin polymorphism data. *J Allergy Clin Immunol* 120: 1406-1412.

20. Henderson J, Northstone K, Lee SP, Liao H, Zhao Y, et al. (2008) The burden of disease associated with filaggrin mutations: a population-based, longitudinal birth cohort study. *J Allergy Clin Immunol* 121: 872-877.
21. Brown SJ, Relton CL, Liao H, Zhao Y, Sandilands A, et al. (2008) Filaggrin null mutations and childhood atopic eczema: a population-based case-control study. *J Allergy Clin Immunol* 121: 940-946.
22. Weidinger S, O'Sullivan M, Illig T, Baurecht H, Depner M, et al. (2008) Filaggrin mutations, atopic eczema, hay fever, and asthma in children. *J Allergy Clin Immunol* 121: 1203-1209.
23. Presland RB, Haydock PV, Fleckman P, Nirunskisiri W, Dale BA (1992) Characterization of the human epidermal profilaggrin gene. Genomic organization and identification of an S-100-like calcium binding domain at the amino terminus. *J Biol Chem* 267: 23772-23781.
24. Sandilands A, Smith FJ, Irvine AD, McLean WH (2007) Filaggrin's fuller figure: a glimpse into the genetic architecture of atopic dermatitis. *J Invest Dermatol* 127: 1282-1284.
25. Heyer E, Tremblay M (1995) Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 56: 970-978.
26. Sriver CR (2001) Human genetics: lessons from Quebec populations. *Annu Rev Genomics Hum Genet* 2: 69-101.
27. Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, et al. (1996) Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am J Hum Genet* 59: 633-643.
28. Engert JC, Berube P, Mercier J, Dore C, Lepage P, et al. (2000) ARSACS, a spastic ataxia common in northeastern Quebec, is caused by mutations in a new gene encoding an 11.5-kb ORF. *Nat Genet* 24: 120-125.
29. Richter A, Rioux JD, Bouchard JP, Mercier J, Mathieu J, et al. (1999) Location score and haplotype analyses of the locus for autosomal recessive spastic ataxia of Charlevoix-Saguenay, in chromosome region 13q11. *Am J Hum Genet* 64: 768-775.

30. Poon AH, Laprise C, Lemire M, Montpetit A, Sinnett D, et al. (2004) Association of vitamin D receptor genetic variants with susceptibility to asthma and atopy. *Am J Respir Crit Care Med* 170: 967-973.
31. Tremblay K, Lemire M, Provost V, Pastinen P, Renaud Y, et al. (2006) Association study between the CX3CR1 gene and asthma. *Genes Immun* 7: 632-639.
32. Official statement of the American Thoracic Society adopted by the ATS Board of Directors (1987) Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. *Am Rev Respir Dis* 136: 225-244.
33. Sub-Committee on Skin Tests of the European Academy of Allergology and Clinical Immunology (1989) Skin tests used in type I allergy testing Position paper. *Allergy* 44 Suppl 10: 1-59.
34. Bouchard G, Roy R, Casgrain B, Hubert M (1989) [Population files and database management: the BALSAC database and the INGRES/INGRID system]. *Hist Mes* 4: 39-57.
35. International HapMap Consortium (2005) A haplotype map of the human genome. *Nature* 437: 1299-1320.
36. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263-265.
37. Oeth P, Beaulieu M, Park C, Kosman D, del Mistro G, et al. (2005) iPLEX assay: Increased plexing efficiency and flexibility for MassArray system through single base primer extension with mass-modified terminators. *SEQUENOM Application Note*
38. Lake SL, Blacker D, Laird NM (2000) Family-based tests of association in the presence of linkage. *Am J Hum Genet* 67: 1515-1525.
39. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, et al. (2004) Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet Epidemiol* 26: 61-69.
40. Laird NM, Horvath S, Xu X (2000) Implementing a unified approach to family-based tests of association. *Genet Epidemiol* 19 Suppl 1: S36-S42.
41. Hudson TJ (2006) Skin barrier function and allergic risk. *Nat Genet* 38: 399-400.
42. Irvine AD, McLean WH (2006) Breaking the (un)sound barrier: filaggrin is a major gene for atopic dermatitis. *J Invest Dermatol* 126: 1200-1202.



43. Spergel JM, Paller AS (2003) Atopic dermatitis and the atopic march. *J Allergy Clin Immunol* 112: S118-S127.
44. Palmer CN, Ismail T, Lee SP, Terron-Kwiatkowski A, Zhao Y, et al. (2007) Filaggrin null mutations are associated with increased asthma severity in children and young adults. *J Allergy Clin Immunol* 120: 64-68.
45. Sasaki T, Kudoh J, Ebihara T, Shiohama A, Asakawa S, et al. (2008) Sequence analysis of filaggrin gene by novel shotgun method in Japanese atopic dermatitis. *J Dermatol Sci* 51: 113-120.
46. Nomura T, Akiyama M, Sandilands A, Nemoto-Hasebe I, Sakai K, et al. (2008) Specific filaggrin mutations cause ichthyosis vulgaris and are significantly associated with atopic dermatitis in Japan. *J Invest Dermatol* 128: 1436-1441.
47. Sandilands A, O'Regan GM, Liao H, Zhao Y, Terron-Kwiatkowski A, et al. (2006) Prevalent and rare mutations in the gene encoding filaggrin cause ichthyosis vulgaris and predispose individuals to atopic dermatitis. *J Invest Dermatol* 126: 1770-1775.
48. Nomura T, Sandilands A, Akiyama M, Liao H, Evans AT, et al. (2007) Unique mutations in the filaggrin gene in Japanese patients with ichthyosis vulgaris and atopic dermatitis. *J Allergy Clin Immunol* 119: 434-440.

#### 8.1.10 FIGURE LEGENDS

**Figure 1:** Schematic representation of the *FLG* genomic structure, the analyzed tagSNPs and the sequencing identified variants.

The location of the tagSNPs is indicated below the gene structure and the sequenced region is indicated above. The sequence used to define the exon 3 *FLG* repeats and to localize the identified variants (1 to 104) has been reported previously [14]. The unsequenced gaps are indicated on the top of the figure. The supplementary Table E1 contains the 104 *FLG* variants detailed description.

**8.1.11 Table 1**

Clinical characteristics of the Saguenay–Lac-Saint-Jean subjects

	Families <sup>a</sup>	
	Probands (n = 208)	Family members (n = 644)
Male : Female ratio	1 : 1.3	1 : 1.1
Mean age (range)	18 (3-46)	40 (2-83)
Smoking status (n (%)) <sup>b</sup>		
Never	172 (84)	278 (44)
Ex-smoker	11 (5)	219 (34)
Smoker	22 (11)	141 (22)
FEV <sub>1</sub> % of predicted value (SD) <sup>c</sup>	92.3 (16.1)	94.6 (21.2)
PC <sub>20</sub> in mg/ml (SD) <sup>d</sup>	2.80 (3.51)	11.38 (5.05)
Serum IgE in µg/L (SD) <sup>e</sup>	226.3 (4.6)	103.3 (4.2)
Number of persons with subphenotypes (%)		
Asthma <sup>f</sup>	208 (100)	266 (41)
Atopy <sup>g</sup>	166 (81)	355 (56)
Atopic asthma	166 (81)	186 (71)
Eczema (AD) <sup>h</sup>	77 (37)	163 (25)
Asthma + AD	77 (37)	85 (32)
Atopy + AD	113 (55)	113 (32)
Atopic asthma + AD	69 (37)	69 (37)

<sup>a</sup> The 208 probands represent the 208 extended families that represent a total of 249 nuclear families for the FBAT analyses.

<sup>b</sup> Evaluated for 206 probands and for 651 family members.

<sup>c</sup> FEV<sub>1</sub> = Geometric mean of the force expiratory volume in one second evaluated for 191 probands and for 553 family members.

<sup>d</sup> PC<sub>20</sub> = Concentration of methacholine inducing a 20% fall in FEV<sub>1</sub>. Evaluated for 165 probands and for 498 family members. Geometric mean and SD were obtained from the log transformed PC<sub>20</sub> values.

<sup>e</sup> IgE = Immunoglobulin E serum concentration evaluated for 196 probands and 574 family members. Geometric mean and SD were obtained from the log transformed IgE values.

<sup>f</sup> Present asthma or past documented clinical history of asthma.

<sup>g</sup> Defined as at least one positive response on skin prick testing (wheal diameter > 3 mm at 10 minutes).

<sup>h</sup> AD = Atopic dermatitis questionnaire self-reported.

**8.1.12 Table 2**  
 Characteristics of the genotyped *FLG* tagSNPs

TagSNP	Chromosome Location <sup>a</sup>	Genomic Feature	HWE <sup>b</sup>
rs2184951	1:150539792	Downstream 3'UTR	0.481
rs12730241	1:150540640	Downstream 3'UTR	0.460
rs11582620	1:150552750	Exon 3 (R412R)	0.711
rs1933063	1:150567845	Upstream 5'UTR	0.843
rs1933064	1:150568200	Upstream 5'UTR	0.944

<sup>a</sup> NCBI SNP build 126, freeze December 2006.

<sup>b</sup> HWE = Hardy-Weinberg Equilibrium p-values (obtained with Haploview software).

**8.1.13 Table 3**

Association of *FLG* tagSNPs with asthma, atopy and atopic asthma phenotypes

TagSNP	Base Change	MAF	Asthma					Atopy					Atopic Asthma				
			Fam	S	E(S)	Z	P	Fam	S	E(S)	Z	P	Fam	S	E(S)	Z	P
rs2184951	T > G	0.18	78	53	68	-2.459	<b>0.014</b>	66	48	62	-2.238	<b>0.025</b>	68	44	54	-1.871	0.0614
rs12730241	G > A	0.18	80	54	70	-2.555	<b>0.011</b>	68	51	66	-2.326	<b>0.020</b>	69	45	56	-1.907	0.0565
rs11582620	A > G	0.13	76	63	63	0.026	0.979	62	58	54	0.730	0.466	65	53	49	0.668	0.5040
rs1933063	G > A	0.09	57	47	46	0.194	0.846	50	37	40	-0.562	0.574	49	34	35	-0.106	0.9156
rs1933064	G > A	0.45	113	166	178	-1.367	0.172	91	133	142	-1.105	0.269	96	120	130	-1.388	0.1652

P-values highlighted in bold are significant (< 0.05).

Abbreviations used: TagSNP = Tagging single nucleotide polymorphism, MAF = Minor allele frequency, Fam = Number of families contributing to the statistic, S = FBAT statistic, E(S) = Expected FBAT statistic, Z = Z-score, P = p-value.

**8.1.14 Table 4**

Repartition of the *FLG* identified genetic variants in regard of the HWE, the single nucleotide type of substitution and the minor allele frequency

Variants *	HWE †		Type ‡		MAF §	
NEW	Yes	47	NS	29	A	21
					B	8
					C	0
			S	18	A	17
					B	1
					C	0
	No	7	NS	7	A	0
					B	1
					C	6
			S	0	A	0
					B	0
					C	0
KNOWN	Yes	34	NS	24	A	14
					B	7
					C	3
			S	10	A	7
					B	3
					C	0
	No	16	NS	12	A	0
					B	1
					C	11
			S	4	A	0
					B	0
					C	4

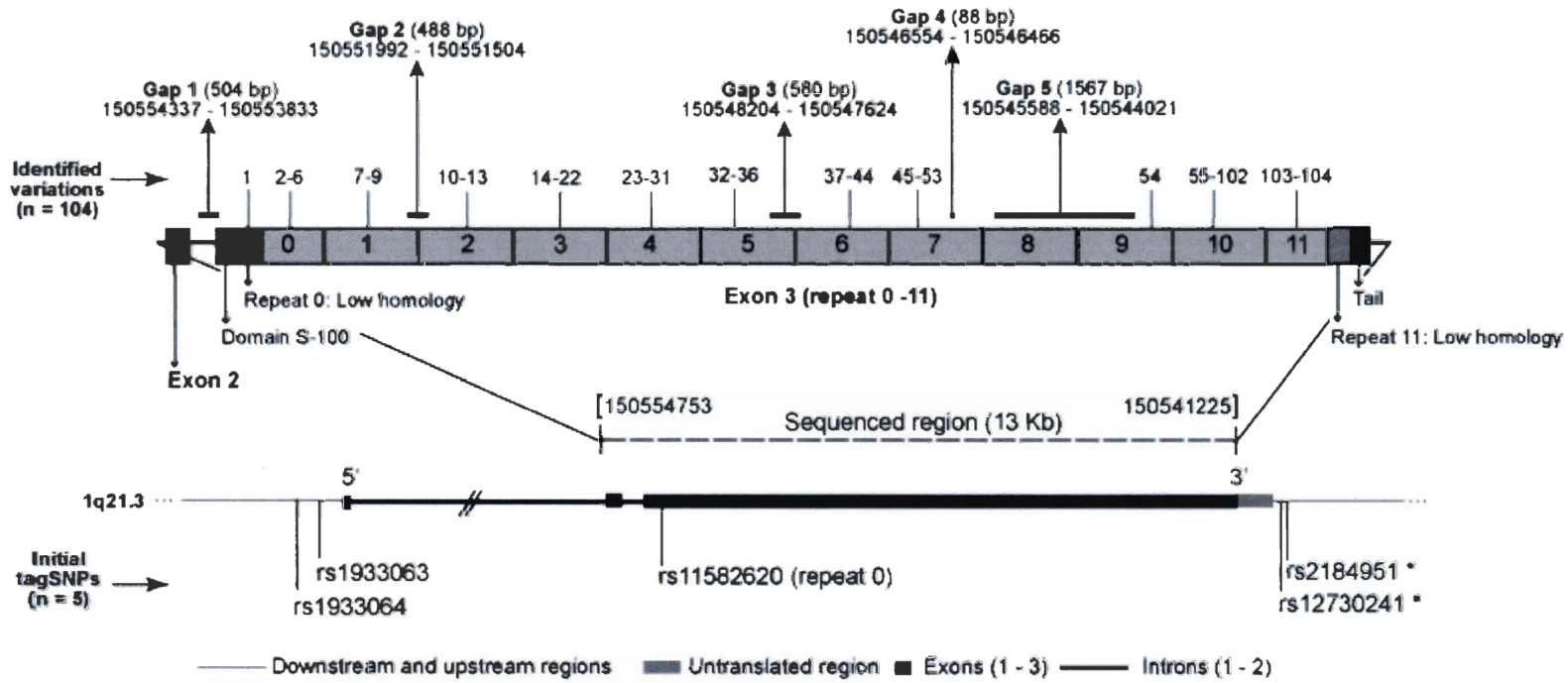
\* NEW refers to newly identified genetic variants; KNOWN refers to the NCBI reported genetic variants.

† HWE = Hardy-Weinberg Equilibrium; Yes = in HWE, No = not in HWE.

‡ NS = non-synonymous substitution; S = synonymous substitution.

§ MAF = Minor allele frequency; A =  $0 < \text{MAF} \leq 0.05$ , B =  $0.05 < \text{MAF} \leq 0.15$ , C =  $0.42 < \text{MAF} \leq 0.50$ .

8.1.16 Figure 1



### 8.1.17 ONLINE SUPPLEMENT

#### Sequencing the Filaggrin Gene in Atopic Asthmatic Subjects From the Saguenay–Lac-Saint-Jean Founder Population

Karine Tremblay, Yohan Bossé, Cecilia Hernández, David Roquis, Thomas J. Hudson and Catherine Laprise

#### LEGEND TO FIGURES

**Figure E1:** Linkage disequilibrium (LD) plot of the five *FLG* tagSNPs genotyped in the Saguenay–Lac-Saint-Jean asthma familial sample.

The location of each tested tagSNP along the chromosome is indicated on top. Each diamond indicates the  $r^2$  between respective pairs of tagSNPs. Diamonds in black represent perfect LD ( $r^2 = 1.0$ ). Strength of LD is depicted by progression of color. For all  $r^2 < 1.0$ , the color moves from black to white as  $r^2$  runs from 1 to 0.

**Figure E2:** Linkage disequilibrium (LD) plot of all *FLG* identified variants in the Saguenay–Lac-Saint-Jean atopic asthmatic subjects.

The location of each variant along the chromosome is indicated below the gene structure. Each diamond indicates the  $r^2$  between respective pairs of SNPs. Diamonds in black represent perfect LD ( $r^2 = 1.0$ ). Strength of LD is depicted by progression of color. For all  $r^2 < 1.0$ , the color moves from black to white as  $r^2$  runs from 1 to 0. A) LD plot for all identified variants ( $n = 104$ ); B) LD plot zoomed section of variants 1 to 52; and C) LD plot zoomed section of variants 53 to 104. Detailed description of genetic variants can be found in Table E3.

**Table E1***FLG* sequencing oligonucleotides and PCR conditions

Sequenced Region	SNP Identified *	Flanking Primer	Sequencing Primer †	Primer Sequence (5' - 3')	Size (bp)	T <sub>m</sub> (°C)	Extension (sec)	Cycle
1	None	<i>FLGE02F</i> <i>FLGE02R</i>	<i>FLGE02F</i>	TGGATTACTTAGAGGAAAATGTAGG GAGAAATAGTTCACAAAGAGCTCAAA TGGATTACTTAGAGGAAAATGTAGG	362	60	90	40
2	1 - 8	<i>FilLR2F</i> <i>RPT1P6</i>	<i>RPT1P6</i> <i>FLG13F</i> <i>FLGe03gF</i>	GTCACTTACCCCATCAAATC ACCTGAGTGTCCAGACCTATT ACCTGAGTGTCCAGACCTATT GCTGATAATGTGATTCTGTCTG GGAGACAATAGGAAGAGGCTAAG	1710	60	120	35
3	9 - 16	<i>FLG13F</i> <i>FLG13R</i>	<i>FILH1F3</i> <i>FLGe03SF</i> <i>RPT2P1F</i>	GCTGATAATGTGATTCTGTCTG GACCCCGATGATTGTTCTCTGT CACGGAAAGGCTGGGCTGA AATCAGGAACTCCTCACGCA AAACAGTCTGAGTCCTCCC	3697	60	240	35
4	17 - 32	<i>FLG35F</i> <i>FLG35R</i>	<i>FLG35F</i> <i>FLG35R</i> <i>FLGe03cF</i>	GCAAGCAGACAAACTCGTAAG ACATCAGACCTTTCCTGGGAC GCAAGCAGACAAACTCGTAAG ACATCAGACCTTTCCTGGGAC TCAGGTCACTAACAGTGAGG	1916	64	240	35
5	104	<i>FLGendF</i> <i>FLGendR</i>	<i>FLGendR</i>	CTAGTACCGCTAAGGAACATGG TGGCTCCTTCGATATTTCTGA TGGCTCCTTCGATATTTCTGA	781	60	90	40
6	103	<i>FLG10F</i> <i>FLG10R</i>	<i>FLG10R</i>	GCCCATGGGCGGACCAGGA CTGCACTACCATAGCTGCC CTGCACTACCATAGCTGCC	1753	62	150	40



7	37 - 50	<i>FLGe03L-F</i>	<i>FLGe03L-F</i> <i>FLGe03M-R</i>	CTGAGTCCACCCATGGACAGT	1401	62	60	35
		<i>FLGe03M-R</i>		AACTGGATCCCCAGTTCC				
8	54 - 102	<i>FLGe03P-F</i>	<i>FLGe03iR</i> <i>FLGe03jR</i>	AAGGCACGGCTCCAGACA	1831	67	60	35
		<i>FLGe03Q-R</i>		CTGGACCCCGCTGATTAC GCTCTGTCTTCGTGATGGGA TCGTGGTAGGATCCCTGTC				
9	33 - 36	<i>FLGe03e_F</i>	<i>FLGe03e_R</i>	CTCCAGGAACCCAGGTC	1401	63	60	35
		<i>FLGe03e_R</i>		CCAAGCAGATCCAAGATGGT CCAAGCAGATCCAAGATGGT				
10	51 - 53	<i>FLG710F</i>	<i>FLG710F</i> <i>FLG78R</i>	CCCAGGACAAGCAGGAACT	751	62	150	33
		<i>FLG78R</i>		CTACCGAATGCTCGTGGTGGT CCCAGGACAAGCAGGAACT CTACCGAATGCTCGTGGTGGT				

\* The numbers refer to the SNPs identified and presented in the Figure 2 and the online Table E1.

† Sequencing reactions: 25 cycles of 96°C 10 sec, 50°C 5 sec and 60°C 4min sec.

**Table E2**Association of *FLG* tagSNPs with asthma, atopy and atopic asthma that co-occur with a personal history of atopic dermatitis

TagSNP	Base Change	MAF	Asthma + AD					Atopy + AD					Atopic Asthma + AD				
			Fam	S	E(S)	Z	P	Fam	S	E(S)	Z	P	Fam	S	E(S)	Z	P
rs2184951	T > G	0.18	29	12	18	-1.860	0.0628	24	8	16	-2.588	<b>0.0097</b>	24	9	15	-2.101	<b>0.0357</b>
rs12730241	G > A	0.18	29	12	18	-1.860	0.0628	24	8	16	-2.588	<b>0.0097</b>	24	9	15	-2.101	<b>0.0357</b>
rs11582620	A > G	0.13	39	30	26	0.982	0.3261	35	32	26	1.666	0.0956	32	26	21	1.379	0.1678
rs1933063	G > A	0.09	20	12	13	-0.343	0.7316	22	11	13	-0.667	0.5050	16	10	10	0	1.0
rs1933064	G > A	0.45	44	43	50	-1.673	0.0943	38	39	44	-1.152	0.2492	35	32	37	-1.387	0.1655

P-values highlighted in bold are significant (&lt; 0.05).

Abbreviations used: AD = Atopic dermatitis; TagSNP = Tagging single nucleotide polymorphism, MAF = Minor allele frequency, Fam = Number of families contributing to the statistic, S = FBAT statistic, E(S) = Expected FBAT statistic, Z = Z-score, P = p-value.

**Table E3**Genetics variants identified in the *FLG* gene

ID *	Description †	Base Change	mRNA Nucleotide Position	Location ‡	Amino Acid Predicted Substitution §		MAF	Observed Genotypes (n = 48)			HWE p-value	
								HMZ <sub>common</sub>	HTZ	HMZ <sub>rare</sub>		
1	New SNP	A > G	472	150553513	NS	Lys (AAA) vs Arg (AGA)	K158R	0.01	46	1	0	1.0
2	rs41267154	G > T	994	150552991	NS	Gly (GGC) vs Val (GTC)	G332V	0.47	3	44	0	9.7E-11
3	New SNP	C > A	1157	150552828	S	Ser (TCC) vs Ser (TCA)	S386S	0.04	43	4	0	1.0
4	rs11582620 (TagSNP)	T > C	1235	150552750	S	Arg (CGT) vs Arg (CGC)	R412R	0.11	37	10	0	1.0
5	rs11588170	G > A	1329	150552656	NS	Gly (GGG) vs Arg (AGG)	G444R	0.46	6	39	2	1.7E-6
6	rs3126081	A > G	1359	150552626	NS	Thr (ACA) vs Ala (GCA)	T454A	0.47	3	44	0	9.7E-11
7	rs11584340	C > T	1431	150552554	NS	Pro (CCT) vs Ser (TCT)	P478S	0.47	3	44	0	9.7E-11
8	R501X	C > T	1500	150552485	NS	Arg (CGA) vs STOP (TGA)	R501X	0.04	43	4	0	1.0
9	New SNP	A > T	1740	150552245	NS	Thr (ACC) vs Ser (TCC)	T581S	0.01	35	1	0	1.0
10	rs3120653	T > C	2507	150551478	S	Asp (GAT) vs Asp (GAC)	D836D	0.47	0	37	2	8.3E-9
11	rs2184952	G > A	2712	150551273	NS	Gly (GGC) vs Ser (AGC)	G905S	0.01	40	1	0	1.0
12	New SNP	G > A	2911	150551074	NS	Arg (CGT) vs His (CAT)	R971H	0.01	40	1	0	1.0
13	rs12756586	C > G	2937	150551048	NS	His (CAT) vs Asp (GAT)	H980D	0.07	35	6	0	1.0
14	rs9436067	C > T	3386	150550599	S	Ser (TCT) vs Ser (TCC)	S1129S	0.48	0	40	2	1.2E-9
15	rs7530018	C > G	3499	150550486	NS	Ala (GCA) vs Gly (GCA)	A1167G	0.48	0	40	2	1.2E-9
16	New SNP	G > T	3608	150550377	NS	Arg (AGG) vs Ser (AGT)	R1203S	0.01	41	1	0	1.0
17	rs11581512	A > G	3793	150550192	NS	His (CAC) vs Arg (CGC)	H1265R	0.07	36	6	0	1.0
18	New SNP	G > C	4060	150549925	NS	Gly (GGA) vs Ala (GCA)	G1354A	0.01	41	1	0	1.0
19	rs11586631	G > A	4078	150549907	NS	Arg (CGC) vs His (CAC)	R1360H	0.43	6	30	1	2.0E-4
20	rs11581451	A > T	4094	150549891	S	Ala (GCA) vs Ala (GCT)	A1365A	0.04	34	3	0	1.0
21	New SNP	T > C	4117	150549868	NS	Ile (ATT) vs Thr (ACT)	I1373T	0.03	35	2	0	1.0

22	rs11581433	A > G	4125	150549860	NS	Arg (AGA) vs Gly (GGA)	R1376G	0.49	1	36	0	0.0000
23	rs12750571	C > T	4308	150549677	NS	Arg (CGT) vs Cys (TGT)	R1437C	0.04	34	3	0	3.0E-9
24	rs12732920	T > C	4409	150549576	S	His (CAT) vs His (CAC)	H1470H	0.09	27	6	0	1.0
25	rs11204978	C > A	4444	150549541	NS	Ser (TCC) vs Tyr (TAC)	S1482Y	0.42	6	24	1	0.0031
26	rs12750253	C > G	4451	150549534	NS	Asp (GAC) vs Glu (GAG)	D1484E	0.08	39	7	0	1.0
27	rs12750081	C > T	4567	150549418	NS	Thr (ACA) vs Ile (ATA)	T1523I	0.08	39	7	0	1.0
28	New SNP	C > T	4677	150549308	NS	Arg (CGT) vs Cys (TGT)	R1560C	0.03	43	3	0	1.0
29	rs12407807	C > T	5050	150548935	NS	Arg (CGC) vs His (CAC)	R1684H	0.44	6	25	2	0.0019
30	rs12405278	G > A	5094	150548891	NS	Arg (CGC) vs Cys (TGC)	R1699C	0.44	6	26	2	0.0012
31	New SNP	G > A	5210	150548775	S	His (CAC) vs His (CAT)	H1737H	0.01	39	1	0	1.0
32	rs12405241	G > A	5413	150548572	NS	Ala (GCG) vs Val (GTG)	A1805V	0.45	6	30	2	2.0E-4
33	New SNP	G > T	5616	150548369	NS	Gln (CAA) vs Lys (AAA)	Q1873K	0.07	38	6	0	1.0
34	rs36006086	G > A	5670	150548315	NS	Arg (CGG) vs Trp (TGG)	R1891W	0.05	43	5	0	1.0
35	rs35359555 & rs12407748	C > T	5671	150548314	NS	Arg (CGG) vs Gln (CAG)	R1891Q	0.47	5	37	2	2.2E-6
36	rs34806697	T > C	5740	150548245	NS	Gln (CAG) vs Arg (CGG)	Q1914R	0.06	38	5	0	1.0
37	rs2184954	A > G	6497	150547488	S	Ser (TCT) vs Ser (TCC)	S2166S	0.47	3	44	0	9.7E-11
38	New SNP	T > G	6573	150547412	NS	Lys (AAA) vs Glu (CAA)	K2192E	0.07	40	7	0	1.0
39	rs2184953	A > G	6579	150547406	NS	Tyr (TAT) vs His (CAT)	Y2194H	0.47	3	44	0	9.7E-11
40	New SNP	C > T	6788	150547197	S	Ala (GCG) vs Ala (GCA)	A2263A	0.02	45	2	0	1.0
41	rs6664985	A > G	6989	150546996	S	His (CAC) vs His (CAT)	H2330H	0.47	0	44	3	9.7E-11
42	New SNP	C > T	7014	150546971	NS	Asp (GAC) vs Asn (AAC)	D2339N	0.47	0	44	3	9.7E-11
43	New SNP	C > G	7096	150546889	NS	Ser (AGT) vs Thr (ACT)	S2366T	0.47	0	44	3	9.7E-11
44	New SNP	C > G	7191	150546794	NS	Glu (GAG) vs Gln (CAG)	E2398Q	0.47	0	44	3	9.7E-11
45	New SNP	C > T	7251	150546734	NS	Val (GTG) vs Met (ATG)	V2418M	0.06	41	6	0	1.0
46	New SNP	C > T	7329	150546656	NS	Lys (AAG) vs Glu (GAG)	K2444E	0.47	3	44	0	9.7E-11
47	R2447X	G > A	7338	150546647	NS	Arg(CGA) vs STOP (TGA)	R2447X	0.02	45	2	0	1.0

48	New SNP	C > T	7397	150546588	S	Pro (CCG) vs Pro (CCA)	P2466P	0.07	40	7	0	1.0
49	New SNP	A > G	7441	150546544	NS	Leu (TTG) vs Ser (TCG)	L2481S	0.47	3	44	0	9.7E-11
50	rs3126074	C > G	7520	150546465	NS	His (CAC) vs Gln (CAG)	H2507Q	0.47	3	44	0	9.7E-11
51	New SNP	G > C	7719	150546266	NS	Gln (CAG) vs Glu (GAG)	Q2574E	0.07	40	6	0	1.0
52	New SNP	T > G	7955	150546030	NS	Glu (GAA) vs Asp (GAC)	E2652D	0.14	34	13	0	0.7468
53	New SNP	C > T	8186	150545799	S	Gly (GGG) vs Gly (GGA)	G2729G	0.01	46	1	0	1.0
54	rs6681433 & rs34927045	A > G	9965	150544020	S	Gln (CAA) vs Gln (CAG)	Q3322Q	0.05	36	2	1	0.1548
55	rs3091276	T > C	10193	150543792	S	Ser (TCT) vs Ser (TCC)	S3398S	0.05	36	2	1	0.1548
56	New SNP	G > A	10202	150543783	S	Gly (GGG) vs Gly (GGA)	G3401G	0.05	36	2	1	0.1548
57	New SNP	G > C	10205	150543780	S	Arg (CGG) vs Arg (CGC)	R3402R	0.05	36	2	1	0.1548
58	New SNP	C > A	10208	150543777	S	Thr (ACC) vs Thr (ACA)	T3403T	0.05	36	2	1	0.1548
59	New SNP	A > G	10209	150543776	NS	Arg (AGG) vs Gly (GGG)	R3404G	0.05	36	2	1	0.1548
60	New SNP	A > C	10212	150543773	NS	Thr (ACC) vs Pro (CCC)	T3405P	0.05	36	2	1	0.1548
61	New SNP	C > G	10224	150543761	NS	Arg (CGA) vs Gly (GGA)	R3409G	0.05	35	2	1	0.1548
62	New SNP	G > A	10233	150543752	NS	Gly (GGA) vs Arg (AGA)	G3412R	0.05	36	2	1	0.1548
63	New SNP	A > G	10240	150543745	NS	His (CAC) vs Arg (CGC)	H3414R	0.05	36	2	1	0.1548
64	New SNP	G > A	10261	150543724	NS	Ser (AGC) vs Asn (AAC)	S3421N	0.08	36	0	3	7.1E-5
65	New SNP	T > C	10317	150543668	NS	Ser (TCA) vs Pro (CCA)	S3440P	0.05	36	2	1	0.1548
66	New SNP	A > T	10326	150543659	NS	Arg (AGA) vs Gly (GGA)	R3443G	0.05	36	2	1	0.1548
67	New SNP	G > A	10334	150543651	S	Arg (AGG) vs Arg (AGA)	R3445R	0.05	36	2	1	0.1548
68	New SNP	A > G	10340	150543645	S	Gly (GGA) vs Gly (GGG)	G3447G	0.05	36	2	1	0.1548
69	New SNP	T > C	10347	150543638	NS	Tyr (TAC) vs His (CAC)	Y3450H	0.05	36	2	1	0.1548
70	New SNP	C > T	10385	150543600	S	Ser (TCC) vs Ser (TCT)	S3462S	0.03	38	0	1	0.026
71	New SNP	C > A	10454	150543531	S	Ala (GCC) vs Ala (GCA)	A3485A	0.04	37	1	1	0.0779
72	New SNP	T > A	10466	150543519	S	Thr (ACT) vs Thr (ACA)	T3489T	0.04	37	1	1	0.0779
73	New SNP	G > A	10468	150543517	NS	Arg (CGT) vs His (CAT)	R3490H	0.04	37	1	1	0.0779

74	New SNP	A > G	10470	150543515	NS	Asn (AAT) vs Asp (GAT)	N3491D	0.04	37	1	1	0.0779
75	rs3126067	T > C	10472	150543513	S	Asn (AAT) vs Asn (AAC)	N3491N	0.04	37	1	1	0.0779
76	New SNP	G > A	10473	150543512	NS	Asp (GAC) vs Asn (AAC)	D3492N	0.04	37	1	1	0.0779
77	rs36006086 & rs35359555	C > G	10475	150543510	NS	Asp (GAC) vs Glu (GAG)	D3492E	0.04	37	1	1	0.0779
78	rs3126066	T > C	10490	150543495	S	Asp (GAT) vs Asp (GAC)	D3497D	0.05	36	2	1	0.1548
79	rs12746409	C > T	10496	150543489	S	Ser (TCC) vs Ser (TCT)	S3499S	0.05	36	2	1	0.1548
80	rs12728908	T > G	10506	150543479	NS	Trp (TGG) vs Gly (GGG)	W3503G	0.05	36	2	1	0.1548
81	rs12733038	A > G	10513	150543472	NS	His (CAT) vs Arg (CGT)	H3505R	0.05	36	2	1	0.1548
82	rs12746250	C > T	10517	150543468	S	His (CAC) vs His (CAT)	H3506H	0.05	36	2	1	0.1548
83	New SNP	A > T	10530	150543455	NS	Thr (ACT) vs Ser (TCT)	T3511S	0.05	36	2	1	0.1548
84	New SNP	C > T	10533	150543452	NS	Gln (CAG) vs STOP (TAG)	Q3512X	0.05	36	2	1	0.1548
85	New SNP	A > G	10534	150543451	NS	Gln (CAG) vs Arg (CGG)	Q3512R	0.05	36	2	1	0.1548
86	New SNP	G > C	10538	150543447	S	Ala (GCG) vs Ala (GCC)	A3513A	0.05	36	2	1	0.1548
87	New SNP	T > G	10560	150543425	NS	Ser (TCC) vs Ala (GCC)	S3521A	0.05	36	2	1	0.1548
88	New SNP	C > A	10562	150543423	S	Ser (TCC) vs Ser (TCA)	S3521S	0.05	36	2	1	0.1548
89	New SNP	G > T	10564	150543421	NS	Gly (GGC) vs Val (GTC)	G3522V	0.05	36	2	1	0.1548
90	rs12746218	C > A	10579	150543406	NS	Ala (GCG) vs Glu (GAG)	A3527E	0.05	36	2	1	0.1548
91	rs12746216	C > T	10584	150543401	NS	Pro (CCC) vs Ser (TCC)	P3529S	0.05	35	2	1	0.1589
92	rs12732870	A > T	10599	150543386	NS	Asn (AAC) vs Tyr (TAC)	N3534Y	0.05	36	2	1	0.1548
93	rs12732869	A > G	10600	150543385	NS	Asn (AAC) vs Ser (AGC)	N3534S	0.05	36	2	1	0.1548
94	rs12728605	T > C	10662	150543323	NS	Trp (TGG) vs Arg (CGG)	W3555R	0.05	36	2	1	0.1548
95	rs7518080	G > T	10690	150543295	NS	Arg (CGT) vs Leu (CTT)	R3564L	0.13	33	0	5	1.1E-6
96	New SNP	C > T	10711	150543274	NS	Ser (TCA) vs Leu (TTA)	S3571L	0.50	0	45	0	6.8E-13
97	rs12072880	C > G	10735	150543250	NS	Thr (ACG) vs Arg (AGG)	T3579R	0.08	39	5	1	0.4451
98	rs12742178	C > T	10763	150543222	S	His (CAC) vs His (CAT)	H3588H	0.06	39	5	0	1.0
99	rs12083389	G > C	10778	150543207	NS	Glu (GAG) vs Asp (GAC)	E3593D	0.08	38	5	1	0.4547

100	New SNP	T > C	10801	150543184	NS	His (CAT) vs Arg (CGT)	H3601R	0.07	38	6	0	1.0
101	New SNP	A > G	10885	150543100	NS	His (CAC) vs Arg (CGC)	H3629R	0.07	38	6	0	1.0
102	New SNP	C > G	11018	150542967	NS	Asp (GAC) vs Glu (GAG)	D3673E	0.07	38	6	0	1.0
103	New SNP	T > G	11782	150542203	NS	His (CAT) vs Pro (CCT)	H3928P	0.02	43	0	1	0.023
104	rs3814299	C > T	11908	150542077	NS	Ser (TCA) vs Leu (TTA)	S3970L	0.01	47	1	0	1.0

Abbreviations used: ID = Identification number, MAF = Minor allele frequency,  $HMZ_{common}$  = Homozygous for the common allele, HTZ = Heterozygous,  $HMZ_{rare}$  = Homozygous for the rare allele, HWE = Hardy-Weinberg Equilibrium.

Variants highlighted are not in HWE ( $p \leq 0.001$ ).

\* Identified variant number

† UCSC SNP reference number (March 2006)

‡ Position in UCSC Genome Browser (March 2006)

§ NS = non-synonymous ; S = Synonymous

Figure E1

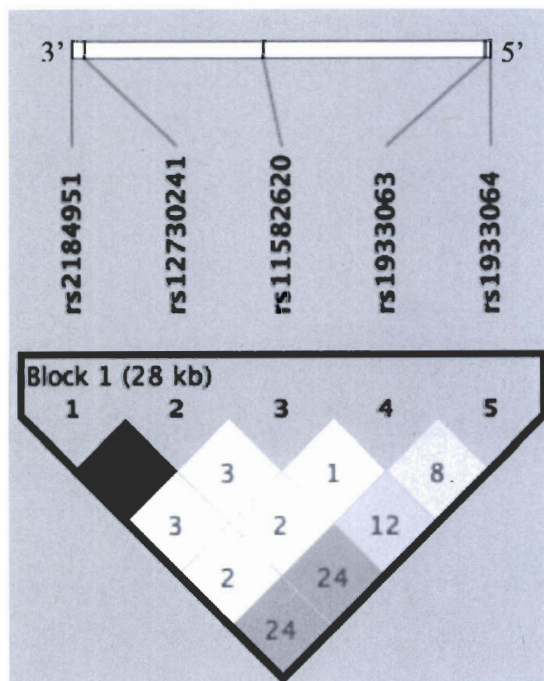




Figure E2

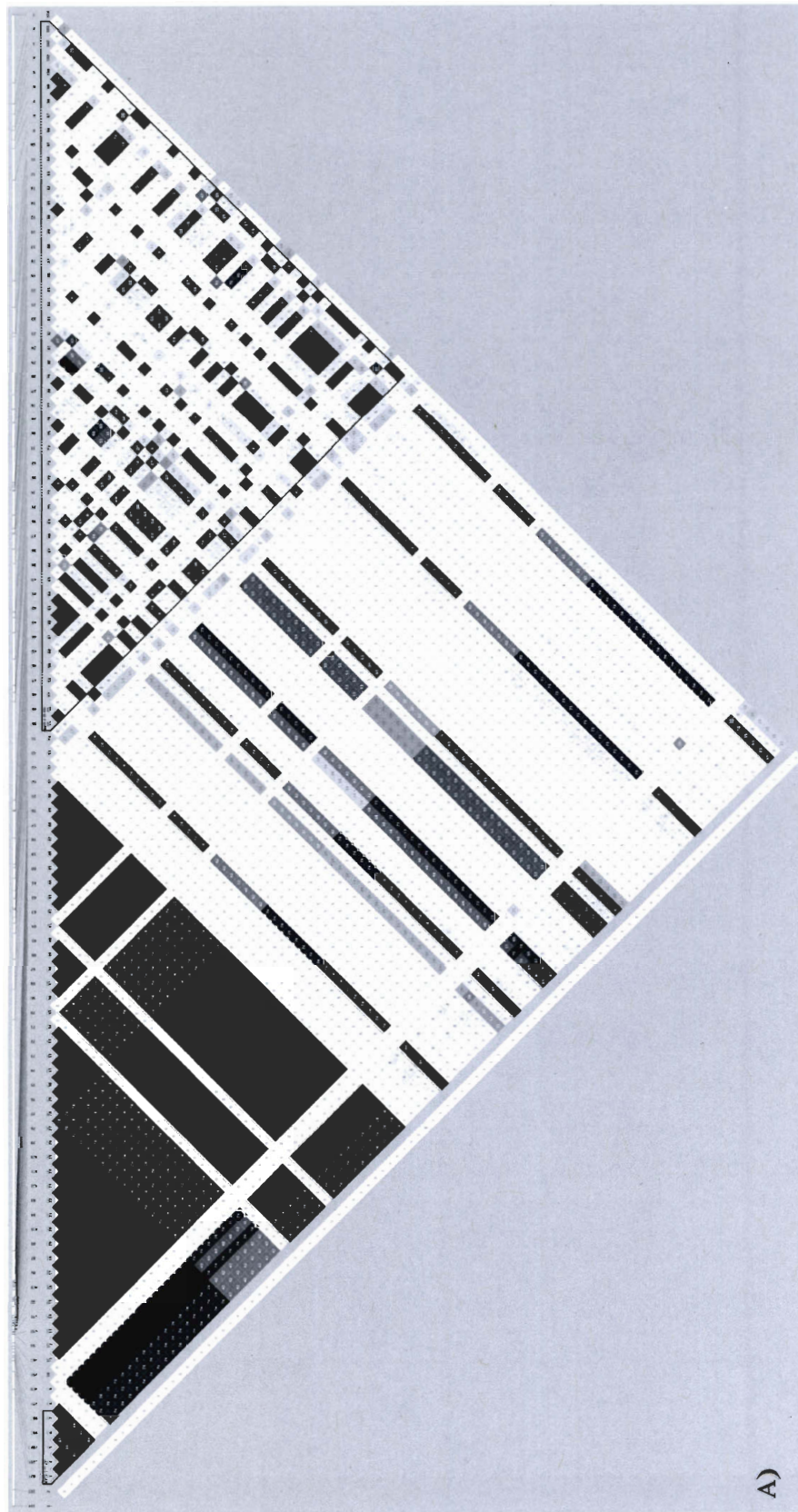


Figure E2 (continued)

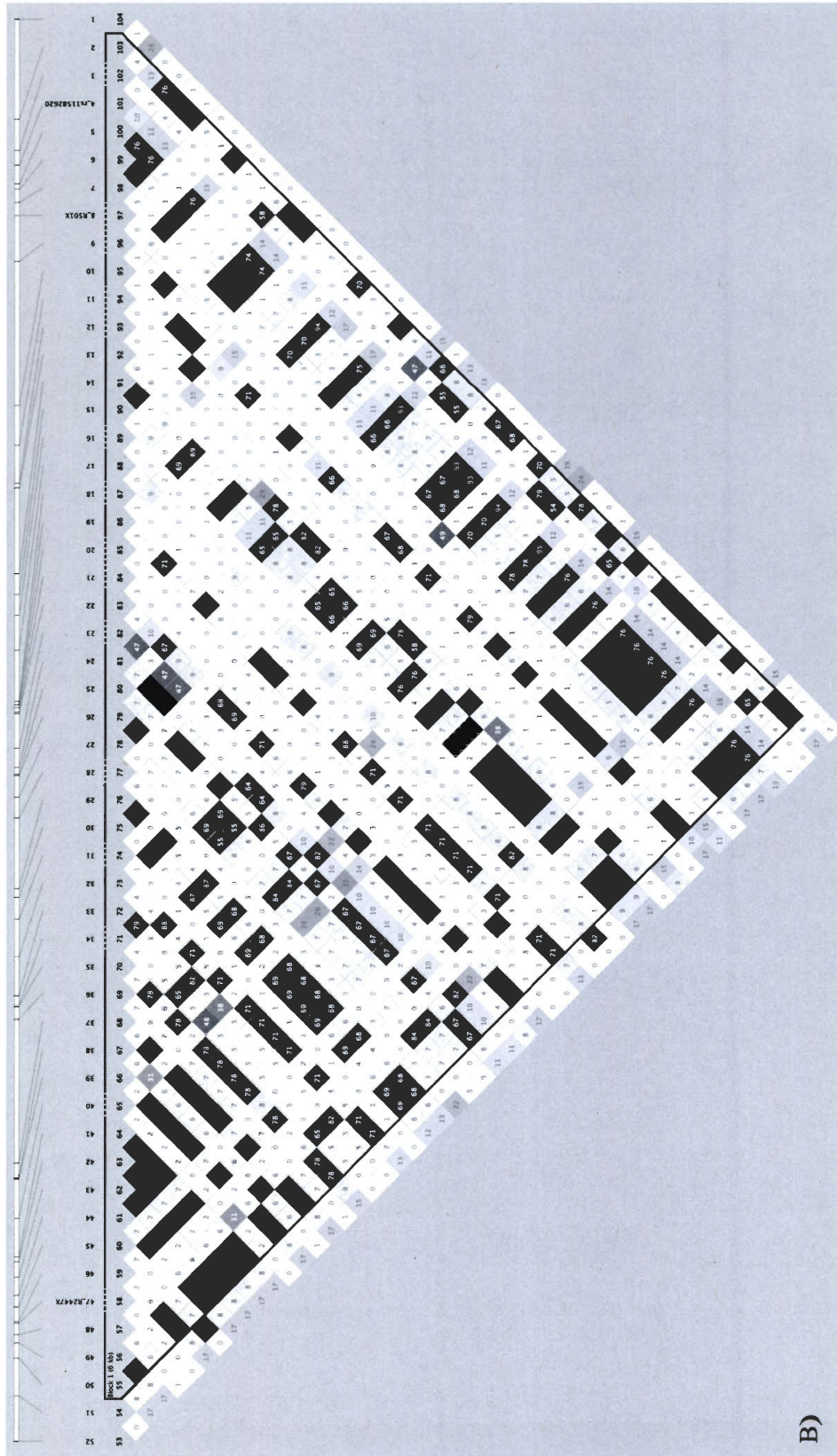
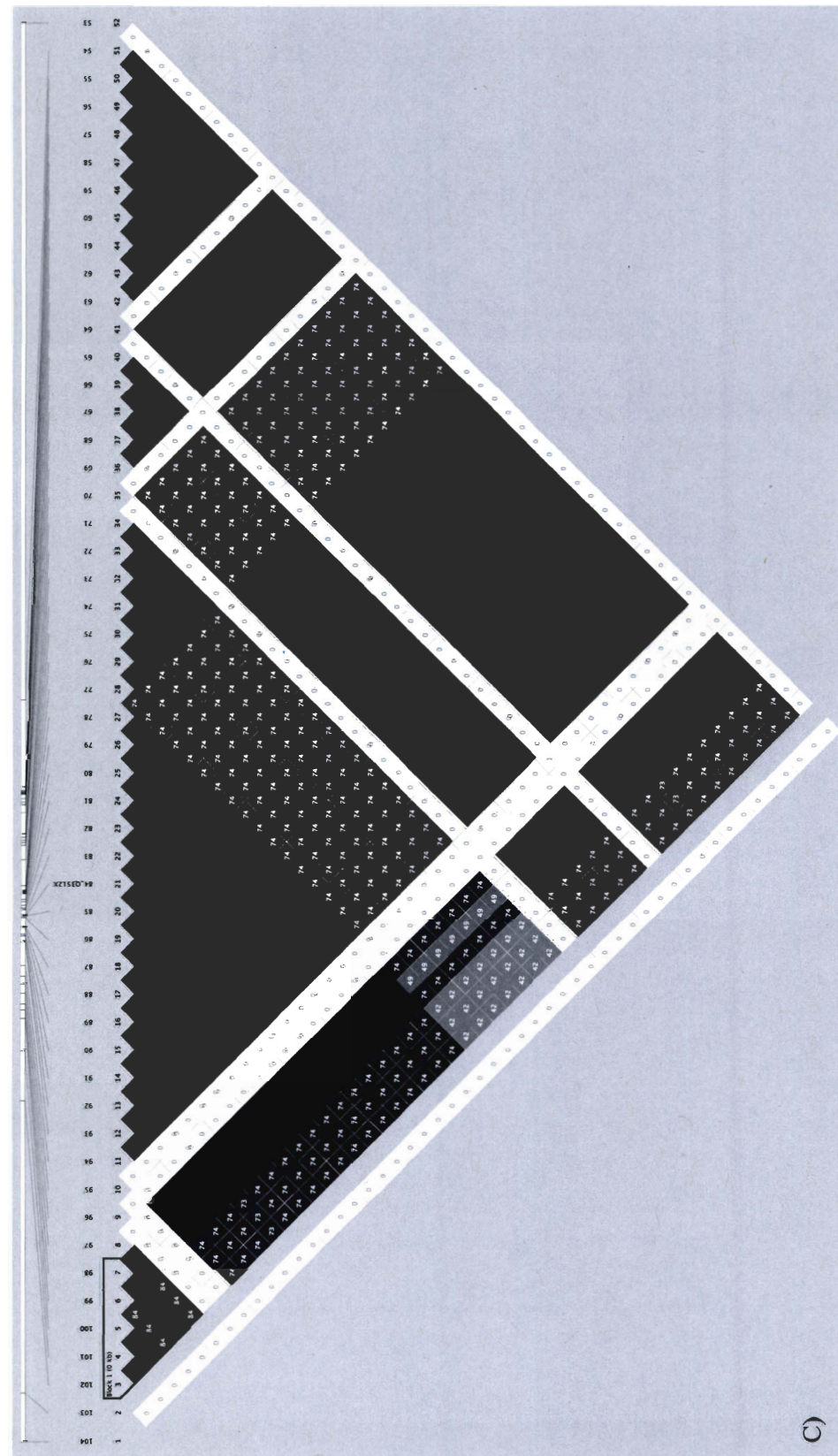


Figure E2 (continued)



6)

## Chapitre 9

### Discussion

Le présent chapitre propose quatre thèmes de discussion en lien avec les objectifs et les résultats présentés dans les articles des Chapitres 5 à 8. Pour commencer, une brève synthèse de l'information génétique recueillie, c'est-à-dire les gènes associés à l'asthme ou à ses phénotypes, sera effectuée afin de rappeler les principales connaissances acquises par le présent projet de recherche. Ensuite, une comparaison des approches employées pour la sélection des gènes candidats sera effectuée, ce qui permettra de répondre à l'avant-dernier objectif des présents travaux. Finalement, les deux derniers thèmes abordés dans cette discussion concernent la principale difficulté rencontrée lors de la réalisation des études d'association d'un trait complexe tel que l'asthme, c'est-à-dire la réplification des découvertes dans des études indépendantes. Comme il sera vu, un nombre considérable de facteurs influencent la distribution des associations parmi les études génétiques, ce qui explique pourquoi certaines associations génétiques peuvent ne pas être répliquées.

#### 9.1 Synthèse de l'information génétique recueillie

L'objectif principal de la présente thèse était l'identification de déterminants génétiques de susceptibilité ou de protection pour l'asthme et ses phénotypes. Pour y arriver, six études d'association, basées sur quatre approches de sélection des gènes candidats, ont été réalisées dans un ou plusieurs échantillons de sujets asthmatiques. Ainsi, un total de 22 gènes candidats, sélectionnés selon une des quatre approches méthodologiques employées, ont été étudiés. À la lumière des résultats présentés dans les Chapitres 5 à 8, six de ces gènes (7/22 ~ 32 % d'efficacité) ont été associés à l'asthme ou à ses phénotypes. Il s'agit des gènes *CX3CR1* (Tremblay et al., 2006), *ALOX15* et *CD14* (Tremblay et al., 2008a), *PLAU* (Begin et al., 2007), *PTPRE* (Tremblay et al., 2008b), *ORMDL3* (Madore et al., 2008) et *FLG* (Tremblay et al., 2008c).

Pour les gènes *CX3CR1*, *ALOX15*, *PLAU* et *PTPRE*, il s'agissait de la toute première association génétique rapportée dans la littérature. D'ailleurs, les gènes *CX3CR1* et *PLAU* font déjà partie de la liste mise à jour des 173 gènes connus pour être associés à

l'asthme ou à ses phénotypes dans au moins une étude d'association génétique (Zhang et al., 2008). En ce qui concerne le gène *ORMDL3*, il s'agissait de la réplication des SNPs, présents au locus 17q21 associés à l'asthme dans une GWAS (Moffatt et al., 2007). Finalement, pour le gène *FLG*, l'association trouvée dans les trios familiaux du SLSJ est l'une des premières concernant le phénotype de l'asthme et sa séquence complète est également la première effectuée chez des sujets asthmatiques (Tremblay et al., 2008c). Le génotypage de toutes les mutations identifiées et validées par le séquençage du gène *FLG* permettra d'identifier la ou les mutations responsables de l'effet protecteur observé.

Un autre objectif des présents travaux était de démontrer l'efficacité des expérimentations de puces à ADN dans l'identification des principaux biomarqueurs différemment exprimés dans les biopsies bronchiques de sujets asthmatiques comparés à des sujets témoins (Chamberland et al., 2008). Cette étude a effectivement permis de démontrer que différentes expérimentations de puces à ADN réalisées avec différents sujets permettent de faire ressortir les biomarqueurs importants dans l'inflammation et le remodelage du tissu bronchique de l'asthmatique tels *CX3CR1* et *ALOX15*, pour lesquels des associations génétiques positives ont également été obtenues.

Les études d'association présentées ont donc contribué à augmenter les connaissances sur les déterminants génétiques de l'asthme et à répondre aux six premiers objectifs spécifiques (Chapitre 4). Toutefois, la découverte de nouveaux déterminants génétiques de susceptibilité ou de protection pour l'asthme n'est que la première étape dans un processus encore plus long (Grant et Hakonarson, 2008). En effet, c'est une ouverture pour les études fonctionnelles qui elles, permettront d'élucider les mécanismes d'action des variants génétiques associés dans la physiopathologie de l'asthme et peut-être même, dans sa pharmacothérapie.

## **9.2 Comparaison des approches méthodologiques employées pour la sélection des gènes candidats étudiés**

L'avant-dernier objectif de la présente thèse est la comparaison des approches méthodologiques employées dans la sélection des gènes candidats nécessaire à la réalisation des études d'association génétique. Étant donné que chacune d'elles a

permis, au bout du compte, de trouver au moins une association génétique positive, l'accent sera porté sur leurs avantages et limites et ce, dans le contexte des résultats obtenus et des observations notées en cours de processus. La comparaison sera traitée de manière plus qualitative puisque le nombre de gènes candidats de départ pour chacune des approches ne permet pas une comparaison quantitative (nombre d'associations positives obtenues sur le nombre de gènes étudiés).

Parmi les approches de sélection des gènes candidats employées dans le présent projet de recherche, quelques-unes sont reconnues comme étant classiques en épidémiologie génétique des traits complexes parce qu'elles sont la « norme » actuelle et ont été employées à grande échelle au cours des dernières décennies. Dans cette catégorie se retrouvent le clonage positionnel, la fonction biologique connue du gène et son association génétique connue avec l'asthme ou à un trait complexe qui partage une même composante physiopathologique. La démonstration de leur efficacité n'est pas à faire, puisqu'un nombre important d'études génétiques basées sur ces approches ont été publiées et qu'il est généralement admis qu'elles sont les méthodes standardisées. Toutefois, ces approches comportent certaines limites, la preuve étant qu'elles n'ont pas permis d'identifier tous les gènes de tous les traits complexes, y compris l'asthme. C'est pourquoi de nouvelles approches, qui combinent les connaissances des approches classiques aux avancées technologiques, sont déployées afin d'optimiser les chances de cibler des gènes. Ces nouvelles approches sont ici considérées comme novatrices, car leur utilisation n'est que récente et que la démonstration de leur efficacité reste à faire dans l'ensemble de la communauté génétique. Dans cette catégorie se retrouvent les expérimentations de puces à ADN et l'outil bioinformatique G2D.

### **9.2.1 Approches classiques**

Pour les approches classiques, elles ont toutes été utilisées en combinaison avec d'autres approches sauf pour deux d'entre elles (Madore et al., 2008; Tremblay et al., 2008c), où l'hypothèse de départ pour la sélection du gène candidat étudié (*ORMDL3* ou *FLG*) étaient basées sur les connaissances génétiques et fonctionnelles des gènes candidats. Ces deux études sont les moins originales des travaux de la présente thèse, mais leur utilité est sans équivoque puisqu'elles ont permis l'identification d'associations génétiques positives. Les approches classiques sont limitées aux

connaissances biologiques, génétiques et fonctionnelles documentées, tandis que leur plus grand avantage réside dans la possibilité de les combiner avec d'autres approches. Par exemple, dans les présents travaux, le clonage positionnel a été combiné aux résultats de l'approche basée sur l'étude de puces à ADN (Begin et al., 2007) et de l'approche basée sur l'outil bioinformatique G2D (Tremblay et al., 2008b), qui seront abordées dans la prochaine section. Un des avantages offert par le clonage positionnel est la possibilité de sélectionner des gènes candidats dont la fonction n'est pas connue et qui n'aurait pas été suspectés autrement que par l'identification d'une région génétique liée au phénotype étudié (Zhang et al., 2008), tel qu'illustré par les gènes *NOX3*, *SNX9* ou *GPR123* (Tremblay et al., 2008b). Toutefois, la plus grande limite à l'utilisation du clonage positionnel réside dans sa dépendance aux connaissances acquises par les études de liaisons préalables. Or, il est reconnu que les études de liaison ont une puissance statistique réduite en ce qui concerne l'identification des régions chromosomiques contenant des gènes qui ont un effet modeste dans la susceptibilité à un trait complexe (Hattersley et McCarthy, 2005). Par conséquent, les gènes présentant un effet modeste sur la susceptibilité à la maladie ont pu ne pas être détectés dans l'étude de liaison effectuée dans des familles du SLSJ. Une puissance statistique plus élevée (par exemple un plus grand nombre de familles incluses dans l'étude de liaison) aurait peut-être permis l'identification de certaines autres régions chromosomiques liées à l'asthme ou à ses phénotypes, régions qui auraient contenu d'autres gènes candidats potentiels.

D'après les résultats obtenus grâce à l'utilisation des approches classiques et de leur combinaison aux autres approches dans la sélection des gènes candidats pour les études d'association, il est possible d'affirmer qu'elles représentent de bonnes hypothèses de départ afin de mener à bien une étude d'association génétique.

### **9.2.2 Approches novatrices**

Pour ce qui est des approches novatrices dans la sélection des gènes candidats, elles ont contribué aux quatre études d'association les plus originales de la présente thèse et elles ont permis de découvrir les premières associations génétiques entre *CX3CR1* (Tremblay et al., 2006), *ALOX15* (Tremblay et al., 2008a), *PLAU* (Begin et al., 2007) et *PTPRE* (Tremblay et al., 2008b) et les phénotypes de l'asthme.

### 9.2.2.1 Profil d'expression génique

L'approche basée sur la sélection des gènes selon leur différence d'expression dans les biopsies bronchiques de sujets asthmatiques comparés aux témoins par puces à ADN (Laprise et al., 2004) s'est avérée un franc succès sur plusieurs aspects. D'abord, pour l'identification de biomarqueurs potentiellement impliqués dans les mécanismes physiopathologiques de l'asthme et de la réponse au traitement anti-inflammatoire (Laprise et al., 2004), dont certains furent répliqués dans une expérimentation de puces indépendante (Chamberland et al., 2008). Ensuite, pour la découverte de trois nouvelles associations génétiques pour l'asthme avec les gènes *CX3CR1* (Tremblay et al., 2006), *ALOX15* (Tremblay et al., 2008a) et *PLAU* (Begin et al., 2007) qui n'auraient pas été sélectionnés comme gènes candidats sans cette connaissance.

Les études du profil d'expression génique dans les tissus de sujets atteints de la maladie étudiée ont révolutionné la recherche biomédicale, car grâce à sa technologie à très haut débit, qui permet de vérifier l'expression de milliers de gènes en une seule expérimentation, il est désormais possible d'identifier de nouveaux biomarqueurs dans les maladies qui représentent de nouvelles cibles thérapeutiques pour l'industrie pharmaceutique ou de nouveaux gènes candidats pour les généticiens (Halapi et Hakonarson, 2004). Toutefois, les études du profil d'expression génique chez l'humain sont limitées par les échantillons biologiques prélevés chez les sujets. Par exemple, dans l'étude de Laprise et collaborateurs (2004), ce sont des biopsies bronchiques qui ont été prélevées chez les sujets asthmatiques et témoins. Or, les biopsies bronchiques asthmatiques, en plus de présenter un épithélium bronchique qui se désagrège plus facilement qu'un épithélium sain à la bronchoscopie (Jeffery et al., 1989), sont un mélange hétérogène de cellules structurales et immunitaires/inflammatoires infiltrées (Rolph et al., 2005). L'utilisation des biopsies bronchiques pourrait donc fournir un portrait partiel de l'expression des gènes impliqués dans l'asthme puisqu'il ne tient pas compte du profil d'expression des cellules présentes dans la lumière bronchique et qui participent aux événements inflammatoires. Cette observation suggère que pour avoir un portrait plus réel de l'activité des gènes dans les poumons des sujets asthmatiques, il y aurait peut-être avantage à combiner les profils d'expression des biopsies bronchiques à celui d'un lavage broncho-alvéolaire, ou même d'un échantillon d'expectoration induite, effectué simultanément chez un même sujet. Ainsi, l'expression des gènes de



tous les types cellulaires structuraux et circulants, y compris les cellules épithéliales détachées à la bronchoscopie, serait prise en considération. D'ailleurs, à ce propos, Rolph et collaborateurs suggèrent que les profils d'expression génétiques employant un tissu particulier pourrait être améliorés en combinant les profils d'expression de types cellulaires spécifiques (Rolph et al., 2005). Une telle approche n'a pas encore été employée dans les études de profil d'expression pour l'asthme, mais sa valeur a été clairement démontrée dans d'autres maladies (Rolph et al., 2005). Elle permettrait peut-être d'observer des différences d'expression significatives des gènes de cytokines pro-inflammatoires exprimés par les cellules circulantes (comme l'IL-4 ou l'IL-13), par exemple, pour lesquelles aucune différence ne fut observée dans le profil d'expression des biopsies bronchiques de sujets asthmatiques comparé à celui des témoins (Laprise et al., 2004). À cet effet, une étude comparative du patron d'expression des macrophages alvéolaires prélevés dans les lavages broncho-alvéolaires de sujets asthmatiques et témoins est actuellement en cours et devrait permettre de compléter, du moins en partie, le profil d'expression obtenu avec les biopsies bronchiques (Laprise et al., 2004).

Malgré cette limite, l'utilisation des biopsies bronchiques pour les expérimentations de puces à ADN a permis de mettre en évidence 79 gènes différemment exprimés chez les asthmatiques (Laprise et al., 2004). À défaut de ne pouvoir identifier leur source cellulaire dans les présentes études autrement que par la revue de littérature, ils peuvent être globalement identifiés comme des biomarqueurs d'inflammation ou de remodelage dans la physiopathologie de l'asthme. Leur sélection comme gènes candidats pour les études génétiques reposent sur l'hypothèse que leur différence d'expression s'explique par des déterminants génétiques. Des neuf gènes candidats investigués selon cette hypothèse et selon leur fonction biologique, deux gènes de molécules immunes, *CX3CR1* et *ALOX15*, ont montré une association génétique positive qu'il reste à caractériser au niveau fonctionnel (Tremblay et al., 2006; Tremblay et al., 2008a). Pour les sept autres gènes, si des facteurs génétiques ne peuvent expliquer cette différence, alors d'autres facteurs le peuvent (consulter la discussion (Chamberland et al., 2008)). Par ailleurs, un autre gène différemment exprimé dans l'étude de puces de Laprise et collaborateurs (2004), *SERPINB2*, n'a montré aucune association génétique avec l'asthme (résultats non présentés). Mais un gène impliqué dans sa voie biologique et localisé dans une région chromosomique liée à l'asthme, *PLAU*, a été associé à l'asthme

(Begin et al., 2007). Sans la connaissance préalable de la différence d'expression de *SERPINB2* et de sa voie biologique, le gène *PLAU* n'aurait pas été ciblé, démontrant une fois de plus l'utilité des expérimentations de puces à ADN dans la sélection de gènes candidats pour l'asthme.

#### 9.2.2.2 Outil G2D

L'utilisation de l'outil G2D s'est également révélée une bonne approche de sélection de gènes candidats, puisqu'elle a permis de prioriser des gènes localisés dans deux régions chromosomiques liées à l'asthme et à l'atopie et ainsi, d'identifier la première association génétique du *PTPRE* avec l'asthme allergique au SLSJ (Tremblay et al., 2008b).

L'outil G2D a été développé *a priori* pour les maladies monogéniques disponibles dans OMIM (Perez-Iratxeta et al., 2002) mais aujourd'hui, il est applicable à tous les phénotypes multigéniques répertoriés dans OMIM (Perez-Iratxeta et al., 2007). En ce qui concerne l'asthme, ce fut le premier trait complexe étudié avec G2D et la preuve du principe fut démontrée à l'aide du gène *ADAM33* (Tableau I, page 50) (Perez-Iratxeta et al., 2002), situé dans la région 20p13 liée à l'asthme et à l'hyperréactivité bronchique (Van Eerdewegh et al., 2002) et associé aux phénotypes de l'asthme dans au moins cinq études indépendantes (Vercelli, 2008). Bien que l'outil G2D a permis de fournir une liste de gènes candidats au potentiel biologique non négligeable dans la physiopathologie de l'asthme et de permettre la première association du gène *PTPRE* avec l'asthme allergique dans nos travaux (Tremblay et al., 2008b), la méthode G2D présente certaines limites qui peuvent affecter la qualité de la liste de gènes produite, qui pourrait être incomplète ou contenir des gènes non pertinents.

En premier lieu, l'algorithme G2D est construit à l'aide des données obtenues des analyses de liaison de différentes maladies monogéniques et disponibles grâce à OMIM. Donc, lorsque l'on interroge G2D pour un phénotype particulier, ces régions génomiques connues et implémentées constituent la référence. Pour l'asthme, plusieurs régions chromosomiques sont connues pour être liées à la maladie, puisqu'au moins 19 études de liaison ont déjà été complétées (Hoffjan et Ober, 2002; Blumenthal, 2005; Steinke et al., 2008). Pour composer avec cette information, les concepteurs de

l'algorithme G2D se sont basés sur une autre banque de données d'études de liaison spécifique pour l'asthme et ses phénotypes, appelée *Asthma and Allergy Gene Database* (Immervoll et Wjst, 1999). Cette base de données n'est maintenant plus en service et donc, les régions chromosomiques liées à l'asthme ne sont pas à jour dans les registres utilisés par G2D. Lorsqu'un gène est localisé dans une région nouvellement liée à l'asthme, de nouvelles données de fonctions biologiques du gène sont ajoutées dans les bases de données publiques interrogées par G2D comme celle du Medline, de OMIM ou de RefSeq (Perez-Iratxeta et al., 2002; Perez-Iratxeta et al., 2005). Par conséquent, si l'information est trop récente, en interrogeant G2D pour des régions chromosomiques et phénotypes d'intérêt, ces nouvelles connaissances biologiques ne seront pas prises en considération et des gènes qui auraient eu une fonction potentielle n'apparaîtront pas dans la liste de gènes candidats priorisés par G2D. Dans le présent cas, cette remarque signifie que pour les régions 6q26 et 10q26.3, il y aurait une possibilité que la liste de gènes priorisés par G2D soit incomplète et que d'autres gènes auraient pu être ciblés avec une banque de données plus à jour.

En second lieu, la méthode des phénotypes de G2D base son recoupage sur la littérature biomédicale via PubMed et est limitée aux études d'association publiées. Ce qui peut poser problème sachant que les études d'association répertoriées dans cette base de données littéraires ne sont pas jugées sur leur qualité par l'algorithme G2D, c'est-à-dire qu'il n'y a aucune évaluation du respect ou non des critères standardisés de la réalisation des études d'association. En fait, comme il sera vu dans les prochaines sections de ce chapitre, les associations génétiques sont variables d'une étude à l'autre en raison de plusieurs facteurs. De nos jours, le plus haut niveau de confiance pour une association génétique est sa réplique dans plusieurs études indépendantes (Newton-Cheh et Hirschhorn, 2005; Holloway et Koppelman, 2007; McCarthy et al., 2008). Or, la plupart des études d'association répertoriées dans PubMed avant 2004-2005 n'étaient pas soumises aux mêmes standards qu'aujourd'hui. En effet, avant cette période, les corrections statistiques pour tests multiples et la réplique dans une étude indépendante n'étaient pas des critères exigés par la plupart des journaux pour publier une étude d'association. Par conséquent, un grand nombre d'études d'association qui se retrouvent actuellement dans les bases de données se révèlent être de faux-positifs, puisque non-répliquées (Ioannidis et al., 2001; Hirschhorn et al., 2002). Ces articles font tout de même partie des banques de données que G2D utilise pour faire son recoupage.

Et l'inverse est aussi vrai. Avant 2004-2005, il n'était pas rare de voir des études d'association présentant des associations négatives avec les phénotypes de l'asthme, ce qui n'est maintenant pratiquement plus le cas (tel qu'il sera vu plus loin). La publication de ces résultats négatifs permet en particulier de confirmer une association positive précédente ou de l'infirmer, permettant la mise en évidence d'associations faux-positifs dans la littérature. Pour ce qui est de l'algorithme G2D, il ne distingue pas la nature de l'association (positive ou négative) et il ne tient pas compte de sa réplification dans une étude indépendante ainsi que de l'origine ethnique des populations étudiées. Par conséquent, ce problème d'évolution des critères standardisés pour la réalisation des études d'association ou de publications manquantes, peut soit empêcher la mise en évidence de gènes importants localisés dans la région chromosomique interrogée pour un phénotype donné ou soit cibler des gènes qui n'ont pas une si grande pertinence pour le phénotype étudié. Compte tenu de cette limite et afin d'optimiser l'étude d'association basée sur l'outil G2D de la présente thèse (Tremblay et al., 2008b), une revue de littérature mise à jour a été effectuée pour chaque gène priorisé afin de sélectionner les meilleurs candidats en fonction de leur rôle biologique potentiel dans la physiopathologie de l'asthme.

Ces limites de l'utilisation de l'outil G2D indiquent la nécessité d'une certaine vigilance dans le choix des gènes priorisés pour le génotypage. Dans le cas présent, comme c'était la première utilisation du G2D dans la sélection des gènes candidats, le réflexe fut de concentrer les efforts sur les dix meilleurs potentiels biologiques selon la connaissance actuelle de la physiopathologie de la maladie. Cette première étape a permis l'identification du gène *PTPRE* (Tremblay et al., 2008b). Il reste maintenant à savoir si les autres gènes qui n'ont pas été retenus pour l'étude d'association auraient également été associés. Comme leur probabilité d'être impliqués dans la biologie de la maladie est non nulle, ils représentent tous des hypothèses de départ pour des études génétiques subséquentes et le potentiel de nouveaux déterminants génétiques de l'asthme. Tout comme pour le profil d'expression génique, lorsque tous les gènes de la liste initiale produite par G2D pour la présente étude seront évalués, une efficacité de la méthode pourra être mesurée et comparée aux autres approches.

Pour terminer, l'approche par gènes candidats employée dans les présentes études d'association est maintenant considérée comme une « vieille » approche

comparativement à l'approche génomique des GWAS (Moffatt, 2008). Ces dernières présentent l'avantage d'évaluer tous les gènes à la fois, mais en contrepartie, elles nécessitent un échantillon de très grande taille pour que la puissance à détecter les associations soit suffisante et ce, à des seuils de signification très petits. Dans le contexte actuel, l'obtention de ces échantillons de milliers de sujets n'est possible que par la collaboration entre chercheurs afin de combiner les différents échantillons existants. Toutefois, tel qu'il en sera discuté dans les prochaines sous-sections, plusieurs facteurs expliquent la variabilité dans la distribution des associations. Ces sources de variation peuvent être difficiles à contrôler dans la combinaison d'échantillons de sujets construits différemment pour des fins d'analyses combinées. L'information spécifique à chaque échantillon se retrouve alors diluée dans un échantillon plus puissant pour l'échelle génomique, mais qui est, en revanche, moins informatif à l'échelle génétique. De plus, la réalisation d'une GWAS est très coûteuse et le ratio d'un gène associé (exemples (Moffatt et al., 2007; Ober et al., 2008) ) sur les 25 000 à 30 000 possibles est très peu bénéfique en terme de génération de nouvelles hypothèses pour les études fonctionnelles subséquentes (Shriner et al., 2007). Par exemple, en se basant sur la connaissance que l'asthme implique plus qu'un gène dans son développement et sa susceptibilité (Elston, 1995; Sandford et al., 1996; Silverman et Palmer, 2000; Cookson, 2002) et qu'un minimum de 173 gènes sont connus pour être associés à l'asthme et à ses phénotypes (Zhang et al., 2008), il est logique de s'attendre à ce qu'une GWAS identifie au moins plus d'un gène associé. Une telle performance pourrait conduire les chercheurs à réutiliser les millions de données génotypiques produites pour stratifier les analyses en fonction des échantillons indépendants sur la base de leurs différences environnementales, phénotypiques ou génétiques. La formulation des hypothèses logiques et plausibles en fonction de la physiopathologie de la maladie sera donc requise pour sélectionner de bons gènes candidats qui serviront à réaliser les études d'association. Ainsi, même si aujourd'hui l'approche par gènes candidats est considérée limitée comparativement aux GWAS, elle s'avère toujours utile, dans un esprit de complémentarité, pour la recherche des gènes des traits complexes.

En somme, il n'existe aucun paradigme relatif à la découverte des gènes pas plus qu'il n'existe un type d'étude unique ou une approche analytique idéale (Palmer et Cardon, 2005). En dépit des affirmations *ex cathedra* de la communauté scientifique sur les conceptions d'étude d'association optimales, il est clair que des approches combinées et

flexibles selon la physiopathologie de la maladie et des technologies de pointe disponibles seront les plus souhaitables dans la recherche et la compréhension des déterminants génétiques des traits complexes tel que l'asthme (Palmer et Cardon, 2005). Les résultats apportés par les travaux de la présente thèse ont soulevé quelques observations et remarques qui pourraient modifier la façon d'aborder l'approche par gènes candidats. En fait, comme chacune des approches employées a contribué à l'identification ou à la réplication d'une association génétique pour l'asthme, le constat final est qu'il y a avantage à exploiter toutes les connaissances et outils disponibles plutôt que de confiner les efforts sur une seule approche. Bref, une approche multidisciplinaire devrait être privilégiée tant en termes d'approches méthodologiques variées que de domaines d'expertise diversifiés (clinique, biomédicale, statistique, génomique, transcriptomique, protéomique, etc.).

### **9.3 Difficultés rencontrées dans les études d'association pour l'asthme**

La recherche de déterminants génétiques pour l'asthme fait actuellement face à plusieurs éléments de discordance dans la communauté de généticiens. D'un côté, certains clament le besoin de développer des méthodes standardisées, car la littérature abonde de gènes associés à l'asthme et la limite pour savoir lesquels sont de véritables gènes n'est pas franchement définie. D'un autre côté, la complexité inhérente à la maladie, son hétérogénéité clinique et sa génétique complexe ne facilitent pas les recherches. En fait, la difficulté n'est pas tant l'identification des associations génétiques, mais bien la réplication dans des études indépendantes qui sont le plus souvent réalisées sous l'effet de multiples facteurs affectant la distribution des associations au sein des diverses populations. Ces facteurs sont responsables des trois principaux degrés de réplication observés dans la littérature, à savoir la réplication stricte, la réplication partielle et la non-réplication (Tableau IV). En ce qui concerne la réplication stricte ou partielle, elles peuvent être évaluées en regard du génotype ou du phénotype, ou des deux.

Les études d'association présentées dans cette thèse ont fait l'objet d'étude de réplication indépendante et les résultats se partagent entre les trois degrés de réplication, résumés dans le Tableau V, tel qu'il sera abordé dans les prochaines sous-sections.

**Tableau IV**

Matrice de réplification utilisée pour évaluer le degré de réplification d'une étude d'association effectuée pour répliquer des observations initiales (traduit de Holloway et Koppelman, 2007).

		Phénotype	
		Stricte	Partielle
Génotype	Stricte	Même SNP, même phénotype	Même SNP, différent phénotype
	Partielle	Différent SNP, même phénotype	Différent SNP, différent phénotype

**Tableau V**

Résumé des associations et de leur réplification au sein des divers échantillons étudiés.

Gène	Association			Type de réplification *	Facteurs de variation possible †
	Échantillon	Phénotype	SNP		
<i>CX3CR1</i> (Tremblay et al., 2006)	SLSJ	Asthme	rs2669849, V249I	Stricte	Conception de l'échantillon
	Vancouver	Asthme	rs2669849, V249I		
<i>ALOX15</i> (Tremblay et al., 2008a)	CAPPS	HRB	rs7217186	Partielle	Conception de l'échantillon Structure génétique des populations
	SAGE	Asthme allergique	rs2664593		
	SLSJ	N/A	N/A	Non répliquée	Interaction G X G Interaction G X E Hétérogénéité phénotypique
	Busselton	N/A	N/A		
<i>CD14</i> (Tremblay et al., 2008a)	SAGE	Asthme	rs9414	Non répliquée	Conception de l'échantillon Structure génétique des populations Interaction G X G Interaction G X E Hétérogénéité phénotypique
	SLSJ, CAPPS, Busselton	N/A	N/A		
<i>PLAU</i> (Begin et al., 2007)	SLSJ	Asthme et HRB	rs2227564, rs2227566	Partielle	Structure génétique des populations Interaction G X G Interaction G X E
	SAGE	Atopie	rs4065		
	CAPPS	N/A	N/A	Non répliquée	Hétérogénéité phénotypique
<i>PTPRE</i> (Tremblay et al., 2008b)	SLSJ	Asthme allergique	rs7081735	Non répliquée	Structure génétique des populations Interaction G X G Interaction G X E Hétérogénéité phénotypique
	CAMP	N/A	N/A		
<i>ORMDL3</i> (Madore et al., 2008)	SLSJ	Asthme	rs4378650, rs8076131	Stricte	N/A
	GWAS (Moffatt et al., 2007)	Asthme	rs4378650, rs8076131		

HRB = Hyperréactivité bronchique, G X G = Gène-Gène, G X E = Gène-Environnement et N/A = Non applicable.

\* Tel que défini dans le Tableau IV et décrit dans les sections 9.3.1 à 9.3.3.

† Tel que décrit dans la section 9.4.

### 9.3.1 La réplication stricte

Le plus haut niveau de confiance pour une association génétique est sa réplication dans une étude indépendante. Idéalement, les plus fortes évidences pour un rôle potentiel du marqueur génétique dans la pathogénèse de la maladie, sont les réplifications des associations d'un même allèle, pour un même phénotype, dont l'effet est le même (de susceptibilité ou de protection) dans tous les échantillons indépendants testés et ce, avec des valeurs de signification des tests statistiques suffisamment petites pour résister à des corrections statistiques pour tests d'hypothèses multiples à l'échelle génomique (Newton-Cheh et Hirschhorn, 2005; Holloway et Koppelman, 2007; McCarthy et al., 2008). Toutefois, la réalité est tout autre. En dépit des efforts et des avancements réalisés dans la recherche de déterminants génétiques de l'asthme, ce seuil de « perfection » est rarement atteint. Même les gènes les plus répliqués, comme *IL13*, *IL4* ou *IL4RA* (Tableau I, page 50), échouent ce test de réplication parfaite dans toutes les études d'association indépendantes effectuées (Ober et Hoffjan, 2006; Hersh et al., 2007).

Cette observation d'une réplication stricte non-obtenue dans toutes les études de réplication indépendantes ne fait pas exception pour les gènes associés dans la présente thèse. En fait, parmi les études d'association présentées, deux réplifications strictes ont été observées en regard du génotype et du phénotype. La première est celle obtenue pour le gène *CX3CR1*. En effet, l'association observée pour l'asthme dans les trios familiaux du SLSJ a été répliquée pour les SNPs rs2669849 et V249I dans un échantillon indépendant provenant de Vancouver (Tremblay et al., 2006). Suivant cette publication, l'association génétique observée pour les variants du *CX3CR1* a été vérifiée dans deux autres études réalisées indépendamment de l'originale. C'est ainsi qu'en 2007, un groupe de chercheurs allemands rapportait une association positive pour les deux mutations de *CX3CR1*, T280M et V249I, avec l'atopie et non l'asthme dans une cohorte d'enfants asthmatiques d'origine allemande et vivant à Munich (n = 1 159) ou à Dresden (n = 1 940) (Depner et al., 2007). Il s'agit dans ce cas d'une réplication partielle de l'association initiale (même SNP, différents phénotypes). Encore plus récemment, une étude d'association pour l'asthme effectuée dans deux échantillons de sujets asthmatiques d'origine tunisienne, un de cas-témoins (121 enfants asthmatiques et 226 enfants non-atteints appariés pour l'âge) et un de trios familiaux (99 familles nucléaires, 297 individus), n'a révélé aucune association positive (Chelbi et al., 2008).



Il s'agit donc d'une non-réplication. D'autres études d'association génétique indépendantes sont actuellement requises pour déterminer l'importance du *CX3CR1* comme facteur de prédisposition génétique pour l'asthme.

La deuxième réplication stricte observée dans les travaux de la présente thèse est celle de l'étude de réplication des SNPs du locus 17q21 (Madore et al., 2008), contenant le gène *ORMDL3* associé à l'asthme dans une GWAS réalisée dans deux échantillons, un familial d'origine britannique et un cas-témoins d'origine allemande, et répliquée dans deux cohortes indépendantes (Moffatt et al., 2007). La réplication stricte des huit mêmes SNPs (incluant les rs4378650 et rs8076131 du gène *ORMDL3*), pour le même phénotype, dont la direction de l'effet est le même et dont les valeurs-p résistent à une correction Bonferroni confirme l'importance de ce locus dans la susceptibilité génétique à l'asthme (Madore et al., 2008). Malgré ce succès, d'autres études sont souhaitables pour définir la structure moléculaire du gène *ORMDL3* et le rôle de ses mutations dans la physiopathologie de l'asthme.

### 9.3.2 La réplication partielle

Dans les travaux de la présente thèse, deux des études ont permis d'observer la distribution des associations génétiques au sein de trois ou de quatre échantillons indépendants (Begin et al., 2007; Tremblay et al., 2008a). En ce qui concerne l'association positive du *ALOX15*, elle a été observée dans deux échantillons sur quatre (dans CAPPS et SAGE et non dans SLSJ et Busselton), où une réplication partielle a également été observée entre les échantillons CAPPS et SAGE (Tremblay et al., 2008a). Dans les deux études, la réplication est partielle autant en ce qui concerne le génotype que le phénotype, c'est-à-dire que les associations des gènes *PLAU* et *ALOX15*, lorsque répliquées, concernent différents SNPs et différents phénotypes dans les différents échantillons étudiés. Ces études d'association ayant été récemment publiées, aucune étude indépendante n'est encore répertoriée dans la littérature comme ce fut le cas pour *CX3CR1*. Pour l'association positive du gène *PLAU* avec l'asthme et l'hyperréactivité bronchique observée dans l'échantillon de trios familiaux du SLSJ, une réplication partielle a été observée dans l'échantillon SAGE, et une non-réplication a été observée dans l'échantillon CAPPS (Begin et al., 2007). D'autres études d'association génétique

indépendantes sont donc requises pour déterminer l'importance des gènes *PLAU* et *ALOX15* comme facteurs de prédisposition génétique pour l'asthme.

### 9.3.3 La non-réplication

L'étude d'association de la présente thèse qui a permis d'identifier l'association génétique pour le gène *PTPRE* dans l'échantillon de trios familiaux du SLSJ fut non répliquée dans l'échantillon CAMP (Tremblay et al., 2008b). Tout comme pour *CX3CR1*, *ALOX15* ou *PLAU*, des études génétiques subséquentes avec le gène *PTPRE* sont nécessaires pour élucider son rôle en tant que facteur de risque génétique pour l'asthme ou ses phénotypes.

Le manque de constance globalement relevé dans la réplication des associations génétiques assombri les observations scientifiques dans ce domaine de recherche et ce, peu importe le trait complexe étudié (Chanock et al., 2007; Vercelli, 2008). D'ailleurs, l'éditorial du journal *Nature Genetics* exprimait certaines inquiétudes vis-à-vis l'augmentation croissante des études génétiques non-répliquées (Nature Genetics Editorial, 1999). Il a même été avancé que les études d'association génétique des traits complexes ont acquis une mauvaise réputation au fil des années (Hall et Blakey, 2005), en raison de ce manque de reproductibilité (Hirschhorn et al., 2002; Newton-Cheh et Hirschhorn, 2005). Ce problème suscite une telle mobilisation, que la littérature a vu apparaître des articles consacrés à la définition des plus hauts standards et exigences pour publier des études d'association considérées comme pertinentes et rigoureuses (Nature Genetics Editorial, 1999; Hall et Blakey, 2005; Chanock et al., 2007). Le problème de la non-réplication dans de multiples populations est donc considéré comme récurrent dans les études génétiques de l'asthme (Steinke et al., 2008) et des traits complexes en général (Ioannidis et al., 2001; Hirschhorn et al., 2002).

Ainsi, la réplication des associations génétiques positives est l'une des étapes les plus difficiles à franchir pour confirmer le statut de déterminants génétiques de susceptibilité ou de protection d'un gène associé dans une étude initiale. En fait, le nombre d'études de réplication indépendantes et le nombre de populations différentes requis pour que les associations génétiques soient considérées strictement répliquées font encore aujourd'hui l'objet de réflexions, car aucun consensus n'a été établi à ce sujet. La

quantité imposante de réplifications partielles ou de non-réplifications témoigne de cette difficulté dans la recherche de déterminants génétiques des traits complexes (Ioannidis et al., 2001; Hirschhorn et al., 2002). Bien qu'une réplification partielle ou une non-réplification soit généralement considérée comme révélatrice d'une association initiale faux-positif (McCarthy et al., 2008), d'autres considèrent, au contraire, que les associations initiales ne sont pas nécessairement de faux-positifs (Shriner et al., 2007). Cette difficulté de la réplification révèle plutôt la présence d'une multitude de facteurs responsables, individuellement ou collectivement, d'une distribution variable des associations génétiques au sein des diverses études et qui expliquent comment une association initiale peut être partiellement ou même non répliquée. Ces facteurs font l'objet de la prochaine sous-section.

#### **9.4 Facteurs influençant la distribution des associations génétiques parmi les études indépendantes**

Toute association génétique initiale non-répliquée dans une étude indépendante peut être le résultat de trois raisons fondamentales : 1) l'association initiale est un faux-positif et est correctement non-répliquée dans une étude indépendante ; 2) l'association initiale est vraie et est faussement non-répliquée dans une étude indépendante (faux-négatif); et 3) l'association initiale est vraie dans la population étudiée et non vraie dans la seconde population (Newton-Cheh et Hirschhorn, 2005). L'apparition des faux-positifs et des faux-négatifs est en lien avec des facteurs de variabilité (ou à des biais) liés à la nature expérimentale des études d'association (comme la conception de l'étude, les erreurs de génotypage, la puissance statistique des échantillons, les corrections statistiques employées ou les biais de publication), tandis que les associations spécifiques aux populations sont en lien avec des facteurs de variabilité liés à la structure génétique des populations étudiées. Les autres facteurs qui peuvent être responsables de l'apparition de réplifications variables au sein des études d'association indépendantes font appel à la complexité inhérente du trait complexe étudié, à savoir ses interactions gène-gène et gène-environnement ainsi que son hétérogénéité phénotypique.

#### 9.4.1 Conception de l'échantillon

La conception des échantillons cas-témoins ou de trios familiaux est soumise à plusieurs biais de sélection ou de classification des sujets si le recrutement et l'évaluation n'est pas effectuée selon des critères standardisés, puisque la manière dont les sujets sont évalués est d'une importance capitale pour permettre la comparaison des associations entre les différentes études (Cordell et Clayton, 2005; Ober et Hoffjan, 2006; McCarthy et al., 2008). De plus, il faut tenir compte de la structure génétique particulière des différents groupes ethniques des sujets participant à l'étude sinon, des biais liés à la stratification des populations résulteront en des associations confondantes (Lander et Schork, 1994; Howson et al., 2005; McCarthy et al., 2008). Ces différents biais ont été abordés dans le Chapitre 3. En ce qui concerne les études d'association de la présente thèse, deux des échantillons étudiés, celui de Vancouver (Tremblay et al., 2006) et de Busselton (Tremblay et al., 2008a), pourraient contenir un ou plusieurs de ces biais, ce qui pourrait expliquer une association génétique confondante.

Pour l'échantillon cas-témoins de Vancouver, utilisé pour répliquer l'association du *CX3CR1* avec l'asthme dans l'échantillon de trios familiaux du SLSJ (Tremblay et al., 2006), les sujets inclus dans cette étude sont les parents des enfants à haut risque de développer de l'asthme recrutés dans un programme d'intervention et de prévention de l'asthme et leur statut asthmatique a été évalué par un questionnaire (Chan-Yeung et al., 2000). Un biais de classification est donc possible car aucun pneumologue n'a confirmé le diagnostic et aucune mesure physiologique comme le VEMS n'a été effectuée, ce qui fait que des personnes asthmatiques ont pu être classifiées comme témoins et vice versa. Cette situation peut également entraîner un biais de sélection des sujets : des individus mal classifiés se retrouvent dans les mauvais groupes de sujets (cas ou témoins), résultant en des associations faux-positifs (Chanock et al., 2007). Ces biais de classification et de sélection potentiels pourraient être responsables d'une répllication faux-positif, en ce sens où la répllication positive de l'association du *CX3CR1* observée dans l'échantillon de Vancouver pourrait ne pas avoir lieu. Une autre source potentielle de biais liée à cet échantillon est celui de la stratification de la population (Lander et Schork, 1994; Howson et al., 2005; McCarthy et al., 2008). En fait, pour minimiser au maximum ce type de biais dans une étude cas-témoins, les individus doivent être appariés selon leur origine ethnique ou géographique (Cordell et Clayton, 2005) où des contrôles génomiques, c'est-à-dire des marqueurs génétiques ancestraux informatifs des

groupes ethniques ou raciaux, doivent être génotypés afin de vérifier la présence de la stratification et de la corriger s'il y a lieu (Pritchard et Rosenberg, 1999; Devlin et al., 2001). Dans le présent cas, des sous-groupes de la population sont présents à l'intérieur de l'échantillon de Vancouver, car 20 % de la cohorte est d'une autre origine ethnique que caucasienne (Chan-Yeung et al., 2000). Les analyses statistiques effectuées avec cet échantillon n'ont pas tenu compte des diverses origines ethniques présentes dans le groupe de sujets cas et dans le groupe de sujets témoins pour les analyses. L'association positive pourrait donc être confondante, c'est-à-dire qu'elle pourrait être représentative de la fréquence allélique d'un groupe ethnique particulier et non représenter une différence en fonction de la maladie. Donc, la réplication effectuée dans cet échantillon, même si elle est considérée comme stricte, pourrait être un faux-positif, ce qui fait que l'association positive du *CX3CR1* observée dans l'échantillon familial du SLSJ aurait pu ne pas être répliquée, ce qui concorderait avec les observations de Chelbi et collaborateurs (Chelbi et al., 2008).

Pour l'échantillon cas-témoins de Busselton, un des quatre échantillons utilisés dans l'étude d'association des gènes candidats identifiés par les puces à ADN et qui a permis d'associer le gène *ALOX15* (Tremblay et al., 2008a), les individus sont tous caucasiens d'origine européenne. La stratification de la population n'est donc pas en cause pour expliquer la non-réplication, même si l'utilisation des contrôles génomiques aurait permis d'en être certain hors de tout doute (Pritchard et Rosenberg, 1999; Devlin et al., 2001). Aucune association positive n'a été trouvée avec cet échantillon. Peut-être que ces résultats sont de faux-négatifs et que s'il n'y avait pas ce possible biais de sélection des sujets témoins, une association positive aurait pu être observée pour *ALOX15* dans l'échantillon Busselton. En effet, dans cet échantillon, certains sujets servent à la fois de cas ou de témoins selon le sous-phénotype étudié. Par exemple, un individu non-asthmatique mais allergique, sert de sujet témoin lorsque le phénotype asthme est analysé et sert de sujet cas lorsque le phénotype allergie est analysé. Dans l'analyse pour l'asthme, si l'individu témoin présente un phénotype intermédiaire (dans ce cas, l'allergie) comment une association significative pourrait être mise en lumière si les deux groupes de sujets comparés présentent chacun les phénotypes relatifs à la maladie? Différents phénotypes, certes, mais qui sont potentiellement influencés par des mêmes déterminants génétiques qu'il reste à découvrir.

En ce qui concerne le biais de classification ou de sélection dans les autres études d'association de la présente thèse, autant les cas-témoins que les trios familiaux, les sujets ont été évalués selon des critères stricts de définition des phénotypes, évitant au maximum ces deux biais potentiels. Pour ce qui est de la stratification des populations, l'utilisation des échantillons de trios familiaux évite ce biais (Spielman et Ewens, 1996). Le fait que des non-réplifications ou des réplifications partielles aient été observées dans ces autres échantillons doit donc faire appel à d'autres facteurs.

#### **9.4.2 Erreur de génotypage**

Les erreurs de génotypage peuvent modifier de façon notable les fréquences alléliques observées et les puissances statistiques de l'échantillon de sujets, ce qui peut avoir un impact sur le résultat des analyses statistiques et les conclusions qui en seront tirées (Rousseau et Laflamme, 2003). Ces erreurs peuvent résulter de la contamination des ADN dans les plaques ou les tubes de stockage, d'erreurs d'étiquetage, d'erreurs d'interprétation des résultats ou de la composante environnementale des laboratoires, surtout pour les nouvelles technologies à haut débit de plus en plus sensibles (Xu et al., 2002). Pour s'assurer de la qualité du génotypage, trois types de vérification sont effectués.

Une première vérification concerne le taux de réussite du génotypage qui dépend de la technologie employée. En fait, toutes ces méthodologies s'accompagnent d'un taux de faux génotypes ainsi que d'un taux d'efficacité (% des essais qui ne produisent pas de résultats - souvent entre 5 et 10 %). Cela peut affecter la puissance d'une étude (prochaine sous-section) tout en augmentant exponentiellement l'ampleur du problème à mesure que les analyses se complexifient et que plusieurs loci sont analysés simultanément. Une méthode de génotypage qui donne un résultat pour 95 % des échantillons (ce qui peut paraître satisfaisant à première vue) signifie que 5 % des échantillons n'ont pas de génotypes à chaque fois qu'un marqueur est analysé. Comme ce ne sont pas nécessairement toujours les mêmes ADN qui ne produisent pas de résultats, la proportion des ADN n'ayant pas de génotype pour tous les marqueurs augmente donc avec le nombre de marqueurs étudiés (Rousseau et Laflamme, 2003). Malgré cette limite, le taux de succès du génotypage avec les technologies employées dans les études d'association de la présente thèse doit être supérieur à 95 %, ce qui est

un seuil généralement fixé pour limiter les biais provoqués par les erreurs de type I. Toutes les études d'association respectent ce seuil.

Une deuxième vérification, spécifique aux échantillons de trios familiaux, est le décompte des erreurs mendéliennes. Pour les études de la présente thèse faisant appel à un échantillon de trios familiaux, les génotypes des familles présentant ces erreurs n'ont pas été utilisés pour les analyses et les SNPs présentant un total de plus de deux erreurs mendéliennes n'ont pas été considérées dans les analyses, respectant ce critère.

Finalement, une dernière vérification consiste à vérifier l'EHW (Rousseau et Laflamme, 2003). Une déviation de cet équilibre populationnel est signe que des erreurs de génotypage sont survenues (donc la fréquence d'un allèle ou d'un génotype est surreprésentée par rapport à celle attendue) ou qu'un des génotypes est favorisé en fonction de la maladie (Xu et al., 2002). Dans toutes les études de la présente thèse, une déviation de l'EHW a été observée dans les échantillons de trios familiaux utilisés pour l'association positive du gène *PLAU* (Begin et al., 2007). Étant donné que les erreurs de génotypage sont peu probables vue la haute qualité des génotypes obtenus, l'observation d'une déviation récurrente de l'EHW pour les SNPs associés dans les trois échantillons étudiés suggère plutôt la présence d'une association entre le locus étudié et l'asthme, ce qui engendrerait la sélection préférentielle de certains génotypes. Cette suggestion est en accord avec les observations de Xu et collaborateurs sur le fait qu'une déviation de l'EHW peut être représentative d'une association génétique (Xu et al., 2002). Pour les échantillons cas-témoins de la présente thèse, aucune déviation de l'EHW a été mise en évidence, ce qui confirme la qualité du génotypage.

Les critères de la qualité du génotypage ont été respectés pour toutes les études d'association présentées et ce, peu importe la technologie utilisée. Il serait donc très surprenant que ce facteur soit en cause pour expliquer des résultats faux-positifs et par conséquent, la variabilité observée parmi les études d'association et leur réplique.

#### **9.4.3 Puissance statistique des échantillons**

Une des étapes principales dans la conception et le recrutement des sujets pour une étude d'association est de déterminer la taille de l'échantillon qui permet une puissance

statistique suffisante pour détecter l'association génétique (Purcell et al., 2003). Le calcul de la puissance statistique et de la taille de l'échantillon nécessaire est effectué par simulation à l'aide de différentes méthodes (Cordell et Clayton, 2005). Ces méthodes tiennent compte de plusieurs paramètres, définis par les chercheurs et les biostatisticiens, comme la fréquence allélique, la pénétrance et la prévalence de la maladie dans la population étudiée ainsi que les probabilités d'erreurs de type I (faux-positifs :  $\alpha$ ) et de type II (faux-négatifs :  $\beta$ ) (Purcell et al., 2003). Une puissance statistique ( $1 - \beta$ ) de 80 % et un  $\alpha$  de 5 % sont généralement employés dans la détermination de la taille de l'échantillon. Ces paramètres assument un certain taux d'erreurs, représentant un compromis afin d'obtenir des tailles d'échantillons réalistes qui tiennent compte de la difficulté à recruter de très grands échantillons (plusieurs milliers) de sujets humains. Cette acceptation générale d'une puissance réduite et de l'exposition conséquente à des résultats faux-positifs est contrebalancée par la nécessité de répliquer toutes associations positives dans des études indépendantes. C'est, en effet, le seul moyen de contrer les faux-positifs possibles. Ainsi, selon les paramètres définis, l'estimation du nombre de trios ou de sujets cas et témoins minimal à inclure dans une étude est calculée afin de permettre la détection des associations entre le trait complexe et les marqueurs génétiques étudiés. Pour les présentes études d'association, Mathieu Lemire (biostatisticien à l'*Ontario Institute for Cancer Research*) et Denise Daley (épidémiologiste à l'Université de la Colombie-Britannique), coauteurs dans certains articles de la présente thèse, ont effectué ces calculs de puissance pour tous les échantillons (sauf celui de CAMP – réalisé par son investigateur principal, le Dr Raby) pour s'assurer de respecter ce critère. Par exemple, pour l'échantillon de trios familiaux du SLSJ, les calculs ont été effectués à l'aide du modèle de simulation développé par Monte Carlo (Borenstein et al., 1987) et démontrent que l'utilisation d'au moins 200 trios offre une puissance satisfaisante en utilisant une prévalence d'asthme de 15 % (observée dans la population générale) et une fréquence allélique de 2,5 % (Lemire, 2004).

Tous les échantillons utilisés pour les présentes études d'association ont la taille d'échantillons de départ nécessaire pour atteindre une puissance statistique suffisante pour détecter les associations. Donc, les non-répliqués et répliqués partiels observés dans les présents travaux ont moins de probabilités d'être attribuables à un



manque de puissance des échantillons employés. Ce qui n'est pas le cas pour l'étude de Chelbi et collaborateurs lors de leur réplique de l'association du *CX3CR1* (Chelbi et al., 2008). En effet, cette étude a été effectuée dans un échantillon cas-témoins, comportant seulement 121 cas et 226 témoins, et dans un échantillon de 99 trios familiaux. Que ce soit l'un ou l'autre des échantillons, la puissance statistique, insuffisante pour détecter l'association, aurait pu causer un résultat faux-négatif, c'est-à-dire l'observation d'une association négative alors qu'elle est positive. Un tel manque de puissance pourrait expliquer pourquoi cette étude (Chelbi et al., 2008) n'a pas répliqué l'association observée pour *CX3CR1* dans l'échantillon du SLSJ (Tremblay et al., 2006).

La détection d'association entre génotype et phénotype pour les traits complexes nécessite d'autant plus de puissance que le génotype étudié est rare, ou que son effet est modeste (Rousseau et Laflamme, 2003). En fait, les petits échantillons sont un problème fréquent en épidémiologie génétique car la puissance statistique qui en résulte est insuffisante pour détecter les contributions mineures de un ou plusieurs allèles (Chanock et al., 2007), ce qui se traduit par l'apparition de plusieurs non-répliques dans la littérature (Hirschhorn et al., 2002).

#### **9.4.4 Corrections statistiques**

Les associations confondantes peuvent également apparaître suite aux fluctuations statistiques et à l'utilisation étendue d'une valeur-p seuil de 0,05 inappropriée comme critère pour déclarer le succès d'un test. En effet, en l'absence d'une vraie association génétique ou de biais, 5 % des études d'association auront une valeur-p inférieure à 0,05, seulement par l'effet du hasard (Newton-Cheh et Hirschhorn, 2005). Dans la littérature, rares sont les études d'association des traits complexes, y compris l'asthme, qui ont corrigé les valeurs-p selon le nombre de tests statistiques effectués, faisant en sorte que le taux d'associations faux-positifs est plus élevé que le niveau nominal (Hoffjan et al., 2003). D'ailleurs, des méta-analyses des associations documentées suggèrent que la plupart des associations publiées sont incorrectes et que ces faux-positifs sont probablement responsables des échecs de réplique des associations entre des variants génétiques et les traits complexes (Ioannidis et al., 2001).

Donc pour minimiser ces faux-positifs liés à la statistique du test d'association, plusieurs types de corrections existent, comme la correction Bonferroni, les tests de permutation ou l'approche bayésienne (Newton-Cheh et Hirschhorn, 2005). Aucun standard n'est actuellement défini pour favoriser l'une ou l'autre des corrections statistiques. Pour ce qui est de la correction Bonferroni, employée dans toutes les études d'association de la présente thèse, elle considère la probabilité d'observer une valeur-p en tenant compte des multiples hypothèses testées (Newton-Cheh et Hirschhorn, 2005). Ce type de correction statistique est approprié lorsqu'un petit nombre de tests est effectué (Rao, 2008). Toutefois, dès qu'un nombre plus élevé de tests est réalisé (par exemple dans l'étude basée sur les puces à ADN où 77 TagSNPs et quatre phénotypes ont été étudiés (Tremblay et al., 2008a)), cette correction statistique est considérée comme très punitive, car elle ne tient pas compte des corrélations qui existent entre les variations génétiques (dû au déséquilibre de liaison) et entre les différents phénotypes intermédiaires de l'asthme (par exemple l'atopie et l'hyperréactivité bronchique) (Newton-Cheh et Hirschhorn, 2005). Dans ce contexte, la correction Bonferroni réduit au maximum les faux-positifs, mais en contrepartie, elle expose les analyses à l'apparition de faux-négatifs (Rao, 2008). Comme les tests statistiques ne sont pas totalement indépendants, des méthodes existent pour calculer cette corrélation qui existent entre les variables étudiées (entre les SNPs et entre les phénotypes) (Nyholt, 2004; Li et Ji, 2005), ce qui permet d'estimer un nombre de tests indépendants total inférieur à celui considéré par une correction Bonferroni classique. Par conséquent, la correction statistique est moins punitive, ce qui permet de trouver un juste équilibre entre les taux de faux-positifs et de faux-négatifs.

Même en adaptant la correction Bonferroni selon la corrélation qui existe entre les SNPs et entre les phénotypes étudiés, la probabilité que les associations génétiques trouvées soient de faux-négatifs est non nulle. Par exemple, dans l'étude G2D (Tremblay et al., 2008b), le nombre de tests total a été considéré comme facteur multiplicatif de la valeur-p. Cette correction très punitive assure l'association avec le gène *PTPRE*, mais en contrepartie, laisse un doute sur d'autres gènes candidats, comme *GPR123*, qui auraient peut-être résisté aux corrections statistiques s'il avait été étudié seul. Bref, la correction statistique permet un certain contrôle des associations faux-positifs. Toutefois, il ne faut pas oublier les hypothèses biologiques de départ et la possibilité d'identifier de nouvelles voies biologiques ainsi que de nouvelles cibles

thérapeutiques potentielles comme c'est le cas pour *GPR123*, un récepteur couplé à une protéine G dont la fonction n'est pas encore caractérisée. Ce gène possède peut-être un effet très minime sur le développement de la maladie et par conséquent, une taille d'échantillons plus grande serait peut-être nécessaire pour découvrir une association positive avec un tel candidat.

#### **9.4.5 Biais de publication**

Les études d'association génétiques sont soumises au biais de publication, c'est-à-dire à la tendance générale de favoriser la publication des associations positives (Hirschhorn et al., 2002; Bosse et Hudson, 2007; Lin et al., 2007). En fait, de nombreuses études d'association négatives ne sont jamais rapportées dans la littérature (Hirschhorn et al., 2002; Hoffjan et al., 2003) et ce, probablement parce que les chercheurs sont moins enclins à rapporter des résultats négatifs (Lawrence, 2003; Lin et al., 2007), sentiment amplifié par la connaissance que les journaux et éditeurs ne sont pas réceptifs à de tels résultats (Lawrence, 2003). Les résultats négatifs sont donc difficiles à publier en raison des haut standards actuels que les journaux préconisent, mais ils sont cruciaux pour distinguer une vraie association positive d'une fausse (Chanock et al., 2007; Lin et al., 2007) et pour fournir une information essentielle à un chercheur qui décidera si oui ou non, il poursuivra les études de réplification ainsi que les études fonctionnelles avec un gène candidat donné (Chanock et al., 2007).

Au lieu de proposer que les journaux publient toutes les études d'association négatives qui sont effectuées dans tous les traits complexes, ce qui serait impensable, certains auteurs se penchent sur le biais de publication et proposent des solutions pour y remédier (Colhoun et al., 2003; Chanock et al., 2007). Pour Chanock et collaborateurs (2007), la solution passe par la voie classique de publication. En fait, ces auteurs suggèrent que seulement les études d'association menées selon les plus haut critères de qualité et qui ont une « négativité utile », c'est-à-dire qui permet la mise en lumière d'une association faux-positif publiée, devraient être publiées sous forme d'article court et succinct (Chanock et al., 2007). Par contre, pour Colhoun et collaborateurs, la solution passe par une voie alternative de publication basée sur les bases de données électroniques (Colhoun et al., 2003). Une telle alternative, pour réussir et être employée à grande échelle, devra donner la reconnaissance, par leurs pairs, aux chercheurs qui

publieraient selon cette voie (Colhoun et al., 2003). Malgré l'utilité potentielle de telles bases de données pour l'avancement de la recherche, la communauté de chercheurs en génétique des traits complexes n'est pas encore mobilisée en ce sens. Comme tout changement de mœurs, la motivation et le temps sont nécessaires pour instaurer une nouvelle façon de faire.

Pour ce qui est des associations génétiques de la présente thèse, le suivi des études de réplication qui seront publiées dans la littérature classique ou électronique permettra, en plus de confirmer les associations positives ou négatives observées, d'élucider si certaines associations positives (comme *CX3CR1*, *ALOX15* ou *PTPRE*) sont en fait de faux-positifs et inversement, si certaines associations négatives (comme *GPR123*, *IL2RB* ou *ADAM8*) sont en fait, de faux-négatifs. De telles informations sont essentielles pour la poursuite des travaux génétiques et fonctionnels sur ces candidats.

#### **9.4.6 Structure génétique des populations**

Les non-répliqués ou les répliqués partiels, si elles ne sont pas le résultat d'une association faux-positif, peuvent être expliquées par la structure génétique des populations qui ont été utilisées pour construire les échantillons à l'étude. La structure des populations fait appel à la notion de déséquilibre de liaison (Chapitre 3). Il existe un patron de déséquilibre de liaison spécifique à chaque population responsable des différentes fréquences alléliques (hétérogénéité allélique) et haplotypiques observées entre les diverses populations (Halapi et Hakonarson, 2004; Ober et Hoffjan, 2006). Si un SNP associé dans une première population n'est pas répliqué dans une seconde population, le premier n'est pas nécessairement un faux-positif, il peut être expliqué par une différence dans la composition des blocs de déséquilibre de liaison qui serait responsable de la présence ou de l'absence du SNP ainsi que de sa fréquence allélique variable.

Pour comparer la structure génétique des échantillons utilisés dans les études d'association de la présente thèse, deux méthodes ont été employées. La première consiste à mesurer le déséquilibre de liaison et à le comparer entre les différentes populations étudiées lorsque possible. La deuxième consiste, à partir des patrons de déséquilibre de liaison mesurés, à définir des haplotypes pour les analyses (Figure 6).

L'utilisation des haplotypes dans les analyses offre l'avantage de couvrir l'ensemble de la région d'un déséquilibre de liaison. L'utilisation des haplotypes est donc comparable à l'utilisation des TagSNPs, en ce sens où une association positive entre un haplotype ou un TagSNP et le phénotype de la maladie est indirecte, c'est-à-dire représentative d'une région de déséquilibre de liaison, et non d'une variation causale qui elle, sera identifiée par le séquençage subséquent de la région génique associée. L'analyse des haplotypes a donc été effectuée dans les études d'association où les variants étudiés n'étaient pas des TagSNPs.

Les variations des patrons de déséquilibre de liaison d'une population à l'autre sont mises en évidence par un autre phénomène génétique, l'effet fondateur. Parmi les populations étudiées, seule celle du SLSJ est reconnue comme une population à effet fondateur (Lander et Schork, 1994; Heyer et Tremblay, 1995; Labuda et al., 1996; Richter et al., 1999; Engert et al., 2000; Scriver, 2001). Ce type de population se caractérise par le partage de longs blocs haplotypiques ancestraux contenant un nombre plus restreint d'allèles associés à une maladie donnée (Malerba et Pignatti, 2005). Comparativement à une population plus hétérogène, comme celle des échantillons SAGE, CAPPS ou CAMP, où ces blocs sont plus courts suite aux événements de recombinaison, il est possible qu'un variant génétique donné s'y retrouve avec une plus grande fréquence allélique (ou plus faible) et par conséquent, permettant l'association ou non. Par exemple, le gène *ALOX15* est associé dans deux populations hétérogènes (CAPPS et SAGE) mais ne l'est pas dans la population du SLSJ (Tremblay et al., 2008a). L'effet fondateur pourrait être responsable d'une structure génétique unique dans cette population, faisant en sorte que certaines associations y sont uniques ou au contraire, totalement absentes. La structure génétique unique peut également être mise en lumière par l'étude de liaison qui a permis d'identifier deux régions chromosomiques, 6q26 et 10q26.3 (Tremblay et al., 2008b), qui n'ont jamais été documentées comme liées à l'asthme ou à ses phénotypes dans les études de liaison réalisées dans les autres populations (Hoffjan et Ober, 2002; Blumenthal, 2005). La structure génétique particulière du SLSJ rend d'autant plus essentielle la réplification dans des échantillons d'origine différente avant de généraliser les associations génétiques trouvées pour l'asthme à l'ensemble des populations.

Mais ce n'est pas parce qu'une association trouvée dans la population du SLSJ n'est pas répliquée dans une autre population plus hétérogène que l'association est un faux-positif. Par exemple, pour le gène *PTPRE*, l'association positive dans l'échantillon du SLSJ peut être vraie et la non-association dans l'échantillon CAMP peut être vraie aussi. Dans ce cas, il est possible que l'association génétique soit unique à la population du SLSJ due à sa structure génétique particulière.

Dans certaines circonstances, il pourrait donc, théoriquement, être impossible d'effectuer le suivi des études d'association parce que la population étudiée est unique ou la disponibilité des sujets est trop limitée pour permettre une autoréplication des associations positives (Chanock et al., 2007). Dans ces cas particuliers où la réplication n'est pas une option, l'interprétation des associations génétiques trouvées pourraient quand même être valable si des évidences biologiques sont fournies par les études fonctionnelles (Chanock et al., 2007).

#### **9.4.7 Interactions gène-gène et gène-environnement**

Des études d'association positives peuvent également ne pas être répliquées parce que le modèle de susceptibilité génétique de la maladie est complexe et implique plusieurs déterminants génétiques à effet modeste sur l'expression du phénotype (Hirschhorn et al., 2002) dont l'amplitude est influencée par d'autres gènes (interactions gène-gène) (Howard et al., 2002; Vercelli, 2008) et par les facteurs environnementaux (interactions gène-environnement) (Ober et Thompson, 2005). En fait, certains variants génétiques peuvent conférer une susceptibilité seulement en combinaison avec d'autres gènes ou dans certains environnements particuliers. Parce que la structure génétique et les expositions environnementales diffèrent entre les populations, il n'est pas surprenant que les associations de SNPs, des TagSNPs ou des haplotypes puissent être variables entre les études (Hoffjan et al., 2003; Ober et Hoffjan, 2006). En fait, selon cette observation, il est logique de penser qu'une association génétique trouvée dans une population à structure génétique particulière dans un environnement précis, soit seulement observée dans d'autres populations partageant les mêmes caractéristiques (Hirschhorn et al., 2002). Or, cette logique idéaliste n'est pas la règle en épidémiologie génétique de l'asthme. C'est pourquoi les études d'interaction entre les facteurs génétiques et environnementaux dans le développement de l'asthme prennent de plus en

plus d'ampleur depuis les dernières années (Wong et al., 2004), afin de tenir compte de cet important facteur de variabilité des associations génétiques entre les diverses études.

Une des limites à la réalisation, à grande échelle, des études d'interactions est le fait qu'elles sont confrontées à la puissance statistique des échantillons étudiés (Luan et al., 2001; Vercelli, 2008), mobilisant les efforts à l'estimation des tailles d'échantillons nécessaires pour réaliser de telles études (Luan et al., 2001; Gauderman, 2002; Wang et Zhao, 2003). Les gènes et l'environnement sont nécessaires pour le développement de l'asthme mais pris isolément, sont insuffisants pour induire l'expression de la maladie (Barnes, 1999; Vercelli, 2008). En fait, une variation dans un seul gène ne peut résulter à l'apparition des phénotypes de l'asthme ou de l'atopie. C'est l'interaction des gènes entre eux et avec l'environnement qui déterminera l'expression de la maladie (Hakonarson et Halapi, 2002). L'importance grandissante de ces interactions est reconnue (Barnes, 1999), mais les modèles mathématiques considérant l'effet cumulatif des facteurs de risque génétiques et environnementaux sont actuellement en exploration.

Étudier les interactions entre les gènes représente donc un défi, c'est pourquoi tout un secteur de la recherche en épidémiologie génétique y est consacré. Tout comme les études d'association, ces études d'interaction doivent sélectionner des gènes candidats dont le rôle physiologique est connu afin de déterminer quelle est leur relation avec le développement de la maladie. Dans l'asthme, rares sont ces études d'interaction. Un exemple d'étude d'interaction gène-gène, réalisée par Kasbech et collaborateurs, concerne des variations dans les gènes de la signalisation immunitaire et de la différenciation des cellules Th2. Dans cette étude réalisée dans un échantillon cas-témoins originaire d'Allemagne, il a été montré que lorsque des SNPs dans les gènes *IL4*, *IL13*, *IL4RA* et *STAT6* (Tableau I, page 50) sont combinés, le risque d'un haut niveau d'IgE sériques augmente de 10,8 fois et le risque de développer de l'asthme augmente de 16,8 fois comparé à l'effet maximal des SNPs pris un à un (Kasbech et al., 2006). Pour ce qui est des interactions gène-environnement, un exemple classique est celui de l'antigène CD14. Dans l'étude réalisée par Simpson et collaborateurs, il a été démontré que les individus porteurs d'un génotype donné d'un des SNP du gène *CD14* ont moins de risque de développer une sensibilisation allergique lorsque leur exposition aux endotoxines est élevée (Simpson et al., 2006). Par contre, cette même exposition

élevée aux endotoxines augmente le risque d'apparition des symptômes de l'asthme chez les enfants non-atopiques et porteurs du génotype (Simpson et al., 2006). Ces exemples d'études d'interaction montrent l'utilité de telles connaissances sur la compréhension de l'asthme en plus de témoigner de l'importance et du besoin pressant de développer de nouvelles approches, tant génétiques que fonctionnelles, pour les réaliser (Hoffjan et al., 2003; Vercelli, 2008).

#### **9.4.8 Hétérogénéité phénotypique**

Un dernier facteur qui peut expliquer la non-réplication d'une association génétique positive concerne l'hétérogénéité phénotypique de l'asthme (Hoffjan et al., 2003; Malerba et Pignatti, 2005). En fait, cette très grande hétérogénéité est doublée d'une très grande complexité (Bosse et Hudson, 2007; Guerra et Martinez, 2008) et d'une histoire naturelle reposant sur des mécanismes pathologiques variables (Reed, 2006; Koh et Irving, 2007). Ainsi, l'asthme peut se développer à n'importe quel âge, être déclenché par plusieurs stimuli (spécifiques ou non), se présenter sous plusieurs formes (professionnel, allergique, etc), se manifester selon plusieurs degrés de sévérité (léger, modéré ou sévère) et être associé avec plusieurs phénotypes intermédiaires comme l'atopie, les niveaux d'IgE, l'hyperréactivité bronchique ou l'éosinophilie, lesquels peuvent tous être assortis selon différentes combinaisons (Guerra et Martinez, 2008). La caractérisation phénotypique est donc une étape méthodologique essentielle afin d'établir le bon phénotype à comparer, car l'hétérogénéité phénotypique peut grandement contribuer à la variation des associations au sein de diverses études de réplication, particulièrement si les différents phénotypes intermédiaires impliquent une voie génétique distincte (Chanock et al., 2007; Guerra et Martinez, 2008). D'autant plus que différents paramètres sont employés pour définir les phénotypes de l'asthme, comme pour les provocations bronchiques, les valeurs normales seuil pour les tests de fonction respiratoire ou du niveau d'IgE sériques, pour ne nommer que ces exemples, ce qui augmente les sources de biais possibles dans l'établissement du phénotype et donc, de l'interprétation des résultats (Halapi et Hakonarson, 2004).

Malgré l'élaboration de définitions claires et la réalisation d'un phénotypage rigoureux, les études d'association de la présente thèse n'échappent pas à une variabilité phénotypique potentielle, découlant de la construction même des échantillons étudiés.



Pour la non-réplication du *PTPRE*, une des raisons évoquée pour expliquer les non-réplifications est justement l'hétérogénéité du phénotype asthme qui dans un cas, est plus un asthme à l'enfance (CAMP = moyenne d'âge de 8 ans), tandis que pour l'échantillon du SLSJ, il pourrait s'agir d'un asthme persistant (moyenne d'âge = 18 ans) (Tremblay et al., 2008b). Des mécanismes génétiques différents, incluant les interactions gène-gène et gène-environnement, pourraient être à l'origine de la persistance de l'asthme chez les sujets du SLSJ comparativement aux sujets de la CAMP pour lequel il n'est pas possible de savoir quel sujet aurait ou non eu un asthme persistant jusqu'à l'âge adulte. Deux autres études mettent en lumière l'hétérogénéité phénotypique pour expliquer les différences dans les réplifications des associations (Begin et al., 2007; Tremblay et al., 2008a). Dans ces deux études, le fait que le recrutement des sujets diffère entre les échantillons peut résulter en différents stades de la maladie et par conséquent, différents types d'asthme. En effet, dans les échantillons CAPPS et SAGE, il s'agit d'un asthme chez les enfants, pour l'échantillon du SLSJ, il s'agit plutôt d'un asthme apparu à l'enfance et qui a persisté à l'âge adulte et finalement, pour l'échantillon de Busselton, il s'agit d'un asthme chez l'adulte (Tremblay et al., 2008a). Comme différents déterminants génétiques peuvent être impliqués dans la prédisposition aux différents types d'asthme, il est plausible que cette source d'hétérogénéité phénotypique soit responsable de la variabilité des associations observées dans ces quatre échantillons.

En résumé, les plus grands défis qui seront à affronter dans les prochaines années ne sera pas l'identification des gènes associés à l'asthme, mais plutôt leur réplication dans les études indépendantes (Chanock et al., 2007), la caractérisation des mécanismes moléculaires qui influencent le risque génétique et l'expression du phénotype ainsi que la mise en évidence de toutes les interactions gène-gène et gène-environnement possibles (McCarthy et al., 2008). Ces connaissances fourniront une meilleure compréhension de la physiopathologie de l'asthme, incluant la réponse du patient à la médication actuelle (Hakonarson et Halapi, 2002), et pourront être utilisées dans le développement de nouvelles stratégies thérapeutiques allant du diagnostic, à la prise en charge du patient et même, au développement de nouveaux médicaments plus sécuritaires et plus efficaces pour tous les patients (Halapi et Hakonarson, 2004; Holloway et Koppelman, 2007).

## Chapitre 10

### Perspectives de recherche

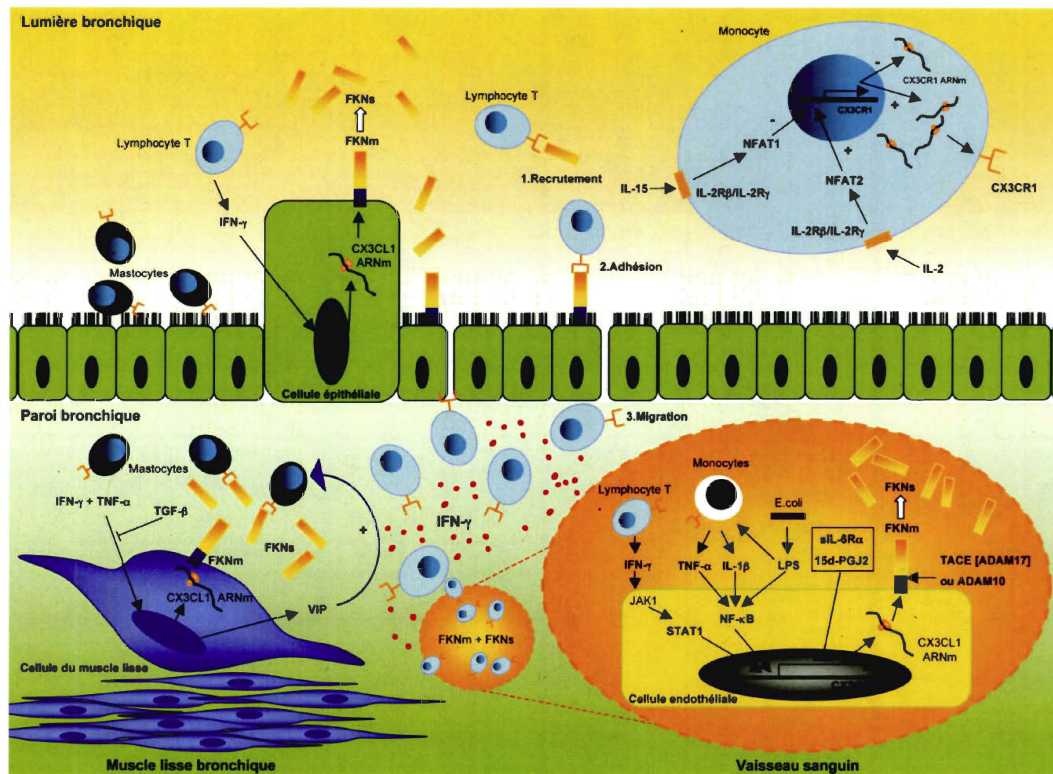
Les tests statistiques réalisés pour identifier une association génétique permettent d'obtenir une valeur-p indicative de la probabilité de découvrir un tel événement, c'est-à-dire une association positive entre un phénotype et un génotype. Cette relation statistique n'est nullement représentative de l'importance biologique de cette association, à savoir l'impact réel des variations génétiques associées dans la physiopathologie de la maladie (Williams et al., 2007). C'est pourquoi les études d'association doivent nécessairement être suivies d'études fonctionnelles afin de vérifier l'implication du gène et de ses produits dans les mécanismes physiopathologiques (Zhang et al., 2008). Pour ce faire, une revue de littérature exhaustive des études génétiques et fonctionnelles publiées pour chacun des gènes associés doit être effectuée. Cet exercice permet de rassembler les connaissances biologiques du gène et de ses produits, en plus de les situer dans le contexte de la pathologie de l'asthme. Une telle intégration des connaissances est facilitée par l'élaboration de schémas permettant une meilleure visualisation du rôle potentiel du gène dans la physiopathologie de la maladie ainsi que l'élaboration de nouvelles hypothèses de recherche. Une fois cette étape accomplie, le chercheur possède les bases nécessaires pour établir sa conception de l'étude fonctionnelle. Dans le cadre du dernier objectif de cette thèse, cet exercice d'intégration a été effectué pour deux des gènes associés, *CX3CR1* et *ALOX15*, afin de démontrer le principe et d'illustrer les nouvelles questions et hypothèses de recherche mises en lumière suivant une étude d'association.

#### 10.1 Exemples de modèle d'intégration biologique

Dans le cadre de la présente thèse, les deux modèles présentés ont été faits avec les gènes *CX3CR1* et *ALOX15*, mais sont applicables aux autres gènes associés. Le choix de ces deux gènes repose sur des évidences biologiques dans les mécanismes inflammatoires qui s'inscrivent dans le contexte de la physiopathologie de l'asthme.

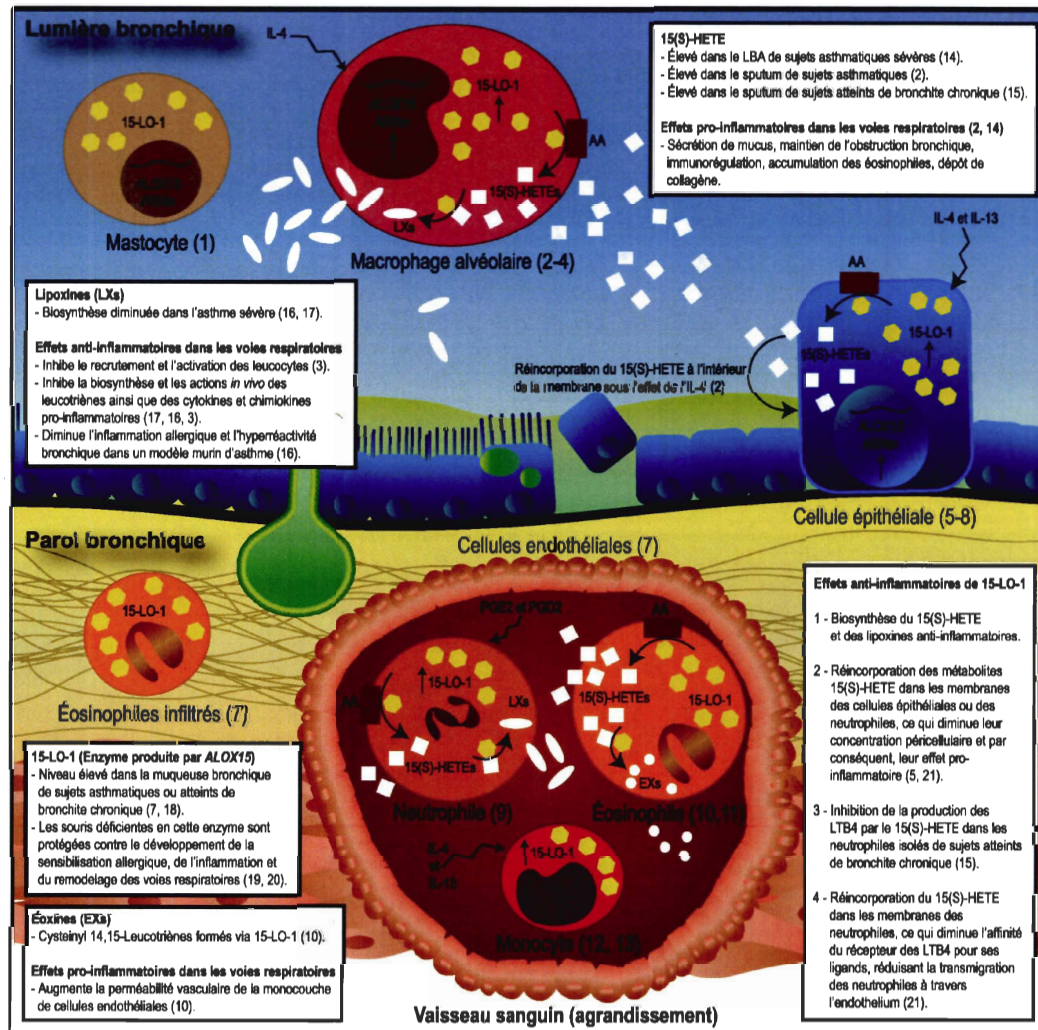
Pour ce qui est du gène *CX3CR1*, la revue de littérature de la voie biologique ainsi que la construction du schéma de son implication potentielle dans les voies respiratoires ont

été effectuées en 2006 lors de la publication de l'article. La Figure 9 résume les principales interrelations du récepteur CX3CR1 avec divers médiateurs et types cellulaires.



**Figure 9 : Représentation schématique et hypothétique de la voie biologique de CX3CR1 et de son ligand dans les voies respiratoires chez l'humain.** La fractalkine (FKN ou CX3CL1) est exprimée dans différents types cellulaires dans les voies respiratoires incluant les cellules endothéliales, épithéliales et musculaires bronchiques. Dans les cellules endothéliales, l'expression de CX3CL1 est induite par plusieurs médiateurs (IFN- $\gamma$ , TNF- $\alpha$ , IL-1 $\beta$  et LPS) via l'activation des facteurs de transcription STAT1 et NF- $\kappa$ B, mais cette expression est inhibée par d'autres médiateurs comme le IL-6R $\alpha$  soluble (sIL-6R $\alpha$ ) ou le métabolite de la prostaglandine PGD2, 15d-PGJ2. L'ARN<sub>m</sub> de la fractalkine est traduit en sa forme membranaire (FKNm), laquelle sera clivée par les enzymes de conversion ADAM17 ou ADAM10 pour produire sa forme soluble (FKNs). La FKNs se lie à son récepteur, CX3CR1, exprimé par différents types de leucocytes (cellules tueuses naturelles, monocytes, lymphocytes T, mastocytes), ce qui initie le recrutement, l'adhésion et la migration de ces cellules dans le foyer d'inflammation tissulaire. Au niveau des cellules épithéliales, CX3CL1 agit également comme chimioattractant pour les leucocytes exprimant le CX3CR1, mais son expression est induite par l'IFN- $\gamma$ . Au niveau des cellules du muscle lisse, elles expriment CX3CL1 lorsqu'elles sont stimulées par l'IFN- $\gamma$  et le TNF- $\alpha$  simultanément (effet atténué par le TGF- $\beta$ ), mais son effet chimioattractant est reconnu seulement pour les mastocytes exprimant le CX3CR1 et activés par le VIP. Ce médiateur vasoactif est également produit par les cellules du muscle lisse et son effet d'activation des mastocytes est nécessaire pour la chimioattraction induite par la FKN. Finalement, il a été montré que l'expression de CX3CR1 dans les leucocytes mononucléaires est sur-régulée par la voie de signalisation intracellulaire de l'IL-2 qui implique l'IL-2R $\beta$ , l'IL-2R $\gamma$  et le NFAT2. En contrepartie, cette expression est inhibée par la voie de signalisation de l'IL-15 qui implique également l'IL-2R $\beta$  et l'IL-2R $\gamma$  ainsi que le NFAT1. La description des molécules et de leur gène ainsi que les références utilisées pour construire cette figure sont disponibles dans le Tableau VII (page 335).

Le même principe employé pour *CX3CR1* a été répété pour *ALOX15*. La revue de littérature et le schéma récapitulatif ont été construits en septembre 2008. La Figure 10 résume l'essentiel de la littérature concernant l'implication de la voie biologique de la 15-LO dans la physiopathologie de l'asthme.



**Figure 10 : Représentation schématique des voies respiratoires asthmatiques montrant l'expression du gène *ALOX15* (ARN<sub>m</sub>) et de son enzyme, la 15-lipoxygénase (15-LO-1).** Pour chacune des cellules structurales ou inflammatoires impliquées, la 15-LO-1 transforme l'acide arachidonique (AA) membranaire en divers métabolites pro- et/ou anti-inflammatoires, tels l'acide 15-hydroxyeicosatétraénoïque (15(S)-HETE), la lipoxine (LX) et l'éoxine (EX). Références utilisées : <sup>1</sup>(Gulliksson et al., 2007), <sup>2</sup>(Profita et al., 2000a), <sup>3</sup>(Levy et al., 2002), <sup>4</sup>(Levy et al., 1993), <sup>5</sup>(Profita et al., 1999), <sup>6</sup>(Brinckmann et al., 1996), <sup>7</sup>(Bradding et al., 1995), <sup>8</sup>(Sigal et al., 1992), <sup>9</sup>(Levy et al., 2001), <sup>10</sup>(Feltenmark et al., 2008), <sup>11</sup>(Turk et al., 1982), <sup>12</sup>(Nassar et al., 1994), <sup>13</sup>(Conrad et al., 1992), <sup>14</sup>(Chu et al., 2002), <sup>15</sup>(Profita et al., 2000b), <sup>16</sup>(Levy et al., 2005), <sup>17</sup>(Planaguma et al., 2008), <sup>18</sup>(Shannon et al., 1993), <sup>19</sup>(Andersson et al., 2008), <sup>20</sup>(Hajek et al., 2008) et <sup>21</sup>(Takata et al., 1994).

## 10.2 Nouvelles questions et hypothèses de recherche

Suivant ces deux exemples de schémas récapitulatifs, de nouvelles questions et hypothèses de recherche peuvent être élaborées (Tableau VI). Ces questions sont regroupées dans un diagramme conceptuel (Figure 11) afin de mettre en contexte et d'illustrer les diverses étapes qui suivent l'identification d'un déterminant génétique d'un trait complexe et qui seront abordées dans les prochains paragraphes.

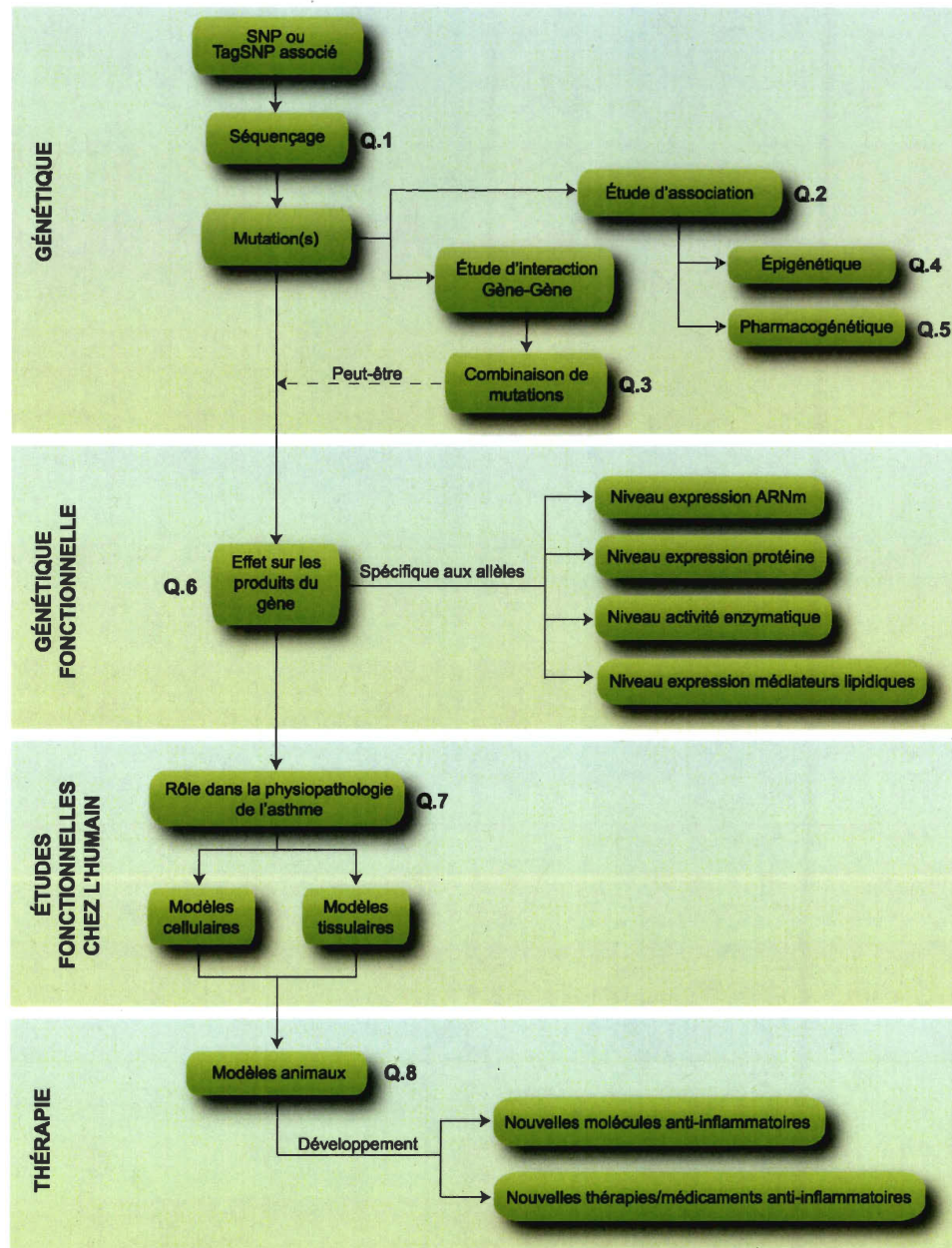
### Tableau VI

Questions de recherche soulevées par les résultats d'une étude d'association.

Questions de recherche	
Q.1	Quelle mutation est responsable de l'association génétique ?
Q.2	Cette mutation est-elle un facteur de risque génétique pour le trait complexe étudié ?
Q.3	Est-ce que la mutation interagit avec d'autres gènes ?
Q.4	Y a-t-il une transmission préférentielle des allèles mutés en fonction des parents ?
Q.5	Est-ce que la mutation affecte la réponse à la médication ?
Q.6	Est-ce que la mutation affecte le niveau d'expression des produits du gène ?
Q.7	Quel est l'impact de la mutation sur les composantes physiopathologiques du trait complexe étudié ?
Q.8	L'induction de la mutation dans un modèle animal fera-t-elle apparaître les caractéristiques physiopathologiques et symptomatologiques du trait complexe étudié ?

Les premières questions soulevées sont de nature génétique. Par exemple, pour le gène *ALOX15* (Tremblay et al., 2008a), étant donné que l'association trouvée concerne un TagSNP, une première question sera de savoir quelle est la mutation causale (Tableau VI - Q.1). Pour y répondre, le séquençage du gène est requis. Une fois la ou les mutations causales identifiées, d'autres études d'association seront effectuées afin de préciser la nature du risque génétique pour les phénotypes étudiés (Tableau VI - Q.2). Ensuite, d'autres types d'études comme les études épigénétiques ou pharmacogénétiques pourront être entreprises. Dans les premières études, le chercheur sera intéressé à savoir si les mutations identifiées sont transmises préférentiellement par l'un ou l'autre des deux parents (phénomène appelé empreinte parentale) (Tableau VI - Q.4) ou dans des environnements particuliers (interaction gène-environnement). Dans les secondes études, le chercheur veut savoir si la ou les mutations identifiées sont responsables d'une réponse variable à la médication chez les sujets atteints (Tableau VI - Q.5). Finalement, le portrait génétique ne sera pas final tant que les études d'interaction gène-gène n'auront pas été complétées (Tableau VI - Q.3). Cette étape permettra de savoir si la ou les mutations agissent seules ou en combinaison avec des

variations génétiques d'autres gènes impliqués dans la voie biologique du gène associé. Le Tableau VII présente, à titre d'exemple, la liste de gènes importants dans la voie biologique du gène associé *CX3CR1* (Tremblay et al., 2006) et plausibles pour des analyses d'interaction.



**Figure 11 : Diagramme conceptuel des étapes logiques suivant une étude d'association génétique.** Suivant l'identification d'un gène associé, des études de nature génétique sont nécessairement suivies d'études fonctionnelles afin de déterminer si le gène représente une cible thérapeutique potentielle. Les abréviations Q.1 à Q.8 correspondent aux questions de recherche spécifiques (Tableau VI).

**Tableau VII**  
Gènes impliqués dans la voie biologique du *CX3CR1*.

Gène	Locus	Nom officiel <sup>a</sup>	Fonction <sup>b</sup>	Sélection
<i>CX3CR1</i>	3p21.3	<i>Chemokine (C-X3-C motif) receptor 1</i>	Exprimé par plusieurs cellules. Rôle dans la chimioattraction, l'adhésion et la migration des cellules. Rôle également dans la défense cellulaire et dans la signalisation intracellulaire.	Moins exprimé dans les bronches d'asthmatiques (Laprise et al., 2004). Associé à l'asthme (Tremblay et al., 2006).
<i>IL2</i>	4q26-q27	<i>Interleukin 2 (IL-2)</i>	Cytokine produite par les lymphocytes T en réponse à des stimulations antigéniques ou mitogéniques. Requisite pour la prolifération des lymphocytes T et pour d'autres activités régulatrices dans le système immunitaire.	Stimule le niveau d'expression du <i>CX3CR1</i> via NFAT2 (Barlic et al., 2004).
<i>IL15</i>	4q31	<i>Interleukin 15 (IL-15)</i>	L'IL-15 est une cytokine structurellement et fonctionnellement liée à l'IL-2. Elle est impliquée dans la différenciation des tueuses naturelles et stimule la prolifération des lymphocytes T. La stimulation par l'IL-15 requiert l'interaction des sous-unités beta et gamma du récepteur de l'IL-2.	Réduit l'activité du promoteur et la transcription du gène <i>CX3CR1</i> via NFAT1 (Barlic et al., 2004; Sechler et al., 2004).
<i>IL2RB</i>	22q13.1	<i>Interleukin 2 receptor, beta (IL-2Rβ)</i>	Cette sous-unité est impliquée dans l'endocytose du récepteur et dans la transduction du signal de l'IL-2.	Requis pour les activités de l'IL-2 et de l'IL-15 (Barlic et al., 2004).
<i>IL2RG</i>	Xq13.1	<i>Interleukin 2 receptor, gamma (IL-2Rγ)</i>	Cette sous-unité est une composante importante pour la signalisation de plusieurs récepteurs de cytokines comme ceux de l'IL-2, l'IL-4, l'IL-7, l'IL-9 et l'IL-15.	Requis pour les activités de l'IL-2 et de l'IL-15 (Barlic et al., 2004).
<i>NFATC2 (NFAT1)</i>	20q13.2-q13.3	<i>Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2 (T cell transcription factor NFAT1)</i>	Impliqué dans l'expression de gènes de la réponse immunitaire comme ceux des cytokines IL-2, IL-3, IL-4 ou TNF-α.	Impliqué dans l'activité de l'IL-2 (Barlic et al., 2004).
<i>NFATC1 (NFAT2)</i>	18q23	<i>Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1 (T cell transcription factor NFAT2)</i>	Rôle dans l'induction de l'expression des gènes de cytokines, notamment de l'IL-2 et de l'IL-4. Régule l'activation, la prolifération, la différenciation et la mort programmée des lymphocytes T ainsi que d'autres types cellulaires.	Impliqué dans l'activité de l'IL-2 (Barlic et al., 2004).
<i>CX3CL1</i>	16q13	<i>Chemokine (C-X3-C motif) ligand 1 (Fractalkine, neurotactine)</i>	La forme soluble est un chimioattractant pour les cellules T et les monocytes. La forme membranaire a un rôle dans l'adhésion et la migration des leucocytes dans l'endothélium.	Ligand principal du récepteur <i>CX3CR1</i> .
<i>ADAM17 (TACE)</i>	2p25	<i>ADAM metallopeptidase domain 17 (TNF alpha, converting enzyme - TACE)</i>	Protéines membranaires impliquées dans une variété de processus biologiques nécessitant des interactions cellule-cellule ou cellule-tissu. Fonctionne comme une enzyme de conversion responsable de la protéolyse de plusieurs protéines situées à la surface des cellules.	Enzyme de conversion de la FKN membranaire (FKNm) en FKN soluble (FKNs) (Umehara et al., 2004).
<i>ADAM10</i>	15q22	<i>ADAM metallopeptidase domain 10</i>	Protéine membranaire de type I qui possède une activité protéase et clive plusieurs protéines comme le TNF-α ou cadhérine-E.	Enzyme de conversion de la FKNm en FKns (Umehara et al., 2004)

<i>IL1B</i>	2q14	<i>Interleukin 1, beta (IL-1β)</i>	Cytokine produite par les macrophages activés. Stimule la maturation et la prolifération des lymphocytes B, ainsi que la croissance des fibroblastes. Rôle dans la réponse inflammatoire. Facteur de susceptibilité pour la maladie inflammatoire de l'intestin.	Stimule l'expression de <i>CX3CL1</i> (Brand et al., 2006).
<i>TNF (TNFA)</i>	6p21.3	<i>Tumor necrosis factor (TNF superfamily, member 2)</i>	Cytokine pro-inflammatoire principalement sécrétée par les macrophages et impliquée dans plusieurs processus biologiques tels que la prolifération, la différenciation et l'apoptose cellulaires, le métabolisme des lipides et la coagulation. Impliquée dans plusieurs maladies incluant les maladies auto-immunes, la résistance à l'insuline et le cancer.	Stimule l'expression de <i>CX3CL1</i> (Sukkar et al., 2004; Yoshikawa et al., 2004; Brand et al., 2006).
<i>IFNG</i>	12q14	<i>Interferon, gamma (IFN-γ)</i>	Produit par les lymphocytes activés par des antigènes ou mitogènes spécifiques. Possède une activité antivirale et d'importantes fonctions immunorégulatrices.	Stimule l'expression de <i>CX3CL1</i> (Fujimoto et al., 2001; Sukkar et al., 2004).
<i>TGFB1 (TGFB)</i>	19q13.2-19q13.1	<i>Transforming growth factor, beta 1 (TGF-β)</i>	Peptide multifonctionnel contrôlant la prolifération et la différenciation de plusieurs types cellulaires.	Atténue la production de FKN induite par le TNF-α et l'IFN-γ dans les cellules du muscle lisse bronchique (Sukkar et al., 2004).
<i>IL6R (IL6RA)</i>	1q21	<i>Interleukin 6 receptor alpha (IL-6Rα)</i>	Une faible concentration de la forme soluble de ce récepteur agit comme un agoniste de l'IL-6, cytokine qui régule la croissance et la différenciation cellulaire et qui joue un rôle important dans la réponse immunitaire.	Inhibe l'expression de la FKN dans les cellules endothéliales (Imaizumi et al., 2004).
<i>IL8</i>	4q13-q21	<i>Interleukin 8 (IL-8)</i>	Chimiokine importante dans la réponse inflammatoire sécrétée par plusieurs types de cellules en réponse à des stimuli inflammatoires. Effet chimioattractant pour les neutrophiles, les basophiles et les cellules T. Impliquée dans l'activation des neutrophiles.	Le niveau d'expression de l'IL-8 est en corrélation avec celui de la FKN (Brand et al., 2006). Associé avec l'inflammation des intestins et up-régulé par la forme soluble de l'IL-6Rα (Imaizumi et al., 2004; Brand et al., 2006).
<i>NFKB1</i>	4q24	<i>Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (NF-κB)</i>	Régulateur de la transcription activé par stimuli intra- et extracellulaires comme les cytokines, les radicaux libres, les radiations UV et les produits bactériens ou viraux. Stimule l'expression de gènes impliqués dans une variété de processus biologiques.	Sous-régule l'expression de <i>CX3CL1</i> (Imaizumi et al., 2002; Imaizumi et al., 2004).
<i>JUN (API)</i>	1p32-p31	<i>Jun activation domain binding protein, activator protein (AP-1)</i>	Régule l'expression génique.	Sous-régule l'expression de <i>CX3CL1</i> (Imaizumi et al., 2002; Imaizumi et al., 2004).
<i>STAT1</i>	2q32.2	<i>Signal transducer and activator of transcription 1 (STAT1)</i>	Activé par plusieurs ligands, comme l'IL-6, ce facteur de transcription régule l'expression d'une variété de gènes importants pour la viabilité de la cellule en réponse à divers stimuli et agents pathogènes.	Sous-régule l'expression de <i>CX3CL1</i> (Imaizumi et al., 2002; Imaizumi et al., 2004).



<i>PTGS2</i> ( <i>COX2</i> )	1q25.2- q25.3	<i>Prostaglandin- endoperoxide synthase 2</i> ( <i>Cyclooxygenase</i> - <i>COX2</i> )	Enzyme clé dans la biosynthèse des prostaglandines (médiateurs inflammatoires) à partir de l'acide arachidonique.	La prostaglandine PGD2 (principal métabolite de l'acide arachidonique produit par la COX2) est métabolisée en 15d-PGJ2, lequel inhibe l'expression de la FKN dans les cellules endothéliales par inhibition de la signalisation des facteurs de transcription NF-κB, AP-1 et STAT1 (Imaizumi et al., 2002; Imaizumi et al., 2004).
<i>PPARG</i>	3p25	<i>Peroxisome proliferator- activated receptor gamma</i> (PPAR-γ)	Ce facteur nucléaire forme un hétérodimère avec les récepteurs rétinoïque X (RXRs) pour réguler la transcription d'une variété de gènes. Augment l'expression de l'inhibiteur des activateurs de plasminogène (PAI-1) dans les cellules endothéliales.	Hypothèse à vérifier : Le 15d-PGJ2 pourrait sous-réguler l'expression du <i>CX3CL1</i> , (Imaizumi et al., 2002) via l'activation du PPAR-γ, puisqu'il s'agit de son ligand naturel.
<i>VIP</i>	6q26	<i>Vasoactive intestinal peptide</i>	Protéine sécrétée de la famille des glucagons. Stimule la contraction du myocarde, cause la vasodilatation, augmente la glycolyse, diminue la pression artérielle et relaxe le muscle lisse de la trachée, de l'estomac et de la vessie.	Amplifie les effets de la FKN sur les mastocytes (El-Shazly et al., 2006).

<sup>a</sup> Source : NCBI (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene>). Le symbole de la protéine est indiqué entre parenthèses.

<sup>b</sup> Obtenue des bases de données publiques : NCBI Gene, OMIM et PubMed ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)); Genecards (<http://bioinformatics.weizmann.ac.il/cards/>); et Swiss-Prot (<http://ca.expasy.org/sprot/sprot-top.html>).

Suite à ces études génétiques, les études fonctionnelles peuvent être initiées. Ces études sont basées sur le développement d'essais *in vitro* nécessaires pour étudier la fonction du gène dans les processus biologiques pour lesquels il est impliqué, ce qui permet par la suite de vérifier si les variations génétiques associées modulent cette fonction (Holloway et Koppelman, 2007). Plus spécifiquement, ces études permettent de vérifier les niveaux d'expression du gène et de ses différents produits dans les cellules impliquées en fonction de la présence ou de l'absence du variant génétique associé (Tableau VI - Q.6 et Figure 11). Cette étape permet de caractériser l'effet des allèles spécifiques sur l'expression des produits du gène. Ensuite, cet effet observé pourra être étudié dans les modèles cellulaires ou tissulaires du contexte physiopathologique étudié (Tableau VI - Q.7). Par exemple, pour le gène *CX3CR1*, des études structurales et fonctionnelles approfondies sont actuellement en cours sous la direction de la Dre Catherine Laprise et sous la codirection de la Dre Valérie Legendre-Guillemain afin de caractériser la dynamique intracellulaire du récepteur CX3CR1, incluant son internalisation et sa signalisation. En fait, il a été vu que le récepteur CX3CR1 est impliqué dans des processus physiopathologiques de diverses maladies inflammatoires,

mais aucune étude sur la dynamique de ce récepteur n'a été rapportée à ce jour. En d'autres termes, les mécanismes par lesquels ce récepteur est régulé suivant la stimulation par son ligand ne sont pas encore caractérisés. Les résultats de ces travaux permettront de comprendre la régulation de son expression de surface et comment elle peut être modulée, ce qui apportera de nouvelles connaissances dans les mécanismes physiopathologiques impliqués dans certaines maladies inflammatoires telle l'asthme.

Finalement, les connaissances acquises par les études fonctionnelles pourront mener à l'identification de gènes importants dans les processus inflammatoires de l'asthme et par conséquent, permettre l'identification de cibles thérapeutiques potentielles pour le développement de nouveaux médicaments anti-inflammatoires (Tableau VI - Q.8). Par exemple, la caractérisation de la signalisation intracellulaire du récepteur CX3CR1 pourrait aboutir à l'identification d'une nouvelle cible thérapeutique pour le traitement anti-inflammatoire. En effet, le récepteur CX3CR1 étant couplé à une protéine G, et cette famille de récepteurs étant des cibles dans le développement de médicaments (Hopkins et Groom, 2002), il est possible d'envisager que de telles études fonctionnelles fourniront les connaissances nécessaires pour des études pharmacologiques. Même le gène *ALOX15*, si l'on pense à l'implication de la voie biologique de la 15-LO dans l'inhibition de l'action des leucotriènes LTB<sub>4</sub> (Takata et al., 1994; Profita et al., 2000b), pourrait se révéler comme une cible thérapeutique intéressante pour le développement de nouveaux médicaments anti-inflammatoires.

Bref, un tel modèle conceptuel de perspectives de recherche démontre l'utilité des études génétiques dans l'identification de cibles potentielles pour des études fonctionnelles, tel que pour le gène *CX3CR1* et, potentiellement, pour tous les gènes associés dans le cadre des présents travaux de doctorat. Ainsi seulement, il sera possible de définir l'impact réel des variants génétiques associés dans les mécanismes inflammatoires ou de remodelage impliqués dans la pathogénèse de l'asthme. Ultimement, certains gènes pourraient s'avérer des cibles intéressantes pour le développement de nouveaux médicaments anti-inflammatoires. Bien entendu, les études fonctionnelles, autant dans les modèles cellulaires que animaux, sont nécessaires avant de conclure sur l'implication biologique ou pharmacologique d'un gène associé à une maladie : les études d'association génétique déterminent une relation statistique que seules les évidences biologiques pourront infirmer ou confirmer.

## Chapitre 11

### Conclusion

La recherche des déterminants génétiques de l'asthme évolue très rapidement, bien qu'elle soit confrontée à plusieurs difficultés. Plus de 170 gènes sont maintenant associés à ce trait complexe et la liste continue de s'allonger, tel que démontré par les études d'association présentées dans cette thèse. Plusieurs facteurs ont contribué à cette remarquable évolution de l'épidémiologie génétique, comme les connaissances acquises grâce au projet Génome Humain ainsi que le développement des biotechnologies de plus en plus performantes en termes de débit et de réduction des coûts du génotypage. Il faut également considérer les efforts des chercheurs dans la mise en place de critères de standardisation et le développement d'approches méthodologiques diversifiées. Les travaux de la présente thèse s'inscrivent au cœur de cette problématique et à la lumière des résultats présentés, trois grands constats ressortent.

D'abord, l'approche par gènes candidats et une méthode efficace pour réaliser les études d'association puisque l'évaluation de 22 gènes candidats a permis d'identifier sept associations positives. Parmi celles-ci, la découverte de quatre nouvelles associations génétiques est un avancement considérable des connaissances de la génétique de l'asthme. Ainsi, à l'ère des GWAS, l'approche par gènes candidats, dans une optique de complémentarité, sera encore largement employée.

Ensuite, les travaux ont permis l'avancement des connaissances dans la méthodologie de recherche par l'utilisation de différentes approches méthodologiques classiques et innovatrices dans la sélection de gènes candidats pour la réalisation des études d'association. Les approches classiques ont permis la validation de gènes déjà associés (*CD14*, *ORMDL3*, *FLG*) tandis que les approches innovatrices, basées sur la sélection de gènes différemment exprimés suivant une étude du profil d'expression génique ou priorisés par l'outil bioinformatique G2D, ont permis l'identification de quatre associations génétiques encore non documentées pour l'asthme et ses phénotypes (*CX3CR1*, *ALOX15*, *PLAU* et *PTPRE*). Ces résultats, en plus de montrer l'utilité des ces nouvelles approches méthodologiques dans la recherche des déterminants génétiques de l'asthme, montrent le potentiel d'une utilisation à plus grande échelle, c'est-à-dire avec

davantage de gènes et par plusieurs autres groupes de recherche. Bref, ces approches de sélection des gènes candidats ont été utilisées pour la première fois pour des études d'association pour l'asthme et suivant leur utilisation à grande échelle, leur efficacité sera démontrée. Ainsi, par cette méthodologie de recherche originale, les présents travaux ont contribué de façon significative à l'avancement de la génétique de l'asthme.

Finalement, le dernier constat est que les gènes associés dans les présents travaux ont soulevé l'élaboration de nouvelles questions et hypothèses de recherche. Que ce soit l'un ou l'autre des gènes associés, ils représentent tous un potentiel d'études génétiques et fonctionnelles subséquentes. En fait, de telles études permettront de caractériser leur rôle dans la physiopathologie de l'asthme, ce qui permettra d'augmenter la compréhension de la maladie. Si tel est le cas, ces gènes associés ainsi que leur voie biologique pourraient représenter de nouveaux biomarqueurs de l'asthme, ce qui pourrait mener au développement d'outils de diagnostic ou de prédiction, ou représenter de nouvelles cibles thérapeutiques, ce qui pourrait mener au développement de nouveaux traitements anti-inflammatoires ou anti-allergiques.

Ce projet de doctorat a donc permis d'identifier quatre nouveaux gènes associés à l'asthme et d'élaborer de nouvelles perspectives de recherche. Même si beaucoup de travaux sont encore nécessaires afin de comprendre tous les mécanismes génétiques et physiopathologiques impliqués dans le développement de l'asthme, la découverte des gènes de susceptibilité ou de protection pour l'asthme est un pas en avant, essentiel puisqu'il aidera, ultimement, à l'amélioration du diagnostic, au développement de meilleurs programmes de prévention et d'intervention et finalement, au développement de nouveaux traitements pour les personnes atteintes de l'asthme.

## Chapitre 12

### Références

\*Excluant celles citées dans les articles qui forment les chapitres 5 à 8 inclusivement.

- Allen M, Heinzmann A, Noguchi E, Abecasis G, Broxholme J, Ponting CP, Bhattacharyya S, Tinsley J, Zhang Y, Holt R, Jones EY, Lench N, Carey A, Jones H, Dickens NJ, Dimon C, Nicholls R, Baker C, Xue L, Townsend E, Kabesch M, Weiland SK, Carr D, von Mutius E, Adcock IM, Barnes PJ, Lathrop GM, Edwards M, Moffatt MF and Cookson WO (2003) Positional cloning of a novel gene influencing asthma from chromosome 2q14. *Nat Genet* 35:258-263.
- Anderson GG and Cookson WO (1999) Recent advances in the genetics of allergy and asthma. *Mol Med Today* 5:264-273.
- Andersson CK, Claesson HE, Rydell-Tormanen K, Swedmark S, Hallgren A and Erjefalt JS (2008) Mice Lacking 12/15-Lipoxygenase have Attenuated Airway Allergic Inflammation and Remodeling. *Am J Respir Cell Mol Biol* 39(6):648-656.
- Antonicelli (2001) Relationship between lipoprotein(a) levels, oxidative stress, and blood pressure levels in patients with essential hypertension. *Clin Exp Med* 1:145-150.
- Ardlie KG, Kruglyak L and Seielstad M (2002) Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* 3:299-309.
- Association pulmonaire du Canada (2008) <http://poumon.ca>.
- Association pulmonaire du Québec (2007) <http://www.pq.poumon.ca>.
- Ast G (2004) How did alternative splicing evolve? *Nat Rev Genet* 5:773-782.
- ATS (1987) Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. *Am Rev Respir Dis* 136 (1):225-44.
- ATS (1995) Standardization of Spirometry, 1994 Update. *Am J Respir Crit Care Med* 152:1107-1136.
- Badour K, McGavin MK, Zhang J, Freeman S, Vieira C, Filipp D, Julius M, Mills GB and Siminovitch KA (2007) Interaction of the Wiskott-Aldrich syndrome protein with sorting nexin 9 is required for CD28 endocytosis and cosignaling in T cells. *Proc Natl Acad Sci U S A* 104:1593-1598.

- Barbato A, Turato G, Baraldo S, Bazzan E, Calabrese F, Panizzolo C, Zanin ME, Zuin R, Maestrelli P, Fabbri LM and Saetta M (2006) Epithelial damage and angiogenesis in the airways of children with asthma. *Am J Respir Crit Care Med* 174:975-981.
- Barlic J, McDermott DH, Merrell MN, Gonzales J, Via LE and Murphy PM (2004) Interleukin (IL)-15 and IL-2 reciprocally regulate expression of the chemokine receptor CX3CR1 through selective NFAT1- and NFAT2-dependent mechanisms. *J Biol Chem* 279:48520-48534.
- Barnes KC (1999) Gene-environment and gene-gene interaction studies in the molecular genetic analysis of asthma and atopy. *Clin Exp Allergy*. 29 Suppl 4: 47-51.
- Beasley R (2004) The Global Burden of Asthma Report, Global Initiative for Asthma (GINA). Available from <http://www.ginasthma.org>.
- Becker A, Watson W, Ferguson A, Dimich-Ward H and Chan-Yeung M (2004) The Canadian asthma primary prevention study: outcomes at 2 years of age. *J Allergy Clin Immunol* 113:650-656.
- Begin P, Tremblay K, Daley D, Lemire M, Claveau S, Salesse C, Kacel S, Montpetit A, Becker A, Chan-Yeung M, Kozyrskyj AL, Hudson TJ and Laprise C (2007) Association of urokinase-type plasminogen activator with asthma and atopy. *Am J Respir Crit Care Med* 175:1109-1116.
- Benayoun L and Pretolani M (2003) Le remodelage bronchique dans l'asthme: mécanismes et enjeux thérapeutiques. *Medecine/Sciences* 19:319-26.
- Berend N, Salome CM and King GG (2008) Mechanisms of airway hyperresponsiveness in asthma. *Respirology* 13:624-631.
- Beuther DA, Weiss ST and Sutherland ER (2006) Obesity and asthma. *Am J Respir Crit Care Med* 174:112-119.
- Bjarnadottir (2004) The human and mouse repertoire of the adhesion family of G-protein-coupled receptors. *Genomics* 84:23-33.
- Bleecker ER, Postma DS and Meyers DA (1997) Evidence for multiple genetic susceptibility loci for asthma. *Am J Respir Crit Care Med* 156:S113-116.
- Blumenthal MN (2005) The role of genetics in the development of asthma and atopy. *Curr Opin Allergy Clin Immunol* 5:141-145.
- Bodmer W and Bonilla C (2008) Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* 40:695-701.
- Borenstein M, Kane J and Buchbinder J (1987) Monte Carlo simulation of statistical power. *Psychopharmacol Bull* 23:300-302.

- Borish L (1999) Genetics of allergy and asthma. *Ann Allergy Asthma Immunol* 82(5):413-24; quiz 424-6.
- Bosse Y and Hudson TJ (2007) Toward a comprehensive set of asthma susceptibility genes. *Annu Rev Med* 58:171-184.
- Bosse Y, Pare PD and Seow CY (2008) Airway wall remodeling in asthma: from the epithelial layer to the adventitia. *Curr Allergy Asthma Rep* 8:357-366.
- Boulet LP, Becker A, Berube D, Beveridge R and Ernst P (1999) [Summary of the recommendations of the Canadian Consensus Conference on Asthma 1999. Canadian Asthma Consensus Group]. *CMAJ* 161:SF1-14.
- Boulet LP (2003) Asymptomatic airway hyperresponsiveness: a curiosity or an opportunity to prevent asthma? *Am J Respir Crit Care Med* 167:371-378.
- Bousquet J, Knani J, Dhivert H, Richard A, Chicoye A, Ware JE, Jr. and Michel FB (1994) Quality of life in asthma. I. Internal consistency and validity of the SF-36 questionnaire. *Am J Respir Crit Care Med* 149:371-375.
- Bousquet J, Demoly P, Vignola AM, Godard P and Michel FB (1999) Comprendre la maladie asthmatique. *Médecine/Sciences* 15: 823-832.
- Bousquet J, Jeffery PK, Busse WW, Johnson M and Vignola AM (2000) Asthma. From bronchoconstriction to airways inflammation and remodeling. *Am J Respir Crit Care Med* 161(5):1720-45.
- Bradding P, Redington AE, Djukanovic R, Conrad DJ and Holgate ST (1995) 15-lipoxygenase immunoreactivity in normal and in asthmatic airways. *Am J Respir Crit Care Med* 151:1201-1204.
- Bradding P, Walls AF and Holgate ST (2006) The role of the mast cell in the pathophysiology of asthma. *J Allergy Clin Immunol* 117:1277-1284.
- Brand S, Hofbauer K, Dambacher J, Schnitzler F, Staudinger T, Pfennig S, Seiderer J, Tillack C, Konrad A, Goke B, Ochsenkuhn T and Lohse P (2006) Increased expression of the chemokine fractalkine in Crohn's disease and association of the fractalkine receptor T280M polymorphism with a fibrostenosing disease Phenotype. *Am J Gastroenterol* 101:99-106.
- Bresciani M, Parisi C, Menghi G and Bonini S (2005) The hygiene hypothesis: does it function worldwide? *Curr Opin Allergy Clin Immunol* 5:147-151.
- Brinckmann R, Topp MS, Zalan I, Heydeck D, Ludwig P, Kuhn H, Berdel WE and Habenicht JR (1996) Regulation of 15-lipoxygenase expression in lung epithelial cells by interleukin-4. *Biochem J* 318 (Pt 1):305-312.
- Broide D (2008) New perspectives on mechanisms underlying chronic allergic inflammation and asthma in 2007. *J Allergy Clin Immunol* 122:475-480.

- Burton PR, Tobin MD and Hopper JL (2005) Key concepts in genetic epidemiology. *Lancet* 366:941-951.
- Busse W, Elias J, Sheppard D and Banks-Schlegel S (1999) Airway remodeling and repair. *Am J Respir Crit Care Med* Sep;160(3):1035-42.
- Busse WW and Lemanske RF, Jr. (2001) Asthma. *N Engl J Med* 344:350-362.
- Busse WW and Rosenwasser LJ (2003) Mechanisms of asthma. *J Allergy Clin Immunol* 111:S799-804.
- CAMP (1999) The Childhood Asthma Management Program (CAMP): design, rationale, and methods. Childhood Asthma Management Program Research Group. *Control Clin Trials* 20:91-120.
- Carter CJ (2007) Convergence of genes implicated in Alzheimer's disease on the cerebral cholesterol shuttle: APP, cholesterol, lipoproteins, and atherosclerosis. *Neurochem Int* 50:12-38.
- Chamberland A, Madore AM, Tremblay K, Laviolette M and Laprise C (2008) A comparison of two sets of microarray experiments to define allergic asthma expression pattern. *Exp Lung Res* [Sous presse].
- Chan-Yeung M, Manfreda J, Dimich-Ward H, Ferguson A, Watson W and Becker A (2000) A randomized controlled study on the effectiveness of a multifaceted intervention program in the primary prevention of asthma in high-risk infants. *Arch Pediatr Adolesc Med* 154:657-663.
- Chan-Yeung M, Ferguson A, Watson W, Dimich-Ward H, Rousseau R, Lilley M, Dybuncio A and Becker A (2005) The Canadian Childhood Asthma Primary Prevention Study: outcomes at 7 years of age. *J Allergy Clin Immunol* 116:49-55.
- Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Abecasis G, Altshuler D, Bailey-Wilson JE, Brooks LD, Cardon LR, Daly M, Donnelly P, Fraumeni JF, Jr., Freimer NB, Gerhard DS, Gunter C, Guttmacher AE, Guyer MS, Harris EL, Hoh J, Hoover R, Kong CA, Merikangas KR, Morton CC, Palmer LJ, Phimister EG, Rice JP, Roberts J, Rotimi C, Tucker MA, Vogan KJ, Wacholder S, Wijsman EM, Winn DM and Collins FS (2007) Replicating genotype-phenotype associations. *Nature* 447:655-660.
- Chelbi H, Ghadiri A, Lacheb J, Ghandil P, Hamzaoui K, Hamzaoui A and Combadiere C (2008) A polymorphism in the CCL2 chemokine gene is associated with asthma risk: a case-control and a family study in Tunisia. *Genes Immun* 9(7):575-581.
- Chu EK, Cheng J, Foley JS, Mecham BH, Owen CA, Haley KJ, Mariani TJ, Kohane IS, Tschumperlin DJ and Drazen JM (2006) Induction of the plasminogen activator system by mechanical stimulation of human bronchial epithelial cells. *Am J Respir Cell Mol Biol* 35:628-638.



- Chu HW, Balzar S, Westcott JY, Trudeau JB, Sun Y, Conrad DJ and Wenzel SE (2002) Expression and activation of 15-lipoxygenase pathway in severe asthma: relationship to eosinophilic phenotype and collagen deposition. *Clin Exp Allergy* 32:1558-1565.
- Chung KF and Barnes PJ (1999) Cytokines in asthma. *Thorax* 54(9):825-857.
- Cockcroft DW (2001) How best to measure airway responsiveness. *Am J Respir Crit Care Med* 163:1514-1515.
- Colhoun HM, McKeigue PM and Davey Smith G (2003) Problems of reporting genetic associations with complex outcomes. *Lancet* 361:865-872.
- Conrad DJ, Kuhn H, Mulkins M, Highland E and Sigal E (1992) Specific inflammatory cytokines regulate the expression of human monocyte 15-lipoxygenase. *Proc Natl Acad Sci U S A* 89:217-221.
- Cooke R and van der Veer A (1916) Human sensitization. *J Immunol* 1:201-205.
- Cookson W (2004) The immunogenetics of asthma and eczema: a new focus on the epithelium. *Nat Rev Immunol* 4:978-988.
- Cookson WO (1999) The alliance of genes and environment in asthma and allergy. *Nature* 402:B5-11.
- Cookson WO and Moffatt MF (2002) The genetics of atopic dermatitis. *Curr Opin Allergy Clin Immunol* 2:383-387.
- Cookson WO (2002) Asthma genetics. *Chest* 121(3 Suppl):7S-13S.
- Cordell HJ, Barratt BJ and Clayton DG (2004) Case/pseudocontrol analysis in genetic association studies: A unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. *Genet Epidemiol* 26:167-185.
- Cordell HJ and Clayton DG (2005) Genetic association studies. *Lancet* 366:1121-1131.
- Daniels SE, Bhattacharria S, James A, Leaves NI, Young A, Hill MR, Faux JA, Ryan GF, le Souef PN, Lathrop GM, Musk AW and Cookson WO (1996) A genome-wide search for quantitative trait loci underlying asthma. *Nature* 383:247-250.
- Davies DE, Wicks J, Powell RM, Puddicombe SM and Holgate ST (2003) Airway remodeling in asthma: new insights. *J Allergy Clin Immunol* 111:215-225; quiz 226.
- Dehmel T, Janke A, Hartung HP, Goebel HH, Wiendl H and Kieseier BC (2007) The cell-specific expression of metalloproteinase-disintegrins (ADAMs) in inflammatory myopathies. *Neurobiol Dis* 25:665-674.

- Depner M, Kormann MS, Klopp N, Illig T, Vogelberg C, Weiland SK, von Mutius E, Combadiere C and Kabesch M (2007) CX3CR1 polymorphisms are associated with atopy but not asthma in German children. *Int Arch Allergy Immunol* 144:91-94.
- Devlin B, Roeder K and Wasserman L (2001) Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol* 60:155-166.
- Dreyfuss G, Hentze M and Lamond AI (1996) From transcript to protein. *Cell* 28 (85):963-972.
- Dubé J and Boulet LP (1996) Rôles de l'inflammation et des modifications des structures bronchiques dans l'asthme allergique. *Médecine/Sciences* 3: 351-57.
- Duffy DL, Martin NG, Battistutta D, Hopper JL and Mathews JD (1990) Genetics of asthma and hay fever in Australian twins. *Am Rev Respir Dis* 142:1351-1358.
- El-Shazly (2006) Fraktalkine produced by airway smooth muscle cells contributes to mast cell recruitment in asthma. *J Immunol* 176:1860-1868.
- El-Shazly A, Berger P, Girodet PO, Ousova O, Fayon M, Vernejoux JM, Marthan R and Tunon-de-Lara JM (2006) Fraktalkine produced by airway smooth muscle cells contributes to mast cell recruitment in asthma. *J Immunol* 176:1860-1868.
- Eley TC and Rijdsdijk F (2005) Introductory guide to the statistics of molecular genetics. *J Child Psychol Psychiatry* 46:1042-1044.
- Elston RC (1995) The genetic dissection of multifactorial traits. *Clin Exp Allergy* 25 Suppl 2:103-106.
- Engert JC, Berube P, Mercier J, Dore C, Lepage P, Ge B, Bouchard JP, Mathieu J, Melancon SB, Schalling M, Lander ES, Morgan K, Hudson TJ and Richter A (2000) ARSACS, a spastic ataxia common in northeastern Quebec, is caused by mutations in a new gene encoding an 11.5-kb ORF. *Nat Genet* 24:120-125.
- Falliers CJ, de Cardoso RR, Bane HN, Coffey R and Middleton E, Jr. (1971) Discordant allergic manifestations in monozygotic twins: genetic identity versus clinical, physiologic, and biochemical differences. *J Allergy* 47:207-219.
- Feltenmark S, Gautam N, Brunnstrom A, Griffiths W, Backman L, Edenius C, Lindbom L, Bjorkholm M and Claesson HE (2008) Eoxins are proinflammatory arachidonic acid metabolites produced via the 15-lipoxygenase-1 pathway in human eosinophils and mast cells. *Proc Natl Acad Sci U S A* 105:680-685.
- Folli C, Descalzi D, Scordamaglia F, Riccio AM, Gamalero C and Canonica GW (2008) New insights into airway remodelling in asthma and its possible modulation. *Curr Opin Allergy Clin Immunol* 8:367-375.
- Frostegard (2002) Autoimmunity, oxidized LDL and cardiovascular disease. *Autoimmun Rev* 1:233-237.

- Fujimoto K, Imaizumi T, Yoshida H, Takanashi S, Okumura K and Satoh K (2001) Interferon-gamma stimulates fractalkine expression in human bronchial epithelial cells and regulates mononuclear cell adherence. *Am J Respir Cell Mol Biol* 25:233-238.
- Galli SJ, Tsai M and Piliponsky AM (2008) The development of allergic inflammation. *Nature* 454:445-454.
- Gauderman WJ (2002) Sample size requirements for association studies of gene-gene interaction. *Am J Epidemiol* 1;155(5):478-84.
- Global initiative for asthma (2007) Global strategy for asthma management and prevention. *Rapport mis à jour téléchargé le 20 Décembre 2007* (<http://www.ginasthma.com/>).
- Goldstein DB, Tate SK and Sisodiya SM (2003) Pharmacogenetics goes genomic. *Nat Rev Genet* 4:937-947.
- Grant SF and Hakonarson H (2008) Microarray technology and applications in the arena of genome-wide association. *Clin Chem* 54:1116-1124.
- Gray IC, Campbell DA and Spurr NK (2000) Single nucleotide polymorphisms as tools in human genetics. *Hum Mol Genet* 9:2403-2408.
- Guerra S and Martinez FD (2008) Asthma genetics: from linear to multifactorial approaches. *Annu Rev Med* 59:327-341.
- Guilbert TW, Morgan WJ, Zeiger RS, Bacharier LB, Boehmer SJ, Krawiec M, Larsen G, Lemanske RF, Liu A, Mauger DT, Sorkness C, Szeffler SJ, Strunk RC, Taussig LM and Martinez FD (2004) Atopic characteristics of children with recurrent wheezing at high risk for the development of childhood asthma. *J Allergy Clin Immunol* 114:1282-1287.
- Gulliksson M, Brunnstrom A, Johannesson M, Backman L, Nilsson G, Harvima I, Dahlen B, Kumlin M and Claesson HE (2007) Expression of 15-lipoxygenase type-1 in human mast cells. *Biochim Biophys Acta* 1771:1156-1165.
- Hajek AR, Lindley AR, Favoreto S, Jr., Carter R, Schleimer RP and Kuperman DA (2008) 12/15-Lipoxygenase deficiency protects mice from allergic airways inflammation and increases secretory IgA levels. *J Allergy Clin Immunol* 22(3):6336-399e3.
- Hakonarson H and Halapi E (2002) Genetic analyses in asthma: current concepts and future directions. *Am J Pharmacogenomics* 2:155-166.
- Halapi E and Hakonarson H (2004) Recent development in genomic and proteomic research for asthma. *Curr Opin Pulm Med* 10:22-30.
- Hall IP and Blakey JD (2005) Genetic association studies in Thorax. *Thorax* 60:357-359.

- Hanson B, McGue M, Roitman-Johnson B, Segal NL, Bouchard TJ, Jr. and Blumenthal MN (1991) Atopic disease and immunoglobulin E in twins reared apart and together. *Am J Hum Genet* 48:873-879.
- Hartl DL and Clark AG (1997) Principles of population genetics. *Sunderland, thrih edition*:Sinauer, 481 p.
- Hattersley AT and McCarthy MI (2005) What makes a good genetic association study? *Lancet* 366:1315-1323.
- Hersh CP, Raby BA, Soto-Quiros ME, Murphy AJ, Avila L, Lasky-Su J, Sylvia JS, Klanderman BJ, Lange C, Weiss ST and Celedon JC (2007) Comprehensive Testing of Positionally Cloned Asthma Genes in Two Populations. *Am J Respir Crit Care Med*.176(9):849-857.
- Heyer E and Tremblay M (1995) Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. *Am J Hum Genet* 56:970-978.
- Heyer E, Tremblay M and Desjardins B (1997) Seventeenth-century European origins of hereditary diseases in the Saguenay population (Quebec, Canada). *Hum Biol* 69:209-225.
- Higuchi (2002) CD156 transgenic mice. Different responses between inflammatory types. *Pathobiology* 70:47-54.
- Hill MR, James AL, Faux JA, Ryan G, Hopkin JM, le Souef P, Musk AW and Cookson WO (1995) Fc epsilon RI-beta polymorphism and risk of atopy in a general population sample. *BMJ* 311:776-779.
- Hill RA, Standen PJ and Tattersfield AE (1989) Asthma, wheezing, and school absence in primary schools. *Arch Dis Child* 64:246-251.
- Hirschhorn JN and Altshuler D (2002) Once and again-issues surrounding replication in genetic association studies. *J Clin Endocrinol Metab* 87:4438-4441.
- Hirschhorn JN, Lohmueller K, Byrne E and Hirschhorn K (2002) A comprehensive review of genetic association studies. *Genet Med* 4:45-61.
- Hoffjan S and Ober C (2002) Present status on the genetic studies of asthma. *Curr Opin Immunol* 14:709-717.
- Hoffjan S, Nicolae D and Ober C (2003) Association studies for asthma and atopic diseases: a comprehensive review of the literature. *Respir Res* 4:14.
- Holgate ST (1999) The epidemic of allergy and asthma. *Nature* 402:B2-4.
- Holgate ST, Lackie P, Wilson S, Roche W and Davies D (2000) Bronchial epithelium as a key regulator of airway allergen sensitization and remodeling in asthma. *Am J Respir Crit Care Med* 162(3 Pt 2):S113-117.

- Holgate ST, Davies DE, Powell RM, Howarth PH, Haitchi HM and Holloway JW (2007) Local genetic and environmental factors in asthma disease pathogenesis: chronicity and persistence mechanisms. *Eur Respir J* 29:793-803.
- Holgate ST (2008) Pathogenesis of asthma. *Clin Exp Allergy* 38:872-897.
- Holloway JW and Koppelman GH (2007) Identifying novel genes contributing to asthma pathogenesis. *Curr Opin Allergy Clin Immunol* 7:69-74.
- Holt PG, Macaubas C, Stumbles PA and Sly PD (1999) The role of allergy in the development of asthma. *Nature* 402:B12-17.
- Hopkins AL and Groom CR (2002) The druggable genome. *Nat Rev Drug Discov* 1:727-730.
- Hopp RJ, Bewtra AK, Watt GD, Nair NM and Townley RG (1984) Genetic analysis of allergic disease in twins. *J Allergy Clin Immunol* 73:265-270.
- Horvath S, Xu X and Laird NM (2001) The family based association test method: strategies for studying general genotype-phenotype associations. *Eur J Hum Genet* 9(4):301-306.
- Howard TD, Koppelman GH, Xu J, Zheng SL, Postma DS, Meyers DA and Bleeker ER (2002) Gene-gene interaction in asthma: IL4RA and IL13 in a Dutch population with asthma. *Am J Hum Genet* 70(1):230-236.
- Howson JM, Barratt BJ, Todd JA and Cordell HJ (2005) Comparison of population- and family-based methods for genetic association analysis in the presence of interacting loci. *Genet Epidemiol* 29:51-67.
- Huss K and Huss RW (2000) Genetics of asthma and allergies. *Nurs Clin North Am* 35(3):695-705.
- Ichinose (1999) Airway autonomic nervous system dysfunction and asthma. *Nihon Koryuiki Gakkai Zhassi* 37:3-9.
- Illig T and Wjst M (2002) Genetics of asthma and related phenotypes. *Paediatr Respir Rev* 3(1):47-51.
- Imaizumi T, Matsumiya T, Tamo W, Shibata T, Fujimoto K, Kumagai M, Yoshida H, Cui XF, Tanji K, Hatakeyama M, Wakabayashi K and Satoh K (2002) 15-Deoxy-D12,14-prostaglandin J2 inhibits CX3CL1/fractalkine expression in human endothelial cells. *Immunol Cell Biol* 80:531-536.
- Imaizumi T, Yoshida H and Satoh K (2004) Regulation of CX3CL1/fractalkine expression in endothelial cells. *J Atheroscler Thromb* 11:15-21.
- Immervoll T and Wjst M (1999) Current status of the Asthma and Allergy Database. *Nucleic Acids Res* 27:213-214.

- International HapMap Consortium (2003) The International HapMap Project. *Nature* 426:789-796.
- International HapMap Consortium (2005) A haplotype map of the human genome. *Nature* 437:1299-1320.
- International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431:931-945.
- Ioannidis JP, Ntzani EE, Trikalinos TA and Contopoulos-Ioannidis DG (2001) Replication validity of genetic association studies. *Nat Genet* 29:306-309.
- Ishikawa (2004) ADAM8 as a novel serological and histochemical marker for lung cancer. *Clin Cancer Res* 10:8363-8370.
- Ito K, Chung KF and Adcock IM (2006) Update on glucocorticoid action and resistance. *J Allergy Clin Immunol* 117:522-543.
- James A (2005) Airway remodeling in asthma. *Curr Opin Pulm Med* 11:1-6.
- James AL, Palmer LJ, Kicic E, Maxwell PS, Lagan SE, Ryan GF and Musk AW (2005) Decline in lung function in the Busselton Health Study: the effects of asthma and cigarette smoking. *Am J Respir Crit Care Med* 171:109-114.
- Jansen DF, Timens W, Kraan J, Rijcken B and Postma DS (1997) (A)symptomatic bronchial hyper-responsiveness and asthma. *Respir Med* 91:121-134.
- Jeffery PK, Wardlaw AJ, Nelson FC, Collins JV and Kay AB (1989) Bronchial biopsies in asthma. An ultrastructural, quantitative study and correlation with hyperreactivity. *Am Rev Respir Dis* 140:1745-1753.
- Jorgensen (1994) Modulation of the immune response by the neuro-endocrine axis in rheumatoid arthritis. *Clin Exp Rheumatol.* 12:435-441.
- Kabesch M, Schedel M, Carr D, Woitsch B, Fritzsich C, Weiland SK and von Mutius E (2006) IL-4/IL-13 pathway genetics strongly influence serum IgE levels and childhood asthma. *J Allergy Clin Immunol* 117:269-274.
- Kay AB (2005) The role of eosinophils in the pathogenesis of asthma. *Trends Mol Med* 11:148-152.
- Kicic A, Sutanto EN, Stevens PT, Knight DA and Stick SM (2006) Intrinsic biochemical and functional differences in bronchial epithelial cells of children with asthma. *Am J Respir Crit Care Med* 174:1110-1118.
- Kiley J, Smith R and Noel P (2007) Asthma phenotypes. *Curr Opin Pulm Med* 13:19-23.
- Kiyohara C, Tanaka K and Miyake Y (2008) Genetic susceptibility to atopic dermatitis. *Allergol Int* 57:39-56.

- Koh MS and Irving LB (2007) The natural history of asthma from childhood to adulthood. *Int J Clin Pract* 61:1371-1374.
- Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, Weissenbach J, Popowska E, Pronicka E, Root AW and Glorieux FH (1996) Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am J Hum Genet* 59:633-643.
- Laird NM and Lange C (2008) Family-based methods for linkage and association analysis. *Adv Genet* 60:219-252.
- Laitinen T, Rasanen M, Kaprio J, Koskenvuo M and Laitinen LA (1998) Importance of genetic factors in adolescent asthma: a population-based twin-family study. *Am J Respir Crit Care Med* 157:1073-1078.
- Laitinen T, Daly MJ, Rioux JD, Kauppi P, Laprise C, Petays T, Green T, Cargill M, Haahtela T, Lander ES, Laitinen LA, Hudson TJ and Kere J (2001) A susceptibility locus for asthma-related traits on chromosome 7 revealed by genome-wide scan in a founder population. *Nat Genet* 28:87-91.
- Laitinen T, Polvi A, Rydman P, Vendelin J, Pulkkinen V, Salmikangas P, Makela S, Rehn M, Pirskanen A, Rautanen A, Zucchelli M, Gullsten H, Leino M, Alenius H, Petays T, Haahtela T, Laitinen A, Laprise C, Hudson TJ, Laitinen LA and Kere J (2004) Characterization of a common susceptibility locus for asthma-related traits. *Science* 304:300-304.
- Lander ES and Schork NJ (1994) Genetic dissection of complex traits. *Science* 21;266(5184):353.
- Laoukili (2001) IL-13 alters mucociliary differentiation and ciliary beating of human respiratory epithelial cells. *J Clin Invest* 108:1817-1824.
- Laprise C and Boulet LP (1996) Airway responsiveness and atopy in families of patients with asthma. *Clin Invest Med* 19:461-469.
- Laprise C, Laviolette M, Boutet M and Boulet LP (1999) Asymptomatic airway hyperresponsiveness: relationships with airway inflammation and remodelling. *Eur Respir J* 14:63-73.
- Laprise C, Sladek R, Ponton A, Bernier MC, Hudson TJ and Laviolette M (2004) Functional classes of bronchial mucosa genes that are differentially expressed in asthma. *BMC Genomics* 5:21-31.
- Lawrence PA (2003) The politics of publication. *Nature* 422:259-261.
- Lemanske RF, Jr. and Busse WW (2003) 6. Asthma. *J Allergy Clin Immunol* 111:S502-519.

- Lemanske RF, Jr. and Busse WW (2006) 6. Asthma: Factors underlying inception, exacerbation, and disease progression. *J Allergy Clin Immunol* 117:S456-461.
- Lemire M, Roslin NM, Laprise C, Hudson TJ and Morgan K (2004) Transmission-ratio distortion and allele sharing in affected sib pairs: a new linkage statistic with reduced bias, with application to chromosome 6q25.3. *Am J Hum Genet* 75:571-586.
- Lemire M (2004) Calcul de puissance et analyses d'interaction. *Communications personnelles*, Centre d'innovation Génome Québec et Université McGill, Montréal (Canada).
- Levy BD, Romano M, Chapman HA, Reilly JJ, Drazen J and Serhan CN (1993) Human alveolar macrophages have 15-lipoxygenase and generate 15(S)-hydroxy-5,8,11-cis-13-trans-eicosatetraenoic acid and lipoxins. *J Clin Invest* 92:1572-1579.
- Levy BD, Clish CB, Schmidt B, Gronert K and Serhan CN (2001) Lipid mediator class switching during acute inflammation: signals in resolution. *Nat Immunol* 2:612-619.
- Levy BD, De Sanctis GT, Devchand PR, Kim E, Ackerman K, Schmidt BA, Szczeklik W, Drazen JM and Serhan CN (2002) Multi-pronged inhibition of airway hyper-responsiveness and inflammation by lipoxin A(4). *Nat Med* 8:1018-1023.
- Levy BD, Bonnans C, Silverman ES, Palmer LJ, Marigowda G and Israel E (2005) Diminished lipoxin biosynthesis in severe asthma. *Am J Respir Crit Care Med* 172:824-830.
- Li J and Ji L (2005) Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 95:221-227.
- Lin PI, Vance JM, Pericak-Vance MA and Martin ER (2007) No gene is an island: the flip-flop phenomenon. *Am J Hum Genet* 80:531-538.
- Los H, Postmus PE and Boomsma DI (2001) Asthma genetics and intermediate phenotypes: a review from twin studies. *Twin Res* 4:81-93.
- Luan JA, Wong MY, Day NE and Wareham NJ (2001) Sample size determination for studies of gene-environment interaction. *Int J Epidemiol* 30:1035-1040.
- Maddox L and Schwartz DA (2002) The pathophysiology of asthma. *Annu Rev Med* 53:477-498.
- Madore AM, Tremblay K, Hudson TJ and Laprise C (2008) Replication of an association between 17q21 SNPs and asthma in a French-Canadian familial collection. *Hum Genet* 123:93-95.
- Malerba G and Pignatti PF (2005) A review of asthma genetics: gene expression studies and recent candidates. *J Appl Genet* 46:93-104.



- Marion RJ, Creer TL and Reynolds RV (1985) Direct and indirect costs associated with the management of childhood asthma. *Ann Allergy* 54:31-34.
- Masoli M, Fabian D, Holt S and Beasley R (2004) The global burden of asthma: executive summary of the GINA Dissemination Committee report. *Allergy* 59:469-478.
- Matsuno (2005) Role of ADAM8 in experimental asthma. *Immunol Lett.* 102:67-73.
- Matsuno O, Miyazaki E, Nureki S, Ueno T, Ando M, Ito K, Kumamoto T and Higuchi Y (2007) Elevated Soluble ADAM8 in Bronchoalveolar Lavage Fluid in Patients with Eosinophilic Pneumonia. *Int Arch Allergy Immunol* 142:285-290.
- McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP and Hirschhorn JN (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 9:356-369.
- McRobert (2003) The amino-terminal domains of the ezrin, radixin, and moesin (ERM) proteins bind advanced glycation end products, an interaction that may play a role in the development of diabetic complications. *J Biol Chem* 278:25783-25789.
- Moffatt MF and Cookson WO (1998) Gene identification in asthma and allergy. *Int Arch Allergy Immunol* 116:247-252.
- Moffatt MF, Schou C, Faux JA, Abecasis GR, James A, Musk AW and Cookson WO (2001) Association between quantitative traits underlying asthma and the HLA-DRB1 locus in a family-based population sample. *Eur J Hum Genet* 9:341-346.
- Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, Depner M, von Berg A, Bufe A, Rietschel E, Heinzmann A, Simma B, Frischer T, Willis-Owen SA, Wong KC, Illig T, Vogelberg C, Weiland SK, von Mutius E, Abecasis GR, Farrall M, Gut IG, Lathrop GM and Cookson WO (2007) Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 448:470-473.
- Moffatt MF (2008) Genes in asthma: new genes and new ways. *Curr Opin Allergy Clin Immunol* 8:411-417.
- Moore WC and Peters SP (2007) Update in asthma 2006. *Am J Respir Crit Care Med* 175:649-654.
- Murthy V, Julien P and Gagne C (1996) Molecular pathobiology of the human lipoprotein lipase gene. *Pharmacol Ther* 70:101-135.
- Nassar GM, Morrow JD, Roberts LJ, 2nd, Lakkis FG and Badr KF (1994) Induction of 15-lipoxygenase by interleukin-13 in human blood monocytes. *J Biol Chem* 269:27631-27634.
- Nature Genetics Editorial (1999) Freely associating. *Nat Genet* 22:1-2.

- Naus (2006) Identification of candidate substrates for ectodomain shedding by the metalloprotease-disintegrin ADAM8. *Biol Chem.* 387:337-346.
- Nauta AJ, Engels F, Knippels LM, Garssen J, Nijkamp FP and Redegeld FA (2008) Mechanisms of allergy and asthma. *Eur J Pharmacol* 585:354-360.
- Neale BM and Sham PC (2004) The future of association studies: gene-based analysis and replication. *Am J Hum Genet* 75:353-362.
- Newton-Cheh C and Hirschhorn JN (2005) Genetic association studies of complex traits: design and analysis issues. *Mutat Res* 573:54-69.
- Nicolae D, Cox NJ, Lester LA, Schneider D, Tan Z, Billstrand C, Kuldanek S, Donfack J, Kogut P, Patel NM, Goodenbour J, Howard T, Wolf R, Koppelman GH, White SR, Parry R, Postma DS, Meyers D, Bleecker ER, Hunt JS, Solway J and Ober C (2005) Fine mapping and positional candidate studies identify HLA-G as an asthma susceptibility gene on chromosome 6p21. *Am J Hum Genet* 76:349-357.
- Nieminen MM, Kaprio J and Koskenvuo M (1991) A population-based study of bronchial asthma in adult twin pairs. *Chest* 100:70-75.
- Noguchi E, Yokouchi Y, Zhang J, Shibuya K, Shibuya A, Bannai M, Tokunaga K, Doi H, Tamari M, Shimizu M, Shirakawa T, Shibasaki M, Ichikawa K and Arinami T (2005) Positional identification of an asthma susceptibility gene on human chromosome 5q33. *Am J Respir Crit Care Med* 172:183-188.
- Nyholt DR (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 74:765-769.
- Ober C, Abney M and McPeck MS (2001) The genetic dissection of complex traits in a founder population. *Am J Hum Genet* 69:1068-1079.
- Ober C and Thompson EE (2005) Rethinking genetic models of asthma: the role of environmental modifiers. *Curr Opin Immunol* 17:670-678.
- Ober C and Hoffjan S (2006) Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 7:95-100.
- Ober C, Tan Z, Sun Y, Possick JD, Pan L, Nicolae R, Radford S, Parry RR, Heinzmann A, Deichmann KA, Lester LA, Gern JE, Lemanske RF, Jr., Nicolae DL, Elias JA and Chupp GL (2008) Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. *N Engl J Med* 358:1682-1691.
- Packard KA and Khan MM (2003) Effects of histamine on Th1/Th2 cytokine balance. *Int Immunopharmacol* 3:909-920.
- Palmer LJ and Cookson WO (2000) Genomic approaches to understanding asthma. *Genome Res* 10:1280-1287.

- Palmer LJ, Knuiiman MW, Divitini ML, Burton PR, James AL, Bartholomew HC, Ryan G and Musk AW (2001) Familial aggregation and heritability of adult lung function: results from the Busselton Health Study. *Eur Respir J* 17:696-702.
- Palmer LJ and Cookson WO (2001) Using single nucleotide polymorphisms as a means to understanding the pathophysiology of asthma. *Respir Res* 2:102-112.
- Palmer LJ and Cardon LR (2005) Shaking the tree: mapping complex disease genes with linkage disequilibrium. *Lancet* 366:1223-1234.
- Panettieri RA, Jr., Covar R, Grant E, Hillyer EV and Bacharier L (2008) Natural history of asthma: persistence versus progression-does the beginning predict the end? *J Allergy Clin Immunol* 121:607-613.
- Pascual RM and Peters SP (2005) Airway remodeling contributes to the progressive loss of lung function in asthma: an overview. *J Allergy Clin Immunol* 116:477-486; quiz 487.
- Peltonen L, Palotie A and Lange K (2000) Use of population isolates for mapping complex traits. *Nat Rev Genet* 1:182-90.
- Perez-Iratxeta C, Bork P and Andrade MA (2002) Association of genes to genetically inherited diseases using data mining. *Nat Genet* 31:316-319.
- Perez-Iratxeta C, Wjst M, Bork P and Andrade MA (2005) G2D: a tool for mining genes associated with disease. *BMC Genet* 6:45.
- Perez-Iratxeta C, Bork P and Andrade-Navarro MA (2007) Update of the G2D tool for prioritization of gene candidates to inherited diseases. *Nucleic Acids Res* 35:W212-216.
- Permutt MA, Wasson J and Cox N (2005) Genetic epidemiology of diabetes. *J Clin Invest* 115:1431-1439.
- Planaguma A, Kazani S, Marigowda G, Haworth O, Mariani TJ, Israel E, Bleecker ER, Curran-Everett D, Erzurum SC, Calhoun WJ, Castro M, Chung KF, Gaston B, Jarjour NN, Busse WW, Wenzel SE and Levy BD (2008) Airway lipoxin A4 generation and lipoxin A4 receptor expression are decreased in severe asthma. *Am J Respir Crit Care Med* 178:574-582.
- Postma DS (2007) Gender differences in asthma development and progression. *Genet Med* 4 Suppl B:S133-146.
- Pritchard JK and Rosenberg NA (1999) Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 65:220-228.
- Profita M, Vignola AM, Sala A, Mirabella A, Siena L, Pace E, Folco G and Bonsignore G (1999) Interleukin-4 enhances 15-lipoxygenase activity and incorporation of 15(S)-HETE into cellular phospholipids in cultured pulmonary epithelial cells. *Am J Respir Cell Mol Biol* 20:61-68.

- Profita M, Sala A, Riccobono L, Paterno A, Mirabella A, Bonanno A, Guerrera D, Pace E, Bonsignore G, Bousquet J and Vignola AM (2000a) 15-Lipoxygenase expression and 15(S)-hydroxyeicoisatetraenoic acid release and reincorporation in induced sputum of asthmatic subjects. *J Allergy Clin Immunol* 105:711-716.
- Profita M, Sala A, Riccobono L, Pace E, Paterno A, Zarini S, Siena L, Mirabella A, Bonsignore G and Vignola AM (2000b) 15(S)-HETE modulates LTB(4) production and neutrophil chemotaxis in chronic bronchitis. *Am J Physiol Cell Physiol* 279:C1249-1258.
- Purcell S, Cherny SS and Sham PC (2003) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 19:149-150.
- Rao DC (2008) An overview of the genetic dissection of complex traits. *Adv Genet* 60:3-34.
- Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW and Hurles ME (2006) Global variation in copy number in the human genome. *Nature* 444:444-454.
- Reed CE (2006) The natural history of asthma. *J Allergy Clin Immunol* 118:543-548; quiz 549-550.
- Reich DE, Gabriel SB and Altshuler D (2003) Quality and completeness of SNP databases. *Nat Genet* 33:457-458.
- Richens J, Fairclough L, Ghaemmaghani AM, Mahdavi J, Shakib F and Sewell HF (2007) The detection of ADAM8 protein on cells of the human immune system and the demonstration of its expression on peripheral blood B cells, dendritic cells and monocyte subsets. *Immunobiology* 212:29-38.
- Richter A, Rioux JD, Bouchard JP, Mercier J, Mathieu J, Ge B, Poirier J, Julien D, Gyapay G, Weissenbach J, Hudson TJ, Melancon SB and Morgan K (1999) Location score and haplotype analyses of the locus for autosomal recessive spastic ataxia of Charlevoix-Saguenay, in chromosome region 13q11. *Am J Hum Genet* 64:768-775.
- Rijcken B, Schouten JP, Weiss ST, Speizer FE and van der Lende R (1987) The relationship of nonspecific bronchial responsiveness to respiratory symptoms in a random population sample. *Am Rev Respir Dis* 136:62-68.
- Risch N and Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* 273:1516-1517.

- Rolph MS, Sisavanh M, Liu SM and Mackay CR (2005) Clues to asthma pathogenesis from microarray expression studies. *Pharmacol Ther* 109:284-294.
- Rolph MS, Sisavanh M, Liu SM and Mackay CR (2006) Clues to asthma pathogenesis from microarray expression studies. *Pharmacol Ther* 109:284-294.
- Rommens J, Iannuzzi M, Kerem B, Drumm M, Melmer G, Dean M, Rozmahel R, Cole J, Kennedy D, Hidaka N and et al. (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* 245:1059-1065.
- Rosenwasser LJ (1996) Genetics of asthma and atopy. *Toxicol Lett* 86:73-77.
- Rousseau F and Laflamme N (2003) Génétique moléculaire: des maladies monogéniques aux maladies complexes. *Médecine/Sciences* 19: 950-954.
- Sandford A, Weir T and Pare P (1996) The genetics of asthma. *Am J Respir Crit Care Med* 153:1749-1765.
- Schleimer RP and Bochner BS (1998) The role of adhesion molecules in allergic inflammation and their suitability as targets of antiallergic therapy. *Clin Exp Allergy* 28 Suppl 3:15-23.
- Schork NJ, Wessel J and Malo N (2008) DNA sequence-based phenotypic association analysis. *Adv Genet* 60:195-217.
- Schumman (1998) Cytokine-induced transcription of protein-tyrosine-phosphatases in human astrocytoma cells. *Brain Res Mol Brain Res* 62:56-64.
- Scriver CR (2001) Human genetics: lessons from Quebec populations. *Annu Rev Genomics Hum Genet* 2:69-101.
- Sechler JM, Barlic J, Grivel JC and Murphy PM (2004) IL-15 alters expression and function of the chemokine receptor CX3CR1 in human NK cells. *Cell Immunol* Aug 230:99-108.
- Shannon VR, Chanez P, Bousquet J and Holtzman MJ (1993) Histochemical evidence for induction of arachidonate 15-lipoxygenase in airway disease. *Am Rev Respir Dis* 147:1024-1028.
- Shriner D, Vaughan LK, Padilla MA and Tiwari HK (2007) Problems with genome-wide association studies. *Science* 316:1840-1842.
- Sibbald B, Horn ME, Brain EA and Gregg I (1980) Genetic factors in childhood asthma. *Thorax* 35:671-674.
- Sigal E, Dicharry S, Highland E and Finkbeiner WE (1992) Cloning of human airway 15-lipoxygenase: identity to the reticulocyte enzyme and expression in epithelium. *Am J Physiol* 262:L392-398.

- Silverman EK and Palmer LJ (2000) Case-control association studies for the Genetics of Complex REspiratory diseases. *Am. J. Respir. Cell Mol. Biol.* 22; 645-648.
- Simpson A, John SL, Jury F, Niven R, Woodcock A, Ollier WE and Custovic A (2006) Endotoxin exposure, CD14, and allergic disease: an interaction between genes and the environment. *Am J Respir Crit Care Med* 174:386-392.
- Skadhauge LR, Christensen K, Kyvik KO and Sigsgaard T (1999) Genetic and environmental influence on asthma: a population-based study of 11,688 Danish twin pairs. *Eur Respir J* 13:8-14.
- Smith DH, Malone DC, Lawson KA, Okamoto LJ, Battista C and Saunders WB (1997) A national estimate of the economic costs of asthma. *Am J Respir Crit Care Med* 156:787-793.
- Smith DM (1988) The effect of respiratory disease on working life. *J R Soc Med* 81:258-260.
- Sotiriou (2006) Lipoprotein(a) in atherosclerotic plaques recruits inflammatory cells through interaction with Mac-1 integrin. *FASEB J* 20:559-561.
- Spielman RS, McGinnis RE and Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52:506-516.
- Spielman RS and Ewens WJ (1996) The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet* 59:983-989.
- Steemers FJ, Chang W, Lee G, Barker DL, Shen R and Gunderson KL (2006) Whole-genome genotyping with the single-base extension assay. *Nat Methods* 3:31-33.
- Steinke JW, Rich SS and Borish L (2008) 5. Genetics of allergic disease. *J Allergy Clin Immunol* 121:S384-387; quiz S416.
- Sub-Committee on Skin Tests of the European Academy of Allergology and Clinical Immunology (1989) Skin tests used in type I allergy testing Position paper. *Allergy* 44 Suppl 10:1-59.
- Sukkar MB, Issa R, Xie S, Oltmanns U, Newton R and Chung KF (2004) Fractalkine/CX3CL1 production by human airway smooth muscle cells: induction by IFN-gamma and TNF-alpha and regulation by TGF-beta and corticosteroids. *Am J Physiol Lung Cell Mol Physiol* 287:L1230-1240.
- Sully (2001) Functional abnormalities in protein tyrosine phosphatase epsilon-deficient macrophages. *Biochem Biophys Res Commun.* 286:184-188.
- Sumi Y and Hamid Q (2007) Airway remodeling in asthma. *Allergol Int* 56:341-348.
- Szczeklik A and Sanak M (2006) The broken balance in aspirin hypersensitivity. *Eur J Pharmacol* 533:145-155.

- Szema (2006) Mice lacking the VIP gene show airway hyperresponsiveness and airway inflammation, partially reversible by VIP. *Am J Physiol Lung Cell Mol Physiol* 291:L880-886.
- Tagaya E and Tamaoki J (2007) Mechanisms of airway remodeling in asthma. *Allergol Int* 56:331-340.
- Takata S, Matsubara M, Allen PG, Janmey PA, Serhan CN and Brady HR (1994) Remodeling of neutrophil phospholipids with 15(S)-hydroxyicosatetraenoic acid inhibits leukotriene B4-induced neutrophil migration across endothelium. *J Clin Invest* 93:499-508.
- Tattersfield AE, Knox AJ, Britton JR and Hall IP (2002) Asthma. *Lancet* 360:1313-1322.
- The Childhood Asthma Management Program Research Group (2000) Long-term effects of budesonide or nedocromil in children with asthma. *N Engl J Med* 343:1054-1063.
- Thorisson GA, Smith AV, Krishnan L and Stein LD (2005) The International HapMap Project Web site. *Genome Res* 15:1592-1593.
- Tian (2003) Lipoprotein(a) level and lipids in type 2 diabetic patients and their normoglycemic first-degree relatives in type 2 diabetic pedigrees. *Diabetes Res Clin Prat* 59:63-69.
- Tremblay K, Lemire M, Provost V, Pastinen P, Renaud Y, Sandford AJ, Laviolette M, Hudson TJ and Laprise C (2006) Association study between the CX3CR1 gene and asthma. *Genes Immun* 7:632-639.
- Tremblay K, Daley D, Chamberland A, Lemire M, Montpetit A, Laviolette M, Musk AW, James AL, Chan-Yeung M, Becker A, Kozyrskyj AL, Sandford AJ, Hudson TJ, Pare PD and Laprise C (2008a) Genetic variation in immune signaling genes differentially expressed in asthmatic lung tissues. *J Allergy Clin Immunol* 122:529-536.
- Tremblay K, Lemire M, Potvin C, Tremblay A, Hunninghake GM, Raby BA, Hudson TJ, Perez-Iratxeta C, Andrade-Navarro MA and Laprise C (2008b) Genes to diseases (G2D) computational method to identify asthma candidate genes. *PLoS ONE* 3:e2907.
- Tremblay K, Bossé Y, Hernandez C, Roquis D, Hudson TJ and Laprise C (2008c) Sequencing the filaggrin gene in atopic asthmatic subjects from the Saguenay-Lac-St-Jean founder population.[En préparation].
- Turk J, Maas RL, Brash AR, Roberts LJ, 2nd and Oates JA (1982) Arachidonic acid 15-lipoxygenase products from human eosinophils. *J Biol Chem* 257:7068-7076.

- Umehara H, Bloom ET, Okazaki T, Nagano Y, Yoshie O and Imai T (2004) Fractalkine in vascular biology: from basic research to clinical disease. *Arterioscler Thromb Vasc Biol* 24:34-40.
- Umemoto (2003) Serum levels of vasoactive intestinal peptide are elevated in patients with atopic dermatitis. *J Dermatol Sci* 31:161-164.
- Van Eerdewegh P, Little RD, Dupuis J, Del Mastro RG, Falls K, Simon J, Torrey D, Pandit S, McKenny J, Braunschweiger K, Walsh A, Liu Z, Hayward B, Folz C, Manning SP, Bawa A, Saracino L, Thackston M, Benchekroun Y, Capparell N, Wang M, Adair R, Feng Y, Dubois J, FitzGerald MG, Huang H, Gibson R, Allen KM, Pedan A, Danzig MR, Umland SP, Egan RW, Cuss FM, Rorke S, Clough JB, Holloway JW, Holgate ST and Keith TP (2002) Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* 25:426-430.
- van Haren EH, Lammers JW, Festen J, Heijerman HG, Groot CA and van Herwaarden CL (1995) The effects of the inhaled corticosteroid budesonide on lung function and bronchial hyperresponsiveness in adult patients with cystic fibrosis. *Respir Med* 89:209-214.
- Vercelli D (2008) Discovering susceptibility genes for asthma and allergy. *Nat Rev Immunol* 8:169-182.
- Walter MJ, Morton JD, Kajiwarra N, Agapov E and Holtzman MJ (2002) Viral induction of a chronic asthma phenotype and genetic segregation from the acute response. *J Clin Invest* Jul;110(2):165-75.
- Wang S and Zhao H (2003) Sample size needed to detect gene-gene interactions using association designs. *Am J Epidemiol* 158:899-914.
- Warner SM and Knight DA (2008) Airway modeling and remodeling in the pathogenesis of asthma. *Curr Opin Allergy Clin Immunol* 8:44-48.
- Weeks DE and Lathrop GM (1995) Polygenic disease: methods for mapping complex disease traits. *Trends Genet* 11:513-519.
- Williams SM, Canter JA, Crawford DC, Moore JH, Ritchie MD and Haines JL (2007) Problems with genome-wide association studies. *Science* 316:1840-1842.
- Wills-Karp M and Ewart SL (2004) Time to draw breath: asthma-susceptibility genes are identified. *Nat Rev Genet* 5:376-387.
- Wong MY, Day NE, Luan JA and Wareham NJ (2004) Estimation of magnitude in gene-environment interactions in the presence of measurement error. *Stat Med* 23:987-998.
- Woolcock AJ, Peat JK, Salome CM, Yan K, Anderson SD, Schoeffel RE, McCowage G and Killalea T (1987) Prevalence of bronchial hyperresponsiveness and asthma in a rural adult population. *Thorax* 42:361-368.



- Xu J, Turner A, Little J, Bleecker ER and Meyers DA (2002) Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? *Hum Genet* 111:573-574.
- Yamamoto (1999) ADAM family proteins in the immune system. *Immunol today* 20(6):278-284.
- Yoshikawa M, Nakajima T, Matsumoto K, Okada N, Tsukidate T, Iida M, Otori N, Haruna S, Moriyama H, Imai T and Saito H (2004) TNF-alpha and IL-4 regulate expression of fractalkine (CX3CL1) as a membrane-anchored proadhesive protein and soluble chemotactic peptide on human fibroblasts. *FEBS Lett* 561:105-110.
- Yu B (2009) Role of in silico tools in gene discovery. *Mol Biotechnol* 41:296-306.
- Zhang (2006) Toll-like receptor 4 deficiency causes pulmonary emphysema. *J Clin Invest* 116:3050-3059.
- Zhang J, Pare PD and Sandford AJ (2008) Recent advances in asthma genetics. *Respir Res* 9:4.
- Zhang Y, Leaves NI, Anderson GG, Ponting CP, Broxholme J, Holt R, Edser P, Bhattacharyya S, Dunham A, Adcock IM, Pulleyn L, Barnes PJ, Harper JJ, Abecasis G, Cardon L, White M, Burton J, Matthews L, Mott R, Ross M, Cox R, Moffatt MF and Cookson WO (2003) Positional cloning of a quantitative trait locus on chromosome 13q14 that influences immunoglobulin E levels and asthma. *Nat Genet* 34:181-186.
- Zhu X, Tang H and Risch N (2008) Admixture mapping and the role of population structure for localizing disease genes. *Adv Genet* 60:547-569.