



**Profils de risque génétiques de la sensibilisation aux aéroallergènes intérieurs et extérieurs
dans la cohorte familiale d'asthme du Saguenay–Lac-Saint-Jean**

Par Marie-Pier Bouchard

**Mémoire présenté à l'Université du Québec à Chicoutimi en vue de l'obtention du grade de
Maître ès sciences (M. Sc.) en Sciences cliniques et biomédicales**

Québec, Canada

© Marie-Pier Bouchard, 2024

RÉSUMÉ

La sensibilisation aux aéroallergènes représente un facteur de risque pour le développement de l'atopie. Il est aussi connu que les aéroallergènes d'origine intérieure (squames d'animaux, acariens, poussière) sont majoritairement associés à l'asthme allergique, tandis que les aéroallergènes d'origine extérieure (moisissure, pollens, mauvaise herbe) sont majoritairement associés à la rhinite allergique. La présente étude vise à identifier des variants génétiques communs et distincts entre la sensibilisation aux aéroallergènes d'intérieur et d'extérieur. Ceci a été possible en effectuant une étude d'association génomique (de l'anglais *Genome-wide association study* [GWAS]) pour la sensibilisation aux aéroallergènes d'intérieur, d'extérieur ou combinée dans la cohorte familiale d'asthme du Saguenay–Lac-Saint-Jean. Les variants associés pour chacun des types de sensibilisation ont ensuite été utilisés pour effectuer des analyses de locus quantitatif (de l'anglais *Quantitative trait locus* [QTL]) pour le niveau d'expression (eQTL) et de méthylation (mQTL) des gènes pour les lymphocytes T CD4⁺ naïfs et pour les éosinophiles provenant d'échantillons sanguins. Ainsi, les GWAS ont permis d'identifier plusieurs associations significatives ($p < 1 \times 10^{-5}$) entre des variants et les types de sensibilisations aux aéroallergènes (intérieur : 259 variants, extérieur : 38 variants et combiné : 57 variants). Ensuite, les analyses fonctionnelles (eQTL et mQTL) ont permis d'identifier des variants génétiques, préalablement associés aux trois types de sensibilisations, pour lesquels il y avait des résultats significatifs (FDR < 0,05) en cis-eQTL et en cis-mQTL pour les mêmes gènes, *RNASET2* et *MIR3939* pour la sensibilisation intérieure et *FGD3* et *CENPP* pour la sensibilisation combinée, chez les deux types de cellules immunitaires à l'étude. De plus, les analyses de corrélation ont permis de mettre en évidence une corrélation significative (FDR < 0,05) entre la différence d'expression et de méthylation de certains gènes chez les lymphocytes T CD4⁺ naïfs. Ainsi, cette étude a permis d'identifier un profil génétique spécifique à chacun des types de sensibilisation aux aéroallergènes et de cibler les voies biologiques impliquées dans deux types de cellules immunitaires.

TABLE DES MATIÈRES

RÉSUMÉ	ii
TABLE DES MATIÈRES.....	iii
LISTE DES TABLEAUX.....	vi
LISTE DES FIGURES	vii
LISTE DES ABRÉVIATIONS.....	viii
REMERCIEMENTS	xi
INTRODUCTION	1
CHAPITRE 1.....	4
IRRITANTS RESPIRATOIRES ET AÉROALLERGÈNES.....	4
1.1 Irritants respiratoires.....	4
1.1.1 Polluants atmosphériques.....	4
1.1.2 Aéroallergènes	5
1.2 Aéroallergènes d'intérieur.....	8
1.2.1 Animaux	9
1.2.2 Acariens	10
1.2.3 Moisissures	11
1.3 Aéroallergènes d'extérieur.....	13
1.3.1 Pollens	13
1.3.1.1 Arbres et arbustes.....	14
1.3.1.2 Graminées.....	14
1.3.1.3 Herbacées.....	15
1.3.2 Moisissures	15
CHAPITRE 2.....	17
HISTOIRE NATURELLE DE L'ASTHME ALLERGIQUE ET DE LA RHINITE ALLERGIQUE	17
2.1 Processus inflammatoire à la suite d'une exposition allergénique chez une personne sensibilisée	17
2.1.1 Sensibilisation allergique.....	17
2.2 L'asthme allergique	20
2.2.1 Épidémiologie de l'asthme allergique	20
2.2.2 Phénotype de l'asthme allergique.....	21
2.2.3 Inflammation bronchique dans l'asthme allergique	23
2.2.3.1 Réaction précoce	23
2.2.3.2 Réaction tardive	25
2.3 La rhinite allergique	27
2.3.1 Épidémiologie de la rhinite allergique	27

2.3.2 Phénotype de la rhinite allergique.....	28
2.3.3 Inflammation causée par la rhinite allergique	28
2.3.3.1 Réaction précoce	29
2.3.3.2 Réaction tardive	30
CHAPITRE 3.....	32
GÉNÉTIQUE ET ÉPIGÉNÉTIQUE DE L'ASTHME ALLERGIQUE ET DE LA RHINITE ALLERGIQUE.....	32
3.1 Trait complexe : définition.....	32
3.1.2 Facteurs génétiques.....	32
3.1.1 Facteurs environnementaux.....	33
3.2 Génétique de la sensibilisation allergique	34
3.2.1 Étude d'association pangénomique	34
3.2.2 Variants de susceptibilité	35
3.3 Épigénétique.....	38
3.3.1 Étude de locus à caractère quantitatif.....	40
CHAPITRE 4.....	42
GENETIC RISK PROFILING FOR INDOOR AND OUTDOOR AEROALLERGENS SENSITIZATION IN A FRENCH-CANADIAN ASTHMA FAMILY COHORT	42
4.1 Avant-propos	42
4.2 Résumé	44
4.3 Abstract	45
4.4 Introduction.....	46
4.5 Materials and Methods	47
4.5.1 Discovery sample.....	49
4.5.2 Replication sample.....	52
4.5.3 Atopic status.....	52
4.5.4 Genotyping and imputation	53
4.5.5 Association analyses.....	54
4.5.6 Replication analyses	54
4.5.7 eQTL and mQTL analyses	55
4.6 Results	56
4.7 Discussion	67
4.8 Conclusion.....	71
4.9 References	72
DISCUSSION	79
CONCLUSION.....	86
BIBLIOGRAPHIE	88

CERTIFICATION ÉTHIQUE	101
ANNEXE 1	102
Supplementary information.....	102
Supplementary Tables.....	107

LISTE DES TABLEAUX

Tableau 1. Prévalence de la sensibilisation allergénique à 15 aéroallergènes de six sites au Canada	6
Tableau 2. Exemple de composantes allergéniques provenant de <i>Dermatophagoides pteronyssinus</i>	8
Tableau 3. Fréquence de la sensibilisation à 11 aéroallergènes chez des individus atopiques de la ville d'Edmonton	10
Tableau 4. Loci, gènes et variants associés à la sensibilisation allergénique	37
Tableau 5. Exemple de la méthylation impliquée dans l'asthme et la rhinite.....	40
Table 6. Phenotypic characteristics of individuals included in the association analyses for the indoor, the outdoor and both types of aeroallergens sensitization	50
Table 7. 200kb regions replicated in the EGEA cohort and corresponding variant's information for the best association with indoor, outdoor and both types of aeroallergens sensitization	60
Table 8. GWAS and cis-QTL information on variants associated with significant eQTL and mQTL targeting the same gene for which there is a significant correlation between gene expression and methylation levels in naïve CD4 ⁺ T cells.....	66
Table 9. Phenotypic characteristics of individuals included in the analyses performed on naïve CD4 ⁺ T cells and eosinophils.....	107
Table 10. Complete characterization of allergens sensitization of individuals included in the association analyses for the indoor, the outdoor and both types of aeroallergens sensitization	109
Table 11. All significant genetic variants associated with the sensitization to indoor aeroallergens in the SLSJ cohort (P<1E-05) and association results for the EGEA cohort.....	112
Table 12. All significant genetic variants associated with the sensitization to outdoor aeroallergens in the SLSJ cohort (P<1E-05) and association results for the EGEA cohort	123
Table 13. All significant genetic variants associated to the sensitization to both types of aeroallergens in the SLSJ cohort (P<1E-05) and association results for the EGEA cohort	125
Table 14. Biological function of proximal genes of genetic variants identified by GWAS for the indoor sensitization	129
Table 15. Biological function of proximal genes of genetic variants identified by GWAS for the outdoor sensitization.....	133
Table 16. Biological function of proximal genes of genetic variants identified by GWAS for both types of sensitization	136
Table 17. Significant cis-eQTL results including genetic variants associated with each sensitization type in naïve CD4 ⁺ T cells and eosinophils	139
Table 18. Significant cis-mQTL results including genetic variants associated with the indoor sensitization in naïve CD4 ⁺ T cells and eosinophils	141
Table 19. Significant cis-mQTL results including genetic variants associated to the outdoor sensitization in naïve CD4 ⁺ T cells and eosinophils	146
Table 20. Significant cis-mQTL results including genetic variants associated with both types of sensitization in naïve CD4 ⁺ T cells and eosinophils	148
Table 21. Correlation between gene expression and methylation levels in CD4 ⁺ T cells in indoor and both types of sensitization	153
Table 22. Genes targets of <i>MIR3939</i> and their biological functions.....	155

LISTE DES FIGURES

Figure 1. Physiopathologie du développement de la sensibilisation à un aéroallergène	19
Figure 2. Les divers phénotypes de l'asthme regroupés en deux endotypes	22
Figure 3. Réaction inflammatoire précoce de l'asthme allergique	24
Figure 4. Réaction inflammatoire tardive de l'asthme allergique	26
Figure 5. Réaction inflammatoire précoce de la rhinite allergique	30
Figure 6. Réaction inflammatoire tardive de la rhinite allergique	31
Figure 7. Schematic representation of this study's layout.....	48
Figure 8. Manhattan plot of significant loci associated ($P < 1E-05$) with a) the indoor sensitization, b) outdoor sensitization, and c) both types of sensitization	58
Figure 9. Venn diagram of the shared loci between the sensitization to indoor, outdoor and both types of aeroallergens	59
Figure 10. Biological functions of genes identified by GWAS for the a) indoor, b) outdoor and c) both types of aeroallergens sensitization.....	64

LISTE DES ABRÉVIATIONS

AA : Asthme allergique
ABPA : Aspergillose bronchopulmonaire allergique
ADAD1 : *Adenosine deaminase domain containing 1*
C11orf30 : *Chromosome 11 open reading frame 30*
CAMK4 : *Calcium/calmodulin dependent protein kinase IV*
CD : Cellule dendritique
CDC42 : *Cell Division Cycle 42*
CDX1 : *Caudal Type Homeobox 1*
CENP-A : *Centromere Protein A*
CENP-P : *Centromere Protein P* (protéine)
CENPP : *Centromere Protein P* (gène)
CH3 : Groupement méthyle
CpG : Cytosine précédant une guanine
CRY2 : *Cryptochrome Circadian Regulator 2*
CYP26A1 : *Cytochrome P450 family 26 subfamily A member 1*
D. farinae : *Dermatophagoides farinae*
D. pteronyssinus : *Dermatophagoides pteronyssinus*
EGEA : Étude Génétique Environnement Asthme
ELK4 : *ETS Transcription Factor ELK4*
EMSY : *EMSY Transcriptional Repressor, BRCA2 Interacting*
eQTL : Locus de caractères quantitatifs d'expression
FASLG : *Fas ligand*
FcεRI : Récepteurs de haute affinité pour les IgE
FGD3 : *FYVE, RhoGEF and PH Domain Containing 3*
FOXA1 : *Forkhead box A1*
FOXE-3 : *Forkhead Box Protein E3*
FOXP3 : *Forkhead box P3*
GATA3 : *GATA binding protein 3*
GM-CSF : Facteur de croissance hématopoïétique
GRN : *Granulin Precursor*
GSDMB : *Gasdermin B*
GTE_x : *Genotype-Tissus Expression*
GWAS : Étude d'association pangénomique
HLA : Antigène des leucocytes humains
HLA-B : *Major Histocompatibility Complex, Class I, B*
HLA-C : *Major Histocompatibility Complex, Class I, C*
HLA-DQA1 : *Major Histocompatibility complex, Class II DQ Alpha 1*
HLA-DQB1 : *Major Histocompatibility complex, Class II DQ Beta 1*
ICAM1 : *Intercellular adhesion molecule 1*
Ig : Immunoglobuline
IgE : Immunoglobuline E
IgM : Immunoglobuline M
IL : Interleukine
IL-10 : Interleukine-10
IL-13 : Interleukine-13 (protéine)
IL-16 : Interleukine-16
IL-25 : Interleukine-25
IL-3 : Interleukine-3
IL-33 : Interleukine-33 (protéine)
IL-4 : Interleukine-4 (protéine)
IL-5 : Interleukine-5

IL-6 : Interleukine-6
 IL-8 : Interleukine-8
 IL-9 : Interleukine-9
IL13 : Interleukine 13 (gène)
IL18R1 : Interleukine 18 receptor 1
IL1RL1 : Interleukine 1 receptor like 1
IL1RL2 : Interleukine 1 receptor like 2
IL2 : Interleukine 2
IL33 : Interleukine 33 (gène)
IL4 : Interleukine 4 (gène)
 ILC : Cellule lymphoïde innée
 ILC2 : Cellule lymphoïde innée de type 2
 ILC3 : Cellule lymphoïde innée de type 3
 Indel : Insertion ou délétion
 IUIS : *International Union of Immunological Societies*
 LD : Déséquilibre de liaison
LINC00299 : Long intergenic non-protein coding RNA 299
LPP : LIM domain containing preferred translocation partner in lipoma
LRRC32 : Leucine Rich Repeat Containing 32
MEIS1 : Meis homeobox 1
 MHC : Complexe majeur d'histocompatibilité
 MIC : MHC Class I polypeptide-Related Sequence
MICA : MHC Class I polypeptide-Related Sequence A
MICB : MHC Class I polypeptide-Related Sequence B
MIR3939 : MicroRNA 3939
 mQTL : Locus de caractères quantitatifs de méthylation
MRPL4 : Mitochondrial ribosomal protein L4
MYC : MYC Proto-Oncogene, BHLH Transcription Factor
NFATC2 : Nuclear factor of activated T cells 2
NFKB1 : Nuclear factor kappa B subunit 1
 NKT : Lymphocyte T tueuse naturelle
NPC2 : NPC intracellular cholesterol transporter 2
ORMDL3 : *ORMDL Sphingolipid Biosynthesis regulator 3*
P. citrinum : *Penicillium citrinum*
PGD2 : Prostaglandine D2
PLCL1 : Phospholipase C like 1 (inactive)
 PR-10 : Pathogenesis-related protein de type 10
 PSA : Antigène prostatique spécifique
PTGER4 : Prostaglandin E receptor 4
PVT1 : Plasmacytoma Variant Translocation 1 Oncogene
 QTL : Locus de caractères quantitatifs
 RA : Rhinite allergique
RANBP6 : RAN binding protein 6
RNASET2 : Ribonuclease T2
RUNX3 : RUNX family transcription factor 3
SIK1 : Salt Inducible Kinase 1
SLC25A38P1 : Solute carrier family 25 member 38 pseudogene 1
SLC25A46 : Solute carrier family 25 member 46
 SLSJ : Saguenay–Lac-Saint-Jean
SMAD3 : SMAD Family Member 3
 SNP : Polymorphisme d'un seul nucléotide
STAT6 : Signal transducer and activator of transcription 6
 Th : Lymphocyte T
 Th0 : Lymphocyte T CD4⁺ naïf
 Th1 : Lymphocyte T auxiliaire de type 1
 Th17 : Lymphocyte T auxiliaire de type 17
 Th2 : Lymphocyte T auxiliaire de type 2

TIGIT : T cell immunoreceptor with Ig and ITIM domains
TLR : Toll Like Receptor
TLR1 : Toll like receptor 1
TLR10 : Toll Like Receptor 10
TLR6 : Toll Like Receptor 6
TMEM232 : transmembrane protein 232
TNF- α : Facteur de nécrose tumorale alpha
TSLP : Lymphopoïétine stromale thymique
TTC6 : Tetratricopeptide repeat domain 6
WDR36 : WD repeat domain 36
WHO : World Health Organization

REMERCIEMENTS

Je tiens à remercier ma directrice de projet, la Pre Catherine Laprise. Merci beaucoup de m'avoir offert cette opportunité de travailler dans ton laboratoire et de m'avoir permis d'acquérir de l'expertise sous ta supervision d'abord pendant mon baccalauréat puis durant mon cheminement à la maîtrise. Merci de m'avoir encouragé, motivé et de m'avoir donné tous les outils nécessaires pour conduire à la réussite incluant le soutien financier par le biais de ta Chaire de recherche du Canada en génomique de l'asthme et des maladies allergiques.

Je tiens également à remercier Anne-Marie Madore, une formidable professionnelle de recherche dans le laboratoire de la Pre Catherine Laprise. Merci d'être restée patiente et d'avoir répondu à toutes mes questions à développement qui parfois menaient à quelque chose que tu n'aurais jamais imaginé! Aussi, je te suis grandement reconnaissante de m'avoir transmis une passion pour la bio-informatique. Merci à Anne-Marie Boucher-Lafleur et Marie-Ève Lavoie, techniciennes de laboratoire également dans l'équipe de la Pre Catherine Laprise, pour votre aide, votre soutien et votre patience pour mes mille et une questions.

Je veux également remercier Rébecca Dionne-Gagné, étudiante à la maîtrise, pour ton amitié, ton aide, tes encouragements, les expériences de vie étudiante partagées et pour m'avoir fait rire tout au long de la maîtrise. Un gros merci à l'ensemble de l'équipe étudiante de la Pre Catherine Laprise pour tout ce que vous avez fait pour moi durant mon parcours. Chacun d'entre vous m'a offert leur amitié et leur soutien tout au long de la maîtrise en créant une ambiance de respect mutuel où il était agréable de se retrouver. J'ai adoré partager de merveilleux moments avec vous tous.

Finalement, je tiens à remercier mon fiancé et ma famille pour m'avoir appuyée, encouragée et même poussée à me dépasser tout au long de ma maîtrise. Sans vous, je ne serais sûrement pas là où je suis aujourd'hui. Merci à Hope et Abby Ouellet, mes deux humaines de soutien émotionnel pour avoir été deux rayons de soleil durant ma maîtrise.

INTRODUCTION

Au cours des dernières décennies, l'industrialisation, l'urbanisation et les changements climatiques ont mené à accroître la prévalence des maladies allergiques dans le monde (Eguiluz-Gracia et al., 2020). L'industrialisation est liée à un grand nombre de particules polluantes dans l'air qui sont responsables, du moins en partie, de l'augmentation de la production de pollens par les végétaux, augmentant aussi l'allergénicité des aéroallergènes (Sierra-Heredia et al., 2018). Le réchauffement climatique est associé à une augmentation de la concentration des aéroallergènes dans l'air, précisément en modifiant la période ainsi que la durée de la saison des pollens et de leur diversité. La combinaison des impacts du réchauffement climatique et de l'industrialisation fait en sorte d'exacerber les symptômes liés aux allergies respiratoires et d'augmenter le nombre de visites à l'urgence auprès des populations à risque (D'Amato et al., 2020; D'Amato et al., 2015; Sierra-Heredia et al., 2018; Tham et al., 2016).

Par ailleurs, bien que les allergies respiratoires aient une composante environnementale, celles-ci possèdent aussi une composante génétique. Ainsi, une sensibilisation à un ou des aéroallergène(s) serait un facteur de risque pour le développement de l'asthme allergique (AA) et/ou de la rhinite allergique (RA) chez une personne ayant une prédisposition génétique (Sierra-Heredia et al., 2018). Des études ont montré que l'AA était plus associé à une sensibilisation allergénique aux aéroallergènes de type intérieur, comme les squames d'animaux, les acariens et la poussière, tandis que la RA était associée à une sensibilisation aux aéroallergènes de type extérieur, comme le pollen, les mauvaises herbes et la moisissure (Boulet et al., 1997; J. Li et al., 2009). Cette observation s'explique en partie par la nature physicochimique des particules allergéniques, mais elle suggère également qu'il existe une composante génétique favorisant le développement d'un type spécifique de sensibilisation et des manifestations cliniques associées (Stewart & Robinson, 2017). C'est dans cette optique que des études ont tenté d'identifier les déterminants génétiques d'une ou plusieurs sensibilisations allergéniques indépendamment du type, ciblant un type ou pour une sensibilisation combinée (aussi nommée polysensibilisation; Amaral et al., 2015; Bønnelykke et al., 2013; M. A. R.

Ferreira et al., 2011; Ramasamy et al., 2011). Cependant, il y a peu d'études qui ont essayé d'identifier des variants génétiques ou des loci prédisposant à une sensibilisation allergique spécifique à ce jour (Bønnelykke et al., 2015).

Les études d'associations pangénomiques (GWAS) sont fréquemment utilisées en génétique pour l'étude des traits complexes. Elles permettent notamment d'identifier des variants génétiques communs, majoritairement des polymorphismes d'un seul nucléotide (SNP), mais également de petites insertions ou délétions (indels), qui sont présents en plus ou moins grande fréquence chez les personnes atteintes du phénotype étudié contrairement à celles sans ce phénotype (Cano-Gamez & Trynka, 2020). Un GWAS est efficace pour identifier de nombreux variants ayant des impacts variés sur le phénotype, mais ils identifieront plus fréquemment des variants qui ont un impact mineur sur le phénotype (Ober & Yao, 2011). En effet, plusieurs variants se retrouvent dans des régions non codantes, nécessitant par la suite une investigation plus poussée sur ces variants afin de comprendre les gènes qu'ils régulent ou dans quel type cellulaire ils agissent (Liu & Montgomery, 2020). Par conséquent, ce type d'étude est souvent suivi d'une étude de locus de caractères quantitatifs (QTL) pour le niveau d'expression (eQTL) et le niveau de méthylation (mQTL). Par la suite, des analyses d'intégration, incluant notamment des analyses d'enrichissement, pour comprendre la nature des liens moléculaires peuvent être réalisées. Ces analyses combinent les données épigénétiques ou transcriptomiques à celles de la génétique afin d'associer une différence de méthylation ou d'expression d'un gène à un variant dans un tissu ou type cellulaire précis (Cano-Gamez & Trynka, 2020; Nica & Dermitzakis, 2013). Par conséquent, il est possible de mieux comprendre comment le variant peut contribuer au développement du phénotype. L'identification de variants associés au phénotype et la compréhension de leur implication dans le mécanisme pathophysiologique de l'allergie rendront possible l'utilisation des résultats en médecine translationnelle. Par exemple, il sera possible d'identifier les individus ayant un risque accru de développer le phénotype et d'entreprendre des démarches de prévention et d'intervention ciblées ou encore de peaufiner les approches thérapeutiques utilisées (B. Li & Ritchie, 2021).

Les objectifs de cette étude étaient donc :

- i) D'identifier des profils de risque génétique différents selon le type d'aéroallergène auquel la personne est sensibilisée, soit aux aéroallergènes d'intérieur, d'extérieur et combinés ;

- ii) D'identifier de possibles voies biologiques impliquées dans la sensibilisation aux aéroallergènes d'intérieur, d'extérieur et combinés pour deux types de cellules immunitaires.

CHAPITRE 1

IRRITANTS RESPIRATOIRES ET AÉROALLERGÈNES

1.1 Irritants respiratoires

Il existe plusieurs types d'irritants respiratoires pouvant avoir des effets néfastes sur la santé respiratoire d'un individu comme les produits chimiques, le climat (vague de chaleur, sécheresse, inondation, etc.) ou encore la température (D'Amato et al., 2015). Cependant, ceux provenant de la pollution et des aéroallergènes ont un impact considérable sur les maladies allergiques respiratoires, telles que l'AA et la RA (Cecchi et al., 2018). Par ailleurs, la concentration de polluants atmosphériques et d'aéroallergènes à laquelle un individu est exposé varie beaucoup selon le pays, la région, que ce soit entre les zones rurales (milieux agricoles) ou urbaines (proximité des usines, densité de la population), selon la saison de l'année, la période de la journée et entre l'intérieur et l'extérieur d'une habitation (humidité et température; Cecchi et al., 2018). Notamment, les polluants peuvent agir seuls ou en combinaison avec les aéroallergènes pour mener à une réaction inflammatoire exacerbée, tout comme pour les aéroallergènes seuls. Par ailleurs, les polluants peuvent aussi influencer la production d'aéroallergènes et leur allergénicité en interagissant avec leur source d'origine (ex. plantes; D'Amato et al., 2020).

1.1.1 Polluants atmosphériques

Il existe plusieurs polluants atmosphériques comme les matières particulaires (hydrocarbures aromatiques polycycliques, carbone noir, composés organiques volatils), le dioxyde de carbone, le méthane et le dioxyde d'azote (Brandt et al., 2015; Cecchi et al., 2018). Ces polluants peuvent se retrouver autant à l'intérieur de la maison qu'à l'extérieur, et peuvent provenir de plusieurs sources comme les usines, les moyens de transport, la production agricole et alimentaire, les sources de chauffage, les éruptions volcaniques et les feux de forêt (Annesi-Maesano et al., 2021; Cecchi et al., 2018). Les effets des polluants sur le système respiratoire sont variés, tels que la diminution de l'activité des cellules ciliées de l'épithélium, l'augmentation de la perméabilité de la barrière épithéliale

(facilitation de la pénétrance des aéroallergènes), des modifications dans la réaction inflammatoire des cellules, la régulation du cycle cellulaire et de la mort cellulaire (création de lésions dans le tissu), le remodelage et la favorisation de la sensibilisation allergénique (Cecchi et al., 2018; D'Amato et al., 2015).

1.1.2 Aéroallergènes

Un aéroallergène est une substance normalement inoffensive qui peut se retrouver en suspension dans l'air et qui peut provoquer une réaction inflammatoire une fois inhalé par des individus sensibilisés à celui-ci (i.e. allergique à l'allergène, défini par la présence de symptômes respiratoires et un test d'allergie cutané avec une induration de plus de 3mm). Ces allergènes sont normalement divisés en deux groupes, soit ceux que l'on retrouve à l'intérieur des bâtiments et ceux que l'on retrouve à l'extérieur des bâtiments. Plus d'informations sur les divers aéroallergènes composant ces deux catégories se trouveront dans les sections suivantes de ce chapitre. Il est intéressant de savoir que le profil de sensibilisation aux aéroallergènes varie beaucoup selon l'endroit, le climat et le degré d'urbanisation d'où l'on se situe dans le monde puisque le couvert végétal et les espèces animales auxquels les personnes sont exposées ne seront pas les mêmes (Tham et al., 2016). Ceci est principalement causé par le fait que la concentration et la diversité en aéroallergènes soient grandement influencées par la température. Par exemple, il y aura une plus grande concentration d'acariens dans les endroits chauds et humides (Acevedo et al., 2019; Tham et al., 2016). De plus, les espèces produisant les aéroallergènes (acariens, plantes, moisissures, etc.) varient entre les endroits étant donné que certains d'entre eux ont des températures plus propices à la prolifération d'une espèce au détriment d'une autre (Acevedo et al., 2019). Par conséquent, il existe des variations dans la prévalence d'individus sensibilisés à différents aéroallergènes selon la région géographique (Tableau 1; P.-J. Bousquet et al., 2007). Ce phénomène peut être causé par plusieurs facteurs comme le climat, la localisation (zones rurales ou urbaines), la diversité génétique d'une population (i.e. la fréquence des polymorphismes de susceptibilité

associés), des facteurs socio-économiques et le contact avec des animaux à la maison pour n'en citer que quelques exemples (Lei & Grammer, 2019; Miller, 2019; Tham et al., 2016).

Tableau 1. Prévalence de la sensibilisation allergénique à 15 aéroallergènes de six sites au Canada (Traduit et adapté de Chan-Yeung et al., 2010)

	Van	Win	Ham	Mon	Hal	ÎPÉ	Prévalence totale
Taille de l'échantillon	417	540	412	471	232	523	2595
Chats	18	22	20	22	26	17	21
Acariens*	43	43	45	46	47	41	44
<i>Dermatophagoides pteronyssinus*</i>	39	38	38	43	43	35	39
<i>Dermatophagoides farinae</i>	38	33	38	41	41	33	37
Moisissures confondues	15	20	16	17	18	10	16
<i>Aspergillus</i>	1	4	2	1	1	1	2
<i>Penicillium*</i>	3	4	3	5	2	2	3
<i>Alternaria alternata*</i>	10	13	12	11	9	7	10
<i>Cladosporium herbarum</i>	5	10	8	8	12	5	8
Arbres	26	18	21	27	20	13	21
Arbres mixtes	21	14	17	21	16	8	16
Bouleau	23	10	16	22	14	9	15
Olivier*	3	4	2	5	5	4	4
Graminées	30	19	28	28	24	14	23
Fléole des prés	29	17	28	26	21	13	22
Pâturin des prés	26	15	23	24	21	11	20
Petite herbe à poux	12	7	28	33	5	5	15

Toutes les prévalences indiquées dans le tableau sont en pourcentage. *Aucune différence significative entre les sites étudiés. Abréviations : Hal = Halifax, Ham = Hamilton, ÎPÉ = Île-du-Prince-Édouard, Mon = Montréal, Van = Vancouver, Win = Winnipeg.

Les sources (ex. plantes, animaux, etc.) desquelles proviennent les aéroallergènes possèdent plusieurs protéines qui forment les composantes allergéniques. La composition protéique de l'aéroallergène est spécifique à l'espèce qui le produit (Sastre, 2010). Par exemple, les *Dermatophagoides (D) farinae* possèdent 37 composantes allergéniques reconnues, alors que les *D. pteronyssinus* en possèdent 32 selon la nomenclature des allergènes produite par la *World Health Organization (WHO)/ International Union of Immunological Societies (IUIS; World Health Organization/ International Union of Immunological Societies, 2007)*. Il est à noter que ce ne sont pas toutes les composantes allergéniques d'un aéroallergène qui sont forcément reconnues par les anticorps. Effectivement, un aéroallergène comporte des composantes majeures (Ex. Der p 1 chez *D. pteronyssinus*), c'est-à-dire que plus de 50 % des individus sensibilisés réagissent à celles-ci, et des composantes mineures (Ex. Der p 10 chez *D. pteronyssinus*), où c'est moins de 50 % des individus sensibilisés qui réagissent à ces dernières (Tableau 2; García-Mozo, 2017; World Health Organization/ International Union of Immunological Societies, 2007). Par contre, certains épitopes sont spécifiques à la famille et au genre des organismes, faisant en sorte que les composantes allergéniques des aéroallergènes de la même espèce et/ou du même genre seront reconnues par des anticorps communs (Tableau 2; F. Ferreira et al., 2004; Sastre, 2010). De plus, il existe des composantes structurellement similaires ou des séquences similaires entre des composantes allergéniques ou protéines de différentes espèces. Celles-ci peuvent donc être reconnues par le même anticorps, permettant ainsi une réaction croisée, c'est-à-dire la reconnaissance de deux protéines ou aéroallergènes différents par un même anticorps (Tableau 2; Sastre, 2010). Ce type de réaction croisée peut même survenir entre une composante d'un aéroallergène et d'une protéine produite par l'homme. Par exemple, l'une des protéines allergéniques de l'aéroallergène du chien peut avoir des réactions croisées avec l'antigène prostatique humain (PSA). Un autre exemple est la cyclophiline produite par l'*Aspergillus fumigatus* qui peut entraîner des réactions croisées avec la cyclophiline produite chez l'humain (Curin & Hilger, 2017; Matricardi et al., 2016; Twaroch et al., 2015).

Tableau 2. Exemple de composantes allergéniques provenant de *Dermatophagoides pteronyssinus* (Adapté de Matricardi et al., 2016; Popescu, 2015; Sastre, 2010; World Health Organization/ International Union of Immunological Societies, 2007)

Allergène	Prévalence de sensibilisation	Type de composante	Type de molécule	Réaction croisée
Der p 1	75-92%	Majeure	Cystéine protéase	Der f 1, Eur m 1
Der p 2	71-100%	Majeure	Protéine de la famille des NPC2	Der f 2, Eur m 2
Der p 9*	92%	Majeure	Sérine protéase collagénolytique	–
Der p 10	6%	Mineure	Tropomyosine	Bla g 7, Blo t 10, Der f 10, Pen a 1

*La composante allergénique n'est pas produite par une autre espèce d'acarien que celle-ci. Abréviations : Bla g = *Blattella germanica* (coquerelle germanique), Blo t = *Blomia tropicalis* (acarien), Der f = *Dermatophagoides farinae* (acarien), Der p = *Dermatophagoides pteronyssinus* (acarien), Eur m = *Euroglyphus maynei* (acarien), Pen a = *Penaeus aztecus* (crevette).

1.2 Aéroallergènes d'intérieur

Il est estimé qu'un individu passe en moyenne 80 % de son temps à l'intérieur d'un bâtiment, que ce soit au travail, à l'école ou à la maison (Eguiluz-Gracia et al., 2020). Par conséquent, il est exposé pendant de très longues périodes à l'environnement intérieur, ce qui, s'il est génétiquement prédisposé, le met plus à risque de développer une sensibilisation allergénique aux aéroallergènes d'intérieur (Sierra-Heredia et al., 2018). Le groupe des aéroallergènes d'intérieur est composé d'animaux, d'acariens et de certaines moisissures (Gray-Ffrench et al., 2022). Ces sources d'aéroallergènes, mentionnées précédemment et qui seront abordées dans cette section, sont celles les plus courantes. Cependant, il en existe d'autres, moins courantes comme les oiseaux, les lapins et les sauterelles qui peuvent occasionner des problèmes respiratoires (Gray-Ffrench et al., 2022; Stewart & Robinson, 2017). De plus, le degré d'exposition aux aéroallergènes d'intérieur est variable et dépend de facteurs comme la présence d'animaux, la localisation, le type d'habitation et le statut socio-économique entre autres (Akar-Ghinril & Phipatanakul, 2020). La sensibilisation aux aéroallergènes d'intérieur est un facteur de risque pour le développement de l'AA (Boulet et al., 1997; J. Li et al., 2009).

1.2.1 Animaux

Les animaux produisant des aéroallergènes sont principalement les animaux à poils, tels que les chats, les chiens et les souris (Stewart & Robinson, 2017). Les aéroallergènes provenant des animaux sont généralement produits par le foie et les glandes sécrétoires de ces derniers et se retrouvent sur la peau et le poil, dans la salive, dans l'urine et dans le sang (Pomés et al., 2016). Ces aéroallergènes se retrouvent en suspension dans l'air, ce qui permet leur inhalation (Akar-Ghinril & Phipatanakul, 2020; Zahradnik & Raulf, 2014). Les aéroallergènes peuvent être en suspension dans l'air pendant plusieurs minutes, et ils peuvent rester à l'intérieur d'une maison longtemps après la disparition de l'animal (Gray-Ffrench et al., 2022; Lei & Grammer, 2019).

Le chat est l'un des aéroallergènes ayant la plus haute prévalence allergénique dans la population humaine (Tableau 3; Ahmed et al., 2019). La sensibilisation aux animaux est un facteur de risque pour le développement de l'asthme (Akar-Ghinril & Phipatanakul, 2020). Environ 67 % des enfants ayant de l'asthme sont sensibilisés aux chats ou aux chiens (Singh & Hays, 2016). De plus, l'exposition à des concentrations élevées d'aéroallergènes de provenance animale est associée à un mauvais contrôle de l'asthme et à une réduction de la fonction respiratoire chez les personnes sensibilisées à ces aéroallergènes (Gray-Ffrench et al., 2022; Singh & Hays, 2016).

Tableau 3. Fréquence de la sensibilisation à 11 aéroallergènes chez des individus atopiques de la ville d'Edmonton (Adapté de Ahmed et al., 2019)

Aéroallergènes	Fréquence¹
Chat	53
Chien	17
<i>Dermatophagoides</i> ²	50
<i>Aspergillus</i>	10
<i>Penicillium</i>	6
<i>Alternaria</i>	24
<i>Cladosporium cladosporioides</i>	8
Bouleau	24
Peuplier	12
Cèdre	10
Fléole des prés	39

¹Les Fréquences indiquées correspondent au pourcentage d'individu atopique ayant le type de sensibilisation indiqué. ²Les deux espèces d'acariens testées sont *Dermatophagoides pteronyssinus* et *Dermatophagoides farinae*.

1.2.2 Acariens

Les acariens de maison ou de poussière font partie de la classe des *Arachnida*. Ils sont très petits, d'une longueur entre 250 et 350 µm, et se retrouvent dans des milieux humides à l'intérieur des bâtiments (Miller, 2019; Seguel et al., 2016). Ils se nourrissent principalement de kératine (peaux mortes, ongles, cheveux), de cellulose (tissus) et de chitine (champignon, insecte), et l'humidité d'un milieu leur permet de répondre à leurs besoins en eau (Lei & Grammer, 2019; Tham et al., 2016). Par conséquent, ils résident dans les tissus et les endroits dans lesquels il peut y avoir des résidus de peau morte (matelas, couvertures, vêtements; Seguel et al., 2016). Ce n'est pas l'acarien qui occasionne les réactions allergéniques chez les individus sensibilisés, mais plutôt ses excréments. Les substances fécales de ces insectes contiennent diverses composantes allergéniques composées d'enzymes digestives et d'une membrane chitineuse péritrophique (Seguel et al., 2016; Stewart &

Robinson, 2017). Ces déjections ont un diamètre de 10 à 40 µm et se retrouvent en suspension dans l'air (Pomés et al., 2016).

Il existe plusieurs espèces d'acariens dont la répartition géographique dépend de la température et de l'humidité. Par exemple, *Blomia tropicalis* est plus présent dans les endroits chauds comme les régions tropicales et subtropicales, alors que *D. pteronyssinus* et *D. farinae* sont capables de survivre dans des endroits froids et humides. Ceux-ci se retrouvent dans toutes les maisons et sont par conséquent les deux espèces d'acariens les plus communes (Acevedo et al., 2019; Pomés et al., 2016).

En 2019, une étude réalisée sur la population d'Edmonton au Canada a évalué que 50% des individus atopiques étaient sensibilisés aux acariens, plaçant les acariens en deuxième position après *Felis domesticus* parmi les 11 aéroallergènes étudiés (Tableau 3; Ahmed et al., 2019). Cependant, dans certaines populations ou régions, cet aéroallergène est celui pour lequel les gens sont le plus sensibilisés. Par exemple, lors de l'étude sur la sensibilisation allergénique réalisée par Chan-Yeung et al., la sensibilisation avec la prévalence la plus élevée pour chacune des six villes canadiennes était celle pour les acariens (Tableau 1). La sensibilisation à *D. pteronyssinus* et *D. farinae* peut toucher de 80 à 90 % des individus atteints d'allergies respiratoires (Acevedo et al., 2019; Tham et al., 2016).

1.2.3 Moisissures

Les champignons qui forment la moisissure sont des organismes eucaryotes sous le règne des *Fungi* (Seguel et al., 2016). Ceux-ci peuvent se retrouver à l'extérieur comme à l'intérieur des habitations puisqu'ils peuvent être transportés par les animaux de compagnie, les vêtements, ou encore être aéroportés (Gray-French et al., 2022). Les composantes allergéniques des champignons se trouvent dans les spores et dans les hyphes de ceux-ci (Stewart & Robinson, 2017). Le diamètre de ces aéroallergènes peut varier entre 2 et 50 µm, mais la majorité possèdent une taille entre 2 et

10 µm, ce qui leur permet de se rendre dans les voies respiratoires inférieures, lorsqu'inhalés (Burge & Rogers, 2000).

Les genres de moisissures les plus fréquemment retrouvés à l'intérieur des habitations sont *Aspergillus* et *Penicillium* (Gray-Ffrench et al., 2022; Lei & Grammer, 2019). Dans le genre *Aspergillus*, c'est *Aspergillus fumigatus* qui est l'espèce la plus étudiée, alors que dans le genre *Penicillium*, ce sont *Penicillium (P.) chrysogenum* et *P. citrinum* (Pomés et al., 2016).

La prévalence de la sensibilisation allergénique pour l'ensemble des moisissures (intérieures et extérieures) est de 16 %. Pour *Aspergillus* et *Penicillium*, cette prévalence varie entre 2 et 10 % et entre 3 et 6 % respectivement au Canada (Tableau 1, Tableau 3; Ahmed et al., 2019; Chan-Yeung et al., 2010). Certaines moisissures, comme *Aspergillus* et *Penicillium*, sont capables de coloniser les voies respiratoires. Par conséquent, le champignon permettra une stimulation allergénique continue par le biais des spores et hyphes créés dans les voies respiratoires et maintiendra la réaction allergique chez les individus sensibilisés à ceux-ci (Lombardi et al., 2017). Un exemple de ce type d'impact est l'aspergillose bronchopulmonaire allergique (ABPA) qui peut se développer chez les individus asthmatiques allergiques ou atteints de fibrose kystique et qui présentent également de l'atopie. Chez ces individus, il y aura plus de mucus produit et moins de clairance pulmonaire, ce qui contribue à créer un milieu optimal pour la croissance d'*Aspergillus*. Par conséquent, il y aura maintien de la réaction inflammatoire, menant à l'exacerbation de leur maladie et à la dégradation de leur fonction respiratoire (Patterson & Strek, 2010). La colonisation des voies respiratoires par les moisissures peut aussi mener à des sinusites allergiques fongiques (Lei & Grammer, 2019). De plus, une forte exposition à la moisissure est associée à l'augmentation du risque de développer de l'asthme et à l'accroissement de sa morbidité chez les individus asthmatiques (Akar-Ghinril & Phipatanakul, 2020; Pomés et al., 2016). Il existe plusieurs facteurs de risques prédisposant à un haut degré d'exposition aux aéroallergènes provenant de la moisissure. Parmi ceux-ci, on retrouve la chaleur, la durée d'ouverture des fenêtres d'une maison pour l'aération, les dégâts d'eau et le taux d'humidité (Akar-Ghinril & Phipatanakul, 2020; Gray-Ffrench et al., 2022).

1.3 Aéroallergènes d'extérieur

Les changements climatiques occasionnés par l'activité humaine ont contribué à l'augmentation de l'exposition aux aéroallergènes d'extérieur, tels que le pollen et la moisissure (D'Amato et al., 2020). Ces changements climatiques entraînent notamment une augmentation de la concentration de pollens et de spores dans l'air et une prolongation de leur présence (D'Amato et al., 2015). Par conséquent, cette forte exposition accroît les risques de développer une sensibilisation aux aéroallergènes d'extérieur. Ce groupe d'aéroallergènes comprend le pollen des arbres, des graminées et des mauvaises herbes, ainsi que les moisissures (Gray-French et al., 2022). Cette catégorie d'allergènes n'est pas présente toute l'année pour un même endroit, avec une concentration variable selon le climat (température, chaleur, pluie, inondation, etc.; Stewart & Robinson, 2017). De plus, la sensibilisation aux aéroallergènes d'extérieur est un facteur de risque pour le développement de la RA (Boulet et al., 1997; J. Li et al., 2009).

1.3.1 Pollens

Le pollen est un aéroallergène produit par les végétaux afin de se reproduire et celui-ci est aéroporté. Cette dissémination est influencée par divers facteurs comme le vent, l'humidité, les précipitations et la température. La majorité des pollens ont un diamètre situé entre 20 et 60 μm , faisant en sorte que ceux-ci restent coincés dans les voies respiratoires supérieures (Hughes et al., 2022; Lei & Grammer, 2019). Par contre, en présence d'une forte humidité dans l'air, il est possible que le grain de pollen se fragmente par choc osmotique et qu'il relâche des particules de dimension inférieure à 5 μm dans l'air. Une fois inhalées, ces petites particules peuvent atteindre les voies respiratoires inférieures et provoquer de l'asthme chez les individus à risque (Hughes et al., 2022; Singh & Hays, 2016). Dans des situations particulières, la dimension des particules fragmentées peut être inférieure à 3 μm et celles-ci pourront même atteindre les alvéoles. Lorsque les grains de pollen sont au niveau de la muqueuse nasale, ceux-ci vont s'hydrater, ce qui occasionne le relâchement de composantes allergéniques et de protéines dans le milieu (Stewart & Robinson, 2017).

1.3.1.1 Arbres et arbustes

Au Québec, le printemps (de mars à juin au Québec) est la saison pollinique des arbres. Les aéroallergènes produits par les arbres ont un diamètre allant de 20 à 60 μm (Demers, 2013; Gouvernement du Québec ; Lei & Grammer, 2019). Il existe plusieurs sortes de pollens provenant des arbres, mais ceux du genre des bouleaux (bouleau à feuilles de peuplier, bouleau jaune et bouleau papier) font partie de ceux causant le plus d'allergies respiratoires au Québec et sont la cause principale de la rhinite allergique saisonnière dans l'hémisphère nord durant cette saison (Demers, 2013; Gray-Ffrench et al., 2022; Matricardi et al., 2016; Pablos et al., 2016). La prévalence de la sensibilisation allergénique aux pollens d'arbre est de 21 % dans la population canadienne et le pollen provenant des bouleaux est celui dont le nombre d'individus y étant sensibilisés est le plus élevé parmi les divers pollens d'arbre au Canada (Tableau 1, Tableau 3; Ahmed et al., 2019; Chan-Yeung et al., 2010).

1.3.1.2 Graminées

La saison pollinique des graminées s'étend de mai à octobre au Québec et les aéroallergènes produits pendant cette période ont un diamètre situé entre 20 et 25 μm (Demers, 2013; Gouvernement du Québec ; Lei & Grammer, 2019). Les deux espèces de graminées qui causent le plus d'allergie au Québec sont la fléole des prés et l'ivraie vivace (Demers, 2013; Gray-Ffrench et al., 2022; Lei & Grammer, 2019). Il est intéressant de savoir qu'une protéine, spécifiquement du groupe 7, des composantes allergéniques de la fléole des prés a déjà été associée à des symptômes plus sévères de l'allergie, à une prévalence accrue d'asthme et à une susceptibilité de développer une maladie allergique (Matricardi et al., 2016). La prévalence de la sensibilisation aux pollens de graminées dans la population canadienne varie entre 23 et 39 %, ce qui lui confère la prévalence la plus élevée pour les pollens, tous types confondus (Tableau 1, Tableau 3; Ahmed et al., 2019; Chan-Yeung et al., 2010). La sensibilisation aux aéroallergènes provenant des graminées est associée au développement d'allergie respiratoire, mais ce type de sensibilisation joue aussi un rôle dans la pathogénèse de la dermite atopique (Singh & Hays, 2016).

1.3.1.3 Herbacées

Au Québec, la saison pollinique des herbacées a lieu du début juillet jusqu'en octobre et les aéroallergènes produits ont un diamètre situé entre 16 et 20 μm (Demers, 2013; Gouvernement du Québec ; Lei & Grammer, 2019). Le genre d'herbacée qui cause le plus de sensibilisation allergénique en Amérique du Nord et au Québec est l'herbe à poux et l'espèce la plus commune dans ce genre est la petite herbe à poux (Demers, 2013; Gadermaier et al., 2004; Gray-Ffrench et al., 2022; Matricardi et al., 2016). Un autre herbacé allergénique important est l'armoise commune (Gadermaier et al., 2004). La prévalence de la sensibilisation allergénique aux pollens de la petite herbe à poux est de 15 % dans la population canadienne (Tableau 1; Chan-Yeung et al., 2010). Quant à l'armoise commune, elle possède une prévalence de sensibilisation allergénique située entre 10 et 14 % chez les individus atteints d'allergies aux pollens (Gadermaier et al., 2004). Ainsi, la sensibilisation aux pollens d'herbacées est la moins prévalente parmi les trois types de sensibilisation aux pollens (Tableau 1; Chan-Yeung et al., 2010).

1.3.2 Moisissures

Comme mentionné précédemment, la moisissure est formée de champignons du règne des *Fungi*. Elle peut se retrouver autant à l'intérieur qu'à l'extérieur des habitations et ses composantes allergéniques sont contenues dans les spores et les hyphes (Gray-Ffrench et al., 2022; Seguel et al., 2016; Stewart & Robinson, 2017). Le diamètre de ces aéroallergènes peut varier entre 2 et 50 μm mais généralement, ils possèdent une taille entre 2 et 10 μm (Burge & Rogers, 2000). La concentration en aéroallergènes dans l'air des moisissures extérieures est très élevée et peut même dépasser de 100 à 1 000 fois la concentration en pollen (Abel-Fernández et al., 2023). La concentration de la moisissure extérieure atteint son maximum durant la moitié de l'été où les conditions sont favorables alors qu'elle diminue pendant l'hiver (Gray-Ffrench et al., 2022). Il existe plusieurs genres et espèces de moisissures, mais celles qui sont les plus abondantes à l'extérieur sont *Alternaria alternata* et *Cladosporium herbarum* (Lei & Grammer, 2019). Le genre *Alternaria* est plus abondant dans les régions au climat sec et chaud et se situe dans le sol, les graines ou les

plantes, alors que *Cladosporium* est plus abondant dans les régions au climat tempéré et se situe dans les plantes et les matières végétales (Gray-Ffrench et al., 2022).

L'allergie à la moisissure extérieure est souvent confondue avec celle au pollen étant donné que les symptômes sont similaires et que le pic de la concentration en spores dans l'air atteint son maximum pendant la saison du pollen des graminées et des mauvaises herbes (Abel-Fernández et al., 2023; Twaroch et al., 2015). Dans la population canadienne, la prévalence pour la sensibilisation allergénique à *Alternaria alternata* et à *Cladosporium herbarum* est de 10 à 24 % et 8 % respectivement (Tableau 1, Tableau 3; Ahmed et al., 2019; Chan-Yeung et al., 2010). L'exposition aux moisissures peut entraîner le développement ou l'exacerbation de l'AA, de la dermatite atopique et de la RA en plus d'une diminution de la fonction respiratoire et de l'augmentation de l'inflammation des voies respiratoires (Abel-Fernández et al., 2023; Hughes et al., 2022; Sánchez et al., 2022). De plus, une exposition à *Alternaria* et/ou à *Cladosporium* serait associée aux visites aux urgences chez les personnes asthmatiques allergiques (Abel-Fernández et al., 2023). Il est à noter qu'une sensibilisation à une espèce de moisissure est associée à un risque élevé de développer des symptômes d'allergies respiratoires tandis qu'une sensibilisation à plusieurs espèces de moisissure est associée à un moins bon contrôle de l'asthme (Sánchez et al., 2022).

CHAPITRE 2

HISTOIRE NATURELLE DE L'ASTHME ALLERGIQUE ET DE LA RHINITE ALLERGIQUE

2.1 Processus inflammatoire à la suite d'une exposition allergénique chez une personne sensibilisée

Le processus inflammatoire associé à une réaction allergique à la suite du contact avec un allergène se divise en deux phases dont la première est la phase de sensibilisation à l'allergène et la deuxième est la phase effectrice où il y a une réaction inflammatoire qui est initiée (Komlósi et al., 2022). Bien que l'AA soit une inflammation des voies respiratoires inférieures et la RA soit une inflammation des voies respiratoires supérieures, tous deux ont la même phase de sensibilisation, mais une phase effectrice qui est spécifique (Jean Bousquet et al., 2001). Le processus de sensibilisation sera d'abord décrit, puis l'inflammation spécifique à chaque condition (AA et RA) sera ensuite détaillée.

2.1.1 Sensibilisation allergique

La première étape initiant le processus physiopathologique de l'AA et de la RA est la sensibilisation allergénique (Figure 2). La sensibilisation allergénique à un aéroallergène surviendra généralement durant l'enfance à travers un épithélium défectueux. En effet, l'épithélium des individus atteints de dermite atopique sera une moins bonne barrière face aux particules externes comme les aéroallergènes (Czarnowicki et al., 2017). Par conséquent, ceux-ci pourront plus facilement traverser la peau grâce à un relâchement des cellules épithéliales causé par la dermite atopique. Cette pénétration de l'aéroallergène dans l'organisme à travers la barrière épithéliale occasionnera la sécrétion d'alarmines (i.e. médiateurs de l'inflammation) par les cellules épithéliales. Ces médiateurs seront l'interleukine (IL) -25, l'IL-33 et la lymphopoïétine stromale thymique (TSLP; Han et al., 2017; Hill & Spergel, 2018). Ces alarmines activeront des cellules dendritiques (CD) et des cellules lymphoïdes innées (ILC) de type 2 (ILC2). Les ILC2, une fois activées, relâcheront de l'IL-4, l'IL-5 et

l'IL-13 qui permettront la différenciation des lymphocytes T CD4⁺ naïfs (Th0). Les CD actives, quant à elles, se fixeront à l'aéroallergène et l'internaliseront afin de le couper en plusieurs peptides, puis migreront vers un nœud lymphoïde (Holgate, 2012). Une fois dans ce nœud, elles présenteront les peptides de l'aéroallergène aux Th0 et en présence d'IL-4, ceci permettra la différenciation des Th0 en lymphocytes T auxiliaires (Th) de type 2 (Th2) spécifiques à l'allergène (Komlósi et al., 2022; Werfel et al., 2016). Ensuite, les Th2 présenteront l'antigène de l'aéroallergène aux lymphocytes B. En plus des cytokines relâchées par les ILC2, les Th2 relâcheront de l'IL-4 et l'IL-13 afin de stimuler la commutation des IgM à la classe des IgE et la différenciation des lymphocytes B en plasmocytes (Holgate, 2012; Palomares et al., 2017). Les plasmocytes relâcheront, dans le milieu, des IgE spécifiques à l'allergène qui se dirigeront dans la circulation sanguine (J Bousquet et al., 2020). Ces IgE seront reconnues et se lieront aux récepteurs de haute affinité pour les IgE (FcεRI) sur la surface des mastocytes et des basophiles. Il est à noter que pendant ces étapes, il y aura expansion clonale des Th2 et des lymphocytes B spécifiques à l'aéroallergène, ce qui produira un groupe de cellules mémoires Th2 et B pour l'aéroallergène (Brough et al., 2020; Palomares et al., 2017). Ainsi, l'individu aura développé des cellules immunitaires capables de reconnaître l'aéroallergène et son système immunitaire sera donc capable de réagir lors d'une future exposition à celui-ci.

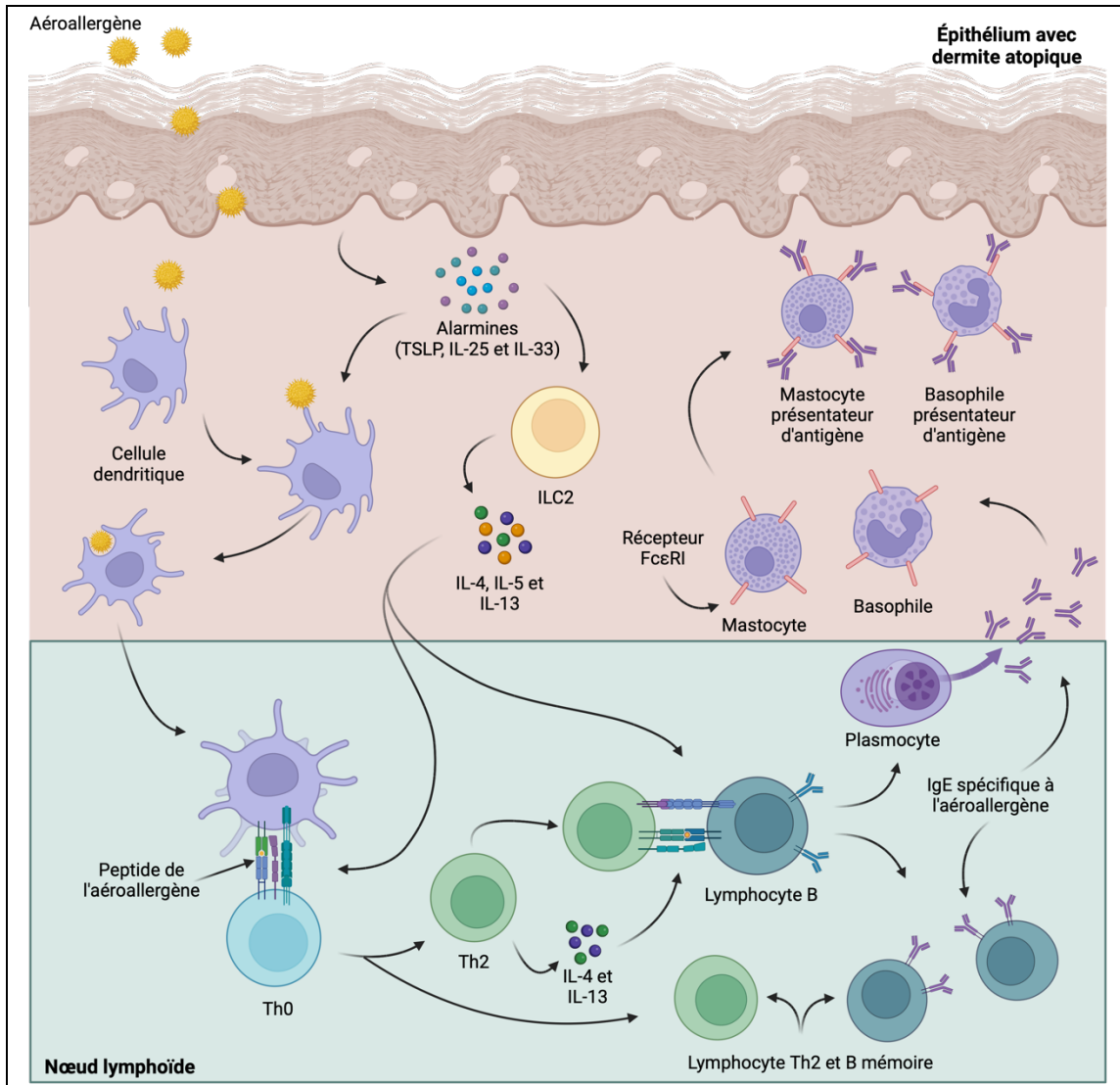


Figure 1. Physiopathologie du développement de la sensibilisation à un aéroallergène

La pénétration de l'aéroallergène au niveau de l'épithélium cutané entraîne une activation des CD et des ILC2. Les CD fixeront l'aéroallergène, le dégraderont en peptides et permettront la différenciation des Th0 en Th2. Puis, les Th2 présenteront l'antigène aux lymphocytes B afin de permettre leur différenciation en plasmocytes. Les plasmocytes produiront et libéreront dans la circulation des IgE spécifiques à l'aéroallergène. Ces IgE se fixeront sur les mastocytes et les basophiles afin que ces cellules puissent reconnaître l'aéroallergène dans le futur (Figure réalisée avec BioRender®, 2023).

2.2 L'asthme allergique

2.2.1 Épidémiologie de l'asthme allergique

L'asthme est une maladie inflammatoire chronique des voies respiratoires qui se traduit par de l'inflammation, de l'hyperréactivité bronchique et un remodelage des voies respiratoires (Holgate et al., 2015; L. Yang et al., 2020). Plus de 358 millions d'individus dans le monde sont atteints d'asthme (A. Ismaila et al., 2019). En 2020, il était estimé qu'environ 2,8 millions (8,7 %) de Canadiens âgés de 12 ans et plus avaient un diagnostic d'asthme (Statistique Canada, 2023). Cependant, cette prévalence n'est pas la même partout dans le monde et sera beaucoup moins élevée dans les pays non industrialisés contrairement aux pays en voie d'industrialisation et industrialisés (Holgate et al., 2015). De plus, elle peut varier par région à l'intérieur d'un même territoire. Par exemple, la région du Saguenay–Lac-Saint-Jean (SLSJ) possède la plus haute prévalence pour l'asthme parmi toutes les régions de la province du Québec (Doucet et al., 2020). L'appartenance à un groupe ethnique influence la prévalence de la maladie, ce qui souligne également l'importance des déterminants génétiques dans le développement de cette maladie (Braman, 2006). À cet effet, les Afro-Américains et les Hispaniques ont une prévalence d'asthme supérieure à celle des populations d'origine européenne française par exemple (Braman, 2006; Toskala & Kennedy, 2015).

L'âge et le sexe sont deux autres facteurs qui influencent la prévalence de l'asthme. Ainsi, durant l'enfance, la dermatite atopique et l'asthme seraient plus fréquents chez les garçons jusqu'à la puberté (< 14 ans), alors que leurs fréquences seraient plus grandes chez les filles à partir de l'adolescence (> 14 ans; Shah & Newcomb, 2018; Yoshida et al., 2017). De plus, l'asthme serait plus sévère chez les garçons avant la puberté (< 14 ans), alors que chez les femmes la sévérité de l'atteinte est plus importante à partir de l'âge adulte (> 23 ans; Shah & Newcomb, 2018). L'incidence de l'AA diminue avec le temps et la plupart des cas d'asthme diagnostiqués après 40 ans sont non associés à l'allergie (Pakkasela et al., 2020). L'asthme est donc un trait complexe résultant de

l'interaction entre des facteurs génétiques et environnementaux qui touchent des personnes de tous les statuts socioéconomiques et durant tout le parcours de vie.

L'asthme a des répercussions sur la qualité de vie d'une personne et celle-ci décroît avec l'accroissement de la sévérité des manifestations de l'asthme (A. S. Ismaila et al., 2013). De plus, l'asthme est un lourd fardeau économique pour le Canada par les divers coûts directs (médicaments, hospitalisation, etc.) et indirects (absentéisme, perte de la productivité, etc.) occasionnés. Notamment, une étude sur les coûts directs et indirects de l'asthme au Québec a permis d'estimer que ceux-ci totaliseraient un fardeau économique d'environ 5,8 milliards de dollars annuellement (A. Ismaila et al., 2019).

2.2.2 Phénotype de l'asthme allergique

L'asthme est associé à des difficultés respiratoires, à une respiration sifflante, à un souffle court, à un serrement de la poitrine, à des expectorations et à la présence de toux dont la sévérité peut varier avec le temps (Holgate et al., 2015). De plus, l'asthme est caractérisé par une hyperréactivité bronchique, une hypertrophie et hyperplasie des muscles lisses (i.e. augmentation de la taille et augmentation du nombre de cellules), un remodelage des voies respiratoires, une bronchoconstriction des voies respiratoires inférieures et une hypersécrétion de mucus (Hamilton & Lehman, 2020; Komlósi et al., 2022; Samitas et al., 2018).

Il existe plusieurs facteurs déclencheurs de l'asthme qui peuvent varier d'un individu à un autre, engendrant une grande variabilité dans les symptômes et leur sévérité (Holgate et al., 2015). Cette variabilité est désignée par l'hétérogénéité clinique de l'asthme. Par conséquent, il existe plusieurs phénotypes d'asthme et leur classification peut se faire suivant certains critères. Ainsi, l'asthme peut être classifié par endotype, c'est-à-dire selon les mécanismes cellulaires et moléculaires impliqués dans sa physiopathologie. Le premier endotype de l'asthme se distingue par l'action des Th2, des ILC2 et des éosinophiles (Kuruvilla et al., 2019). Ce type d'inflammation est dit de type Th2. Le deuxième endotype quant à lui est marqué par une inflammation de type non-Th2 dont les

mécanismes physiopathologiques sont engendrés principalement par l'action des neutrophiles, des cellules Th1 ou Th17 et des ILC3 (Kim et al., 2014; Kuruvilla et al., 2019). La classification phénotypique de l'asthme se base sur des caractéristiques observables incluant l'âge, le sexe, les symptômes, la sévérité et les éléments déclencheurs de l'asthme (Figure 3; Agache & Akdis, 2019; Kuruvilla et al., 2019). Parmi les divers phénotypes de l'asthme, c'est l'AA qui est le plus fréquent chez les individus de jeune âge. Notamment, elle représente environ 75 % des individus asthmatiques (Hamilton & Lehman, 2020; Toskala & Kennedy, 2015). L'AA est caractérisé par une inflammation de type Th2, un taux d'IgE totale élevé, une sensibilisation allergénique et peut s'accompagner d'un taux élevé d'éosinophiles dans le sang (Schatz & Rosenwasser, 2014). La réaction inflammatoire de ce phénotype est généralement causée par la présence d'un ou de plusieurs allergène(s) au(x)quel(s) l'individu est sensibilisé (Hamilton & Lehman, 2020). C'est l'endotype Th2 et le phénotype d'AA qui est d'ailleurs considéré dans ce mémoire.

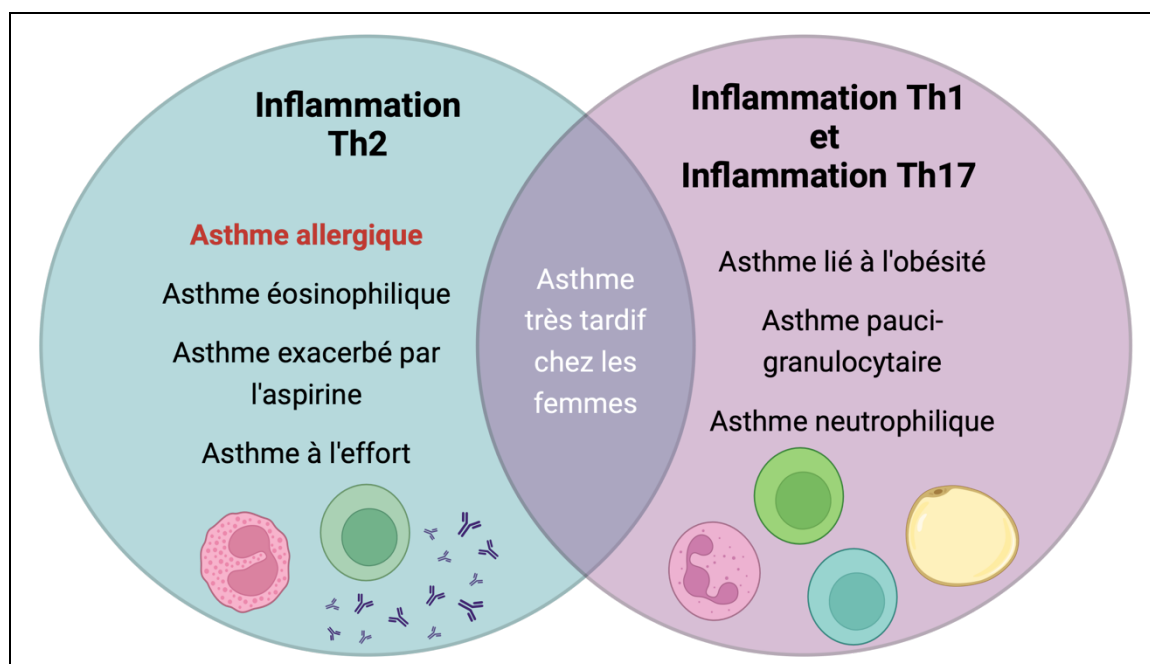


Figure 2. Les divers phénotypes de l'asthme regroupés en deux endotypes

L'asthme peut être classé selon deux endotypes regroupant plusieurs phénotypes. Les phénotypes sont regroupés en deux endotypes selon les similarités dans leurs mécanismes physiopathologiques (Figure réalisée avec BioRender®, 2023).

2.2.3 Inflammation bronchique dans l'asthme allergique

La réaction inflammatoire dans l'AA survient une fois que l'individu est sensibilisé à l'aéroallergène et qu'il est exposé à celui-ci. Cette réaction peut se faire en deux phases, soit la phase précoce et la phase tardive. La phase précoce est une réaction d'hypersensibilité de type I médiée par les mastocytes, les basophiles et les IgE et qui survient après quelques minutes d'exposition et qui se résout environ 1 à 3 heures après. La phase tardive est quant à elle une réaction d'hypersensibilité de type IVb médiée par les Th2 et les éosinophiles qui survient de 3 à 4 heures après la phase précoce et qui atteint son pic d'intensité après 6 à 12 heures suivant la phase précoce (F. Ferreira et al., 2004; Gauvreau et al., 2015; Uzzaman & Cho, 2012). Il est à noter que ce ne sont pas tous les individus asthmatiques qui subissent la phase tardive étant donné qu'elle survient chez environ 80 % des enfants et 60 % des adultes (M. A. R. Ferreira, 2004; Gauvreau et al., 2015).

2.2.3.1 Réaction précoce

Durant la phase précoce (Figure 4), l'aéroallergène traverse l'épithélium bronchique et occasionne la sécrétion de médiateurs de l'inflammation par les cellules épithéliales. Ces médiateurs sont l'IL-25, l'IL-33 et la TSLP, qui activeront les ILC2 qui sécréteront de l'IL-4, l'IL-5, l'IL-9 et de l'IL-13 dans le milieu (Hall & Agrawal, 2014; Holgate et al., 2015). L'aéroallergène se fixera aux IgE spécifiques fixées sur les récepteurs FcεRI des mastocytes présentateurs d'antigènes préalablement créés lors de la phase de sensibilisation allergénique (Bloemen et al., 2007). En plus de la présence de l'IL-4, de l'IL-9 et de l'IL-13, cela occasionnera l'activation des mastocytes, provoquant ainsi leur dégranulation (Holgate et al., 2015). Leurs granules relâcheront alors de l'histamine, des protéases, des protéoglycanes, de la prostaglandine D2 (PGD2), des leucotriènes, des chimiokines et des cytokines (facteurs de nécrose tumorale alpha [TNF-α], IL-3, IL-4, IL-5, IL-6, IL-8 et IL-16; Bloemen et al., 2007). La même chose se déroulera pour les basophiles présentateurs d'antigènes (Gauvreau et al., 2015; Hall & Agrawal, 2014). Par contre, leurs granules déverseront dans le milieu de l'histamine, des leucotriènes, de l'IL-4 et de l'IL-13 (Bloemen et al., 2007). Ainsi, la PGD2 stimulera

la bronchoconstriction des muscles lisses, la vasodilatation, l'augmentation de la perméabilité vasculaire et le recrutement des éosinophiles au site de l'inflammation. Les leucotriènes stimuleront la bronchoconstriction, l'hypersécrétion de mucus, l'augmentation de la perméabilité vasculaire et l'activation des éosinophiles. L'histamine quant à elle stimulera la constriction des muscles lisses des bronches et la sécrétion de mucus par les cellules à mucus (Bloemen et al., 2007; Hall & Agrawal, 2014). De plus, l'IL-5 servira dans l'activation des éosinophiles et l'IL-13 stimulera l'hypersécrétion de mucus par les cellules à mucus (Holgate et al., 2015). Ainsi, la phase précoce provoquera la bronchoconstriction, l'hypersécrétion de mucus, la vasodilatation, la perméabilité vasculaire, l'hyperréactivité bronchique et le recrutement de cellules inflammatoires (Bloemen et al., 2007).

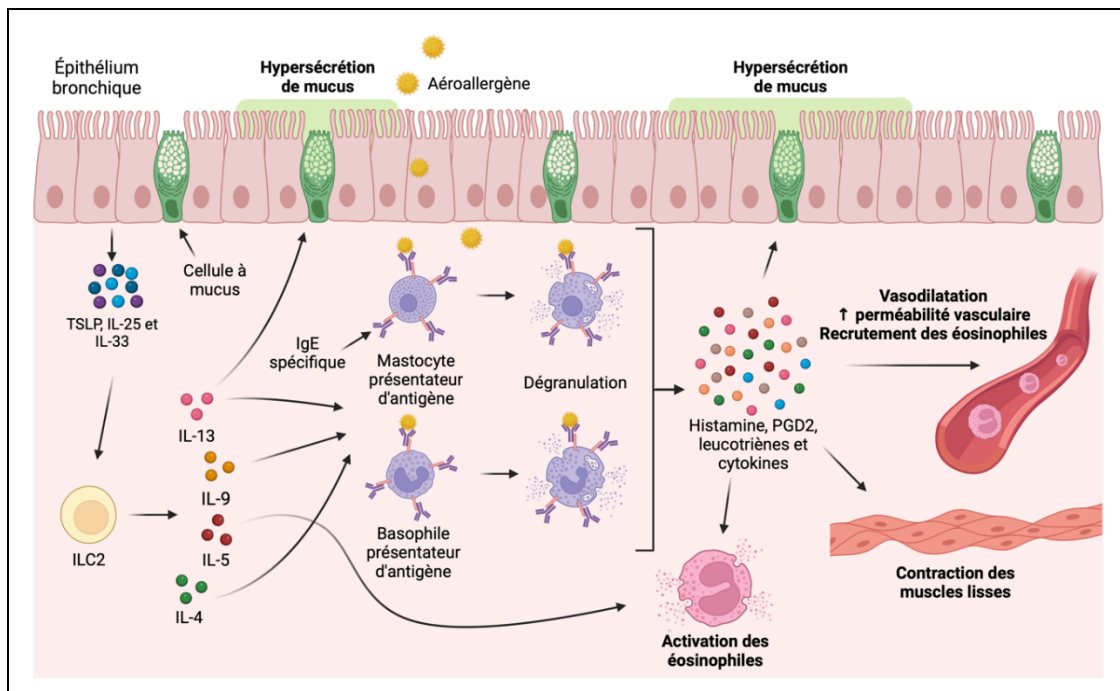


Figure 3. Réaction inflammatoire précoce de l'asthme allergique

L'aéroallergène se fixera sur les IgE spécifiques sur les mastocytes et les basophiles présentateurs d'antigènes, ce qui entraînera leur activation. Celle-ci entraînera ensuite le relâchement de médiateurs de l'inflammation. Ainsi, il y aura stimulation de l'hypersécrétion de mucus, la contraction des muscles lisses, la vasodilatation et l'augmentation de la perméabilité des vaisseaux sanguins, ainsi que le recrutement et l'activation des éosinophiles (Figure réalisée avec BioRender®, 2023).

2.2.3.2 Réaction tardive

Les médiateurs de l'inflammation relâchés précédemment lors de la réaction précoce permettront l'enclenchement de la réaction tardive par l'activation des CD. Une fois activés les CD, fixeront un aéroallergène et iront présenter les peptides de l'aéroallergène aux Th2 mémoires présents dans le tissu et aux Th0 présents dans le nœud lymphoïde en présence d'IL-4 relâchées par les ILC2 et les éosinophiles. Cela activera les Th2 présents dans le tissu, alors qu'il y aura différenciation en Th2 dans le nœud lymphoïde et migration de celles-ci vers le site de l'inflammation (M. A. R. Ferreira, 2004; Holgate, 2012). Ainsi, les Th2 sécréteront dans le milieu des interleukines (IL-3, IL-4, IL-5, IL-6, IL-9, IL-10, IL-13), des facteurs de croissance hématopoïétique (GM-CSF) et des chimiokines. Ces molécules seront impliquées dans plusieurs processus nécessaires au développement de la phase tardive de l'AA. Notamment, les GM-CSF permettront le recrutement des macrophages et des neutrophiles au site de l'inflammation (Hall & Agrawal, 2014). L'IL-4 et l'IL-13 stimuleront la production des IgE spécifiques par les plasmocytes, et ceux-ci migreront dans le tissu bronchique. Par la suite, les IgE se fixeront sur les aéroallergènes, favorisant leur phagocytose par les macrophages et les neutrophiles (M. A. R. Ferreira, 2004). L'IL-5 permettra le recrutement et l'activation des éosinophiles au site de l'inflammation, ce qui causera le déversement de leurs contenus cytotoxiques dans le milieu. Ainsi, il y aura déversement de molécules comme les protéines basiques majeures, des radicaux libres, des protéines cationiques et des neurotoxines dérivées des éosinophiles, des peroxydases, des leucotriènes, des cytokines (Ex. IL-4, IL-5, IL-13, GM-CSF) et des chimiokines. Ces molécules endommageront la barrière épithéliale, favoriseront le maintien de l'inflammation, stimuleront la bronchoconstriction, augmenteront l'hypersécrétion de mucus, la perméabilité vasculaire et le recrutement d'éosinophiles et de Th2 (Bloemen et al., 2007; Hall & Agrawal, 2014). De plus, les cellules épithéliales endommagées sécréteront des alarmines, ce qui stimulera le maintien de la réponse inflammatoire (Eifan & Durham, 2016). Il y aura une boucle de recrutement et d'activation de cellules inflammatoires qui se mettra en place, contribuant ainsi à l'accumulation de ces cellules dans le tissu. Cela provoquera le remodelage du tissu bronchique par un épaissement apparent de la membrane basale (dépôt de collagènes de type I, III, V et VI, de

périostines, ténascines et de fibronectines), la fibrose épithéliale, l'hyperplasie des cellules à mucus et l'hyperplasie ainsi que l'hypertrophie (i.e. l'accroissement de la taille) des cellules musculaires lisses (Bloemen et al., 2007; Holgate et al., 2015; Komlósi et al., 2022). La Figure 5 schématise les principales étapes de la réaction inflammatoire tardive spécifique à l'AA.

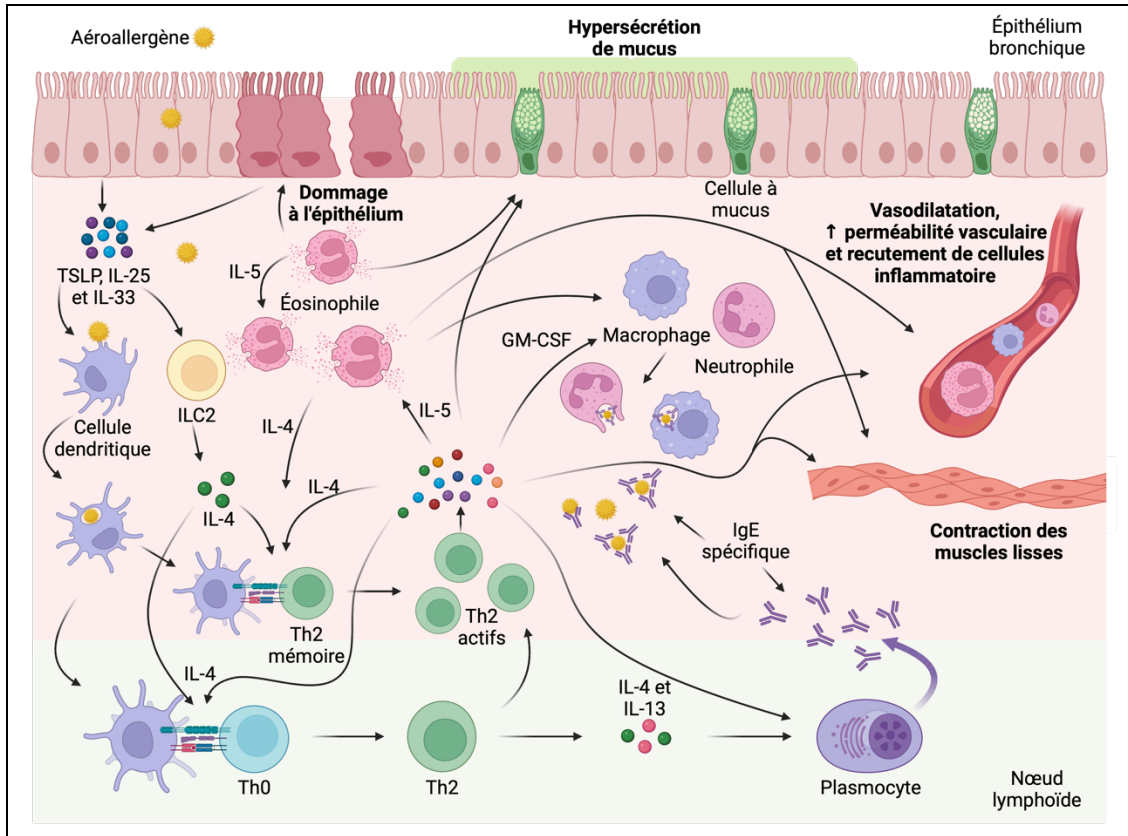


Figure 4. Réaction inflammatoire tardive de l'asthme allergique

L'activation des CD entraînera la présentation de l'aéroallergène aux Th2 mémoires et aux Th0. Ceci causera une migration des cellules Th2 et des IgE ainsi que le recrutement et l'activation de cellules immunitaires au site de l'inflammation. L'aéroallergène sera ensuite phagocyté par les cellules immunitaires. Les Th2 et les éosinophiles relâcheront des médiateurs de l'inflammation qui occasionneront des dommages à l'épithélium, l'hypersécrétion de mucus, de la bronchoconstriction, l'augmentation de la perméabilité vasculaire et le recrutement de cellules inflammatoires (Figure réalisée avec BioRender®, 2023).

2.3 La rhinite allergique

2.3.1 Épidémiologie de la rhinite allergique

La RA est une maladie inflammatoire des muqueuses nasales résultant d'une sensibilisation allergénique et qui, tout comme l'asthme, possède une prévalence qui peut varier selon les pays. Ainsi, la prévalence de cette maladie se situe entre 5 à 50 % dépendamment d'où l'on se trouve dans le monde (Wise et al., 2023). Dans la population canadienne, la prévalence de la RA se situe entre 20 à 25 % puisqu'elle est variable d'une province à l'autre (Keith et al., 2012). Cette variation de la prévalence peut être causée par plusieurs facteurs comme la diversité génétique, la présence de polluant ou encore le mode de vie. La prévalence de la RA est aussi influencée par l'âge et le sexe, étant plus élevée chez les garçons avant la puberté pour devenir plus élevée chez les filles après la puberté (Keller et al., 2018; Pinart et al., 2017). Globalement, la prévalence de la RA augmente avec l'âge, peu importe le sexe (Czarnowicki et al., 2017).

La RA peut avoir des répercussions sur la qualité de vie des individus atteints et de leur entourage (Wise et al., 2023). Notamment, en 2012, une étude sur la gestion et l'impact de la RA sur la population canadienne a permis de mettre en évidence que les symptômes de la RA avaient un impact sur la vie quotidienne d'environ 72 % des personnes atteintes de RA pendant la saison du pollen. Les symptômes rapportés qui perturbaient le plus les individus au quotidien étaient la fatigue (46 %) et les maux de tête (37 %) alors que ceux qui perturbaient le moins étaient une mauvaise concentration (32 %) et une productivité réduite (23 %; Keith et al., 2012). La RA représente un lourd fardeau économique pour la société par ses divers coûts directs et indirects. Cependant, il est difficile d'évaluer précisément à combien ce fardeau s'élève. Par contre, une étude effectuée en 2005 sur les coûts occasionnés par l'allergie aux pollens d'*Ambrosia artemisiifolia* au Québec a démontré que les coûts directs associés au traitement de la RA se situeraient entre 157 et 230 millions de dollars (Tardif, 2008; Tardif & Masson, 2008).

2.3.2 Phénotype de la rhinite allergique

La RA est une maladie des voies respiratoires supérieures qui résulte de la sensibilisation allergénique, dont la symptomatologie et la sévérité peuvent varier dans le temps (Meng et al., 2020; Wise et al., 2023). Ainsi, les principaux symptômes, qui peuvent survenir seuls ou combinés, sont les éternuements, l'écoulement nasal, la congestion ainsi que l'inflammation des muqueuses nasales (Eifan & Durham, 2016). La réaction d'inflammation de la RA est de type Th2, ce qui veut dire qu'elle implique les cellules Th2, les ILC2 et leurs cytokines. De plus, comme l'AA, elle peut se faire en deux temps, soit la phase précoce et la phase tardive (Wise et al., 2023). La phase dite précoce est une réaction d'hypersensibilité de type I qui survient après quelques minutes d'exposition à un allergène et atteint son maximum cinq minutes après l'exposition à l'allergène (G. W. Scadding et al., 2015; Uzzaman & Cho, 2012; Wise et al., 2023). Ce type de réaction occasionne l'apparition d'écoulement nasal, de congestion nasale et de démangeaisons chez un individu (Wise et al., 2018). En ce qui concerne la phase dite tardive, elle est une réaction d'hypersensibilité de type IVb, elle survient entre 4 et 12 heures après l'exposition à un allergène (Drazdauskaitė et al., 2021; Uzzaman & Cho, 2012). Cette réaction provoque une obstruction nasale, une augmentation de la sécrétion de mucus, une perte de l'odorat, une inflammation des muqueuses nasales et une hyperactivité nasale (Wise et al., 2018). De plus, l'hyperactivité nasale occasionnée lors de la RA rend la muqueuse nasale très réactive non seulement aux allergènes, mais à d'autres stimuli comme de fortes odeurs ou d'autres substances irritantes (Nur Husna et al., 2022).

2.3.3 Inflammation causée par la rhinite allergique

Tout comme l'AA, la réaction inflammatoire de la RA survient une fois que l'individu est sensibilisé à l'aéroallergène et qu'il est exposé à celui-ci. La RA partage les mêmes processus inflammatoires précoces et tardifs que l'AA, mais ceux-ci se déroulent dans les voies respiratoires supérieures, soit dans les voies nasales. Dans la RA, il y a peu de remodelage des voies respiratoires supérieures comparativement à ce qui est observé dans l'AA au niveau du tissu bronchique. Cette

absence de remodelage est attribuable à plusieurs paramètres, dont l'absence de muscle lisse, et à une meilleure capacité de protection et de régénération de l'épithélium (réduction de la desquamation) dans les voies respiratoires supérieures (Gelfand, 2004; Samitas et al., 2018; Vignola et al., 2003). De plus, les voies respiratoires supérieures ont une plus grande vascularisation ce qui contribue à l'obstruction nasale observée dans la RA par l'action des médiateurs de l'inflammation relâchés par les mastocytes et les basophiles (Vignola et al., 2003). La réaction précoce de la RA survient après quelques minutes d'exposition, alors que la réaction tardive débute après 3 à 4 heures et atteint son pic entre 6 et 9 heures après l'exposition (G. Scadding, 2014). La phase précoce survient chez tous les individus atteints de RA, mais la phase tardive ne survient que chez 50 % des individus atteints de RA (Peters-Golden et al., 2006).

2.3.3.1 Réaction précoce

Dans la phase précoce de la RA, l'histamine occasionnera, en plus de l'hypersécrétion de mucus par les cellules à mucus, une sensation de démangeaison et des éternuements en agissant sur les nerfs sensoriels présents dans la cavité nasale (Eifan & Durham, 2016). Les leucotriènes et la PGD2 sécrétés par les mastocytes et les basophiles augmenteront la perméabilité vasculaire des vaisseaux sanguins, la sécrétion de mucus par les glandes muqueuses et les cellules à mucus, et stimuleront aussi le recrutement et l'activation des éosinophiles dans le tissu nasal. Ainsi, ces molécules provoqueront un œdème et de la rhinorrhée (i.e. écoulement nasal). Cet œdème des parois nasales et la surproduction de mucus occasionnée par ces médiateurs provoqueront la congestion nasale observée chez les personnes atteintes de la RA (Eifan & Durham, 2016; Gelfand, 2004; Peters-Golden et al., 2006). En résumé, la réaction précoce de la RA se différencie de celle observée pour l'AA par son absence de contraction des muscles lisses et par son action sur les glandes muqueuses et les nerfs sensoriels. La Figure 7 résume les principales étapes des mécanismes de l'inflammation précoce de la RA. Elle permet de mettre en lumière les différences avec ce processus dans le contexte de l'AA (Figure 4).

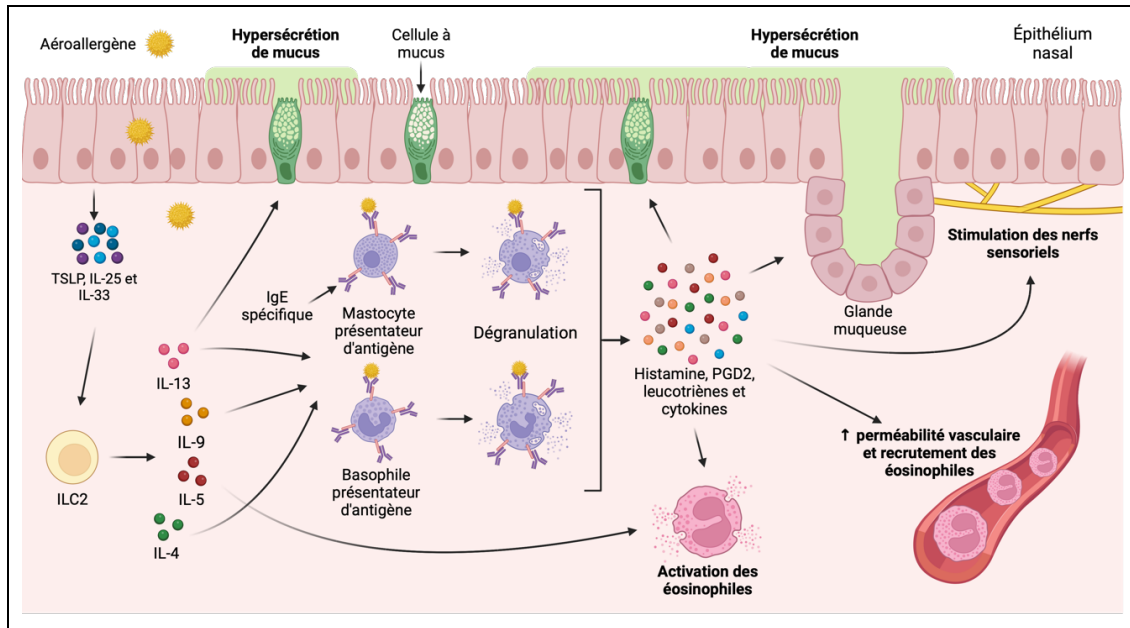


Figure 5. Réaction inflammatoire précoce de la rhinite allergique

L'aéroallergène se fixera sur les IgE spécifiques sur les mastocytes et les basophiles présentateurs d'antigènes, ce qui les active. Leur activation entraînera le relâchement de médiateurs de l'inflammation. Ainsi, il y aura stimulation de l'hypersécrétion de mucus par les cellules à mucus et par les glandes muqueuses, la stimulation des nerfs sensoriels, l'augmentation de la perméabilité des vaisseaux sanguins ainsi que le recrutement et l'activation des éosinophiles (Figure réalisée avec BioRender®, 2023).

2.3.3.2 Réaction tardive

Dans la phase tardive de la RA, il y aura absence de remodelage des voies respiratoires supérieures et de la stimulation des muscles lisses. Par contre, il y aura l'ajout de la stimulation des glandes muqueuses et des nerfs sensoriels au processus inflammatoire de la phase tardive. Les cellules qui prédomineront dans le tissu lors de la phase tardive seront les Th2 et les éosinophiles. Ces cellules relâcheront des médiateurs de l'inflammation qui contribueront au phénotype de la rhinorrhée, de la démangeaison, de l'enflure et de l'obstruction nasale par les mécanismes suivant : la stimulation des cellules à mucus, des glandes muqueuses et des nerfs sensoriels, la dilatation et l'augmentation de la perméabilité des vaisseaux sanguins, et la favorisation de l'infiltration du tissu nasal par des cellules immunitaires (J Bousquet et al., 2020; Settipane & Schwindt, 2013). Il est à noter que parmi les médiateurs inflammatoires relâchés par les Th2 et les éosinophiles on retrouve

des cytokines permettant leurs propres recrutement et activation en plus d'occasionner la réactivation des processus de la phase précoce (Drazdauskaitė et al., 2021; Hansen et al., 2004). De plus, le relâchement du contenu cytotoxique des granules des éosinophiles occasionnera des dommages à la barrière épithéliale, ce qui contribuera à la sécrétion d'alarmines et au maintien de l'inflammation (Eifan & Durham, 2016). La Figure 8 résume les principales étapes des mécanismes de l'inflammation tardive de la RA. Elle permet de mettre en lumière les différences avec ce processus dans le contexte de la phase tardive de l'AA (Figure 5).

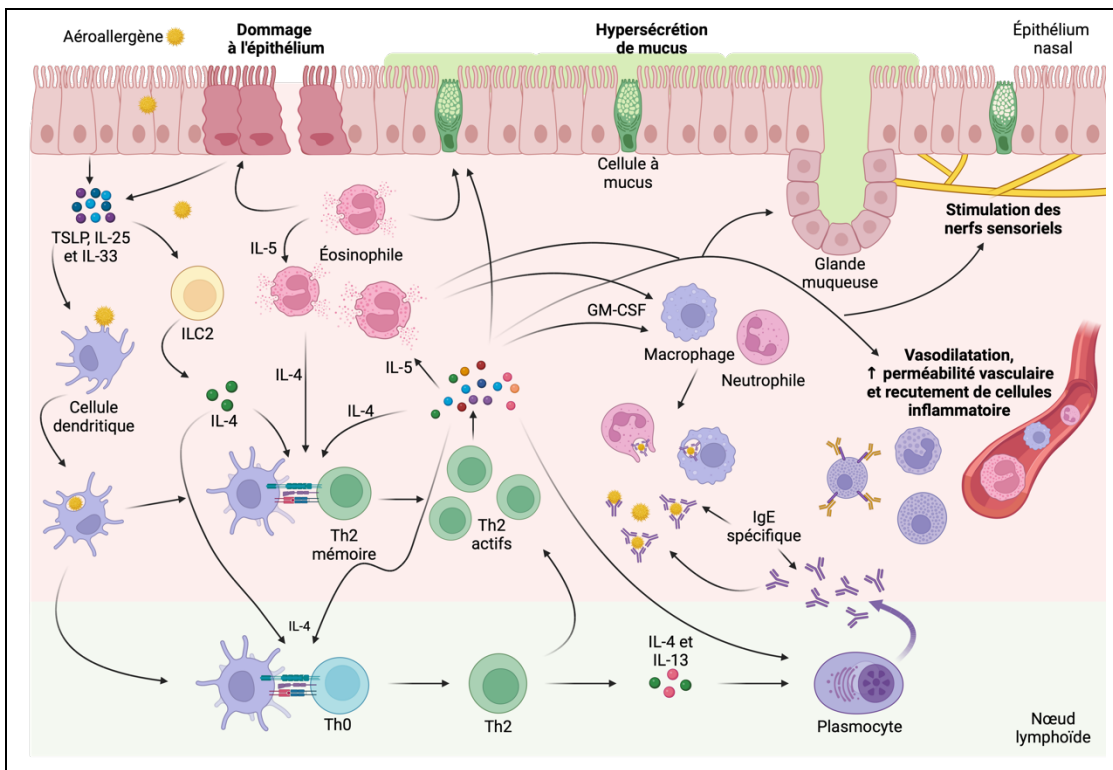


Figure 6. Réaction inflammatoire tardive de la rhinite allergique

L'activation des CD entraînera la présentation de l'aéroallergène aux Th2 mémoires et aux Th0. Ceci causera une migration des cellules Th2 et des IgE ainsi que le recrutement et l'activation de cellules immunitaires au site de l'inflammation. L'aéroallergène sera ensuite phagocyté par les cellules immunitaires. Les médiateurs de l'inflammation relâchés par les Th2 et les éosinophiles occasionneront des dommages à l'épithélium, l'hypersécrétion de mucus, la stimulation des nerfs sensoriels, l'augmentation de la dilation et de la perméabilité vasculaire ainsi que le recrutement de cellules inflammatoires dans la muqueuse nasale (Figure réalisée avec BioRender®, 2023).

CHAPITRE 3

GÉNÉTIQUE ET ÉPIGÉNÉTIQUE DE L'ASTHME ALLERGIQUE ET DE LA RHINITE ALLERGIQUE

3.1 Trait complexe : définition

Les maladies allergiques sont des traits complexes, ce qui signifie que leur phénotype est modulé à la fois par des facteurs génétiques et des facteurs environnementaux (Wang et al., 2023). Les facteurs génétiques sont propres à l'individu et correspondent à des variants génétiques transmis par les parents. Pour chaque trait complexe, il existe plusieurs variants de susceptibilité dont chacun contribue au risque total de développer la maladie. Il est à noter que les combinaisons de ces variants de susceptibilité sont distinctes entre individus, ce qui fait en sorte que le risque de développer le trait est différent d'un individu à l'autre (Virolainen et al., 2023). En ce qui concerne les facteurs environnementaux, ils sont propres aux habitudes de vie d'un individu ou à l'environnement auquel l'individu est exposé, et ce même *in utero* (Figure 9; Potaczek et al., 2017; Virolainen et al., 2023). Des exemples de facteurs environnementaux sont l'exposition à des polluants, à des aéroallergènes et les infections respiratoires entre autres (Potaczek et al., 2017). La combinaison de facteurs environnementaux et génétiques ainsi que leurs interactions (i.e. mécanismes épigénétiques) permet de moduler le risque de développer un trait précis et donc d'éventuellement mener au développement d'un trait complexe comme l'AA et la RA (Virolainen et al., 2023).

3.1.2 Facteurs génétiques

Certaines variations dans la séquence nucléotidique de l'ADN peuvent être plus ou moins fréquentes pour les personnes atteintes comparativement aux personnes non atteintes d'un phénotype spécifique. Lorsqu'un tel déséquilibre existe, il est suggéré que cette variation soit associée au trait étudié (Ntontsi et al., 2021). Les SNP qui se retrouvent sur l'ensemble du génome servent de marqueurs moléculaires pour les études d'association génétique (Wang et al., 2023). Ces SNP peuvent se retrouver à l'intérieur d'un gène, mais sont majoritairement situés dans des régions

de l'ADN dites « non codantes », c'est-à-dire qu'elles ne codent pas pour la production de protéines. Lorsque la mutation survient dans un gène, celle-ci peut modifier la structure ou la fonction de la protéine codée et contribuer au développement du phénotype. En ce qui concerne un variant génétique survenant dans une région « non codante », celui-ci peut aussi contribuer au développement du phénotype. Ceci s'explique par l'implication potentielle des régions « non codantes » dans les mécanismes de régulation des gènes et ultimement du phénotype (Virolainen et al., 2023).

Les variants génétiques sont héréditaires, c'est-à-dire qu'ils sont transmis à la descendance, mais ceux-ci peuvent aussi survenir suite à une mutation spontanée dans le génome d'un individu. Ainsi, la présence d'individus ayant des allergies dans la famille est un facteur de risque pour le développement de maladies allergiques (Baloh & Mathias, 2023). Des études ont été capables d'évaluer l'héritabilité des maladies allergiques, c'est-à-dire l'estimation de la partie du phénotype qui est attribuable à la génétique dans une population donnée. Ainsi, l'héritabilité estimée de l'asthme se situe entre 35 et 95 %, celle de la RA entre 33 et 91 % et celle pour la sensibilisation allergique entre 40 et 85 % selon la population étudiée (Bønnelykke et al., 2013; Ober & Yao, 2011). Il est à noter que ce ne sont pas tous les variants génétiques qui ont la même importance pour le développement de la maladie. Effectivement, certains variants contribuent plus que d'autres au risque génétique, mais cela n'empêche pas ceux contribuant le moins de s'accumuler pour contribuer au développement de la maladie (Ober & Yao, 2011).

3.1.1 Facteurs environnementaux

L'exposome, soit l'ensemble des facteurs environnementaux auxquels un individu est exposé, se divise en trois types. Le premier exposome correspond aux facteurs environnementaux externes d'un milieu, tels que le climat, la pollution, le type de zone d'habitation (urbaine ou rurale), et l'éducation. Le deuxième type d'exposome est externe et lié aux habitudes de vie d'un individu, comme l'alimentation, le tabagisme, la consommation de cannabis, la consommation d'alcool et l'activité physique. Le dernier type d'exposome est l'environnement interne, soit biologique,

correspondant aux manifestations biologiques et métaboliques dans le corps, résultant des deux autres types d'exposomes (Burbank et al., 2017). De ce fait, une multitude de facteurs environnementaux peuvent avoir un impact variable selon la maladie. Il existe des facteurs protecteurs, comme certaines bactéries du microbiote intestinal qui protège contre le développement d'allergie alimentaire par exemple, alors que d'autres sont néfastes, comme le virus respiratoire syncytial qui est un facteur de risque pour le développement de l'asthme (Virolainen et al., 2023). L'impact des facteurs environnementaux sur le développement des traits complexes se fait par la régulation de l'expression des gènes par l'épigénétique. L'épigénétique est l'ensemble des changements biochimiques de l'ADN, héréditaires ou acquis, qui ne modifient pas la séquence nucléotidique de celui-ci (DeVries & Vercelli, 2015; Potaczek et al., 2017).

3.2 Génétique de la sensibilisation allergique

3.2.1 Étude d'association pangénomique

Une étude GWAS permet l'association de variants génétiques à un phénotype. Lors de ce type d'étude, le génome complet d'individus atteints d'un phénotype précis est comparé à celui d'individus ne le présentant pas. Cette comparaison permet d'identifier tous les variants génétiques qui sont présents en plus ou moins grande fréquence chez les individus atteints du phénotype étudié (Cano-Gamez & Trynka, 2020). Par conséquent, les variants qui ressortent sont significativement plus ou moins fréquents chez les individus ayant le phénotype et sont donc associés au trait, s'ils passent le seuil de significativité du GWAS. Ce type d'analyse a permis d'identifier de nouvelles associations entre des variants et un phénotype ainsi que des loci de susceptibilité et de voies biologiques impliquées dans plusieurs traits (Sollis et al., 2023; Tam et al., 2019). De plus, les données obtenues par GWAS peuvent avoir d'autres applications plus larges. Notamment, elles peuvent être utilisées pour l'estimation de l'héritabilité d'une maladie ainsi que pour la création de scores de risque polygénique pour le dépistage préventif de maladies (Tam et al., 2019). Les GWAS sont largement utilisés comme stratégie d'analyse en génétique et en 2022, le *GWAS Catalog* regroupait environ

400 000 associations SNP-maladie provenant de plus de 45 000 GWAS différents (Sollis et al., 2023). Notamment, un article de revue mentionne qu'environ 267 loci associés à l'asthme et/ou à la RA auraient été identifiés dans la littérature à partir de 31 GWAS (J Bousquet et al., 2023).

3.2.2 Variants de susceptibilité

Dans les dernières décennies, la recherche sur l'AA et la RA a permis d'identifier plusieurs variants génétiques et gènes impliqués dans leur développement. Beaucoup d'entre eux sont associés à la fois à l'AA et à la RA. Ceci n'est pas surprenant étant donné que les maladies allergiques partagent une partie de leur composante génétique et qu'elles possèdent des mécanismes inflammatoires similaires (B. Y. Choi et al., 2021). Notamment, le locus 11q13 contenant le gène *EMSY transcriptional repressor, BRCA2 interacting (EMSY)* et le gène *leucine rich repeat containing 32 (LRRC32)* sont associés à l'asthme, la dermatite atopique et la RA (Portelli et al., 2015). Le gène *EMSY* est impliqué dans les fonctions de la barrière épithéliale et de sa régulation, alors que le gène *LRRC32* est impliqué dans les fonctions des lymphocytes T et dans la tolérance immunitaire (B. Y. Choi et al., 2021). Parmi les nombreux gènes et variants génétiques associés à l'AA, c'est le locus 17q12-21 qui est l'un des plus associés et répliqués dans différentes études (Bønnelykke et al., 2014; Demenais et al., 2018; Madore et al., 2008; Stern et al., 2020). Cette région contient les gènes *gasdermin B (GSDMB)* et *ORMDL sphingolipid biosynthesis regulator 3 (ORMDL3)* (Stern et al., 2020). Le gène *GSDMB* est impliqué dans la régulation de la pyroptose des cellules épithéliales par le biais d'un clivage par les lymphocytes T cytotoxiques et les cellules T tueuses naturelles (NKT; Das et al., 2017; Zhou et al., 2020). Quant au gène *ORMDL3*, il est impliqué dans la régulation de la production des sphingolipides et dans le repliement des protéines suite au stress du réticulum endoplasmique (Das et al., 2017; Jones et al., 2019; Portelli et al., 2015). L'un des gènes associés à la RA est le gène *mitochondrial ribosomal protein L4 (MRPL4)* situé sur le locus 19p13.2 (B. Y. Choi et al., 2021). Cette association a été identifiée via GWAS dans la population de Singapour en Chine par Andiappan et al. (2011) et elle a été répliquée dans la population Han en Chine par Wei et al. (2013). Ainsi, le gène *MRPL4* est impliqué dans le maintien de l'intégrité des ribosomes, dans la

traduction des protéines mitochondriales et pourrait aussi être impliqué dans les processus d'inflammation allergique (Andiappan et al., 2011). Une autre hypothèse de son implication dans l'allergie provient de sa grande proximité (~ 10,8 kb) avec le gène *intercellular adhesion molecule 1* (*ICAM1*) situé dans le même locus et impliqué dans la régulation de la réaction allergique au niveau du tissu nasal (Andiappan et al., 2011; Wei et al., 2013). Ainsi, le gène *MRPL4* pourrait avoir un impact sur le niveau d'expression du gène *ICAM1* puisque l'un de ses SNP (rs8111930) a déjà été associé à l'expression de la protéine soluble produite par *ICAM1* (Andiappan et al., 2011).

Bien que plusieurs variants génétiques, loci et gènes aient été associés aux diverses maladies allergiques, peu de GWAS se sont concentrées sur la génétique de la sensibilisation allergénique (Bønnelykke et al., 2013). Le Tableau 4 présente un résumé de ces associations. Parmi les gènes identifiés, plusieurs d'entre eux sont impliqués dans des mécanismes inflammatoires. Notamment, les gènes *toll like receptor (TLR) 1 (TLR1)*, *TLR6* et *TLR10* qui sont impliqués dans la reconnaissance de motifs moléculaires lors de la réponse immunitaire contre les pathogènes. Il est suggéré qu'une mutation dans ces gènes peut favoriser le développement de maladies allergiques (Bønnelykke et al., 2013; Tesse et al., 2011). Le gène *TSLP* est impliqué quant à lui dans le développement de l'inflammation de type Th2 par l'activation des CD et des ILC2 lors de la réaction allergique (Figures 2, 4-5 et 7-8; Han et al., 2017; Hill & Spergel, 2018). Les gènes du *major histocompatibility complex*, *class II DQ alpha 1 (HLA-DQA1)* et *HLA-DQB1* sont quant à eux impliqués dans la reconnaissance des allergènes et la réaction immunitaire de type Th2 (Waage et al., 2018). Les gènes *MYC proto-oncogene*, *BHLH transcription factor (MYC)* et *plasmacytoma variant translocation 1 oncogene (PVT1)* seraient notamment impliqués dans l'activation des lymphocytes T (Bønnelykke et al., 2013).

Tableau 4. Loci, gènes et variants associés à la sensibilisation allergénique

Locus	Gène à proximité	Variant	Référence
1q24.3	<i>FASLG</i> et <i>SLC25A38P1</i>	rs78037977	(Waage et al., 2018)
2p25.1	<i>LINC00299</i>	rs891058	(Waage et al., 2018)
2q12.1	<i>IL1RL1</i> , <i>IL1RL2</i> et <i>IL18R1</i>	rs3771175, rs10189629*, rs2160203 et rs950881	(Bønnelykke et al., 2013; Hinds et al., 2013; Šaulienė et al., 2015; Waage et al., 2018)
2q33.1	<i>PLCL1</i>	rs10497813	(Hinds et al., 2013)
3q28	<i>LPP</i>	rs9865818* et rs9860547	(Bønnelykke et al., 2013; Hinds et al., 2013; Waage et al., 2018)
4p14	<i>TLR1</i> , <i>TLR6</i> et <i>TLR10</i>	rs17616434**, rs2101521* et rs5743618	(Bønnelykke et al., 2013; Hinds et al., 2013; Nilsson et al., 2014; Waage et al., 2018)
4q24	<i>NFKB1</i>	rs4648050 et rs12509403	(Waage et al., 2018)
4q27	<i>IL2</i> et <i>ADAD1</i>	rs17454584, rs2069772** et rs17388568*	(Bønnelykke et al., 2013; Hinds et al., 2013; Ramasamy et al., 2011; Šaulienė et al., 2015; Waage et al., 2018)
5p13.1	<i>PTGER4</i>	rs7720838*	(Hinds et al., 2013; Šaulienė et al., 2015)
5q22.1	<i>SLC25A46</i> , <i>TSLP</i> , <i>TMEM232</i> , <i>WDR36</i> et <i>CAMK4</i>	rs10056340*, rs17513503, rs1438673* et rs12657787	(Bønnelykke et al., 2013; Hinds et al., 2013; Nilsson et al., 2014; Ramasamy et al., 2011; Waage et al., 2018)
6p21.33	<i>HLA-B</i> , <i>HLA-C</i> et <i>MICA</i>	rs6932730 et rs9266772	(Bønnelykke et al., 2013; Hinds et al., 2013)
6p21.32	<i>HLA-DQA1</i> et <i>HLA-DQB1</i>	rs6906021** et rs7775228*	(Bønnelykke et al., 2013; Hinds et al., 2013; Nilsson et al., 2014; Ramasamy et al., 2011; Šaulienė et al., 2015)
8q24.21	<i>MYC</i> et <i>PVT1</i>	rs4410871*	(Bønnelykke et al., 2013; Šaulienė et al., 2015)
9p24.1	<i>RANBP6</i> et <i>IL33</i>	rs7032572	(Hinds et al., 2013)
10p14	<i>GATA3</i>	rs962993	(Hinds et al., 2013)

11q13.5	<i>C11orf30</i> et <i>LRRC32</i>	rs2155219***, rs7936070 et rs7936323	(Andiappan et al., 2013; Bønnelykke et al., 2013; Hinds et al., 2013; Ramasamy et al., 2011; Waage et al., 2018)
12q13.3	<i>STAT6</i>	rs1059513**	(Bønnelykke et al., 2013; Šaulienė et al., 2015; Waage et al., 2018)
14q21.1	<i>FOXA1</i> et <i>TTC6</i>	rs1998359*	(Hinds et al., 2013; Nilsson et al., 2014)
15q22.33	<i>SMAD3</i>	rs17228058	(Hinds et al., 2013)
17q12	<i>GSDMB</i>	rs9303280	(Hinds et al., 2013)
20q13.2	<i>NFATC2</i>	rs6021270	(Hinds et al., 2013)

*Variant répliqué dans une autre étude **Variant répliqué dans trois études indépendantes

***Variant répliqué dans quatre études indépendantes

3.3 Épigénétique

L'épigénétique correspond à la modification biochimique de l'ADN qui n'altère pas sa séquence, dont certaines de ces modifications sont modulées par des facteurs environnementaux, et ayant un impact sur l'expression des gènes (Potaczek et al., 2017). L'épigénétique régule cette expression par différents mécanismes, dont l'acétylation, la méthylation, l'ubiquitination et la phosphorylation. Ces modifications se font soit sur l'ADN, sur les histones ou sur les protéines produites afin de les réguler (Potaczek et al., 2017; Wang et al., 2023). Ces marques épigénétiques surviennent tout au long de la vie à commencer par la période prénatale (Ntontsi et al., 2021). La marque épigénétique la plus étudiée est celle de la méthylation de l'ADN. Celle-ci correspond à l'ajout d'un groupement méthyle (-CH₃) sur une cytosine précédant une guanine (CpG). Ce groupe dinucléotidique seul est appelé un site CpG et la majorité d'entre eux sont méthylés (Blomen & Boonstra, 2011; Potaczek et al., 2017). Les régions entre les gènes, dites intergéniques, et les introns, c'est-à-dire les parties ne codant pas pour une protéine dans un gène, sont dits pauvres en CpG (Miranda & Jones, 2007). Par contre, il existe des zones riches en CpG nommés îlots CpG qui sont généralement non méthylés (Blomen & Boonstra, 2011; Potaczek et al., 2017). Ceux-ci sont

normalement situés dans les régions régulatrices des gènes comme le promoteur ou l'amplificateur afin d'en réguler la transcription. Le mécanisme de régulation de cette marque épigénétique fait en sorte que plus un gène est méthylé et moins il est exprimé, et vice versa (Potaczek et al., 2017). La méthylation sert au maintien de l'identité cellulaire, à la régulation de l'expression des gènes et à la formation de l'hétérochromatine (i.e. enroulement de l'ADN condensé rendant impossible la transcription; Blomen & Boonstra, 2011; Miranda & Jones, 2007). De plus, cette marque épigénétique est héritable, réversible et son patron diffère selon le type cellulaire étudié (Blomen & Boonstra, 2011; Ntontsi et al., 2021).

En étudiant l'épigénétique de maladie, il est possible d'identifier les gènes qui sont hypométhylés ou hyperméthylés dans les voies biologiques impliquées dans la maladie étudiée. Ce type d'étude permet de mieux comprendre les mécanismes épigénétiques derrière une maladie et d'identifier de nouvelles voies biologiques impliquées dans la maladie. Notamment, une étude a été capable d'identifier 81 loci différemment méthylés entre des enfants (6 à 12 ans) atteints d'AA et sans AA. Parmi les loci identifiés, 73 étaient hypométhylés et huit étaient hyperméthylés chez les enfants atteints d'AA. De plus, plusieurs gènes qui étaient hypométhylés étaient impliqués dans la voie de signalisation de la réponse immunitaire de type Th2, comme *IL4* et *IL13* (I. V. Yang et al., 2015). Dans le Tableau 5, il est possible de retrouver d'autres gènes pour lesquels une différence de méthylation a déjà été associée à une différence dans son expression pour l'asthme et la RA dans la littérature.

Tableau 5. Exemple de la méthylation impliquée dans l'asthme et la rhinite (Adapté de Nadeau et al., 2010; Ntontsi et al., 2021; Pascual et al., 2011; Tan et al., 2020; Wang et al., 2023; I. V. Yang et al., 2015)

Gène	Méthylation	Association
<i>CYP26A1</i>	Hyperméthylation (promoteur)	Asthme allergique
<i>FOXP3</i>	Hyperméthylation (promoteur et/ou intron)	Asthme et rhinite allergique
<i>IL13</i>	Hypométhylation (exon, intron, 5'-UTR)	Asthme allergique à l'enfance
<i>MEIS1</i>	Hyperméthylation (intron)	Asthme allergique à l'enfance
<i>RUNX3</i>	Hypométhylation (intron)	Asthme allergique à l'enfance
<i>TIGIT</i>	Hypométhylation (exon et 5'-UTR)	Asthme allergique à l'enfance

Abréviation: 5'-UTR= région 5' non traduite

3.3.1 Étude de locus à caractère quantitatif

Les GWAS permettent d'identifier une multitude de variants génétiques associés à des traits complexes, mais la majorité de ces variants, soit environ 90 %, sont situés dans des régions régulatrices ou « non codantes » (B. Li & Ritchie, 2021). Il devient donc difficile d'interpréter leur implication dans la pathogenèse de la maladie. Une autre difficulté d'interprétation des GWAS est que certains des variants identifiés peuvent ne pas être le variant causal de la maladie, mais un variant se retrouvant en déséquilibre de liaison (LD) avec celui-ci. Afin d'identifier des variants et des gènes candidats solides, il est donc nécessaire de combiner des données génomiques à des données transcriptomiques et/ou épigénomiques (Akiyama, 2021). Ainsi, l'étude des QTL, développée pour l'analyse de traits continus, permet de combiner des données de GWAS à des données omiques comme le niveau d'expression (eQTL) et la méthylation (mQTL). Ce type d'étude permet de cibler des mécanismes fonctionnels impliqués dans la pathogenèse de la maladie étudiée par l'association de variants génétiques à une différence d'expression de gènes ou de méthylation de l'ADN (Gleason et al., 2020; Ye et al., 2020). Notamment, Morin et al. (2017) ont été capables d'identifier une association significative pour la RA entre un SNP préalablement identifié par GWAS

(rs888989) et le niveau de méthylation d'un CpG (cg18424208). Cette association était située dans le promoteur du gène *caudal type homeobox 1 (CDX1)*, et en utilisant le portail *Genotype-Tissue Expression (GTEx)*, ils ont été capables d'identifier que l'expression de ce gène était influencée par le SNP identifié (eQTL). Par conséquent, en utilisant des données épigénétiques et en les combinant aux données génétiques obtenues par GWAS (mQTL), les auteurs ont réussi à identifier un gène d'intérêt impliqué dans la RA (Morin et al., 2017). De plus, l'intégration des données de GWAS à plus d'un type de données omiques permet d'obtenir une meilleure interprétation de l'implication des variants génétiques dans la pathogenèse de la maladie (Gleason et al., 2020). Il est à noter que ce type d'étude permet aussi l'identification des mécanismes biologiques propre à des cellules étant donné que les données omiques sont uniques au type cellulaire ou au tissu duquel elles proviennent (Ye et al., 2020). Par la suite, des études d'enrichissement fonctionnel permettront de cibler des gènes ayant un fort impact dans les voies biologiques préalablement identifiées. Dans le cadre de ce projet, les données omiques qui seront combinées aux données génétiques seront des données d'expression des gènes et de méthylation de l'ADN issu de lymphocytes T CD4⁺ naïfs et d'éosinophiles.

CHAPITRE 4

GENETIC RISK PROFILING FOR INDOOR AND OUTDOOR AEROALLERGENS SENSITIZATION IN A FRENCH-CANADIAN ASTHMA FAMILY COHORT

Marie-Pier Bouchard^{1,2}, Christophe Linhard³, Anne-Marie Boucher-Lafleur^{1,2}, Anne-Marie-Madore^{1,2}, Emmanuelle Bouzigon³, Catherine Laprise^{*1,2}

1Centre intersectoriel en santé durable, Université du Québec à Chicoutimi, Saguenay, QC, Canada, G7H 2B1

2Département des Sciences Fondamentales, Université du Québec à Chicoutimi, Saguenay, QC, Canada, G7H 2B1

3UMRS 1124, INSERM, Université de Paris, 75006, Paris, France

Corresponding author:

Catherine Laprise, PhD

Université du Québec à Chicoutimi

555 Boulevard de l'Université, Saguenay, Québec, CANADA, G7H 2B1

Tel: +418-545-5011 ext. 5659

Fax: +418-615-1203

Email: catherine.laprise@uqac.ca

4.1 Avant-propos

Cet article est en cours de correction par les co-auteurs et n'a donc pas encore été soumis dans un journal.

Remerciements : Les auteurs de l'étude souhaitent remercier les participants et leurs familles qui ont été recrutés dans la cohorte d'asthme familial du Saguenay–Lac-Saint-Jean pour leur participation à cette étude. Ils remercient également Lucile Pain pour avoir isolé les lymphocytes T

CD4⁺ naïfs et les éosinophiles à partir d'échantillons sanguins. Les données de GWAS ont été rendues disponibles par la Commission européenne dans le cadre du projet GABRIEL (Étude multidisciplinaire pour identifier les causes génétiques et environnementales de l'asthme dans une communauté européenne, #018996) sous le programme intégré LSH-2004-1.2.5-1 Approche post-génomique pour comprendre la base moléculaire de l'asthme pour le contrôle préventif ou thérapeutique. Cette étude a utilisé des données qui ont été produites par le Consortium UK10K pour l'imputation des génotypes. Une liste complète des chercheurs ayant contribué à la production des données est disponible sur www.UK10K.org. Le financement de la UK10K a été fourni par la Wellcome Trust award WT091310.

Contribution des auteurs à l'article : M.-P. B. : Nettoyage des données, analyses, interprétation des résultats, rédaction du manuscrit ; C. Linhard : Analyses de réplication ; A.-M. B.-L. : Isolation des lymphocytes T naïfs et des éosinophiles, révision et édition du manuscrit; A.-M. M. : Supervision des analyses, révision et édition du manuscrit ; E. B. : Supervision des analyses de réplication ; C. Laprise : Conception, supervision, acquisition du financement, révision et édition du manuscrit.

Financement : Marie-Pier Bouchard a été financé par le Réseau de recherche en santé respiratoire du Québec (RSRQ). Ce projet a été soutenu par une subvention d'opérations des Instituts de recherche en Santé du Canada (IRSC, Catherine Laprise, #133605). Pre Laprise est la directrice de la Chaire de recherche du Canada sur la génomique de l'asthme et des maladies allergiques (<https://www.chairs-chaire.gc.ca/chairholders-titulaires/profile-eng.aspx?profileId=1540>), la co-directrice de la Chaire de recherche en santé durable du Québec, la co-directrice de l'axe de recherche Environnement/Génétiques/Cancer du RSRQ (<https://rsr-qc.ca/environnement-genetique-cancer/>), chercheuse membre de l'étude CHILD et chercheuse membre du réseau des allergies, des gènes et de l'environnement (AllerGen NCE Inc.).

Conflits d'intérêts : Les auteurs n'ont aucun conflit d'intérêts à déclarer.

Matériel supplémentaire : Le matériel supplémentaire peut être retrouvé dans l'annexe 1.

4.2 Résumé

La sensibilisation aux aéroallergènes est un facteur de risque courant pour le développement d'allergies respiratoires. Il est connu que les aéroallergènes d'intérieur sont majoritairement associés à l'asthme allergique (AA), alors que les aéroallergènes d'extérieur sont plus souvent associés à la rhinite allergique (AR). Ainsi, il serait possible d'identifier des variants génétiques communs et distincts entre la sensibilisation aux aéroallergènes d'intérieur et d'extérieur. Ceci a été effectué par la réalisation d'une étude d'association pangénomique pour la sensibilisation aux aéroallergènes intérieurs, extérieurs et combinés dans la cohorte familiale d'asthme du Saguenay–Lac-Saint-Jean. Par la suite, les variants identifiés pour chacun des types de sensibilisation ont été utilisés pour la réalisation d'analyses de locus quantitatif (QTL) pour le niveau d'expression et de méthylation de gènes dans des échantillons de lymphocyte T CD4⁺ naïfs et d'éosinophiles. Les résultats ont montré plusieurs associations significatives ($P < 1E-05$) avec chacun des types de sensibilisation (259 variants dans 16 loci pour les aéroallergènes d'intérieur, 38 variants dans 13 loci pour les aéroallergènes d'extérieur et 57 variants dans 15 loci pour la sensibilisation combinée). Seulement un locus était commun entre le profil génétique de la sensibilisation aux aéroallergènes d'intérieur et d'extérieur. Un variant (rs3798307) pour l'analyse de la sensibilisation aux aéroallergènes d'intérieur était significativement associé avec le niveau d'expression ainsi que le niveau de méthylation d'un même gène dans les lymphocytes T CD4⁺ naïfs. Ceci a aussi été observé pour quatre variants (rs10761191, rs4743880, rs10682558 et rs6479446) pour l'analyse de la sensibilisation combinée. Cette recherche a permis d'identifier un profil génétique spécifique à chacun des types de sensibilisation aux aéroallergènes et a étudié les mécanismes biologiques impliqués derrière ces sensibilisations dans deux types de cellules immunitaires.

Mots clés : Asthme allergique, rhinite allergique, allergie respiratoire, variant génétique, sensibilisation aux aéroallergènes, lymphocyte T CD4⁺ naïfs, éosinophiles, eQTL, mQTL

4.3 Abstract

Aeroallergens sensitization is a common risk factor for allergic respiratory diseases development. It is also known that indoor aeroallergens are majorly associated to allergic asthma (AA) while the outdoor aeroallergens are more associated to allergic rhinitis (AR). So, it could be possible to identify shared and specific genetic variants between the indoor and the outdoor aeroallergens sensitization. This was achieved by doing a genome-wide association study (GWAS) in the Saguenay–Lac-Saint-Jean (SLSJ) asthma family cohort for the sensitization to indoor, outdoor or both types of aeroallergens. Afterwards, the variants identified for each type of sensitization were used for quantitative trait locus (QTL) analyses with gene expression and methylation levels in naïve CD4⁺ T cells and eosinophil samples. Results showed many significant associations ($P < 1E-05$) with each sensitization type (259 variants in 16 loci for the indoor aeroallergens, 38 variants in 13 loci for the outdoor aeroallergens and 57 variants in 15 loci for both types of sensitization). Only one locus was common between the genetic profile of indoor and outdoor aeroallergen sensitization. Interestingly, one variant (rs3798307) for the indoor sensitization analysis was significantly associated with gene expression as well as methylation levels of a same gene in the naïve CD4⁺ T cells. This was also observed for four variants (rs10761191, rs4743880, rs10682558 and rs6479446) for the analysis of both types of sensitization. This research identified a genetic profile specific for each type of aeroallergens sensitization and studied the biological mechanisms behind aeroallergens sensitization in two immune cell types.

Keywords: allergic asthma, allergic rhinitis, allergic respiratory disease, genetic variant, aeroallergen sensitization, naïve CD4⁺ T cells, eosinophils, eQTL, mQTL

4.4 Introduction

During the last few decades, an increase in prevalence of allergic respiratory diseases (ARD) such as allergic asthma (AA) and rhinitis (AR) has been observed around the world. This phenomenon can be explained in part by the industrialization and by climate change that have a strong impact on the respiratory health.¹ Consequences of these parameters, such as air pollution and global warming, all contribute to the augmentation of aeroallergens concentration, the modification of the period and duration of pollination and the changes in pollen distribution.²⁻⁵ Also, air pollution can enhance the allergenicity of certain aeroallergens.^{1,2} All of these factors contribute to an increase in prevalence and exacerbation of ARD, including AA and AR.² They are also associated with an increased risk of developing an outdoor aeroallergens sensitization which can explain the increase of its prevalence.⁶ As for the increased indoor aeroallergens sensitization prevalence, it can be partly explained by the fact that almost 80% of the westernized population spend most of their time indoor.¹ Moreover, the air pollutant levels found indoor can be 100 times higher than the ones found outdoor, which can increase the allergenicity of certain aeroallergens.⁷

Studies have shown that sensitization to indoor aeroallergens were more associated to AA, while outdoor sensitization was more associated to AR.^{8,9} Furthermore, it has been known for a few decades that allergens sensitization is a risk factor for AR and AA development. Many studies have proved this association and tried to identify which aeroallergen was associated with AR or AA.⁹⁻¹² Only a few tried to identify genetic variants associated with the type of allergens sensitization. For example, Ramasamy et al., identified a single nucleotide polymorphism (SNP; rs2155219) that was associated with the risk of grass pollen sensitization ($P=9.4E-09$; $OR=1.22$) and to AR ($P=3.8E-08$; $OR=1.17$).¹³

Thus, it would be possible to identify genetic variants predisposing for an allergenic sensitization to aeroallergens classified according to their origin, either indoor (mite, dust, animal dander), outdoor (pollen, mold, weed) or for both types. This study aims at identifying a genetic profile associated to an allergic sensitization to indoor and outdoor aeroallergens in the Saguenay–Lac-

Saint-Jean (SLSJ) asthma family cohort. To study the biological mechanisms underlying the associations found with aeroallergen sensitization, quantitative trait locus analysis (QTL) for the expression (eQTL) and the methylation (mQTL) levels in two immune cell types, naïve CD4⁺ T cells and eosinophils, were also performed.

4.5 Materials and Methods

For a better understanding of the overall method, a schematic representation is shown in Figure 10. This figure contains a short description of cohorts, analyses, thresholds and covariates used throughout all steps performed.

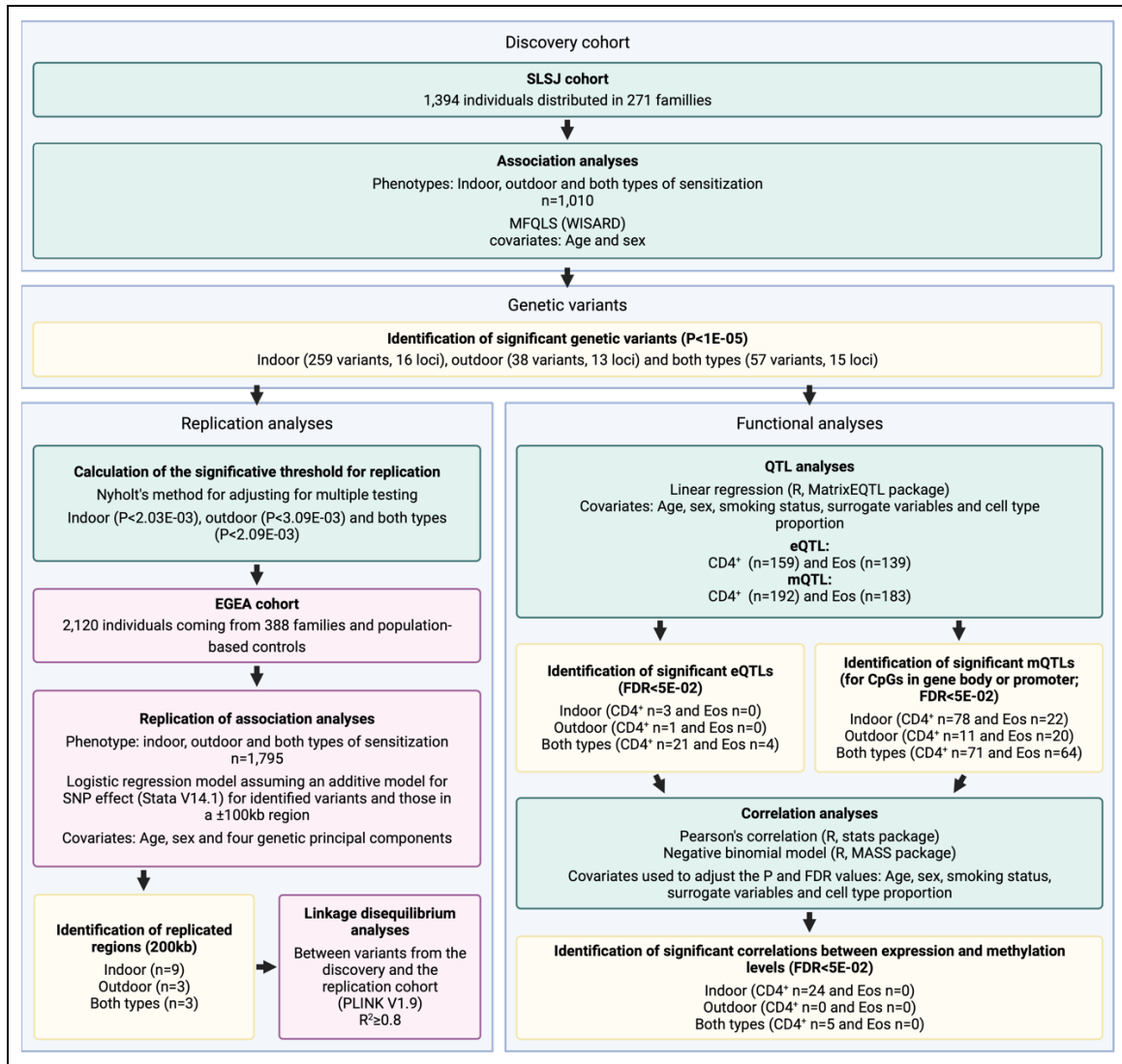


Figure 7. Schematic representation of this study's layout

This figure contains a short description of the discovery and replication cohorts. It also includes the analyses performed, their covariates and the threshold used to identify significant results, as well as the number of significant results obtained for each step. Abbreviation: CD4⁺=Naïve CD4⁺ T cells, EGEA cohort=The French Epidemiological study on the Genetics and Environment of Asthma cohort, Eos=Eosinophils, FDR=False discovery rate, MASS=Modern applied Statistics with S, MFQLS=Multiple Family-based Quasi-likelihood Score, QTL=Quantitative trait locus, SLSJ cohort=The Saguenay–Lac-Saint-Jean asthma family cohort, WISARD=Workbench for Integrated Superfast Association Study with Related Data (Created with BioRender®, 2023).

4.5.1 Discovery sample

The sample used for this study came from the SLSJ asthma family cohort that includes data for 1,394 individuals distributed in 271 independent families from northeastern Québec in Canada. This cohort has a well described phenotypic profile for each of its participant with more than 75 phenotypic characteristics about general health, respiratory health and allergic status.¹⁴ Individuals recruited in this cohort are from a young founder population with European ancestry characterized by a founder effect, making it easier to study complex traits such as ARD.¹⁵ Individuals with AA were recruited through probands, then through their parents and grandparents if it was possible. All probands needed to have four grandparents that originated from the SLSJ. More information on recruitment, clinical evaluation and phenotypic data from the SLSJ cohort can be found in Laprise 2014.¹⁴ From the 1,394 individuals in this cohort, 1,198 have imputed genotypic data available.^{16,17} Among them, 215 individuals were recruited for naïve CD4⁺ T cells and eosinophils isolation to measure expression and methylation levels by next generation sequencing.¹⁶ After filtering, a total of 1,010 individuals with genotypic and phenotypic data remained for genome-wide association study (GWAS) and a total of 192 and 183 samples remained for QTL analyses in the naïve CD4⁺ T cells and eosinophil samples respectively. Out of the 1,010 individuals in this cohort, 564 (56%) had an indoor sensitization, 445 (44%) had an outdoor sensitization and 370 (37%) had both types of sensitizations. Their relevant phenotypic characteristics are presented in Table 6 (see supplementary Table 9 for more details on their relevant phenotypic characteristics). All participants gave informed written consent, and all experimental procedures were approved by the *Centre intégré universitaire de santé et de services sociaux du SLSJ* ethics committee (projects #0002-001 and #2015-035).

Table 6. Phenotypic characteristics of individuals included in the association analyses for the indoor, the outdoor and both types of aeroallergens sensitization

	Indoor		Outdoor		Both types	
	Non-allergic (n= 446)	Allergic (n= 564)	Non-allergic (n= 565)	Allergic (n= 445)	Non-allergic (n= 640)	Allergic (n= 370)
M:F ratio	1.00:1.32	1.00:0.99	1.00:1.28	1.00:0.95	1.00:1.26	1.00:0.92
Age, mean (range)	38 (0-85)	29 (2-79)	36 (0-85)	29 (4-79)	36 (0-85)	28 (5-79)
Age, median	40	28	39	30	39	28
IgE, mean (SD)¹	60 (4)	258 (4)	79 (4.21)	268 (4)	83 (4)	316 (4)
Atopy, n (%)²	81 (18)	564 (100)	200 (35)	445 (100)	275 (43)	370 (100)
Indoor allergy, n (%)	0 (0)	564 (100) ^a	194 (34)	370 (83) ^b	194 (30)	370 (100) ^a
Outdoor allergy, n (%)	75 (17)	370 (66) ^a	0 (0)	445 (100) ^b	75 (12)	370 (100) ^b
Both types of allergies, n (%)	0 (0)	370 (66) ^a	0 (0)	370 (83) ^b	0 (0)	370 (100) ^c
Other allergy, n (%)³	15 (4)	100 (21)	17 (4)	98 (26)	26 (5)	89 (29)
Atopic dermatitis, n (%)	31 (7)	260 (46)	77 (14)	214 (48)	106 (17)	185 (50)
Dermatitis, n (%)⁴	155 (35)	260 (46)	201 (36)	214 (48)	230 (36)	185 (50)
Asthma, n (%)⁵	174 (39)	371 (66)	255 (46)	290 (65)	293 (46)	252 (68)
Rhinitis, n (%)⁶	74 (17)	255 (45) ^a	113 (20)	216 (49) ^{a, b}	130 (20)	199 (54) ^b
Rhinitis + asthma, n (%)	44 (10)	182 (33)	74 (13)	152 (34)	84 (13)	142 (39)
Sensitization to aeroallergens						
Animal allergy, n (%)	0 (0)	431 (76) ^a	125 (22)	306 (69) ^b	125 (20)	306 (83) ^a
House dust allergy, n (%)	0 (0)	240 (43) ^a	49 (9)	191 (43) ^a	49 (8)	191 (52) ^b
House dust mite allergy, n (%)	0 (0)	389 (69) ^{a, b}	112 (20)	277 (62) ^a	112 (18)	277 (75) ^b
Grass pollen allergy, n (%)	37 (8)	286 (51) ^a	0 (0)	323 (73) ^b	37 (6)	286 (78) ^b

Tree pollen allergy, n (%)	43 (10)	282 (50) ^a	0 (0)	325 (73) ^b	43 (7)	282 (76) ^b
Mold allergy, n (%)	14 (3)	113 (20) ^a	0 (0)	127 (29) ^b	14 (2)	113 (31) ^b
Allergic monosensitization, n (%)	38 (9)	66 (12) ^a	72 (13)	32 (7) ^b	104 (16)	0 (0) ^c
Allergic polysensitization, n (%)	43 (10)	498 (88) ^a	128 (23)	413 (93) ^b	171 (27)	370 (100) ^c
Number of sensitizations	1-10	1-28	1-7	1-28	1-10	1-28
One category of allergens, n (%)	57 (13)	110 (20) ^a	116 (21)	51 (11) ^b	167 (26)	0 (0) ^c
Two categories of allergens, n (%)	20 (4)	111 (20) ^a	66 (12)	65 (15) ^{a, b}	86 (13)	45 (12) ^b
Three categories of allergens, n (%)	4 (1)	100 (18)	17 (3)	87 (20)	21 (3)	83 (22)
Four categories of allergens, n (%)	0 (0)	97 (17) ^a	1 (0)	96 (22) ^{a, b}	1 (0)	96 (26) ^b
Five categories of allergens, n (%)	0 (0)	79 (14) ^a	0 (0)	79 (18) ^{a, b}	0 (0)	79 (21) ^b
Six categories of allergens, n (%)	0 (0)	43 (8)	0 (0)	43 (10)	0 (0)	43 (12)
Seven categories of allergens, n (%)⁷	0 (0)	24 (4)	0 (0)	24 (5)	0 (0)	24 (6)

The significant difference between each group is noted with the letters a, b and c. If phenotype have different letters, it means there was a statistical difference between them. If groups share the same letter or none of the groups have a letter it means that there was no statistical difference between them. ¹Geometric mean and standard deviation (SD) of blood immunoglobulin E (IgE) concentration. ²Defined as at least one positive response on skin prick testing (wheal diameter ≥ 3 mm more than the negative control at 10 min). ³The other allergy category includes all food allergens that were tested in the SLSJ cohort. ⁴Defined as personal history of urticaria, eczema or both. ⁵Present or past documented clinical history of asthma. ⁶Present or past personal history of rhinitis. ⁷The seventh category of allergens correspond to the food allergens.

All statistical comparisons for the phenotypic characteristics were made using the SPSS statistics software (version 29.0.1.0). Comparisons were executed between the non-allergic and allergic groups within each sensitization type and between the allergic group of each sensitization type. All comparisons for categorical variables were made using a chi-square test and the significant differences between groups were identified using the Bonferroni adjustment. Since the continuous variables had a non-parametric distribution, their comparisons were made using the Kruskal-Wallis test for non-parametric data and a multiple comparison test for post-hoc analysis.

4.5.2 Replication sample

The replication sample came from the French Epidemiological study on the Genetics and Environment of Asthma (EGEA).¹⁸ This cohort is a family-based study on asthma combined with a case-control study made of three surveys in a period of 20 years (1991-2013). The whole cohort is built from 388 asthmatic probands and their 1,317 immediate family members and 415 population-based controls for a total of 2,120 individuals. The EGEA cohort is a well described cohort for which all participants have an extensive phenotypic characterization (standardized questionnaires, physiological and biological tests, data on lifestyle and environment) as described elsewhere.¹⁸ The sample size for each analysis was 1,795 individuals for the indoor and 1,789 individuals for the outdoor and both types of aeroallergens sensitization. The number of cases in this cohort was 40% (720/1,795) for indoor sensitization, 35% (635/1,789) for outdoor sensitization and 24% (431/1,789) for both types of sensitizations.

4.5.3 Atopic status

Atopy was considered if there were one or more positive response(s) in the skin prick test (SPT). To be positive, the induration for an allergen had to be ≥ 3 mm than the negative control after 10 minutes of testing. All tests were done during the absence of any allergic symptoms during the day-to-day life of the participants and each participant had to stop the use of anti-histaminic 72 hours

prior to the SPT. Twenty-five common aeroallergens, divided into six categories, were tested for the SLSJ asthma family cohort: 1) animal (cat, dog, horse and cattle dander and bird feathers); 2) house dust; 3) house dust mite (HDM; *Dermatophagoides [D] farinae* and *D. pteronyssinus*); 4) grass pollen (grasses, weeds, ragweed, timothy and ryegrass); 5) tree pollen (tree mix, birch, maple, oak and elm); 6) mold (*Cladosporium*, *Alternaria*, *Aspergillus* and *Penicillium*). There was also a seventh category of allergen which was composed of 16 food allergens, but that category was not used in any of the analyses for aeroallergens sensitization. More information on each category of aeroallergens tested can be found in Table 6 (see supplementary information section Allergens and supplementary Table 10 for more information on each tested aeroallergen). Eleven aeroallergens were tested in the EGEA cohort which are cat dander, *D. pteronyssinus*, *Blattella germanica* (German cockroach), olive birch pollen, *Parietaria judaica* (spreading pellitory) pollen, timothy grass pollen, ragweed pollen, *Aspergillus*, *Cladosporium herbarum* and *Alternaria tenuis*.

To be considered sensitized to indoor aeroallergens in the SLSJ cohort, it was necessary to have a positive reaction to at least one of the aeroallergens from the animal, dust or HDM categories. The indoor sensitization in the replication cohort was considered for a positive reaction to one of the following allergens: cat, *D. pteronyssinus* and *Blattella germanica*. To be considered sensitized to outdoor aeroallergens in the SLSJ cohort and the EGEA cohort, it was necessary to have a positive reaction to at least one of the aeroallergens from the grass pollen, tree pollen or mold category. For both cohorts, sensitization for both types of aeroallergens was considered when at least one positive reaction was present for both indoor and outdoor aeroallergen types.

4.5.4 Genotyping and imputation

For each participant, a blood sample was collected and the DNA extraction was performed using the Blood and Cell Midi Kit following the manufacturer's instructions (Qiagen, Valencia, CA, USA). Genotypes were extracted from genome-wide data obtained by Illumina Human610-Quad BeadChip (Illumina, San Diego, CA, USA). Afterward, these data were imputed with the UK10K and the 1000 Genomes reference sequences using Shapeit²¹⁹ and Impute²²⁰ softwares. To be included

in the analyses, the imputed data had to have an imputation quality score ≥ 0.9 , a minor allele frequency (MAF) ≥ 0.01 , a P value for Hardy-Weinberg equilibrium $\geq 1E-05$ and a genotype and individual call rates $\geq 95\%$. After quality control, the data of 7,829,249 variants was retained for analysis. More information on genotyping, quality control filtering and imputations of the SLSJ data can be found in Madore et al.¹⁶

The EGEA samples were genotyped using the Illumina Human610-Quad BeadChip (Illumina, San Diego, CA, USA). Genotype imputations were carried out using the Minimac4²¹ software and the 1000 Genomes phase 3 reference panel. To be included in the analyses, the imputed data had to have an imputation quality score ≥ 0.5 and $MAF \geq 0.02$. After quality control a total of 7,040,000 variants were available for analysis.²²

4.5.5 Association analyses

A GWAS with the three phenotypes was carried out in the discovery cohort by using the Multiple Family-based Quasi-likelihood Score (MFQLS)²³ implemented in the Workbench for Integrated Superfast Association Study with Related Data (WISARD; Version 1.3.2)²⁴ toolkit. MFQLS is a powerful quasi-likelihood method of estimation extended from the Cochran-Armitage test that considers kinship coefficients and is suited for family cohorts with extended pedigrees.^{23,24} Considering the sample size and the complex trait studied for this project, variants with P values under the suggestive significance threshold of $1E-05$ were identified as significant. The analyses included age and sex as covariates. No principal component was used for the analyses since all of the individuals from the SLSJ cohort came from the same population. Manhattan plots were built using the qqman package (version 0.1.8)²⁵ in R (version 3.5.1).

4.5.6 Replication analyses

As association signals in different cohorts may be driven by different genetic variants, regions of 200kb were defined around associated genetic variants or clusters of associated genetic variants

that are proximal to each other (see supplementary information section Regions for replication). The analyses were done on all genetic variants located in these 200kb regions using a logistic regression model assuming an additive effect for the SNP. The covariates used were age, sex and the first four genetic principal components computed using EIGENSTRAT2.0²⁶ software to adjust for the population stratification.²⁶ The familial dependency was considered using the cluster and robust options of the regress function in Stata (v14.1)²⁷. Testing of SNP effect was based on a Wald-test.

The significance threshold used for the replication analysis was calculated with Nyholt's method²⁸ using PLINK1.9 (www.cog-genomics.org/plink/1.9/)²⁹ and R. This approach was used to calculate the threshold for each sensitization type according to the number of independent genetic variants associated in the discovery cohort.²⁸ The significance thresholds obtained for the indoor, outdoor and both types of sensitization were $P < 2.03E-03$, $P < 3.09E-03$ and $P < 2.09E-03$ respectively.

Linkage disequilibrium (LD) analysis was calculated between the significant variants found in the SLSJ cohort and the ones validated for the same 200kb region in the EGEA cohort. This was done using the pairwise LD measures for multiple SNP in PLINK1.9 with $r^2 \geq 0.8$ as threshold.³⁰

4.5.7 eQTL and mQTL analyses

The eQTL and mQTL analyses were performed in naïve CD4⁺ T cells and eosinophils to get a better understanding of the genetic mechanisms behind associations for the three types of aeroallergen sensitization. Naïve CD4⁺ T cells were defined as CD3⁺, CD4⁺, CD45RA⁺, CD45RO⁻ cells. These cells were chosen for these analyses because of their known involvement in asthma and allergy.³¹ Immune cells were isolated from 200 ml blood samples as described in Madore et al (see supplementary information section Immune cell isolation).¹⁶ Phenotypic characteristics of the individuals for which expression and/or methylation data was available for the naïve CD4⁺ T cells and eosinophils can be found in the supplementary Table 9. For a complete description of sequencing and data preparation (alignment, quality control and filtering) for methylation and expression data refer to supplementary information section, DNA methylation sequencing and levels and RNA

sequencing and levels, see also Madore et al.¹⁶ Prior to analyses, the gene expression counts were normalized on the library size and log2 transformed using DESeq2³² package in R. Methylation level at cytosine preceding a guanine (CpG) sites undertook an arcsine square root transformation.

The eQTL and mQTL analyses were carried out using the package MatrixEQTL (version 2.3)³³ in R. The covariates used were age, sex, smoking status, first five surrogate variables and proportion of naïve CD4⁺ T cells or eosinophils depending on which sample type was used for analyses. For more information on the calculation of the surrogate variables and the cell type proportion, please see supplemental data section Cell-type contribution and surrogate variables. The QTL analyses were performed for cis-eQTLs and cis-mQTLs (distance between a gene or a CpG and a genetic variant $\leq 1,000\text{kb}$) and for trans-eQTLs and trans-mQTLs (distance between a gene or a CpG and a genetic variant $> 1,000\text{kb}$). Associations between genetic variants and gene expression or CpGs were considered significant for a false discovery rate (FDR) ≤ 0.05 . The R package CHIPpeakAnno (version 3.30.1)^{34,35} was used to find the nearest gene for the CpGs included in significant mQTLs. To be considered linked to a specific gene, a CpG had to be located within the gene body or in the promoter, which was defined as the region located within 1.5kb from the gene transcription start site (TSS).

Correlation analyses between gene expression and DNA methylation levels included in significant eQTLs and mQTLs were made with the Pearson's correlations. Regressions using a negative binomial model were also performed using the R package MASS (version 7.3.50)³⁶. The same covariates than the eQTL and mQTL analyses were used. Correlations were considered significant for a $\text{FDR} \leq 0.05$.

4.6 Results

The purpose of this study was to identify a specific genetic profile for each type of aeroallergen sensitization studied, which were: 1) indoor aeroallergens, 2) outdoor aeroallergens and 3) both types of aeroallergens. Association analyses identified significant associations for 259, 38 and 57 variants

($P < 1E-05$) for the sensitization to indoor, outdoor and both types of aeroallergens respectively. Among them, no variant was common between the indoor and outdoor sensitization. However, the analysis for both types of sensitizations shared two and 11 variants with the analysis on indoor and outdoor sensitization respectively. Considering the results for the indoor sensitization, 229 of the 259 genetic variants identified were located in the locus 5q12.3 and 59 of them were in or near (± 10 kb) the ring finger protein 180 (*RNF180*) gene in this locus. A description of all variants associated can be found in the supplementary Tables 11 to 13.

Manhattan plots obtained for the indoor (Figure 11A) and the outdoor (Figure 11B) sensitization highlighted the differences in genetic association patterns. Genetic variants identified for the indoor and outdoor sensitization were located in 16 and 13 loci respectively with only one shared locus between both analyses, the 9q31.3. The genetic variants associated in the analysis for both types of aeroallergens sensitization were located in 15 loci (Figure 11C). From these, the 9q31.1 locus was identified and thus was common to all analyses. Three other loci, 5q12.3, 8q12.1 and 13q14.12, were common with the analysis for indoor aeroallergens. Another three loci, 2p21, 8p23.1 and 15q11.2, were shared with the analysis for outdoor aeroallergens (Figure 12).

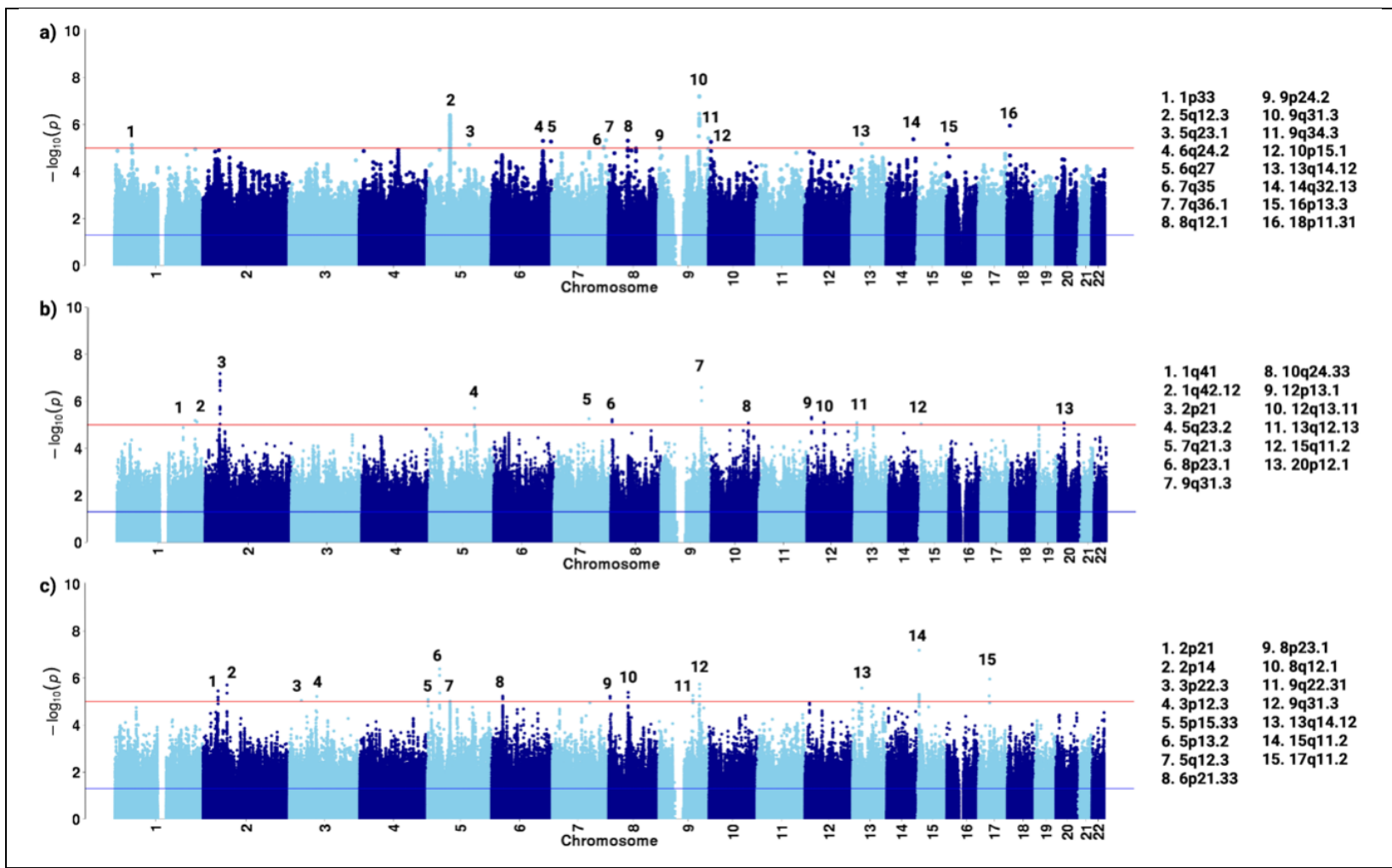


Figure 8. Manhattan plot of significant loci associated ($P < 1E-05$) with a) the indoor sensitization, b) outdoor sensitization, and c) both types of sensitization

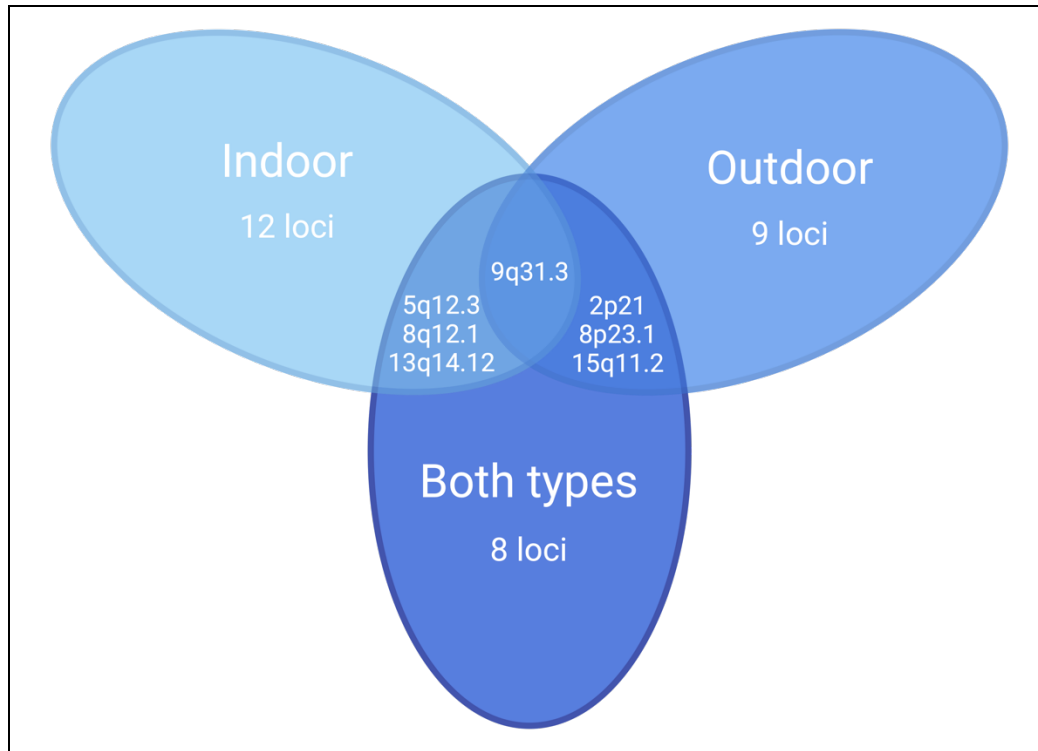


Figure 9. Venn diagram of the shared loci between the sensitization to indoor, outdoor and both types of aeroallergens (Created with BioRender®, 2023)

A replication analysis was performed in the EGEA cohort for each sensitization type considering a 200kb region around each genetic variant or cluster of genetic variants associated in the discovery cohort. None of the associated variants identified for each sensitization type in the SLSJ cohort were directly replicated in the EGEA cohort according to the significance threshold calculated (supplementary Table 11-13). However, it was possible to replicate nine 200 kb regions for the indoor aeroallergen sensitization that were located in eight loci (1p33, 5q23.1, 7q35, 8q12.1, 9p24.2, 9q31.3, 14q32.13, a16p13.3). It was also possible to replicate three regions for the outdoor aeroallergens sensitization in loci 12p13.1, 13q12.13 and 15q11.2 and three regions for both types of aeroallergens sensitization located in loci 2p14, 5q12.3 and 6p21.33 (Table 7). A total of 167 out of the 169 variants passing the replication threshold for both types of sensitizations in the replication cohort were located in the same locus, 6p21.33.

Table 7. 200kb regions replicated in the EGEA cohort and corresponding variant's information for the best association with indoor, outdoor and both types of aeroallergens sensitization

Associated variant by 200kb regions in SLSJ ^a	Number of associated variants in EGEA ^b	Best variant in EGEA ^c	SE ^c	OR ^c	P value ^c	Locus ^c	Nearest gene ^d	Variant information ^e	Distance to the nearest variant identified in SLSJ ^f
Indoor									
rs75285112	4	rs80076649	3.18E-01	3.10	3.79E-04	1p33	<i>TRABD2B</i>	Intron variant	5.7kb
rs370819375	3	rs201018956	3.67E-01	2.91E-01	7.56E-04	5q23.1	<i>LINC02208</i>	Intron variant	-1.8kb
rs12531771	1	rs147759759	4.00E-01	3.58	1.44E-03	7q35	–	Intergenic variant	-33.0kb
rs6998445	4	rs1436001	8.18E-02	7.64E-01	1.02E-03	8q12.1	–	Intergenic variant	-43.5kb
rs10979982	2	rs186338108	4.13E-01	2.43E-01	6.05E-04	9q31.3	–	Intergenic variant	39.5kb
rs10980680/ rs10980683/ rs1121570/ rs1030006/ rs2110548/ rs723591/ rs12345138/ rs976181/ rs7039431/ rs141132389/ rs62573262/ rs10817132/ rs35663630/ rs10980699	2	rs10817138	8.25E-02	7.65E-01	1.17E-03	9q31.3	–	Intergenic variant	-48.1kb
rs1331456	1	rs112780050	3.74E-01	3.55	6.92E-04	9p24.2	<i>VLDLR-AS1</i>	Intron variant	-62.3kb

rs1599965	1	rs117585110	3.52E-01	3.50	3.77E-04	14q32.13	<i>SYNE3</i>	Intron variant	-23.3kb
rs13336442	3	rs10903009	1.05E-01	1.41	9.17E-04	16p13.3	<i>CCDC154, CLCN7</i>	1.5kb upstream variant, exon variant	43.7kb
Outdoor									
rs55750092/ rs76153631/ rs78376850/ rs55954298/ rs55816732/ rs55928877/ rs4763934/ rs55729386	3	rs74816398	2.82E-01	3.07	6.94E-05	12p13.1	<i>FAM234B</i>	4.03kb upstream variant	59.5kb
rs1924783/ rs7993071/ rs17082242/ rs7336497/ rs7320645	3	rs117347415	2.37E-01	2.04	2.61E-03	13q12.13	<i>ATP8A2</i>	Intron variant	-27.5kb
rs28788274	1	rs149013529	3.22E-01	2.65	2.51E-03	15q11.2	–	Intergenic variant	-32.6kb
Both types									
rs75269698/ rs76173900	1	rs114055510	2.97E-01	2.56	1.56E-03	2p14	–	Intergenic variant	21.0kb
rs6449729	1	rs148849598	4.01E-01	3.44	2.06E-03	5q12.3	<i>RGS7BP</i>	8.33kb upstream variant	-90.6kb

rs13437096/ rs7756426/ rs79028790/ rs113473106	167	rs2596501	8.48E-02	7.22E-01	1.20E-04	6p21.33	<i>HLA-B,</i> <i>MIR6891</i>	Intron variant, 1.79kb downstream variant	34.1kb
---	-----	-----------	----------	----------	----------	---------	---------------------------------	---	--------

^aAll significant ($P < 1E-05$) variants identified in the SLSJ cohort grouped by 200kb regions delimited for replication analyses. ^b Number of variants with significant association in the EGEA cohort for each 200kb region (significant thresholds: $P_{\text{indoor}} < 2.03E-03$, $P_{\text{outdoor}} < 3.09E-03$ and $P_{\text{both types}} < 2.09E-03$) ^cVariants located in the 200kb region with the best P value in the EGEA cohort and its association values. ^dGene in which the variant with the best association is located in or nearest to (≤ 10 kb on the right and left sides) according to the UCSC Genomes browser (hg19). For variant which are intergenic (further than 10kb), no information is provided. ^eVariant position toward the gene according to UCSC. ^fDistance between the variant with the best association in the EGEA cohort and to the closest significant variant identified in the SLSJ cohort for this locus. If the EGEA variant is situated after the closest SLSJ variant, the distance number will be negative. However, the variant with the best association in the EGEA cohort might not be the nearest one to all of the SLSJ variants for this locus. Abbreviation: EGEA=The French Epidemiological study on the Genetics and Environment of Asthma cohort, OR=Odd ratio, SE=Standard error, SLSJ=The Saguenay–Lac-Saint-Jean asthma family cohort.

Afterwards, a LD analysis was done to assess any linkage between significant variants identified in the SLSJ cohort and variants passing the replication threshold in the EGEA cohort for corresponding 200kb regions of replication because it was expected that not the same variant would be found in both cohorts. LD analyses of the 6p21.33 locus for both types of aeroallergens sensitization identified four variants in the SLSJ cohort (rs13437096, rs7756426, rs79028790 and rs113473106) that were in LD ($r^2=1$), with one variant identified in the EGEA cohort (rs7751923). Those genetic variants were located from 31,355,287bp to 31,361,058bp on chromosome 6 and were 1.0kb from the 3' end of the MICA antisense RNA 1 (*MICA-AS1*) gene and 6.5kb from the 5' end of the MHC class I polypeptide-related sequence A (*MICA*) gene. The analyses showed that none of the variants identified in the SLSJ cohort were in LD with a variant associated in the EGEA cohort for any of the 200kb regions identified for the indoor and outdoor aeroallergens sensitization.

The nearest genes of each genetic variants associated in the SLSJ cohort in the three analyses were identified and classified into biological functions using the Protein Analysis Through Evolutionary Relationships (PANTHER 17.0)³⁷ classification system, GeneCards (<https://www.genecards.org>)^{38,39} and UniProt⁴⁰ (Figure 13, supplementary Table 14-16). As shown in Figure 13, each sensitization type had at least one gene associated with immune signalization and response. For the indoor sensitization, these genes were the ribonuclease T2 (*RNASET2*), the lysophosphatidic acid receptor 1 (*LPAR1*) and the small nucleolar RNA host gene 7 (*SNHG7*; supplementary Table 14). For the outdoor sensitization, the only gene was the histone deacetylase 7 (*HDAC7*; supplementary Table 15). As for both types of sensitizations, the genes were the C1q and TNF related 3 (*C1QTNF3*), the transmembrane protein 98 (*TMEM98*) and the FYVE, RhoGEF and PH domain containing 3 (*FGD3*; supplementary Table 16).

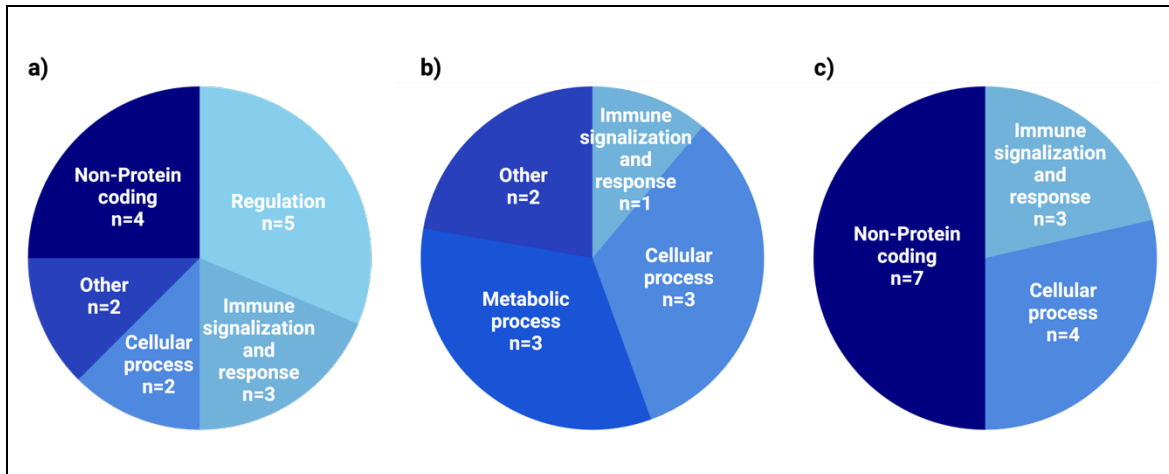


Figure 10. Biological functions of genes identified by GWAS for the a) indoor, b) outdoor and c) both types of aeroallergens sensitization

Genes identified were the ones proximal to a genetic variant (variant in gene body or $\leq 10\text{kb}$ from gene) significantly associated in one of the three GWAS. They were classified by biological function by the Protein Analysis Through Evolutionary Relationships (PANTHER 17.0) classification system, GeneCards and UniProt.

To get a better understanding of the biological mechanisms behind the aeroallergen sensitization, eQTL and mQTL analyses were performed in naïve CD4^+ T cells and eosinophils for genetic variants associated to each sensitization type in the SLSJ cohort. For the sensitization to indoor aeroallergens, there were three significant cis-eQTLs and 78 cis-mQTLs in the naïve CD4^+ T cells and 22 significant cis-mQTLs in the eosinophils (supplementary Tables 17, 18). For the sensitization to outdoor aeroallergens, there were one cis-eQTL and 11 cis-mQTLs in the naïve CD4^+ T cells and 20 significant cis-mQTLs in the eosinophils (supplementary Tables 17, 19). As for the combination of both types of sensitizations, there were 21 significant cis-eQTLs and 71 cis-mQTLs in the naïve CD4^+ T cells and four significant cis-eQTLs and 64 cis-mQTLs in the eosinophils (supplementary Tables 17, 20).

By combining the eQTL and mQTL results (Table 8), it was possible to identify one variant (rs3798307) associated to the sensitization to indoor aeroallergens for which there was a significant association with gene expression as well as methylation (23 and 1 CpG(s) respectively) levels for the *RNASET2* gene and for the microRNA 3939 (*MIR3939*) gene in the naïve CD4^+ T cells. It was also

possible to identify four variants (rs4743880, rs6479446, rs10682558 and rs10761191) associated to both types of aeroallergens that showed significant associations with gene expression as well as methylation (15 CpGs) levels for the *FGD3* gene in the naïve CD4⁺ T cells. Furthermore, one of those variants (rs10761191) was also associated with changes in gene expression as well as methylation (1 CpG) level for the centromere protein P (*CENPP*) gene in the naïve CD4⁺ T cells.

Table 8. GWAS and cis-QTL information on variants associated with significant eQTL and mQTL targeting the same gene for which there is a significant correlation between gene expression and methylation levels in naïve CD4⁺ T cells

Variant ID	Variant position ^a	P value GWAS	Gene	FDR eQTL	Number of CpGs	Best CpG's position ^b	Best CpG's position in gene ^c	Best CpG's FDR mQTL ^d	rPearson ^d	FDR Corr ^d
Indoor sensitization										
rs3798307	6:167,363,948	5.31E-06	<i>MIR3939</i>	1.82E-04	1	6:167,411,355	Body	1.60E-02	-0.545	9.47E-16
			<i>RNASET2</i>	7.14E-07	23	6:167,365,888	Body	6.78E-06	-0.596	1.43E-37
Both types of sensitization										
rs10761191	9:95,781,040	5.51E-06	<i>CENPP</i>	3.52E-03	1	9:95,140,430	Body	4.80E-02	-0.543	1.55E-16
			<i>FGD3</i>	7.56E-03	4	9:95,796,054	Body	6.60E-13		
rs4743880	9:95,777,594	8.44E-06	<i>FGD3</i>	1.13E-02	4	9:95,796,054	Body	2.27E-12	0.259	9.63E-03
rs10682558	9:95,778,654	8.63E-06	<i>FGD3</i>	1.13E-02	4	9:95,796,054	Body	2.27E-12		
rs6479446	9:95,782,991	8.82E-06	<i>FGD3</i>	4.61E-03	4	9:95,796,054	Body	7.27E-13		

^aGenetic variant's position on the genome according to the human genome reference 19 (hg19). ^bGenomic position of the CpG with the lowest FDR in the mQTL result according to the hg19. ^cPosition of the CpG with the best association and located in the targeted gene (body or promoter) according to the hg19. ^dValues indicated here are for the CpG with the best association. Abbreviations: Corr=Correlation, CpGs=Cytosine preceding Guanine sites, eQTL=expression quantitative trait loci, FDR=False discovery rate, GWAS=Genome-wide association study, mQTL=methylation quantitative trait loci.

Correlation analyses were performed between expression and methylation levels for *RNASET2*, *MIR3939*, *FGD3* and *CENNP* genes in naïve CD4⁺ T cells. The correlation analyses showed significant results for 23, one, four and one CpGs that were significantly correlated to the gene expression levels of *RNASET2*, *MIR3939*, *FGD3* and *CENNP* genes respectively (Table 8, supplementary Table 21). Thus, these results showed interactions between genetic variants as well as expression and methylation levels of same genes.

4.7 Discussion

The principal objective of this study was to compare genetic profiles for the allergic sensitization to indoor, outdoor and both types of aeroallergens in the SLSJ cohort. First, GWAS analyses for each sensitization type was performed in order to achieve that aim. These analyses identified 16 (259 variants), 13 (38 variants) and 15 (57 variants) loci with variants associated to the indoor, the outdoor and both types of aeroallergens sensitization respectively. Among them, only the 9q31.1 locus was common to each aeroallergen sensitization type, showing specificity of the genetic profile identified for each of them. Moreover, combining genomics, epigenomics and transcriptomics data in immune cell types also allowed to partly describe the genetic mechanisms underlying the association of one genetic variant with indoor sensitization and four variants with both types of sensitization.

None of the identified variants in this study had previously been associated to AA, AR or other ARD in the literature. However, 11 out of the 37 proximal genes (± 10 kb from genetic variant) and 33 out of the 36 loci in which associated variants were located were already associated to allergic diseases and/or ARD in the literature. Among the variants associated to the indoor sensitization, an association signal carried by 229 genetic variants located in a 394,849bp region of the 5q12.3 locus was identified. Fifty-nine of these genetic variants were located near (≈ 10 kb) or in the only gene of this region, *RNF180*. Interestingly, Morin et al. already found a significant association with a variant (rs7713289) in the *RNF180* gene and AR in the SLSJ cohort. Furthermore, they found that this variant had a significant cis-mQTL result with one CpG site (cg17370163), located in the *RNF180* gene's

promoter, which had been significantly associated to AR through an epigenome-wide association study on whole blood samples in the SLSJ cohort.⁴¹ In this study, QTL analyses were performed on eosinophils and naïve CD4⁺ T cells. One variant (rs356593) also had a significant cis-mQTL with another CpG site located in the *RNF180* gene's promoter in the naïve CD4⁺ T Cells (supplementary Table 18). This variant was different than the one identified by Morin et al. and were 179kb apart, even if both seemed to modulate methylation levels in the promoter of *RNF180*.⁴¹

The replication analysis allowed validation of 15 out of the 45 200kb regions surrounding genetic variants identified in the SLSJ cohort. These regions were located in 14 loci and only one of these was not among the 33 that were already associated to allergic disease and/or ARD in the literature. These results strengthened the implication of these loci in the development of an allergic sensitization. None of the variants associated in the SLSJ cohort were directly replicated in the European cohort. However, one of the replicated 200kb regions, which is located in the 6p21.33 locus, had variants in LD between the SLSJ and the EGEE cohorts. An association between this 200kb region and allergic sensitization had previously been identified by Bønnelykke et al. for common food and inhaled allergens (rs6932730) and by Hinds et al. for self-reported pollen, dust-mite and cat allergies (rs9266772).^{42,43} In this study, four variants of this 200kb region in the SLSJ cohort and 167 in the EGEE cohort were associated with both types of aeroallergens sensitization. Interestingly, the variant rs6932730 was among the 167 variants identified in the EGEE cohort. It is important to note that the 6p21.33 locus contains many genes like the major histocompatibility complex, class I, B (*HLA-B*) and *MICA* which are involved in immune function.⁴³ This data shows the importance of the 6p21.33 locus as a candidate for the study of allergic sensitization, without specificity to indoor or outdoor aeroallergens. It should be further studied to determine which genes are implicated in the process of sensitization.

The excellent phenotyping of the aeroallergen sensitization of each individual from the SLSJ cohort is one of this study's strength. The individuals of this cohort had to undergo a SPT of 24 aeroallergens. This yielded a well-defined sensitization phenotype compared to other cohorts used in GWAS studies on allergic sensitization, for which only a few aeroallergens were tested or had only

self-reported allergy status. For example, Hinds et al.'s study used self-reported diagnosis for cat, dust-mite and pollen allergies for assessing allergic sensitization while Ramasamy et al., only used grass pollen sensitization for their analyses.^{13,43} On the other hand, this makes replication difficult since few cohorts have tested as many aeroallergens for their participants. In this study, the number of aeroallergens, especially for the indoor category that includes only three aeroallergens in the EGEA cohort, was different between the SLSJ and the EGEA cohorts. This could explain why only a third of the 200kb regions were replicated, all for the outdoor or both types of aeroallergens sensitization.

Combining genetic data with epigenetic and transcriptomic data made it possible to identify one variant (rs3798307) for the indoor sensitization and four variants (rs10761191, rs4743880, rs10682558 and rs6479446) for both types of sensitization that were associated with the gene expression as well as methylation levels of the same genes in the naïve CD4⁺ T cells. Those genes were *RNASET2* and *MIR3939* for the indoor sensitization and *FGD3* and *CENPP* in both types of sensitization. Moreover, expression and methylation levels of these genes were also correlated in the naïve CD4⁺ T cells. The availability of the expression and methylation data for purified naïve CD4⁺ T cells and eosinophils samples for more than 100 individuals of the SLSJ cohort is one strength of this study. However, the eQTL and mQTL analyses performed could not be replicated as these data were not available in the replication cohort.

Previously, Ferreira et al. had also found an intergenic variant (rs72033857) between *RNASET2* and *MIR3939* that was associated with the risk of allergic disease in individuals with European ancestry.⁴⁴ The *RNASET2* gene is located in the 6q27 locus and is coding for a ribonuclease that recognizes and degrades microbial RNAs. The *RNASET2* activity leads to the stimulation of the innate immune system by the activation of toll-like receptors 8 (TLR8s) in immune cells, promoting a non-atopic immune reaction of lymphocytes and macrophages.⁴⁵⁻⁴⁷ Interestingly, some studies even showed that the immune reaction induced by *RNASET2* could be used to reduce symptoms of allergic disease.^{48,49} As for *MIR3939*, it codes for a non-coding RNA that functions as a post-transcriptional regulator of gene expression. Nothing is known for its involvement in allergies. Some of *MIR3939* gene targets were identified using miRTarBase (version 9.0 beta;

<https://mirtarbase.cuhk.edu.cn>)⁵⁰ and can be found in supplementary Table 22 with their most relevant biological function identified with PANTHER³⁷, GeneCards^{38,39} and UniProt⁴⁰. Certain of its targets are involved in immunity, such as ETS transcription factor ELK4 (*ELK4*) and salt inducible kinase 1 (*SIK1*) which are involved in macrophages polarization, the granulin precursor (*GRN*; previously known as *PGRN*) which is involved in the pathogenesis of AA and the cryptochrome circadian regulator 2 (*CRY2*) which is implicated in the nuclear factor kappa B (NF-κB) pathway.⁵¹⁻⁵⁴ Thus, *MIR3939* is an interesting micro-RNA for the study of allergic disease through its interaction with its gene targets.

Considering the sensitization to both types of aeroallergens, all four variants were located in the gene body of *FGD3* on locus 9q22.31. These variants were significantly associated with the expression and methylation levels of *FGD3* in the naïve CD4⁺ T cells. Among them, rs10761191 also showed significant associations with the gene expression and methylation levels of *CENPP*, another gene of the 9q22.31 locus, still in the naïve CD4⁺ T cells. None of these variants and genes were previously associated with allergic diseases in the literature. The *FGD3* gene is involved in cell shape and inhibition of migration and is normally highly expressed in cells implicated in immune reaction. This may indicate that under normal conditions, these cells have their motility inhibited.^{55,56} Further studies are needed to confirm this mechanism of action in allergies. The *CENPP* gene is a member of the centromere protein (CENP)-O subunit of the CENP-H-I complex. The CENPP protein is involved in processes leading to chromosome segregation during mitosis.⁵⁷⁻⁵⁹ The link between this gene and allergy is unclear, but some studies have found autoantibodies for the CENP proteins in autoimmune diseases like systemic sclerosis.^{60,61} Interestingly, one of those autoantibodies targeting the CENP proteins (anti-CENP-A) is able to recognize another protein (forkhead box E3; FOXE-3).⁶⁰ Considering this information, CENPP autoantibodies could probably have a cross-reaction with certain proteins found in allergens, which could explain its association to both types of aeroallergens sensitization.

4.8 Conclusion

This study was able to establish a specific genetic profile associated to the sensitization to indoor, outdoor as well as for both types of aeroallergens in a French-Canadian cohort. Combination of genetic, transcriptomic and epigenetic data in two immune cell types was also performed in order to get a better understanding of the biological mechanisms behind allergic sensitization. Further research is needed to completely confirm the biological mechanisms identified here and their impact on allergic sensitization and ARD. However, these results show promising avenues for the development of genetic and/or epigenetic predicting tools to identify risk for allergens sensitization and prevent their development, leading to the reduction of allergies as well as AA and AR prevalence.

4.9 References

1. Eguiluz-Gracia I, Mathioudakis AG, Bartel S, Vijverberg SJH, Fuertes E, Comberiat P, Cai YS, Tomazic PV, Diamant Z, Vestbo J, et al. The need for clean air: The way air pollution and climate change affect allergic rhinitis and asthma. *Allergy*. 2020;75(9):2170-2184. doi: 10.1111/all.14177
2. D'Amato G, Chong-Neto HJ, Monge Ortega OP, Vitale C, Ansotegui I, Rosario N, Haahtela T, Galan C, Pawankar R, Murrieta-Aguttes M, et al. The effects of climate change on respiratory allergy and asthma induced by pollen and mold allergens. *Allergy*. 2020;75(9):2219-2228. doi: 10.1111/all.14476
3. Sierra-Heredia C, North M, Brook J, Daly C, Ellis AK, Henderson D, Henderson SB, Lavigne É, Takaro TK. Aeroallergens in Canada: Distribution, Public Health Impacts, Opportunities for Prevention. *Int J Environ Res Public Health*. 2018;15(8):1577. doi: 10.3390/ijerph15081577
4. Tham EH, Lee AJ, Bever HV. Aeroallergen sensitization and allergic disease phenotypes in Asia. *Asian Pac J Allergy Immunol*. 2016;34(3):181-189. doi: 10.12932/AP0770
5. D'Amato G, Holgate ST, Pawankar R, Ledford DK, Cecchi L, Al-Ahmad M, Al-Enezi F, Al-Muhsen S, Ansotegui I, Baena-Cagnani CE, et al. Meteorological conditions, climate change, new emerging factors, and asthma and related allergic disorders. A statement of the World Allergy Organization. *World Allergy Organ J*. 2015;8(1):25. doi: 10.1186/s40413-015-0073-0
6. Annesi-Maesano I, Maesano CN, Biagioni B, D'Amato G, Cecchi L. Call to action: Air pollution, asthma, and allergy in the exposome era. *J Allergy Clin Immunol*. 2021;148:70-72. doi: 10.1016/j.jaci.2021.05.026
7. Seguel JM, Merrill R, Seguel D, Campagna AC. Indoor Air Quality. *Am J Lifestyle Med*. 2016;11(4):284-295. doi: 10.1177/1559827616653343
8. Li J, Sun B, Huang Y, Lin X, Zhao D, Tan G, Wu J, Zhao H, Cao L, Zhong N, et al. A multicentre study assessing the prevalence of sensitizations in patients with asthma and/or rhinitis in China. *Allergy*. 2009;64(7):1083-1092. doi: 10.1111/j.1398-9995.2009.01967.x

9. Boulet L-P, Turcotte H, Laprise C, Lavertu C, Bedard P-M, Lavoie A, Hébert J. Comparative degree and type of sensitization to common indoor and outdoor allergens in subjects with allergic rhinitis and/or asthma. *Clin Exp Allergy*. 1997;27(1):52-59. doi: 10.1111/j.1365-2222.1997.tb00672.x
10. Halken S. Early sensitisation and development of allergic airway disease – risk factors and predictors. *Paediatr Respir Rev*. 2003;4(2):128-134. doi: 10.1016/S1526-0542(03)00026-5
11. Sheffield PE, Weinberger KR, Kinney PL. Climate Change, Aeroallergens, and Pediatric Allergic Disease. *Mt Sinai J Med*. 2011;78(1):78-84. doi: 10.1002/msj.20232
12. Warm K, Hedman L, Lindberg A, Lötvall J, Lundbäck B, Rönmark E. Allergic sensitization is age-dependently associated with rhinitis, but less so with asthma. *J Allergy Clin Immunol*. 2015;136(6):1559-1565.E1552. doi: 10.1016/j.jaci.2015.06.015
13. Ramasamy A, Curjuric I, Coin LJ, Kumar A, McArdle WL, Imboden M, Leynaert B, Kogevinas M, Schmid-Grendelmeier P, Pekkanen J, et al. A genome-wide meta-analysis of genetic variants associated with allergic rhinitis and grass sensitization and their interaction with birth order. *J Allergy Clin Immunol*. 2011;128(5):996-1005. doi: 10.1016/j.jaci.2011.08.030
14. Laprise C. The Saguenay-Lac-Saint-Jean asthma familial collection: the genetics of asthma in a young founder population. *Genes Immun*. 2014;15(4):247-255. doi: 10.1038/gene.2014.12
15. Morin A, Madore A-M, Kwan T, Ban M, Partanen J, Rönnblom L, Syvänen A-C, Sawcer S, Stunnenberg H, Lathrop M, et al. Exploring rare and low-frequency variants in the Saguenay–Lac-Saint-Jean population identified genes associated with asthma and allergy traits. *Eur J Hum Genet*. 2019;27(1):90-101. doi: 10.1038/s41431-018-0266-4
16. Madore A-M, Pain L, Boucher-Lafleur A-M, Morin A, Meloche J, Simon M-M, Ge B, Kwan T, Cheung WA, Pastinen T, et al. Asthma-associated polymorphisms in 17q12-21 locus modulate methylation and gene expression of GSDMA in naïve CD4+ T cells. *J Genet Genomics*. 2020;47(3):171-174. doi: 10.1016/j.jgg.2020.03.002

17. Madore A-M, Bossé Y, Margaritte-Jeannin P, Vucic E, Lam WL, Bouzigon E, Bourbeau J, Laprise C. Analysis of GWAS-nominated loci for lung cancer and COPD revealed a new asthma locus. *BMC Pulm Med*. 2022;22(1):155. doi: 10.1186/s12890-022-01890-7
18. Kauffmann F, Dizier MH, Pin I, Paty E, Gormand F, Vervloet D, Bousquet J, Neukirch F, Annesi I, Orszcyn MP, et al. Epidemiological study of the genetics and environment of asthma, bronchial hyperresponsiveness, and atopy: phenotype issues. *Am J Respir Crit Care Med*. 1997;156(4 Pt 2):S123-S129. doi: 10.1164/ajrccm.156.4.12tac9
19. Delaneau O, Marchini J, Zagury J-F. A linear complexity phasing method for thousands of genomes. *Nat Methods*. 2012;9(2):179-181. doi: 10.1038/nmeth.1785
20. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009;5(6):e1000529. doi: 10.1371/journal.pgen.1000529
21. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48(10):1284-1287. doi: 10.1038/ng.3656
22. Vernet R, Matran R, Zerimech F, Madore A-M, Lavoie M-E, Gagnon P-A, Mohamdi H, Margaritte-Jeannin P, Siroux V, Dizier M-H, et al. Identification of novel genes influencing eosinophil-specific protein levels in asthma families. *J Allergy Clin Immunol*. 2022;150(5):1168-1177. doi: 10.1016/j.jaci.2022.05.017
23. Won S, Kim W, Lee S, Lee Y, Sung J, Park T. Family-based association analysis: a fast and efficient method of multivariate association analysis with multiple variants. *BMC Bioinformatics*. 2015;16:46. doi: 10.1186/s12859-015-0484-5
24. Lee S, Choi S, Qiao D, Cho M, Silverman EK, Park T, Won S. WISARD: workbench for integrated superfast association studies for related datasets. *BMC Med Genomics*. 2018;11(Suppl 2):39. doi: 10.1186/s12920-018-0345-y
25. Turner SD. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J Open Source Softw*. 2018;3(25):731. doi: 10.21105/joss.00731

26. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006;38(8):904-909. doi: 10.1038/ng1847
27. StataCorp. *Stata Statistical Software: Release 14*; College Station, TX: StataCorp LLC; 2015.
28. 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. doi: 10.1086/383251
29. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience.* 2015;4(1):7. doi: 10.1186/s13742-015-0047-8
30. Purcell S. LD calculations. <https://zzz.bwh.harvard.edu/plink/ld.shtml#ld2>. 2017. Accessed 17 October 2023.
31. Komlósi ZI, van de Veen W, Kovács N, Szűcs G, Sokolowska M, O'Mahony L, Akdis M, Akdis CA. Cellular and molecular mechanisms of allergic asthma. *Mol Aspects Med.* 2022;85:100995. doi: 10.1016/j.mam.2021.100995
32. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550. doi: 10.1186/s13059-014-0550-8
33. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics.* 2012;28(10):1353-1358. doi: 10.1093/bioinformatics/bts163
34. Zhu LJ. Integrative analysis of ChIP-chip and ChIP-seq dataset. *Methods Mol Biol.* 2013;1067:105-124. doi: 10.1007/978-1-62703-607-8_8
35. Zhu LJ, Gazin C, Lawson ND, Pagès H, Lin SM, Lapointe DS, Green MR. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics.* 2010;11:237. doi: 10.1186/1471-2105-11-237
36. Venables WN, Ripley BD. *Modern Applied Statistics with S*, Fourth Edition. Springer; 2002.
37. Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albu L-P, Mi H. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* 2022;31(1):8-22. doi: 10.1002/pro.4218

38. Safran M, Rosen N, Twik M, BarShir R, Stein TI, Dahary D, Fishilevich S, Lancet D. The GeneCards Suite. In: Abugessaisa I, Kasukawa T, eds. Practical Guide to Life Science Databases. Singapore: Springer; 2021:27-56.
39. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I, Mazor Y, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr Protoc Bioinformatics*. 2016;54:1.30.31-31.30.33. doi: 10.1002/cpbi.5
40. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res*. 2023;51(D1):D523-D531. doi: 10.1093/nar/gkac1052
41. Morin A, Laviolette M, Pastinen T, Boulet L-P, Laprise C. Combining omics data to identify genes associated with allergic rhinitis. *Clin Epigenetics*. 2017;9:3. doi: 10.1186/s13148-017-0310-1
42. Bønnelykke K, Matheson MC, Pers TH, Granel R, Strachan DP, Alves AC, Linneberg A, Curtin JA, Warrington NM, Standl M, et al. Meta-analysis of genome-wide association studies identifies ten loci influencing allergic sensitization. *Nat Genet*. 2013;45:902-906. doi: 10.1038/ng.2694
43. Hinds DA, McMahon G, Kiefer AK, Do CB, Eriksson N, Evans DM, St Pourcain B, Ring SM, Moutain JL, Francke U, et al. A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci. *Nat Genet*. 2013;45(8):907-911. doi: 10.1038/ng.2686
44. Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, Helmer Q, Tillander A, Ullemer V, van Dongen J, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat Genet*. 2017;49(12):1752-1757. doi: 10.1038/ng.3985
45. Greulich W, Wagner M, Gaidt MM, Stafford C, Cheng Y, Linder A, Carell T, Hornung V. TLR8 is a sensor of RNase T2 degradation products. *Cell*. 2019;179(6):1264-1275.e1213. doi: 10.1016/j.cell.2019.11.001

46. Vierbuchen T, Stein K, Heine H. RNA is taking its Toll: Impact of RNA-specific Toll-like receptors on health and disease. *Allergy*. 2019;74(2):223-235. doi: 10.1111/all.13680
47. Scaldaferri D, Bosi A, Fabbri M, Pedrini E, Inforzato A, Valli R, Frattini A, De Vito A, Noonan DM, Taramelli R, et al. The human RNASET2 protein affects the polarization pattern of human macrophages in vitro. *Immunol Lett*. 2018;203:102-111. doi: 10.1016/j.imlet.2018.09.005
48. Horak F. VTX-1463, a novel TLR8 agonist for the treatment of allergic rhinitis. *Expert Opin Investig Drugs*. 2011;20(7):981-986. doi: 10.1517/13543784.2011.583237
49. Stein K, Brand S, Jenckel A, Sigmund A, Chen ZJ, Kirschning CJ, Kauth M, Heine H. Endosomal recognition of *Lactococcus lactis* G121 and its RNA by dendritic cells is key to its allergy-protective effects. *J Allergy Clin Immunol*. 2017;139(2):667-678. doi: 10.1016/j.jaci.2016.06.018
50. Huang H-Y, Lin Y-C-D, Cui S, Huang Y, Tang Y, Xu J, Bao J, Li Y, Wen J, Zuo H, et al. miRTarBase update 2022: an informative resource for experimentally validated miRNA–target interactions. *Nucleic Acids Res*. 2022;50(D1):D222-D230. doi: 10.1093/nar/gkab1079
51. Xie L. MKL1/2 and ELK4 co-regulate distinct serum response factor (SRF) transcription programs in macrophages. *BMC Genomics*. 2014;15:301. doi: 10.1186/1471-2164-15-301
52. Darling NJ, Toth R, Arthur JSC, Clark K. Inhibition of SIK2 and SIK3 during differentiation enhances the anti-inflammatory phenotype of macrophages. *Biochem J*. 2017;474(4):521-537. doi: 10.1042/BCJ20160646
53. Choi J-P, Park SY, Moon K-A, Ha EH, Woo YD, Chung DH, Kwon H-S, Kim T-B, Park H-S, Moon H-B, et al. Macrophage-derived progranulin promotes allergen-induced airway inflammation. *Allergy*. 2020;75(5):1133-1145. doi: 10.1111/all.14129
54. Narasimamurthy R, Hatori M, Nayak SK, Liu F, Panda S, Verma IM. Circadian clock protein cryptochrome regulates the expression of proinflammatory cytokines. *Proc Natl Acad Sci U S A*. 2012;109(31):12662-12667. doi: 10.1073/pnas.1209965109
55. Hayakawa M, Matsushima M, Hagiwara H, Oshima T, Fujino T, Ando K, Kikugawa K, Tanaka H, Miyazawa K, Kitagawa M. Novel insights into FGD3, a putative GEF for Cdc42, that

- undergoes SCFFWD1/ β -TrCP-mediated proteasomal degradation analogous to that of its homologue FGD1 but regulates cell morphology and motility differently from FGD1. *Genes Cells*. 2008;13(4):329-342. doi: 10.1111/j.1365-2443.2008.01168.x
56. Willis S, Sun Y, Abramovitz M, Fei T, Young B, Lin X, Ni M, Achua J, Regan MM, Gray KP, et al. High Expression of FGD3, a Putative Regulator of Cell Morphology and Motility, Is Prognostic of Favorable Outcome in Multiple Cancers. *JCO Precis Oncol*. 2017;1:PO.17.00009. doi: 10.1200/po.17.00009
 57. Okada M, Cheeseman IM, Hori T, Okawa K, McLeod IX, Yates 3rd JR, Desai A, Fukagawa T. The CENP-H-I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. *Nat Cell Biol*. 2006;8(5):446-457. doi: 10.1038/ncb1396
 58. Foltz DR, Jansen LET, Black BE, Bailey AO, Yates 3rd JR, Cleveland DW. The human CENP-A centromeric nucleosome-associated complex. *Nat Cell Biol*. 2006;8(5):458-469. doi: 10.1038/ncb1397
 59. Hori T, Shang W-H, Takeuchi K, Fukagawa T. The CCAN recruits CENP-A to the centromere and forms the structural core for kinetochore assembly. *J Cell Biol*. 2013;200(1):45-60. doi: 10.1083/jcb.201210106
 60. Perosa F, Prete M, Di Lernia G, Ostuni C, Favoino E, Valentini G. Anti-centromere protein A antibodies in systemic sclerosis: Significance and origin. *Autoimmun Rev*. 2016;15(1):102-109. doi: 10.1016/j.autrev.2015.10.001
 61. Kajio N, Takeshita M, Suzuki K, Kaneda Y, Yamane H, Ikeura K, Sato H, Kato S, Shimizu H, Tsunoda K, et al. Anti-centromere antibodies target centromere-kinetochore macrocomplex: a comprehensive autoantigen profiling. *Ann Rheum Dis*. 2021;80(5):651-659. doi: 10.1136/annrheumdis-2020-218881
 62. Song G, Hu C, Zhu H, Wang L, Zhang F, Li Y, Wu L. New Centromere Autoantigens Identified in Systemic Sclerosis Using Centromere Protein Microarrays. *J Rheumatol*. 2013;40(4):461-468. doi: 10.3899/jrheum.120264
 63. BioRender. Scientific illustration software. <https://www.biorender.com>. 2023.

DISCUSSION

L'objectif principal de cette étude était d'identifier le profil génétique spécifique à chacun des types de sensibilisation aux aéroallergènes étudiés soit d'intérieur, d'extérieur et combinée. Suite à des analyses de GWAS pour chacun des phénotypes, il a été possible d'établir des profils spécifiques à chacun des types de sensibilisation dans un sous-échantillon (n=1 010) d'individus ayant des données génomique et phénotypique pour l'allergie de la cohorte familiale d'asthme du SLSJ. Pour la sensibilisation aux aéroallergènes d'intérieurs, 16 loci ont été identifiés parmi lesquels 12 sont spécifiques à ce type de sensibilisation, c'est-à-dire qu'ils ne sont pas partagés avec la sensibilisation aux aéroallergènes d'extérieur et/ou combinée. Pour la sensibilisation aux aéroallergènes d'extérieur, neuf loci lui sont spécifiques parmi un total de 13 loci identifiés dans la cohorte SLSJ. En ce qui concerne la sensibilisation combinée, huit loci lui sont spécifiques parmi les 15 loci identifiés dans la cohorte SLSJ.

La réplication des résultats d'associations obtenues pour chacun des types de sensibilisation aux aéroallergènes dans la cohorte du SLSJ a été faite dans la cohorte de l'Étude Génétique Environnement Asthme (EGEA). Cela a permis la réplication, dans la cohorte EGEA, de neuf régions de 200 kb situées dans huit des 16 loci associés à la sensibilisation aux aéroallergènes d'intérieur dans la cohorte du SLSJ. Pour la sensibilisation aux aéroallergènes d'extérieur, trois régions de 200 kb situées dans trois des 13 loci identifiés, pour la sensibilisation aux aéroallergènes d'extérieur dans la cohorte du SLSJ, ont été répliqués dans la cohorte EGEA. En ce qui concerne la sensibilisation aux aéroallergènes combinée, trois régions de 200 kb situées dans trois des 15 loci mis en évidence pour la sensibilisation aux aéroallergènes combinée dans la cohorte du SLSJ, ont été répliqués dans la cohorte EGEA. Le peu de régions répliquées pour la sensibilisation aux aéroallergènes d'extérieur et combinée pourrait être dû au fait que des aéroallergènes d'extérieur différents ont été utilisés pour le diagnostic de ce type de sensibilisation entre les deux cohortes. Un autre facteur qui peut expliquer ce peu de réplication est la différence de taille des blocs de LD (plus grand pour EGEA que pour

SLSJ), ce qui fait que les blocs ne sont pas forcément les mêmes pour les deux populations malgré leur même origine ethnique (Madore et al., 2020).

Dans la littérature, seulement un des loci identifiés dans ce projet a déjà été associé à la sensibilisation allergénique. Il s'agit du locus 6p21.33 qui a uniquement été associé à la sensibilisation combinée dans ce projet et dont l'une des régions de 200 kb répliquées dans la cohorte EGEA est située à l'intérieur de ce locus (6p21.33). La première étude à avoir identifié un variant dans ce locus comme étant associé à la sensibilisation allergénique est celle de Bønnelykke et al. (2013). Ils ont identifié le variant rs6932730 ($p_{\text{combiné}} = 4,20 \times 10^{-8}$; $OR_{\text{combiné}} = 1,14$) qui se situe près des gènes *major histocompatibility complex, class I, B (HLA-B)* et *MHC class I polypeptide-related sequence A (MICA)*; Bønnelykke et al., 2013). Une autre étude par Hinds et al. (2013) a quant à elle identifié le variant rs9266772 ($p = 3,2 \times 10^{-12}$; $OR = 1,11$) se situant près des gènes *HLA-C* et *MICA*. Ainsi, lorsque l'on compare les résultats obtenus à ceux rapportés dans les études de Bønnelykke et al. (2013) et de Hinds et al. (2013), il est possible de constater que quatre nouveaux variants du locus 6p21.33 ont été mis en évidence dans la cohorte du SLSJ. Parmi ces quatre variants, le variant rs13437096 est le plus près des deux déjà rapportés dans la littérature alors que rs6932730 et rs9266772 ont une distance respective de 1,11 kb et 3,17 kb. Cependant, l'analyse du LD a montré que le variant rs6932730 n'était pas en LD avec aucun des quatre variants identifiés dans la cohorte du SLSJ. En ce qui concerne le variant rs9266772, aucune information n'était disponible pour ce variant lors de la répllication et de l'analyse du LD. Il a aussi été possible d'identifier 167 variants du locus 6p21.33 qui passent le seuil de significativité de l'analyse de répllication, pour la sensibilisation aux aéroallergènes combinée dans la cohorte EGEA. Il est à noter que le variant rs6932730 est l'un des variants identifiés dans cette région génétique. Le locus 6p21.33 renferme des gènes codant pour les complexes majeurs d'histocompatibilités (MHC) de classe I qui font partie des antigènes des leucocytes humains (HLA) et les *MHC Class I polypeptide-Related Sequence (MIC)*. Ainsi, le locus 6p21.33 regroupe des gènes comme *HLA-B*, *HLA-C*, *MICA* et *MICB* qui ont des fonctions importantes dans le système immunitaire par le biais de leur implication dans la présentation d'antigènes et de la réponse au stress (Neefjes et al., 2011; Stojanovic et al., 2018). Ainsi, des études supplémentaires

de ce locus permettront de déterminer si un ou plusieurs des gènes qu'il renferme pourrait avoir un impact sur le développement d'une sensibilisation allergénique.

Le deuxième objectif de cette étude était d'identifier de possibles voies biologiques impliquées dans la sensibilisation aux aéroallergènes. Ainsi, des analyses d'eQTL, de mQTL et de corrélation entre les niveaux d'expression et de méthylation ont été réalisées afin d'intégrer les données génomiques, transcriptomiques et épigénétiques provenant de lymphocytes T CD4⁺ naïfs et d'éosinophiles. Comme mentionné au chapitre 3, la majorité des variants identifiés par GWAS sont situés dans des régions « non codantes » de l'ADN et il est donc difficile d'interpréter leurs impacts sur la maladie (B. Li & Ritchie, 2021; Liu & Montgomery, 2020). Bien évidemment, plusieurs variants identifiés dans cette étude ne font pas exception et se retrouvent donc dans des régions intergéniques, en amont ou en aval d'un gène ou encore dans un gène dont la fonction est encore inconnue. Ainsi, les analyses génétiques fonctionnelles réalisées dans cette étude ont permis d'identifier plusieurs gènes d'intérêt par le biais des analyses d'eQTL et de mQTL. Cela a permis d'identifier 12 variants ayant une association significative avec le niveau d'expression de neuf gènes chez les lymphocytes T CD4⁺ naïfs et trois variants ayant une association significative avec le niveau d'expression de quatre autres gènes pour les éosinophiles. De plus, les résultats de mQTL montrent que certains variants sont associés à une différence de méthylation pour le(s) même(s) site(s) CpG dans un gène ou différent(s) site(s) CpG dans un même gène à la fois dans les lymphocytes T CD4⁺ et les éosinophiles pour chacun des trois types de sensibilisations aux aéroallergènes. Ces résultats permettent de mettre en évidence qu'il existe de possibles voies communes et spécifiques aux lymphocytes T CD4⁺ naïfs et aux éosinophiles dans le développement d'une sensibilisation aux aéroallergènes.

Les analyses génétiques fonctionnelles ont mis en évidence un variant ayant une différence d'expression significative pour deux gènes dans les lymphocytes T CD4⁺ naïfs seulement et une différence de méthylation significative pour plusieurs CpG pour les mêmes gènes dans les lymphocytes T CD⁺ naïfs (23 CpG) et les éosinophiles (3 CpG) pour la sensibilisation aux aéroallergènes d'intérieur. Par ailleurs, les analyses de corrélation ont permis d'établir une corrélation

significative entre le niveau d'expression et de méthylation de deux gènes dans les lymphocytes T CD4⁺ naïfs pour la sensibilisation aux aéroallergènes d'intérieur. Ces gènes sont la *ribonuclease T2 (RNASET2)* et le *microRNA 3939 (MIR3939)* qui sont tous les deux situés sur le locus 6q27. *RNASET2* est un gène impliqué dans la reconnaissance de pathogènes. Il est aussi impliqué dans la voie de signalisation conduisant à la polarisation des macrophages en M1 ainsi qu'à la différenciation des Th0 en Th1 et en lymphocytes T folliculaires (Greulich et al., 2019; Scaldaferri et al., 2018; Vierbuchen et al., 2019). Ce processus entraîne une réduction de la polarisation des macrophages vers le type M2 et la différenciation des Th0 vers la voie Th2 (Scaldaferri et al., 2018; Vierbuchen et al., 2019). Ainsi, le mécanisme d'action de *RNASET2* favorise une réaction immunitaire de type non allergique et laisse envisager qu'une inhibition de ce gène pourrait favoriser une réaction de type allergique. En ce qui concerne *MIR3939*, il s'agit d'un gène codant pour un micro-ARN impliqué dans la régulation de l'expression de certains gènes (voir le Tableau 22 dans l'annexe). Ce micro-ARN régule entre autres les quatre gènes suivants : *ETS transcription factor ELK4 (ELK4)*, *salt inducible kinase 1 (SIK1)*, *granulin precursor (GRN)* et *cryptochrome circadian regulator 2 (CRY2)*. Ces gènes sont impliqués dans la régulation des réactions inflammatoires (J.-P. Choi et al., 2020; Darling et al., 2017; Narasimamurthy et al., 2012; Xie, 2014). Notamment, J.-P. Choi et al. (2020) ont démontré que le gène *GRN* est impliqué dans le développement d'une inflammation allergique des voies respiratoires par l'induction de la production de cytokines de type Th2 par les NKT et les cellules épithéliales suites à sa sécrétion par les macrophages exposés à un allergène. Ainsi, *MIR3939* pourrait être impliqué dans le développement de maladies allergiques par son interaction indirecte sur l'un ou plusieurs de ses gènes cibles impliqués dans les réactions inflammatoires.

Les analyses génétiques fonctionnelles ont aussi mis en évidence quatre variants ayant une différence d'expression significative pour deux gènes dans les lymphocytes T CD4⁺ naïfs seulement et une différence de méthylation significative pour plusieurs CpG pour les mêmes gènes dans les lymphocytes T CD⁺ naïfs (16 CpG) et les éosinophiles (12 CpG) pour la sensibilisation aux aéroallergènes combinée. Par contre, les analyses de corrélation ont identifié une corrélation significative entre le niveau d'expression et de méthylation de ces deux gènes dans les lymphocytes

T CD4⁺ naïfs pour la sensibilisation aux aéroallergènes combinés. Ces deux gènes sont *FYVE*, *RhoGEF and PH domain containing 3 (FGD3)* et *centromere protein P (CENPP)*, qui sont tous deux situés dans le locus 9q22.31. *FGD3* est un gène codant pour une protéine impliquée dans la forme et la migration cellulaire par son interaction avec la protéine *cell division cycle 42 (CDC42)*. Cela conduira à la production de lamellipodes, inhibant la migration de la cellule (Hayakawa et al., 2008). En conditions normales, la protéine CDC42 est fortement exprimée dans les lymphocytes T, NKT et les cellules myéloïdes, ce qui indique que leur migration est normalement inhibée (Willis et al., 2017). Par conséquent, ce mécanisme laisse penser que l'inhibition de *FGD3* pourrait occasionner l'augmentation de la migration des cellules inflammatoires lors de réactions allergiques. Cependant, plus d'études sur l'impact de *FGD3* seraient nécessaires pour confirmer cette hypothèse. En ce qui concerne le gène *CENPP*, il est impliqué dans l'intégration de la protéine CENP-A dans le centromère des chromosomes, l'assemblage des kinétochores et la ségrégation des chromosomes lors de l'interphase de la mitose (Foltz et al., 2006; Hori et al., 2013; Okada et al., 2006). Bien que l'implication de *CENPP* dans l'allergie ne soit pas évidente, une étude réalisée par Song et al. (2013) a permis d'identifier la présence d'auto-anticorps contre la protéine CENP-P chez des patients atteints de la sclérose systémique. Notamment, une étude a déjà démontré qu'il existe une réaction croisée entre les anti-CENP-A et les auto-anticorps de *forkhead box protein E3 (FOXE-3)* (Perosa et al., 2016). Or, comme mentionné dans la section 1.2.1 Animaux et la section 1.2.3 Moisissures du chapitre 1, il existe des composantes allergéniques pour lesquelles les anticorps spécifiques peuvent avoir des réactions croisées avec des protéines d'origine humaine. Par exemple, les anticorps reconnaissant les antigènes de la cyclophiline de *Aspergillus fumigatus* pourront reconnaître la cyclophiline produite chez l'humain (Matricardi et al., 2016; Twaroch et al., 2015). Un autre exemple, est la réaction croisée possible entre l'une des composantes allergéniques du chien et de l'antigène prostatique spécifique (PSA) humain (Curin & Hilger, 2017). Par conséquent, il pourrait être possible que les auto-anticorps contre CENP-P puissent reconnaître une composante allergénique d'un aéroallergène. Ceci n'est qu'une hypothèse permettant d'interpréter l'implication de *CENPP* dans la sensibilisation allergénique et elle devrait être étudiée plus en profondeur.

L'une des forces de cette étude repose sur l'utilisation d'une cohorte pour laquelle il y avait des données phénotypiques pour la sensibilisation allergénique à 24 aéroallergènes. Ces données ont ainsi permis d'étudier la sensibilisation aux aéroallergènes selon l'origine de ceux-ci (intérieur, extérieur ou combinée) et d'identifier un profil génétique spécifique à chacun des types de sensibilisation aux aéroallergènes. Grâce à ces données phénotypiques, cette étude se démarque des autres qui n'ont utilisé des données phénotypiques que pour un petit nombre d'aéroallergènes ou une seule catégorie (par exemple des animaux ou pollens exclusivement) ou des données pour l'allergie auto-rapportée (Hinds et al., 2013; Ramasamy et al., 2011). Par contre, cet avantage entraîne aussi l'une des limites de cette étude puisqu'il rend la réplique des données d'association plus difficile, puisque peu de cohortes possèdent des données phénotypiques pour autant d'aéroallergènes ou pour les mêmes aéroallergènes. Ainsi, le fait que seulement un tiers des régions de 200 kb aient été répliquées dans la cohorte EGEA pourrait être dû au fait qu'il n'y avait pas le même nombre et/ou les mêmes aéroallergènes considérés lors des tests. Toutefois, il n'est pas exclu que ces différences soient dues aux populations étudiées puisque certaines associations génétiques dans l'étude des traits complexes sont spécifiques aux populations étudiées (Gurdasani et al., 2019).

Une autre limite de cette étude est la difficulté à répliquer les résultats identifiés dans la littérature puisque beaucoup de cohortes vont privilégier une caractérisation plus superficielle (autorapporté, etc.) de l'AA, de la RA et de l'allergie en général contrairement à celle du SLSJ où le phénotype a été établi pour tous les participants par un spécialiste clinique. Une autre limite est l'impossibilité d'atteindre le seuil de significativité à l'échelle du génome ($5E-08$) qui est normalement utilisé pour les GWAS. Cette limite découle du fait que l'allergie a été séparée en trois groupes, soit intérieur, extérieur et combiné, ce qui a contribué à réduire la taille de l'échantillon analysé pour chaque type d'allergie. De plus, cela a contribué à rendre difficile l'identification de variant ayant un fort impact sur le trait étudié.

Une autre force de cette étude est l'utilisation de données d'expression et de méthylation pour deux types de cellules immunitaires impliquées dans la pathogenèse de l'allergie. L'utilisation de données transcriptomique et épigénomique pour deux types de cellules immunitaires a permis

d'identifier des voies biologiques d'intérêt qui sont potentiellement impliquées dans la pathophysiologie de l'AA et de l'AR. L'utilisation de types cellulaires spécifiques à une maladie lors de QTL permet d'avoir une meilleure puissance d'analyse contrairement à l'utilisation d'un mixte cellulaire (Kim-Hellmuth et al., 2020). Cela est dû au fait que le transcriptome et l'épigénome sont spécifiques aux cellules et que les variations génétiques vont souvent avoir un effet au niveau de l'expression et de la méthylation d'un type cellulaire précis (Kim-Hellmuth et al., 2020; Westra et al., 2015). Une autre des limites de cette étude est l'impossibilité de répliquer les résultats d'eQTL et de mQTL dans une autre cohorte, étant donné qu'aucune autre cohorte d'asthme et d'allergies ne possède de données d'expression et de méthylation dans les lymphocytes T CD4⁺ naïfs et dans les éosinophiles.

CONCLUSION

Cette étude a permis d'atteindre le premier objectif, soit d'établir un profil génétique différent pour la sensibilisation aux aéroallergènes d'intérieur, d'extérieur et combinée dans la cohorte d'asthme familiale du Saguenay–Lac-Saint-Jean. Les analyses génétiques ont permis d'associer 12 loci spécifiques à la sensibilisation aux aéroallergènes d'intérieur, neuf loci spécifiques à la sensibilisation aux aéroallergènes d'extérieur et huit loci spécifiques à la sensibilisation combinée. Notamment, aucun variant génétique n'était commun entre la sensibilisation aux aéroallergènes d'intérieur et d'extérieur, mais un seul locus (9q31.3) l'était pour les trois types de sensibilisation. Par la suite, il a été possible de répliquer dans la cohorte EGEA, huit, trois et trois loci ayant été associé dans la cohorte du SLSJ à la sensibilisation aux aéroallergènes d'intérieur, d'extérieur et combinée respectivement. De plus, il a été possible de répliquer une association, déjà établie dans la littérature, entre le locus 6p21.33 et la sensibilisation allergénique. Ainsi, cette première partie de l'étude a permis de mettre en avant de nouveaux variants génétiques et loci candidats pour l'étude de la sensibilisation allergénique. Le deuxième objectif de cette étude, soit d'identifier de possibles voies biologiques impliquées dans chacun des trois types de sensibilisation aux aéroallergènes étudiés, a aussi été atteint. Notamment, il a été possible de mettre en lumière certains gènes (*RNASET2*, *MIR3939*, *FGD3* et *CENPP*) dont il serait intéressant de poursuivre la recherche avec d'autres analyses fonctionnelles afin de mieux comprendre les voies biologiques dans lesquels ils sont impliqués et d'accroître les connaissances existantes sur la sensibilisation allergénique et le développement des maladies allergiques respiratoires. De plus, des analyses additionnelles où l'intégration des données omiques seraient poursuivies avec plus de participants ou en incluant les données d'autres cellules activées dans la réaction allergique et contribuant au processus inflammatoire dont les lymphocytes T CD4⁺ naïfs, viendraient également bonifier la compréhension sur les maladies allergiques.

Ainsi, des recherches seront nécessaires afin de pouvoir confirmer les mécanismes biologiques qui ont été identifiés dans cette étude et qui sont potentiellement impliqués dans la

sensibilisation allergénique et de mieux comprendre l'impact de ceux-ci sur le développement de l'AA et de la RA. Les résultats obtenus montrent qu'il sera éventuellement possible de créer des outils génétiques et/ou épigénétiques permettant d'évaluer le niveau de risque d'un individu pour le développement d'une sensibilisation allergénique et donc de prévenir le développement de celle-ci ou de mieux intervenir par des thérapies plus ciblées.

BIBLIOGRAPHIE

- Abel-Fernández, E., Martínez, M. J., Galán, T., & Pineda, F. (2023). Going over Fungal Allergy: *Alternaria alternata* and Its Allergens. *Journal of Fungi (Basel)*, 9(5), 582. <https://doi.org/10.3390/jof9050582>
- Acevedo, N., Zakzuk, J., & Caraballo, L. (2019). House Dust Mite Allergy Under Changing Environments. *Allergy, Asthma & Immunology Research*, 11(4), 450-469. <https://doi.org/10.4168/aair.2019.11.4.450>
- Agache, I., & Akdis, C. A. (2019). Precision medicine and phenotypes, endotypes, genotypes, regiotypes, and theratypes of allergic diseases. *The Journal of Clinical Investigation*, 129(4), 1493-1503. <https://doi.org/10.1172/JCI124611>
- Ahmed, H., Ospina, M. B., Sideri, K., & Vliagoftis, H. (2019). Retrospective analysis of aeroallergen's sensitization patterns in Edmonton, Canada. *Allergy, Asthma & Clinical Immunology*, 15, 6. <https://doi.org/10.1186/s13223-019-0320-y>
- Akar-Ghinril, N., & Phipatanakul, W. (2020). The Indoor Environment and Childhood Asthma. *Current Allergy and Asthma Reports*, 20(9), 43. <https://doi.org/10.1007/s11882-020-00941-5>
- Akiyama, M. (2021). Multi-omics study for interpretation of genome-wide association study. *Journal of Human Genetics*, 66, 3-10. <https://doi.org/10.1038/s10038-020-00842-5>
- Amaral, A. F. S., Minelli, C., Guerra, S., Wjst, M., Probst-Hensch, N., Pin, I., Svanes, C., Janson, C., Heinrich, J., & Jarvis, D. L. (2015). The locus C11orf30 increases susceptibility to poly-sensitization. *Allergy*, 70(3), 328-333. <https://doi.org/10.1111/all.12557>
- Andiappan, A. K., Wang de, Y., Anantharaman, R., Parate, P. N., Suri, B. K., Low, H. Q., Li, Y., Zhao, W., Castagnoli, P., Liu, J., & Chew, F. T. (2011). Genome-Wide Association Study for Atopy and Allergic Rhinitis in a Singapore Chinese Population. *PLoS One*, 6(5), e19719. <https://doi.org/10.1371/journal.pone.0019719>
- Andiappan, A. K., Wang de, Y., Anantharaman, R., Suri, B. K., Lee, B. T. K., Rotzschke, O., Liu, J., & Chew, F. T. (2013). Replication of genome-wide association study loci for allergic rhinitis and house dust mite sensitization in an Asian population of ethnic Chinese in Singapore. *Journal of Allergy and Clinical Immunology*, 131(5), 1431-1433.e1438. <https://doi.org/10.1016/j.jaci.2012.11.001>
- Annesi-Maesano, I., Maesano, C. N., Biagioni, B., D'Amato, G., & Cecchi, L. (2021). Call to action: Air pollution, asthma, and allergy in the exposome era. *Journal of Allergy and Clinical Immunology*, 148(1), 70-72. <https://doi.org/10.1016/j.jaci.2021.05.026>
- Baloh, C. H., & Mathias, R. A. (2023). Recent progress in the genetic and epigenetic underpinnings of atopy. *Journal of Allergy and Clinical Immunology*, 151(1), 60-69. <https://doi.org/10.1016/j.jaci.2022.10.027>
- BioRender®. (2023). *Outil d'illustrations scientifiques*. <https://www.biorender.com>
- Bloemen, K., Verstraelen, S., Van Den Heuvel, R., Witters, H., Nelissen, I., & Schoeters, G. (2007). The allergic cascade: Review of the most important molecules in the asthmatic lung. *Immunology Letters*, 113(1), 6-18. <https://doi.org/10.1016/j.imlet.2007.07.010>
- Blomen, V. A., & Boonstra, J. (2011). Stable transmission of reversible modifications: maintenance of epigenetic information through the cell cycle. *Cellular and Molecular Life Sciences*, 68(1), 27-44. <https://doi.org/10.1007/s00018-010-0505-5>

- Bønnelykke, K., Matheson, M. C., Pers, T. H., Granell, R., Strachan, D. P., Alves, A. C., Linneberg, A., Curtin, J. A., Warrington, N. M., Standl, M., Kerkhof, M., Jonsdottir, I., Bukvic, B. K., Kaakinen, M., Sleimann, P., Thorleifsson, G., Thorsteinsdottir, U., Schramm, K., Baltic, S., ... EARly Genetics and Lifecourse Epidemiology (EAGLE) Consortium. (2013). Meta-analysis of genome-wide association studies identifies ten loci influencing allergic sensitization. *Nature Genetics*, 45(8), 902-906. <https://doi.org/10.1038/ng.2694>
- Bønnelykke, K., Sleiman, P., Nielsen, K., Kreiner-Møller, E., Mercader, J. M., Belgrave, D., den Dekker, H. T., Husby, A., Sevelsted, A., Faura-Tellez, G., Mortensen, L. J., Paternoster, L., Flaaten, R., Mølgaard, A., Smart, D. E., Thomsen, P. F., Rasmussen, M. A., Bonàs-Guarch, S., Holst, C., ... Bisgaard, H. (2014). A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. *Nature Genetics*, 46(1), 51-55. <https://doi.org/10.1038/ng.2830>
- Bønnelykke, K., Sparks, R., Waage, J., & Milner, J. D. (2015). Genetics of allergy and allergic sensitization: common variants, rare mutations. *Current Opinion in Immunology*, 36, 115-126. <https://doi.org/10.1016/j.coi.2015.08.002>
- Boulet, L.-P., Turcotte, H., Laprise, C., Lavertu, C., Bedard, P.-M., Lavoie, A., & Hébert, J. (1997). Comparative degree and type of sensitization to common indoor and outdoor allergens in subjects with allergic rhinitis and/or asthma. *Clinical & Experimental Allergy*, 27(1), 52-59. <https://doi.org/10.1111/j.1365-2222.1997.tb00672.x>
- Bousquet, J., Anto, J. M., Bachert, C., Baiardini, I., Bosnic-Anticevich, S., Canonica, G. W., Melén, E., Palomares, O., Scadding, G. K., Togias, A., & Toppila-Salmi, S. (2020). Allergic rhinitis. *Nature Reviews Disease Primers*, 6(1), 95. <https://doi.org/10.1038/s41572-020-00227-0>
- Bousquet, J., Melén, E., Haahtela, T., Koppelman, G. H., Togias, A., Valenta, R., Akdis, C. A., Czarlewski, W., Rothenberg, M., Valiulis, A., Wickman, M., Akdis, M., Aguilar, D., Bedbrook, A., Bindslev-Jensen, C., Bosnic-Anticevich, S., Boulet, L. P., Brightling, C. E., Brussino, L., ... Anto, J. M. (2023). Rhinitis associated with asthma is distinct from rhinitis alone: The ARIA-MeDALL hypothesis. *Allergy*, 78(5), 1169-1203. <https://doi.org/10.1111/all.15679>
- Bousquet, J., Van Cauwenberge, P., Khaltaev, N., & World Health Organization. (2001). Allergic Rhinitis and its Impact on Asthma. *The Journal of Allergy and Clinical Immunology*, 108(5 Suppl), S147-S334. <https://doi.org/10.1067/mai.2001.118891>
- Bousquet, P.-J., Chinn, S., Janson, C., Kogevinas, M., Burney, P., Jarvis, D., & European Community Respiratory Health Survey I. (2007). Geographical Variation in the Prevalence of Positive Skin Tests to Environmental Aeroallergens in the European Community Respiratory Health Survey I. *Allergy*, 62(3), 301-309. <https://doi.org/10.1111/j.1398-9995.2006.01293.x>
- Braman, S. S. (2006). The Global Burden Of Asthma. *CHEST*, 130(1 Suppl), 4S-12S. https://doi.org/10.1378/chest.130.1_suppl.4S
- Brandt, E. B., Myers, J. M. B., Ryan, P. H., & Hershey, G. K. K. (2015). Air pollution and allergic diseases. *Current Opinion in Pediatrics*, 27(6), 724-735. <https://doi.org/10.1097/MOP.0000000000000286>
- Brough, H. A., Nadeau, K. C., Sindher, S. B., Alkotob, S. S., Chan, S., Bahnson, H. T., Leung, D. Y. M., & Lack, G. (2020). Epicutaneous sensitization in the development of food allergy: what is the evidence and how can this be prevented? *Allergy*, 75(9), 2185-2205. <https://doi.org/10.1111/all.14304>

- Burbank, A. J., Sood, A. K., Kesic, M. J., Peden, D. B., & Hernandez, M. L. (2017). Environmental determinants of allergy and asthma in early life. *Journal of Allergy and Clinical Immunology*, 140(1), 1-12. <https://doi.org/10.1016/j.jaci.2017.05.010>
- Burge, H. A., & Rogers, C. A. (2000). Outdoor allergens. *Environmental Health Perspectives*, 108, 653-659. <https://doi.org/10.1289/ehp.00108s4653>
- Cano-Gamez, E., & Trynka, G. (2020). From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. *Frontiers in Genetics*, 11, 424. <https://doi.org/10.3389/fgene.2020.00424>
- Cecchi, L., D'Amato, G., & Annesi-Maesano, I. (2018). External exposome and allergic respiratory and skin diseases. *Journal of Allergy and Clinical Immunology*, 141(3), 846-857. <https://doi.org/10.1016/j.jaci.2018.01.016>
- Chan-Yeung, M., Anthonisen, N. R., Becklake, M. R., Bowie, D., Sonia Buist, A., Dimich-Ward, H., Ernst, P., Sears, M. R., Siersted, H. C., Sweet, L., Van Til, L., & Manfreda, J. (2010). Geographical variations in the prevalence of atopic sensitization in six study sites across Canada. *Allergy*, 65(11), 1404-1413. <https://doi.org/10.1111/j.1398-9995.2010.02399.x>
- Choi, B. Y., Han, M., Kwak, J. W., & Kim, T. H. (2021). Genetics and Epigenetics in Allergic Rhinitis. *Genes (Basel)*, 12(12), 2004. <https://doi.org/10.3390/genes12122004>
- Choi, J.-P., Park, S. Y., Moon, K.-A., Ha, E. H., Woo, Y. D., Chung, D. H., Kwon, H.-S., Kim, T.-B., Park, H.-S., Moon, H.-B., Song, W.-J., & Cho, Y. S. (2020). Macrophage-derived progranulin promotes allergen-induced airway inflammation. *Allergy*, 75(5), 1133-1145. <https://doi.org/10.1111/all.14129>
- Curin, M., & Hilger, C. (2017). Allergy to pets and new allergies to uncommon pets. *Allergologie Select*, 1(2), 214-221. <https://doi.org/10.5414/ALX01842E>
- Czarnowicki, T., Krueger, J. G., & Guttman-Yassky, E. (2017). Novel concepts of prevention and treatment of atopic dermatitis through barrier and immune manipulations with implications for the atopic march. *Journal of Allergy and Clinical Immunology*, 139(6), 1723-1734. <https://doi.org/10.1016/j.jaci.2017.04.004>
- D'Amato, G., Chong-Neto, H. J., Monge Ortega, O. P., Vitale, C., Ansotegui, I., Rosario, N., Haahtela, T., Galan, C., Pawankar, R., Murrieta-Aguttes, M., Cecchi, L., Bergmann, C., Ridolo, E., Ramon, G., Diaz, S. G., D'Amato, M., & Annesi-Maesano, I. (2020). The effects of climate change on respiratory allergy and asthma induced by pollen and mold allergens. *Allergy*, 75(9), 2219-2228. <https://doi.org/10.1111/all.14476>
- D'Amato, G., Holgate, S. T., Pawankar, R., Ledford, D. K., Cecchi, L., Al-Ahmad, M., Al-Enezi, F., Al-Muhsen, S., Ansotegui, I., Baena-Cagnani, C. E., Baker, D. J., Bayram, H., Bergmann, K. C., Boulet, L.-P., Buters, J. T. M., D'Amato, M., Dorsano, S., Douwes, J., Finlay, S. E., ... Annesi-Maesano, I. (2015). Meteorological conditions, climate change, new emerging factors, and asthma and related allergic disorders. A statement of the World Allergy Organization. *World Allergy Organization Journal*, 8(1), 1-52. <https://doi.org/10.1186/s40413-015-0073-0>
- Darling, N. J., Toth, R., Arthur, J. S. C., & Clark, K. (2017). Inhibition of SIK2 and SIK3 during differentiation enhances the anti-inflammatory phenotype of macrophages. *Biochemical Journal*, 474(4), 521-537. <https://doi.org/10.1042/BCJ20160646>
- Das, S., Miller, M., & Broide, D. H. (2017). Chapter One - Chromosome 17q21 Genes ORMDL3 and GSDMB in Asthma and Immune Diseases. Dans F. W. Alt (Éd.), *Advances in Immunology* (Vol. 135, pp. 1-52). Academic Press. <https://doi.org/10.1016/bs.ai.2017.06.001>

- Deménais, F., Margaritte-Jeannin, P., Barnes, K. C., Cookson, W. O. C., Altmüller, J., Ang, W., Barr, R. G., Beaty, T. H., Becker, A. B., Beilby, J., Bisgaard, H., Björnsdóttir, U. S., Bleecker, E., Bønnelykke, K., Boomsma, D. I., Bouzigon, E., Brightling, C. E., Brossard, M., Brusselle, G. G., ... Australian Asthma Genetics Consortium (AAGC) collaborators. (2018). Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. *Nature Genetics*, 50(1), 42-53. <https://doi.org/10.1038/s41588-017-0014-7>
- Demers, I. (2013). *État des connaissances sur le pollen et les allergies : les assises pour une gestion efficace*. Institut national de santé publique du Québec. <https://www.inspq.qc.ca/publications/1678>
- DeVries, A., & Vercelli, D. (2015). Epigenetics in allergic diseases. *Current Opinion in Pediatrics*, 27(6), 719-723. <https://doi.org/10.1097/MOP.0000000000000285>
- Doucet, M., Beaudoin, C., Laouan Sidi, Anassour, E., & Bureau d'information et d'études en santé des populations. (2020). *Portrait québécois de l'asthme : prévalence, incidence et mortalité de 2001 à 2016 (publication no 2669)*. Institut national de santé publique du Québec. https://www.inspq.qc.ca/sites/default/files/publications/2671_asthme_prevalence_incidence_mortalite.pdf
- Drazdauskaitė, G., Layhadi, J. A., & Shamji, M. H. (2021). Mechanisms of Allergen Immunotherapy in Allergic Rhinitis. *Current Allergy and Asthma Reports*, 21(1), 2. <https://doi.org/10.1007/s11882-020-00977-7>
- Eguiluz-Gracia, I., Mathioudakis, A. G., Bartel, S., Vijverberg, S. J. H., Fuertes, E., Comberiati, P., Cai, Y. S., Tomazic, P. V., Diamant, Z., Vestbo, J., Galan, C., & Hoffmann, B. (2020). The need for clean air: The way air pollution and climate change affect allergic rhinitis and asthma. *Allergy*, 75(9), 2170-2184. <https://doi.org/10.1111/all.14177>
- Eifan, A. O., & Durham, S. R. (2016). Pathogenesis of rhinitis. *Clinical & Experimental Allergy*, 46(9), 1139-1151. <https://doi.org/10.1111/cea.12780>
- Ferreira, F., Hawranek, T., Gruber, P., Wopfner, N., & Mari, A. (2004). Allergic cross-reactivity: from gene to the clinic. *Allergy*, 59(3), 243-267. <https://doi.org/10.1046/j.1398-9995.2003.00407.x>
- Ferreira, M. A. R. (2004). Inflammation in allergic asthma: Initiating events, immunological response and risk factors. *Respirology*, 9(1), 16-24. <https://doi.org/10.1111/j.1440-1843.2003.00516.x>
- Ferreira, M. A. R., Matheson, M. C., Duffy, D. L., Marks, G. B., Hui, J., Le Souëf, P., Danoy, P., Baltic, S., Nyholt, D. R., Jenkins, M., Hayden, C., Willemsen, G., Ang, W., Kuokkanen, M., Beilby, J., Cheah, F., de Geus, E. J. c., Ramasamy, A., Vedantam, S., ... Australian Asthma Genetics Consortium. (2011). Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *The Lancet*, 378(9795), 1006-1014. [https://doi.org/10.1016/S0140-6736\(11\)60874-X](https://doi.org/10.1016/S0140-6736(11)60874-X)
- Foltz, D. R., Jansen, L. E. T., Black, B. E., Bailey, A. O., Yates 3rd, J. R., & Cleveland, D. W. (2006). The human CENP-A centromeric nucleosome-associated complex. *Nature Cell Biology*, 8(5), 458-469. <https://doi.org/10.1038/ncb1397>
- Gadermaier, G., Dedic, A., Obermeyer, G., Frank, S., Himly, M., & Ferreira, F. (2004). Biology of Weed Pollen Allergens. *Current Allergy and Asthma Reports*, 4(5), 391-400. <https://doi.org/10.1007/s11882-004-0090-5>
- García-Mozo, H. (2017). Poaceae pollen as the leading aeroallergen worldwide: A review. *Allergy*, 72(12), 1849-1858. <https://doi.org/10.1111/all.13210>

- Gauvreau, G. M., El-Gammal, A. I., & O'Byrne, P. M. (2015). Allergen-induced airway responses. *European Respiratory Journal*, 46(3), 819-831. <https://doi.org/10.1183/13993003.00536-2015>
- Gelfand, E. W. (2004). Inflammatory mediators in allergic rhinitis. *Journal of Allergy and Clinical Immunology*, 114(5 Suppl), S135-S138. <https://doi.org/10.1016/j.jaci.2004.08.043>
- Gleason, K. J., Yang, F., Pierce, B. L., He, X., & Chen, L. S. (2020). Primo: integration of multiple GWAS and omics QTL summary statistics for elucidation of molecular mechanisms of trait-associated SNPs and detection of pleiotropy in complex traits. *Genome Biology*, 21, 236. <https://doi.org/10.1186/s13059-020-02125-w>
- Gouvernement du Québec. (13/06/2023). *Rhinite saisonnière (rhume des foins)*. Repéré le 13/03/2024 à <https://www.quebec.ca/sante/problemes-de-sante/a-z/rhinite-saisonniere-rhume-des-foins>
- Gray-French, M., Fernandes, R. M., Sinha, I. P., & Abrams, E. M. (2022). Allergen Management in Children with Type 2-High Asthma. *Journal of Asthma and Allergy*, 15, 381-394. <https://doi.org/10.2147/JAA.S276994>
- Greulich, W., Wagner, M., Gaidt, M. M., Stafford, C., Cheng, Y., Linder, A., Carell, T., & Hornung, V. (2019). TLR8 is a sensor of RNase T2 degradation products. *Cell*, 179(6), 1264-1275.e1213. <https://doi.org/10.1016/j.cell.2019.11.001>
- Gurdasani, D., Barroso, I., Zeggini, E., & Sandhu, M. S. (2019). Genomics of disease risk in globally diverse populations. *Nature Reviews Genetics*, 20(9), 520-535. <https://doi.org/10.1038/s41576-019-0144-0>
- Hall, S., & Agrawal, D. K. (2014). Key Mediators in the Immunopathogenesis of Allergic Asthma. *International Immunopharmacology*, 23(1), 316-329. <https://doi.org/10.1016/j.intimp.2014.05.034>
- Hamilton, D., & Lehman, H. (2020). Asthma Phenotypes as a Guide for Current and Future Biologic Therapies. *Clinical Reviews in Allergy & Immunology*, 59(2), 160-174. <https://doi.org/10.1007/s12016-019-08760-x>
- Han, H., Roan, F., & Ziegler, S. F. (2017). The atopic march: current insights into skin barrier dysfunction and epithelial cell-derived cytokines. *Immunological Reviews*, 278(1), 116-130. <https://doi.org/10.1111/imr.12546>
- Hansen, I., Klimek, L., Mösges, R., & Hörmann, K. (2004). Mediators of inflammation in the early and late phase of allergic rhinitis. *Current Opinion in Allergy and Clinical Immunology*, 4(3), 159-163. <https://doi.org/10.1097/00130832-200406000-00004>
- Hayakawa, M., Matsushima, M., Hagiwara, H., Oshima, T., Fujino, T., Ando, K., Kikugawa, K., Tanaka, H., Miyazawa, K., & Kitagawa, M. (2008). Novel insights into FGD3, a putative GEF for Cdc42, that undergoes SCFFWD1/β-TrCP-mediated proteasomal degradation analogous to that of its homologue FGD1 but regulates cell morphology and motility differently from FGD1. *Genes to Cells*, 13(4), 329-342. <https://doi.org/10.1111/j.1365-2443.2008.01168.x>
- Hill, D. A., & Spergel, J. M. (2018). The atopic march: Critical evidence and clinical relevance. *Annals of Allergy, Asthma & Immunology*, 120(2), 131-137. <https://doi.org/10.1016/j.anai.2017.10.037>
- Hinds, D. A., McMahon, G., Kiefer, A. K., Do, C. B., Eriksson, N., Evans, D. M., St Pourcain, B., Ring, S. M., Moutain, J. L., Francke, U., Davey-Smith, G., Timpson, N. J., & Tung, J. Y. (2013). A

- genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci. *Nature Genetics*, 45(8), 907-911. <https://doi.org/10.1038/ng.2686>
- Holgate, S. T. (2012). Innate and adaptive immune responses in asthma. *Nature Medicine*, 18(5), 673-683. <https://doi.org/10.1038/nm.2731>
- Holgate, S. T., Wenzel, S., Postma, D. S., Weiss, S. T., Renz, H., & Sly, P. D. (2015). Asthma. *Nature Reviews Disease Primers*, 1(1), 15025. <https://doi.org/10.1038/nrdp.2015.25>
- Hori, T., Shang, W.-H., Takeuchi, K., & Fukagawa, T. (2013). The CCAN recruits CENP-A to the centromere and forms the structural core for kinetochore assembly. *Journal of Cell Biology*, 200(1), 45-60. <https://doi.org/10.1083/jcb.201210106>
- Hughes, K. M., Price, D., & Suphioglu, C. (2022). Importance of allergen–environment interactions in epidemic thunderstorm asthma. *Therapeutic Advances in Respiratory Disease*, 16, 17534666221099733. <https://doi.org/10.1177/17534666221099733>
- Ismaila, A., Blais, L., Dang-Tan, T., Zhang, S., Forget, A., Kettani, F.-Z., Corriveau, D., Lemiere, C., & Registre de données en Santé Pulmonaire (RESP) investigators. (2019). Direct and indirect costs associated with moderate and severe asthma in Quebec, Canada. *Canadian Journal of Respiratory, Critical Care, and Sleep Medicine*, 3(3), 134-142. <https://doi.org/10.1080/24745332.2018.1544839>
- Ismaila, A. S., Sayani, A. P., Marin, M., & Su, Z. (2013). Clinical, economic, and humanistic burden of asthma in Canada: a systematic review. *BMC Pulmonary Medicine*, 13, 70. <https://doi.org/10.1186/1471-2466-13-70>
- Jones, B., Milstien, S., & Spiegel, S. (2019). ORMDL3 and allergic asthma: From physiology to pathology. *Journal of Allergy and Clinical Immunology*, 144(3), 634-640. <https://doi.org/10.1016/j.jaci.2019.07.023>
- Keith, P. K., Desrosiers, M., Laister, T., Schellenberg, R. R., & Wasserman, S. (2012). The burden of allergic rhinitis (AR) in Canada: perspectives of physicians and patients. *Allergy, Asthma & Clinical Immunology*, 8(1), 7. <https://doi.org/10.1186/1710-1492-8-7>
- Keller, T., Hohmann, C., Standl, M., Wijga, A. H., Gehring, U., Melén, E., Almqvist, C., Lau, S., Eller, E., Wahn, U., Christiansen, E. S., Von Berg, A., Heinrich, J., Lehmann, I., Maier, D., Postma, D. S., Antó, J. M., Bousquet, J., Keil, T., & Roll, S. (2018). The sex-shift in single disease and multimorbid asthma and rhinitis during puberty - a study by MeDALL. *Allergy*, 73(3), 602-614. <https://doi.org/10.1111/all.13312>
- Kim, H. Y., Lee, H. J., Chang, Y.-J., Pichavant, M., Shore, S. A., Fitzgerald, K. A., Iwakura, Y., Israel, E., Bolger, K., Faul, J., DeKruyff, R. H., & Umetsu, D. T. (2014). IL-17 producing innate lymphoid cells and the NLRP3 inflammasome facilitate obesity-associated airway hyperreactivity. *Nature Medicine*, 20(1), 54-61. <https://doi.org/10.1038/nm.3423>
- Kim-Hellmuth, S., Aguet, F., Oliva, M., Muñoz-Aguirre, M., Kasela, S., Wucher, V., Castel, S. E., Hamel, A. R., Viñuela, A., Roberts, A. L., Mangul, S., Wen, X., Wang, G., Barbeira, A. N., Garrido-Martín, D., Nadel, B. B., Zou, Y., Bonazzola, R., Quan, J., ... Lappalainen, T. (2020). Cell type–specific genetic regulation of gene expression across human tissues. *Science*, 369(6509), eaaz8528. <https://doi.org/10.1126/science.aaz8528>
- Komlósi, Z. I., van de Veen, W., Kovács, N., Szűcs, G., Sokolowska, M., O'Mahony, L., Akdis, M., & Akdis, C. A. (2022). Cellular and molecular mechanisms of allergic asthma. *Molecular Aspects of Medicine*, 85, 100995. <https://doi.org/10.1016/j.mam.2021.100995>

- Kuruville, M. E., Lee, F. E.-H., & Lee, G. B. (2019). Understanding Asthma Phenotypes, Endotypes, and Mechanisms of Disease. *Clinical Reviews in Allergy & Immunology*, 56(2), 219-233. <https://doi.org/10.1007/s12016-018-8712-1>
- Lei, D. K., & Grammer, L. C. (2019). An Overview of Allergens. *Allergy and Asthma Proceedings*, 40(6), 362-365. <https://doi.org/10.2500/aap.2019.40.4247>
- Li, B., & Ritchie, M. D. (2021). From GWAS to Gene: Transcriptome-Wide Association Studies and Other Methods to Functionally Understand GWAS Discoveries. *Frontiers in Genetics*, 12, 713230. <https://doi.org/10.3389/fgene.2021.713230>
- Li, J., Sun, B., Huang, Y., Lin, X., Zhao, D., Tan, G., Wu, J., Zhao, H., Cao, L., Zhong, N., & China Alliance of Research on Respiratory Allergic Disease (CARRAD). (2009). A multicentre study assessing the prevalence of sensitizations in patients with asthma and/or rhinitis in China. *Allergy*, 64(7), 1083-1092. <https://doi.org/10.1111/j.1398-9995.2009.01967.x>
- Liu, B., & Montgomery, S. B. (2020). Identifying causal variants and genes using functional genomics in specialized cell types and contexts. *Human Genetics*, 139(1), 95-102. <https://doi.org/10.1007/s00439-019-02044-2>
- Lombardi, C., Savi, E., Ridolo, E., Passalacqua, G., & Canonica, G. W. (2017). Is allergic sensitization relevant in severe asthma? Which allergens may be culprits? *World Allergy Organization Journal*, 10(1), 2. <https://doi.org/10.1186/s40413-016-0138-8>
- Madore, A.-M., Pain, L., Boucher-Lafleur, A.-M., Morin, A., Meloche, J., Simon, M.-M., Ge, B., Kwan, T., Cheung, W. A., Pastinen, T., & Laprise, C. (2020). Asthma-associated polymorphisms in 17q12-21 locus modulate methylation and gene expression of GSDMA in naïve CD4+ T cells. *Journal of Genetics and Genomics*, 47(3), 171-174. <https://doi.org/10.1016/j.jgg.2020.03.002>
- Madore, A.-M., Tremblay, K., Hudson, T. J., & Laprise, C. (2008). Replication of an association between 17q21 SNPs and asthma in a French-Canadian familial collection. *Human Genetics*, 123(1), 93-95. <https://doi.org/10.1007/s00439-007-0444-x>
- Matricardi, P. M., Kleine-Tebbe, J., Hoffmann, H. J., Valenta, R., Hilger, C., Hofmaier, S., Aalberse, R. C., Agache, I., Asero, R., Ballmer-Weber, B., Barber, D., Beyer, K., Biedermann, T., Bilò, M. B., Blank, S., Bohle, B., Bosshard, P. P., Breiteneder, H., Brough, H. A., ... Ollert, M. (2016). EAACI Molecular Allergology User's Guide. *Pediatric Allergy and Immunology*, 27(Suppl 23), 1-250. <https://doi.org/10.1111/pai.12563>
- Meng, Y., Wang, C., & Zhang, L. (2020). Advances and novel developments in allergic rhinitis. *Allergy*, 75(12), 3069-3076. <https://doi.org/10.1111/all.14586>
- Miller, J. D. (2019). The Role of Dust Mites in Allergy. *Clinical Reviews in Allergy & Immunology*, 57(3), 312-329. <https://doi.org/10.1007/s12016-018-8693-0>
- Miranda, T. B., & Jones, P. A. (2007). DNA methylation: The nuts and bolts of repression. *Journal Of Cellular Physiology*, 213(2), 384-390. <https://doi.org/10.1002/jcp.21224>
- Morin, A., Laviolette, M., Pastinen, T., Boulet, L.-P., & Laprise, C. (2017). Combining omics data to identify genes associated with allergic rhinitis. *Clinical Epigenetics*, 9, 3. <https://doi.org/10.1186/s13148-017-0310-1>
- Nadeau, K., McDonald-Hyman, C., Noth, E. M., Pratt, B., Hammond, S. K., Balmes, J., & Tager, I. (2010). Ambient air pollution impairs regulatory T-cell function in asthma. *Journal of Allergy and Clinical Immunology*, 126(4), 845-852.e810. <https://doi.org/10.1016/j.jaci.2010.08.008>

- Narasimamurthy, R., Hatori, M., Nayak, S. K., Liu, F., Panda, S., & Verma, I. M. (2012). Circadian clock protein cryptochrome regulates the expression of proinflammatory cytokines. *Proceedings of the National Academy of Sciences of the United States of America*, 109(31), 12662-12667. <https://doi.org/10.1073/pnas.1209965109>
- Neefjes, J., Jonqsmas, M. L. M., Paul, P., & Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews Immunology*, 11(12), 823-836. <https://doi.org/10.1038/nri3084>
- Nica, A. C., & Dermitzakis, E. T. (2013). Expression quantitative trait loci: present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1620), 20120362. <https://doi.org/10.1098/rstb.2012.0362>
- Nilsson, D., Henmyr, V., Halldén, C., Säll, T., Kull, I., Wickman, M., Melén, E., & Cardell, L. O. (2014). Replication of genomewide associations with allergic sensitization and allergic rhinitis. *Allergy*, 69(11), 1506-1514. <https://doi.org/10.1111/all.12495>
- Ntontsi, P., Photiades, A., Zervas, E., Xanthou, G., & Samitas, K. (2021). Genetics and Epigenetics in Asthma. *International Journal of Molecular Sciences*, 22(5), 2412. <https://doi.org/10.3390/ijms22052412>
- Nur Husna, S. M., Tan, H.-T. T., Md Shukri, N., Mohd Ashari, N. S., & Wong, K. K. (2022). Allergic Rhinitis: A Clinical and Pathophysiological Overview. *Frontiers in Medicine (Lausanne)*, 9, 874114. <https://doi.org/10.3389/fmed.2022.874114>
- Ober, C., & Yao, T.-C. (2011). The Genetics of Asthma and Allergic Disease: A 21st Century Perspective. *Immunological Reviews*, 242(1), 10-30. <https://doi.org/10.1111/j.1600-065X.2011.01029.x>
- Okada, M., Cheeseman, I. M., Hori, T., Okawa, K., McLeod, I. X., Yates 3rd, J. R., Desai, A., & Fukagawa, T. (2006). The CENP-H-I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. *Nature Cell Biology*, 8(5), 446-457. <https://doi.org/10.1038/ncb1396>
- Pablos, I., Wildner, S., Asam, C., Wallner, M., & Gadermaier, G. (2016). Pollen Allergens for Molecular Diagnosis. *Current Allergy and Asthma Reports*, 16, 31. <https://doi.org/10.1007/s11882-016-0603-z>
- Pakkasela, J., Ilmarinen, P., Honkamäki, J., Tuomisto, L. E., Andersén, H., Piirilä, P., Hisinger-Mölkänen, H., Sovijärvi, A., Backman, H., Lundbäck, B., Rönmark, E., Kankaanranta, H., & Lehtimäki, L. (2020). Age-specific incidence of allergic and non-allergic asthma. *BMC Pulmonary Medicine*, 20(1), 9. <https://doi.org/10.1186/s12890-019-1040-2>
- Palomares, O., Akdis, M., Martín-Fontecha, M., & Akdis, C. A. (2017). Mechanisms of immune regulation in allergic diseases: the role of regulatory T and B cells. *Immunological Reviews*, 278(1), 219-236. <https://doi.org/10.1111/imr.12555>
- Pascual, M., Suzuki, M., Isidoro-Garcia, M., Padrón, J., Turner, T., Lorente, F., Dávila, I., & Grealley, J. M. (2011). Epigenetic changes in B lymphocytes associated with house dust mite allergic asthma. *Epigenetics*, 6(9), 1131-1137. <https://doi.org/10.4161/epi.6.9.16061>
- Patterson, K., & Streck, M. E. (2010). Allergic Bronchopulmonary Aspergillosis. *Proceedings of the American Thoracic Society*, 7(3), 237-244. <https://doi.org/10.1513/pats.200908-086AL>

- Perosa, F., Prete, M., Di Lernia, G., Ostuni, C., Favoino, E., & Valentini, G. (2016). Anti-centromere protein A antibodies in systemic sclerosis: Significance and origin. *Autoimmunity Reviews*, 15(1), 102-109. <https://doi.org/10.1016/j.autrev.2015.10.001>
- Peters-Golden, M., Gleason, M. M., & Togias, A. (2006). Cysteinyl leukotrienes: multi-functional mediators in allergic rhinitis. *Clinical & Experimental Allergy*, 36(6), 689-703. <https://doi.org/10.1111/j.1365-2222.2006.02498.x>
- Pinart, M., Keller, T., Reich, A., Frölich, M., Cabieses, B., Hohmann, C., Postma, D. S., Bousquet, J., Antó, J. M., & Keil, T. (2017). Sex-Related Allergic Rhinitis Prevalence Switch from Childhood to Adulthood: A Systematic Review and Meta-Analysis. *International Archives of Allergy and Immunology*, 172(4), 224-235. <https://doi.org/10.1159/000464324>
- Pomés, A., Chapman, M. D., & Wünschmann, S. (2016). Indoor Allergens and Allergic Respiratory Disease. *Current Allergy and Asthma Reports*, 16(6), 43. <https://doi.org/10.1007/s11882-016-0622-9>
- Portelli, M. A., Hodge, E., & Sayers, I. (2015). Genetic risk factors for the development of allergic disease identified by genome-wide association. *Clinical & Experimental Allergy*, 45(1), 21-31. <https://doi.org/10.1111/cea.12327>
- Potaczek, D. P., Harb, H., Michel, S., Alhamwe, B. A., Renz, H., & Tost, J. (2017). Epigenetics and allergy: from basic mechanisms to clinical applications. *Epigenomics*, 9(4), 539-571. <https://doi.org/10.2217/epi-2016-0162>
- Ramasamy, A., Curjuric, I., Coin, L. J., Kumar, A., McArdle, W. L., Imboden, M., Leynaert, B., Kogevinas, M., Schmid-Grendelmeier, P., Pekkanen, J., Wjst, M., Bircher, A. J., Sovio, U., Rochat, T., Hartikainen, A.-L., Balding, D. J., Jarvelin, M.-R., Probst-Hensch, N., Strachan, D. P., & Jarvis, D. L. (2011). A genome-wide meta-analysis of genetic variants associated with allergic rhinitis and grass sensitization and their interaction with birth order. *Journal of Allergy and Clinical Immunology*, 128(5), 996-1005. <https://doi.org/10.1016/j.jaci.2011.08.030>
- Samitas, K., Carter, A., Kariyawasam, H. H., & Xanthou, G. (2018). Upper and lower airway remodelling mechanisms in asthma, allergic rhinitis and chronic rhinosinusitis: The one airway concept revisited. *Allergy*, 73(5), 993-1002. <https://doi.org/10.1111/all.13373>
- Sánchez, P., Vélez-del-Burgo, A., Suñén, E., Martínez, J., & Postigo, I. (2022). Fungal Allergen and Mold Allergy Diagnosis: Role and Relevance of *Alternaria alternata* Alt a 1 Protein Family. *Journal of Fungi (Basel)*, 8(3), 277. <https://doi.org/10.3390/jof8030277>
- Sastre, J. (2010). Molecular diagnosis in allergy. *Clinical & Experimental Allergy*, 40(10), 1442-1460. <https://doi.org/10.1111/j.1365-2222.2010.03585.x>
- Šaulienė, I., Greičiuvienė, J., Šukienė, L., Juškevičiūtė, N., Benner, C., Zinkevičienė, A., Ripatti, S., Donner, K., & Kainov, D. (2015). Genetic Loci Associated with Allergic Sensitization in Lithuanians. *PLoS ONE*, 10(7), e0134188. <https://doi.org/10.1371/journal.pone.0134188>
- Scadding, G. (2014). Cytokine Profiles in Allergic Rhinitis. *Current Allergy and Asthma Reports*, 14(5), 435. <https://doi.org/10.1007/s11882-014-0435-7>
- Scadding, G. W., Eifan, A., Penagos, M., Dumitru, A., Switzer, A., McMahon, O., Phippard, D., Togias, A., Durham, S. R., & Shamji, M. H. (2015). Local and systemic effects of cat allergen nasal provocation. *Clinical & Experimental Allergy*, 45(3), 613-623. <https://doi.org/10.1111/cea.12434>

- Scaldeferri, D., Bosi, A., Fabbri, M., Pedrini, E., Inforzato, A., Valli, R., Frattini, A., De Vito, A., Noonan, D. M., Taramelli, R., Mortara, L., & Acquati, F. (2018). The human RNASET2 protein affects the polarization pattern of human macrophages in vitro. *Immunology Letters*, 203, 102-111. <https://doi.org/10.1016/j.imlet.2018.09.005>
- Schatz, M., & Rosenwasser, L. (2014). The allergic asthma phenotype. *The Journal of Allergy and Clinical Immunology: In Practice*, 2(6), 645-648. <https://doi.org/10.1016/j.jaip.2014.09.004>
- Seguel, J. M., Merrill, R., Seguel, D., & Campagna, A. C. (2016). Indoor Air Quality. *American Journal of Lifestyle Medicine*, 11(4), 284-295. <https://doi.org/10.1177/1559827616653343>
- Settipane, R. A., & Schwindt, C. (2013). Chapter 15: Allergic rhinitis. *American Journal of Rhinology & Allergy*, 27(3 suppl 1), S52-S55. <https://doi.org/10.2500/ajra.2013.27.3928>
- Shah, R., & Newcomb, D. C. (2018). Sex Bias in Asthma Prevalence and Pathogenesis. *Frontier in Immunology*, 9, 2997. <https://doi.org/10.3389/fimmu.2018.02997>
- Sierra-Heredia, C., North, M., Brook, J., Daly, C., Ellis, A. K., Henderson, D., Henderson, S. B., Lavigne, É., & Takaro, T. K. (2018). Aeroallergens in Canada: Distribution, Public Health Impacts, Opportunities for Prevention. *International Journal of Environmental Research and Public Health*, 15(8), 1577. <https://doi.org/10.3390/ijerph15081577>
- Singh, M., & Hays, A. (2016). Indoor and Outdoor Allergies. *Primary Care: Clinics in Office Practice*, 43(3), 451-463. <https://doi.org/10.1016/j.pop.2016.04.013>
- Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., Groza, T., Güneş, O., Hall, P., Hayhurst, J., Ibrahim, A., Ji, Y., John, S., Lewis, E., MacArthur, J. A. L., McMahon, A., Osumi-Sutherland, D., Panoutsopoulou, K., Pendlington, Z., ... Harris, L. W. (2023). The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Research*, 51(D1), D977-D985. <https://doi.org/10.1093/nar/gkac1010>
- Song, G., Hu, C., Zhu, H., Wang, L., Zhang, F., Li, Y., & Wu, L. (2013). New Centromere Autoantigens Identified in Systemic Sclerosis Using Centromere Protein Microarrays. *The Journal of Rheumatology*, 40(4), 461-468. <https://doi.org/10.3899/jrheum.120264>
- Statistique Canada. (2023). *Tableau 13-10-0096-01 Caractéristiques de la santé, estimations annuelles*. Repéré le 20/06/2023 à <https://doi.org/10.25318/1310009601-fra>
- Stern, J., Pier, J., & Litonjua, A. A. (2020). Asthma epidemiology and risk factors. *Seminars in Immunopathology*, 42(1), 5-15. <https://doi.org/10.1007/s00281-020-00785-1>
- Stewart, G. A., & Robinson, C. (2017). Chapter 4 - Indoor and Outdoor Allergens and Pollutants. Dans R. E. O'Hehir, S. T. Holgate, & A. Sheikh (Éds.), *Middleton's Allergy Essentials* (1, pp. 73-116). Elsevier. <https://doi.org/10.1016/B978-0-323-37579-5.00004-0>
- Stojanovic, A., Correia, M. P., & Cerwenka, A. (2018). The NKG2D/NKG2DL Axis in the Crosstalk Between Lymphoid and Myeloid Cells in Health and Disease. *Frontier in Immunology*, 9, 827. <https://doi.org/10.3389/fimmu.2018.00827>
- Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*, 20(8), 467-484. <https://pubmed.ncbi.nlm.nih.gov/31068683/>
- Tan, L., Qiu, T., Xiang, R., Cao, C., Deng, Y., Tao, Z., & Xu, Y. (2020). Down-regulation of Tet2 is associated with Foxp3 TSDR hypermethylation in regulatory T cell of allergic rhinitis. *Life Sciences*, 241, 117101. <https://doi.org/10.1016/j.lfs.2019.117101>

- Tardif, I. (2008). *Portrait des coûts de santé associés à l'allergie au pollen de l'herbe à poux*. Agence de la santé et des services sociaux de la Montérégie.
- Tardif, I., & Masson, E. (2008). *Prévalence et coûts de santé de l'allergie au pollen de l'herbe à poux*. Institut national de santé du Québec. <https://www.inspq.qc.ca/bise/prevalence-et-couts-de-sante-de-l-allergie-au-pollen-de-l-herbe-poux>
- Tesse, R., Pandey, R. C., & Kabesch, M. (2011). Genetic variations in toll-like receptor pathway genes influence asthma and atopy. *Allergy*, 66(3), 307-316. <https://doi.org/10.1111/j.1398-9995.2010.02489.x>
- Tham, E. H., Lee, A. J., & Van Bever, H. (2016). Aeroallergen sensitization and allergic disease phenotypes in Asia. *Asian Pacific Journal of Allergy and Immunology*, 34(3), 181-189. <https://doi.org/10.12932/AP0770>
- Toskala, E., & Kennedy, D. W. (2015). Asthma risk factors. *International Forum of Allergy & Rhinology*, 5(Suppl 1), S11-S16. <https://doi.org/10.1002/alar.21557>
- Twaroch, T. E., Curin, M., Valenta, R., & Swoboda, I. (2015). Mold Allergens in Respiratory Allergy: From Structure to Therapy. *Allergy, Asthma & Immunology Research*, 7(3), 205-220. <https://doi.org/10.4168/aair.2015.7.3.205>
- Uzzaman, A., & Cho, S. H. (2012). Chapter 28: Classification of hypersensitivity reactions. *Allergy and Asthma Proceedings*, 33(3 Suppl 1), S96-S99. <https://doi.org/10.2500/aap.2012.33.3561>
- Vierbuchen, T., Stein, K., & Heine, H. (2019). RNA is taking its Toll: Impact of RNA-specific Toll-like receptors on health and disease. *Allergy*, 74(2), 223-235. <https://doi.org/10.1111/all.13680>
- Vignola, A. M., Chanez, P., & Bousquet, J. (2003). The relationship between asthma and allergic rhinitis: exploring the basis for a common pathophysiology. *Clinical & Experimental Allergy Reviews*, 3(2), 63-68. <https://doi.org/10.1046/j.1472-9725.2003.00015.x>
- Virolainen, S. J., VonHandorf, A., Viel, K. C. M. F., Weirauch, M. T., & Kottyan, L. C. (2023). Gene-environment interactions and their impact on human health. *Genes and Immunity*, 24(1), 1-11. <https://doi.org/10.1038/s41435-022-00192-6>
- Waage, J., Standl, M., Curtin, J. A., Jessen, L. E., Thorsen, J., Tian, C., Schoettler, N., 23andMe Research Team, AAGC collaborators, Flores, C., Abdellaoui, A., Ahluwalia, T. S., Alves, A. C., Amaral, A. F. S., Antó, J. M., Arnold, A., Barreto-Luis, A., Baurecht, H., van Beijsterveldt, C. E. M., ... Bønnelykke, K. (2018). Genome-wide association and HLA fine-mapping studies identify risk loci and genetic pathways underlying allergic rhinitis. *Nature Genetics*, 50(8), 1072-1080. <https://doi.org/10.1038/s41588-018-0157-1>
- Wang, J., Zhou, Y., Zhang, H., Hu, L., Liu, J., Wang, L., Wang, T., Zhang, H., Cong, L., & Wang, Q. (2023). Pathogenesis of allergic diseases and implications for therapeutic interventions. *Signal Transduction and Targeted Therapy*, 8, 138. <https://doi.org/10.1038/s41392-023-01344-4>
- Wei, X., Zhang, Y., Fu, Z., & Zhang, L. (2013). The Association between Polymorphisms in the MRPL4 and TNF- α Genes and Susceptibility to Allergic Rhinitis. *PLoS One*, 8(3), e57981. <https://doi.org/10.1371/journal.pone.0057981>
- Werfel, T., Allam, J.-P., Biedermann, T., Eyerich, K., Gilles, S., Guttman-Yassky, E., Hoetzenecker, W., Knol, E., Simon, H.-U., Wollenberg, A., Bieber, T., Lauener, R., Schmid-Grendelmeier, P., Traidl-Hoffmann, C., & Akdis, C. A. (2016). Cellular and molecular immunologic

- mechanisms in patients with atopic dermatitis. *Journal of Allergy and Clinical Immunology*, 138(2), 336-349. <https://doi.org/10.1016/j.jaci.2016.06.010>
- Westra, H.-J., Arends, D., Esko, T., Peters, M. J., Schurmann, C., Schramm, K., Kettunen, J., Yaghootkar, H., Fairfax, B. P., Andiappan, A. K., Li, Y., Fu, J., Karjalainen, J., Platteel, M., Visschedijk, M., Weersma, R. K., Kasela, S., Milani, L., Tserel, L., ... Franke, L. (2015). Cell Specific eQTL Analysis without Sorting Cells. *PLoS Genetics*, 11(5), e1005223. <https://doi.org/10.1371/journal.pgen.1005223>
- Willis, S., Sun, Y., Abramovitz, M., Fei, T., Young, B., Lin, X., Ni, M., Achua, J., Regan, M. M., Gray, K. P., Gray, R., Wang, V., Long, B., Kammler, R., Sparano, J. A., Williams, C., Goldstein, L. J., Salgado, R., Loi, S., ... Leyland-Jones, B. (2017). High Expression of FGD3, a Putative Regulator of Cell Morphology and Motility, Is Prognostic of Favorable Outcome in Multiple Cancers. *JCO Precision Oncology*, 1, PO.17.00009. <https://doi.org/10.1200/po.17.00009>
- Wise, S. K., Damask, C., Roland, L. T., Ebert, C., Levy, J. M., Lin, S., Luong, A., Rodriguez, K., Sedaghat, A. R., Toskala, E., Villwock, J., Abdullah, B., Akdis, C., Alt, J. A., Ansotegui, I. J., Azar, A., Baroody, F., Benninger, M. S., Bernstein, J., ... Zhang, L. (2023). International consensus statement on allergy and rhinology: Allergic rhinitis – 2023. *International Forum of Allergy & Rhinology*, 13(4), 293-859. <https://doi.org/10.1002/alr.23090>
- Wise, S. K., Lin, S. Y., Toskala, E., Orlandi, R. R., Akdis, C. A., Alt, J. A., Azar, A., Baroody, F. M., Bachert, C., Canonica, G. W., Chacko, T., Cingi, C., Ciprandi, G., Corey, J., Cox, L. S., Creticos, P. S., Custovic, A., Damask, C., DeConde, A., ... Zacharek, M. (2018). International Consensus Statement on Allergy and Rhinology: Allergic Rhinitis. *International Forum of Allergy & Rhinology*, 8(2), 108-352. <https://doi.org/10.1002/alr.22073>
- World Health Organization/ International Union of Immunological Societies. (2007). *Allergen Nomenclature*. <http://www.allergen.org/index.php>
- Xie, L. (2014). MKL1/2 and ELK4 co-regulate distinct serum response factor (SRF) transcription programs in macrophages. *BMC Genomics*, 15, 301. <https://doi.org/10.1186/1471-2164-15-301>
- Yang, I. V., Pedersen, B. S., Liu, A., O'Connor, G. T., Teach, S. J., Kattan, M., Misiak, R. T., Gruchalla, R., Steinbach, S. F., Szeffler, S. J., Gill, M. A., Calatroni, A., David, G., Hennessy, C. E., Davidson, E. J., Zhang, W., Gergen, P., Togias, A., Busse, W. W., & Schwartz, D. A. (2015). DNA methylation and childhood asthma in the inner city. *Journal of Allergy and Clinical Immunology*, 136(1), 69-80. <https://doi.org/10.1016/j.jaci.2015.01.025>
- Yang, L., Fu, J., & Zhou, Y. (2020). Research Progress in Atopic March. *Frontiers in Immunology*, 11, 1907. <https://doi.org/10.3389/fimmu.2020.01907>
- Ye, Y., Zhang, Z., Liu, Y., Diao, L., & Han, L. (2020). A Multi-Omics Perspective of Quantitative Trait Loci in Precision Medicine. *Trends in Genetics*, 36(5), 318-336. <https://doi.org/10.1016/j.tig.2020.01.009>
- Yoshida, K., Adachi, Y., Sasaki, M., Akashi, M., Itazawa, T., Ohya, Y., Odajima, H., & Akasawa, A. (2017). Ultraviolet index is associated with childhood eczema prevalence: Sex and age differences. *Pediatric Allergy and Immunology*, 28(4), 391-393. <https://doi.org/10.1111/pai.12706>
- Zahradnik, E., & Raulf, M. (2014). Animal allergens and their presence in the environment. *Frontiers in Immunology*, 5, 76. <https://doi.org/10.3389/fimmu.2014.00076>

Zhou, Z., He, H., Wang, K., Shi, X., Wang, Y., Su, Y., Wang, Y., Li, D., Liu, W., Zhang, Y., Shen, L., Han, W., Shen, L., Ding, J., & Shao, F. (2020). Granzyme A from cytotoxic lymphocytes cleaves GSDMB to trigger pyroptosis in target cells. *Science*, 368(6494), eaaz7548. <https://doi.org/10.1126/science.aaz7548>

CERTIFICATION ÉTHIQUE

Le présent mémoire a fait l'objet de deux certifications éthiques. Les numéros des certificats sont : 0002-001 et 2015-035.

ANNEXE 1

Supplementary information

Genetic Risk Profiling for Indoor and Outdoor Aeroallergens Sensitization in A French-Canadian Asthma Family Cohort

Marie-Pier Bouchard, Christophe Linhard, Anne-Marie Madore, Emmanuelle Bouzigon and Catherine Laprise

Materials and Methods

Allergens

The 26 allergens tested during the skin prick test were classified into six categories which are animal, house dust mites (HDM), house dust, grass pollen, tree pollen and mold. The animal category was made up of the following allergens: cat (*Felis catus*; n=397 allergic individuals/1,010 individuals for which the data is available for this allergen), dog (*Canis familiaris*; n=192/1,010), horse (*Equus ferus caballus*; n=136/1,004) and cattle (*Bos taurus*; n=28/254) dander and bird feathers (n=51/985). The house dust only includes one general allergen (n=240/1,004). The HDM category included allergens for the *Dermatophagoides (D.) farinae* (n=346/1,009) and *D. pteronyssinus* (n=332/1,009). The grass pollen category was composed of grasses (*Poaceae*; n=271/1,009), weeds (n=143/1,004), ragweed (*Ambrosia*; n=126/1,009), timothy (*Phleum pratense*; n=227/999) and perennial ryegrass (*Lolium perenne*; n=189/1,002) pollen as allergens. As for the tree pollen category, the allergens tested were tree mix (n=181/1,001), birch (*Betula*; n=276/1,010), maple (*Acer*; n=91/1,008), oak (*Quercus*; n=95/999) and elm (*Ulmus*; n=71/1,008) pollen. For the mold category, the allergens tested were *Chladosporium* (n=58/1,004), *Alternaria* (n=63/1,004), *Aspergillus* (n=50/1,007) and *Penicillium* (n=49/765). There was a seventh category made of 16 food allergens, but that category was not used for any of the analyses of this study.

Immune cell isolation

The immune cells isolated for this study were the eosinophils and the naïve CD4⁺ T cells which was defined as units with CD3⁺, CD4⁺, CD45RA⁺, CD45RO⁻ markers. These cell types were chosen

considering their known involvement in asthma and allergy.¹ Naïve CD4⁺ T cells and eosinophils were isolated from 200 ml blood sample according to protocols already described.^{2,3} The blood samples came from a subset of 215 individuals from the Saguenay–Lac-Saint-Jean asthma cohort. Afterwards, a proportion of the cells (1E-06) was snap-frozen until DNA extraction and the same number of cells were suspended in Trizol for RNA extraction. DNA and RNA extraction from both CD4⁺ T cells and eosinophils were done using the DNeasy Blood & Tissue Kit and the RNeasy Mini Kit respectively following the company's instructions (Qiagen, Toronto, ON, Canada). Further details can be found in the supplemental data from Madore et al.⁴

Cell-type contribution and surrogate variables

In order to estimate the variation in purity of naïve CD4⁺ T cell and eosinophil samples, the cell-type contribution approach was used using whole genome methylation data for each sample. The method employed was developed by Houseman⁵ and implemented in RnBeads^{6,7} R package. Afterwards, proportion of naïve CD4⁺ T cells or eosinophils was used as covariates for eQTL and mQTL analyses. Also, surrogate variables were calculated using the sva package⁸ in R in order to take into account hidden confounders and relatedness between samples.

DNA methylation sequencing and levels

As previously described, bisulfite sequencing was performed for a custom methyl capture panel of already identified immune-linked genetic regions.^{9,10} A total of 4,609,564 Cytosine preceding Guanine (CpGs) were covered with this panel for a total of 822,884 regions and 119,089,296pb sequenced.¹⁰ In collaboration with R&D at Roche NimbleGen, the methyl-capture sequencing (MCC-Seq) methylation protocol was developed and optimized. CpGs were retained if they fulfilled several criteria: 1) being located in a region targeted by the custom panel (defined as each targeted region ± 100 kb), 2) having more than 5 reads, 3) not overlapping with a genetic variant location with a minor allele frequency (MAF) ≥ 0.01 and 4) not overlapping with a DAC Blacklisted Regions or Duke Excluded Regions generated by the ENCODE project. Samples with more than 1.5M CpGs and CpGs with data for a minimum of 120 samples were considered for analyses. Once quality control filtering was done, data were available for 192 samples for the naïve CD4⁺ T cells and 4,703,225 CpGs for

analyses. While the eosinophil data were available for 183 samples and 4,625,332 CpGs for analyses. The methylation levels were calculated as total (forward and reverse) non-converted C reads over (forward and reverse) reads. Detailed methods can be found in the supplemental data from Madore et al.⁴

RNA sequencing and levels

RNA sequencing was performed using 500 ng of RNA from each sample at the McGill University and Génome Québec Innovation Centre. The generated raw reads were filtered for quality (phred33 \geq 30) and length ($n\geq$ 32). Also, adapter sequences were removed with Trimmomatic v.0.32¹¹. After passing filters, the reads were then aligned to the human genome reference hg19 using TopHat v.2.0.10¹² and Bowtie v.2.1.0¹³. Only genes with a minimum of 10 reads in all samples were considered as expressed in each cell type. Once quality control filtering done and retaining only samples data that were available for all covariates, there were transcriptomic data available for 159 naïve CD4⁺ T cell samples and 17,366 genes for analyses. While the eosinophil data were available for 139 samples and 17,513 genes for analyses. More information on this method can be found in Madore et al.⁴

Regions for replication

The 200kb regions were calculated differently if there was only one significant variant in the locus or more. If there was only one significant variant in the locus, the regions were defined as \pm 100kb around the variant position. However, if there was more than one significant variant in the locus, the regions were defined as \pm 100kb around the first and last variant position mean. The only exception for this was the region located in the 5q12.3 locus in which 229 variants were associated to the indoor sensitization. Since there was a difference of 394,849bp between the first (rs356593) and the last (rs13162093) significant variant in this locus, this region was defined as starting at 63,340,373bp and ending at 63,735,222bp on the chromosome 5. It is important to note that one locus could have more than one region in it. This is due to the fact that certain variant had more than 400kb of distance between them and thus were separated into two regions.

References

1. Komlósi ZI, van de Veen W, Kovács N, Szűcs G, Sokolowska M, O'Mahony L, Akdis M, Akdis CA. Cellular and molecular mechanisms of allergic asthma. *Mol Aspects Med.* 2022;85:100995. doi: 10.1016/j.mam.2021.100995
2. Guilbert M, Ferland C, Bossé M, Flamand N, Lavigne S, Laviolette M. 5-Oxo-6,8,11,14-eicosatetraenoic acid induces important eosinophil transmigration through basement membrane components: comparison of normal and asthmatic eosinophils. *Am J Respir Cell Mol Biol.* 1999;21(1):97-104. doi: 10.1165/ajrcmb.21.1.3517
3. Ferland C, Guilbert M, Davoine F, Flamand N, Chakir J, Laviolette M. Eotaxin promotes eosinophil transmigration via the activation of the plasminogen-plasmin system. *J Leukoc Biol.* 2001;69(5):772-778.
4. Madore A-M, Pain L, Boucher-Lafleur A-M, Morin A, Meloche J, Simon M-M, Ge B, Kwan T, Cheung WA, Pastinen T, et al. Asthma-associated polymorphisms in 17q12-21 locus modulate methylation and gene expression of GSDMA in naïve CD4+ T cells. *J Genet Genomics.* 2020;47(3):171-174. doi: 10.1016/j.jgg.2020.03.002
5. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012;13:86. doi: 10.1186/1471-2105-13-86
6. Assenov Y, Müller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with RnBeads. *Nat Methods.* 2014;11(11):1138-1140. doi: 10.1038/nmeth.3115
7. Müller F, Scherer M, Assenov Y, Lutsik P, Walter J, Lengauer T, Bock C. RnBeads 2.0: comprehensive analysis of DNA methylation data. *Genome Biol.* 2019;20(1):55. doi: 10.1186/s13059-019-1664-9
8. Leek JT, Storey JD. Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis. *PLoS Genet.* 2007;3(9):1724-1735. doi: 10.1371/journal.pgen.0030161

9. Allum F, Shao X, Guénard F, Simon MM, Busche S, Caron M, Lambourne J, Lessard J, Tandre K, Hedman ÅK, et al. Characterization of functional methylomes by next-generation capture sequencing identifies novel disease-associated variants. *Nat Commun.* 2015;6:7211. doi: 10.1038/ncomms8211
10. Cheung WA, Shao X, Morin A, Siroux V, Kwan T, Ge B, Aïssi D, Chen L, Vasquez L, Allum F, et al. Functional variation in allelic methylomes underscores a strong genetic contribution and reveals novel epigenetic alterations in the human epigenome. *Genome Biol.* 2017;18(1):50. doi: 10.1186/s13059-017-1173-7
11. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114-2120. doi: 10.1093/bioinformatics/btu170
12. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009;25(9):1105-1111. doi: 10.1093/bioinformatics/btp120
13. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9:357-359. doi: 10.1038/nmeth.1923

Supplementary Tables

Table 9. Phenotypic characteristics of individuals included in the analyses performed on naïve CD4⁺ T cells and eosinophils

	Naïve CD4 ⁺ T cells		Eosinophils	
	Expression (n=159)	Methylation (n=192)	Expression (n=139)	Methylation (n=183)
M:F ratio	1.00:1.06	1.00:1.04	1.00:1.01	1.00:1.01
Age, mean (range)	49 (18-81)	49 (18-81)	46 (18-81)	48 (18-81)
Age, median	52	53	47	50
IgE, µg/L, mean (SD) ¹	135 (5)	131 (5)	161 (5)	143 (5)
Atopy, n (%) ²	113 (71)	139 (72)	107 (77)	140 (77)
Indoor allergy, n (%)	98 (62)	120 (63)	94 (68)	121 (66)
Outdoor allergy, n (%)	85 (53)	106 (55)	80 (58)	106 (58)
Both types of allergies, n (%)	73 (46)	90 (47)	70 (50)	90 (49)
Atopic dermatitis, n (%)	69 (43)	87 (45)	70 (50)	89 (49)
Asthma, n (%) ³	98 (62)	122 (64)	91 (65)	121 (66)
Rhinitis, n (%) ⁴	71 (45)	91 (47)	64 (46)	86 (47)
Rhinitis+asthma, n (%)	51 (32)	66 (34)	47 (34)	65 (36)
Smoking status				
Ever smoked, n (%) ⁵	78 (49)	99 (52)	66 (47)	93 (51)
Never smoked, n (%)	81 (51)	93 (48)	73 (53)	90 (49)

All statistical comparisons between the data for the expression in the naïve CD4⁺ T cells and in the eosinophils and between the data for the methylation in the naïve CD4⁺ T cells and in the eosinophils showed that there was no significant difference between them. All statistical comparisons between the data for the expression and the methylation in the naïve CD4⁺ T cells and the eosinophils showed

that there was no significant difference between them. ¹Geometric mean and standard deviation (SD) of blood immunoglobulin E (IgE) concentration. ²Defined as at least one positive response on skin prick testing (wheal diameter ≥ 3 mm more than the negative control at 10 min). ³Present or past documented clinical history of asthma. ⁴Present or past personal history of rhinitis. ⁵The smoking status of ever smoked regroups every individual that is smoking or has stopped smoking.

Table 10. Complete characterization of allergens sensitization of individuals included in the association analyses for the indoor, the outdoor and both types of aeroallergens sensitization

	Indoor		Outdoor		Both types		P value
	Non-allergic (n= 446)	Allergic (n= 564)	Non-allergic (n= 565)	Allergic (n= 445)	Non-allergic (n= 640)	Allergic (n= 370)	
Animal allergy, n (%)	0 (0)	431 (76) ^a	125 (22)	306 (69) ^b	125 (20)	306 (83) ^a	2.02E-05
Cat dander, n (%)	0 (0)	397 (70) ^a	109 (19)	288 (65) ^a	109 (17)	288 (78) ^b	2.31E-04
Dog dander, n (%)	0 (0)	192 (34)	43 (8)	149 (33)	43 (7)	149 (40)	8.30E-02
Horse dander, n (%)	0 (0)	136 (24)	23 (4)	113 (26)	23 (4)	113 (31)	7.70E-02
Cattle dander, n (%)	0 (0)	28 (16)	2 (2)	26 (19)	2 (2)	26 (21)	6.02E-01
Bird feather, n (%)	0 (0)	51 (9)	9 (2)	42 (10)	9 (1)	42 (12)	4.63E-01
House dust allergy, n (%)	0 (0)	240 (43) ^a	49 (9)	191 (43) ^a	49 (8)	191 (52) ^b	1.40E-02
House dust mite allergy, n (%)	0 (0)	389 (69) ^{a, b}	112 (20)	277 (62) ^a	112 (18)	277 (75) ^b	5.45E-04
Dermatophagoides farinae, n (%)	0 (0)	346 (61) ^{a, b}	95 (17)	251 (56) ^a	95 (15)	251 (68) ^b	3.77E-03
Dermatophagoides pteronyssinus, n (%)	0 (0)	332 (59) ^a	84 (15)	248 (56) ^a	84 (13)	248 (67) ^b	3.52E-03
Grass pollen allergy, n (%)	37 (8)	286 (51) ^a	0 (0)	323 (73) ^b	37 (6)	286 (78) ^b	1.26E-19
Grasses, n (%)	25 (6)	246 (44) ^a	0 (0)	271 (61) ^b	25 (4)	246 (67) ^b	7.28E-13
Weeds, n (%)	6 (1)	137 (25) ^a	0 (0)	143 (32) ^b	6 (1)	137 (37) ^b	9.48E-05
Ragweed, n (%)	12 (3)	114 (20) ^a	0 (0)	126 (28) ^b	12 (2)	114 (31) ^b	4.03E-04
Timothy, n (%)	15 (3)	212 (38) ^a	0 (0)	227 (52) ^b	15 (2)	212 (59) ^b	1.59E-09
Rye grass, n (%)	11 (2)	178 (32) ^a	0 (0)	189 (43) ^b	11 (2)	178 (49) ^b	7.40E-07
Tree pollen allergy, n (%)	43 (10)	282 (50) ^a	0 (0)	325 (73) ^b	43 (7)	282 (76) ^b	7.46E-20
Tree mix, n (%)	22 (5)	159 (29) ^a	0 (0)	181 (41) ^b	22 (3)	159 (44) ^b	1.02E-06

Birch, n (%)	32 (7)	244 (43) ^a	0 (0)	276 (62) ^b	32 (5)	244 (66) ^b	2.32E-13
Maple, n (%)	7 (2)	84 (15) ^a	0 (0)	91 (21) ^{a, b}	7 (1)	84 (23) ^b	6.02E-03
Oak, n (%)	5 (1)	90 (16) ^a	0 (0)	95 (22) ^{a, b}	5 (1)	90 (25) ^b	4.16E-03
Elm, n (%)	6 (1)	65 (12) ^a	0 (0)	71 (16) ^{a, b}	6 (1)	65 (18) ^b	2.20E-02
Mold allergy, n (%)	14 (3)	113 (20) ^a	0 (0)	127 (29) ^b	14 (2)	113 (31) ^b	3.35E-04
Cladosporium, n (%)	3 (1)	55 (10) ^a	0 (0)	58 (13) ^a	3 (0)	55 (15) ^a	5.10E-02
Alternaria, n (%)	10 (2)	53 (9) ^a	0 (0)	63 (14) ^a	10 (2)	53 (14) ^a	2.60E-02
Aspergillus fumigatus, n (%)	2 (0)	48 (9)	0 (0)	50 (11)	2 (0)	48 (13)	8.20E-02
Penicillium, n (%)	1 (0)	48 (11)	0 (0)	49 (14)	1 (0)	48 (17)	1.15E-01
Food allergy, n (%)	15 (4)	100 (21)	17 (4)	98 (26)	26 (5)	89 (29)	6.60E-02
Allergic monosensitization, n (%)	38 (9)	66 (12) ^a	72 (13)	32 (7) ^b	104 (16)	0 (0) ^c	8.60E-11
Allergic polysensitization, n (%)	43 (10)	498 (88) ^a	128 (23)	413 (93) ^b	171 (27)	370 (100) ^c	8.60E-11
Number of sensitizations	1-10	1-28	1-7	1-28	1-10	1-28	-
One category of allergens, n (%)	57 (13)	110 (20) ^a	116 (21)	51 (11) ^b	167 (26)	0 (0) ^c	1.25E-18
Two categories of allergens, n (%)	20 (4)	111 (20) ^a	66 (12)	65 (15) ^{a, b}	86 (13)	45 (12) ^b	5.60E-03
Three categories of allergens, n (%)	4 (1)	100 (18)	17 (3)	87 (20)	21 (3)	83 (22)	2.08E-01
Four categories of allergens, n (%)	0 (0)	97 (17) ^a	1 (1)	96 (22) ^{a, b}	1 (0.16)	96 (26) ^b	5.32E-03
Five categories of allergens, n (%)	0 (0)	79 (14) ^a	0 (0)	79 (18) ^{a, b}	0 (0)	79 (21) ^b	1.35E-02
Six categories of allergens, n (%)	0 (0)	43 (8)	0 (0)	43 (10)	0 (0)	43 (12)	1.17E-01
Seven categories of allergens, n (%)	0 (0)	24 (4)	0 (0)	24 (5)	0 (0)	24 (6)	3.19E-01

All statistical comparisons between non-allergic and allergic individuals were significant for the exception of the allergic monosensitization for the indoor phenotype ($P=9.85E-02$) and the allergic sensitization to two categories of allergens for the outdoor phenotype ($P=1.70E-01$) and for both phenotypes ($P=5.61E-01$). The significant difference between each phenotype is noted with the letters a, b and c. If phenotypes have different letters, it means there is a statistical difference between them. If groups share the same letter or none of the groups have a letter, it means that there is no statistical difference between them.

Table 11. All significant genetic variants associated with the sensitization to indoor aeroallergens in the SLSJ cohort ($P < 1E-05$) and association results for the EGEA cohort

Variant ID ^a	Chr	Position ^b	Minor allele	MAF	Stat	P _{SLSJ}	P _{EGEA} ^c	Locus	Nearest gene ^d	Variant information ^e
rs75285112	1	48,410,577	T	2.15E-02	20.07	7.45E-06	3.03E-01	1p33	<i>TRABD2B</i>	Intron
rs356593	5	63,340,373	G	3.80E-01	22.75	1.84E-06	N/A	5q12.3	–	Intergenic variant
rs36160782	5	63,365,225	T	3.87E-01	19.73	8.94E-06	8.86E-01	5q12.3	–	Intergenic variant
rs36173573	5	63,365,227	TA	3.87E-01	19.73	8.94E-06	N/A	5q12.3	–	Intergenic variant
rs356605	5	63,366,254	T	3.87E-01	19.73	8.94E-06	8.86E-01	5q12.3	–	Intergenic variant
rs369133	5	63,369,324	A	3.87E-01	19.63	9.39E-06	8.69E-01	5q12.3	–	Intergenic variant
rs441080	5	63,369,409	A	3.87E-01	19.73	8.94E-06	8.87E-01	5q12.3	–	Intergenic variant
rs416463	5	63,369,878	C	3.87E-01	19.73	8.94E-06	8.88E-01	5q12.3	–	Intergenic variant
rs356635	5	63,375,421	T	3.65E-01	24.20	8.67E-07	8.17E-01	5q12.3	–	Intergenic variant
rs4562001	5	63,376,657	A	3.65E-01	24.20	8.67E-07	8.78E-01	5q12.3	–	Intergenic variant
rs10051553	5	63,377,895	T	3.65E-01	24.20	8.67E-07	8.77E-01	5q12.3	–	Intergenic variant
rs1587067	5	63,380,385	T	3.65E-01	24.20	8.67E-07	8.77E-01	5q12.3	–	Intergenic variant
rs11286839	5	63,382,997	G	3.78E-01	20.80	5.10E-06	N/A	5q12.3	–	Intergenic variant
rs6449700	5	63,383,370	G	3.65E-01	24.20	8.67E-07	9.59E-01	5q12.3	–	Intergenic variant
rs356632	5	63,383,651	G	3.65E-01	24.20	8.67E-07	9.13E-01	5q12.3	–	Intergenic variant
rs356631	5	63,383,849	G	3.65E-01	24.20	8.67E-07	9.13E-01	5q12.3	–	Intergenic variant
rs356629	5	63,384,513	A	3.65E-01	24.20	8.67E-07	9.12E-01	5q12.3	–	Intergenic variant
rs356625	5	63,385,467	C	3.65E-01	24.20	8.67E-07	9.13E-01	5q12.3	–	Intergenic variant
rs356624	5	63,386,246	G	3.65E-01	24.20	8.67E-07	9.13E-01	5q12.3	–	Intergenic variant
rs356622	5	63,386,908	C	3.66E-01	24.25	8.48E-07	9.14E-01	5q12.3	–	Intergenic variant
rs356621	5	63,387,184	T	3.66E-01	24.19	8.71E-07	9.12E-01	5q12.3	–	Intergenic variant
rs356617	5	63,388,546	G	3.65E-01	24.20	8.67E-07	8.76E-01	5q12.3	–	Intergenic variant

rs168558	5	63,390,398	C	3.63E-01	23.87	1.03E-06	9.13E-01	5q12.3	–	Intergenic variant
rs34792252	5	63,392,364	A	3.63E-01	23.87	1.03E-06	N/A	5q12.3	–	Intergenic variant
rs1494629	5	63,394,447	A	3.63E-01	23.87	1.03E-06	9.66E-01	5q12.3	–	Intergenic variant
rs434513	5	63,394,831	T	3.63E-01	23.87	1.03E-06	9.57E-01	5q12.3	–	Intergenic variant
rs453931	5	63,396,156	A	3.63E-01	23.87	1.03E-06	9.57E-01	5q12.3	–	Intergenic variant
rs200072450	5	63,397,194	A	3.63E-01	20.20	6.98E-06	9.57E-01	5q12.3	–	Intergenic variant
rs10042853	5	63,397,729	A	3.63E-01	20.20	6.98E-06	9.60E-01	5q12.3	–	Intergenic variant
rs77685870	5	63,398,014	G	3.63E-01	20.20	6.98E-06	9.57E-01	5q12.3	–	Intergenic variant
rs10073616	5	63,398,939	G	3.63E-01	20.20	6.98E-06	9.81E-01	5q12.3	–	Intergenic variant
rs7725793	5	63,399,494	T	3.63E-01	20.20	6.98E-06	9.60E-01	5q12.3	–	Intergenic variant
rs112734978	5	63,401,001	T	3.61E-01	20.58	5.71E-06	N/A	5q12.3	–	Intergenic variant
rs17796121	5	63,401,236	G	3.63E-01	20.20	6.98E-06	9.60E-01	5q12.3	–	Intergenic variant
rs392360	5	63,402,195	G	3.63E-01	20.20	6.98E-06	9.22E-01	5q12.3	–	Intergenic variant
rs9763077	5	63,402,805	A	3.63E-01	20.20	6.98E-06	9.60E-01	5q12.3	–	Intergenic variant
rs421980	5	63,404,744	T	3.63E-01	20.20	6.98E-06	9.57E-01	5q12.3	–	Intergenic variant
rs356555	5	63,406,866	A	3.63E-01	20.20	6.98E-06	9.22E-01	5q12.3	–	Intergenic variant
rs7718578	5	63,410,697	G	3.63E-01	20.20	6.98E-06	9.57E-01	5q12.3	–	Intergenic variant
rs356569	5	63,412,673	G	3.63E-01	20.20	6.98E-06	9.57E-01	5q12.3	–	Intergenic variant
rs356567	5	63,414,400	C	3.61E-01	20.01	7.72E-06	8.78E-01	5q12.3	–	Intergenic variant
rs356564	5	63,416,788	T	3.61E-01	20.47	6.04E-06	8.83E-01	5q12.3	–	Intergenic variant
rs62369051	5	63,419,229	T	3.62E-01	20.14	7.19E-06	9.21E-01	5q12.3	–	Intergenic variant
rs445966	5	63,423,487	C	3.62E-01	20.14	7.19E-06	9.18E-01	5q12.3	–	Intergenic variant
rs6888798	5	63,425,122	A	3.62E-01	20.14	7.19E-06	N/A	5q12.3	–	Intergenic variant
rs1082153	5	63,426,420	T	3.62E-01	20.14	7.19E-06	9.19E-01	5q12.3	–	Intergenic variant
rs185851237	5	63,427,475	G	3.62E-01	20.14	7.19E-06	9.21E-01	5q12.3	–	Intergenic variant
rs356589	5	63,431,701	A	3.62E-01	20.14	7.19E-06	8.83E-01	5q12.3	–	Intergenic variant

rs356590	5	63,433,593	G	3.71E-01	23.67	1.14E-06	N/A	5q12.3	–	Intergenic variant
rs55723551	5	63,438,060	A	3.69E-01	23.95	9.90E-07	9.21E-01	5q12.3	–	Intergenic variant
rs356592	5	63,438,751	T	3.69E-01	23.95	9.90E-07	9.19E-01	5q12.3	–	Intergenic variant
rs431517	5	63,440,291	T	3.71E-01	23.67	1.14E-06	9.18E-01	5q12.3	–	Intergenic variant
rs356578	5	63,441,800	T	3.71E-01	23.67	1.14E-06	9.18E-01	5q12.3	–	Intergenic variant
rs5868370	5	63,443,062	C	3.77E-01	25.68	4.03E-07	N/A	5q12.3	–	Intergenic variant
rs181523	5	63,443,188	G	3.69E-01	24.88	6.10E-07	9.38E-01	5q12.3	–	Intergenic variant
rs356576	5	63,444,044	C	3.66E-01	25.49	4.46E-07	9.38E-01	5q12.3	–	Intergenic variant
rs356575	5	63,444,052	G	3.66E-01	25.49	4.46E-07	9.38E-01	5q12.3	–	Intergenic variant
rs356574	5	63,444,260	C	3.69E-01	24.88	6.10E-07	9.38E-01	5q12.3	–	Intergenic variant
rs356573	5	63,445,065	G	3.67E-01	23.75	1.09E-06	9.17E-01	5q12.3	–	Intergenic variant
rs356570	5	63,447,789	G	3.69E-01	24.88	6.10E-07	9.25E-01	5q12.3	–	Intergenic variant
rs356560	5	63,448,488	T	3.81E-01	23.86	1.04E-06	8.19E-01	5q12.3	–	Intergenic variant
rs356559	5	63,448,661	A	3.69E-01	24.88	6.10E-07	8.90E-01	5q12.3	–	Intergenic variant
rs356558	5	63,448,916	C	3.69E-01	24.88	6.10E-07	9.25E-01	5q12.3	–	Intergenic variant
rs356557	5	63,448,979	C	3.69E-01	24.88	6.10E-07	9.25E-01	5q12.3	–	Intergenic variant
rs35249983	5	63,449,036	C	3.69E-01	24.88	6.10E-07	N/A	5q12.3	–	Intergenic variant
rs356640	5	63,454,211	G	3.69E-01	24.88	6.10E-07	9.32E-01	5q12.3	<i>RNF180</i>	7.5kb upstream variant
rs356639	5	63,454,322	T	3.69E-01	24.88	6.10E-07	9.25E-01	5q12.3	<i>RNF180</i>	7.3kb upstream variant
rs356637	5	63,455,392	C	3.69E-01	24.88	6.10E-07	9.25E-01	5q12.3	<i>RNF180</i>	6.3kb upstream variant
rs6449702	5	63,466,256	A	3.69E-01	24.88	6.10E-07	9.55E-01	5q12.3	<i>RNF180</i>	Intron variant
rs3104058	5	63,469,539	G	3.68E-01	24.43	7.72E-07	9.10E-01	5q12.3	<i>RNF180</i>	Intron variant
rs1603090	5	63,471,824	G	3.69E-01	24.88	6.10E-07	9.38E-01	5q12.3	<i>RNF180</i>	Intron variant
rs951960	5	63,473,068	T	3.69E-01	24.88	6.10E-07	9.38E-01	5q12.3	<i>RNF180</i>	Intron variant

rs6898951	5	63,477,547	A	3.70E-01	23.20	1.46E-06	9.75E-01	5q12.3	<i>RNF180</i>	Intron variant
rs17796598	5	63,482,115	T	3.69E-01	24.88	6.10E-07	9.41E-01	5q12.3	<i>RNF180</i>	Intron variant
rs2055135	5	63,486,821	T	3.69E-01	24.88	6.10E-07	9.76E-01	5q12.3	<i>RNF180</i>	Intron variant
rs2934804	5	63,489,273	A	3.69E-01	24.88	6.10E-07	9.77E-01	5q12.3	<i>RNF180</i>	Intron variant
rs2962997	5	63,491,100	C	3.69E-01	24.88	6.10E-07	9.77E-01	5q12.3	<i>RNF180</i>	Intron variant
rs1389864	5	63,497,180	G	3.69E-01	24.88	6.10E-07	9.05E-01	5q12.3	<i>RNF180</i>	Intron variant
rs2962998	5	63,497,836	A	3.69E-01	24.88	6.10E-07	9.47E-01	5q12.3	<i>RNF180</i>	Intron variant
rs2366281	5	63,501,533	C	3.69E-01	24.88	6.10E-07	9.51E-01	5q12.3	<i>RNF180</i>	Intron variant
rs1994931	5	63,503,249	C	3.69E-01	24.88	6.10E-07	9.51E-01	5q12.3	<i>RNF180</i>	Intron variant
rs2968294	5	63,512,436	G	3.69E-01	24.88	6.10E-07	9.48E-01	5q12.3	<i>RNF180</i>	Intron variant
rs2963002	5	63,513,099	G	3.69E-01	24.88	6.10E-07	9.44E-01	5q12.3	<i>RNF180</i>	Intron variant
rs11311375	5	63,515,051	G	3.69E-01	24.88	6.10E-07	N/A	5q12.3	<i>RNF180</i>	Intron variant
rs7729392	5	63,518,268	G	3.69E-01	24.88	6.10E-07	9.42E-01	5q12.3	<i>RNF180</i>	Intron variant
rs10676352	5	63,519,384	CAGA G	3.69E-01	24.88	6.10E-07	N/A	5q12.3	<i>RNF180</i>	Intron variant
rs11743978	5	63,519,662	G	3.69E-01	24.88	6.10E-07	9.42E-01	5q12.3	<i>RNF180</i>	Intron variant
rs10805384	5	63,521,079	G	3.69E-01	24.88	6.10E-07	9.42E-01	5q12.3	<i>RNF180</i>	Intron variant
rs7712048	5	63,534,780	T	3.67E-01	23.00	1.62E-06	9.34E-01	5q12.3	<i>RNF180</i>	Intron variant
rs13360622	5	63,543,066	C	3.65E-01	23.58	1.20E-06	9.08E-01	5q12.3	<i>RNF180</i>	Intron variant
rs10041491	5	63,544,087	A	3.68E-01	23.61	1.18E-06	9.08E-01	5q12.3	<i>RNF180</i>	Intron variant
rs6892297	5	63,546,871	A	3.68E-01	23.61	1.18E-06	9.46E-01	5q12.3	<i>RNF180</i>	Intron variant
rs7720768	5	63,566,130	C	3.68E-01	24.18	8.75E-07	9.47E-01	5q12.3	<i>RNF180</i>	Intron variant
rs1947094	5	63,572,677	A	3.68E-01	24.18	8.75E-07	9.09E-01	5q12.3	<i>RNF180</i>	Intron variant
rs1155286	5	63,574,428	G	3.67E-01	23.89	1.02E-06	9.02E-01	5q12.3	<i>RNF180</i>	Intron variant
rs1813341	5	63,577,504	T	3.68E-01	24.18	8.75E-07	9.49E-01	5q12.3	<i>RNF180</i>	Intron variant
rs10076464	5	63,579,416	C	3.66E-01	23.35	1.35E-06	9.57E-01	5q12.3	<i>RNF180</i>	Intron variant
rs57614366	5	63,586,009	G	3.68E-01	24.18	8.75E-07	9.47E-01	5q12.3	<i>RNF180</i>	Intron variant

rs10063060	5	63,592,613	G	3.67E-01	24.03	9.49E-07	N/A	5q12.3	<i>RNF180</i>	Intron variant
rs6449707	5	63,598,077	G	3.68E-01	24.18	8.75E-07	9.16E-01	5q12.3	<i>RNF180</i>	Intron variant
rs1112246	5	63,600,018	A	3.65E-01	22.34	2.28E-06	9.40E-01	5q12.3	<i>RNF180</i>	Intron variant
rs11958206	5	63,600,233	T	3.66E-01	22.34	2.28E-06	N/A	5q12.3	<i>RNF180</i>	Intron variant
rs1445959	5	63,600,853	T	3.66E-01	22.34	2.28E-06	9.12E-01	5q12.3	<i>RNF180</i>	Intron variant
rs10039444	5	63,606,450	T	3.65E-01	21.56	3.43E-06	9.40E-01	5q12.3	<i>RNF180</i>	Intron variant
rs7719273	5	63,611,491	A	3.67E-01	23.37	1.33E-06	9.13E-01	5q12.3	<i>RNF180</i>	Intron variant
rs12522604	5	63,613,639	A	3.67E-01	23.37	1.33E-06	9.74E-01	5q12.3	<i>RNF180</i>	Intron variant
rs6449710	5	63,613,720	A	3.67E-01	23.37	1.33E-06	9.29E-01	5q12.3	<i>RNF180</i>	Intron variant
rs6449711	5	63,616,077	T	3.77E-01	21.29	3.94E-06	9.74E-01	5q12.3	<i>RNF180</i>	Intron variant
rs6449715	5	63,622,181	C	3.67E-01	23.37	1.33E-06	9.74E-01	5q12.3	<i>RNF180</i>	Intron variant
rs6870974	5	63,622,420	C	3.67E-01	23.37	1.33E-06	9.74E-01	5q12.3	<i>RNF180</i>	Intron variant
rs1348715	5	63,625,119	T	3.74E-01	23.13	1.51E-06	9.30E-01	5q12.3	<i>RNF180</i>	Intron variant
rs12518161	5	63,628,359	C	3.67E-01	23.50	1.25E-06	9.74E-01	5q12.3	<i>RNF180</i>	Intron variant
rs10524410	5	63,640,358	AAAA G	3.67E-01	23.50	1.25E-06	N/A	5q12.3	<i>RNF180</i>	Intron variant
rs6449718	5	63,640,649	A	3.67E-01	23.50	1.25E-06	9.85E-01	5q12.3	<i>RNF180</i>	Intron variant
rs35590919	5	63,641,815	ATCT T	3.67E-01	23.50	1.25E-06	N/A	5q12.3	<i>RNF180</i>	Intron variant
rs6449719	5	63,646,283	C	3.67E-01	23.50	1.25E-06	9.98E-01	5q12.3	<i>RNF180</i>	Intron variant
rs1526875	5	63,656,845	A	3.67E-01	23.50	1.25E-06	9.34E-01	5q12.3	<i>RNF180</i>	Intron variant
rs1526874	5	63,657,109	G	3.67E-01	23.50	1.25E-06	9.85E-01	5q12.3	<i>RNF180</i>	Intron variant
rs7732201	5	63,659,492	A	3.68E-01	22.16	2.51E-06	9.33E-01	5q12.3	<i>RNF180</i>	Intron variant
rs6870956	5	63,663,031	A	3.67E-01	23.50	1.25E-06	9.95E-01	5q12.3	<i>RNF180</i>	Intron variant
rs10056417	5	63,663,330	T	3.66E-01	22.44	2.17E-06	9.93E-01	5q12.3	<i>RNF180</i>	Intron variant
rs56066106	5	63,665,294	T	3.88E-01	24.41	7.79E-07	N/A	5q12.3	<i>RNF180</i>	Intron variant
rs28615727	5	63,665,335	A	3.67E-01	23.50	1.25E-06	9.95E-01	5q12.3	<i>RNF180</i>	Intron variant

											6.8kb downstream variant
rs1348686	5	63,675,479	C	3.72E-01	25.08	5.51E-07	7.87E-01	5q12.3	<i>RNF180</i>		
rs10805385	5	63,679,947	A	3.73E-01	23.98	9.72E-07	7.94E-01	5q12.3	–		Intergenic variant
rs1445759	5	63,682,604	G	3.71E-01	25.50	4.42E-07	7.94E-01	5q12.3	–		Intergenic variant
rs6449723	5	63,684,731	C	3.73E-01	23.98	9.72E-07	7.94E-01	5q12.3	–		Intergenic variant
rs7735885	5	63,686,428	C	3.73E-01	23.98	9.72E-07	7.94E-01	5q12.3	–		Intergenic variant
rs6449724	5	63,688,248	G	3.73E-01	22.88	1.72E-06	7.93E-01	5q12.3	–		Intergenic variant
rs7715314	5	63,690,244	G	3.73E-01	23.98	9.72E-07	7.93E-01	5q12.3	–		Intergenic variant
rs6864510	5	63,691,564	A	3.69E-01	23.61	1.18E-06	8.52E-01	5q12.3	–		Intergenic variant
rs6897021	5	63,695,563	T	3.73E-01	23.98	9.72E-07	7.87E-01	5q12.3	–		Intergenic variant
rs10471613	5	63,696,064	A	3.75E-01	24.17	8.80E-07	8.69E-01	5q12.3	–		Intergenic variant
rs1809291	5	63,697,753	C	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–		Intergenic variant
rs898519	5	63,698,034	G	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–		Intergenic variant
rs12517575	5	63,701,870	A	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–		Intergenic variant
rs28690156	5	63,702,076	C	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–		Intergenic variant
rs10706640	5	63,702,474	G	3.73E-01	23.98	9.72E-07	N/A	5q12.3	–		Intergenic variant
rs6449726	5	63,702,539	A	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–		Intergenic variant
rs6449727	5	63,702,613	T	3.73E-01	23.98	9.72E-07	6.42E-01	5q12.3	–		Intergenic variant
rs6449728	5	63,702,685	G	3.72E-01	23.58	1.20E-06	6.60E-01	5q12.3	–		Intergenic variant
rs6449729	5	63,702,811	A	3.64E-01	24.25	8.44E-07	6.60E-01	5q12.3	–		Intergenic variant
rs77908216	5	63,702,853	A	3.73E-01	23.98	9.72E-07	6.55E-01	5q12.3	–		Intergenic variant
rs78665401	5	63,702,854	A	3.73E-01	23.98	9.72E-07	6.55E-01	5q12.3	–		Intergenic variant
rs77334418	5	63,702,855	G	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–		Intergenic variant
rs11744919	5	63,703,382	C	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–		Intergenic variant
rs6449730	5	63,704,583	G	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–		Intergenic variant
rs6868317	5	63,704,912	T	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–		Intergenic variant

rs62369400	5	63,704,942	C	3.73E-01	23.98	9.72E-07	6.39E-01	5q12.3	–	Intergenic variant
rs62369401	5	63,704,988	C	3.73E-01	23.98	9.72E-07	6.43E-01	5q12.3	–	Intergenic variant
rs34211306	5	63,706,005	CA	3.73E-01	23.98	9.72E-07	N/A	5q12.3	–	Intergenic variant
rs59667389	5	63,706,281	G	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–	Intergenic variant
rs62369403	5	63,706,455	C	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–	Intergenic variant
rs11959389	5	63,706,888	G	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–	Intergenic variant
rs7734057	5	63,707,632	T	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–	Intergenic variant
rs7733937	5	63,707,641	T	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–	Intergenic variant
rs6449731	5	63,707,723	A	3.71E-01	23.98	9.73E-07	6.60E-01	5q12.3	–	Intergenic variant
rs6449732	5	63,707,849	T	3.79E-01	22.99	1.63E-06	6.60E-01	5q12.3	–	Intergenic variant
rs6449733	5	63,707,864	A	3.79E-01	22.99	1.63E-06	6.60E-01	5q12.3	–	Intergenic variant
rs6896012	5	63,708,132	T	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–	Intergenic variant
rs6896434	5	63,708,203	C	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–	Intergenic variant
rs1037843	5	63,708,707	G	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–	Intergenic variant
rs11954734	5	63,708,791	T	3.67E-01	22.13	2.55E-06	6.60E-01	5q12.3	–	Intergenic variant
rs62369417	5	63,708,830	T	3.68E-01	21.97	2.77E-06	6.61E-01	5q12.3	–	Intergenic variant
rs4288085	5	63,708,856	A	3.68E-01	21.97	2.77E-06	6.61E-01	5q12.3	–	Intergenic variant
rs4631133	5	63,708,993	A	3.73E-01	23.98	9.72E-07	6.61E-01	5q12.3	–	Intergenic variant
rs11955741	5	63,710,070	T	3.73E-01	23.98	9.72E-07	6.60E-01	5q12.3	–	Intergenic variant
rs11949865	5	63,710,339	C	3.73E-01	23.98	9.72E-07	6.53E-01	5q12.3	–	Intergenic variant
rs6878245	5	63,712,022	T	3.70E-01	22.25	2.39E-06	6.61E-01	5q12.3	–	Intergenic variant
rs10059274	5	63,712,604	C	3.68E-01	21.20	4.13E-06	6.61E-01	5q12.3	–	Intergenic variant
rs10074447	5	63,712,734	T	3.69E-01	20.05	7.55E-06	6.61E-01	5q12.3	–	Intergenic variant
rs10074537	5	63,713,007	A	3.69E-01	20.05	7.55E-06	6.61E-01	5q12.3	–	Intergenic variant
rs4084594	5	63,714,445	C	3.51E-01	21.39	3.76E-06	4.57E-01	5q12.3	–	Intergenic variant
rs6875989	5	63,714,773	A	3.51E-01	21.39	3.76E-06	4.57E-01	5q12.3	–	Intergenic variant

rs1584164	5	63,715,916	A	3.50E-01	22.11	2.57E-06	4.78E-01	5q12.3	–	Intergenic variant
rs7711925	5	63,716,244	T	3.51E-01	21.39	3.76E-06	4.57E-01	5q12.3	–	Intergenic variant
rs7715439	5	63,716,369	G	3.52E-01	21.07	4.42E-06	4.53E-01	5q12.3	–	Intergenic variant
rs7715449	5	63,716,398	G	3.53E-01	20.96	4.69E-06	4.19E-01	5q12.3	–	Intergenic variant
rs7735729	5	63,716,429	G	3.52E-01	21.58	3.39E-06	4.18E-01	5q12.3	–	Intergenic variant
rs3919739	5	63,716,468	G	3.51E-01	21.27	3.99E-06	4.23E-01	5q12.3	–	Intergenic variant
rs1823021	5	63,716,851	A	3.51E-01	21.27	3.99E-06	4.23E-01	5q12.3	–	Intergenic variant
rs28877318	5	63,716,943	T	3.51E-01	21.27	3.99E-06	4.54E-01	5q12.3	–	Intergenic variant
rs1900147	5	63,717,733	A	3.51E-01	21.39	3.75E-06	4.57E-01	5q12.3	–	Intergenic variant
rs10075731	5	63,718,378	T	3.51E-01	22.02	2.70E-06	N/A	5q12.3	–	Intergenic variant
rs10075740	5	63,718,438	A	3.51E-01	22.02	2.70E-06	4.57E-01	5q12.3	–	Intergenic variant
rs10075915	5	63,718,674	A	3.51E-01	22.02	2.70E-06	4.57E-01	5q12.3	–	Intergenic variant
rs60252979	5	63,718,911	T	3.51E-01	22.02	2.70E-06	4.57E-01	5q12.3	–	Intergenic variant
rs10039637	5	63,719,296	C	3.51E-01	22.02	2.70E-06	4.58E-01	5q12.3	–	Intergenic variant
rs7736758	5	63,720,345	C	3.51E-01	22.02	2.70E-06	4.58E-01	5q12.3	–	Intergenic variant
rs7737492	5	63,720,551	C	3.51E-01	22.02	2.70E-06	N/A	5q12.3	–	Intergenic variant
rs7704895	5	63,721,607	T	3.51E-01	22.02	2.70E-06	4.57E-01	5q12.3	–	Intergenic variant
rs10039406	5	63,722,306	T	3.51E-01	22.02	2.70E-06	5.25E-01	5q12.3	–	Intergenic variant
rs10036568	5	63,722,514	G	3.51E-01	22.02	2.70E-06	5.25E-01	5q12.3	–	Intergenic variant
rs10036642	5	63,722,814	G	3.51E-01	22.02	2.70E-06	5.24E-01	5q12.3	–	Intergenic variant
rs7710819	5	63,722,983	A	3.51E-01	22.02	2.70E-06	5.24E-01	5q12.3	–	Intergenic variant
rs6875564	5	63,723,673	C	3.51E-01	22.02	2.70E-06	5.24E-01	5q12.3	–	Intergenic variant
rs35041124	5	63,723,963	C	3.51E-01	21.60	3.36E-06	N/A	5q12.3	–	Intergenic variant
rs57798228	5	63,725,456	G	3.51E-01	22.02	2.70E-06	N/A	5q12.3	–	Intergenic variant
rs7381057	5	63,725,572	G	3.51E-01	22.02	2.70E-06	5.24E-01	5q12.3	–	Intergenic variant
rs70983616	5	63,725,906	G	3.51E-01	22.02	2.70E-06	N/A	5q12.3	–	Intergenic variant

rs7378783	5	63,726,874	G	3.50E-01	21.57	3.42E-06	5.24E-01	5q12.3	–	Intergenic variant
rs7378757	5	63,726,936	A	3.50E-01	21.57	3.42E-06	5.24E-01	5q12.3	–	Intergenic variant
rs13160954	5	63,727,271	G	3.51E-01	22.02	2.70E-06	5.24E-01	5q12.3	–	Intergenic variant
rs75526518	5	63,727,514	A	3.51E-01	20.21	6.93E-06	N/A	5q12.3	–	Intergenic variant
rs12514524	5	63,727,576	C	3.51E-01	21.28	3.96E-06	5.24E-01	5q12.3	–	Intergenic variant
rs12514555	5	63,727,721	C	3.51E-01	21.28	3.96E-06	5.24E-01	5q12.3	–	Intergenic variant
rs59861578	5	63,728,335	C	3.47E-01	21.08	4.39E-06	5.72E-01	5q12.3	–	Intergenic variant
rs1380772	5	63,728,494	A	3.75E-01	21.09	4.39E-06	5.24E-01	5q12.3	–	Intergenic variant
rs1380773	5	63,728,683	G	3.54E-01	21.32	3.89E-06	5.24E-01	5q12.3	–	Intergenic variant
rs1380774	5	63,728,743	G	3.54E-01	21.32	3.89E-06	5.24E-01	5q12.3	–	Intergenic variant
rs4643896	5	63,729,128	C	3.51E-01	21.28	3.96E-06	4.73E-01	5q12.3	–	Intergenic variant
rs4290964	5	63,729,295	C	3.51E-01	21.28	3.96E-06	5.24E-01	5q12.3	–	Intergenic variant
rs12332245	5	63,729,397	C	3.51E-01	21.28	3.96E-06	5.23E-01	5q12.3	–	Intergenic variant
rs4304047	5	63,729,599	A	3.51E-01	21.28	3.96E-06	5.24E-01	5q12.3	–	Intergenic variant
rs4359994	5	63,729,610	A	3.51E-01	21.28	3.96E-06	5.24E-01	5q12.3	–	Intergenic variant
rs4533846	5	63,729,725	T	3.51E-01	21.28	3.96E-06	5.24E-01	5q12.3	–	Intergenic variant
rs12516166	5	63,730,207	G	3.51E-01	21.28	3.96E-06	5.24E-01	5q12.3	–	Intergenic variant
rs12516711	5	63,730,410	C	3.51E-01	21.28	3.96E-06	5.24E-01	5q12.3	–	Intergenic variant
rs59383911	5	63,730,453	AG	3.51E-01	21.28	3.96E-06	N/A	5q12.3	–	Intergenic variant
rs10939990	5	63,731,700	A	3.51E-01	21.28	3.96E-06	N/A	5q12.3	–	Intergenic variant
rs10039575	5	63,732,142	C	3.51E-01	21.28	3.96E-06	5.18E-01	5q12.3	–	Intergenic variant
rs4557363	5	63,732,305	G	3.51E-01	21.28	3.96E-06	4.71E-01	5q12.3	–	Intergenic variant
rs60327423	5	63,732,809	TA	3.51E-01	21.28	3.96E-06	N/A	5q12.3	–	Intergenic variant
rs4432836	5	63,733,748	C	3.51E-01	21.28	3.96E-06	5.18E-01	5q12.3	–	Intergenic variant
rs4327567	5	63,733,973	G	3.51E-01	21.28	3.96E-06	5.18E-01	5q12.3	–	Intergenic variant
rs4470714	5	63,734,171	C	3.51E-01	21.28	3.96E-06	5.18E-01	5q12.3	–	Intergenic variant

rs4594806	5	63,734,870	G	3.53E-01	21.19	4.16E-06	N/A	5q12.3	–	Intergenic variant
rs11955926	5	63,734,900	G	3.53E-01	21.19	4.16E-06	5.18E-01	5q12.3	–	Intergenic variant
rs11955953	5	63,735,005	G	3.53E-01	21.19	4.16E-06	5.18E-01	5q12.3	–	Intergenic variant
rs13162093	5	63,735,222	G	3.52E-01	21.86	2.94E-06	5.18E-01	5q12.3	–	Intergenic variant
rs370819375	5	117,876,341	G	5.53E-02	20.14	7.19E-06	N/A	5q23.1	<i>LINC02208</i>	Intron variant
rs2502625	6	144,591,955	A	3.61E-01	20.85	4.97E-06	6.25E-01	6q24.2	–	Intergenic variant
rs3798307	6	167,363,948	A	2.04E-01	20.72	5.31E-06	6.57E-01	6q27	<i>RNASET2</i>	Intron variant
rs12531771	7	145,165,023	C	1.08E-01	19.72	8.95E-06	8.15E-01	7q35	–	Intergenic variant
rs75632282	7	151,175,033	A	2.09E-02	20.98	4.64E-06	8.81E-01	7q36.1	<i>RHEB</i>	Intron variant
rs79861276	7	151,183,780	A	2.09E-02	20.98	4.64E-06	8.63E-01	7q36.1	<i>RHEB</i>	Intron variant
rs6998445	8	58,651,175	T	4.43E-01	20.89	4.86E-06	1.76E-01	8q12.1	–	Intergenic variant
rs1331456	9	2,451,397	C	3.10E-01	19.52	9.96E-06	2.89E-01	9p24.2	<i>VLDLR-AS1</i>	Intron variant
rs10979982	9	112,413,523	C	9.31E-02	21.71	3.16E-06	2.67E-01	9q31.3	<i>PALM2AKAP2, MIR548 (RF01061-187)</i>	Intron variant, 8.7kb upstream variant
rs10980680	9	113,749,166	C	4.70E-01	29.20	6.54E-08	5.35E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs10980683	9	113,753,524	C	4.66E-01	25.94	3.53E-07	4.96E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs1121570	9	113,753,770	C	4.68E-01	29.29	6.23E-08	5.40E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs1030006	9	113,779,011	G	4.85E-01	23.98	9.73E-07	7.02E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs2110548	9	113,782,934	C	4.85E-01	24.46	7.59E-07	7.25E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs723591	9	113,783,435	T	4.85E-01	24.46	7.59E-07	7.44E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs12345138	9	113,783,796	A	4.85E-01	24.46	7.59E-07	7.73E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs976181	9	113,785,792	G	4.86E-01	25.07	5.53E-07	7.21E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs7039431	9	113,785,983	T	4.86E-01	25.05	5.59E-07	7.17E-01	9q31.3	<i>LPAR1</i>	Intron variant

rs141132389	9	113,786,826	TATA TATA TACA CAC A C	4.88E-01	24.90	6.03E-07	N/A	9q31.3	<i>LPAR1</i>	Intron variant
rs62573262	9	113,787,141	C	4.86E-01	25.07	5.53E-07	7.40E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs10817132	9	113,789,688	A	4.85E-01	23.80	1.07E-06	7.65E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs35663630	9	113,789,836	AT	4.85E-01	24.30	8.23E-07	N/A	9q31.3	<i>LPAR1</i>	Intron variant
rs10980699	9	113,790,030	A	4.85E-01	23.70	1.13E-06	7.86E-01	9q31.3	<i>LPAR1</i>	Intron variant
rs117148242	9	139,624,985	A	6.85E-03	21.34	3.84E-06	3.20E-01	9q34.3	<i>SNHG7, LCN10</i>	2.3kb upstream variant, 7.6kb downstream variant
rs2892077	10	5,653,677	G	2.17E-01	20.68	5.43E-06	1.44E-01	10p15.1	–	Intergenic variant
rs7997863	13	45,369,779	G	2.51E-01	20.29	6.67E-06	3.02E-01	13q14.12	<i>LINC00330</i>	3.9kb downstream variant
rs1599965	14	95,874,407	C	3.04E-01	21.15	4.24E-06	4.17E-02	14q32.13	<i>SYNE3</i>	Exon variant
rs13336442	16	1,539,722	A	7.98E-02	20.23	6.87E-06	6.98E-01	16p13.3	<i>PTX4, TELO2</i>	1.3kb upstream variant, 3.6kb upstream variant
rs79102289	18	7,025,917	T	2.77E-02	23.71	1.12E-06	7.07E-01	18p11.31	<i>LAMA1</i>	Intron variant

^aVariant ID in bold character can also be found in the significant variant associated to both types of aeroallergens sensitization. ^bGenetic variant's position on the genome according to the human genome reference 19 (hg19). ^cIf a variant has a P value of NA in the EGEEA cohort it means that this particular variant wasn't found during the replication analysis. ^dGene in which the variant is located in or nearest to (≤ 10 kb on the right or left sides) according to the UCSC Genome browser (hg19). For variants which are intergenic (further than 10kb from a gene), no information is provided. ^eVariant position toward the gene according to UCSC. Abbreviation: Chr=Chromosome, EGEEA=The French Epidemiological study on the Genetics and Environment of Asthma cohort, MAF=Minor allele frequency, SLSJ=The Saguenay–Lac-Saint-Jean asthma family cohort, Stat=Statistic.

Table 12. All significant genetic variants associated with the sensitization to outdoor aeroallergens in the SLSJ cohort ($P < 1E-05$) and association results for the EGEA cohort

Variant ID ^a	Chr	Position ^b	Minor allele	MAF	Stat	P _{SLSJ}	P _{EGEA} ^c	Locus	Nearest gene ^d	Variant information ^e
rs1690716	1	220,927,511	A	3.05E-02	20.33	6.51E-06	1.98E-01	1q41	<i>MTARC2</i>	Intron variant
rs74969563	1	226,202,298	T	9.58E-02	20.04	7.58E-06	7.14E-01	1q42.12	–	Intergenic variant
rs62139377	2	41,976,649	T	2.68E-01	29.16	6.68E-08	3.55E-01	2p21	–	Intergenic variant
rs62139378	2	41,976,763	T	2.68E-01	29.16	6.68E-08	3.64E-01	2p21	–	Intergenic variant
rs5830698	2	41,983,319	AT	2.69E-01	27.83	1.32E-07	N/A	2p21	–	Intergenic variant
rs11693207	2	41,984,278	A	3.08E-01	19.61	9.48E-06	8.65E-01	2p21	–	Intergenic variant
rs4952510	2	41,986,768	A	2.65E-01	26.89	2.15E-07	2.88E-01	2p21	–	Intergenic variant
rs4952374	2	41,997,648	C	2.53E-01	27.46	1.61E-07	4.10E-01	2p21	–	Intergenic variant
rs2005174	2	41,998,764	A	2.47E-01	22.72	1.87E-06	3.28E-01	2p21	–	Intergenic variant
rs755009	2	41,999,338	T	2.45E-01	21.56	3.42E-06	3.38E-01	2p21	–	Intergenic variant
rs72788670	2	41,999,848	C	2.53E-01	25.98	3.44E-07	3.58E-01	2p21	–	Intergenic variant
rs17530511	2	42,000,393	C	2.48E-01	22.38	2.24E-06	2.56E-01	2p21	–	Intergenic variant
rs11689227	2	42,001,225	A	2.47E-01	22.55	2.05E-06	3.18E-01	2p21	–	Intergenic variant
rs11695476	2	42,001,819	C	2.47E-01	22.89	1.72E-06	4.38E-01	2p21	–	Intergenic variant
rs248727	5	127,037,482	C	4.55E-01	22.66	1.94E-06	2.42E-01	5q23.2	<i>CCDC192</i>	1.6kb upstream variant
rs79507752	7	97,826,050	C	1.46E-02	20.64	5.54E-06	N/A	7q21.3	<i>LMTK2</i>	Intron variant
rs28396474	8	8,762,110	G	6.36E-02	20.06	7.50E-06	3.40E-01	8p23.1	–	Intergenic variant
rs57779176	8	8,763,310	G	6.41E-02	20.47	6.07E-06	4.92E-01	8p23.1	–	Intergenic variant
rs139588645	8	8,771,761	GA	6.41E-02	20.47	6.07E-06	N/A	8p23.1	–	Intergenic variant
rs10981012	9	114,453,006	A	1.20E-01	26.55	2.57E-07	5.73E-01	9q31.3	<i>SHOC1</i>	Intron variant
rs10981094	9	114,586,573	T	1.37E-01	24.04	9.44E-07	5.73E-01	9q31.3	–	Intergenic variant
rs3781361	10	105,375,817	T	4.79E-01	19.87	8.31E-06	2.86E-01	10q24.33	<i>SH3PXD2A</i>	Intron variant

rs55750092	12	13,252,806	T	6.18E-02	20.62	5.60E-06	3.19E-01	12p13.1	<i>GSG1</i>	Intron variant
rs76153631	12	13,257,151	C	6.18E-02	20.62	5.60E-06	3.17E-01	12p13.1	<i>GSG1</i>	0.5kb upstream variant
rs78376850	12	13,257,973	G	6.23E-02	20.95	4.72E-06	3.17E-01	12p13.1	<i>GSG1</i>	1.3kb upstream variant
rs55954298	12	13,263,429	ATT	6.19E-02	20.62	5.61E-06	N/A	12p13.1	<i>GSG1</i>	6.8kb upstream variant
rs55816732	12	13,264,236	T	6.18E-02	20.62	5.60E-06	3.85E-01	12p13.1	<i>GSG1</i>	7.6kb upstream variant
rs55928877	12	13,265,549	T	6.18E-02	20.62	5.60E-06	3.85E-01	12p13.1	<i>GSG1</i>	8.9kb upstream variant
rs4763934	12	13,266,755	A	6.18E-02	20.62	5.60E-06	3.79E-01	12p13.1	–	Intergenic variant
rs55729386	12	13,269,208	G	6.23E-02	20.95	4.72E-06	3.85E-01	12p13.1	–	Intergenic variant
rs2544037	12	48,215,180	G	3.34E-01	19.92	8.06E-06	5.19E-01	12q13.11	<i>HDAC7</i>	1.4kb upstream variant
rs1924783	13	25,985,873	G	3.27E-01	19.90	8.16E-06	8.48E-01	13q12.13	<i>ATP8A2</i>	Intron variant
rs7993071	13	25,986,112	G	3.27E-01	19.90	8.16E-06	8.48E-01	13q12.13	<i>ATP8A2</i>	Intron variant
rs17082242	13	25,987,881	A	3.28E-01	19.88	8.25E-06	8.42E-01	13q12.13	<i>ATP8A2</i>	Intron variant
rs7336497	13	25,988,454	A	3.27E-01	19.90	8.16E-06	8.26E-01	13q12.13	<i>ATP8A2</i>	Intron variant
rs7320645	13	25,988,656	C	3.27E-01	19.90	8.16E-06	8.28E-01	13q12.13	<i>ATP8A2</i>	Intron variant
rs28788274	15	25,010,316	A	3.59E-01	19.66	9.26E-06	7.45E-01	15q11.2	–	Intergenic variant
rs1578401	20	17,210,565	A	4.91E-01	19.85	8.37E-06	3.00E-01	20p12.1	<i>PCSK2</i>	Intron variant

^aVariant ID in bold character can also be found in the significant variant associated with both types of aeroallergens sensitization. ^bGenetic variant's position on the genome according to the human genome reference 19 (hg19). ^cIf a variant has a P value of NA in the EGEA cohort it means that this particular variant wasn't found during replication analysis. ^dGene in which the variant is located in or nearest to (≤ 10 kb on the right and left sides) according to the UCSC Genome browser (hg19). For variants which are intergenic (further than 10kb), no information is provided. ^eVariant position toward the gene according to UCSC. Abbreviation: Chr=Chromosome, EGEA=The French Epidemiological study on the Genetics and Environment of Asthma cohort, MAF=Minor allele frequency, SLSJ=The Saguenay–Lac-Saint-Jean asthma family cohort, Stat=Statistic.

Table 13. All significant genetic variants associated to the sensitization to both types of aeroallergens in the SLSJ cohort ($P < 1E-05$) and association results for the EGEA cohort

Variant ID ^a	Chr	Position ^b	Minor allele	MAF	Stat	P _{SLSJ}	P _{EGEA} ^c	Locus	Nearest gene ^d	Variant information ^e
rs62139377	2	41,976,649	T	2.68E-01	21.49	3.55E-06	4.60E-01	2p21	–	Intergenic variant
rs62139378	2	41,976,763	T	2.68E-01	21.49	3.55E-06	4.74E-01	2p21	–	Intergenic variant
rs5830698	2	41,983,319	AT	2.69E-01	20.27	6.72E-06	N/A	2p21	–	Intergenic variant
rs4952510	2	41,986,768	A	2.65E-01	19.79	8.66E-06	3.71E-01	2p21	–	Intergenic variant
rs4952374	2	41,997,648	C	2.53E-01	20.34	6.47E-06	4.11E-01	2p21	–	Intergenic variant
rs75269698	2	67,080,118	T	2.28E-02	21.10	4.36E-06	6.18E-01	2p14	–	Intergenic variant
rs76173900	2	67,177,740	G	2.10E-02	22.61	1.98E-06	7.55E-01	2p14	<i>LINC01799</i>	Intron variant
rs115461272	3	33,597,658	T	1.98E-02	19.71	9.02E-06	6.54E-01	3p22.3	<i>CLASP2</i>	Intron variant
rs4586862	3	77,103,310	C	1.24E-01	20.47	6.05E-06	7.83E-01	3p12.3	<i>ROBO2</i>	Intron variant
rs4975729	5	1,775,198	T	2.78E-01	19.88	8.23E-06	7.05E-01	5p15.33	–	Intergenic variant
rs115933755	5	34,118,989	C	3.82E-02	25.65	4.09E-07	1.20E-01	5p13.2	<i>C1QTNF3</i> , <i>C1QTNF3-AMACR</i>	Intron variant, intron variant
rs199871444	5	34,123,038	C	2.88E-02	24.43	7.71E-07	N/A	5p13.2	<i>C1QTNF3</i> , <i>C1QTNF3-AMACR</i>	Intron variant, intron variant
rs116762404	5	34,538,653	T	4.37E-02	21.13	4.30E-06	2.22E-01	5p13.2	–	Intergenic variant
rs6449729	5	63,702,811	A	3.64E-01	19.57	9.69E-06	8.16E-01	5q12.3	–	Intergenic variant
rs13437096	6	31,355,287	A	2.53E-01	20.01	7.70E-06	2.66E-03	6p21.33	<i>MICA-AS1</i>	6.8kb downstream variant
rs7756426	6	31,360,694	A	2.50E-01	20.24	6.84E-06	2.61E-03	6p21.33	<i>MICA-AS1</i>	1.4kb downstream variant

rs79028790	6	31,360,861	A	2.53E-01	20.01	7.70E-06	N/A	6p21.33	<i>MICA-AS1</i>	1.2kb downstream variant
rs113473106	6	31,361,058	G	2.50E-01	20.52	5.91E-06	2.74E-03	6p21.33	<i>MICA-AS1</i>	1.0kb downstream variant
rs28396474	8	8,762,110	G	6.36E-02	20.15	7.17E-06	2.11E-01	8p23.1	–	Intergenic variant
rs57779176	8	8,763,310	G	6.41E-02	20.49	6.00E-06	2.66E-01	8p23.1	–	Intergenic variant
rs139588645	8	8,771,761	GA	6.41E-02	20.49	6.00E-06	N/A	8p23.1	–	Intergenic variant
rs62512006	8	59,610,183	T	9.10E-02	21.23	4.08E-06	8.53E-01	8q12.1	–	Intergenic variant
rs62512007	8	59,610,523	G	9.10E-02	21.23	4.08E-06	8.54E-01	8q12.1	–	Intergenic variant
rs115793380	8	59,612,502	T	9.10E-02	21.23	4.08E-06	N/A	8q12.1	–	Intergenic variant
rs11774021	8	59,612,992	T	9.10E-02	21.23	4.08E-06	9.61E-01	8q12.1	–	Intergenic variant
rs116905394	8	59,690,382	G	7.33E-02	20.35	6.46E-06	9.60E-01	8q12.1	–	Intergenic variant
rs117022993	8	59,691,136	C	7.33E-02	20.35	6.46E-06	9.60E-01	8q12.1	–	Intergenic variant
rs62505091	8	59,694,627	G	7.33E-02	20.35	6.46E-06	9.99E-01	8q12.1	–	Intergenic variant
rs4743880	9	95,777,594	T	4.66E-01	19.84	8.44E-06	1.89E-01	9q22.31	<i>FGD3, LOC101927954</i>	Intron variant, 0.4kb upstream variant
rs10682558	9	95,778,654	G	4.66E-01	19.79	8.63E-06	N/A	9q22.31	<i>FGD3, LOC101927954</i>	Intron variant, 1.4kb upstream variant
rs10761191	9	95,781,040	T	4.69E-01	20.65	5.51E-06	1.99E-01	9q22.31	<i>FGD3, LOC101927954</i>	Intron variant, 3.8kb upstream variant
rs6479446	9	95,782,991	T	4.69E-01	19.75	8.82E-06	1.96E-01	9q22.31	<i>FGD3, LOC101927954</i>	Intron variant, 5.8kb upstream variant
rs10981012	9	114,453,006	A	1.20E-01	22.77	1.83E-06	2.46E-01	9q31.3	<i>SHOC1</i>	Intron variant

rs72754210	9	114,581,300	A	4.78E-01	20.56	5.77E-06	1.51E-01	9q31.3	–	Intergenic variant
rs10981094	9	114,586,573	T	1.37E-01	21.89	2.89E-06	3.94E-01	9q31.3	–	Intergenic variant
rs7997863	13	45,369,779	G	2.51E-01	22.04	2.68E-06	3.17E-01	13q14.12	<i>LINC00330</i>	3.9kb downstream variant
rs28788274	15	25,010,316	A	3.59E-01	29.18	6.60E-08	4.88E-01	15q11.2	–	Intergenic variant
rs67801593	15	25,011,361	G	2.89E-01	19.86	8.34E-06	6.47E-01	15q11.2	–	Intergenic variant
rs67799280	15	25,011,443	T	2.89E-01	19.86	8.34E-06	6.46E-01	15q11.2	–	Intergenic variant
rs200054505	15	25,012,281	AAT	2.89E-01	19.86	8.34E-06	N/A	15q11.2	–	Intergenic variant
rs113986838	15	25,012,336	A	2.71E-01	20.84	4.99E-06	7.29E-01	15q11.2	–	Intergenic variant
rs7498117	15	25,012,842	G	2.89E-01	19.86	8.34E-06	6.27E-01	15q11.2	–	Intergenic variant
rs4340312	15	25,013,023	G	2.89E-01	19.86	8.34E-06	6.27E-01	15q11.2	–	Intergenic variant
rs2123683	15	25,013,195	G	2.88E-01	20.76	5.21E-06	6.46E-01	15q11.2	–	Intergenic variant
rs34980104	15	25,013,533	A	2.89E-01	19.86	8.34E-06	6.27E-01	15q11.2	–	Intergenic variant
rs8028920	15	25,013,804	T	2.89E-01	19.86	8.34E-06	6.27E-01	15q11.2	–	Intergenic variant
rs8029309	15	25,014,094	G	2.89E-01	19.86	8.34E-06	6.46E-01	15q11.2	–	Intergenic variant
rs8033617	15	25,014,871	A	2.89E-01	19.86	8.34E-06	6.27E-01	15q11.2	–	Intergenic variant
rs8035301	15	25,015,163	T	2.89E-01	19.86	8.34E-06	6.27E-01	15q11.2	–	Intergenic variant
rs1597994	15	25,015,805	G	2.85E-01	20.37	6.37E-06	6.27E-01	15q11.2	–	Intergenic variant
rs4254306	15	25,017,123	C	2.89E-01	19.86	8.34E-06	6.29E-01	15q11.2	–	Intergenic variant
rs12910736	15	25,025,458	T	2.90E-01	19.92	8.06E-06	6.26E-01	15q11.2	–	Intergenic variant
rs1603731	15	25,027,457	C	2.90E-01	19.92	8.06E-06	6.26E-01	15q11.2	–	Intergenic variant
rs7180394	15	25,029,728	G	2.90E-01	19.92	8.06E-06	6.01E-01	15q11.2	–	Intergenic variant
rs7163591	15	25,030,127	G	2.90E-01	19.92	8.06E-06	6.02E-01	15q11.2	–	Intergenic variant

rs9896435	17	29,872,405	T	1.22E-02	20.58	5.72E-06	N/A	17q11.2	<i>RAB11FIP4, MT-TA, TRT-CGT4-1</i>	7.2kb downstream variant, 4.7kb upstream variant, 4.7kb upstream variant
rs148593533	17	31,258,657	A	2.06E-02	23.73	1.11E-06	N/A	17q11.2	<i>TMEM98</i>	Exon variant

^aVariant ID in bold character can also be found in the significant variant associated to indoor or outdoor aeroallergens sensitization. ^bGenetic variant's position on the genome according to the human genome reference 19 (hg19). ^cIf a variant has a P value of NA in the EGEA cohort it means that this particular variant wasn't found during replication analysis. ^dGene in which the variant is located in or nearest to (≤ 10 kb on the right and left sides) according to the UCSC Genome browser (hg19). For variants which are intergenic (further than 10kb), no information is provided. ^eVariant position toward the gene according to UCSC. Abbreviation: Chr=Chromosome, EGEA=The French Epidemiological study on the Genetics and Environment of Asthma cohort, MAF=Minor allele frequency, SLSJ=The Saguenay–Lac-Saint-Jean asthma family cohort, Stat=Statistic.

Table 14. Biological function of proximal genes of genetic variants identified by GWAS for the indoor sensitization

Gene name	Gene symbol	Cytoband	Protein class	Function	Association to disease	References
Regulation						
TraB domain containing 2B	<i>TRABD2B</i>	1p33	Metalloprotease	Negative regulation of Wnt/ β -catenin signaling pathway by inactivation of Wnt proteins	Asthma (indirect association)	1, 2
Ras homolog, MTORC1 binding	<i>RHEB</i>	7q36.1	Small GTPase	Involved in TOR signaling activation by its binding to mTORC1. Involved in regulating macrophage polarisation	Allergic asthma	3, 4
Telomere maintenance 2	<i>TELO2</i>	16p13.3	–	Positive regulation of TORC1 and TORC2 signaling by stabilizing and promoting mTORC1 and mTORC2 complexes assembly	–	5, 6
Ring finger protein 180	<i>RNF180</i>	5q12.3	Ubiquitin-protein ligase	Positive regulation of protein ubiquitination and of proteasomal ubiquitin-dependent protein catabolic process, protein polyubiquitination and regulation of catalytic activity	Allergic asthma and rhinitis	7, 8

Laminin subunit alpha 1	<i>LAMA1</i>	18p11.31-18p11.23	Cell adhesion molecule	Involved in the lung formation during embryogenic development through its mediation of attachment, migration and organisation of cells into tissues. Positive regulation of integrin-mediated pathway	Pulmonary fibrosis	9, 10
-------------------------	--------------	-------------------	------------------------	---	--------------------	-------

Immune signalization and response

Ribonuclease T2	<i>RNASET2</i>	6q27	Endoribonuclease	Recognition and degradation of microbial pathogens RNA and RNA-dependent activation of TLR8	Allergic disease (variant)	11, 12
Lysophosphatidic acid receptor 1	<i>LPAR1</i>	9q31.3	G-protein coupled receptor	Contributes to the response to tissue damage and pathogens by modulating certain functions like proliferation, platelet aggregation, cell migration, smooth muscle aggregation and chemotaxis. Role in triggering inflammation in response to bacterial lipopolysaccharide	Lung fibrosis, asthma	13-16
Small nucleolar RNA host gene 7	<i>SNHG7</i>	9q34.3	–	Promote M2 polarization of macrophages and involved in regulation of autophagy in cancer cells	Resistance to docetaxel in lung adenocarcinoma cells	17

Cellular process

Spectrin repeat containing nuclear envelope family member 3	<i>SYNE3</i>	14q32.13	–	Cytoskeleton organization, regulation of cell shape, protein localization to membrane, part of the meiotic nuclear membrane microtubule tethering complex and a component in the linker of nucleoskeleton and cytoskeleton (LINC) complex	–	18
Lipocalin 10	<i>LCN10</i>	9q34.3	Transfer/Carrier protein	Binding and transportation in specific cells of small hydrophobic ligands. Lipocalins are involved in inflammation and immunity	–	19, 20

Non-protein coding

VLDLR antisense RNA 1	<i>VLDLR-AS1</i>	9p24.2	–	Unknown function and is affiliated to the lncRNA class	Susceptibility to pneumonia	21
Long intergenic non-protein coding RNA 2208	<i>LINC02208</i>	5q23.1	–	Unknown function and is affiliated to the lncRNA class	Response to infections (variant)	22
Long intergenic non-protein coding RNA 330	<i>LINC00330</i>	13q14.12	–	Unknown function and is affiliated to the lncRNA class	–	–
Micro RNA 548	<i>MIR548</i> (<i>RF01061-187</i>)	9q31.3	–	miRNA	–	–

Other

PALM2 and AKAP2 fusion	<i>PALM2AKAP2</i>	9q31.3	–	Unknown function	–	–
Pentraxin 4	<i>PTX4</i>	16p13.3	Scaffold/adaptor protein	Unknown function, but is hypothesized to have a role in innate immunity by structural homology with short pentraxins	–	23

References:

1: Zhang X, *et al. Cell.* 2012;149(7):1565-1577. **2:** Kwak HJ, *et al. Exp Mol Med.* 2015;47(12):e198. **3:** Long X, *et al. Curr Bio.* 2005;15(8):702-713. **4:** Li K, *et al. Sci Rep.* 2017;7:42655. **5:** Kaizuka T, *et al. J Biol Chem.* 2010;285(26):20109-20116. **6:** Takai H, *et al. Genes Dev.* 2010;24(18):2019-2030. **7:** Ogawa M, *et al. Genes Cells.* 2008;13(4):397-409. **8:** Morin A, *et al. Clin Epigenetics.* 2017;9:3. **9:** Christie PE, *et al. Eur Respir J.* 2004;24:107-115. **10:** Lee C-M, *et al. JCI Insight.* 2018;3(18):e99574. **11:** Greulich W, *et al. Cell.* 2019;179(6):1264-1275.e.13. **12:** Ferreira MA, *et al. Nat Genet.* 2017; 49(12):1752-1757. **13:** Tager AM, *et al. Nat Med.* 2008;14:45-54. **14:** George J, *et al. J Periodontol.* 2009;80(8):1338-1347. **15:** Song HY, *et al. Biochim Biophys Acta.* 2010;1801(1):23-30. **16:** Zhao Y and Natarajan V. *Cell Signal.* 2009;21(3):367-377. **17:** Talotta R, *et al. Biochim Biophys Acta Mol Basis Dis.* 2022;1868(2):166291. **18:** Morgan JT, *et al. Mol Biol Cell.* 2011;22(22):4324-4334. **19:** Suzuki K, *et al. Gene.* 2004;339:49-59. **20:** Jensen-Jarolim E, *et al. Allergy.* 2016;71(3):286-294. **21:** Hayden LP, *et al. Am J Respir Cell Mol Biol.* 2017;56(1):20-28. **22:** Zhang K, *et al. Cancer Letters.* 2022;526:142-154. **23:** Wang Z, *et al. Front Immunol.* 2020;11:1757.

Table 15. Biological function of proximal genes of genetic variants identified by GWAS for the outdoor sensitization

Gene name	Gene symbol	Cytoband	Protein class	Function	Association to disease	References
Immune signaling and response						
Histone deacetylase 7	<i>HDAC7</i>	12q13.11	Histone modifying enzyme	Negative regulation of interleukin-2 production, negative regulation of NIK/NF- κ B signaling, positive regulation of the transcriptional repressor activity of FOXP3	Allergic disease (variant), adult-onset asthma	1-3
Metabolic process						
Proprotein convertase subtilisin/kexin type 2	<i>PCSK2</i>	20p12.1	Serine protein	Protein, peptide hormone and insulin processing. Proteolytic activation of polypeptide hormones like α -Melanocyte-stimulating hormone (α -MSH) and neuropeptide precursors. Involved in secretory pathway	Atopic dermatitis (indirect association)	4, 5
Mitochondrial amidoxime reducing component 2	<i>MTARC2</i>	1q41	–	Member of the N-reductive enzyme system which is involved in nitrogen compound metabolic process	Immune escape in hepatocellular carcinoma	6
SH3 and PX domains 2A	<i>SH3PXD2A</i>	10q24.33	Scaffold/adaptor protein	Involved in invadopodia and podosome formation and extracellular matrix degradation. Allows NOX1- or NOX-3 dependent reactive oxygen species (ROS) generation and ROS localisation	Diisocyanate-induced asthma (variant)	7, 8

Cellular process

Shortage in chiasmata 1	<i>SHOC1</i>	9q31.3	ERCC4-(HhH) ₂ -like protein	Involved in crossing over formation during meiosis prophase I in germ cells	Lung cancer survival (variant)	9-12
Germ cell associated 1	<i>GSG1</i>	12p13.1	Cytoskeletal protein	Predicted to enable RNA polymerase binding activity and its promoter interacts with HEBP1's promoter which allows HEBP1 expression in T-cell	Lung squamous cell carcinoma survival	13, 14
ATPase phospholipid transporting 8A2	<i>ATP8A2</i>	13q12.13	Primary active transporter	Aminophospholipid translocase. Is involved in regulating the number of intestinal intraepithelial lymphocytes	Lung adenocarcinoma	15-17

Other

Coiled-coil domain containing 192	<i>CCDC192</i>	5q23.2	–	Unknown function	–	–
Lemur tyrosine kinase 2	<i>LMTK2</i>	7q21.3	Non-receptor serine/threonine protein kinase	Involved in the nerve growth factor (NGF)-TrkA signaling	Cystic fibrosis	18, 19

References:

1: Li B, *et al. Proc Natl Acad Sci U S A.* 2007;104(11):4571-4576. **2:** Ferreira MA, *et al. Nat Genet.* 2017;49(12):1752-1757. **3:** Ferreira MAR, *et al. Am J Hum Genet.* 2019;104(4):665-684. **4:** Chen Y-C, *et al. Diabetes Obes Metab.* 2018;20(suppl. 2):64-76. **5:** Andoh T, *et al. Am J Pathol.* 2019;189(9):1775-1785. **6:** Wu D, *et al. Front Genet.* 2021;12:790093. **7:** Saini P and Courtneidge SA. *J Cell Sci.* 2018;131(1):jcs203661. **8:** Yucesoy B, *et al. Toxicol Sci.* 2015;146(1):192-201. **9:** Guiraldelli MF, *et al. PloS Genet.*

2018;14(5):e1007381. **10:** Yao C, et al. *J Med Genet.* 2021;58(10):679-686. **11:** Wang W, et al. *Mol Hum Reprod.* 2022;28(6):gaac015. **12:** Brhane Y, et al. *Cancer Epidemiol Biomarkers Prev.* 2020;29(10):1983-1992. **13:** Chandra V, et al. *Nat Genet.* 2021;53(1):110-119. **14:** Qi L, et al. *Peer J.* 2019;7:e7727. **15:** Coleman JA, et al. *J Biol Chem.* 2009;284(47):32670–32679. **16:** Ishifune C, et al. *PLoS Biol.* 2019;17(5):e3000262. **17:** Zhang X, et al. *Front Cell Dev Biol.* 2021;9:740487. **18:** Kawa S, et al. *Genes Cells.* 2004;9(3):219-232. **19:** Cruz DF, et al. *Front Cell Dev Biol.* 2020;8:58.

Table 16. Biological function of proximal genes of genetic variants identified by GWAS for both types of sensitization

Gene name	Gene symbol	Cytoband	Protein class	Function	Association to disease	References
Immune signaling and response						
C1q and TNF related 3	<i>C1QTNF3</i>	5p13.2	Scaffold/adaptor protein	Negative regulation of NIK/NF- κ B signaling, of IL-6 production, of gluconeogenesis, of CCL2 production, of inflammatory response and positive regulation of cytokine production	Allergic rhinitis (down regulated)	1-3
Transmembrane protein 98	<i>TMEM98</i>	17q11.2	Transmembrane protein	Induce T-helper 1 cell differentiation	Non-small cell lung cancer	4, 5
FYVE, RhoGEF and PH domain containing 3	<i>FGD3</i>	9q22.31	Guanyl-nucleotide exchange factor	Activate CDC42 by its GTPase binding activity which is involved in cell migration and adhesion	Prognostic of cancer like breast cancer and lung adenocarcinoma	6, 7
Cellular process						
Cytoplasmic linker associated protein 2	<i>CLASP2</i>	3p22.3	Non-motor microtubule binding protein	Involved in stabilizing and organizing microtubules which are used for processes like migration and can increase wound healing. Involved in the inhibition of the JAK/STAT pathway by its association with SOCS3, and CLIP-170, IQGAP1	Anti-inflammatory effect, non-small cell lung cancer risk	8-11

Roundabout guidance receptor 2	<i>ROBO2</i>	3p12.3	Cell adhesion molecule	Receptor of SLIT2 that allows axon guidance and cell migration	Asthma, allergic rhinitis (down regulated)	3, 12
Shortage in chiasmata 1	<i>SHOC1</i>	9q31.3	ERCC4-(HhH) ₂ -like protein	Involved in crossing over formation during meiosis prophase I in germ cells	Lung cancer survival (variant)	13-16
RAB11 family interacting protein 4	<i>RAB11FIP4</i>	17q11.2	–	Involved in regulation of endocytic traffic and cytokinesis	Asthma (CpG), Cancer	17-19

Non-protein coding

Uncharacterized LOC101927954	<i>LOC101927954</i>	9q22.31	–	Unknown function and is affiliated to the lncRNA class	–	–
Long intergenic non-protein coding RNA 330	<i>LINC00330</i>	13q14.12	–	Unknown function and is affiliated to the lncRNA class	Diisocyanate-induced asthma (variant), lung adenocarcinoma (up-regulated), cystic fibrosis (expression)	20-22
<i>MICA</i> antisense RNA 1	<i>MICA-AS1</i>	6p21.33	–	Unknown function and is affiliated to the lncRNA class	Pneumonia (variant)	23
Long intergenic non-protein coding RNA 1799	<i>LINC01799</i>	2p14	–	Unknown function and is affiliated to the lncRNA class	Milk allergy	24

Transfer RNA-threonine (anticodon CGT) 4-1	<i>TRT-CGT4-1</i>	17q11.2	–	Unknown function and is affiliated to the tRNA class	–	–
Mitochondrially encoded TRNA Ala (GCN)	<i>MT-TA</i>	17q11.2	–	Unknown function and is affiliated to the tRNA class	–	–
<i>C1QTNF3-AMACR</i> readthrough	<i>C1QTNF3-AMACR</i>	5p13.2	–	Unknown function, is affiliated to the lncRNA class and is a candidate for nonsense-mediated mRNA decay	–	–

References:

1: Weigert J, et al. *FEBS Lett.* 2005;579(25):5565-5570. **2:** Li D, et al. *Medicine (Baltimore).* 2015;94(36):e1503. **3:** Wei X, et al. *BMC Med Genomics.* 2021;14:100. **4:** Fu W, et al. *J Interferon Cytokine Res.* 2015;35(9):720-733. **5:** Mao M, et al. *Int J Clin Exp Pathol.* 2015;8(12):15661-15669. **6:** Hayakawa M, et al. *Genes Cells.* 2008;13(4):329-342. **7:** Willis S, et al. *JCO Precis Oncol.* 2017;1:PO.17.00009. **8:** Akhmanova A, et al. *Cell.* 2001;104(6):923-935. **9:** Kim M, et al. *Wound Repair Regen.* 2019;27(4):345-359. **10:** Karki P, et al. *J Biol Chem.* 2021;296:100239. **11:** You D, et al. *BMC Med.* 2022;20:203. **12:** Ding L, et al. *Hum Genomics.* 2013;7(1):16. **13:** Guiraldelli MF, et al. *PloS Genet.* 2018;14(5):e1007381. **14:** Yao C, et al. *J Med Genet.* 2021;58(10):679-686. **15:** Wang W, et al. *Mol Hum Reprod.* 2022;28(6):gaac015. **16:** Brhane Y, et al. *Cancer Epidemiol Biomarkers Prev.* 2020;29(10):1983-1992. **17:** Wallace DME, et al. *Biochem Biophys Res Commun.* 2002;299(5):770-779. **18:** Cardenas A, et al. *Nat Commun.* 2019;10:3095. **19:** Machesky LM. *Biosci Rep.* 2019;39(1):BSR20182219. **20:** Yucesoy B, et al. *Toxicol Sci.* 2015;146(1):192-201. **21:** Hu J, et al. *Transl Cancer Res.* 2019;8(4):1046-1064. **22:** Rehman T, et al. *J Clin Invest.* 2021;131(16):e150398. **23:** Sakaue S, et al. *Nat Genet.* 2021;53(10):1415-1424. **24:** Hong X, et al. *Nat Commun.* 2015;6:6304.

Table 17. Significant cis-eQTL results including genetic variants associated with each sensitization type in naïve CD4⁺ T cells and eosinophils

Variant ID	Gene expression	Stat	P value	FDR	Beta	Cell type
Indoor sensitization						
rs117148242	<i>RNF208</i>	-4.38	2.23E-05	1.06E-02	-4.55E-01	CD4 ⁺
rs3798307	<i>RNASET2</i>	-6.66	5.02E-10	7.14E-07	-2.49E-01	CD4 ⁺
	<i>MIR3939</i>	-5.40	2.55E-07	1.82E-04	-1.70E-01	CD4 ⁺
Outdoor sensitization						
rs1690716	<i>MARK1</i>	-4.35	2.49E-05	1.31E-02	-2.83E-02	CD4 ⁺
Both types of sensitization						
	<i>LOC100128361</i>	4.51/ 4.57/ 4.51/ 4.53	1.30E-05/ 1.01E-05/ 1.30E-05/ 1.23E-05	3.33E-03/ 3.33E-03/ 3.33E-03/ 3.33E-03	1.14E-01/ 1.14E-01/ 1.14E-01/ 1.13E-01	CD4 ⁺
rs10682558/ rs10761191/ rs4743880/ rs6479446	<i>CENPP</i>	4.33/ 4.25/ 4.33/ 3.99	2.79E-05/ 3.78E-05/ 2.79E-05/ 1.04E-04	3.47E-03/ 3.52E-03/ 3.47E-03/ 7.56E-03	1.42E-01/ 1.38E-01/ 1.42E-01/ 1.31E-01	CD4 ⁺
	<i>FGD3</i>	3.83/ 3.97/ 3.83/ 4.16	1.87E-04/ 1.11E-04/ 1.87E-04/ 5.39E-05	1.13E-02/ 7.56E-03/ 1.13E-02/ 4.61E-03	3.85E-02/ 3.92E-02/ 3.85E-02/ 4.11E-02	CD4 ⁺
rs113473106/ rs13437096/ rs7756426/ rs79028790	<i>AIF1</i>	-4.39/ -4.28/ -4.28/ -4.28	2.18E-05/ 3.38E-05/ 3.38E-05/ 3.38E-05	3.47E-03/ 3.47E-03/ 3.47E-03/ 3.47E-03	-4.44E-01/ -4.34E-01/ -4.34E-01/ -4.34E-01	CD4 ⁺
	<i>SKIV2L</i>	-3.61/ -3.57/ -3.57/ -3.57	4.23E-04/ 4.86E-04/ 4.86E-04/ 4.86E-04	2.38E-02/ 2.38E-02/ 2.38E-02/ 2.38E-02	-1.17E-01/ -1.16E-01/ -1.16E-01/ -1.16E-01	CD4 ⁺

rs116762404	<i>DNAJC21</i>	-4.06	7.95E-05	6.27E-03	-1.05E-01	CD4 ⁺
rs148593533	<i>ASIC2</i>	-4.24	4.33E-05	2.15E-02	-1.13E-01	Eos
	<i>SPACA3</i>	-4.30	3.37E-05	2.15E-02	-8.37E-03	Eos
rs4975729	<i>LOC100506688</i>	-3.85	1.86E-04	4.62E-02	-1.05E-02	Eos
rs9896435	<i>SH3GL1P2</i>	-3.98	1.14E-04	3.76E-02	-6.71E-01	Eos

Abbreviations: CD4⁺=Naïve CD4⁺ T cells, Eos=Eosinophils, FDR=False discovery rate, Stat=Statistic.

Table 18. Significant cis-mQTL results including genetic variants associated with the indoor sensitization in naïve CD4⁺ T cells and eosinophils

Variants ID ^a	CpG	Stat	P value	FDR	Beta	CpG's position in gene		Cell type
						Body	Promoter	
rs75285112	1:49,242,583	-4.69	5.41E-06	2.67E-02	-1.26E-01	<i>AGBL4</i>	<i>BEND5</i>	CD4 ⁺
	1:48,937,957	-4.61	7.60E-06	3.29E-02	-1.23E-01	–	<i>SPATA6</i>	CD4 ⁺
rs356593	5:63,461,232	-4.58	8.79E-06	3.66E-02	-6.45E-02	–	<i>RNF180</i>	CD4 ⁺
rs3798307	6:167,365,877	6.92	7.64E-11	6.78E-06	1.91E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,365,888	7.04	3.83E-11	6.78E-06	1.99E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,365,855	6.68	2.89E-10	2.05E-05	1.68E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,371,100	6.52	6.90E-10	3.50E-05	1.96E-01	–	<i>RNASET2</i>	CD4 ⁺
	6:167,070,197	-6.45	9.82E-10	3.88E-05	-1.22E-01	<i>RPS6KA2</i>	–	CD4 ⁺
	6:167,365,340	6.32	1.98E-09	7.03E-05	1.00E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,365,909	6.30	2.21E-09	7.09E-05	1.81E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,366,511	6.27	2.55E-09	7.09E-05	1.48E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,365,308	6.18	4.10E-09	1.04E-04	9.19E-02	<i>RNASET2</i>	–	CD4 ⁺
	6:167,365,386	6.14	5.21E-09	1.20E-04	1.21E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,365,863	6.13	5.40E-09	1.20E-04	1.67E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,070,154	-6.04	8.61E-09	1.51E-04	-1.34E-01	<i>RPS6KA2</i>	–	CD4 ⁺
	6:167,365,847	6.02	9.29E-09	1.51E-04	1.31E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,275,549	-5.95	1.34E-08	1.98E-04	-1.09E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,070,135	-5.85	2.22E-08	3.15E-04	-1.27E-01	<i>RPS6KA2</i>	–	CD4 ⁺	

6:167,070,428	-5.80	2.85E-08	3.89E-04	-1.18E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,070,325	-5.78	3.25E-08	4.22E-04	-1.02E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,366,196	5.77	3.32E-08	4.22E-04	1.36E-01	<i>RNASET2</i>	–	CD4 ⁺
6:167,244,035	5.58	8.52E-08	9.45E-04	1.65E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,070,040	-5.46	1.58E-07	1.70E-03	-1.07E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,371,091	5.44	1.73E-07	1.80E-03	1.64E-01	–	<i>RNASET2</i>	CD4 ⁺
6:167,070,296	-5.36	2.52E-07	2.55E-03	-1.03E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,365,903	5.32	3.10E-07	3.06E-03	1.53E-01	<i>RNASET2</i>	–	CD4 ⁺
6:167,069,572	-5.28	3.59E-07	3.36E-03	-1.25E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,070,309	-5.22	4.95E-07	4.51E-03	-1.11E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,070,607	-5.03	1.18E-06	9.31E-03	-1.13E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,070,127	-5.01	1.30E-06	9.87E-03	-1.10E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,349,958	5.00	1.37E-06	9.90E-03	1.68E-01	<i>RNASET2</i>	–	CD4 ⁺
6:167,070,053	-4.89	2.17E-06	1.51E-02	-1.09E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,070,616	-4.89	2.26E-06	1.55E-02	-1.18E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,411,355	4.87	2.39E-06	1.60E-02	9.84E-02	<i>MIR3939</i>	<i>CEP43, CCR6</i>	CD4 ⁺
6:167,366,359	4.84	2.77E-06	1.76E-02	1.14E-01	<i>RNASET2</i>	–	CD4 ⁺
6:167,705,750	-4.84	2.82E-06	1.76E-02	-9.32E-02	<i>UNC93A</i>	–	CD4 ⁺
6:167,069,618	-4.81	3.11E-06	1.87E-02	-1.02E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,069,582	-4.78	3.66E-06	2.08E-02	-1.09E-01	<i>RPS6KA2</i>	–	CD4 ⁺
6:167,371,869	4.76	3.91E-06	2.17E-02	1.10E-01	–	<i>RNASET2</i>	CD4 ⁺
6:167,070,263	-4.71	4.83E-06	2.56E-02	-1.09E-01	<i>RPS6KA2</i>	–	CD4 ⁺

	6:167,070,327	-4.70	5.09E-06	2.66E-02	-8.06E-02	<i>RPS6KA2</i>	–	CD4 ⁺
	6:167,070,027	-4.69	5.34E-06	2.67E-02	-9.33E-02	<i>RPS6KA2</i>	–	CD4 ⁺
	6:167,366,501	4.70	5.21E-06	2.67E-02	1.27E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,365,735	4.66	6.08E-06	2.92E-02	9.51E-02	<i>RNASET2</i>	–	CD4 ⁺
	6:167,350,137	4.59	8.33E-06	3.52E-02	1.41E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,506,241	-4.55	9.80E-06	3.95E-02	-9.08E-02	<i>CCR6</i>	–	CD4 ⁺
	6:167,368,710	4.55	9.96E-06	3.97E-02	1.57E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,368,630	4.53	1.08E-05	4.21E-02	1.74E-01	<i>RNASET2</i>	–	CD4 ⁺
	6:167,366,232	4.50	1.21E-05	4.68E-02	1.09E-01	<i>RNASET2</i>	–	CD4 ⁺
	7:152,133,539	-6.09	6.51E-09	1.28E-04	-2.11E-01	–	<i>FABP5P3, KMT2C</i>	CD4 ⁺
	7:150,946,081	-5.71	4.57E-08	5.41E-04	-2.03E-01	<i>SMARCD3</i>	–	CD4 ⁺
rs75632282/ rs79861276	7:150,929,683	-5.17	6.11E-07	5.16E-03	-1.11E-01	<i>CHPF2</i>	–	CD4 ⁺
	7:151,217,681	-5.00	1.33E-06	9.87E-03	-2.44E-01	–	<i>RHEB</i>	CD4 ⁺
	7:151,216,510	-4.65	6.27E-06	2.92E-02	-1.31E-01	<i>RHEB</i>	–	CD4 ⁺
	7:151,217,624	-4.57	8.97E-06	3.66E-02	-2.14E-01	–	<i>RHEB</i>	CD4 ⁺
	9:139,010,681	-6.02	9.36E-09	1.51E-04	-2.32E-01	<i>TMEM250</i>	–	CD4 ⁺
	9:139,090,726	-6.05	8.19E-09	1.51E-04	-3.26E-01	<i>LHX3</i>	–	CD4 ⁺
	9:138,987,016	-5.65	6.26E-08	7.17E-04	-1.76E-01	<i>NACC2</i>	–	CD4 ⁺
rs117148242	9:139,137,291	-5.30	3.41E-07	3.27E-03	-2.84E-01	<i>QSOX2</i>	–	CD4 ⁺
	9:138,799,172	-5.20	5.32E-07	4.73E-03	-2.53E-01	–	<i>CAMSAP1</i>	CD4 ⁺
	9:138,852,761	-4.85	2.63E-06	1.70E-02	-1.68E-01	<i>UBAC1</i>	–	CD4 ⁺
	9:138,799,768	-4.81	3.23E-06	1.91E-02	-2.37E-01	–	<i>CAMSAP1</i>	CD4 ⁺

	9:139,223,206	4.62	7.31E-06	3.25E-02	3.70E-01	<i>GPSM1</i>	<i>DKFZP434A062</i>	CD4 ⁺
	9:139,010,709	-4.60	7.96E-06	3.41E-02	-1.49E-01	<i>TMEM250</i>	–	CD4 ⁺
	9:138,799,535	-4.53	1.05E-05	4.16E-02	-1.91E-01	–	<i>CAMSAP1</i>	CD4 ⁺
rs2892077	10:5,932,240	-4.62	7.21E-06	3.24E-02	-4.15E-02	<i>FBH1</i>	<i>ANKRD16</i>	CD4 ⁺
	16:2,246,875	-4.96	1.62E-06	1.15E-02	-7.52E-02	–	<i>CASKIN1</i>	CD4 ⁺
	16:1,361,829	-4.77	3.70E-06	2.08E-02	-1.69E-01	<i>UBE2I</i>	–	CD4 ⁺
rs13336442	16:1,832,561	-4.73	4.51E-06	2.43E-02	-4.50E-02	<i>SPSB3</i>	<i>NUBP2</i>	CD4 ⁺
	16:1,361,900	-4.65	6.33E-06	2.92E-02	-1.51E-01	<i>UBE2I</i>	–	CD4 ⁺
	16:1,673,702	4.66	6.23E-06	2.92E-02	1.26E-01	<i>CRAMP1</i>	–	CD4 ⁺
	16:1,361,767	-4.61	7.55E-06	3.29E-02	-1.31E-01	<i>UBE2I</i>	–	CD4 ⁺
rs75285112	1:47,696,635	-7.95	2.37E-13	8.33E-08	-1.86E-01	<i>STIL, TAL1</i>	–	Eos
	6:167,411,428	5.99	1.18E-08	6.93E-04	1.11E-01	–	<i>MIR3939, CEP43, CCR6</i>	Eos
	6:167,371,091	5.76	3.84E-08	1.35E-03	1.44E-01	–	<i>RNASET2</i>	Eos
	6:167,411,355	5.68	5.51E-08	1.76E-03	1.09E-01	<i>MIR3939</i>	<i>CEP43, CCR6</i>	Eos
	6:167,070,154	-5.23	4.79E-07	9.34E-03	-1.29E-01	<i>RPS6KA2</i>	–	Eos
rs3798307	6:167,069,833	-5.15	7.17E-07	1.24E-02	-1.70E-01	<i>RPS6KA2</i>	–	Eos
	6:167,571,584	5.15	7.07E-07	1.24E-02	1.04E-01	–	<i>GPR31</i>	Eos
	6:167,070,135	-5.06	1.05E-06	1.60E-02	-1.32E-01	<i>RPS6KA2</i>	–	Eos
	6:168,198,137	-4.93	1.96E-06	2.65E-02	-1.20E-01	–	<i>LINC01558</i>	Eos
	6:167,454,660	4.76	4.08E-06	4.93E-02	1.27E-01	<i>CCR6</i>	–	Eos
	7:151,574,854	-5.83	2.62E-08	1.02E-03	-1.48E-01	<i>PRKAG2-AS1</i>	<i>PRKAG2</i>	Eos

rs75632282/ rs79861276	7:151,217,339	-5.35	2.83E-07	6.22E-03	-1.29E-01	–	<i>RHEB</i>	Eos
	7:150,812,585	-5.13	7.77E-07	1.24E-02	-2.13E-01	<i>AGAP3</i>	–	Eos
rs117148242	9:138,986,594	-6.98	6.24E-11	6.99E-06	-1.12E-01	<i>NACC2</i>	–	Eos
	9:138,853,073	-4.94	1.82E-06	2.55E-02	-2.20E-01	<i>UBAC1</i>	–	Eos
rs2892077	10:5,709,119	5.24	4.60E-07	9.34E-03	1.22E-01	–	<i>ASB13</i>	Eos
	10:5,709,177	4.96	1.66E-06	2.42E-02	1.39E-01	–	<i>ASB13</i>	Eos
rs13336442	16:1,838,095	5.45	1.73E-07	4.38E-03	2.09E-01	<i>NUBP2</i>	–	Eos
	16:2,034,383	-4.88	2.36E-06	3.07E-02	-5.40E-02	<i>GFER</i>	–	Eos

^aCpGs and variants shown are those with significant mQTL results (FDR<5.00E-02) in the whole sample. Abbreviations: CD4⁺=Naïve CD4⁺ T cells, CpG=Cytosine preceding Guanine sites, Eos=Eosinophils, FDR=False discovery rate, Stat=Statistic.

Table 19. Significant cis-mQTL results including genetic variants associated to the outdoor sensitization in naïve CD4⁺ T cells and eosinophils

Variants ID ^a	CpG	Stat	P value	FDR	Beta	CpG's position in gene		Cell type
						Body	Promoter	
rs79507752	7:97,910,856	-4.76	3.99E-06	4.53E-02	-1.42E-01	–	<i>BRI3</i>	CD4 ⁺
rs3781361	10:105,437,661	4.71	4.88E-06	4.53E-02	8.50E-02	<i>SH3PXD2A</i>	–	CD4 ⁺
rs4763934/ rs55729386/ rs55750092/ rs55816732/ rs55928877/ rs76153631/ rs78376850/ rs55954298^b	12:12,869,980	-4.73/ -4.72	4.57E-06/ 4.77E-06	4.53E-02	-5.45E-02	–	<i>CDKN1B</i>	CD4 ⁺
rs2544037	12:48,501,168	5.09	8.93E-07	4.53E-02	7.83E-02	<i>PFKM</i>	<i>SENP1</i>	CD4 ⁺
rs1690716	1:221,052,581	-5.05	1.14E-06	2.25E-02	-9.62E-02	–	<i>HLX</i>	Eos
	1:220,445,787	-5.00	1.38E-06	2.34E-02	-1.27E-01	<i>RAB3GAP2</i>	–	Eos
rs74969563	1:221,915,475	-4.63	7.17E-06	4.26E-02	-1.03E-01	<i>DUSP10</i>	–	Eos
	1:225,997,848	-5.38	2.37E-07	7.05E-03	-4.70E-02	<i>EPHX1</i>	–	Eos
rs79507752	7:98,741,384	-7.46	4.15E-12	4.94E-07	-1.19E-01	<i>SMURF1</i>	–	Eos
	7:98,476,013	-5.43	1.89E-07	7.05E-03	-1.21E-01	–	<i>TRRAP</i>	Eos
rs3781361	10:105,437,746	5.60	8.26E-08	4.91E-03	1.07E-01	<i>SH3PXD2A</i>	–	Eos
	10:105,437,758	4.85	2.76E-06	3.86E-02	8.17E-02	<i>SH3PXD2A</i>	–	Eos
	10:105,437,783	4.70	5.28E-06	3.86E-02	8.11E-02	<i>SH3PXD2A</i>	–	Eos
	10:105,437,694	4.67	6.07E-06	4.00E-02	8.35E-02	<i>SH3PXD2A</i>	–	Eos

rs4763934/ rs55729386/ rs55750092/ rs55816732/ rs55928877/ rs76153631/ rs78376850/ rs55954298^b	12:13,153,099	-4.69/ -4.66	5.52E-06/ 6.39E-06	3.86E-02/ 4.00E-02	-7.52E-02/ -7.46E-02	<i>HEBP1</i>	<i>HTR7P1</i>	Eos
rs1578401	20:17,608,075	-5.32	3.14E-07	7.47E-03	-9.40E-02	<i>RRBP1</i>	–	Eos
	20:17,608,080	-4.79	3.53E-06	3.86E-02	-8.87E-02	<i>RRBP1</i>	–	Eos

^aCpGs and variants shown are those with significant mQTL results (FDR<5.00E-02) in the whole sample. ^bValues in bold characters are for the variant rs55954298 which are in bold character too. Abbreviations: CD4⁺=Naïve CD4⁺ T cells, CpG=Cytosine preceding Guanine sites, Eos=Eosinophils, FDR=False discovery rate, Stat=Statistic.

Table 20. Significant cis-mQTL results including genetic variants associated with both types of sensitization in naïve CD4⁺ T cells and eosinophils

Variants ID ^{a,b}	CpG	Stat	P value	FDR	Beta	CpG's position in gene		Cell type	
						Body	Promoter		
rs115461272	3:33,759,237	-5.76	3.59E-08	1.11E-04	-1.02E-01	<i>CLASP2</i>	–	CD4 ⁺	
	3:33,759,541	-4.84	2.75E-06	5.08E-03	-1.15E-01	<i>CLASP2</i>	–	CD4 ⁺	
	3:33,688,381	4.60	8.08E-06	1.42E-02	2.63E-01	<i>CLASP2</i>	–	CD4 ⁺	
rs4975729	5:893,269	-4.42	1.72E-05	2.75E-02	-3.60E-02	<i>TRIP13</i>	<i>BRD9</i>	CD4 ⁺	
rs6479446/ rs10761191/ rs10682558 / rs4743880	9:95,795,992	12.12/ 12.43/ 12.03	3.79E-25/ 4.77E-26/ 6.96E-25	1.27E-20/ 4.81E-21/ 1.76E-20	2.25E-01/ 2.27E-01/ 2.26E-01	<i>FGD3</i>	–	CD4 ⁺	
		9:95,783,223	7.54/ 7.49	2.16E-12/ 2.19E-12/ 2.86E-12	1.30E-08/ 1.60E-08				1.59E-01/ 1.58E-01/ 1.59E-01
			9:95,783,341	4.46/ 4.45/ 4.53	1.42E-05/ 1.51E-05/ 1.05E-05				2.36E-02/ 2.50E-02/ 1.79E-02
	9:95,795,971	10.17/ 10.28/ 10.03	1.71E-19/ 8.07E-20/ 4.09E-19	2.46E-15/ 1.25E-15/ 5.15E-15	1.83E-01/ 1.84E-01/ 1.83E-01	<i>FGD3</i>	–	CD4 ⁺	
	9:95,783,395	7.82/ 7.95/ 7.93	4.27E-13/ 1.95E-13/ 2.19E-13	2.69E-09/ 1.36E-09/ 1.42E-09	1.22E-01/ 1.23E-01/ 1.25E-01				

9:95,783,204	4.99/ 4.97/ 5.16	1.43E-06/ 1.52E-06/ 6.57E-07	2.72E-03/ 2.87E-03/ 1.38E-03	8.43E-02/ 8.37E-02/ 8.73E-02	<i>FGD3</i>	–	CD4 ⁺
9:95,795,919	11.86/ 11.94/ 11.66	2.20E-24/ 1.32E-24/ 8.76E-24	4.44E-20/ 2.95E-20/ 1.47E-19	2.33E-01/ 2.33E-01/ 2.33E-01	<i>FGD3</i>	–	CD4 ⁺
9:95,783,234	5.36/ 5.38/ 5.25	2.54E-07/ 2.34E-07/ 4.23E-07	5.64E-04/ 5.29E-04/ 9.17E-04	1.13E-01/ 1.13E-01/ 1.12E-01	<i>FGD3</i>	–	CD4 ⁺
9:95,783,388	6.32/ 6.39/ 6.44	2.01E-09/ 1.37E-09/ 1.03E-09	7.11E-06/ 4.92E-06/ 3.84E-06	9.81E-02/ 9.85E-02/ 1.00E-01	<i>FGD3</i>	–	CD4 ⁺
9:95,795,942	9.21/ 9.26/ 9.26	7.94E-17/ 6.02E-17/ 6.05E-17	7.27E-13/ 6.42E-13/ 6.42E-13	1.81E-01/ 1.81E-01/ 1.83E-01	<i>FGD3</i>	–	CD4 ⁺
9:95,796,054	9.22/ 9.24/ 9.02	7.77E-17/ 6.55E-17/ 2.71E-16	7.27E-13/ 6.60E-13/ 2.27E-12	2.04E-01/ 2.03E-01/ 2.02E-01	<i>FGD3</i>	–	CD4 ⁺
9:95,783,192	6.30/ 6.41/ 6.54	2.14E-09/ 1.25E-09/ 6.25E-10	7.45E-06/ 4.60E-06/ 2.49E-06	1.31E-01/ 1.32E-01/ 1.36E-01	<i>FGD3</i>	–	CD4 ⁺
9:95,783,056	4.74/ 4.50/ 4.39	4.27E-06/ 1.20E-05/ 1.90E-05	7.54E-03/ 2.03E-02/ 2.91E-02	8.82E-02/ 8.38E-02/ 8.29E-02	<i>FGD3</i>	–	CD4 ⁺

	9:95,795,959	12.29/ 12.59/ 12.20	1.21E-25/ 1.58E-26/ 2.30E-25	8.15E-21/ 3.19E-21/ 9.26E-21	2.06E-01/ 2.08E-01/ 2.07E-01	<i>FGD3</i>	–	CD4 ⁺
	9:95,794,497	5.82/ 5.71/ 5.65	2.69E-08/ 4.50E-08/ 6.29E-08	8.46E-05/ 1.31E-04/ 1.67E-04	1.25E-01/ 1.23E-01/ 1.23E-01	<i>FGD3</i>	–	CD4 ⁺
rs10761191	9:95,140,430	-4.26	3.28E-05	4.80E-02	-1.53E-01	<i>CENPP</i>	–	CD4 ⁺
rs148593533	17:30,846,785	-4.54	1.03E-05	1.78E-02	-1.09E-01	<i>MYO1D</i>	–	CD4 ⁺
	17:31,255,101	-4.42	1.73E-05	2.75E-02	-1.70E-01	<i>TMEM98</i>	–	CD4 ⁺
rs9896435	17:30,334,554	-4.44	1.58E-05	2.58E-02	-9.21E-02	–	<i>LRRC37B</i>	CD4 ⁺
	17:29,036,528	-4.43	1.61E-05	2.62E-02	-1.31E-01	–	<i>SUZ12P1</i>	CD4 ⁺
	17:30,186,284	-4.36	2.19E-05	3.26E-02	-8.64E-02	<i>COPRS</i>	–	CD4 ⁺
	17:30,593,834	-4.27	3.12E-05	4.59E-02	-1.24E-01	<i>RHBDL3</i>	–	CD4 ⁺
rs116905394/ rs117022993/ rs62505091/ rs115793380/ rs11774021/ rs62512006/ rs62512007	8:58,907,473	-4.96/ -4.49	1.65E-06/ 1.29E-05	5.56E-03/ 2.61E-02	-9.45E-02/ -6.97E-02	<i>FAM110B</i>	–	Eos
rs115793380/ rs11774021/ rs62512006/ rs62512007	8:60,030,763	-5.24	4.62E-07	1.70E-03	-7.53E-02	<i>TOX</i>	–	Eos
rs28396474	8:9,008,158	-4.39	1.96E-05	3.82E-02	-2.22E-02	<i>PPP1R3B</i>	–	Eos

rs6479446/ rs10761191/ rs10682558/ rs4743880	9:95,140,430	-4.52/ -4.45/ -4.49	1.16E-05/ 1.56E-05/ 1.28E-05	2.50E-02/ 3.11E-02/ 2.61E-02	-1.66E-01/ -1.63E-01/ -1.65E-01	<i>CENPP</i>	–	Eos
	9:95,709,345	-5.22/ -5.45/ -5.52	5.15E-07/ 1.69E-07/ 1.23E-07	1.86E-03/ 7.47E-04/ 5.55E-04	-2.38E-01/ -2.46E-01/ -2.50E-01	–	<i>FGD3</i>	Eos
	9:95,783,204	6.51/ 6.47/ 6.76	7.96E-10/ 9.81E-10/ 2.05E-10	4.53E-06/ 5.42E-06/ 1.20E-06	1.25E-01/ 1.24E-01/ 1.29E-01	<i>FGD3</i>	–	Eos
	9:95,783,223	4.63/ 4.54/ 4.91	7.13E-06/ 1.08E-05/ 2.10E-06	1.75E-02/ 2.38E-02/ 6.35E-03	7.45E-02/ 7.27E-02/ 7.85E-02	<i>FGD3</i>	–	Eos
	9:95,783,234	4.58/ 4.60/ 4.92	8.87E-06/ 8.32E-06/ 1.98E-06	1.98E-02/ 1.88E-02/ 6.15E-03	9.40E-02/ 9.37E-02/ 1.00E-01	<i>FGD3</i>	–	Eos
	9:95,794,428	4.53/ 4.64/ 4.61	1.11E-05/ 6.85E-06/ 7.84E-06	2.43E-02/ 1.70E-02/ 1.83E-02	1.03E-01/ 1.05E-01/ 1.05E-01	<i>FGD3</i>	–	Eos
	9:95,795,919	10.51/ 10.63/ 10.52	2.89E-20/ 1.35E-20/ 2.66E-20	2.61E-16/ 1.42E-16/ 2.52E-16	2.37E-01/ 2.37E-01/ 2.37E-01	<i>FGD3</i>	–	Eos
	9:95,795,942	15.40/ 15.64/ 15.51	3.30E-34/ 6.97E-35/ 1.65E-34	6.57E-30/ 4.62E-30/ 4.69E-30	3.81E-01/ 3.81E-01/ 3.82E-01	<i>FGD3</i>	–	Eos

	9:95,795,959	16.00/ 16.04/ 15.47	6.89E-36/ 5.33E-36/ 2.16E-34	6.85E-31/ 6.85E-31/ 4.78E-30	3.73E-01/ 3.71E-01/ 3.68E-01	<i>FGD3</i>	–	Eos
	9:95,795,971	15.51/ 15.56/ 14.90	1.63E-34/ 1.19E-34/ 8.95E-33	4.69E-30/ 4.69E-30/ 1.48E-28	3.69E-01/ 3.67E-01/ 3.62E-01	<i>FGD3</i>	–	Eos
	9:95,795,992	14.05/ 14.33/ 13.97	2.33E-30/ 3.64E-31/ 4.07E-30	3.32E-26/ 5.57E-27/ 5.07E-26	3.13E-01/ 3.14E-01/ 3.12E-01	<i>FGD3</i>	–	Eos
	9:95,796,054	10.65/ 10.73/ 10.20	1.19E-20/ 6.80E-21/ 2.06E-19	1.31E-16/ 7.96E-17/ 1.71E-15	2.60E-01/ 2.60E-01/ 2.53E-01	<i>FGD3</i>	–	Eos
	17:29,036,499	-5.61	8.09E-08	3.84E-04	-1.21E-01	–	<i>SUZ12P1</i>	Eos
rs9896435	17:29,058,569	-4.49	1.33E-05	2.67E-02	-7.56E-02	<i>SUZ12P1</i>	–	Eos
	17:29,422,776	-5.32	3.26E-07	1.33E-03	-1.09E-01	<i>NF1</i>	<i>MIR4733</i>	Eos
	17:30,592,521	-4.41	1.83E-05	3.60E-02	-1.72E-01	–	<i>RHBDL3</i>	Eos

^aCpGs and variants shown are those with significant mQTL results (FDR<5.00E-02) in the whole sample. ^bValues in bold characters are for the variant identified in bold character too. Abbreviations: CD4⁺=Naïve CD4⁺ T cells, CpG=Cytosine preceding Guanine sites, Eos=Eosinophils, FDR=False discovery rate, Stat=Statistic.

Table 21. Correlation between gene expression and methylation levels in CD4⁺ T cells in indoor and both types of sensitization

Gene	CpG ^{a,b}	Coeff	SE	OR	rPearson	FDR Corr
Indoor sensitization						
<i>MIR3939</i>	6:167,411,355	-2.06	2.46E-01	1.27E-01	-5.45E-01	9.47E-16
	6:167,349,958	-4.83E-01	6.56E-02	6.17E-01	-4.45E-01	2.13E-12
	6:167,350,137	-6.03E-01	7.45E-02	5.47E-01	-4.43E-01	1.03E-14
	6:167,365,308	-1.42	1.26E-01	2.42E-01	-5.09E-01	1.01E-27
	6:167,365,340	-9.07E-01	1.20E-01	4.04E-01	-4.18E-01	5.91E-13
	6:167,365,386	-9.52E-01	1.01E-01	3.86E-01	-5.38E-01	1.09E-19
	6:167,365,735	-7.23E-01	1.13E-01	4.85E-01	-5.04E-01	1.37E-09
	6:167,365,847	-9.14E-01	8.85E-02	4.01E-01	-5.34E-01	1.75E-23
	6:167,365,855	-9.11E-01	6.98E-02	4.02E-01	-5.79E-01	5.25E-37
<i>RNASET2</i>	6:167,365,863	-8.50E-01	6.54E-02	4.27E-01	-6.13E-01	8.55E-37
	6:167,365,877	-8.68E-01	6.27E-02	4.20E-01	-6.36E-01	4.53E-41
	6:167,365,888	-8.23E-01	6.25E-02	4.39E-01	-5.96E-01	1.43E-37
	6:167,365,903	-7.80E-01	7.15E-02	4.58E-01	-5.65E-01	3.89E-26
	6:167,365,909	-7.07E-01	7.23E-02	4.93E-01	-5.70E-01	3.70E-21
	6:167,366,196	-7.46E-01	8.64E-02	4.74E-01	-3.50E-01	1.33E-16
	6:167,366,232	-6.65E-01	9.75E-02	5.14E-01	-2.63E-01	9.13E-11
	6:167,366,359	-4.13E-01	1.01E-01	6.62E-01	-1.86E-01	2.02E-04

	6:167,366,501	-7.33E-01	7.88E-02	4.81E-01	-4.64E-01	3.65E-19
	6:167,366,511	-8.79E-01	8.50E-02	4.15E-01	-4.92E-01	1.54E-23
	6:167,368,630	-5.82E-01	4.79E-02	5.59E-01	-6.29E-01	3.72E-32
	6:167,368,710	-4.67E-01	5.98E-02	6.27E-01	-5.51E-01	7.75E-14
	6:167,371,091*	-7.52E-01	6.44E-02	4.72E-01	-6.31E-01	9.01E-30
	6:167,371,100*	-6.82E-01	6.28E-02	5.06E-01	-5.93E-01	7.34E-26
	6:167,371,869*	-5.95E-01	1.03E-01	5.52E-01	-4.66E-01	5.49E-08
Both types of sensitization						
<i>CENPP</i>	9:95,140,430	-5.51E-01	6.33E-02	5.76E-01	-5.43E-01	1.55E-16
	9:95,783,223	8.92E-02	2.76E-02	1.09	-3.28E-02	7.52E-03
<i>FGD3</i>	9:95,783,234	8.66E-02	3.04E-02	1.09	-1.24E-01	1.98E-02
	9:95,783,395	1.13E-01	3.47E-02	1.12	1.22E-01	7.31E-03
	9:95,796,054	7.83E-02	2.50E-02	1.08	2.59E-01	9.63E-03

^aGenomic position of CpGs according to the human genome reference 19. ^bCpGs marked with * are located in the gene's promoter while the others are in the gene's body. Abbreviations: Coeff=Coefficient, Corr=Correlation, CpGs=Cytosine preceding Guanine sites, FDR=False discovery rate, OR=Odd ratio, SE=Standard error.

Table 22. Genes targets of *MIR3939* and their biological functions

Gene name	Gene symbol	Cytoband	Protein class	Function	Association to disease	References
ETS transcription factor ELK4	<i>ELK4</i>	1q32.1	Winged helix/forkhead transcription factor	Regulation of stress response in macrophages to external infection. Involved in M2 polarization of macrophages in gastric cancer	–	1, 2
Collagen type VIII alpha 1 chain	<i>COL8A1</i>	3q12.1	Extracellular matrix structural protein	Involved in extracellular matrix organization, cell adhesion, epithelial cell proliferation and angiogenesis	Chronic obstructive pulmonary disease, atopic eczema	3, 4
Salt inducible kinase 1	<i>SIK1</i>	21q22.3	Non-receptor Serine/threonine-protein kinase	Negative regulation of CREB transcription factor activity, involved with other SIK in inhibition of M2 polarisation of macrophage	Eczema (variant), allergic disease (variant)	5-7

Cryptochrome circadian regulator 2	<i>CRY2</i>	11p11.2	DNA photolyase	Negative regulation of transcription by RNA polymerase II, Circadian regulation of gene expression, Regulation of NF- κ B pathway through interaction with the adenylyl cyclase	Lower expression in well-controlled asthmatics with night symptoms. Low expression in the asthmatic group with high proportion of allergic rhinitis and atopic dermatitis and high expression in the asthmatic group with high proportion of atopic dermatitis and low proportion of allergic rhinitis	8, 9
Protein arginine methyltransferase 8	<i>PRMT8</i>	12p13.32	Protein modifying enzyme	Protein methylation, peptidyl-arginine methylation, regulation of protein binding, involved in formation of asymmetric dimethylarginine which is involved in asthma	Pneumonia (variant)	10, 11
LHFPL tetraspan subfamily member 3	<i>LHFPL3</i>	7q22.2-7q22.3	Scaffold/adaptor protein	Sensory perception of sound	Asthma exacerbation despite use of inhaled corticosteroids (variant)	12

Plakophilin 1	<i>PKP1</i>	1q32.1	Intermediate filament binding protein	Facilitate the formation of intermediate filaments, component of desmosome, stabilize desmosomal proteins at the membrane and involved in promoting and stabilizing cell adhesion	Asthma (variant)	13, 14
Granulin precursor	<i>GRN</i>	17q21.31	–	Positive regulation of cell migration, of angiogenesis, of epithelial cell proliferation. Regulation of inflammatory response. Negative regulation of neutrophil activation. Promote differentiation of CD4 ⁺ T cells into Treg cells	Neutrophilic asthma	15, 16
Inka box actin regulator 2	<i>INKA2</i>	1p13.2	–	Involved in regulation of catalytic activity, Inhibitor of the serine/threonine-protein kinase PAK4 which promotes cell growth, proliferation, migration and metastasis in certain types of cancer	Down-regulated in lung fibroblasts in idiopathic pulmonary fibrosis	17
HPS4 biogenesis of lysosomal organelles complex 3 subunit 2	<i>HPS4</i>	22q12.1	–	Protein targeting, vesicle-mediated transport, melanocyte differentiation, involved in regulations of biogenesis and/or function of the lung lamellar body	Development of pulmonary fibrosis in Hermansky-Pudlak syndrome with HPS4 mutation	18, 19

References:

1: Xie L. *BMC Genomics*. 2014;15:301. **2:** Zheng L *et al. J Transl Med*. 2021;19:342. **3:** Strafella C, *et al. Biomed Res Int*. 2019;2019:3457898. **4:** Cosentino J, *et al. Nat Genet*. 2023;55(5):787-795. **5:** Darling NJ, *et al. Biochem J*. 2017;474(4):521-537. **6:** Ferreira MA, *et al. Nat Genet*. 2017;49(12):1752-1757. **7:** Kichaev G, *et al. Am J Hum Genet*. 2019;104(1):65-75. **8:** Narasimamurthy R, *et al. Proc Natl Acad Sci U S A*. 2012;109(31):12662-12667. **9:** Chen H-C, *et al. J Asthma Allergy*. 2021;14:371-380. **10:** Zakrzewicz D, *et al. Int J Mol Sci*. 2012;13(10):12383-12400. **11:** Campos AI, *et al. Twin Res Hum Genet*. 2021;24(3):145-154. **12:** Hernandez-Pacheco N, *et al. Clin Exp Allergy*. 2019;49(6):789-798. **13:** Hofmann I, *et al. J Cell Sci*. 2000;113(Pt 13):2471-2483. **14:** Hatzfeld M. *Biochim Biophys Acta*. 2007;1773(1):69-77. **15:** Pogonowska M, *et al. Cent Eur J Immunol*. 2019;44(1):97-101. **16:** Choi J-P, *et al. Allergy*. 2020;75(5):1133-1145. **17:** Sheu C-C, *et al. Int J Mol Med*. 2019;43(4):1643-1656. **18:** Wei ML. *Pigment Cell Res*. 2006;19(1):19-42. **19:** Yokoyama T and Gochuico BR. *Eur Respir Rev*. 2021;30(159):200193.