

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

Recherche de déterminants génomiques impliqués dans l'hypertension, sur le chromosome X, chez des familles du Saguenay-Lac-Saint-Jean

par
Audrey Noël

Programme de bio-informatique
Faculté de médecine

Mémoire présenté à la Faculté de médecine
en vue de l'obtention du grade de
Maître ès Sciences (M. Sc.)
en bio-informatique

Décembre 2007
© Audrey Noël, 2007



Université de Montréal
Faculté de médecine

Ce mémoire intitulé:
Recherche de déterminants génomiques impliqués dans l'hypertension, sur le chromosome
X, chez des familles du Saguenay-Lac-Saint-Jean

Présenté par :
Audrey Noël

a été évalué par un jury composé des personnes suivantes :

Dr Daniel Sinnett
président-rapporteur

Dr Pavel Hamet
directeur de recherche

Dr Ettore Merlo
codirecteur

Dre Louise Pilote
membre du jury

Résumé Français

L'hypertension, un désordre complexe et multifactoriel, est un des principaux facteurs de risque pour le développement de maladies cardiaques et d'accidents vasculaires cérébraux. À travers le monde, cette maladie multigénique affecte environ 1 milliard de personnes et au Canada, seulement 13% des sujets souffrant d'hypertension sont contrôlés adéquatement. Un effet du sexe sur la pression artérielle est déjà fondé dans la littérature ; néanmoins, bien que l'importance du chromosome X, dans cet effet et dans l'hypertension, soit bien établie, les mécanismes de son rôle restent encore évasifs. De plus, ses caractéristiques génétiques particulières rendent les études génétiques compliquées.

Des tests de liaison et d'association, à l'aide de *single nucleotide polymorphisms*, ont été utilisés dans le but d'investiguer, sur le chromosome X, les déterminants génomiques de l'hypertension chez des familles du Saguenay-Lac-Saint-Jean. Plusieurs associations avec des SNPs connus ont été découvertes lors de ce projet. Le résultat le plus intéressant concerne le rs2116366, localisé dans le gène *fatty acid amide hydrolase 2*. Ce SNP a montré une association spécifique aux femmes avec la pression artérielle prise durant 24 heures, avec une différence de 7,4 mmHg entre les deux groupes homozygotes de femmes. D'après les résultats, ce SNP pourrait contribuer à la régulation de la pression artérielle via une participation de FAAH2.

Ces résultats confirment le rôle du chromosome X dans l'hypertension. Nos découvertes sur ce chromosome aideront à trouver les déterminants génétiques, spécifiques au sexe, impliqués dans la pathogenèse de l'hypertension et son traitement.

Mots clés: chromosome X, hypertension, liaison, association, SNPs, gènes, déséquilibre de liaison

Résumé Anglais

Hypertension, a complex and multifactorial disorder, is one of the major risk factors for the development of cardiovascular diseases and cerebrovascular disease. This polygenic disease affects around a billion of people in the world and in Canada, only 13% of the people suffering from hypertension are controlled properly. A clear gender-specific effect on blood pressure is reported in the literature; nevertheless, although the importance of the X chromosome in this effect and in hypertension is established, its mechanisms are still elusive. Moreover, the particular genetic features of the X chromosome make gene finding studies complicated.

Joint model-free linkage and family based association tests with single nucleotide polymorphisms were used to investigate the genomic determinants underlying the phenotypes pertinent to hypertension among families from the Saguenay-Lac-Saint-Jean region. Many associations with known SNPs were obtained from this project. The more interesting result concerned the rs2116366 which is located in the fatty acid amide hydrolase 2 gene. This SNP showed a female specific association with a 24 hours systolic blood pressure phenotype, with a 7,4 mmHg difference between homozygote female groups. According to the results, this SNP may contribute to blood pressure regulation by a participation of FAAH2.

Those results confirm the implication of the chromosome X in hypertension. Our finding on this chromosome may help to find sex-specific genetic determinants involved in the pathogenesis of hypertension and his treatment.

Key words: chromosome X, hypertension, linkage, association, SNPs, genes, linkage disequilibrium

Tables des Matières

| | |
|--|-----------|
| RÉSUMÉ FRANÇAIS | I |
| RÉSUMÉ ANGLAIS..... | II |
| TABLES DES MATIERES | III |
| LISTE DES TABLEAUX | V |
| LISTE DES FIGURES..... | VI |
| LISTE DES ABRÉVIATIONS..... | VII |
| DEDICACE..... | IX |
| REMERCIEMENTS..... | X |
| INTRODUCTION | I |
| CHAPITRE 1 - DESORDRES GENETIQUES | 2 |
| 1.1 Désordre chromosomique | 2 |
| 1.2 Désordre monogénique | 3 |
| 1.3 Désordre multifactoriel | 4 |
| CHAPITRE 2 - HYPERTENSION..... | 5 |
| 2.1 Pression artérielle..... | 5 |
| 2.1.1 Systèmes régulateurs de la pression artérielle..... | 5 |
| 2.1.2 Distribution de la pression artérielle | 9 |
| 2.2 Hypertension artérielle..... | 10 |
| 2.2.1 Hypertension secondaire..... | 10 |
| 2.2.2 Hypertension essentielle | 11 |
| 2.3 Composants métaboliques de l'hypertension artérielle..... | 13 |
| 2.3.1 Dyslipidémie..... | 13 |
| 2.3.2 Obésité..... | 14 |
| 2.3.3 Syndrome métabolique | 15 |
| CHAPITRE 3 - CHROMOSOME X | 16 |
| 3.1 Les caractéristiques du chromosome X | 16 |
| 3.2 Inactivation du chromosome X | 17 |
| 3.3 Recombinaison | 18 |
| 3.4 Syndromes résultant d'un nombre de chromosomes X anormaux..... | 20 |
| 3.4.1 Syndrome de Turner | 20 |
| 3.4.2 Syndrome de Klinefelter | 20 |
| CHAPITRE 4 - APPROCHE GENETIQUE..... | 21 |
| 4.1 Population du Saguenay-Lac-Saint-Jean du Québec | 21 |
| 4.2 Procédure | 22 |
| 4.2.1 Approches..... | 22 |
| 4.2.2 Analyses génétiques..... | 23 |
| 4.2.3 Impact de l'hémizygoté de l'homme sur les études génétiques..... | 26 |
| 4.3 Marqueurs génétiques | 27 |
| 4.3.1 SNPs | 27 |
| 4.3.2 Blocs d'haplotypes..... | 29 |
| 4.3.3 Correction pour test multiple | 29 |
| MÉTHODOLOGIE | 31 |
| CHAPITRE 6 – OUTILS ET DONNEES UTILISES..... | 32 |
| 6.1 Sujets | 32 |
| 6.2 Phénotypes..... | 34 |

| | |
|--|-----------|
| 6.3 Génotypes | 36 |
| 6.3.1 Construction des blocs d'haplotypes..... | 37 |
| 6.4 Tests statistiques | 37 |
| 6.5 Impact des SNPs et conservation de séquences entre espèces | 40 |
| 6.5.1 Prédiction de la fonction des SNPs..... | 40 |
| 6.5.1 Conservation des séquences entre différentes espèces..... | 41 |
| 6.6 Bio-informatique..... | 42 |
| RÉSULTATS..... | 44 |
| CHAPITRE 7 – RESULTATS SUR LE CHROMOSOME X..... | 45 |
| 7.1 Programme Merlin..... | 45 |
| 7.2 Programme XRC-TDT | 46 |
| 7.3 Relation entre le génotype et le phénotype | 49 |
| 7.4 Déséquilibre de liaison..... | 52 |
| 7.5 Prédiction de l'impact des SNPs..... | 55 |
| 7.6 Conservation des séquences entre les espèces | 57 |
| DISCUSSION | 65 |
| CHAPITRE 8 – SNPs ET GENES D'INTERET DECOUVERTS SUR LE CHROMOSOME X..... | 66 |
| 8.1 SNPs et gènes d'intérêt | 66 |
| 8.2 Impact d'un SNP sur le phénotype | 75 |
| 8.3 Conservation de séquences entre les espèces..... | 77 |
| 8.4 Associations différentes entre les sexes | 80 |
| 8.5 Limites du projet de recherche concernant l'étude de la population du Saguenay-Lac-Saint-Jean .. | 81 |
| CONCLUSION..... | 83 |
| RÉFÉRENCES..... | 86 |

Liste des Tableaux

| | | |
|---------------|---|----|
| TABLEAU I | TYPES ET FRÉQUENCES DES DÉSORDRES GÉNÉTIQUES, ADAPTÉ DE RIMOIN ET AL., 2002 ¹ | 2 |
| TABLEAU II | LES FORMES MENDÉLIENNES D'HYPERTENSION, ADAPTÉ D'HAMET ET AL., 1998 ³⁴ | 12 |
| TABLEAU III | SEUILS DES PHÉNOTYPES DÉFINISSANT LE SYNDROME MÉTABOLIQUE, ADAPTÉ DE KHUNTI AND DAVIES ⁵¹ | 15 |
| TABLEAU IV | CARACTÉRISTIQUES BIOLOGIQUES, CLASSÉES PAR LE SEXE, DES 468 SUJETS GÉNOTYPÉS..... | 33 |
| TABLEAU V | DESCRIPTION DES DIFFÉRENTS LOGICIELS UTILISÉS LORS DU PROJET..... | 39 |
| TABLEAU VI | MILLEURES ASSOCIATIONS OBTENUES POUR LE CHROMOSOME X AVEC LE PROGRAMME XRC-TDT | 47 |
| TABLEAU VII | ASSOCIATIONS ENTRE LE RS1949393 ET LA PAS DURANT 24 HEURES POUR DIFFÉRENTS SOUS-GROUPES À L'AIDE DE XRC-TDT | 48 |
| TABLEAU VIII | ASSOCIATIONS ENTRE LE RS407890 ET LA FRÉQUENCE CARDIAQUE POUR DIFFÉRENTS SOUS-GROUPES À L'AIDE DE XRC-TDT | 48 |
| TABLEAU IX | ASSOCIATIONS ENTRE LE RS2116366 ET LA PAS DURANT 24 HEURES POUR DIFFÉRENTS SOUS-GROUPES À L'AIDE DE XRC-TDT | 48 |
| TABLEAU X | ASSOCIATIONS ENTRE LE RS2116366 ET LA PAS DURANT LA NUIT POUR DIFFÉRENTS SOUS-GROUPES À L'AIDE DE XRC-TDT | 49 |
| TABLEAU XI | ASSOCIATIONS ENTRE LE RS727562 ET LA PAS DURANT LE JOUR POUR DIFFÉRENTS SOUS-GROUPES À L'AIDE DE XRC-TDT | 49 |
| TABLEAU XII | COMPARAISON DES MOYENNES DE PAS DURANT 24 HEURES SELON LES DIFFÉRENTS GÉNOTYPES DU RS2116366 | 51 |
| TABLEAU XIII | COMPARAISON DES MOYENNES DE PAS DURANT LA NUIT SELON LES DIFFÉRENTS GÉNOTYPES DU RS2116366 | 51 |
| TABLEAU XIV | PRÉVALENCE DU STATUT D'HYPERTENSION CHEZ LES FEMMES SELON LES GÉNOTYPES DU RS2116366..... | 52 |
| TABLEAU XV | COMPARAISON DES MOYENNES DE LA FRÉQUENCE CARDIAQUE SELON LES DIFFÉRENTS GÉNOTYPES DU RS407890 | 52 |
| TABLEAU XVI | NOMBRE DE SNPs DE LA POPULATION CAUCASIENNE EN LD AVEC LES SNPs D'INTÉRÊT | 54 |
| TABLEAU XVII | BLOCS D'HAPLOTYPES ET RÉGIONS DE LD COMPLET ASSOCIÉS AU SNPs IDENTIFIÉS | 54 |
| TABLEAU XVIII | PRÉDICTION D'IMPACT DES SNPs IDENTIFIÉS AVEC F-SNP | 55 |
| TABLEAU XIX | PRÉDICTION D'IMPACT DES SNPs IDENTIFIÉS AVEC SNPINSPECTOR | 56 |

Liste des Figures

| | | |
|-----------|--|----|
| FIGURE 1 | INTERACTIONS ENTRE LES GÈNES ET L'ENVIRONNEMENT. TIRÉ DE SEDA ET AL., 2005 ⁵ | 4 |
| FIGURE 2 | RÔLE DE L'ANGIOTENSINE II DANS LA RÉGULATION DE LA PRESSION ARTÉRIELLE | 6 |
| FIGURE 3 | MÉCANISME DU SYSTÈME RÉNINE-ANGIOTENSINE | 7 |
| FIGURE 4 | RÔLE D'ACE2 DANS LE SYSTÈME RÉNINE-ANGIOTENSINE, TIRÉ DE YAGIL ET YAGIL, 2003 ¹³ | 8 |
| FIGURE 5 | PROFIL DE LA PRESSION ARTÉRIELLE SYSTOLIQUE ET DIASTOLIQUE SELON LE SEXE, L'ÂGE ET DIFFÉRENTES ETHNIES, TIRÉ DE BURT ET AL. 1995 ²⁶ | 9 |
| FIGURE 6 | PROFIL DE LA PRESSION ARTÉRIELLE SYSTOLIQUE ET DIASTOLIQUE, SELON LE SEXE ET L'ÂGE, BASÉ SUR DEUX ÉTUDES D'ÉCHANTILLON REPRÉSENTATIF DES ÉTATS-UNIS (NHANES) ET DU CANADA (CHHS), TIRÉ DE JOFFRES ET AL., 2001 ²⁵ | 10 |
| FIGURE 7 | STRUCTURE DU CHROMOSOME X ET LE PROFIL D'INACTIVATION DES GÈNES, TIRÉ DE SPATZ ET AL., 2004 ⁶⁵ | 18 |
| FIGURE 8 | REPRÉSENTATION SCHÉMATIQUE DES HOMOLOGIES MAJEURES ENTRE LES CHROMOSOMES SEXUELS DE L'HUMAIN, TIRÉ DE ROSS ET AL., 2005 ⁵³ | 19 |
| FIGURE 9 | LOCALISATION DE LA POPULATION DU SAGUENAY-LAC-SAINT-JEAN (EN ROUGE) DE LA RÉGION DE QUÉBEC, TIRÉ DE HTTP://EN.WIKIPEDIA.ORG | 21 |
| FIGURE 10 | REPRÉSENTATION SCHÉMATIQUE D'UN SNP..... | 27 |
| FIGURE 11 | RÉSUMÉ DE LA MÉTHODOLOGIE UTILISÉE | 40 |
| FIGURE 12 | ARBRE PHYLOGÉNÉTIQUE DES 28 ESPÈCES DISPONIBLES POUR LES ANALYSES AVEC MULTIZ ET PHASTCONS | 42 |
| FIGURE 13 | CARTE DE DÉSÉQUILIBRE DE LIAISON, OBTENUE AVEC HAPLOVIEW, DANS LA POPULATION DU SLSJ POUR LE RS4911871 ASSOCIÉ À L'INDICE HEATHER..... | 53 |
| FIGURE 14 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE RS10521413 .. | 58 |
| FIGURE 15 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE RS569959..... | 58 |
| FIGURE 16 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE RS233575..... | 59 |
| FIGURE 17 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE RS950463..... | 59 |
| FIGURE 18 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE RS1335614..... | 60 |
| FIGURE 19 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE RS4911871..... | 60 |
| FIGURE 20 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE GÈNE ACE2 .. | 61 |
| FIGURE 21 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE GÈNE RPGR...62 | |
| FIGURE 22 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE GÈNE HTR2C | 62 |
| FIGURE 23 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE GÈNE IL1RAPL1 | 63 |
| FIGURE 24 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE GÈNE FAAH2..... | 63 |
| FIGURE 25 | ALIGNEMENTS ET CONSERVATION DE SÉQUENCES CHEZ 28 ESPÈCES POUR LE GÈNE SPIN2...64 | |
| FIGURE 26 | LIEN POSSIBLE ENTRE OTC ET L'HYPERTENSION | 67 |
| FIGURE 27 | REPRÉSENTATION DE LA SYNTÉNIE, CHEZ LE RAT, LA SOURIS ET L'HUMAIN, POUR LA RÉGION ENTOURANT LE GÈNE KLHL4 (RECTANGLES ROUGES), TIRÉE DE WWW.NCBI.NLM.NIH.GOV | 69 |
| FIGURE 28 | GÈNE FAAH2..... | 71 |
| FIGURE 29 | RÔLE DE L'ENZYME FAAH..... | 72 |
| FIGURE 30 | RÔLE DE LA SÉROTONINE ET RÔLE POSSIBLE DU RÉCEPTEUR HTR2C DANS L'HYPERTENSION; RÉCEPTEUR :R..... | 73 |

Liste des abréviations

- ❖ ACE = *angiotensin converting enzyme*
- ❖ Ang = *angiotensine*
- ❖ Arg = *arginine*
- ❖ ARN = *acide ribonucléique*
- ❖ cGMP = *cyclic guanosine monophosphat*
- ❖ CHD = *maladie coronarienne*
- ❖ CHHS = *Canadian Heart Health Surveys*
- ❖ cM = *centimorgans*
- ❖ ECG = *électrocardiogramme*
- ❖ FAAH2 = *fatty acid amide hydrolase 2*
- ❖ GEE = *generalized estimating equations*
- ❖ GTP = *guanosine triphosphate*
- ❖ GGT = *gamma glutamyl transferase*
- ❖ HDL = *high density lipoprotein*
- ❖ HT = *hypertendu*
- ❖ HTR2C = *5-hydroxytryptamine (serotonin) receptor 2C*
- ❖ HWE = *Hardy-Weinberg equilibrium*
- ❖ IL1RAPL1 = *interleukin 1 receptor accessory protein-like 1*
- ❖ KLHL4 = *kelch-like 4*
- ❖ LD = *linkage disequilibrium*
- ❖ LVH = *left ventricular hypertrophy*
- ❖ LVMI = *left ventricular mass index*
- ❖ Merlin = *multipoint engine for rapid likelihood inference*
- ❖ mmHg = *millimètres de mercure*
- ❖ NHANES III = *the Third National Health and Nutrition Examination Surveys*
- ❖ NO = *nitric oxyde*
- ❖ NOb = *non obèse*
- ❖ NT = *normotendu*
- ❖ Ob = *Obèse*
- ❖ OTC = *ornithine carbamoyltransferase*
- ❖ PAD = *pression artérielle diastolique*
- ❖ PAR = *pseudoautosomal region*
- ❖ PAS = *pression artérielle systolique*
- ❖ Pb = *paires de bases*
- ❖ QTL = *locus à trait quantitatif*
- ❖ RPGR = *retinitis pigmentosa GTPase regulator*
- ❖ SHR = *spontaneously hypertensive rat*
- ❖ SLSJ = *Saguenay-Lac-Saint-Jean*
- ❖ SM = *syndrome métabolique*
- ❖ SNP = *single nucleotide polymorphism*
- ❖ SPIN2 = *spindlin family member 2*
- ❖ SQL = *structured query language*
- ❖ TFBS = *transcription factor binding sites*
- ❖ TGIF2LX = *TGFB-induced factor 2-like, X-linked*

- ❖ WKY = rats normotendus Wistar-Kyoto
- ❖ XIC = *X inactivation center*
- ❖ XIST = *X inactivation-specific transcript*
- ❖ XRC-TDT = *X-linked reconstruction - combination transmission/disequilibrium test*

Dédicace

Que sait du désert celui qui ne regarde qu'un grain de sable?
Erik Orsenna

À mes parents, ma sœur et mon conjoint. Je vous aime.

Remerciements

Tout d'abord, je tiens à remercier mon conjoint Dominique pour son appui et sa compréhension ainsi que mon père Denis, ma sœur Jessie, ma tante Nicole et mon oncle Réal pour leurs encouragements et leur présence. Je voudrais également remercier spécialement ma mère Solange pour son soutien sans borne, son aide précieuse et surtout pour son grand intérêt porté à mon cheminement. Vous avez tous été une source d'énergie pour moi.

Je souhaiterais remercier Dr. Pavel Hamet et Dr. Ettore Merlo, mes deux directeurs de maîtrise, qui ont bien voulu me donner ma chance en m'acceptant au sein de leur équipe réputée. Le support et le soutien qu'ils m'ont démontrés tout au long de ma maîtrise furent énormément appréciés. J'aimerais exprimer ma reconnaissance spécialement au Dr. Hamet, pour sa patience, sa générosité, le partage de son savoir de même que pour m'avoir donné le droit à l'erreur. Il a été un modèle pour moi et m'a grandement permis d'améliorer mes capacités de chercheur.

Je voudrais aussi dire merci à mes collègues de travail, Pierre-Luc Brunelle, Alexandru Gurau, François Gauthier et Louis-Philippe Lemieux-Perreault pour avoir été des compagnons de travail extraordinaires et pour avoir répondu à mes millions de questions. Grâce à eux, mes connaissances en informatique se sont grandement améliorées. Pour terminer, je tiens à remercier les secrétaires France Robin, Érika Fisseler, Francine Allie, Francine Philibert et Diane Legault, pour leur dévouement et tous les services rendus ainsi que tous les membres du jury pour avoir accepté d'évaluer mon mémoire.

Introduction

Chapitre 1 - Désordres génétiques

Les mutations génétiques sont à l'origine de la variation du pool génétique et elles peuvent être séparées en mutations constitutives ou en mutations somatiques. Une mutation constitutive ou de cellules germinales représente une mutation héritable, pouvant être passée à la prochaine génération; à l'opposé, une mutation somatique est une mutation acquise puisqu'elle survient uniquement dans les cellules non germinales. Dans bien des cas, ces mutations peuvent mener à une pathologie dont il est possible d'évaluer le niveau d'expression chez les individus par la notion de pénétrance. Cette dernière se traduit comme étant la portion d'individus ayant un certain génotype et qui exprime le phénotype en question¹; par exemple, si la maladie est exprimée chez tous les individus ayant un génotype donné, la pénétrance sera de 100%.

De leur côté, les gènes représentent les déterminants principaux de la variation humaine et à partir de la proportion jouée par ceux-ci, les désordres génétiques peuvent être séparés en trois groupes : désordres chromosomiques, désordres monogéniques et désordres multifactoriels.

1.1 Désordre chromosomique

Il résulte d'un désordre chromosomique lorsqu'il y a présence d'une altération visible dans le nombre ou la structure des chromosomes². Le phénotype est alors causé par un excès ou une déficience des gènes d'un chromosome entier ou d'un segment chromosomique. Un exemple très connu est la trisomie 21, attribuable à une copie en trop du chromosome 21. Parmi les diverses catégories de désordres génétiques, les désordres chromosomiques sont les moins fréquents (Tableau I).

Tableau I Types et fréquences des désordres génétiques, adapté de Rimoin *et al.*, 2002¹

| Type de désordres | Fréquence avant 25 ans | Fréquence après 25 ans | Fréquence durant la vie |
|------------------------|------------------------|------------------------|-------------------------|
| Chromosomiques | 1,8/1000 | 2/1000 | 3,8/1000 |
| Monogéniques | 3,6/1000 | 16,4/1000 | 20/1000 |
| Multifactoriels | 46,4/1000 | 600/1000 | 646,4/1000 |
| De cellules somatiques | ----- | 240/1000 | 240/1000 |

1.2 Désordre monogénique

Un désordre monogénique, ou maladie à transmission mendélienne, est déterminé par les allèles situés sur un simple locus. Elle est donc causée par des gènes mutants individuels¹. La mutation peut survenir sur un ou les deux allèles d'un gène qui est positionné sur un autosome ou un chromosome sexuel. Ce type de désordre est caractérisé par un profil de transmission observable à l'intérieur des familles¹.

Dans une généalogie, quatre profils sont envisageables selon le chromosome où se trouve le gène affecté (autosome ou chromosome X) et si le phénotype est dominant ou récessif¹. Une transmission autosomique dominante est causée par un gène mutant situé sur un autosome. Dans ce profil, étant donné que l'allèle mutant est dominant, il suffit que l'un des deux chromosomes porte cet allèle pour que l'homme ou la femme soit affecté par la maladie. Pour sa part, la transmission autosomique récessive est aussi caractérisée par un gène localisé sur des autosomes. Comme les allèles de la maladie sont récessifs, celle-ci se manifeste uniquement chez les hommes ou les femmes homozygotes, car les hétérozygotes possèdent un allèle « normal » masquant celui responsable de la maladie. Concernant le mode récessif lié au chromosome X, celui-ci réfère à la transmission liée au sexe et dans ce cas-ci les allèles récessifs se retrouvent évidemment sur le chromosome X. Comme l'homme est hémizyote pour ce chromosome, s'il y a présence de l'allèle récessif sur son unique chromosome X, il présentera sans contredit la maladie; toutefois, chez les femmes, seulement celles homozygotes seront affectées par la maladie. Le dernier profil, la transmission dominante liée au X, survient si l'allèle dominant se situe sur le chromosome X et la maladie apparaîtra chez les hommes hémizyotes de même que chez les femmes hétérozygotes.

L'étude des maladies à transmission mendélienne est simplifiée par le fait que la source génétique de la pathologie est située dans un unique gène et que le phénotype observé peut être directement attribuable au génotype.

1.3 Désordre multifactoriel

Ce désordre, appelé aussi maladie complexe, est le résultat d'une combinaison de petites variations dans plusieurs gènes. Ces variations peuvent toucher un petit nombre de nucléotides aussi peu que le changement d'un seul nucléotide (SNPs) et montrer divers degrés d'importance sur le phénotype correspondant (pénétrance). Une variation dans un seul gène ne peut être directement responsable de la maladie, mais sa présence est nécessaire afin de la produire. Les facteurs environnementaux tels l'alcool, la nutrition, l'activité physique et le poids corporel, y sont souvent impliqués, comme c'est le cas pour l'hypertension³, aggravant ou accélérant le trait en question. En effet, en agissant au niveau de l'expression génétique, ces facteurs environnementaux peuvent mener à une modification ou à une amplification phénotypique^{4,5} (Figure 1). Tout comme les maladies mendéliennes, les maladies multifactorielles témoignent d'une grande variation ethnique et géographique; mais, au sein des familles, aucun profil de transmission spécifique n'est décelable¹. Parmi tous les désordres, ceux multifactoriels sont de loin les plus fréquents; par exemple, après l'âge de 25 ans, la fréquence des désordres monogéniques est de 16.4/1000 naissances d'individus tandis qu'elle est de 600/1000 pour les désordres complexes (Tableau I).

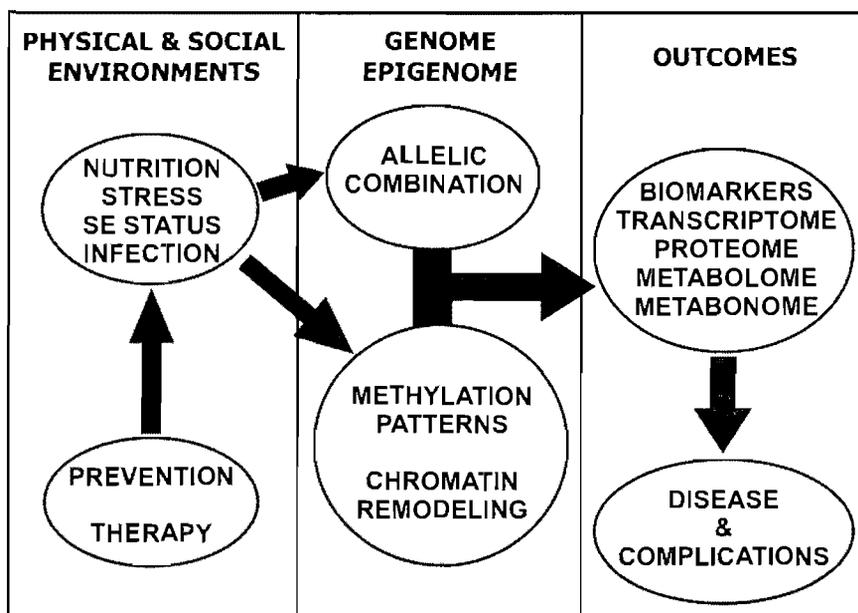


Figure 1 Interactions entre les gènes et l'environnement. SE = socio-économique, tiré de Seda *et al.*, 2005⁵

Chapitre 2 - Hypertension

2.1 Pression artérielle

Cliniquement, la pression artérielle représente la pression sanguine dans la circulation systémique. Cette dernière est la force, par unité de surface, exercée par le sang contre la paroi d'un vaisseau et elle est mesurée en millimètres de mercure (mmHg). La pression artérielle est influencée par la fréquence des contractions cardiaques, l'élasticité des artères, ainsi que le volume de sang propulsé. Elle est exprimée par deux mesures qui sont la pression artérielle systolique (PAS) et la pression artérielle diastolique (PAD). Respectivement, il s'agit de la pression maximale au moment de la contraction du cœur qui expulse le sang dans l'aorte (systole) et la pression minimale au moment du relâchement du cœur et de l'évacuation du sang de l'aorte (diastole).

2.1.1 Systèmes régulateurs de la pression artérielle

Des mécanismes humoraux, rénaux et vasculaires assurent la régulation de la pression artérielle. Les mécanismes humoraux incluent les hormones de la médulla et du cortex surrénal, le facteur natriurétique auriculaire, l'hormone antidiurétique, l'angiotensine II (Ang II), les facteurs endothéliaux et le monoxyde d'azote (NO). Tandis que le système rénine-angiotensine, les facteurs de prolifération, d'inflammation, d'immunité, les cytokines et autres, définissent les mécanismes rénaux et vasculaires. Seuls les mécanismes ayant un lien au projet seront décrits davantage ci-dessous.

2.1.1.1 Mécanismes humoraux

2.1.1.1.1 Angiotensine II

Au moment où l'appareil juxta-glomérulaire des reins libère la rénine pour contrer une perfusion rénale inadéquate, l'angiotensine II est produite par son action sur le substrat, l'angiotensinogène. Cette hormone favorise la libération de l'hormone antidiurétique par la neurohypophyse en plus d'aldostérone par les glandes surrénales (Figure 2). En agissant sur les artérioles systémiques, dans le cas de l'hormone antidiurétique et sur les tubules rénaux dans celui de l'aldostérone, ces hormones causeront, par différents mécanismes, une élévation de la pression artérielle systémique⁶.

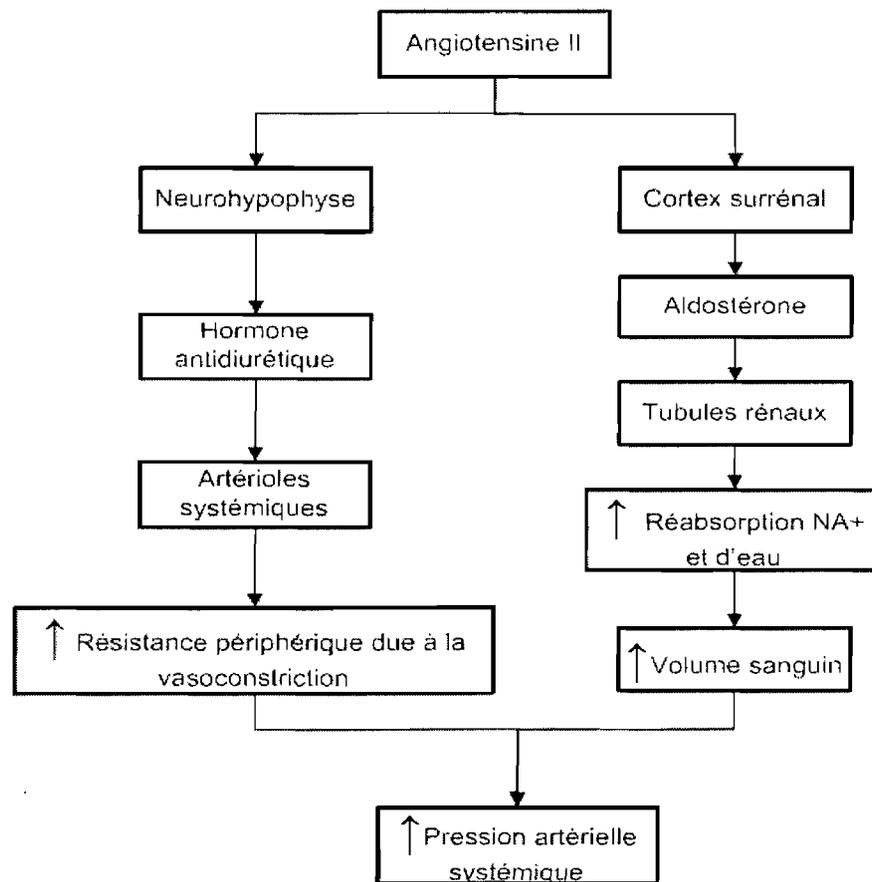


Figure 2 Rôle de l'angiotensine II dans la régulation de la pression artérielle

2.1.1.1.2 Monoxyde d'azote

Au moment où le débit sanguin s'accélère, du monoxyde d'azote est libéré par les cellules endothéliales. Cette substance produira une vasodilatation systémique très localisée, agissant par l'intermédiaire de la cGMP^{7, 8}. Cette dernière est un nucléotide cyclique dérivé de la guanosine triphosphate (GTP) qui agit comme second messenger. La cGMP joue un rôle, entre autres, dans le transport du sodium⁹ et dans la relaxation des muscles lisses vasculaires dont sont faits principalement les vaisseaux sanguins^{10, 11}. Les effets du monoxyde d'azote sont de courte durée, car le monoxyde d'azote est rapidement détruit; toutefois, son rôle vasodilatateur est majeur, car lors de son inhibition la pression artérielle monte en flèche^{6, 12}.

2.1.1.2 Mécanismes rénaux et vasculaires

2.1.1.2.1 Système rénine-angiotensine

En réponse à une baisse de la pression artérielle, les cellules spécialisées de l'appareil juxtaglomérulaire des reins vont produire la rénine dans le sang⁶. Cette enzyme, par son action catalytique, va transformer l'angiotensinogène, une glycoprotéine produite majoritairement dans le foie, en Ang I (Figure 3).

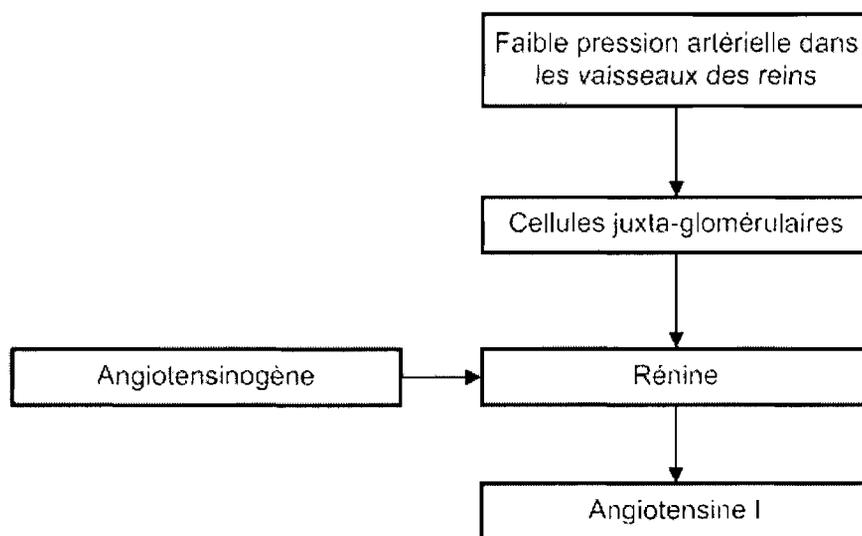


Figure 3 Mécanisme du système rénine-angiotensine

Ensuite, de nombreuses réactions enzymatiques définiront les deux axes de ce système¹³ (Figure 4). Dans le premier axe, découvert en premier, l'Ang I est convertie en un important vasoconstricteur, l'Ang II, par l'enzyme de conversion de l'angiotensine (ACE). Dans le second axe, l'Ang I est hydrolysée par l'enzyme de conversion de l'angiotensine 2 (ACE2), une mono-peptidase nouvellement découverte, pour donner l'Ang 1-9 et à partir de celle-ci ACE va générer l'Ang 1-7, un vasodilatateur. ACE2 participe aussi à une autre voie qui connecte les deux axes dans laquelle elle utilise l'Ang II pour produire l'Ang 1-7¹³. Vu l'effet contraire de ces axes, ceux-ci représentent une balance qui maintient un état normotendu; par conséquent, si un déséquilibre survient entre les deux axes, il peut en résulter un état hypertendu ou hypotendu. Ce système crucial de la pression artérielle serait impliqué dans le développement de l'hypertension de même qu'au niveau des blessures rénales¹⁴.

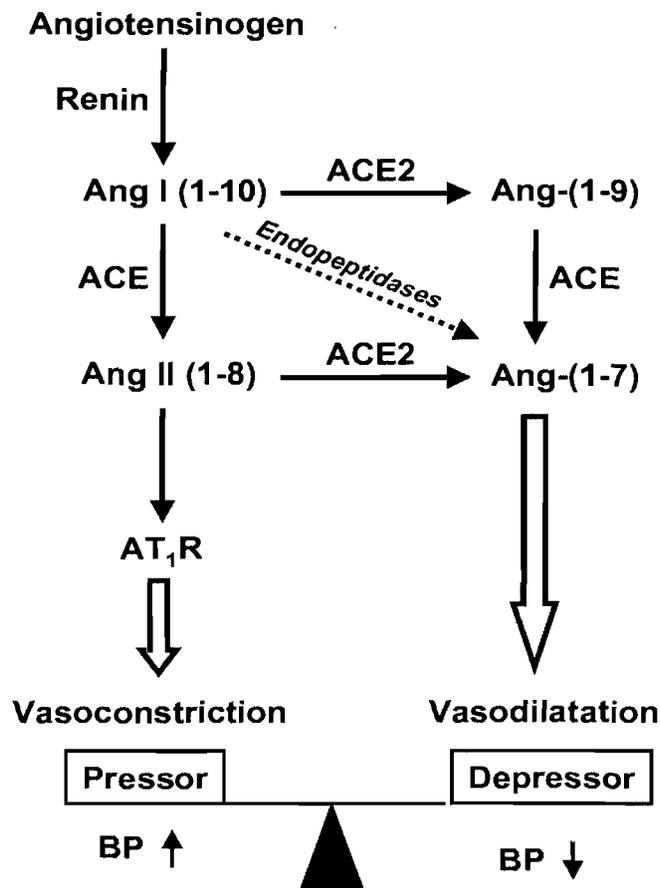


Figure 4 Rôle d'ACE2 dans le système rénine-angiotensine, tiré de Yagil et Yagil, 2003¹³

Le gène ACE2, localisé dans la région Xp22, a été découvert il y a sept ans¹⁵⁻¹⁷ et depuis ce temps, l'attention qui lui est portée ne cesse de grandir. La traduction de ce gène résulte en une protéine membranaire intégrale de type 1 qui fonctionne comme une carboxypeptidase¹⁸. ACE2 est le premier homologue humain du gène ACE connu et comporte 40% d'identité avec ce dernier¹⁹. Un gène ACE2 a également été identifié chez la souris (XF5) ainsi que chez le rat (Xq21). Concernant sa distribution, ACE2 a été identifiée dans le cœur, les reins, les testicules et récemment, dans les poumons, le cerveau et le tractus gastro-intestinal²⁰.

La littérature suggère qu'ACE2 serait impliquée dans la régulation de la fonction cardiaque. En effet, en étudiant des souris knockouts pour ACE2, Crackower *et al.* ont démontré que l'absence de cette enzyme est associée à un défaut de contractilité cardiaque

de même qu'à une augmentation du niveau d'angiotensine II²¹. De plus, des associations entre ce gène et l'hypertension ont été soulevées dans la littérature. Entre autres, chez le rat spontanément hypertendu (SHR), Crackower *et al.* ont montré que les niveaux rénaux d'ACE2 s'avèrent être plus bas que chez les rats normotendus Wistar-Kyoto (WKY)^{20, 21}. Chez l'humain, diverses études rapportent des associations entre des polymorphismes situés dans le gène ACE2 et l'hypertension essentielle^{22, 23}. Même si le rôle joué par ACE2 dans la pathophysiologie de l'hypertension reste à clarifier, celui-ci représente une cible potentielle et intéressante pour le traitement de l'hypertension et des maladies cardiovasculaires²⁴. Dans le projet proposé, une approche par gène candidat a été utilisée afin d'étudier le rôle d'ACE2 dans l'hypertension.

2.1.2 Distribution de la pression artérielle

Le profil des PAS et PAD par rapport à l'âge est bien établi et pour différentes origines ethniques (Figure 5) et populations (Figure 6), celui-ci est bien conservé. Les figures 5 et 6 montrent que la PAS augmente progressivement avec l'âge, alors que la PAD augmente jusqu'à l'âge de 50 ans et diminue par la suite^{25, 26}. Jusqu'à environ 65 ans, la PAS est plus élevée chez les hommes que chez les femmes, mais l'inverse survient vers 70 ans dans la CHHS²⁵.

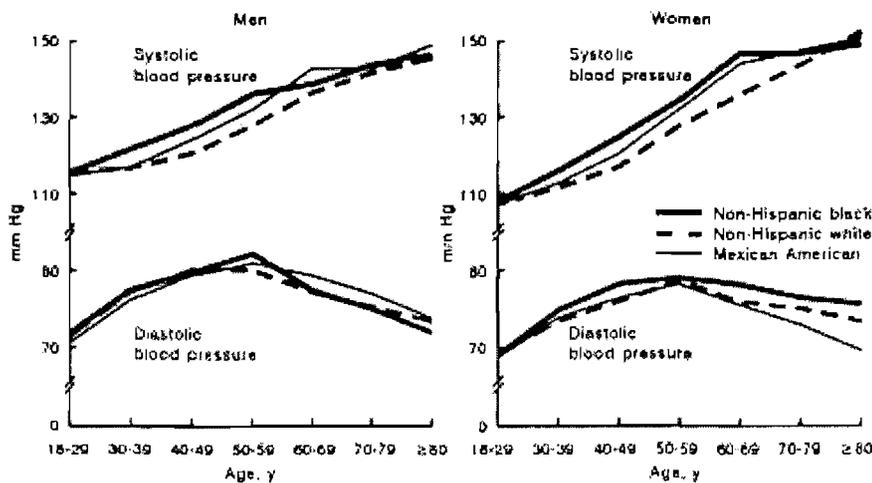


Figure 5 Profil de la pression artérielle systolique et diastolique selon le sexe, l'âge et différentes ethnies, tiré de Burt *et al.* 1995²⁶. Distribution de la PAS et la PAD selon l'âge et l'origine ethnique pour des hommes et des femmes âgés de 18 ans et plus. Ces données sont tirées d'études d'un échantillon représentatif de la population des États-Unis, *the Third National Health and Nutrition Examination Surveys* (NHANES III).

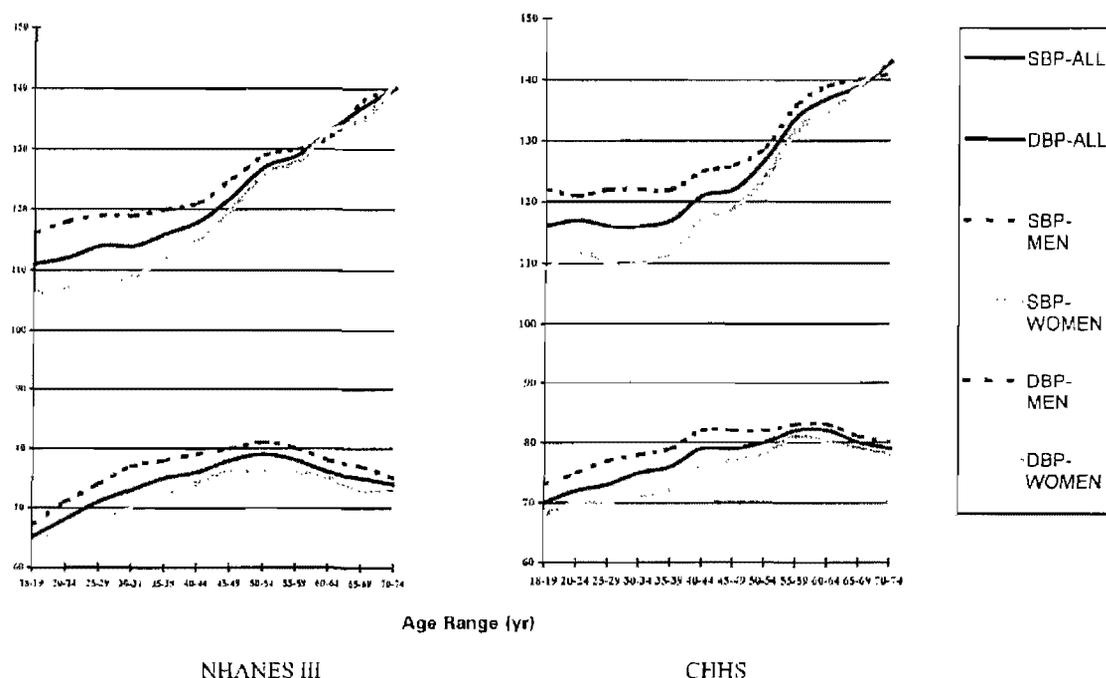


Figure 6 Profil de la pression artérielle systolique et diastolique, selon le sexe et l'âge, basé sur deux études d'échantillon représentatif des États-Unis (NHANES) et du Canada (CHHS), tiré de Joffres *et al.*, 2001²⁵. Distribution de la PAS et la PAD selon l'âge (18 ans à 74 ans) et le sexe pour *the Third National Health and Nutrition Examination Surveys* (NHANES III) et la *Canadian Heart Health Surveys* (CHHS) qui contient des données de dix provinces canadiennes.

2.2 Hypertension artérielle

L'hypertension artérielle est une maladie qui représente le plus important facteur de risque pour le développement de maladies cardiaques, d'accidents vasculaires cérébraux, d'insuffisance cardiaque, d'insuffisance rénale et d'insuffisance coronarienne. Elle se définit par une élévation anormale de la pression artérielle systolique et/ou diastolique. Un individu est hypertendu si sa pression systolique est égale ou supérieure à 140 mmHg et/ou si sa pression diastolique est égale ou supérieure à 90 mmHg. L'hypertension secondaire et l'hypertension essentielle désignent les deux classes d'hypertension artérielle existantes.

2.2.1 Hypertension secondaire

Il est question d'hypertension secondaire lorsque l'augmentation de la pression artérielle est la conséquence d'une cause sous-jacente identifiable. Les causes les plus communes sont la maladie rénoparenchymale, l'apnée du sommeil, l'aldostéronisme primaire, l'hyperthyroïdisme, le syndrome de Cushing, la sténose de l'artère rénale et le phéochromocytome^{27, 28}. Une fois la cause découverte, un traitement adéquat (chirurgie,

diète, exercice, radiothérapie, etc.) peut être envisagé pour ramener la pression artérielle à une valeur normale sans avoir recours à la prise de médicaments antihypertenseurs²⁹. Dans la population, la prévalence de la forme secondaire de l'hypertension est estimée à environ 5%^{28, 30, 31}. Dans le cas des personnes âgées, sa détection représente un défi pour les médecins, car les méthodes de diagnostic sont coûteuses, des facteurs pathophysiologiques peuvent influencer les symptômes et le risque des interventions chirurgicales est davantage élevé chez ces sujets³⁰.

2.2.2 Hypertension essentielle

L'évidence substantielle que la pathologie de l'hypertension consistait en une partie génétique et une partie héréditaire fut amenée par Biron *et al.* en 1976 par l'étude de familles canadiennes-françaises comprenant des enfants naturels et adoptés³². Ils ont évalué le pointage de corrélation de la pression artérielle en premier entre les parents et les enfants naturels, puis entre les parents et les enfants adoptés. Le pointage fut significatif pour le premier type de comparaison et non pour le deuxième et la corrélation entre une paire d'enfants fut seulement significative quand ceux-ci sont naturels. Étant donné que l'enfant adopté et celui naturel vivent au sein de la même famille, ils partagent donc le même environnement, d'où la possibilité pour les auteurs de conclure que l'hérédité explique une grande partie de la ressemblance familiale de la pression artérielle chez les enfants. Contrairement à l'hypertension secondaire, l'hypertension essentielle survient en l'absence de cause secondaire connue. Par les données des deux études référées précédemment, la NHANES III (1988-1994) et la CHHS (1986-1992), la prévalence de l'hypertension essentielle en Amérique du Nord serait d'environ 21% chez les gens âgés de 18 à 74 ans²⁵. En 2000, elle s'élève à 26,4% dans la population mondiale³³. L'hypertension essentielle est divisée en deux catégories : l'hypertension mendélienne et polygénique.

2.2.2.1 Hypertension mendélienne

À l'origine d'un défaut à l'intérieur d'un seul gène, l'hypertension mendélienne regroupe toute forme monogénique d'hypertension qui est héritée selon un des modèles mendéliens (Tableau II). Ce type d'hypertension est plutôt rare; cependant, la compréhension de sa pathologie pourrait fournir de nouvelles connaissances sur la régulation de la pression artérielle et par ce fait, aider à la résolution de la forme complexe.

Tableau II Les formes mendéliennes d'hypertension, adapté d'Hamet *et al.*, 1998³⁴

| Syndrome | Hérédité | Chromosome | Gène |
|---|----------|------------------------------|--------------------------------------|
| Défauts dans le métabolisme des stéroïdes | | | |
| Hyperaldostéronisme sensible aux glucocorticoïdes | ADI | 8q22 | CYP11B1/CYP11B2 |
| Pseudo-hermaphroditisme chez les hommes | ARI | 10q24.3 | CYP17 |
| Pseudo-hermaphroditisme chez les femmes | ARI | 8q22 | CYP11B1 |
| Excédent apparent de minéralocorticoïdes | ARI | 16q22 | HSD11B2 |
| Défauts dans le transport ionique | | | |
| Syndrome Liddle | ADI | 16p12–p13 | SCNN1B SCNN1G |
| Syndrome Gordon | ADI | 1q31-q42 12p13.3 17q21 | PHAI1A (locus) WNK1 WNK4 |
| Surproduction de catécholamines | | | |
| Phéochromocytome isolé | ADI | 1p | ? |
| Néoplasie endocrine multiple de type II A | ADI | 10q11.2 | RET proto-oncogene |
| Néoplasie endocrine multiple de type II B | ADI | 10q11.2 | RET proto-oncogene |
| Syndrome von Hippel-Lindau | ADI | 3p26–25 | VHL gène suppresseur de tumeur |
| Neurofibromatose de type 1 | ADI | 17q11.2 | NF1 |
| Autre | | | |
| Hypertension et brachydactylie | ADI | 12p12.2-p11.2 | HTNB |

ADI : forme d'hérédité autosomique dominante ; ARI : forme d'hérédité autosomique récessive

2.2.2.2 Hypertension polygénique

L'hypertension polygénique représente la forme complexe ou multifactorielle, puisqu'elle fait intervenir plusieurs gènes avec un effet modeste aussi bien que des facteurs environnementaux. En effet, 20-50% de la variation interindividuelle de la pression artérielle serait déterminée génétiquement³⁵. Aucun profil d'hérédité clair n'est observé chez cette forme d'hypertension; par contre, plusieurs caractéristiques cliniques, pathologiques et hémodynamiques aident à mieux cerner cette forme complexe. Une histoire de famille positive et la race en sont des exemples³⁶. Il ne faut pas oublier aussi les habitudes de vie telles que le régime alimentaire, le tabagisme et l'exercice.

Plusieurs approches sont utilisées pour disséquer les bases génétiques de l'hypertension essentielle polygénique. Une de ces approches est d'effectuer des études de liaisons sur une population fondatrice afin de réduire l'hétérogénéité génétique et d'obtenir

des expositions environnementales davantage uniformes³⁷. Il s'agit de celle adoptée récemment par le laboratoire du Dr. Hamet sur une population fondatrice du Canada, le Saguenay-Lac-Saint-Jean, décrite plus loin³⁸. L'aspect unique de leur étude repose sur le fait que plus de 250 phénotypes, reliés à de nombreux facteurs cardiovasculaires, hormonaux, neuraux, métaboliques et rénaux, et impliqués dans la régulation de la pression artérielle, ont été obtenus pour chaque individu de l'échantillon³⁷. Ces phénotypes ont été analysés dans 120 familles dans lesquelles au moins une paire d'enfants de mêmes parents était atteinte d'hypertension de manière prématurée. L'étude visait l'identification et la caractérisation des voies physiologiques par lesquelles les mutations agissent pour augmenter ou maintenir la pression artérielle de même que l'identification de loci de traits quantitatifs de la pression artérielle. Ceci a permis la découverte de phénotypes intermédiaires potentiellement utiles dans la division des sujets hypertendus en sous-groupes plus homogènes pour réduire la variation dans les voies de signalisation de la maladie et faciliter l'identification de gènes. De plus, l'obtention de plusieurs phénotypes mesurables avec une meilleure précision que celle de la pression artérielle a amélioré la puissance des analyses. Cette étude a révélé au total 46 loci, représentant le plus grand nombre de loci contribuant aux traits cardiovasculaires et métaboliques rapporté à date dans une unique population³⁷.

2.3 Composants métaboliques de l'hypertension artérielle

2.3.1 Dyslipidémie

L'hypertension est souvent associée à d'autres facteurs de risques cardiovasculaires tels le diabète, l'obésité et la dyslipidémie³⁹. C'est d'une part, en créant une dysfonction endothéliale des vaisseaux sanguins que la dyslipidémie jouerait un rôle dans le développement de l'hypertension⁴⁰. Prise avec la dyslipidémie et l'obésité, l'hypertension représente une proportion significative du fardeau global et régional des maladies⁴¹. Le fardeau régional représentant, dans cette étude, le nombre d'années de vie perdue à cause d'une mort prématurée ou d'une invalidité (unité DALY) ainsi que le taux de mortalité. Dans cet article, ces mesures ont été évaluées sur 14 régions épidémiologiques sélectionnées dans le monde et en divisant les régions (pauvres, en développement et développées). Dans d'autres cas, le fardeau régional aurait pu se traduire en fardeau économique. Les récentes données apportées par l'étude d'Halperin *et al.*, accomplie chez

les hommes exclusivement, indiquent que la dyslipidémie serait associée au développement de l'hypertension et que les lipides plasmatiques s'avèreraient très utiles dans l'identification des personnes à risque pour cette maladie⁴². De leur côté, Sesso *et al.* ont montré que la dyslipidémie est également impliquée dans la pathologie de l'hypertension chez les femmes⁴³. Même chez différentes ethnies, le niveau de triglycérides constitue un élément de prédiction du développement de l'hypertension⁴⁴.

La sélection de paires de frères et sœurs, basée sur la présence d'hypertension et de dyslipidémie, ont permis au laboratoire du Dr. Hamet de découvrir de nombreux composants du syndrome métabolique sur le chromosome 1 et 3³⁸. De plus, son laboratoire a aussi été le premier à accomplir une analyse de balayage génomique global du syndrome métabolique dans le modèle de rats PD/Cub qui a révélé de nouveaux loci impliqués dans la génétique de la dyslipidémie⁴⁵.

2.3.2 Obésité

Aux États-Unis, durant la dernière décennie, la prévalence de l'obésité est passée de 23% à 30% et environ 66% des adultes ont un excès de poids ou sont obèses⁴⁶. L'étendue de l'obésité au sein des réseaux sociaux (amis, enfants, entourage) représenterait un facteur contribuant à l'épidémie d'obésité et ce phénomène de réseau s'avèrerait être un élément révélateur des traits comportementaux et biologiques de l'obésité⁴⁷. Même si l'obésité est le facteur de risque principal pour le développement de l'hypertension essentielle⁴⁸, elle n'est pas toujours une partie intégrante de cette dernière. En effet, des études de ségrégation proposent que l'obésité et l'obésité associée à l'hypertension partagent certains déterminants génétiques⁴⁹. Un de ceux-ci serait le locus du facteur onconécrosant-alpha, une cytokine pro-inflammatoire, suggéré par l'investigation de gènes candidats au sein de généalogies hypertendues du SLSJ⁴⁹. D'un autre côté, l'hypertension peut affecter les individus obèses aussi bien que ceux non obèses ; ces deux manifestations différant par diverses caractéristiques cliniques dont une réabsorption rénale du sodium plus élevée chez les individus obèses et hypertendus⁴⁸. Pausova *et al.* ont examiné l'homogénéité génétique chez des familles hypertendues en provenance du Saguenay-Lac-Saint-Jean (SLSJ) et des sous-ensembles de familles obèses et non obèses⁵⁰. Leurs résultats ont montré que le degré d'homogénéité génétique, pour cette population, est augmenté chez les familles

sélectionnées pour l'hypertension et davantage élevé au sein des deux sous-ensembles⁵⁰. Ils suggèrent donc que l'hypertension avec et sans obésité représenterait, en partie, des mécanismes pathophysiologiques et des entités génétiques distinctes ; par conséquent, certains des gènes communs aux deux sous-ensembles contribueraient aux différences cliniques observées.

2.3.3 Syndrome métabolique

En plus des mesures d'obésité abdominale et de dyslipidémie athérogène, celles d'hypertension et d'élévation de la glycémie représentent les caractéristiques du syndrome métabolique (SM). Précisément, ce syndrome est constitué d'une combinaison de phénotypes incluant : la circonférence de la taille, la pression artérielle, le niveau de lipoprotéines de haute densité (HDL), les triglycérides et le glucose. D'après *the Adult Treatment Panel III of the US National Cholesterol Education Program*, une personne possède ce syndrome si elle rencontre trois des critères mentionnés dans le tableau III⁵¹. Ce syndrome représente un risque accru pour le développement de maladies cardiovasculaires. En effet, Safar *et al.* ont démontré que le SM contribue au vieillissement artériel accéléré et donc à la progression de la rigidité aortique avec l'âge⁵². Dans leur étude, basée sur les canadiens-français, Hamet *et al.* proposent une approche généalogique quantitative afin de résoudre les maladies complexes tels le syndrome métabolique de même que l'hypertension³⁸. Cette méthode a entre autres permis d'identifier des loci spécifiques pour des composants du syndrome métabolique.

Tableau III Seuils des phénotypes définissant le syndrome métabolique, adapté de Khunti and Davies⁵¹

| Sexe | Circonférence de la taille (cm) | Pression artérielle (mmHg) | HDL (mmol/L) | Triglycéride (mmol/L) | Glucose (mmol/L) |
|--------|---------------------------------|--|--------------|-----------------------|------------------|
| Femmes | >88 | Systolique > 130 Diastolique > 85 ou médicaments | < 1,29 | >1,69 | > 6,1 |
| Hommes | > 102 | Systolique > 130 Diastolique > 85 ou médicaments | <1,04 | >1,69 | > 6,1 |

Chapitre 3 - Chromosome X

3.1 Les caractéristiques du chromosome X

Les chromosomes sexuels ont évolué à partir d'une paire de chromosomes autosomaux dans les derniers 300 millions d'années. Dans le processus, le chromosome X a conservé les éléments fonctionnels tandis que le chromosome Y n'a presque rien préservé des chromosomes autosomaux. Ces deux chromosomes forment ensemble la 23^e paire de chromosomes. Le chromosome X détient des caractéristiques qui lui sont uniques dans le génome humain. Les femmes en possèdent deux et les hommes seulement un et sont donc hémizygotés pour ce chromosome. Son mode de transmission est particulier, la femme en transmet un à sa fille et un à son garçon. Le père, de son côté, en donne un à sa fille et transmet plutôt le chromosome Y à son fils. Le chromosome X comporte environ 155 millions de paires de bases (pb). Ross *et al.* ont déterminé 99,3% de la séquence d'euchromatine de ce chromosome et ils ont annoté 1098 gènes desquels 699 sont des gènes connus, 132 sont des nouvelles séquences codantes, 166 des nouveaux transcrits et 700 des pseudogènes⁵³. Ils ont observé que la densité de gènes, mis à part les pseudogènes, est parmi la plus basse des chromosomes annotés jusqu'à présent, tout comme la longueur des gènes; malgré tout, ce chromosome contient le plus long gène connu du génome, celui de la dystrophine. Les exons des 1 098 gènes représentent seulement 1,7% de la séquence de ce chromosome et en se basant sur la longueur des locus des gènes, 33% du chromosome X est transcrit⁵³.

L'importance du chromosome X dans l'hypertension reste encore très évasive. Les caractéristiques génétiques particulières de ce chromosome (mode de recombinaison, de ségrégation, de transmission et l'hémizygoté de l'homme) compliquent la dissection des attributs génomiques de l'hypertension. Étant donné son comportement distinct, le chromosome X requiert un traitement spécial et adapté lors des analyses génétiques, car si ce chromosome est traité comme un autosome, une différence au niveau du sexe dans le phénotype pourrait mener à une fausse liaison⁵⁴. En raison de ces difficultés, les méthodes d'analyses génétiques focalisent principalement sur les autosomes, le chromosome X demandant davantage d'efforts et de réflexion pour être étudié de manière appropriée. Le développement/conception des logiciels génétiques adaptés au chromosome X devient donc

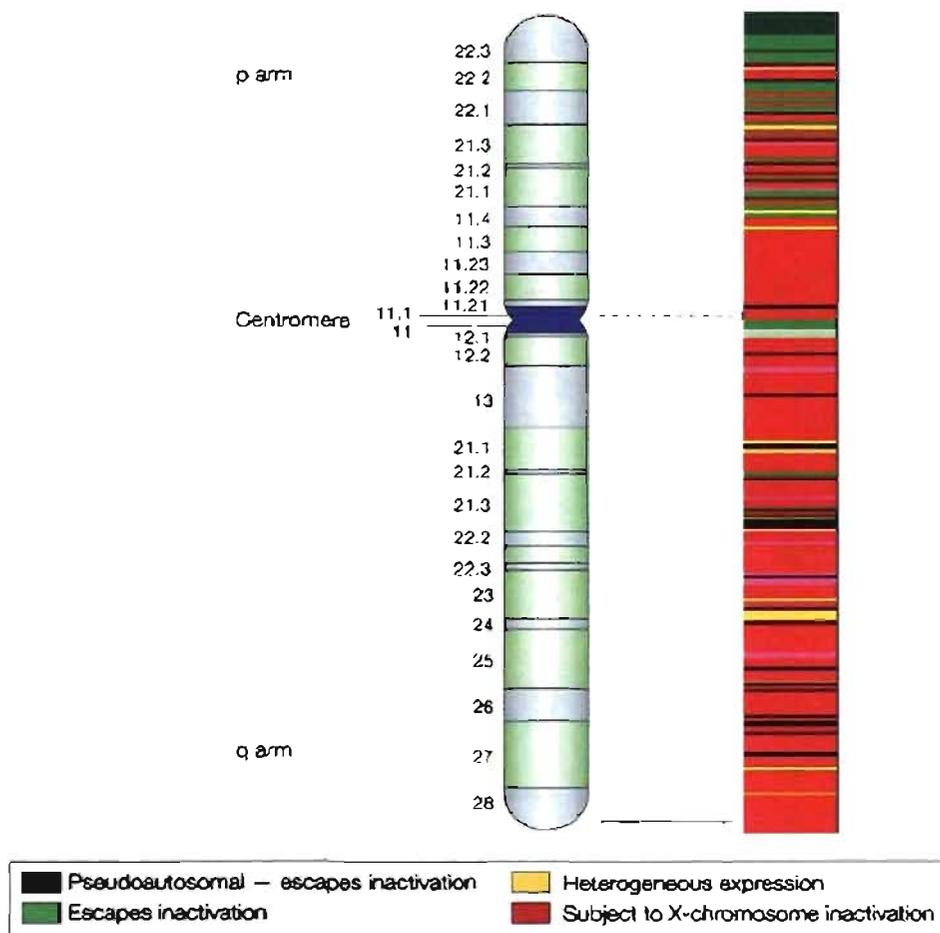
une tâche exigeante et complexe, ce qui justifie la rare quantité de ceux-ci pour les chercheurs.

3.2 Inactivation du chromosome X

Tôt dans le développement embryonnaire de la femme, un des deux chromosomes X est inactivé et reste inactif dans les tissus somatiques⁵⁵. Cependant, dans la lignée germinale de la femme, le chromosome inactif est réactivé afin de subir la recombinaison méiotique avec le deuxième chromosome⁵³. Il n'est pas tout à fait clair si l'inactivation se fait de manière aléatoire ou non. Williams et Wu proposent deux modèles pour ce processus : un aléatoire, mais dont le choix ne se fait pas entre les deux chromosomes X et l'autre non aléatoire, dans lequel le choix arriverait dans la génération des cellules précédant l'inactivation du chromosome X par une marque asymétrique induite ou héritée du parent⁵⁶.

Ce processus d'inactivation sert à doser, entre les femmes et les hommes, les produits des gènes liés au chromosome X. Cette inactivation est initiée au centre d'inactivation du chromosome X (XIC)⁵⁷. Le gène du transcrit spécifique d'inactivation du chromosome X (XIST), qui se trouve dans le XIC, encode un acide ribonucléique (ARN) non traduit qui est requis et suffisant pour l'inactivation^{58, 59}. Ce transcrit est exprimé, au départ, sur les deux chromosomes X et devient finalement exprimé seulement sur celui inactif^{53, 59}. Cette inactivation requiert, sur le chromosome inactivé, un bon nombre de changements au niveau de la chromatine, pour que ce dernier devienne transcriptionnellement inactif⁶⁰⁻⁶². L'ARN XIST, associé au chromosome inactivé, intervient aussi à cette étape en recrutant des complexes qui modifient la chromatine⁶³; cependant, 15% des gènes échappent toujours à cette inactivation et 10% de plus y échappent occasionnellement⁶⁴ (Figure 7). Le fait que des gènes échappent à l'inactivation est sûrement dû au fait que ceux-ci ne subissent pas des modifications épigénétiques suffisantes. Selon les régions du chromosome X, le nombre de gènes échappant à ce phénomène diffère; résultat de l'évolution des chromosomes sexuels. Fait étonnant, le nombre de gènes échappant à ce processus diffère d'une femme à l'autre reflétant une expression hétérogène chez les femmes pour ces gènes⁶⁴. Ce phénomène inclut plusieurs gènes n'ayant pas d'équivalent fonctionnel sur le chromosome Y⁶⁴. Dans le cas de

chromosome surnuméraire à l'intérieur d'un noyau diploïde, chaque chromosome en excès sera inactivé pour n'avoir qu'un seul chromosome X actif.



Nature Reviews | Cancer

Figure 7 Structure du chromosome X et le profil d'inactivation des gènes, tiré de Spatz *et al.*, 2004⁶⁵

3.3 Recombinaison

Dans la lignée germinale de la femme, le chromosome X inactif est réactivé et subit la recombinaison méiotique avec le second⁵³; tout comme les autosomes, la longueur entière du chromosome est impliquée dans le processus. À l'opposé, la recombinaison entre les chromosomes X et Y s'effectue sur une petite portion de ceux-ci située à l'extrémité des bras. Étant donné que les gènes compris dans ces régions sont partagés entre les deux chromosomes sexuels, ces régions sont dites pseudoautosomales. Deux régions

pseudoautosomales sont présentes, l'une sur les bras courts des chromosomes sexuels, PAR1, et l'autre sur les bras longs, PAR2 (Figure 8). L'homologie entre la région PAR1 du chromosome X et Y est maintenue grâce à la recombinaison dans la méiose des hommes. Les gènes de cette région ne subissent pas l'inactivation du chromosome X et se retrouvent donc en deux copies chez les femmes et les hommes. De son côté, PAR2 a été produite par la duplication du matériel du chromosome X au chromosome Y depuis la divergence des lignées de l'humain et du chimpanzé. Certains gènes de cette région sont sujets à l'inactivation du chromosome X⁵³ (Figure 7).

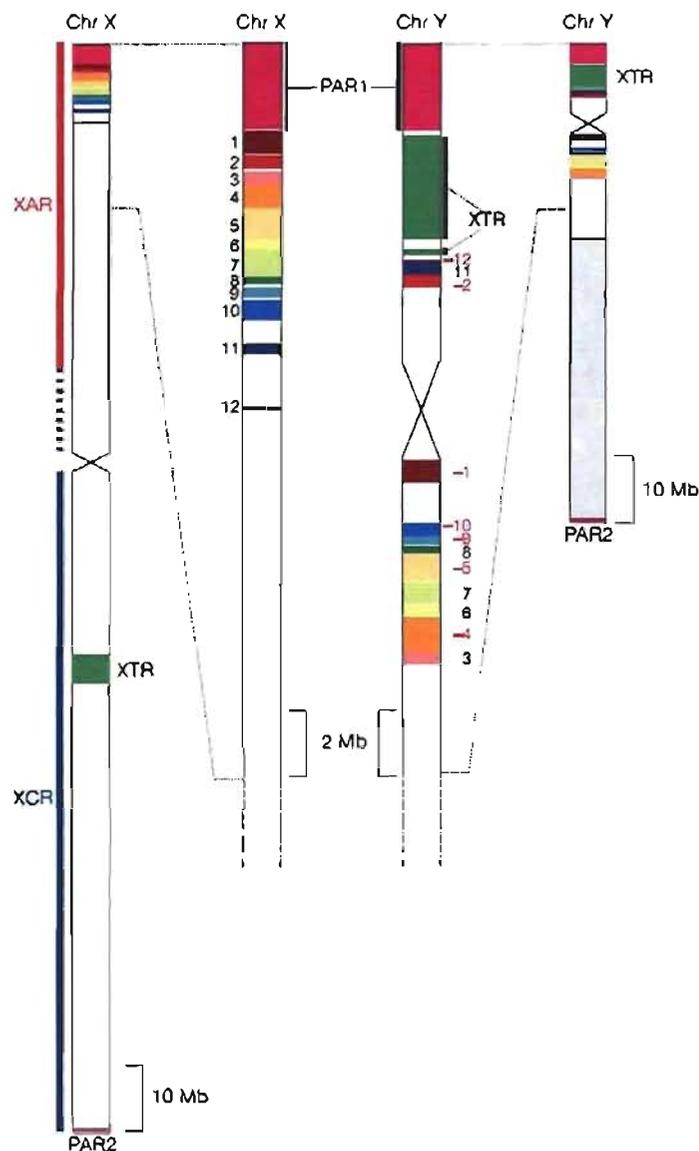


Figure 8 Représentation schématique des homologies majeures entre les chromosomes sexuels de l'humain, tiré de Ross *et al.*, 2005⁵³

3.4 Syndromes résultant d'un nombre de chromosomes X anormaux

3.4.1 Syndrome de Turner

La première évidence d'un rôle génétique du chromosome X dans l'hypertension fut amenée par l'étude d'Hilbert *et al*⁶⁶. Ensuite, d'autres découvertes l'ont confirmée, comme le syndrome de Turner. Ce phénomène qui survient seulement chez les femmes, est la conséquence d'une monosomie complète ou partielle d'un chromosome X. Les femmes viables montrent une petite stature ainsi qu'une dysgénésie gonadique les rendant pour la plupart infertiles. La majorité de ces femmes souffrent de maladies cardiovasculaires, d'ostéoporose et de désordres gastro-intestinaux et rénaux⁶⁷. Les problèmes liés au syndrome métabolique comme la dyslipidémie, l'obésité et en particulier une pression artérielle élevée sont aussi observables chez ces femmes⁶⁸.

3.4.2 Syndrome de Klinefelter

Le syndrome de Klinefelter est le désordre chromosomique sexuel le plus commun chez l'homme et il a été décrit pour la première fois en 1942 par Harry Klinefelter. Un homme positif pour cette maladie montre un caryotype avec un ou plusieurs chromosomes X surnuméraires. Le phénotype peut varier, mais en général l'homme affecté présente un mauvais développement des testicules qui peut mener à l'infertilité, un hypogonadisme, des difficultés d'apprentissage et un niveau élevé de gonadotropine^{69, 70}.

Chapitre 4 - Approche génétique

4.1 Population du Saguenay-Lac-Saint-Jean du Québec

Depuis plus de quarante ans, la population du SLSJ (Figure 9) est la cible de nombreuses études qui résultent en l'accroissement de nos connaissances sur les déterminants génétiques des désordres monogéniques⁷¹. Ce n'est que récemment que cette population a été mise à profit des désordres complexes dont l'hypertension⁴⁹. La disponibilité de documents généalogiques datant de l'époque des fondateurs français au 17^e siècle rend très avantageuse son étude. L'établissement de cette région, dont le territoire couvre environ 11 000 km², a débuté au milieu des années 1800; les fondateurs provenant majoritairement de la région de Charlevoix. L'effet fondateur qui a créé cette population est à l'origine du pool génétique observable au sein de celle-ci. Cet effet se caractérise par la création d'une nouvelle population à partir d'un petit nombre d'individus en provenance d'une population mère. À ce moment-là, la nouvelle population ne contient qu'une parcelle de la variation génétique de la population mère. Ce qui rend la nouvelle population davantage homogène génétiquement que la population mère.



Figure 9 Localisation de la population du Saguenay-Lac-Saint-Jean (en rouge) de la région de Québec, tiré de <http://en.wikipedia.org>

Cette population, géographiquement isolée, comporte diverses particularités génétiques. Effectivement, certaines maladies héréditaires montrent une prévalence inhabituellement élevée ou sont spécifiques à cet endroit comme la tyrosinémie⁷² et l'ataxie spastique de Charlevoix–Saguenay respectivement⁷¹, tandis que d'autres sont extrêmement

rare. En plus de l'effet fondateur, un autre phénomène peut expliquer ces particularités : une forte natalité⁷¹. À une certaine époque, la population du Saguenay-Lac-Saint-Jean a montré une fécondité très importante, ce qui a donné la possibilité à des gènes défectueux de pouvoir diffuser davantage. Le succès des nombreuses études menées sur cette population repose principalement sur son origine spécifique, sa stabilité et sur sa grande accessibilité grâce au projet BALSAC, une base de données généalogiques, dont Gérard Bouchard fut le grand fondateur en 1972⁷³.

4.2 Procédure

4.2.1 Approches

4.2.1.1 Approche par balayage génomique

L'approche par balayage génomique consiste à identifier systématiquement des milliers de marqueurs génétiques polymorphiques à travers le génome entier, entre autres à l'aide de *geneChips*. Ces marqueurs sont ensuite employés dans des analyses génétiques telles que des tests de liaisons et d'associations dans le but d'identifier les facteurs de risque génétique d'une pathologie. Cette méthode possède le bienfait de pouvoir identifier des régions chromosomiques dans lesquelles aucun gène candidat connu n'a encore été répertorié, ce qui apporte le potentiel de découvrir de nouveaux gènes⁷⁴. En choisissant cette approche, il n'est pas essentiel non plus de connaître la physiologie de la maladie, ce qui convient parfaitement à l'étude des maladies complexes. De plus, les résultats obtenus par un balayage génomique aident grandement à guider la recherche de gènes candidats en donnant une meilleure priorité aux candidats localisés dans des régions de liaisons significatives (par exemple un locus à trait quantitatif, QTL).

Le grand nombre de SNPs impliqués est néanmoins la source de certains désavantages tels que le temps de calcul requis pour traiter cette large quantité de SNPs avec la totalité des phénotypes disponibles et les coûts financiers élevés. Ce grand nombre de SNPs se traduit aussi par un nombre élevé de tests d'associations qui peut résulter certainement en un signal hasardeux puisque la probabilité d'obtenir des résultats significatifs par chance (faux positifs) augmente avec le nombre de tests effectués. Le grand nombre de résultats faux positifs rapportés dans les études provient de la situation dans laquelle la stratégie de déclarer une significativité statistique est basée uniquement sur la

valeur p^{75} . Colhoun *et al.* ont estimé que la fraction des résultats faux positifs obtenus dans les études d'associations entre une variante génétique et une maladie est d'au moins 0.95⁷⁶. D'où la nécessité d'appliquer un type de correction (Bonferroni, permutations) lors d'accomplissement de multiples tests de nature statistique sur les mêmes données. Ces tests permettent de se protéger d'une surinterprétation d'un résultat statistiquement significatif qui n'en n'est pas un en vérité.

4.2.1.2 Approche par gène candidat

L'approche par gène candidat a pour objectif l'identification des facteurs de risque génétique d'une maladie, en testant directement les effets des variants génétiques d'un gène pouvant potentiellement contribuer à la pathologie⁷⁷. Elle fournit donc l'avantage de pouvoir trouver des gènes ayant des effets modestes⁷⁷. Des études d'associations serviront à évaluer l'effet des variants sur le phénotype. Pour réaliser cette approche, seul un petit nombre de SNPs est requis, engendrant des coûts financiers raisonnables. Elle offre également le choix d'évaluer de manière critique les SNPs les plus représentatifs (ceux pouvant altérer la fonction protéinique) et les plus utiles pour le génotypage du gène investigué; cependant, elle nécessite avant tout que la séquence et la fonction du gène candidat soient très bien caractérisées de même que la biologie en arrière de la pathologie.

4.2.2 Analyses génétiques

4.2.2.1 Analyse de liaison génétique

Lorsqu'un gène, causant un désordre génétique, et un marqueur sont suffisamment proches sur le même chromosome et tendent à être hérités ensemble lors de la méiose, ils sont en liaison. Il est possible de retracer un gène responsable d'une maladie génétique, en utilisant cette notion de liaison, avec la méthode d'analyse de liaison génétique. Celle-ci permet d'étudier de manière statistique les allèles d'un locus et les phénotypes qu'ils engendrent. Pour se faire, des familles dans lesquelles la maladie est présente sont utilisées et une liaison entre des marqueurs génétiques et la maladie est recherchée. Un élément important affectant la liaison est la recombinaison génétique; plus un gène et un marqueur sont proches l'un de l'autre, moins il y a de chance qu'ils soient séparés par la recombinaison. La recombinaison est utilisée afin de déterminer la distance génétique entre le gène et le marqueur et elle se mesure en centimorgans (cM) où 1 cM correspond à 1% de

recombinaison. Dans ce type d'analyse, la présence de parents hétérozygotes est essentielle afin d'identifier s'il y a recombinaison ou non; de plus, les *single nucleotide polymorphism* (SNPs) y sont favorisés par rapport aux microsatellites en raison de leur grande densité, les rendant davantage utiles et informatifs. Une analyse de liaison génétique s'exprime en un score de logarithme des probabilités (score lod, Z). Un Z de trois ou plus indique une évidence de liaison et que la probabilité d'observer un pedigree donné si les deux loci ne sont pas liés est moins que 1 sur 1000. Un pic de liaison significative couvre souvent une large région, mais en général un gène pouvant expliquer 5% à 10% de la variance du trait se trouverait dans un intervalle de 40 ou 20 cM respectivement⁷⁸. Ces pics représentent en réalité des loci à trait quantitatif (QTLs), une région particulière du génome susceptible de contenir le gène ayant le réel impact sur le trait ciblé. Parce que la majorité des traits sont modulés par plus d'un gène, la recherche de QTL par analyse de liaisons est intéressante, car elle permet d'étudier un locus entier de gènes relié au trait investigué afin de déduire si ces gènes ont un effet significatif sur le phénotype. Par conséquent, un seul trait peut être associé à plusieurs QTLs sur le même chromosome ou sur des chromosomes distincts.

Les études de liaisons s'avèrent être souvent une étape préalable afin de découvrir des gènes candidats. Une fois un QTL identifié, il suffit d'utiliser, par exemple, un test d'association pour réduire et raffiner la région cible. Puis, suite à l'identification d'une région d'ADN d'intérêt, celle-ci peut être séquencée. Les séquences des gènes de cette région peuvent ensuite être comparées à une base de données d'ADN pour chercher des gènes pour lesquels la fonction est connue. Le nombre de QTLs relié à la variation du trait constitue une information importante qui nous renseigne sur l'architecture génétique du trait. Est-ce que ce trait est défini par plusieurs gènes ou non et ceux-ci ont-ils un effet modeste ou non. Puisque l'hypertension n'est pas définie par un mode de transmission particulier, une méthode d'analyse génétique non paramétrique doit être utilisée qui dépendra uniquement sur le fait que deux individus affectés posséderont des allèles prédisposant à la maladie en commun¹.

4.2.2.2 Étude d'association génétique

Les études d'association génétique peuvent être séparées en deux approches : les études basées sur les familles et celles basées sur la population. Pour la première catégorie, les tests de déséquilibre de transmission, qui détectent une association en présence de

liaison, sont employés. Ce test mesure la transmission démesurée d'un allèle des parents hétérozygotes aux enfants affectés.

La deuxième catégorie regroupe les études traditionnelles cas-contrôles. Ce type d'étude regarde la fréquence allélique entre un groupe d'individus contrôles et un groupe d'individus affectés par la maladie; dans ce cas, si un groupe d'allèles en particulier augmente en fréquence dans le groupe affecté, il est possible que ces allèles soient associés à la pathologie. L'avantage de ce type d'étude c'est qu'il est plus facile de recruter un grand nombre de personnes que les études basées sur des familles. Aucune supposition sur le mode de transmission de la maladie n'est nécessaire. Le problème majeur réside dans la possibilité d'avoir de fausses associations si la correspondance entre les cas et les contrôles n'est pas appropriée selon par exemple l'ethnicité des sujets.

Dans le projet proposé, cette avenue n'a pas été envisagée, car au départ notre échantillon n'a pas été conçu dans le but de pratiquer ce genre d'analyse. Il serait donc très difficile de séparer les sujets des familles en cas/contrôle et de faire la correspondance des sujets de ces deux groupes selon l'âge, leur sexe et leur exposition aux multiples facteurs environnementaux. Si le groupe cas comprend une collection hétérogène d'étiologies pour un désordre complexe, comme pour l'hypertension, la puissance à détecter une association significative est aussi réduite⁷⁹. Concernant le groupe contrôle, celui-ci doit inclure des sujets pour lesquels il faut avoir la certitude qu'ils sont non affectés. Dans l'étude de l'hypertension, plusieurs traits complexes représentent des facteurs de risque pour son développement. Le risque d'assigner inadéquatement le statut non affecté à un sujet est grand et réduit la puissance de détection. En effet, au moment du recrutement, les personnes du groupe contrôle ne sont peut-être pas encore diagnostiquées, mais comportent des symptômes (phénotypes intermédiaires) de l'hypertension autres qu'une pression artérielle élevée. Aussi, si les sujets du groupe contrôle sont exposés à des facteurs environnementaux impliqués dans l'hypertension, il n'est pas clair dans quel groupe ceux-ci doivent être placés. Certes, à cause de la complexité de l'hypertension et de la particularité de notre échantillon (grandes familles), une étude cas-contrôle impliquerait trop de sources de confusion pour être réalisée et pouvoir rapporter des associations valides.

À partir d'un test d'association, il est aussi possible aussi d'évaluer la relation phénotype/génotype. Dans ce cas, la différence entre les moyennes des valeurs phénotypiques des sujets, regroupés selon les allèles du marqueur, est considérée. L'étude d'association peut se mesurer de différentes façons : le *Odds ratio*, le risque relatif et la valeur p. Concernant le *Odds ratio*, il s'agit de faire le rapport du nombre de sujets affectés avec l'allèle multiplié par le nombre de contrôles sans l'allèle divisé par le nombre de sujets affectés sans l'allèle multiplié par le nombre de contrôles avec l'allèle. Un *Odds ratio* égale à un signifie que la fréquence allélique est la même pour le groupe contrôle et le groupe affecté. Si la valeur du rapport est supérieure à un, cela veut dire que la chance d'avoir la maladie est ce nombre de fois plus élevé chez les personnes ayant l'allèle. Le risque relatif, pour sa part, compare le risque de développer la maladie lorsqu'on a l'allèle relativement au risque si on ne l'a pas. Si le risque relatif est inférieur à un, l'allèle confère une protection contre le développement de la maladie. Concernant la valeur p, celle-ci est définie comme étant la probabilité qu'une variable assume une valeur plus grande ou égale à celle observée et ce, uniquement par chance. Dans le projet présenté, l'absence d'association correspond à l'hypothèse nulle (H_0) et la présence d'association entre un marqueur génétique et un phénotype est désignée par l'hypothèse alternative (H_1). Généralement, l' H_0 est rejetée si la valeur p est inférieure ou égale à un certain niveau de significativité, représenté par la lettre grecque α , fixé dans notre cas à 5%, c'est-à-dire une chance sur 20 d'avoir tort et de rejeter faussement l' H_0 .

4.2.3 Impact de l'hémizygoté de l'homme sur les études génétiques

Le fait que l'homme soit hémizygoté pour le chromosome X vient simplifier certaines étapes de ces approches. Pour l'approche par gène candidat, puisque l'homme possède un seul chromosome X, cela facilite la charge de travail liée au génotypage. De leur côté, les études de balayage du chromosome X, à l'aide de blocs d'haplotypes, incorporent souvent exclusivement les hommes, car étant hémizygotés, leurs haplotypes sont déjà phasés⁸⁰ et peuvent être extraits directement de leurs génotypes. De plus, l'évidence de transmission de ce chromosome, du père à sa fille, simplifie la conception des logiciels de génétique. Avec un seul patron de transmission possible, la complexité d'incorporation du chromosome X dans les algorithmes génétiques est largement diminuée. Puis, pour les tests d'associations à partir des SNPs, les hommes représentent toujours strictement deux groupes alléliques à comparer; tandis que pour les femmes, la présence

d'hétérozygotie crée un groupe génotypique de plus à comparer, requérant un plus grand nombre de femmes pour les échantillons et de tests statistiques à accomplir.

4.3 Marqueurs génétiques

4.3.1 SNPs

Le SNP a été utilisé comme marqueur génétique principal lors de ce projet (Figure 10). Il s'agit d'un changement d'une base survenant dans plus de 1% de la population. Durant les dernières années, l'utilisation des SNPs afin d'effectuer des analyses génétiques est devenue davantage présente. Plusieurs caractéristiques favorisent leur sélection: ils sont hérités de manière stable, ils sont fréquents (1/1000 pb) et uniformément distribués à travers le génome¹. Ce type de marqueur fournit un génotypage automatisé rapide et son impact sur l'expression des gènes⁸¹ ainsi que dans les maladies complexes a été prouvé⁸²⁻⁸⁵.

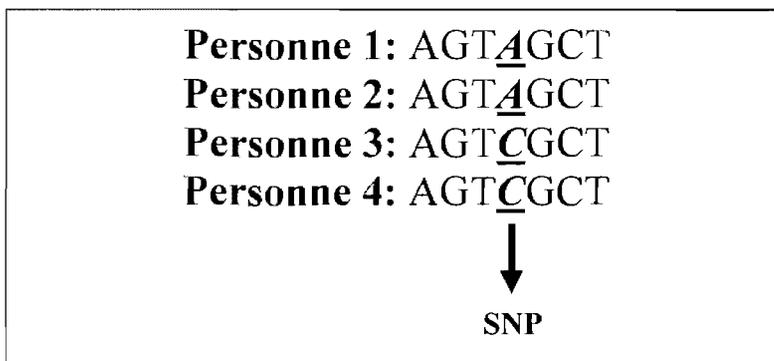


Figure 10 Représentation schématique d'un SNP

4.3.1.1 Équilibre d'Hardy-Weinberg

La loi d'Hardy-Weinberg (HW), utilisée en génétique des populations, lie la fréquence allélique au génotype¹. Cette loi précise qu'à la suite d'une génération de reproduction aléatoire et sans accumulation de mutation, la fréquence des génotypes à un locus génétique deviendra fixée à une valeur en équilibre. Il est donc possible, au sein d'une population respectant certaines conditions, de calculer la relation entre la fréquence allélique et le génotype à l'aide d'une équation mathématique simple. Pour un locus composé des allèles A et a, avec une fréquence allélique de p et q, ce principe prédit que la fréquence d'un génotype homozygote AA et aa est de p^2 et q^2 et que celle d'un génotype hétérozygote Aa est de $2pq$. La première propriété de la loi d'HW postule que la fréquence

de ces trois génotypes est donnée par les termes d'une expansion binomiale, $(p+q)^2 = p^2 + 2pq + q^2$. La deuxième propriété spécifie que peu importe les fréquences alléliques présentes dans la population, les fréquences génotypiques seront $p^2:2pq:q^2$ et ces dernières resteront constantes de génération en génération aussi longtemps que les fréquences alléliques restent constantes¹.

Les violations des considérations d'HW peuvent causer des déviations dans les prévisions attendues. Une de ces considérations est un accouplement aléatoire et la transgresser entraînerait une déviation dans les fréquences génotypiques de la population. Trois transgressions touchent l'accouplement aléatoire: la consanguinité, l'accouplement assorti et la structure de la population. En effet, la consanguinité cause une augmentation de l'homozygoté pour tous les gènes. De son côté, l'accouplement assorti, défini par le choix d'un compagnon selon s'il possède certains traits particuliers, a pour conséquence d'augmenter les génotypes homozygotes, car la tendance est de choisir quelqu'un qui nous ressemble. Concernant la population, il faut que celle-ci soit exempte de stratification c'est-à-dire qu'il y ait absence de sous-groupe distinct génétiquement. Lorsque la sélection d'un compagnon, à l'intérieur d'une population, est restreinte aux membres d'un sous-groupe particulier il résulte un excès d'homozygotes dans cette population. Il peut aussi se produire une fluctuation aléatoire des fréquences génotypiques dans une population de petite taille, phénomène appelé dérive génétique (*genetic drift*). Ces trois violations résultent également en une élévation de la fréquence des maladies autosomiques récessives.

Les autres considérations d'HW à respecter affectent les fréquences alléliques, soient: un taux de mutation négligeable, aucune sélection contre un génotype particulier et aucune immigration. Comme les taux de mutations sont de l'ordre de 10^{-4} à 10^{-8} , ils auront donc un effet peu appréciable sur les fréquences alléliques. Pour sa part, la sélection causera aussi un changement au niveau des fréquences alléliques. Si la sélection mène à la perte d'allèles à l'exception de celle favorisée, il sera question de sélection directionnelle, mais si la sélection aboutit à un équilibre sans la perte d'allèles il s'agira d'une sélection de balancement. De plus, si au sein d'une population, il y a une immigration significative d'individus en provenance d'une population avec des fréquences alléliques différentes de celles de la population endogène, il est certain que ce phénomène modifiera la fréquence des allèles. En général, les changements apportés par les mutations, la sélection et

l'immigration causent une déviation moindre de l'équilibre d'HW que ceux concernant l'accouplement aléatoire, car ceux-ci se font lentement et par petites incréments.

La déviation de l'équilibre d'HW est devenue un test accepté pour détecter les erreurs de génotypage⁸⁶; néanmoins, il faut noter qu'une déviation ne signifie pas nécessairement une erreur de génotypage, mais qu'elle peut survenir par chance ou en raison d'autres facteurs génétiques.

4.3.2 Blocs d'haplotypes

Les SNPs peuvent être regroupés en des blocs d'haplotypes, un ensemble de marqueurs proches et présents sur le même chromosome qui est hérité ensemble. Parce qu'ils sont formés par des marqueurs étroitement liés, les blocs d'haplotypes sont plus informatifs que les marqueurs individuels^{87, 88}. Même si c'est un variant rare, un bloc d'haplotypes peut se révéler primordial dans l'étude des maladies complexes⁸⁹. Comme les SNPs, la construction des blocs d'haplotypes, a été privilégiée par de nombreuses études et montre une implication significative dans les maladies et la modulation des gènes⁹⁰.

4.3.3 Correction pour test multiple

4.3.3.1 Correction de Bonferroni

L'accomplissement de multiples tests de nature statistique sur les mêmes données peut sans doute mener à des résultats significatifs uniquement par chance, d'où la nécessité d'appliquer un type de correction. La méthode de Bonferroni, la plus simple et la plus fréquemment considérée, s'avère être une protection contre ce phénomène. Si le seuil de significativité choisi est à 0,05, il suffit de vérifier si la valeur obtenue, en multipliant la valeur p observée par le test avec le nombre de tests, se trouve sous ou au-dessus du seuil. Si la valeur de la multiplication est sous le seuil de 0,05, alors il peut être conclu que la valeur p observée est réellement significative. Tout en étant rigoureuse, elle est aussi très conservatrice et pourrait rejeter de bons résultats⁹¹ puisqu'elle obtient le plus haut taux de résultats dits faux négatifs. La correction de Bonferroni est une procédure à une seule étape dans laquelle chaque valeur p est corrigée de manière indépendante; d'autre part, si les hypothèses sont corrélées comme c'est le cas lorsque deux SNPs sont en déséquilibre de liaison (LD), cette correction sera beaucoup trop sévère.

4.3.3.2 Correction par permutation

Une autre alternative est la correction par permutations qui s'accomplit à l'aide de plusieurs étapes. Cette correction prendra avantage de la dépendance entre ces SNPs en permutant tous ces SNPs au même moment. Au départ, les valeurs p sont calculées pour chaque SNP à partir de l'ensemble de données originales et sont classées. Pour la valeur p la plus basse, il faut connaître la probabilité d'obtenir cette valeur quand autant de tests sont effectués. La méthode de permutation crée alors des ensembles de données artificielles pour tous les SNPs et une valeur p est calculée pour ces pseudo-ensembles. Le minimum des nouvelles valeurs P est retenu et comparé aux originales. Ce processus est répété des milliers de fois et la valeur p ajustée est définie par la proportion des pseudo-ensembles pour lesquels la pseudo p valeur minimale est plus petite que celle originale⁹². Cette méthode comporte toutefois certaines limites; effectivement, dû aux nombreuses répétitions d'analyses, elle est lente et les calculs informatiques requis sont très demandant voir même impossible à réaliser dans certains cas⁹³.

Méthodologie

Chapitre 6 – Outils et données utilisés

6.1 Sujets

Dans le but d'étudier les déterminants génétiques et environnementaux de la pathogenèse de l'hypertension, des données génotypiques et phénotypiques ont été recueillies préalablement par des équipes médicales du centre hospitalier de l'Université de Montréal et du complexe hospitalier de La Sagamie à Chicoutimi. Ces données ont été obtenues chez des familles canadiennes-françaises originaires de la région du SLSJ du Québec. Au total, 120 familles (900 sujets) ont été sélectionnées, car elles respectaient le critère d'avoir au moins deux enfants, de mêmes parents, atteints d'hypertension et de dyslipidémie avant l'âge de 55 ans³⁸. De cet échantillon, des techniciens du laboratoire ont génotypé 468 sujets avec des *geneChips* d'Affymetrix 50K. La totalité des sujets n'a pu être génotypée à cause du coût financier des *geneChips* d'Affymetrix. Étant donné qu'il s'agissait à ce moment là d'une nouvelle technologie, le coût de génotypage s'élevait à 1000\$ par personne. Les caractéristiques physiologiques de ces sujets sont résumées dans le tableau IV. Puis, avant l'achat des *geneChips* d'Affymetrix, 769 sujets ont été génotypés pour deux SNPs localisés dans le gène ACE2 à l'aide de la technique d'amplification en chaîne par polymérase (PCR) par un membre du laboratoire. Dans ce temps, les 900 individus n'avaient pas tous été recrutés, car le processus de sélection progressive des sujets était toujours en cours. C'est après avoir trouvé deux individus d'une famille respectant les critères mentionnés ci-haut que les proches étaient ensuite contactés pour solliciter leur participation à l'étude. Tous les sujets de cette étude ont lu et signé le formulaire de consentement. Toutes les informations sur les sujets restent confidentielles et ont été enregistrées, par des informaticiens du laboratoire de recherche, dans une base de données non nominative et codée.

Tableau IV Caractéristiques biologiques, classées par le sexe, des 468 sujets génotypés

| Caractéristiques | 2 sexes | | Femmes (n=250) | | Hommes (n=218) | |
|---|---------|------|-------------------|------|-------------------|------|
| | Moyenne | SD | Moyenne | SD | Moyenne | SD |
| Âge | 47 | 14 | 49 | 15 | 45 | 12 |
| PAS (mmHg) | 130 | 20 | 130 | 23 | 130 | 17 |
| PAD (mmHg) | 82 | 12 | 79 | 12 | 85 | 11 |
| Hauteur (cm) | 166 | 10 | 159 | 6 | 173 | 7 |
| Poids (kg) | 73 | 16 | 66 | 15 | 82 | 14 |
| Indice de masse corporel (kg/m ²) | 27 | 5 | 26 | 6 | 27 | 4 |
| Rapport taille-hanches | 0,9 | 0,1 | 0,8 | 0,1 | 1 | 0,07 |
| Circonférence de la taille (cm) | 90 | 14 | 84 | 15 | 96 | 11 |
| Circonférence des hanches (cm) | 100 | 8 | 100 | 10 | 100 | 7 |
| Circonférence de la cuisse au niveau proximal (cm) | 59 | 5 | 59 | 7 | 58 | 6 |
| Circonférence de la cuisse hauteur mi-cuisse (cm) | 53 | 6 | 53 | 6 | 54 | 6 |
| Circonférence de la cuisse au niveau distal (cm) | 40 | 5 | 40 | 5 | 41 | 4 |
| Pli cutané sur biceps (mm) | 21 | 14 | 25 | 14 | 17 | 12 |
| Pli cutané sur triceps (mm) | 29 | 14 | 33 | 13 | 24 | 13 |
| Pli cutané sur omoplate (mm) | 26 | 13 | 27 | 14 | 25 | 11 |
| Pli cutané niveau point iliaque supérieur (mm) | 25 | 13 | 26 | 13 | 25 | 13 |
| Pli cutané sur cuisse (mm) | 34 | 16 | 42 | 14 | 27 | 16 |
| Gras corporel (%) | 32 | 9 | 38 | 8 | 26 | 6 |
| Cholestérol total (mmol/litre) | 5,4 | 1,2 | 5,4 | 1,2 | 5,4 | 1,1 |
| HDL (mmol/ litre) | 1,3 | 0,4 | 1,5 | 0,5 | 1,2 | 0,3 |
| LDL (mmol/ litre) | 3,3 | 1,0 | 3,2 | 1,1 | 3,5 | 1,0 |
| Triglycérides (mmol/ litre) | 1,8 | 1,3 | 1,6 | 1,0 | 2,1 | 1,6 |
| Glucose (mmol/ litre) | 5,3 | 1,3 | 5,2 | 1,2 | 5,5 | 1,3 |
| Créatinine (µmol/ litre) | 78 | 18 | 69 | 15 | 88 | 16 |
| Acide urique (µmol/ litre) | 304 | 80 | 266 | 70 | 348 | 69 |
| Glycérol (mmol/ litre) | 0,07 | 0,05 | 0,09 | 0,05 | 0,06 | 0,05 |
| Réactance de l'eau extracellulaire (Ω) | 633 | 96 | 691 | 86 | 574 | 65 |
| Réactance de l'eau intracellulaire (Ω) | 1207 | 290 | 1389 | 266 | 1025 | 175 |
| Insuline à jeun (pmol/litre) | 113 | 47 | 107 | 41 | 121 | 53 |
| Arginine-vasopressine (pg/ml) | 4,7 | 1,4 | 4,8 | 1,4 | 4,5 | 1,4 |
| Rénine (ng/ml/h) | 1,9 | 1,5 | 2,0 | 1,6 | 1,9 | 1,4 |
| Gamma glutamyl transférase (U/litre) | 26 | 25 | 18 | 15 | 34 | 30 |
| Indice Milwaukee | 28 | 13 | 32 | 13 | 24 | 11 |
| Indice Heather | 6,1 | 2,5 | 7,4 | 2,4 | 4,9 | 1,9 |
| Fréquence cardiaque (n/min) | 70 | 12 | 74 | 12 | 67 | 11 |
| PAS durant le jour (mmHg) | 126 | 16 | 122 | 16 | 130 | 15 |
| PAD durant le jour (mmHg) | 77 | 10 | 75 | 10 | 80 | 10 |
| PAS durant la nuit (mmHg) | 113 | 15 | 109 | 15 | 117 | 14 |
| PAD durant la nuit (mmHg) | 68 | 11 | 65 | 11 | 71 | 10 |
| PAS 24hrs (mmHg) | 119 | 15 | 115 | 15 | 123 | 14 |
| PAD 24hrs (mmHg) | 72 | 10 | 70 | 10 | 75 | 10 |
| Indice de masse ventriculaire gauche (g/m ²) | 1,7 | 0,5 | 1,6 | 0,4 | 1,8 | 0,5 |
| Déflexion maximale d'ICG différencié (mm) | 1,3 | 0,4 | 1,5 | 0,4 | 1,1 | 0,4 |

6.2 Phénotypes

La description, ci-dessous, du prélèvement des phénotypes provient de l'étude de Kotchen *et al.* portant sur les déterminants génétiques de l'hypertension et l'identification de phénotypes candidats⁹⁴. Les données phénotypiques ont été prélevées antérieurement par les équipes médicales du centre hospitalier de l'Université de Montréal et du complexe hospitalier de La Sagamie à Chicoutimi. Le statut d'hypertension fut attribué au sujet si ses mesures standards de pression artérielle étaient $> 140/90$ mmHg à deux occasions ou s'ils prenaient des médicaments antihypertenseurs. Un taux de cholestérol élevé fut défini par un niveau de cholestérol dans le sérum > 5.2 mmol/L ou la prise de médicaments hypolipidémiant. L'histoire familiale a été recueillie afin de valider la relation entre frères et sœurs. Toutes personnes présentant un diabète sucré, une concentration dans le sérum de créatine $> 2,2$ mg/dL, un abus de substance comme l'alcool et une relation de demi-frère/sœur furent exclues de l'étude. Pour les participants qualifiés, des échantillons de sang veineux ont été obtenus afin d'extraire l'ADN des leucocytes et les mensurations anthropomorphiques suivantes ont été obtenues : hauteur, poids, circonférence des hanches et de la taille et l'épaisseur des plis cutanés à plusieurs endroits. Les sujets éligibles ont été considérés comme candidats potentiels pour un protocole de deux jours en milieu hospitalier pour des mesures phénotypiques additionnelles. Avant la sélection des sujets, pour ce protocole, un examen physique a été fait. D'autres critères d'exclusion furent appliqués incluant : une pression diastolique > 100 mmHg durant la prise de médication, un indice de masse corporelle > 34 kg/m², une grossesse, une tumeur maligne et un infarctus du myocarde ou un accident cérébrovasculaire dans les six mois précédant. Avant le début du protocole de deux jours, la médication antihypertensive fut arrêtée durant au moins une semaine et celle hypolipidémiant pour un mois. L'après-midi avant le jour 1, les sujets ont été mis sur une diète permettant le maintien du poids (150 mEq Na⁺/80 mEq K⁺).

Durant le jour 1, des cathéters intraveineux ont été insérés dans les deux bras : un pour l'échantillon de sang et l'autre pour des infusions. Les mensurations de base incluant les mesures standards de pression artérielle, un échocardiogramme à 2 dimensions, des mesures à jeun de concentrations lipidiques, de glucose et d'insuline ont été prises. Du sang du sujet à jeun, anticoagulé avec de l'héparine, a été obtenu pour évaluer le contre-transport Na⁺-Li⁺ et le co-transport Na⁺-K⁺ des globules rouges. Dans une salle tranquille, les

mesures suivantes ont été prises après que le sujet fut couché pour 60 minutes et après qu'il fut 10 minutes en position debout : les mesures standardisées et répétées de la pression artérielle, le débit cardiaque et le débit systolique, des mesures de l'activité de rénine plasmatique et d'aldostérone plasmatique, la norépinéphrine et l'épinéphrine, la cAMP et la cGMP. Des mesures hémodynamiques ont été répétées après que les sujets aient été assis pour 30 minutes et encore après un test de stress mental, de deux minutes, dans lequel les sujets sont amenés à résoudre rapidement une série de problèmes arithmétiques de plus en plus difficiles. La tension artérielle a été mesurée par un appareil automatisé de type *Dinamap*; *Johnson & Johnson Medical, Inc*, le débit cardiaque et le débit systolique par une pléthysmographie à impédance par le *Sorba Medical Systems, Inc*. L'index cardiaque a été calculé comme étant le débit cardiaque/l'air de la surface corporelle et la résistance périphérique totale a été calculée comme étant la moyenne de la pression artérielle/débit cardiaque. À la suite de l'exécution de ces protocoles, la sensibilité à l'insuline fut évaluée avec la méthode d'échantillonnage réduite du modèle de Bergman avec injection d'insuline. Débuté au jour 1 pour une période de 24 heures, la pression artérielle fut mesurée toutes les 20 minutes durant le jour et toutes les 45 minutes durant la nuit avec un moniteur *Accutacker II* de *Sun Tech Medical Instruments, Inc*.

Au jour 2, la pression artérielle, le débit sanguin rénal (clairance PAH) et le taux de filtration glomérulaire (clairance d'inuline) furent mesurés en réponse à deux taux infusionnels de norépinéphrine ($0,01$ et $0,025 \mu\text{g} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$) pour 30 minutes chaque. En même temps que ces mesures, les sujets ont reçu un volume d'eau oralement de 15 mL/kg et ensuite 850 mL de solution physiologique salée IV durant les six heures qui ont suivi. Les concentrations plasmatiques d'arginine et de vasopressine ont été prises avant et après la provocation des fluides. Durant ce même jour 2, la clairance de lithium a été déterminée avec l'administration de 600 mg de carbonate de lithium PO à 11 heure PM dans la nuit précédente et les mesures d'excrétion de lithium plasmatique et dans l'urine durant 2 heures ont été recueillies.

Un total de 1650 phénotypes est accessible à partir d'une base de données nommée *Cardiogen* dont 1321 phénotypes principaux et 329 phénotypes dérivés de ceux-ci. Des informations, sur l'environnement tel que le tabagisme et la consommation d'alcool, sont

aussi disponibles. Pour le projet, des phénotypes de pression artérielle ainsi que 253 phénotypes liés spécifiquement au sexe ont été retenus. Ces derniers ont été classés dans cette catégorie si la différence entre le score LOD de la femme et de l'homme est plus grande ou égale à deux, à une position donnée sur un chromosome⁹⁵.

Étant donné que les phénotypes indice Heather et indice Milwaukee ont soulevé des résultats statistiquement significatifs lors du projet, ils méritent qu'on les définisse davantage. L'impédance cardiographique est utilisée afin d'évaluer les propriétés du flux sanguin au niveau du thorax. Elle représente précisément le changement dans le temps (dZ/dt) de l'impédance transthoracique (Z, Ω) durant le cycle cardiaque. Le ratio dZ/dt se mesure à partir d'une série d'électrodes qui détectent les changements électriques et de résistance dans le thorax du sujet. Ensuite, ces données, en plus des informations provenant de l'électrocardiogramme (ECG), servent à calculer des paramètres hémodynamiques et de contractilité myocardiale incluant l'indice Heather et l'indice Milwaukee. L'indice Heather est obtenu des tracés de l'ECG et de dZ/dt et est égal à $dZ/dt_{max}/QZ1$, où dZ/dt_{max} est la déflexion maximale de l'ondulation initiale et $QZ1$ est le temps depuis le début de l'onde Q (dans le ECG) jusqu'au pic dZ/dt . Pour sa part, l'indice Milwaukee se calcule en faisant $dZ/dt_{max}/REP$, où REP est le temps depuis le début de l'éjection ventriculaire jusqu'au pic dZ/dt .

6.3 Génotypes

La base de données du laboratoire contient différents types de marqueurs génétiques, mais les SNPs ont été choisis pour la réalisation de ce projet. Pour le groupe contenant les 468 sujets, la recherche de SNPs a été réalisée à l'aide de *geneChips* 50K d'Affymétrie, par des techniciens du laboratoire, menant à l'identification de 58 000 SNPs pour tous les chromosomes dont 1204 SNPs sur le chromosome X. Pour les 769 sujets, deux SNPs, localisés dans l'intron 16 du gène *ACE2*, ont été déterminés par la technique de polymérase en chaîne (rs2074192 : T/C, rs233575 : A/G) par un chercheur du laboratoire. Ces SNPs ont permis de former trois haplotypes : T-A, C-G et C-A.

6.3.1 Construction des blocs d'haplotypes

J'ai réalisé la construction du bloc d'haplotypes pour les deux SNPs d'ACE2 en utilisant le paramètre D' de LD, défini comme étant l'association non aléatoire d'allèle. Le programme EMLD existe déjà pour cette tâche (<http://request.mdacc.tmc.edu/~qhuang/Software/pub.htm>). EMLD débute en calculant les fréquences d'haplotypes à l'aide de l'algorithme *expectation-maximisation* pour seulement deux SNPs; après, il trouve les valeurs des paramètres du LD. Les allèles manquants ne sont pas incorporés dans les calculs. Malheureusement, ce programme est construit pour fonctionner avec les autosomes. En effet, EMLD requiert en entrée absolument deux génotypes pour chaque sujet; cependant, dans le cas du chromosome X, les hommes sont hémizygotés. Cette situation incorpore un biais dans le calcul du paramètre D' qui s'effectue à partir de la fréquence allélique, en attribuant un nombre de génotypes pour l'homme plus élevé que supposé. Pour résoudre ce problème, j'ai apporté une modification au programme EMLD afin que celui-ci considère un seul génotype pour l'homme. La définition employée, pour construire le bloc d'haplotypes, a été tirée de l'article de Gabriel *et al.*⁹⁶. Pour les 1204 SNPs, le programme Haploview (version 4.0) qui accepte maintenant le chromosome X a été utilisé⁹⁷. La définition employée pour construire les blocs d'haplotypes fut encore celle de Gabriel *et al.*⁹⁶.

6.4 Tests statistiques

J'ai effectué des analyses de liaisons et d'associations à l'aide de différents logiciels. Acceptant le chromosome X via l'option *minx* (*Merlin in X*), le programme *multipoint engine for rapid likelihood inference* (Merlin) a été utilisé pour accomplir des analyses de liaisons non paramétriques basées sur les familles⁹⁸. Ce logiciel utilise des arbres d'héritabilité partielle pour représenter le flux de gènes à l'intérieur d'un pedigree. Pour les phénotypes continus, l'option *variance components* ou *quantitative trait loci* a été choisie selon la normalité de la distribution phénotypique. L'utilisation de l'option *non-parametric linkage* et *pairs* pour les phénotypes binaires n'a malheureusement pu aboutir à des résultats faute d'avoir des familles informatives.

Devant ces limites, j'ai décidé de sélectionner le programme *X-linked reconstruction-combination transmission/disequilibrium test* (XRC-TDT). Il s'agit de la

première méthode d'association basée sur des familles qui teste des marqueurs du chromosome X spécifiquement. Ce programme teste pour une association entre un marqueur et une maladie, en présence de déséquilibre de liaison entre les deux loci, lorsque l'échantillon consiste en des familles nucléaires pour lesquelles il y a de l'information génotypique incomplète pour les parents. Il accepte, cependant, uniquement les phénotypes qualitatifs et donc, il requiert une transformation des phénotypes continus en phénotypes binaires. Selon Royston *et al.*, cette dichotomisation peut être vue comme une forme d'arrondissement extrême des valeurs quantitatives, engendrant inévitablement une perte d'information et de puissance⁹⁹. Ces auteurs mentionnent également que la dichotomisation équivaut à perdre un tiers des données avec une sérieuse perte de la puissance à détecter une vraie relation et participe à augmenter la probabilité de résultats faux positifs⁹⁹. J'ai transformé les phénotypes continus sous forme binaire en utilisant la médiane comme seuil d'affectation.

Afin d'évaluer la présence d'une relation génotype/phénotype, j'ai analysé les SNPs ayant montrés des résultats significatifs avec le programme XRC-TDT. Puisque l'échantillon est constitué de familles et par ce fait de données corrélées, un test Anova n'est pas approprié. J'ai donc choisi la méthode *generalized estimating equations* (GEE)^{100, 101} qui fournit un outil de régression alternative pour analyser ce type d'échantillon. Les familles n'ont pas à être séparées et des phénotypes binaires aussi bien que continus peuvent être incorporés. J'ai accompli ces analyses avec la version 9.1 du logiciel statistique SAS à l'aide de la procédure GENMOD. Étant donné que la méthode GEE prend en considération une distribution normale des phénotypes, j'ai dû appliquer un logarithme et une racine carrée sur certains phénotypes.

Pour prendre en considération les multiples tests accomplis, la correction de Bonferroni peut être appliquée. Concernant les 1204 SNPs, basé sur l'équilibre d'Hardy-Weinberg, la fréquence de l'allèle mineur et le taux de réussite du génotypage, 445 SNPs sur 1204 SNPs ont été éliminés. Alors, les résultats seront significatifs si la valeur p obtenue est inférieure à 1E-06 et il sera question d'associations suggestives si la valeur p se situe entre 0.05 et 1E-06. Pour le bloc d'haplotypes d'ACE2, les résultats seront significatifs si la valeur p obtenue est inférieure à 1E-03 et il y aura associations

suggestives si la valeur p se trouve entre 0.05 et $1E-03$. J'ai exécuté les analyses chez les deux sexes ensemble et séparément afin de respecter l'architecture du chromosome X. Une brève description des logiciels utilisés lors du projet est présentée dans le tableau V et un résumé de la méthode est schématisé à la figure 11.

Tableau V Description des différents logiciels utilisés lors du projet

| Test | Description | Caractéristique |
|---------|---|---|
| Anova | Analyse de variance (test d'association) | <ul style="list-style-type: none"> ▪ Prend en considération que l'échantillon soumis est constitué de personnes indépendantes et non de familles |
| Merlin | Analyse de liaison non-paramétrique basée sur les familles | <ul style="list-style-type: none"> ▪ Accepte les phénotypes continus et binaires ▪ L'option pour le chromosome X (minx) a été peu testée ▪ Requiert une division des familles en familles nucléaires ▪ En divisant les familles de diverses façons, des résultats différents sont obtenus |
| XRC-TDT | Première méthode d'analyse d'association, en présence de liaison, basée sur les familles qui teste spécifiquement les marqueurs du chromosome X | <ul style="list-style-type: none"> ▪ Requiert une division des familles en familles nucléaires ▪ Accepte uniquement les phénotypes qualitatifs ▪ Prend une heure pour analyser un phénotype sur 1204 SNPs |
| GEE | Test d'association sur les familles | <ul style="list-style-type: none"> ▪ Ne requiert pas une division des familles ▪ Phénotypes binaires et continus acceptés ▪ Pour les phénotypes continus, assume une distribution normale |

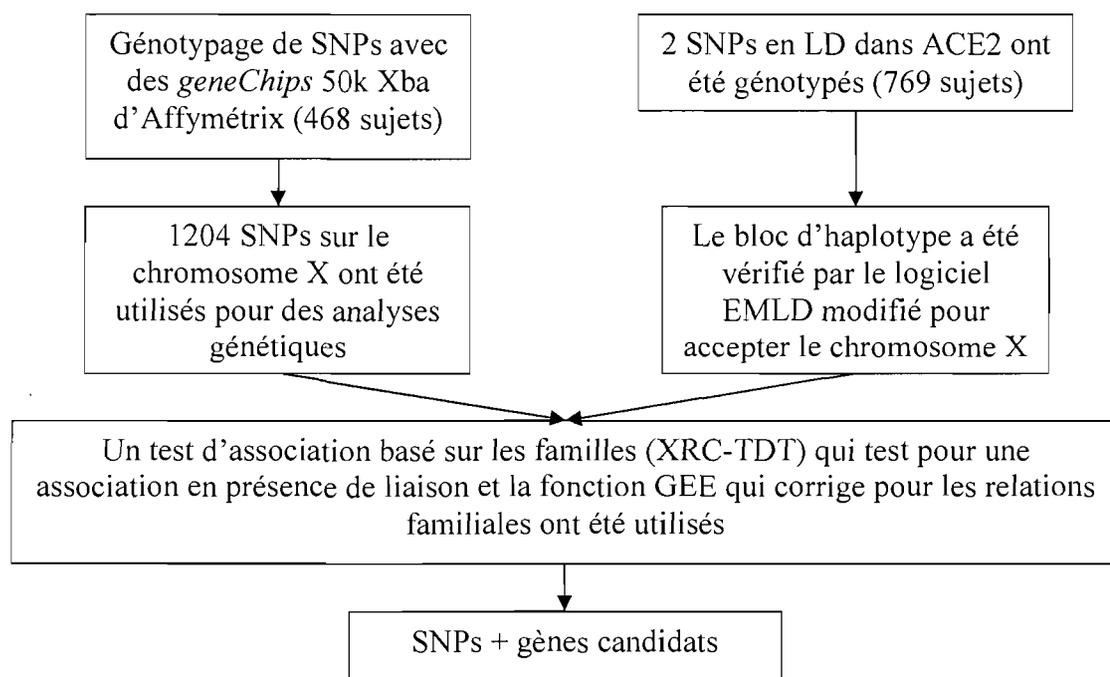


Figure 11 Résumé de la méthodologie utilisée

6.5 Impact des SNPs et conservation de séquences entre espèces

Dans le but de prédire l'impact des SNPs identifiés et de connaître la conservation de leurs séquences et des gènes qui y sont associés, des outils bio-informatiques ont été employés dont certains sont mentionnés dans les articles de Uzun *et al.*¹⁰² et Bhatti *et al.*¹⁰³. Il s'agit des logiciels MAPPER¹⁰⁴, SNAP¹⁰⁵ et le navigateur Web UCSC^{106, 107}. En plus de ces derniers, F-SNP¹⁰⁸ et MatInspector (SNPInspector)^{109, 110} ont également été utilisés. Bien entendu, les logiciels focalisant sur les SNPs codants (par exemple, les SNPs non synonymes) ont été mis de côté dû au fait que les SNPs identifiés lors de ce projet se situent uniquement au niveau des introns et des régions intergéniques.

6.5.1 Prédiction de la fonction des SNPs

Pour prédire l'impact fonctionnel des SNPs non codants, les programmes MAPPER (*Multi-genome Analysis of Positions and Patterns of Elements of Regulation*), F-SNP (*functional SNP*) et MatInspector ont été employés. Ces programmes ont comme objectif commun la recherche de sites de liaison pour les facteurs de transcription (TFBS) dans une séquence d'ADN. Néanmoins, il n'y a pas une forte évidence que ces TFBS soient vraiment fonctionnelles, mais ceci apporte du moins la possibilité qu'un SNP non codant

puisse influencer la régulation de l'expression génique. MAPPER représente une plateforme d'identification computationnelle de TFBS dans plusieurs génomes et ce, à l'intérieur du gène choisi par l'utilisateur. F-SNP, une base de données, a également recours à des programmes permettant de prédire les TFBS tels que TFSearch¹¹¹ et Consite¹¹² en plus de fournir des informations sur les effets fonctionnels des SNPs à l'aide de 16 outils bio-informatiques et base de données. F-SNP analyse la séquence du SNP plutôt que celle du gène entier. Similaire aux deux programmes précédents, MatInspector (SNPInspector) est un logiciel utilisant une large librairie de descriptions matricielles pour les TFBS afin de localiser des appariements dans les séquences d'ADN tout en se basant sur la séquence qui flanque le SNP.

Des outils bio-informatiques ont aussi été cherchés dans le but d'évaluer l'impact des SNPs d'intérêt au niveau de l'épissage. Plusieurs programmes existent à cet effet dont SNAP (*integrated SNP annotation platform*). Malheureusement, ces programmes sont valides uniquement sur des SNPs situés dans les exons. En effet, ils prédisent des sites pour les amplificateurs ou les silencers d'épissage exonique (*Exonic Splicing Enhancers/Silencers*; ESE et ESS) ou évaluent la séquence de jonction exon-intron afin de vérifier si elle représenterait un site donneur ou accepteur adéquat. Comme les ESS/ESE se trouvent exclusivement dans les exons et que les SNPs identifiés ne sont pas localisés dans une jonction exon-intron, l'impact de nos SNPs sur l'épissage n'a pu être évalué.

6.5.1 Conservation des séquences entre différentes espèces

Des outils bio-informatiques sont disponibles afin d'examiner la conservation des régions d'intérêts entre les espèces. Il en existe plusieurs dont : MutDB¹¹³, ECR browser¹¹⁴, Vista¹¹⁵, Ensembl¹¹⁶, PupaSuite¹¹⁷ et UCSC^{106, 107}. Le navigateur Web UCSC est largement utilisé dans la population scientifique et contient une énorme collection de génomes sur lesquels plusieurs éléments peuvent être cherchés et étudiés visuellement comme les séquences, les variations, les répétitions, l'expression, les transcrits, les gènes, les maladies associées à la région ciblée, la génomique comparative, etc. De plus, il est facile et simple d'utilisation, il est complet, il offre une belle visualisation du patron de conservation de séquence entre les espèces et il comporte plusieurs options pour exécuter les analyses souhaitées. UCSC comporte une section de génomique comparative dans laquelle se trouve

l'option de conservation qui est définie par deux analyses : un alignement multiple (Multiz¹¹⁸) et des mesures de conservation dans l'évolution (PhastCons¹¹⁹). Ces analyses peuvent être effectuées sur un total de 28 espèces de vertébrés dont l'arbre phylogénétique est représenté à la figure 12. Pour l'analyse avec PhastCons, deux mesures peuvent être obtenues : une conservation pour les 28 espèces (nommé *Vertebrate Cons* dans l'interface visuelle) et une autre restreinte aux mammifères placentaux (*Mammal Cons*) donc 17 espèces plus l'humain. Si seulement les mammifères apparaissent dans l'alignement, ces deux mesures produisent les mêmes résultats; cependant, si la région est conservée chez les espèces non mammaliennes, le score augmente sinon il diminuera.

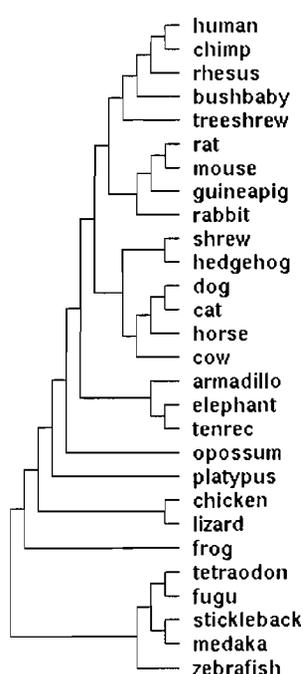


Figure 12 Arbre phylogénétique des 28 espèces disponibles pour les analyses avec Multiz et PhastCons

6.6 Bio-informatique

Afin de gérer cette grande quantité de données génotypiques et phénotypiques, l'utilisation de la bio-informatique fut essentielle. Ma formation en bio-informatique m'a permis de concevoir les programmes (langage JAVA) requis pour : automatiser et optimiser les tests génétiques (GEE et XRC-TDT), produire les fichiers d'entrées ou les scripts sous différents formats (Merlin, XRC-TDT et GEE), transformer des fichiers de sortie en graphiques (XRC-TDT et Merlin), accomplir la correction de covariables (sexe et âge),

vérifier la distribution normale des phénotypes et calculer la fréquence de l'allèle mineur. J'ai aussi employé la bio-informatique pour modifier le programme EMLD afin qu'il puisse accepter le chromosome X. De plus, grâce à ses nombreuses méthodes disponibles, j'ai eu recours au logiciel R à plusieurs reprises afin d'accomplir diverses tâches tout au long des analyses (calcul de l'équilibre d'Hardy-Weinberg) et j'ai construit les requêtes pour la base de données au moyen du langage SQL.

Résultats

Chapitre 7 – Résultats sur le chromosome X

Avant d'entamer une étude génétique, il est primordial d'exclure les SNPs qui résulteraient d'une erreur de génotypage. Pour des données basées sur des familles, ces erreurs peuvent augmenter les erreurs de type I et II¹²⁰ et mener à de fausses conclusions. De plus, les SNPs ayant une fréquence de l'allèle mineur inférieur à 5% ont été éliminés de l'étude, car l'hypertension ne représente pas une maladie rare. Pour prendre en considération les multiples tests accomplis, la correction de Bonferroni peut être appliquée. Les résultats obtenus à partir des 1204 SNPs seront significatifs si la valeur p obtenue est inférieure à 1E-06 et il sera question d'associations suggestives si la valeur p est entre 0.05 et 1E-06. Pour le bloc d'haplotypes d'ACE2, les résultats seront significatifs si celle-ci est inférieure à 1E-03 et il y aura associations suggestives si la valeur p se trouve entre 0.05 et 1E-03.

7.1 Programme Merlin

Des analyses de liaisons non paramétriques basées sur les familles ont été accomplies par Merlin via l'option minx; option conçue spécialement pour le chromosome X. Ce programme a l'avantage d'accepter les phénotypes continus aussi bien que ceux binaires et de s'exécuter dans un temps raisonnable. Par contre, sur le site Internet du logiciel (<http://www.sph.umich.edu/csg/abecasis/Merlin/reference.html>), les auteurs font une mise en garde à savoir qu'aucun article ne décrit la performance de l'option minx ainsi que les algorithmes en détail. Même si les auteurs croient que l'option minx fournit des résultats valables, les méthodes ne sont pas publiées et donc, les auteurs mentionnent qu'il faut les utiliser avec précaution. De plus, en divisant différemment les grandes familles en familles nucléaires, les résultats obtenus par ces diverses séparations varient. En effet, des résultats qui sont statistiquement significatifs lors d'une certaine séparation peuvent ne plus être présents lors d'une nouvelle séparation des familles. Étant donné ce manque de robustesse, de validité et de cohérence, une suite a été donnée uniquement aux résultats provenant du programme XRC-TDT dont la méthode a été publiée et testée par simulations¹²¹.

7.2 Programme XRC-TDT

Plusieurs résultats significatifs ont été obtenus avec les analyses effectuées à l'aide du programme XRC-TDT. Dans cette section, les résultats sont présentés sans avoir été corrigés pour les tests multiples afin de ne pas éliminer des SNPs qui pourraient être importants dans l'hypertension. Le tableau VI résume les meilleurs résultats identifiés avec XRC-TDT; il contient les SNPs avec les phénotypes auxquels ils sont associés ainsi que les gènes d'intérêt. La sélection de ces gènes repose sur deux conditions : le gène contient le SNP identifié ou il est positionné à proximité du SNP, c'est-à-dire à 500 kb avant ou après le SNP. La distance de 500 kb permet l'existence de LD entre le SNP et la variante réellement à l'origine de l'effet pathologique observé¹²². En étudiant le LD de diverses régions génomiques, dans une population du Royaume-Uni, Abecasis *et al.* ont trouvé que la moitié des SNPs étaient en un LD suffisamment utile pour l'identification d'associations à une distance allant jusqu'à 50 kb, avec certaines associations s'étendant aussi loin que 500 kb¹²³.

Chez les hommes, les rs438102, rs407890 ainsi que le bloc d'haplotypes d'ACE2 sont à l'origine d'associations significatives; tandis que chez les femmes, ce sont les rs2116366, rs727562 et rs4911871. L'association entre le rs727562 et la pression artérielle durant le jour est aussi présente dans le groupe des deux sexes réunis. Dans ce même groupe, les rs4131420 et rs1949393 sont associés à l'indice Milwaukee et à six phénotypes de pression artérielle respectivement. Plusieurs gènes sont positionnés près de ces SNPs et parmi ceux-ci, certains représentent des gènes candidats prometteurs.

Tableau VI Meilleures associations obtenues pour le chromosome X avec le programme XRC-TDT

| SNP | cM | Phénotype | Valeur p obtenues avec XRC-TDT | | | Gène à proximité |
|---------------------------------|--------|--|--|--|--|--|
| | | | 2 sexes | Femmes | Hommes | |
| rs4131420 | • 37,0 | • Indice Milwaukee | • 6,90E-06 | • 0,004 | • 0,3 | RPGR, OTC, SC4MOP, TCTE1L, TM4SF2, SRPX, HYPM, MIG12, SYTL5, LOC286441, LOC340549, LOC392442, LOC392443, LOC442446 |
| rs438102 | • 82,8 | • Statut d'hypertension* | • 0,002 | • 0,4 | • 1,48E-04 | LOC139363, FLJ30672, LOC392552 |
| rs1949393 | • 57,8 | • Statut d'hypertension* • PAD durant 24h • PAS durant 24h • PAD dans le jour • PAS dans le jour • PAD dans la nuit | • 9,97E-04 • 1,57E-04 • 4,91E-04 • 2,41E-06 • 1,67E-04 • 7,10E-04 | • 0,06 • 0,03 • 0,04 • 0,002 • 0,02 • 0,1 | • 0,2 • 0,05 • 0,03 • 0,02 • 0,02 • 0,1 | KLHL4, LAMR1P15, bA345E19.2 |
| rs2116366 | • 52,1 | • PAS durant 24hrs • PAS durant la nuit | • 0,008 • 0,01 | • 6,67E-04 • 6,91E-04 | • 0,4 • 0,4 | FLJ31204 (FAAH2), ZXDB, ZXDA, LOC169981, NOD2 4, LOC392477, SPIN2 |
| rs727562 | • 30,6 | • PAS dans le jour | • 5,86E-04 | • 2,07E-04 | • 0,1 | IL1RAPL1 |
| rs4911871 | • 69,1 | • Indice Heather | • 0,5 | • 5,46E-06 | • 0,1 | HTR2C, IL13RA2, TMSL7, LRCH2, FLJ40249, LOC286528 |
| rs407890 | • 59,7 | • Fréquence cardiaque | • 0,4 | • 0,09 | • 1,12E-05 | TGIF2LX |
| rs2074192, rs233575 (haplotype) | • 22,2 | • LVMI | • 0,7 | • 0,5 | • 0,04 | ACE2 |

*Le statut d'hypertension est le seul phénotype qualitatif du tableau; cM : centimorgan

Afin de vérifier si les associations trouvées sont rattachées à un état biologique particulier, des sous-catégories ont été formées et analysées: hypertendu (HT), normotendu (NT), obèse (Ob) et non obèse (NOB). Des associations significatives dans l'un de ces sous-groupes ont été constatées pour les rs1949393, rs2116366, rs727562 et rs407890. À propos du rs1949393, les associations sont hautement significatives chez le groupe renfermant les deux sexes, en particulier chez les hypertendus (Tableau VII). Pour le rs407890, toutes les associations sont concentrées strictement chez les hommes, sans distinction pour les sous-catégories (Tableau VIII). Quant au rs2116366, une forte association, avec la PAS de 24hrs

ainsi que de nuit, a été identifiée parmi les femmes hypertendues et non obèses (Tableaux IX et X). Puis, pour le rs727562 qui était associé dans deux groupes (Tableau VI), un plus grand nombre d'associations est retrouvé chez les femmes et cible en particulier celles hypertendues et non obèses (Tableau XI).

Tableau VII Associations entre le rs1949393 et la PAS durant 24 heures pour différents sous-groupes à l'aide de XRC-TDT

| Sous-groupes | 2 sexes | Hommes | Femmes |
|--------------|----------|--------|--------|
| Non divisé | 4,99E-04 | 0,03 | 0,04 |
| Ob | 0,03 | 0,3 | 0,1 |
| -Ob | 0,01 | 0,2 | 0,2 |
| HT | 5,12E-03 | 0,5 | 0,2 |
| NT | 0,1 | 0,1 | 0,3 |
| HT+Ob | 0,2 | 0,4 | 0,5 |
| HT-Ob | 0,08 | 1 | 0,2 |
| NT+Ob | 0,4 | 0,6 | 1 |
| NT-Ob | 0,2 | 0,4 | 0,7 |

Tableau VIII Associations entre le rs407890 et la fréquence cardiaque pour différents sous-groupes à l'aide de XRC-TDT

| Sous-Groupes | 2 sexes | Hommes | Femmes |
|--------------|---------|----------|--------|
| Non divisé | 0,3 | 1,12E-05 | 0,09 |
| Ob | 0,3 | 7,78E-03 | 0,4 |
| -Ob | 1 | 0,02 | 0,3 |
| HT | 0,09 | 3,27E-03 | 0,5 |
| NT | 1 | 0,03 | 0,3 |
| HT+Ob | 0,2 | 0,1 | 1 |
| HT-Ob | 0,8 | 0,3 | 0,7 |
| NT+Ob | 1 | 0,1 | 1 |
| NT-Ob | 1 | 1 | 0,5 |

Tableau IX Associations entre le rs2116366 et la PAS durant 24 heures pour différents sous-groupes à l'aide de XRC-TDT

| Sous-groupes | 2 sexes | Hommes | Femmes |
|--------------|----------|--------|----------|
| Non divisé | 0,005 | 0,4 | 6,67E-04 |
| Ob | 0,4 | 0,4 | 1 |
| -Ob | 0,05 | 0,5 | 1,38E-03 |
| HT | 6,92E-03 | 0,1 | 7,26E-03 |
| NT | 0,3 | 0,3 | 0,2 |
| HT+Ob | 0,7 | 0,3 | 1 |
| HT-Ob | 0,01 | 0,2 | 0,02 |
| NT+Ob | 0,9 | 1 | 1 |
| NT-Ob | 0,4 | 0,3 | 0,2 |

Tableau X Associations entre le rs2116366 et la PAS durant la nuit pour différents sous-groupes à l'aide de XRC-TDT

| Sous-Groupes | 2 sexes | Hommes | Femmes |
|--------------|----------|--------|----------|
| Non divisé | 0,007 | 0,4 | 6,91E-04 |
| Ob | 0,4 | 0,4 | 1 |
| -Ob | 0,03 | 0,8 | 1,38E-03 |
| HT | 0,02 | 0,2 | 7,26E-03 |
| NT | 0,6 | 0,3 | 0,6 |
| HT+Ob | 0,7 | 0,7 | 1 |
| HT-Ob | 1,90E-03 | 0,1 | 0,02 |
| NT+Ob | 1 | 1 | --- |
| NT-Ob | 0,8 | 0,3 | 0,5 |

Tableau XI Associations entre le rs727562 et la PAS durant le jour pour différents sous-groupes à l'aide de XRC-TDT

| Sous-groupes | 2 sexes | Hommes | Femmes |
|--------------|---------|--------|----------|
| Non divisé | 0,002 | 0,1 | 2,07E-04 |
| Ob | 0,05 | 0,06 | 0,2 |
| -Ob | 0,02 | 0,8 | 1,78E-03 |
| HT | 0,06 | 1 | 2,07E-03 |
| NT | 0,05 | 0,1 | 1 |
| HT+Ob | 0,9 | 0,4 | 0,5 |
| HT-Ob | 0,1 | 1 | 8,30E-03 |
| NT+Ob | 0,2 | 0,3 | 0,4 |
| NT-Ob | 0,9 | 0,4 | 1 |

7.3 Relation entre le génotype et le phénotype

La relation, entre le phénotype et le génotype des SNPs identifiés, a été évaluée à partir d'un test d'association accompli par la méthode GEE, une analyse de régression. Pour ce faire, les sujets ont été regroupés selon les allèles du SNP et les moyennes de phénotype pour ces groupes alléliques ont été comparées entre elles. Lorsque la variable indépendante, ici le génotype, est de type facteur (variable catégorique), une des valeurs du facteur doit être considérée comme la valeur de référence. Dans le script servant à exécuter la fonction GEE (SAS 9.1), la division des sujets en différents sous-groupes alléliques requiert la création de variables fictives (*dummy variable*). Celles-ci représentent des sous-groupes de l'échantillon à l'étude (ici les sous-groupes alléliques) et seront affectées à la valeur 0 ou 1. Par exemple, chez les femmes du tableau XII, la variable fictive fixée à 0 est le sous-groupe allélique (GG) et devient la valeur de référence; tandis que les variables

fictives assignées à 1 correspondent aux restes des sous-groupes alléliques (AA et AG). Alors, les variables établies à 1 seront comparées à la valeur de référence, celle affectée à 0.

Une relation phénotype/génotype significative existe dans le cas des rs2116366 et rs407890. En ce qui concerne le rs2116366, la plus grande différence de moyenne, statistiquement significative, a été observée chez les femmes pour la PAS de 24 heures (Tableau XII). Une différence de 7,4 mmHg est présente entre les deux groupes de femmes homozygotes; les femmes ayant les allèles AA montrant une valeur de PAS plus élevée. Chez les femmes, cet allèle posséderait donc un effet dominant qui n'est pas autant présent chez les hommes; mais globalement, les mesures de PAS, chez les individus ayant au moins un allèle A, sont plus élevées (Tableaux XII et XIII). Concernant la prévalence du statut d'hypertension, 64,6% des femmes avec le génotype AA s'avèrent être hypertendues comparativement à 42,3% chez les femmes homozygotes pour l'allèle G (Tableau XIV). Pour notre population, l'allèle A du rs2116366 est le plus fréquent, ce qui correspond à ce qui est observé dans la population caucasienne sur HapMap, où 76,7% des chromosomes ont l'allèle A à cette position; tandis que, l'allèle G est plus fréquent avec 98,3% chez les africains, où la prévalence de l'hypertension est la plus élevée. Quant au rs407890, la différence de fréquence cardiaque la plus significative se trouve au sein du groupe des deux sexes réunis avec 3,4 battements/minute entre les homozygotes AA+A et GG+G (Tableau XV).

Tableau XII Comparaison des moyennes de PAS durant 24 heures selon les différents génotypes du rs2116366

| Sexe | Génotype du SNP | # sujets | Moyenne | Déviati on standard | Valeur p |
|---------|-----------------|----------|---------|---------------------|-----------|
| Femme | AA | 36 | 118,8 | 14,8 | 0,02 |
| | AG | 60 | 114,7 | 13,6 | 0,2 |
| | GG | 26 | 111,4 | 10,5 | Référence |
| | AA | 36 | 118,8 | 14,8 | 0,2 |
| | AG | 60 | 114,7 | 13,6 | Référence |
| | GG | 26 | 111,4 | 10,5 | 0,2 |
| | AG+GG | 86 | 113,7 | 12,7 | Référence |
| Homme | A | 75 | 123,9 | 12,1 | 0,1 |
| | G | 49 | 120,7 | 12,3 | Référence |
| 2 sexes | AA+A | 111 | 121,0 | 13,0 | 0,03 |
| | AG | 60 | 118,4 | 13,5 | 0,5 |
| | GG+G | 75 | 116,3 | 12,0 | Référence |
| | AA+A | 111 | 121,0 | 13,0 | 0,4 |
| | AG | 60 | 118,4 | 13,5 | Référence |
| | GG+G | 75 | 116,3 | 12,0 | 0,5 |
| | AG+GG+G | 135 | 117,3 | 12,7 | Référence |

Tableau XIII Comparaison des moyennes de PAS durant la nuit selon les différents génotypes du rs2116366

| Sexe | Génotype du SNP | # sujets | Moyenne | Déviati on standard | Valeur p |
|---------|-----------------|----------|---------|---------------------|-----------|
| Femme | AA | 36 | 112,7 | 15,0 | 0,04 |
| | AG | 60 | 106,8 | 13,0 | 0,7 |
| | GG | 27 | 105,7 | 11,0 | Référence |
| | AA | 36 | 112,7 | 15,0 | 0,04 |
| | AG | 60 | 106,8 | 13,0 | Référence |
| | GG | 27 | 105,7 | 11,0 | 0,7 |
| | AG+GG | 87 | 106,5 | 12,3 | Référence |
| Homme | A | 75 | 117,4 | 12,7 | 0,2 |
| | G | 49 | 114,2 | 12,6 | Référence |
| 2 sexes | AA+A | 111 | 114,5 | 13,3 | 0,06 |
| | AG | 60 | 110,7 | 12,9 | 0,9 |
| | GG+G | 76 | 110,2 | 12,5 | Référence |
| | AA+A | 111 | 114,5 | 13,3 | 0,09 |
| | AG | 60 | 110,7 | 12,9 | Référence |
| | GG+G | 76 | 110,2 | 12,5 | 0,9 |
| | AG+GG+G | 136 | 110,4 | 12,6 | Référence |

Tableau XIV Prévalence du statut d'hypertension chez les femmes selon les génotypes du rs2116366

| Sexe | Sujets hypertendus avec le génotype AA (n = 79) | Sujets hypertendus avec le génotype AG (n = 117) | Sujets hypertendus avec le génotype GG (n=52) |
|----------------|---|--|---|
| Femmes (n=248) | 64,6% | 46,2% | 42,3% |

Tableau XV Comparaison des moyennes de la fréquence cardiaque selon les différents génotypes du rs407890

| Sexe | Génotype du SNP | # sujets | Moyenne | Déviati on standard | Valeur p |
|---------|-----------------|----------|---------|---------------------|-----------|
| Femme | AA | 10 | 75,1 | 10,9 | 0,4 |
| | AG | 54 | 73,5 | 12,2 | 0,9 |
| | GG | 69 | 72,9 | 9,6 | Référence |
| | AA | 10 | 75,1 | 10,9 | 0,6 |
| | AG | 54 | 73,5 | 12,2 | Référence |
| | GG | 69 | 72,9 | 9,6 | 0,9 |
| | AA | 10 | 75,1 | 10,9 | 0,5 |
| | AG+GG | 123 | 73,2 | 10,8 | Référence |
| Homme | A | 42 | 69,6 | 10,8 | 0,1 |
| | G | 97 | 65,7 | 10,3 | Référence |
| 2 sexes | AA+A | 52 | 72,6 | 10,8 | 0,02 |
| | AG | 54 | 70,1 | 12,3 | 0,5 |
| | GG+G | 166 | 69,2 | 10,1 | Référence |
| | AA+A | 52 | 72,6 | 10,8 | 0,3 |
| | AG | 54 | 70,1 | 12,3 | Référence |
| | GG+G | 166 | 69,2 | 10,1 | 0,5 |
| | AA+A | 52 | 72,6 | 10,8 | 0,03 |
| | AG+GG+G | 220 | 69,4 | 10,7 | Référence |

7.4 Déséquilibre de liaison

Pour les SNPs d'intérêt, la prochaine étape consiste à regarder s'ils sont en LD avec un SNP localisé dans un gène ou s'ils font partis d'un bloc d'haplotypes. La figure 13 montre un exemple d'une carte de LD produite avec le logiciel Haploview⁹⁷. Dans cet exemple, le LD présent entre les SNPs numérotés 778 à 785 est illustré par des losanges rouges entre ceux-ci. La couleur du losange représente la valeur de LD; plus celle-ci se rapproche de 1 (LD complet) plus elle sera foncée. Les valeurs indiquées dans les losanges rouges représentent le D'. Dans la figure 13, un bloc d'haplotypes est formé par les SNPs numérotés de 778 à 785.

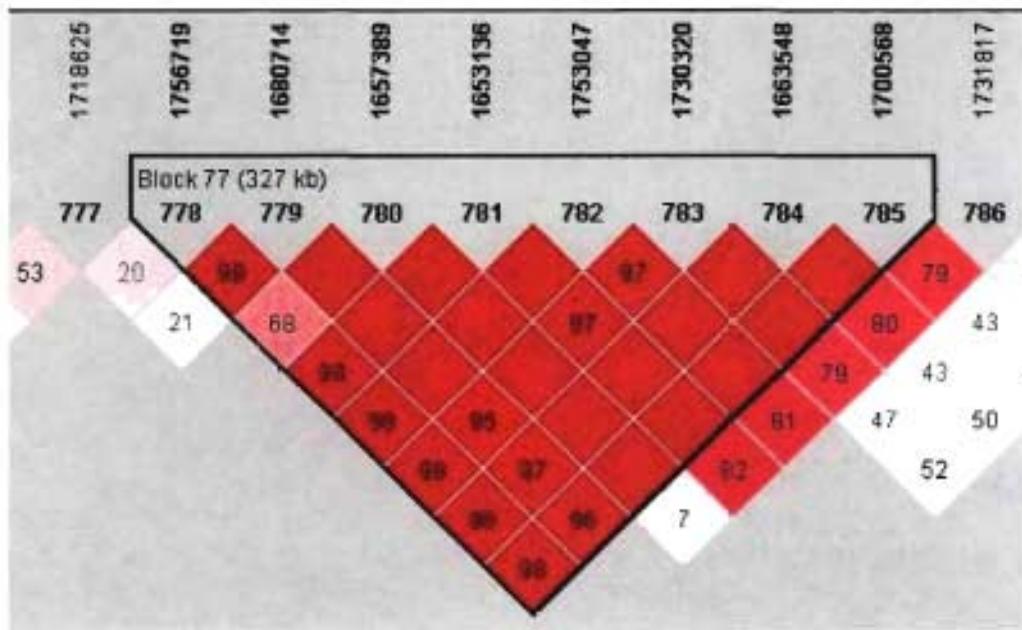


Figure 13 Carte de déséquilibre de liaison, obtenue avec Haploview, dans la population du SLSJ pour le rs4911871 associé à l'indice Heather

Le tableau XVI répertorie le nombre de SNPs de la population caucasienne en LD avec les SNPs d'intérêt identifiés par XRC-TDT et indique si ceux-ci se trouvent dans un gène. Les informations sur les SNPs de la population caucasienne ont été prises sur le site public d'HapMap. Trois des SNPs mis en évidence par XRC-TDT (rs2116366, rs727562, rs4911871) se localisent déjà à l'intérieur d'un gène; alors que, parmi les quatre autres, un seul, le rs4131420, est en LD avec des SNPs de la population caucasienne qui sont situés dans le gène *retinitis pigmentosa GTPase regulator* (RPGR). Au total, quatre SNPs de la population caucasienne, en LD avec ceux identifiés lors du projet, sont situés dans des exons de gènes.

Tableau XVI Nombre de SNPs de la population caucasienne en LD avec les SNPs d'intérêt

Tableau XVI Nombre de SNPs de la population caucasienne en LD avec les SNPs d'intérêt

| SNPs identifiés | # de SNPs caucasiens | SNPs localisés dans un gène |
|-----------------|----------------------|--|
| rs4131420 | 18 | Les SNPs caucasiens sont dans le gène RPGR |
| rs438102 | 0 | Aucun SNP n'est dans un gène |
| rs1949393 | 0 | Aucun SNP n'est dans un gène |
| rs2116366 | 65 | Le SNP identifié ainsi que ceux caucasiens sont localisés dans FLJ31204 (FAAH2). Deux des SNPs caucasiens, rs2516023 et rs1367830, sont situés dans l'exon 1 et l'exon 10 de ce gène, avec une fonction synonyme |
| rs727562 | 20 | Le SNP identifié et ceux caucasiens sont dans IL1RAPL1 |
| rs4911871 | 54 | Le SNP identifié et ceux caucasiens se localisent dans HTR2C. Deux des SNPs caucasiens, rs1801412 et le rs6318, se situent dans l'exon 6 (fonction 3'UTR) et l'exon 4 (fonction non synonyme) |
| rs407890 | 0 | Aucun SNP n'est dans un gène |

À l'aide du programme Haploview, les blocs d'haplotypes correspondants aux SNPs identifiés ont été trouvés. Le tableau XVII offre une description de ces blocs avec les gènes couverts par ceux-ci. De plus, pour les SNPs n'étant pas associés à un bloc, ce tableau indique quels SNPs sont en LD complet avec le SNP identifié lors du projet.

Tableau XVII Blocs d'haplotypes et régions de LD complet associés au SNPs identifiés

| SNP | Bloc d'haplotypes et régions de LD complet | Longueur de la région touchée (kb) | Gène couvert |
|-----------------------|--|------------------------------------|--|
| rs4131420 | rs10521413, rs4131420 | 17 | RPGR (rs10521413 : intron 6) |
| rs438102 | rs438102, rs398262, rs1929364 | 2 | --- |
| rs1949393 | rs1949393, rs1589075 | 0,074 | --- |
| rs2116366 | rs7050676, rs204166, rs10521487, rs2116366 | 449 | SPIN2, FAAH2 (rs2116366 : intron 7) |
| rs727562 | rs727562, rs10521953 | 14 | IL1RAPL1 (rs727562 : intron 5, rs10521953: intron 5) |
| rs4911871 | rs569959, rs539748, rs4911871, rs2497522, rs2428707, rs950463, rs5946229, rs1335614 | 327 | HTR2C (rs569959 : intron 1, rs539748: intron 1, rs4911871: intron 4, rs2497522: intron 4, rs2428707: intron 4), LOC727794 (rs569959: exon 1), SNORA35, MIRN448 |
| rs407890 | rs223724, rs407890 | 3 | --- |
| rs2074192 rs233575 | rs2074192, rs233575 | 0,176 | ACE2 (rs2074192 : intron 16, rs233575 : intron 16) |

Les SNPs en caractère gras sont ceux qui forment réellement un bloc d'haplotypes et les autres sont uniquement en LD complet ($D' = 1$)

7.5 Prédiction de l'impact des SNPs

L'impact des SNPs non codants a ensuite été analysé à l'aide de trois programmes : F-SNP, MAPPER et SNPInspector. Le tableau XVIII décrit les résultats obtenus avec F-SNP. Pour cinq SNPs, le programme a prédit un changement de sites de liaison de facteur de transcription (TFBS) en comparant l'allèle ancestral à l'allèle mutant, dont trois SNPs où les facteurs de transcription se rapportent à l'espèce humaine.

Tableau XVIII Prédiction d'impact des SNPs identifiés avec F-SNP

| SNP | Type de région | Impact prédit | TFBS chez l'humain |
|------------|------------------|---|---|
| rs4131420 | En amont de RPGR | --- | --- |
| rs10521413 | Intronique | --- | --- |
| rs438102 | Intergénique | --- | --- |
| rs398262 | Intergénique | --- | --- |
| rs1929364 | Intergénique | --- | --- |
| rs1949393 | Intergénique | --- | --- |
| rs1589075 | Intergénique | --- | --- |
| rs2116366 | Intronique | --- | --- |
| rs7050676 | Intergénique | --- | --- |
| rs204166 | Intergénique | --- | --- |
| rs10521487 | Intergénique | --- | --- |
| rs727562 | Intronique | --- | --- |
| rs10521953 | Intronique | --- | --- |
| rs4911871 | Intronique | Changement dans la prédiction de TFBS | Avec l'allèle A (ancestral), il y a gain d'un TFBS pour CDPCR |
| rs569959 | Intronique | Changement dans la prédiction de TFBS | --- |
| rs539748 | Intronique | Changement dans la prédiction de TFBS | Avec l'allèle A (ancestral), il y a gain d'un TFBS pour Pbx-1 |
| rs2497522 | Intronique | Changement dans la prédiction de TFBS | Avec l'allèle C (mutant), il y a gain d'un TFBS pour CRE-BP |
| rs2428707 | Intronique | Aucun changement dans la prédiction de TFBS | --- |
| rs950463 | En aval d'HTR2C | --- | --- |
| rs5946229 | En aval d'HTR2C | --- | --- |
| rs1335614 | En aval d'HTR2C | --- | --- |
| rs407890 | Intergénique | --- | --- |
| rs223724 | Intergénique | --- | --- |
| rs2074192 | Intronique | Aucun changement dans la prédiction de TFBS | --- |
| rs233575 | Intronique | Changement dans la prédiction de TFBS | --- |

Le logiciel MAPPER a aussi prédit une liste farouche de TFBS dans les introns des gènes où se situent les SNPs d'intérêt et ce pour différentes espèces. En regardant la

visualisation des résultats de ce programme à l'aide du site UCSC, seul le rs2074192 se situe à l'intérieur d'un TFSB putatif, celui de Pou3F2. Il faut noter que pour le gène HTR2C, le programme n'a pu terminer l'analyse, peut-être en raison d'un trop grand nombre de TFBS possibles, dépassant la capacité du logiciel à tous les afficher.

Par la suite, des analyses à l'aide de SNPInspector ont été accomplies. Pour tous les SNPs situés dans des régions introniques, ce logiciel a prédit des gains ou des pertes de TFBS (Tableau XIX). Pour le rs4911871, tout comme F-SNP, SNPInspector suggère qu'avec l'allèle G il y a perte du TFBS d'une protéine de type CDPCR.

Tableau XIX Prédiction d'impact des SNPs identifiés avec SNPInspector

| SNP | Changement | Gain/Perte de TFBS | Brin | Facteur de transcription |
|------------|------------|--------------------|------|--------------------------|
| rs10521413 | C → T | Gain | - | PAX6 |
| | C → T | Gain | - | E2F4 DP2 |
| rs2116366 | C → T | Perte | + | OCT1 |
| | C → T | Gain | - | PAX3 |
| | C → T | Gain | - | MYT1 |
| | C → T | Gain | - | P53 |
| rs727562 | A → C | Perte | + | E4BP |
| | A → C | Perte | + | VBP |
| | A → C | Perte | + | ATATA |
| | A → C | Gain | - | PAX6 |
| rs10521953 | A → T | Perte | + | STAF |
| | A → T | Gain | - | LTATA |
| rs4911871 | A → G | Perte | + | CDPCR3HD |
| | A → G | Perte | + | HOX PBX |
| | A → G | Gain | - | CDPCR3 |
| | A → G | Gain | - | YY1 |
| | A → G | Gain | - | NEUROD1 |
| rs569959 | A → G | Perte | + | YY1 |
| | A → G | Gain | - | CDPCR3 |
| | A → G | Gain | - | WHN |
| rs539748 | A → G | Perte | + | XVENT2 |
| | A → G | Perte | + | CDP |
| | A → G | Perte | - | BRN2 |
| | A → G | Perte | + | NFY |
| | A → G | Perte | + | HOXC13 |
| | A → G | Gain | + | GATA |
| rs2497522 | C → T | Perte | - | TAXCREB |
| | C → T | Gain | - | TEF1 |
| rs2428707 | A → G | Perte | + | PLZF |
| | A → G | Gain | - | CDX2 |
| rs2074192 | C → T | Perte | - | OTC |
| | C → T | Perte | - | NF1 |
| | C → T | Gain | - | MRF2 |
| rs233575 | C → T | Gain | + | BCL6 |

7.6 Conservation des séquences entre les espèces

En examinant la conservation de séquences entre les espèces via le site UCSC, il peut être observé que les rs4131420, rs438102, rs398262, rs1929364, rs1949393, rs1589075, rs2116366, rs7050676, rs204166, rs10521487, rs727562, rs10521953, rs539748, rs2497522, rs2428707, rs5946229, rs407890, rs223724 et rs2074192 ne se situent pas au niveau d'une région conservée (données non montrées). Les rs10521413, rs569959 et 233575 se localisent à moins de 20 pb d'une région bien conservée chez la plupart des mammifères placentaux et les primates (Figures 14-16). Pour les primates uniquement, le rs950463 survient entre deux régions relativement bien conservées (Figure 17). Puis, deux SNPs, rs1335614 et rs4911871, sont contenus aux extrémités d'une région moyennement conservée (chez le chimpanzé, le Rhesus, le chien et l'éléphant) et peu conservée (chez les primates et six mammifères placentaux) respectivement (Figures 18 et 19).

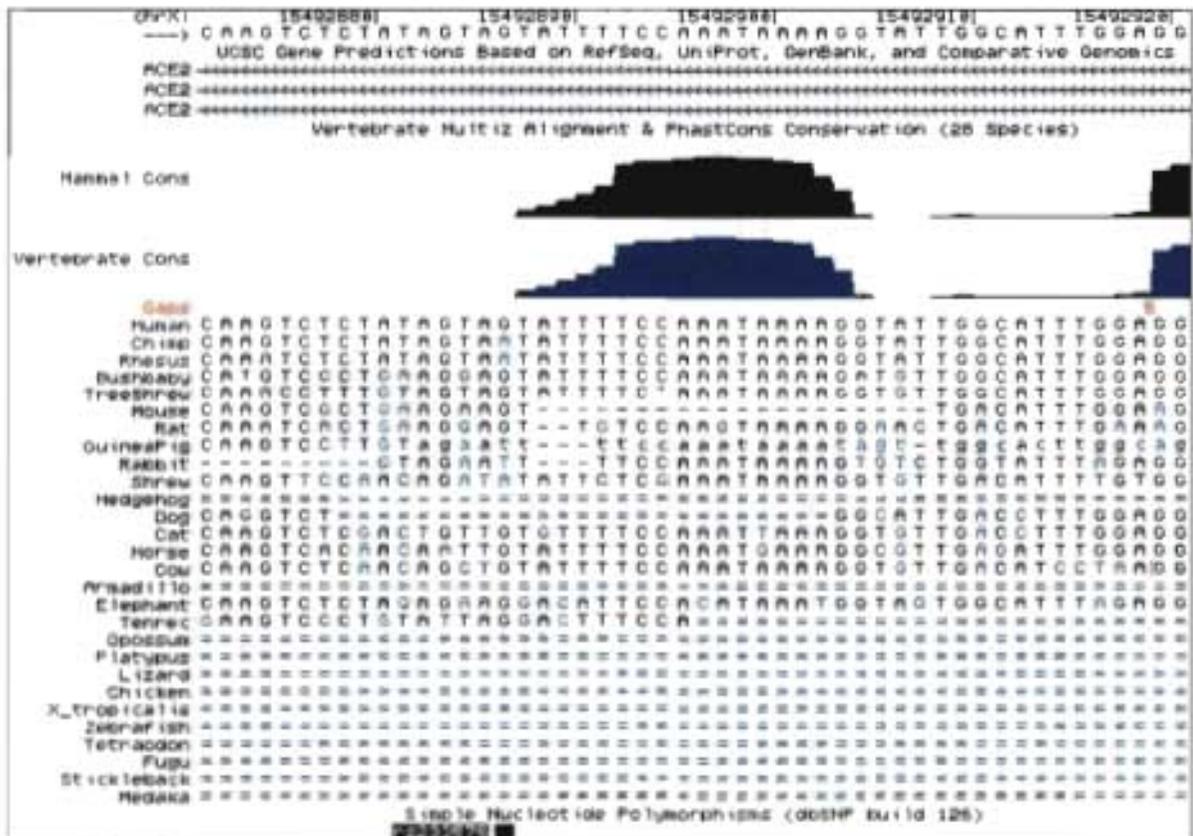


Figure 16 Alignments et conservation de séquences chez 28 espèces pour le rs233575

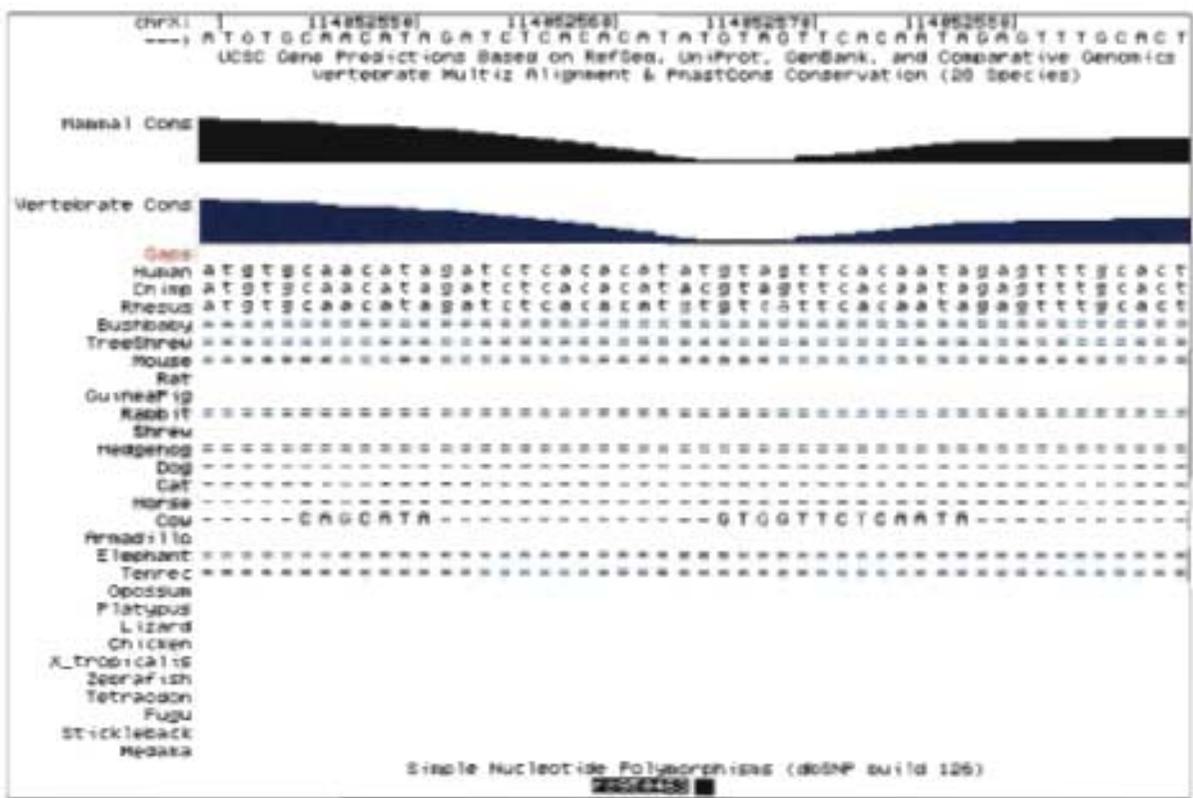


Figure 17 Alignments et conservation de séquences chez 28 espèces pour le rs950463

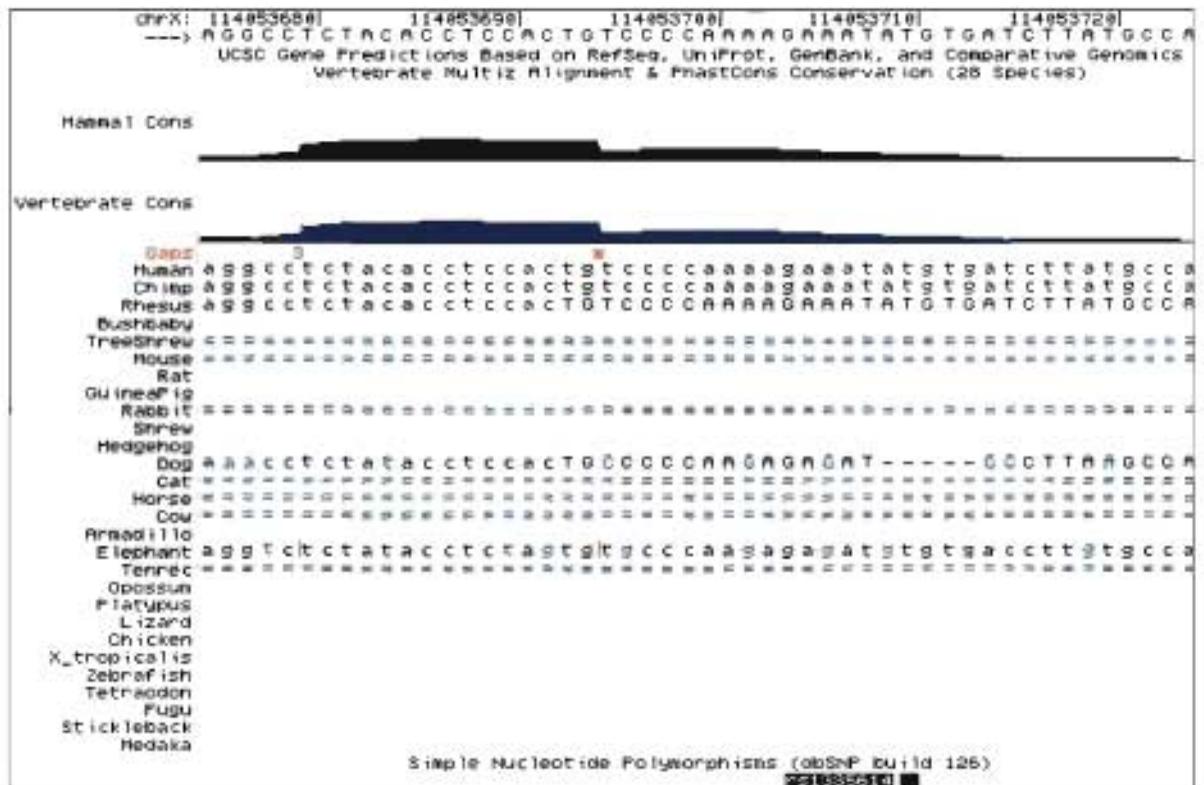


Figure 18 Alignments et conservation de séquences chez 28 espèces pour le rs1335614

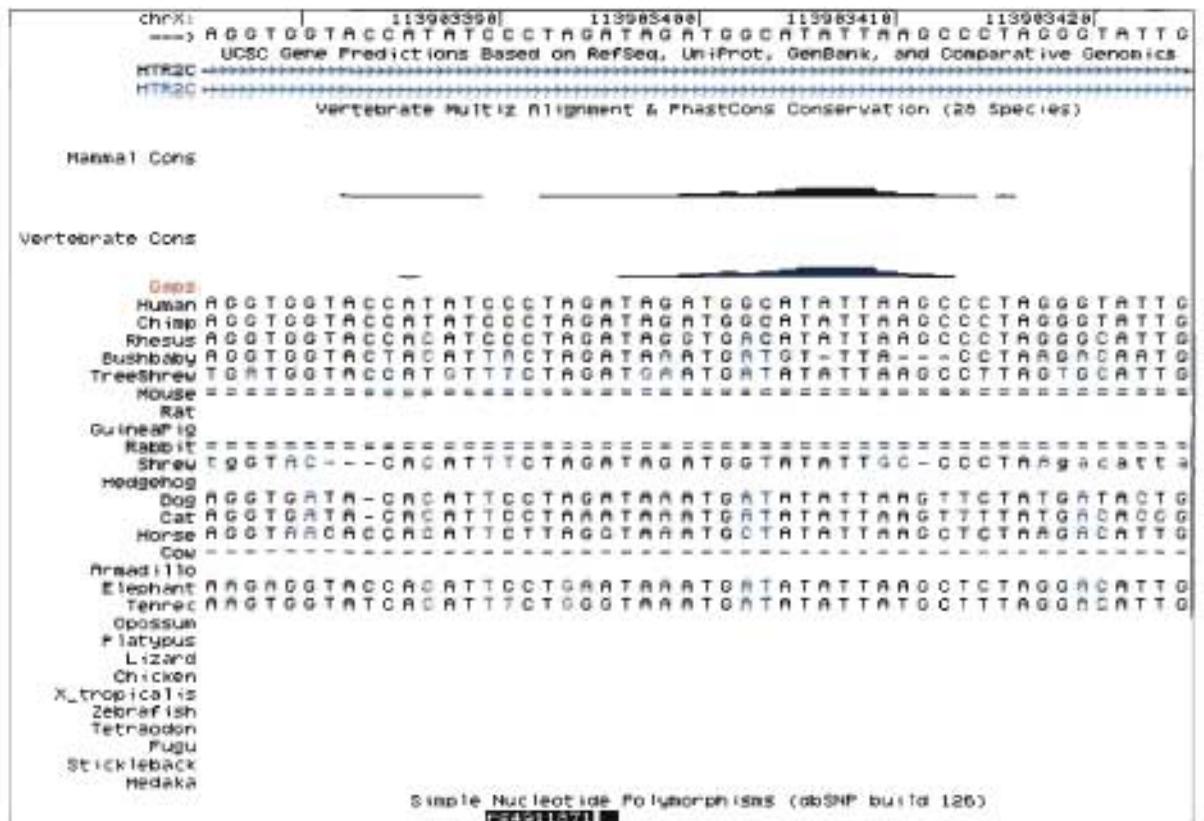


Figure 19 Alignments et conservation de séquences chez 28 espèces pour le rs4911871

Après avoir examiné le niveau de conservation des régions où sont présents les SNPs identifiés, la conservation des gènes couverts par les blocs d'haplotypes a été étudiée (RPGR, FAAH2, IL1RAPL1, HTR2C, ACE2 et SPIN2). Pour ACE2, toutes les séquences des exons démontrent une très bonne homologie surtout chez les mammifères placentaux que chez le reste des vertébrés en plus de la présence d'un haut pic de conservation dans l'intron 2 (Figure 20). Pour le gène RPGR, il peut être constaté que la conservation se trouve chez la plupart des séquences exoniques chez les primates et les mammifères placentaux (Figure 21). Un patron similaire est également observable pour HTR2C pour les mêmes espèces (Figure 22). Pour sa part, IL1RAPL1 montre un taux de conservation sans préférence pour les exons chez les primates et les mammifères placentaux (Figure 23); tandis que, FAAH2 représente l'homologie la plus faible de tous les gènes étudiés, avec quelques bouts de séquences similaires chez le chimpanzé et le Rhesus (Figure 24). De son côté SPIN2, contient un exon hautement conservé chez les primates et environ la moitié des mammifères placentaux. Même des alignements ont pu être établis pour le reste des vertébrés (Figure 25). Malheureusement, le patron de conservation de SNORA35 et MIRN448 n'ont pu être obtenu. En effet, en entrant les noms de ces deux gènes sur le site UCSC, la visualisation affichée concerne le chromosome 7 et non le chromosome X.

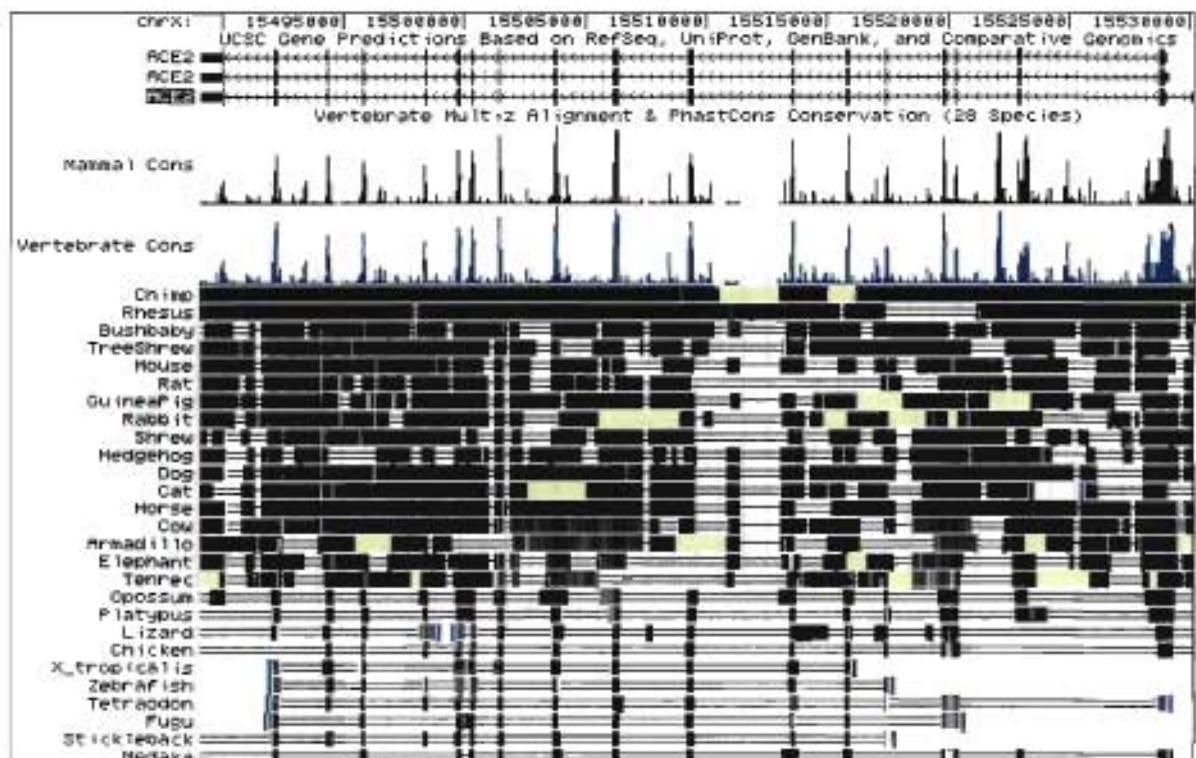


Figure 20 Alignements et conservation de séquences chez 28 espèces pour le gène ACE2

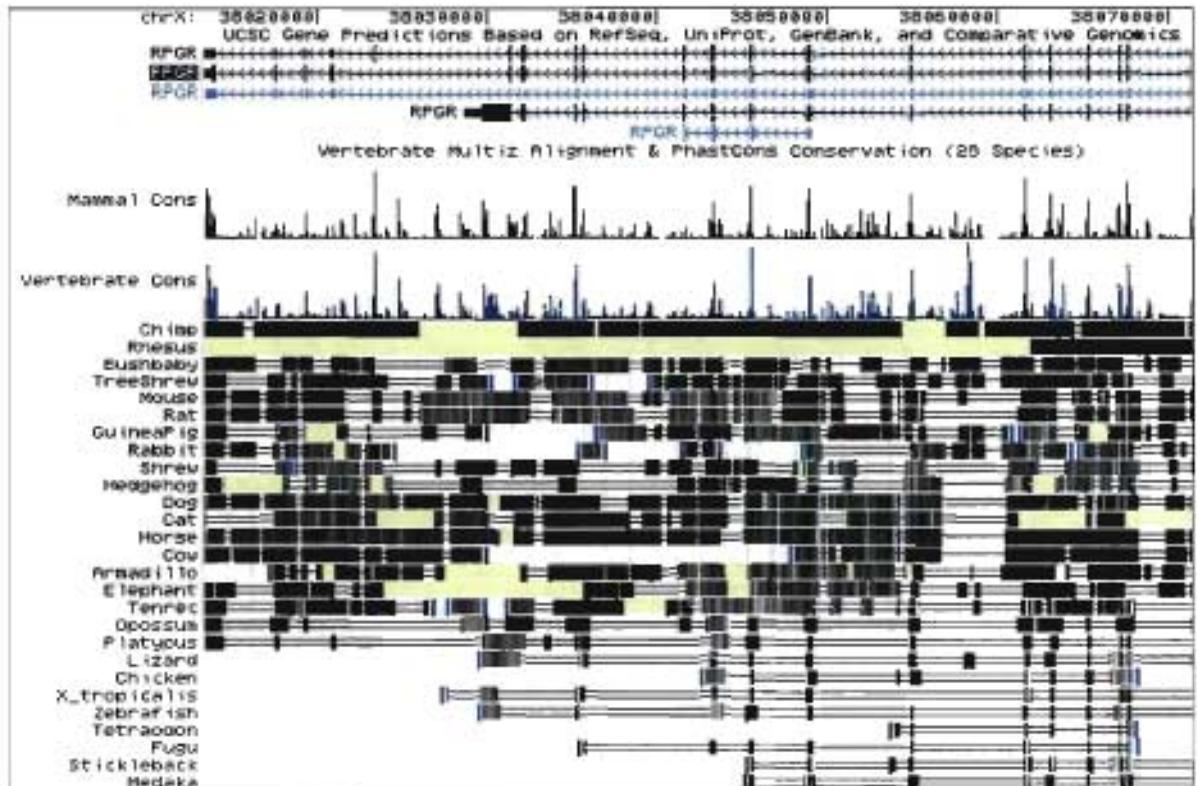


Figure 21 Alignements et conservation de séquences chez 28 espèces pour le gène RPRG

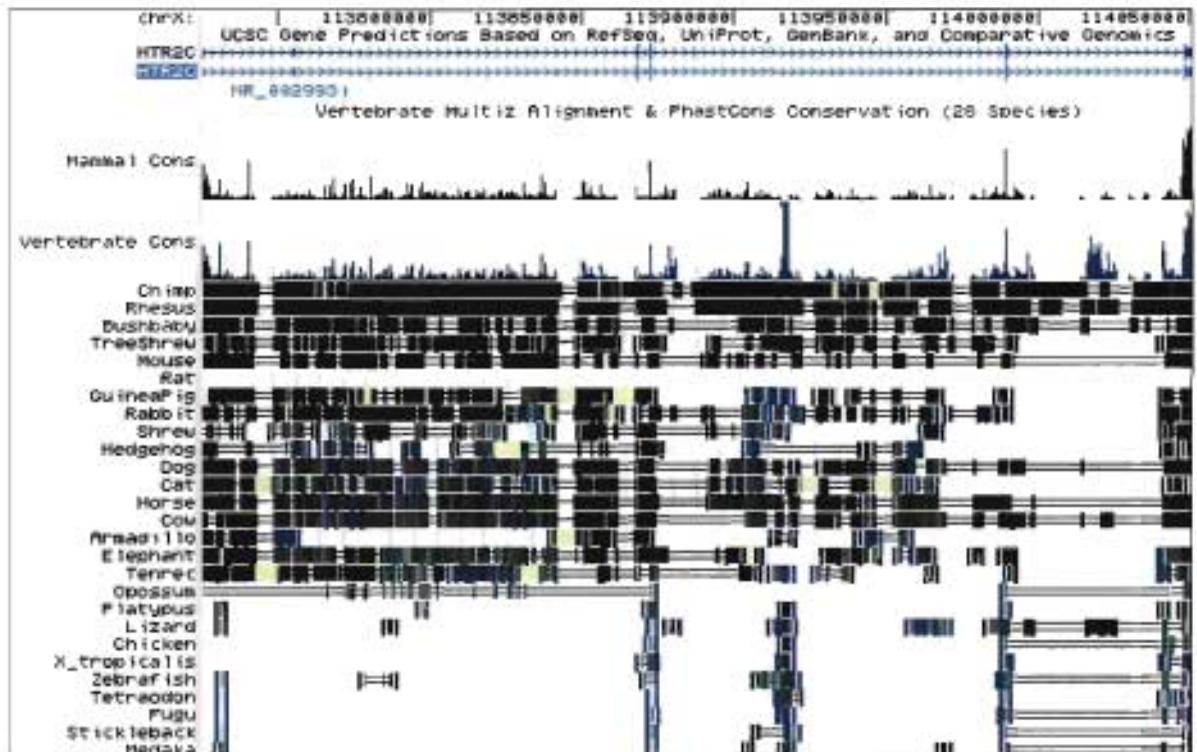


Figure 22 Alignements et conservation de séquences chez 28 espèces pour le gène HTR2C

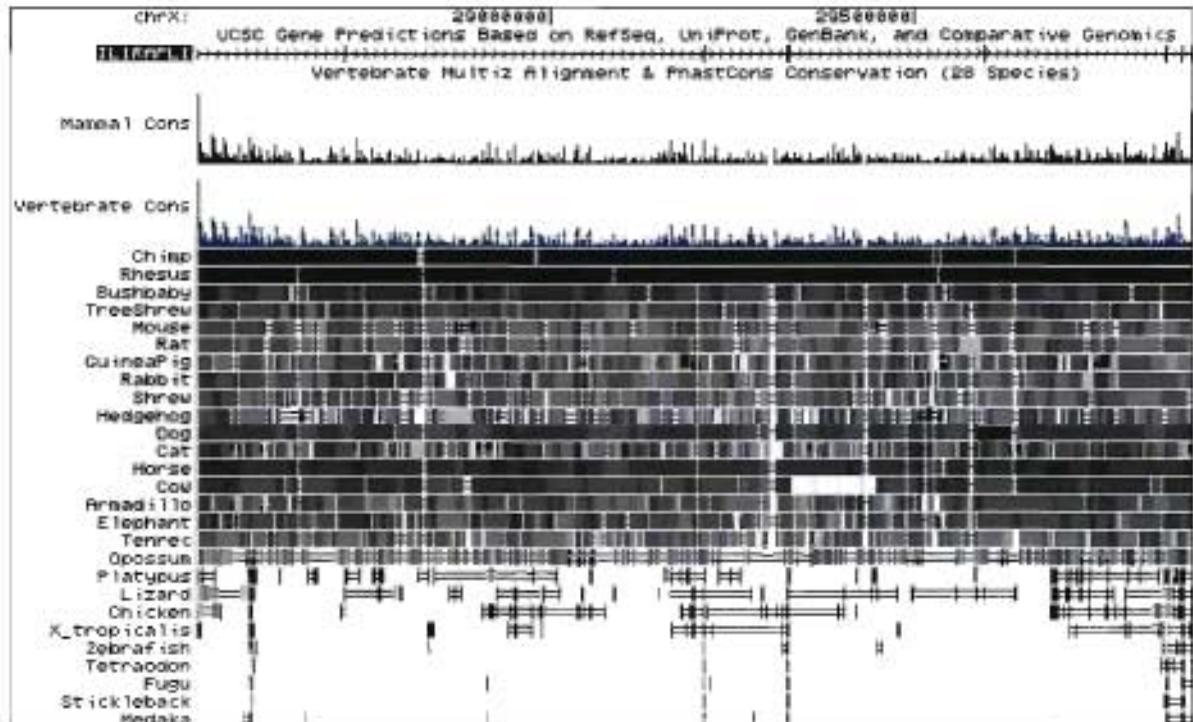


Figure 23 Alignements et conservation de séquences chez 28 espèces pour le gène IL1RAPL1

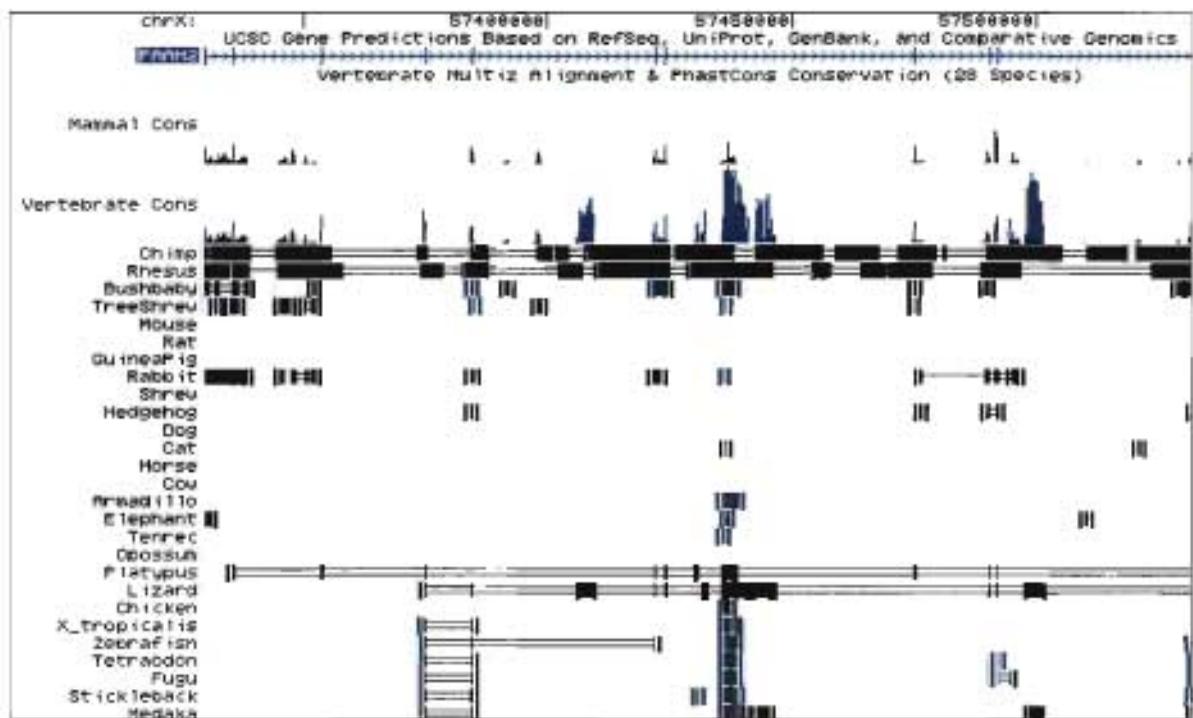


Figure 24 Alignements et conservation de séquences chez 28 espèces pour le gène FAAH2

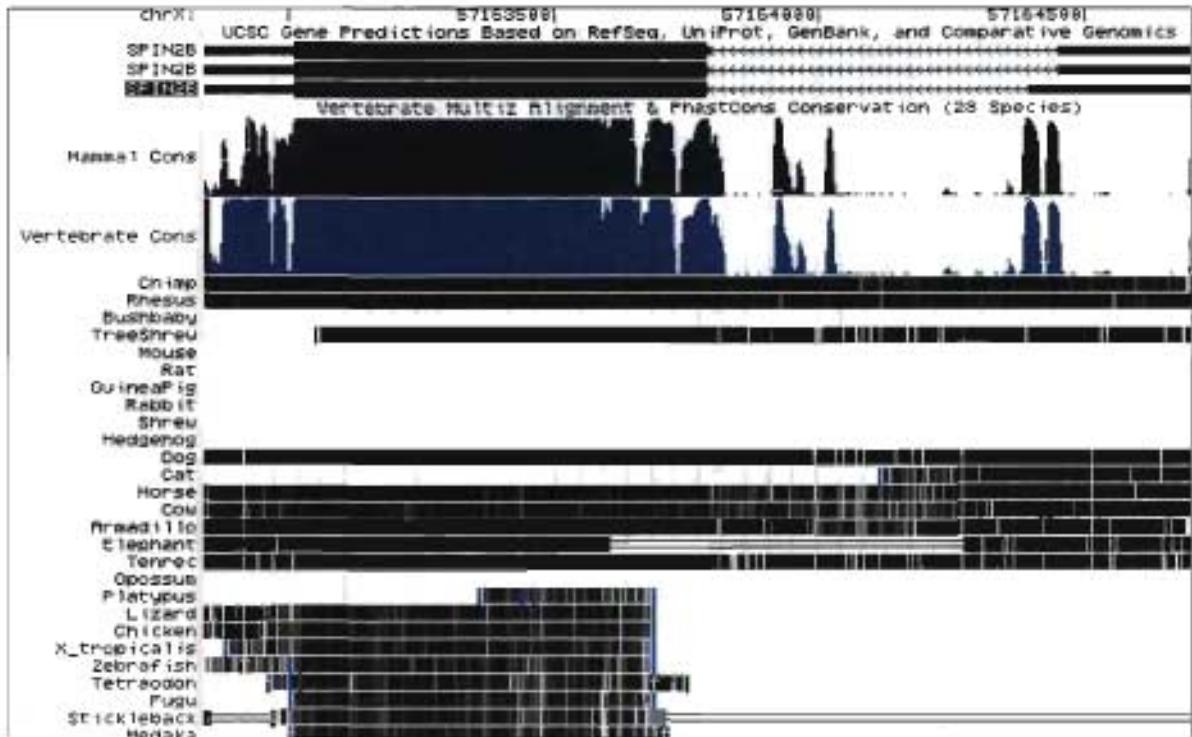


Figure 25 Alignements et conservation de séquences chez 28 espèces pour le gène SPIN2

Discussion

Chapitre 8 – SNPs et gènes d'intérêt découverts sur le chromosome X

8.1 SNPs et gènes d'intérêt

Sept SNPs ont été choisis pour être présentés dans ce mémoire étant donné leur haut niveau de significativité et aux gènes auxquels ils se rapportent. Pour le groupe des deux sexes réunis, les rs4131420 et rs1949393 ont montré une association significative (Tableau VI). Le rs4131420 est en association avec l'indice Milwaukee, une mesure de contractilité cardiaque. Plusieurs SNPs de la population caucasienne, en LD avec le rs4131420, sont localisés dans le gène RPGR de même que le rs10521413 du SLSJ qui forme un bloc d'haplotypes avec celui-ci (Tableaux XVI et XVII). Pour le moment, un défaut à l'intérieur ce gène est impliqué uniquement dans des pathologies de la vision; cependant, son expression dans différents tissus, dont le cœur, inclut la possibilité qu'il puisse jouer un rôle dans des pathologies autres que celles affectant l'œil. Ensuite, en examinant les gènes proches du SNP identifié, celui de l'ornithine carbamoyltransferase (OTC) a été repéré. Ce dernier se trouve à seulement 23 771 pb du rs4131420 et pourrait être impliqué dans la susceptibilité de désordres complexes comme les maladies cardiovasculaires (Figure 26). En effet, en métabolisant l'ornithine en citrulline, l'OTC joue un rôle important dans la voie de synthèse de la L-Arginine (Arg)¹²⁴. Celle-ci est un acide aminé qui agit comme substrat dans la synthèse de NO¹²⁵, un facteur relaxant essentiel dans la régulation du tonus vasculaire et de l'hémodynamique. L'Arg exerce aussi ses actions vasculaires par des mécanismes indépendants du NO tels que la sécrétion d'insuline¹²⁶. Une carence en Arg, un précurseur du NO, sera logiquement à l'origine d'une déficience en ce dernier. Le NO est un radical libre qui agit comme agent antioxydant. Pour plusieurs organes, il fournit une protection contre l'action des superoxydes (O₂⁻), en l'éliminant des tissus biologiques; toutefois, la superoxyde dismutase accomplit principalement la réduction de O₂⁻¹²⁷. L'O₂⁻, un produit du métabolisme cellulaire, a le pouvoir d'endommager les organes. Il est donc possible que la déficience de NO engendre une accumulation croissante d'O₂⁻¹²⁸. L'O₂⁻ peut aussi désactiver le NO pour former du peroxynitrite réduisant la disponibilité de NO¹²⁹. Dans l'hypertension, il a été rapporté que la libération excessive d'O₂⁻ désactive le NO aussitôt qu'il est produit menant au développement d'une pression artérielle élevée¹³⁰.

Cuzzocrea *et al.*, ont enlevé l'O₂⁻ à l'aide d'une superoxyde dismutase synthétique, ce qui a eu pour effet de rétablir la pression artérielle proche des valeurs normales¹³¹.

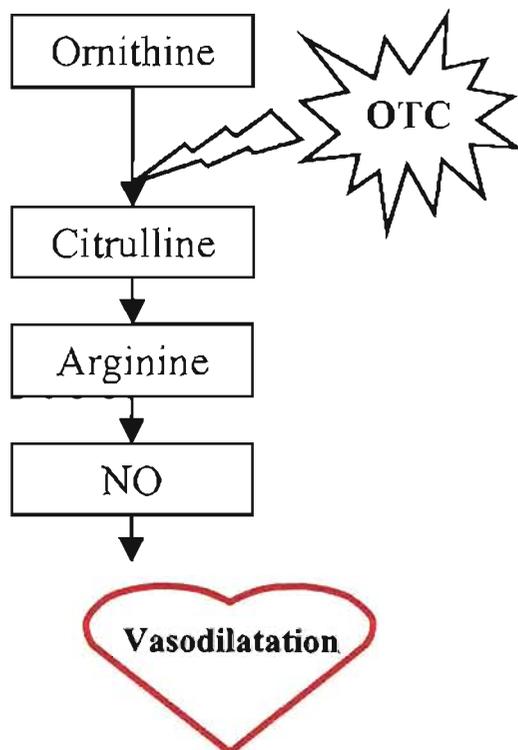


Figure 26 Lien possible entre OTC et l'hypertension

Pour le second SNP, le rs1949393, les associations disparaissent lorsque les deux sexes sont analysés séparément (Tableaux VI et VII) relevant soit d'une cause biologique ou simplement d'une puissance statistique insuffisante. En effet, la séparation par sexe a pour conséquence de diminuer considérablement la taille des échantillons sur lesquels les tests sont faits, diminuant ainsi la précision des estimés. Dès lors, l'étude n'a plus la puissance statistique adéquate pour détecter la différence désirée. Certes, plus la taille de l'effet clinique ciblé est faible, plus le nombre d'individus requis pour établir la significativité doit être grand ; de même que, si l'échantillon est trop petit, le test devient susceptible aux erreurs de type II (faux négatifs)¹³². D'après le tableau VII, le rs1949393 semble avoir davantage un lien avec la pression artérielle qu'avec l'indice de masse corporelle. En effet, en divisant les individus en sous-catégories, la plus forte association se trouve chez les 2 sexes dans la sous-catégorie des hypertendus; mais, lorsque les sujets sont regroupés selon leur statut d'hypertension et d'obésité, les associations disparaissent. Ni ce SNP et ni le rs1589075 du SLSJ en LD complet avec celui-ci ne se situe dans un gène (Tableau XVII). Ceci était prévisible puisque la région Xq21.3 est pauvre en gène et

présente plusieurs points de rupture¹³³. Le registre des microRNAs de l'institut de Sanger (<http://www.sanger.ac.uk/Software/Rfam/mirna/>) ne soulève aucun microRNA pour cette région de même que la littérature n'y rapporte pas l'existence d'un désert de gènes. Toutefois, les travaux de Ross *et al.* indiquent qu'une des homologies entre le chromosome X et Y se trouve dans une large région nommée *X-transposed region* (XTR, Figure 8). Celle-ci correspond justement à Xq21.3 et Yp11.2-p11.3 et fut dupliquée du chromosome X au Y depuis la divergence des lignées de l'humain et du chimpanzé⁵³. Il reste que le gène le plus proche du rs1949393 est le *kelch-like 4* (KLHL4) qui est à 222 837 pb. Ce dernier est caractérisé par des motifs *kelch* répétés, un domaine de liaison aux protéines, POZ/BTB, et est exprimé au niveau du cœur foetal¹³⁴. Ces motifs répétés seraient des domaines de liaison à l'actine. La fonction spécifique de ce gène n'a pas encore été déterminée; cependant, celui-ci est situé dans la région critique de palatoschisis liée au chromosome X¹³⁴. Afin de recueillir plus d'informations sur cette région, les recherches ont été portées chez le rat (*Rattus norvegicus*, Xq32) et la souris (*Mus musculus*, XE1) en cherchant la synténie à l'aide de KLHL4 (Figure 27). Malheureusement, pour ces deux espèces, aucun article ne porte sur cette région; par ailleurs, chez l'humain plusieurs gènes et loci localisés dans cette région sont rattachés à diverses maladies.

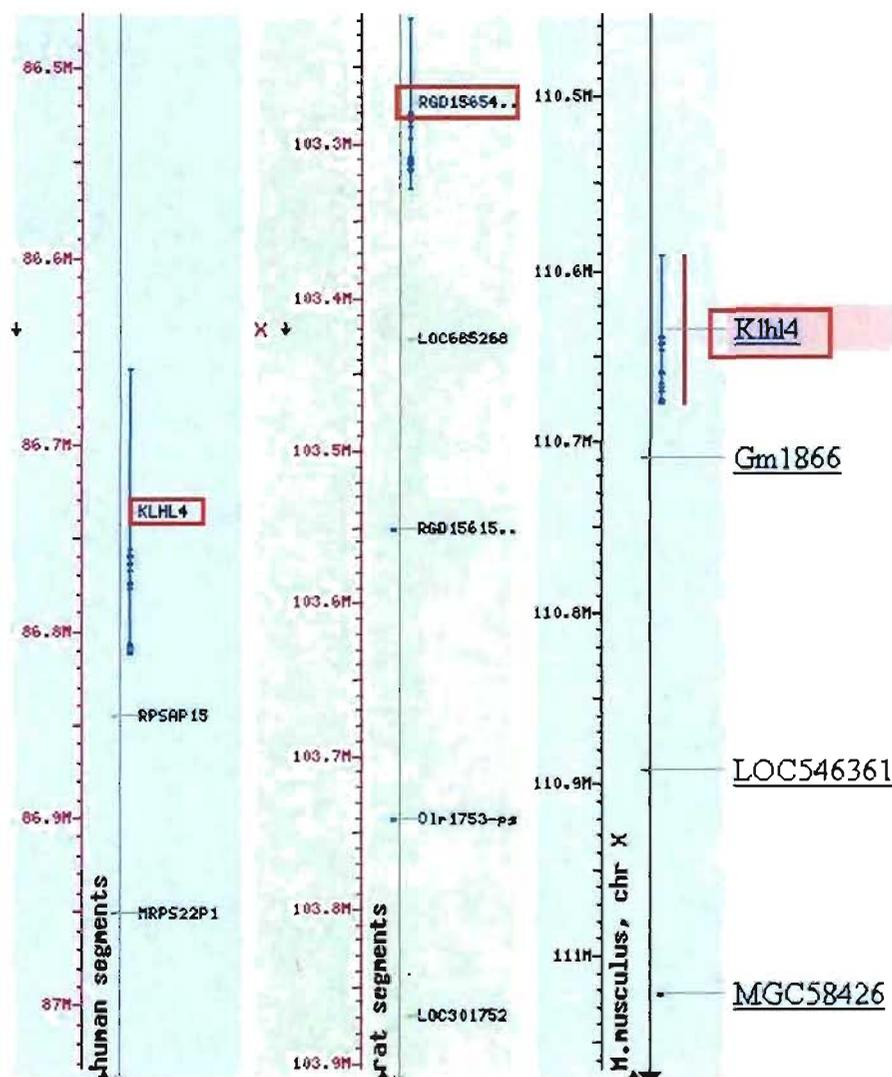


Figure 27 Représentation de la synténie, chez le rat, la souris et l'humain, pour la région entourant le gène KLHL4 (rectangles rouges), tirée de www.NCBI.nlm.nih.gov

Comme l'indique le tableau VI, le rs727562 est associé au même phénotype à l'intérieur de deux groupes : femmes et tous. L'association trouvée chez les deux sexes ensemble représenterait en réalité la forte association présente chez les femmes. Le tableau XI, contenant les données des sous-catégories, démontre bien ce phénomène. D'après ce tableau, il peut être constaté qu'une seule association apparaît chez les deux sexes comparativement à trois chez les femmes et aucune chez les hommes. Donc, l'association chez les femmes serait suffisamment forte pour être détectée même en présence d'hommes dans l'échantillon; par conséquent, l'impact du rs727562, sur la PAS durant le jour, serait spécifique à la femme. Le rs727562 aussi bien que les SNPs caucasiens et le rs10521953 du SLSJ en LD avec celui-ci se situent dans les introns du gène *interleukin 1 receptor*

accessory protein-like 1 (IL1RAPL1) (Tableaux XVI et XVII). Ce gène, de la région Xp21, encode une protéine appartenant à une nouvelle classe de la famille de récepteurs II-1/Toll, sans toutefois être une protéine réceptrice pour Il-1¹³⁵. Les résultats rapportés par Bahi *et al.* suggèrent qu'IL1RAPL1 pourrait représenter un régulateur d'exocytose dépendant du calcium¹³⁵. Ce gène est exprimé dans différents tissus humains dont le rein fœtal et le cœur adulte^{136, 137} et tout comme les récepteurs d'interleukine 1, il contient le domaine de récepteur Toll/interleukine-1. Ce dernier est crucial dans la transduction du signal lors de la réponse immunitaire de l'hôte¹³⁸; cependant, d'autres recherches sont nécessaires afin d'évaluer le rôle d'IL1RAPL1 dans la régulation immunitaire et la défense de l'hôte. Une accumulation de plusieurs évidences impliquerait le système immunitaire dans la pathogenèse de l'hypertension; en effet, les personnes hypertendues démontrent un niveau plus élevé d'IgG dans le sérum¹³⁹, de globules blancs, de certaines sortes de cellules T que celles normotendues¹⁴⁰. Il a été démontré que l'interleukine-6, une cytokine messagère menant à l'amélioration de l'inflammation, serait corrélée positivement avec l'indice de rigidité des larges artères¹⁴¹ et représenterait un facteur de risque indépendant pour une pression artérielle élevée chez des sujets en santé¹⁴². De plus, Dalekos *et al.* ont rapporté que les concentrations plasmatiques d'interleukine-1 β étaient davantage élevées chez les sujets souffrant d'hypertension essentielle¹⁴³. Une évidence de liaison a aussi été rapportée, par Cucca *et al.*, entre la région Xp21-p11 et le diabète de type 1¹⁴⁴.

Strictement chez les femmes, les rs2116366 et rs4911871, sont associés à la PAS (durant 24 heures et la nuit) et l'indice Heather (indice de contractilité cardiaque) respectivement (Tableau VI). D'après les résultats obtenus, les allèles du rs2116366 démontrent une implication importante dans la régulation de la pression artérielle systolique et ce, surtout chez les femmes et en particulier celles hypertendues et non obèses (Tableaux IX et X). Tout comme le rs727562, l'impact du rs2116366, sur les deux phénotypes auxquels il est associé, serait spécifique à la femme. Ce SNP (Xp11.1) est situé dans un des introns du gène *fatty acid amide hydrolase 2* (FAAH2) (Figure 28). Les acides gras amidés, une classe de transmetteurs lipidiques, constituent le substrat de l'enzyme FAAH. Cette classe de transmetteurs renferme l'anandamide cannabinoïde endogène qui possède une activité hypotensive et cardiopressive en plus de l'oléamide, une substance induisant le sommeil. En hydrolysant les acides gras amidés, la FAAH contrôle l'amplitude

et la durée du signal de ceux-ci (Figure 29). L'expression de FAAH2 a été observée dans plusieurs tissus humains, dont le cœur et les ovaires¹⁴⁵. Batkai *et al.*, en étudiant l'effet des endocannabinoïdes (l'anandamide) sur l'hypertension chez les rats spontanément hypertendus (SHR), ont observé une augmentation de l'expression de FAAH1 au niveau du myocarde et par le fait même une diminution des niveaux d'anandamide dans cette souche¹⁴⁶. D'après leurs données, ils concluent que les endocannabinoïdes suppriment la contractilité cardiaque dans l'hypertension et qu'une amélioration des effets vasodilatateurs de l'anandamide en bloquant sa dégradation peut normaliser la pression artérielle¹⁴⁶. Cette découverte cible les endocannabinoïdes comme nouvelle cible thérapeutique dans le traitement de l'hypertension. Il reste à savoir si FAAH2 pourrait réguler de la même manière le signalement des endocannabinoïdes au niveau du cœur chez l'humain. En plus de couvrir le gène FAAH2, la région de LD complète formée du rs2116366 et de trois autres SNPs (rs7050676, rs204166, rs10521487) de la population du SLSJ, englobe aussi le gène SPIN2 (*spindlin family member 2*) (Tableau XVII). Il s'agit d'une nouvelle protéine nucléaire régulant la progression du cycle cellulaire et dont l'activité est liée à l'inhibition de l'apoptose¹⁴⁷; par contre, aucun lien n'est fait dans la littérature entre SPIN2 et l'hypertension.



Figure 28 Gène FAAH2, le rs2116366 se situe dans l'intron 7

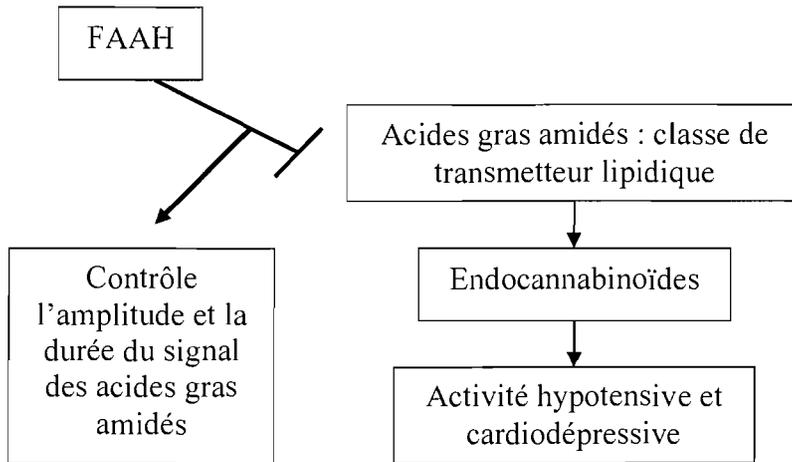


Figure 29 Rôle de l'enzyme FAAH

L'association la plus significative chez les femmes fut celle entre le rs4911871 (Xq23) et l'indice Heather (Tableau VI). Ce dernier se situe dans un des introns du gène *5-hydroxytryptamine (serotonin) receptor 2C* (HTR2C). Le rs4911871 en plus de sept SNPs du SLSJ (rs569959, rs539748, rs2497522, rs2428707, rs950463, rs5946229, rs1335614) font partis d'un large bloc d'haplotypes s'étendant sur 327 kb (Tableau XVII). En plus du rs4911871, quatre SNPs de ce bloc sont localisés dans les régions introniques d'HTR2C. De plus, plusieurs SNPs caucasiens en LD avec le SNP identifié sont aussi contenus à l'intérieur d'HTR2C dont le rs1801412, situé dans l'exon 6 et qui possède une fonction 3' UTR et le rs6318, positionné dans l'exon 4 et qui a une fonction non synonyme (Tableau XVI). Des études ont établi un rôle pour le récepteur HTR2C dans l'obésité et la résistance au diabète de type II¹⁴⁸, deux désordres complexes impliqués dans l'hypertension. De plus, la sérotonine représente une substance vasomotrice qui en liant différents récepteurs mène à une vasoconstriction ou à une vasodilatation. Par des effets périphériques et centraux comme la régulation de la pression artérielle, cette substance pourrait être responsable de certaines formes d'hypertension¹⁴⁹ et ce peut-être en faisant intervenir HTR2C (Figure 30). Des antagonistes des récepteurs de sérotonine, tels que la kétansérine, ont été proposés comme agents thérapeutiques pour les maladies cardiovasculaires incluant l'hypertension¹⁵⁰. Par une analyse de liaison, Pajukanta *et al.* ont découvert deux loci, dont un sur le chromosome Xq23-26 qui serait susceptible de jouer un rôle dans les maladies coronariennes¹⁵¹. Il est connu que les hommes sont à plus haut risque de développer une maladie coronarienne (CHD) que les femmes ; le ratio hommes-femmes étant de 2 :1, pour

une incidence fatale de la maladie et ce, dans plusieurs pays avec des habitudes de vie différentes¹⁵². Tout comme l'hypertension, la présence de CHD chez les femmes en préménopause est moindre que celles postménopausées. En effet, l'étude *Framingham* de Gordon *et al.* indique que le taux d'incidence de CHD chez les femmes postménopausées est plus que le double de celui des femmes en préménopause et du même âge¹⁵³. Déjà impliqué dans l'hypertension, un effet protecteur de l'œstrogène a aussi été suggéré pour ce phénomène, mais reste controversé. Effectivement, l'étude de Lawlor *et al.* indique plutôt que le dimorphisme sexuel dans la CHD résulterait des facteurs environnementaux¹⁵⁴. Un des SNPs (rs569959) du bloc d'haplotypes se retrouve aussi dans l'exon 1 du LOC727794 qui n'est qu'une protéine hypothétique pour le moment. De plus, ce bloc couvre les gènes SNORA35, un petit ARN du nucléole, et MIRN448, un microRNA, pour lesquels aucun article n'est publié ; néanmoins, les microRNAs pourraient être impliqués dans l'hypertension aussi bien que les maladies cardiaques^{155, 156} (Tableau XVII).

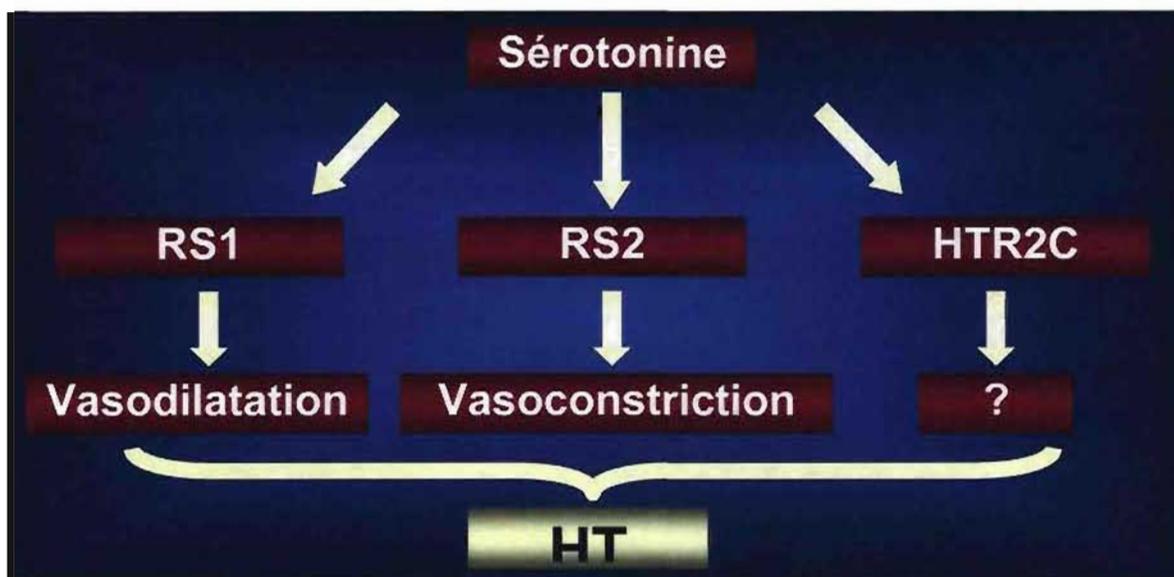


Figure 30 Rôle de la sérotonine et rôle possible du récepteur HTR2C dans l'hypertension; récepteur :R

Trois associations ont été identifiées chez les hommes exclusivement et font intervenir les rs438102, rs407890 et le bloc d'haplotypes d'ACE2 (Tableau VI). L'association la plus significative est celle entre le rs407890 et la fréquence cardiaque. Le statut d'hypertension et d'obésité ne sont pas à l'origine de cette association comme le démontre le tableau VIII, de même qu'il n'y a pas de relation génotype/phénotype chez les hommes. Dans ce cas, l'association entre ce SNP et la fréquence cardiaque ferait peut-être

intervenir un phénotype intermédiaire. Ce SNP forme un bloc d'haplotypes avec le rs223724, mais aucun d'eux n'est compris à l'intérieur d'un gène, ni même en couvre un (Tableau XVII); toutefois, le gène *TGFB-induced factor 2-like, X-linked* (TGIF2LX) s'avère être le plus proche du rs407890, à une distance de 364 318 pb. Il s'agit d'un gène de boîte homéotique situé dans la région Xq21 qui est exclusivement exprimé dans les testicules adultes, ce qui lui suggère un rôle dans la spermatogenèse¹⁵⁷. De son côté, le bloc d'haplotypes constitué du rs438102 et de deux autres SNPs du SLSJ (rs398262 et rs1929364) n'englobe aucun gène (Tableau XVII), mais il est à noter que le rs438102 est à proximité de trois gènes pour lesquels aucune information n'est disponible (LOC139363, LOC392552 et FLJ30672).

Le résultat le plus intéressant, pour les hommes, est sans contredit l'association entre le bloc d'haplotypes du gène ACE2 (rs2074192 et rs233575) et l'indice de masse ventriculaire gauche (LVMI), tous deux ciblés dans les recherches portant sur l'hypertension. ACE2 est impliqué dans le maintien de l'équilibre de la pression artérielle ; tandis que, le LVMI sert de mesure pour l'hypertrophie du ventricule gauche (LVH). Ce dernier réfère à l'épaississement du muscle cardiaque du ventricule gauche qui peut résulter d'une pression artérielle élevée. Au début de l'hypertension, le LVH représente un phénomène physiologique compensatoire, mais après un certain temps le muscle cardiaque s'use et défailit. Dans l'hypertension essentielle, les sujets ayant un LVH montrent un risque plus élevé de développer une complication cardiovasculaire dans le future que ceux ayant une masse du ventricule gauche normale¹⁵⁸. Le LVH est aussi un facteur de risque, indépendant de la pression artérielle, des maladies cardiovasculaires¹⁵⁹. L'étude de Lieb *et al.* en 2006 confirme ce résultat ; en effet, en étudiant cinq SNPs de ACE2 dans une population allemande, ceux-ci ont aussi découvert une association entre un bloc d'haplotypes, contenant le rs233575, et le LVMI ($p = 0,027$) uniquement chez les hommes¹⁶⁰. Concernant uniquement ACE2 notre groupe de recherche a démontré récemment une association entre l'hypertension, des phénotypes liés à l'hypertension et ce gène, dans la population du SLSJ (Labuda *et al.*, article en préparation). En plus de son interaction avec le système rénine-angiotensine, il a été constaté récemment qu'ACE2 hydrolyse trois autres classes de peptides dont l'apeline¹⁶. Il est intéressant de noter qu'ACE2 et le gène APELIN sont tous deux situés sur le chromosome X, au niveau de la

bande cytogénétique Xp22 et Xq25 respectivement. L'apeline représenterait un neuropeptide diurétique potentiel qui agirait contre les actions de la vasopressine par l'inhibition de son activité neuronale et de sa libération¹⁶¹.

8.2 Impact d'un SNP sur le phénotype

Selon les résultats, il peut être constaté que les SNPs ne sont pas tous associés au même phénotype et à l'opposé, deux SNPs différents peuvent être associés au même phénotype. Afin d'explorer la relation entre les désordres génétiques humains et les gènes correspondant à ces maladies, Goh *et al.* ont créé la notion de *diseasome* : combinaison de tous les désordres génétiques humains connus (*disease phenome*) et de la liste complète des gènes associés (*disease genome*)¹⁶². Cette combinaison est représentée par un réseau défini par un graphique bipartite de deux ensembles de nœuds disjoints. À partir de ce graphique, deux projections/réseaux fournissant des informations biologiques très pertinentes sont obtenues. Le *diseasome* offre donc une référence visuelle rapide des liens génétiques entre les désordres et les gènes¹⁶², permet de mieux comprendre les mécanismes génétiques de ces pathologies et révèle des connexions pas encore suspectées¹⁶³. Cet outil pourrait certainement aider à disséquer l'aspect polygénique et hétérogène de l'hypertension.

Un SNP peut résulter en un impact physiologique lors des différentes étapes régulatrices de l'expression d'un gène. Si au départ, la population scientifique était davantage intéressée au SNP à l'intérieur d'un exon, plusieurs études prouvent qu'il faut aussi s'intéresser à la relevance biologique de ceux dans les introns¹⁶⁴. Chacune des étapes de la formation d'une protéine peut être perturbée par un SNP. Menzaghi *et al.*, ont rapporté un SNP qui, positionné dans la région promotrice, altère la transcription et est associé à une pression artérielle systolique élevée¹⁶⁵. Un SNP a aussi le pouvoir de modifier à la hausse aussi bien qu'à la baisse un phénomène biologique. Par exemple, Jabocson *et al.* et Yamaguchi *et al.* ont rapporté des SNPs à l'origine d'effets opposés soit une augmentation et diminution de l'efficacité de la traduction à l'origine d'une pathologie¹⁶⁶,¹⁶⁷. Au niveau protéinique, un SNP peut altérer défavorablement le repliement d'une protéine, par la création de nouvelles liaisons¹⁶⁸. De plus, différentes formes d'un gène ou de variants génétiques peuvent résulter en un effet protecteur ou en l'effet inverse¹⁶⁹. Même la région 3' UTR peut être la cible d'une défaillance due à un SNP; en effet, ces SNPs sont

fonctionnels et peuvent affecter l'appariement de facteur de transcription comme l'indique l'étude de Chen *et al.* sur la maladie coronarienne¹⁷⁰ ainsi que la stabilité de l'ARNm¹⁷¹.

Étant donné que les SNPs identifiés se trouvent au niveau des introns et des régions intergéniques, une prédiction de leur impact a pu être faite uniquement pour les sites de liaison aux facteurs de transcription (TFBS). Les analyses des SNPs introniques, avec F-SNP, MAPPER et SNPInspector, ont révélé que tous ces SNPs altèrent des TFBS (Tableaux XVIII et XIX). Malgré cela, comme mentionné dans la méthodologie, il y a une faible évidence que ces TFBS soient vraiment fonctionnels, mais ceci apporte du moins la possibilité qu'un SNP non codant puisse affecter la régulation de l'expression génique. Il est connu que les TFBS sont d'un contenu pauvre en information à cause de leur courte séquence et que ceux-ci peuvent donc survenir par hasard à plusieurs endroits dans le génome^{172, 173}. Cependant, quelques évidences dans la littérature proposent l'existence de TFBS au niveau des introns. Par exemple, en étudiant l'impact de SNPs situés dans l'intron 1 du gène SNAP-25, Gosso *et al.* ont découvert une liste putative des TFBS modifiés par ces derniers. Ils suggèrent que même si les TFBS représentent des portions constitutives du génome, le contexte cellulaire et génomique déterminera si la séquence d'un facteur de transcription deviendra fonctionnelle ou non¹⁷⁴. Par l'étude d'un type de lupus, Prokunina *et al.*, ont aussi découvert que l'allèle d'un SNP intronique du gène PDCD1 altère un site de liaison pour le facteur de transcription RUNX1⁸⁴; résultats qu'ils ont confirmés par une technique de mobilité électrophorétique.

Concernant les SNPs identifiés lors de ce projet, les trois programmes utilisés (F-SNP, MAPPER et SNPInspector) ont rapporté des résultats assez différents (Tableaux XVIII et XIX). Ceci peut s'expliquer par l'algorithme ou la méthode de recherche employé par le logiciel ou par le fait que ces programmes ne possèdent tout simplement pas les mêmes séquences de TFBS dans leur base de données. Aussi, le logiciel MAPPER prédit uniquement une liste des TFBS pour différentes régions d'un gène choisi et ce, sans tenir compte des changements survenant avec les différents allèles des SNPs ; tandis que, F-SNP et SNPInspector ciblent plutôt la séquence flanquant un SNP sélectionné. Selon les résultats du tableau XIX, il peut être constaté que la présence de l'allèle non ancestral aboutit globalement autant à des gains de TFBS (44% des changements) qu'à des pertes (56%) ;

par contre, en regardant un SNP précis, l'allèle non ancestral peut provoquer, par exemple pour le rs539748, davantage de perte que de gain. Ceci dépend évidemment de la séquence flanquant le SNP en question, mais ce ratio pourrait être un indice sur l'effet néfaste ou le bienfait biologique des allèles d'un SNP. Comme les outils bio-informatiques choisis ne font que des suppositions de changements de TFBS, il reste à valider en premier la véracité de la grande quantité des TFBS prédits. Par la suite, il faudra déterminer l'impact réel, dans un intron, d'une perte/gain des TFBS au niveau de la régulation de la transcription et de la formation des complexes de facteurs de transcription. Puis, vérifier si l'altération des TFBS résultant en une régulation aberrante d'un gène pourrait mener à une pathologie comme l'hypertension. La prédiction d'impact des SNPs à l'aide d'outils bio-informatiques représente une méthode très intéressante qui peut être accomplie avant le génotypage de sujets pour un gène. En effet, elle offre de l'information très utile permettant de bien sélectionner les SNPs à génotyper en se basant sur les résultats de prédiction.

8.3 Conservation de séquences entre les espèces

En examinant la conservation de séquences entre 28 espèces via le site UCSC, il a été observé que 19 des 25 SNPs ne sont pas positionnés dans une région conservée. La totalité des SNPs intergéniques en fait parti. Dans certains cas ceci est simplement le résultat d'un manque de données sur les séquences pour la majorité des espèces. Il se peut également qu'étant de moins grande importance, ces régions intergéniques soient restées propres aux espèces. La sélection naturelle aurait fait en sorte qu'elles ne soient pas conservées au cours de l'évolution et plusieurs mutations ont pu s'y glisser, diminuant l'exactitude de l'appariement pour ces séquences. Chez les femmes, il pourrait s'agir de point de recombinaison; tandis que, chez l'homme aucun des gènes couverts par ces SNPs n'appartient aux 29 gènes contenus dans la région pseudoautosomale du chromosome X, où à lieu la recombinaison entre ce chromosome et le Y¹⁷⁵. Il faut savoir, par contre, que des régions entre les gènes contenant des éléments régulateurs, comme un amplificateur de transcription, ont été identifiées¹⁷⁶.

De leur côté, les rs10521413, rs569959 et rs233575, figurent à une distance de moins de 20 pb d'une région bien conservée chez la plupart des mammifères placentaux et les primates, mais qui ne l'est pas chez le reste des vertébrés (Figures 14-16). Ceci

s'explique par la divergence dans l'évolution de ces groupes d'espèces qui ne se trouvent pas sur les mêmes branchements dans l'arbre phylogénétique (Figure 12). Cette divergence évolutive se traduit donc par un bagage génétique distinct pour certaines régions du génome qui confère à ces groupes leur spécificité et leurs caractéristiques. À l'opposé, pour les régions préservées chez les primates et les mammifères placentaux, ceci peut indiquer qu'il s'agit de régions indispensables et cruciales pour ces espèces et qu'elles sont essentielles à leur survie. Alors, la présence d'un variant à proximité de ces régions conservées peut sans aucun doute altérer leur fonction et leur intégrité et par conséquent, mener à un désordre biologique.

De son côté, le rs950463, situé en aval d'HTR2C, se trouve entre deux régions relativement bien conservées uniquement chez les primates; tandis que, les mammifères placentaux ont plusieurs bases incapables d'être alignées pour cette région (Figure 17). Étant donné que ce gène est aussi présent chez d'autres espèces dont le rat et la souris, ceci est possiblement dû à la distance évolutionnaire entre les espèces. En effet, les primates étant les plus proches des humains, une meilleure similarité de séquences peut survenir entre ces deux espèces. Pour les deux SNPs restants, rs1335614 et rs4911871, ceux-ci sont proches de régions moins bien préservées (Figures 18 et 19). Encore une fois, la distance au niveau de l'évolution peut être évoquée afin d'expliquer cette observation en plus que certaines espèces n'ont pas de séquence disponible pour cette région.

Après avoir examiné le niveau de conservation des régions où sont présents les SNPs identifiés, la conservation des gènes associés aux blocs d'haplotypes a été étudiée. Pour ACE2, tous les exons démontrent une très bonne conservation surtout chez les mammifères placentaux en plus de la présence d'un haut pic de conservation dans l'intron 2. L'importance de ce gène chez les mammifères placentaux aurait favorisé la sauvegarde de sa structure à travers l'évolution, précisément celle des exons qui sont à l'origine de la protéine produite. Quant à l'intron 2, celui-ci pourrait contenir un élément régulateur essentiel à la bonne transcription de la séquence d'ADN (Figure 20).

Pour le gène RPGR, une préservation apparaît chez la plupart des exons chez les vertébrés et chez les mammifères placentaux exclusivement (Figure 21). Kirschner *et al.*,

ont comparé la séquence d'ADN de ce gène entre l'humain et la souris¹⁷⁷. Leurs résultats démontrent que tous les exons de RPGR sont conservés entre ces espèces tout en montrant un degré variable de conservation ainsi que les introns exposent un bon degré d'homologie. Les exons 3-6 détiennent la meilleure homologie, ce qui n'est pas reflété dans notre analyse. Il faut comprendre que le graphique de conservation dans notre cas, est obtenu par la somme des résultats d'appariements chez plusieurs espèces et non uniquement d'un appariement entre l'humain et la souris. De plus, selon eux, une meilleure constance d'homologie est trouvée pour l'intron 14, ce qui correspond à nos résultats. Ils justifient la conservation prononcée des introns par le fait que l'expression tissu spécifique et l'épissage requièrent plusieurs éléments de contrôles importants et conservés durant l'évolution.

Un patron similaire à RPGR est suivi par HTR2C, mais ce gène comporte davantage de petits bouts de régions non conservées (Figure 22). Même si ces deux gènes sont présents chez plusieurs espèces, leur séquence d'ADN peut avoir subi un certain nombre de changements au cours de l'évolution afin de répondre au besoin de chaque espèce. Pour sa part, IL1RAPL1 montre un taux de conservation similaire sur toute sa longueur et les exons ne sont pas la cible principale de la conservation indiquant que la bonne intégrité de tous les constituants de ce gène pourrait être nécessaire à une expression génique adéquate (Figure 23).

FAAH2 représente le gène le moins conservé de tous ceux étudiés durant ce projet; seuls deux primates montrent une homologie intéressante (Figure 24). Ceci était prévisible puisque ce gène vient tout récemment d'être découvert chez l'humain et n'a pas encore été découvert avec certitude chez d'autres espèces; toutefois, sur le site de NCBI, un locus similaire à FAAH2 est rapporté pour l'espèce *Macaca mulatta*, une sorte de primate. De son côté, SPIN2 détient une région hautement conservée autant chez certains mammifères placentaux que chez des vertébrés qui ne sont pas dans cette catégorie (Figure 25). Étant donné que ce gène a persisté durant l'évolution et qu'il renferme un unique exon, il va de soit que celui-ci soit bien préservé.

Pour les régions conservées, il faut faire attention à ne pas mal interpréter ce que cela signifie. Pour une séquence ne s'alignant pas sur la complète longueur du gène à l'étude, il peut s'agir uniquement d'un appariement de similarité de séquence par hasard et qui n'a donc rien à voir avec les éléments constitutionnels d'ADN de ce gène. Puis, même dans la possibilité d'un alignement sur toute la longueur du gène, s'il y a des régions d'incertitude, il peut s'agir d'un gène complètement différent de celui ciblé. D'autres études sont, à l'évidence, requises pour confirmer les régions de conservation et pour déterminer les facteurs à l'origine de leur conservation.

8.4 Associations différentes entre les sexes

Afin de respecter l'architecture du chromosome X, les analyses génétiques ont été exécutées sur trois groupes différents : deux sexes réunis, femmes et hommes. Les associations obtenues lors de ce projet se répartissent au sein de ces groupes. Au départ, il est bien connu que plus l'échantillon est grand, meilleure sera la puissance de détection. D'où le plus grand nombre d'associations rapportées dans le tableau VII pour le groupe réunissant les 2 sexes.

Le fait qu'un seul sexe soulève une association peut provenir de l'effet du genre sur le phénotype étudié. Par exemple, en ce qui concerne la pression artérielle, les deux sexes montrent un profil du phénotype clairement distinct et explicable par des hormones sexuelles (Figures 5 et 6)¹⁷⁸, la stature et l'arbre artériel qui sont plus courts chez la femme. Il a aussi été démontré que la réponse, face aux médicaments antihypertenseurs, n'est pas la même chez les deux sexes¹⁷⁹⁻¹⁸¹. Pour le LVMI, Marcus *et al.*, ont montré qu'une augmentation du LVMI serait le résultat de différents déterminants chez les hommes et les femmes¹⁸² ; par conséquent, ceci pourrait amener différentes associations selon le sexe.

Une interaction entre le sexe et le génotype peut aussi être à l'origine d'une association spécifique au sexe. Dans le but d'identifier les déterminants génomiques spécifiques au sexe de l'hypertension, Hamet *et al.* ont effectué systématiquement l'héritabilité et une analyse de liaison multipoint (microsatellites) spécifique au sexe ainsi qu'une analyse de liaison bipoint (SNPs, blocs d'haplotypes) sur 608 traits à partir de 120 familles du SLSJ⁹⁵. La présence d'une interaction entre le sexe et le génotype a été assurée

par l'exécution de 10 000 permutations à chaque locus montrant une différence de LOD > 2 entre les sexes. Le test de liaison multipoint a révélé une interaction génotype X sexe à 31 loci. L'analyse bipoint a confirmé 24 de ceux-ci dont 9 sont localisés à l'intérieur de gènes connus. Finalement, plusieurs mesures de PAS ont montré une interaction génotype X sexe.

Il peut aussi s'agir de la portée de l'effet du gène sur la pathologie. Il est connu que les maladies polygéniques comme l'hypertension rassemblent plusieurs gènes ayant un effet modeste. Alors, tout dépendant de leur effet, ces gènes seraient détectables dans différents groupes. Le présent projet ne pourrait pas cependant expliquer véritablement les raisons à l'origine des différentes associations observées entre les hommes et les femmes.

8.5 Limites du projet de recherche concernant l'étude de la population du Saguenay-Lac-Saint-Jean

La population canadienne-française du SLSJ représente une cible fréquente et bénéfique pour les études de maladies génétiques. En effet, la disponibilité de documents généalogiques datant de l'époque des fondateurs français au 17^e siècle (projet BALSAC⁷³), sa situation d'isolement géographique, l'effet fondateur à son origine ainsi qu'une forte natalité⁷¹, lui confèrent un pool génétique relativement homogène et stable rendant très avantageuse son étude. De plus, les particularités de cette population simplifient les études de maladies multifactorielles, comme l'hypertension, offrant la possibilité de mieux disséquer les composantes génétiques. À l'évidence, un nombre moindre de gènes devrait contribuer à la pathologie et ceux-ci devraient donc avoir un plus grand impact/importance dans le développement du phénotype, augmentant la puissance de les détecter. De plus, puisque cette population est exempte de sous-groupes ayant été génétiquement séparés, elle risque d'être peu affectée par la stratification^{183, 184}, c'est-à-dire l'observation d'un changement dans la fréquence des allèles due à des différences dans l'ascendance plutôt qu'à des associations de gènes avec la maladie¹.

L'étude de la population canadienne-française comporte aussi certaines limites. Par exemple, certaines maladies héréditaires y sont extrêmement rares, de sorte que cette population ne peut pas servir d'échantillon à toutes les études génétiques. Puis, sans aucun doute, le plus grand désavantage réside dans le fait que cette population n'est

malheureusement pas représentative des autres populations. À vrai dire, il est connu qu'au sein de distinctes populations et d'ethnies, les SNPs, la fréquence allélique, le LD et la longueur des blocs d'haplotypes diffèrent beaucoup aussi bien que les désordres complexes impliqués dans les maladies cardiovasculaires qui varient selon la distribution géographique^{87, 185-187}. Ces observations découlent de l'histoire de la population (en terme d'âge, du nombre de fondateurs, d'évolution et de migration), de l'humain aussi bien que des événements de recombinaisons¹⁸⁵⁻¹⁸⁷. C'est pour ces raisons que nous suggérons en conclusion qu'une validation de nos résultats devra être confirmée dans une autre cohorte.

Conclusion

Des gènes d'intérêt reliés à la pression artérielle et spécifiques au sexe ont été découverts sur le chromosome X à l'aide de tests d'associations effectués sur des familles en provenance du Saguenay-Lac-Saint-Jean. Pour certains SNPs, une relation significative existe entre le phénotype et les allèles du SNPs et pour d'autres, le statut d'hypertension et d'obésité semblent influencer l'association identifiée. Les données obtenues lors de ce projet confirment l'implication déjà rapportée du chromosome X dans l'hypertension. Les résultats sur ce chromosome pourraient aider à trouver des déterminants génétiques, spécifiques au sexe, contribuant à la pathologie de l'hypertension de même que valider des résultats rapportés dans la littérature antérieurement à l'exemple de l'association entre ACE2 et le LVMI.

La prochaine étape consiste à effectuer de nouveau ces analyses génétiques en regroupant les 1204 SNPs (*geneChips* d'Affymetrix) en blocs d'haplotypes puisque ceux-ci sont plus informatifs que les SNPs individuels. Ensuite, il faudrait appliquer la correction pour test multiple à l'aide des permutations. Puis, par la suite, faire la validation fonctionnelle de ces gènes pour établir les voies de signalisations présentes entre ceux-ci et la régulation de la pression artérielle dans le but de comprendre leur implication dans la pathologie de l'hypertension artérielle. Cette validation devra aussi être confirmée dans une autre cohorte. Les résultats de ce travail permettront d'établir l'importance des déterminants génétiques spécifiques au sexe impliqués dans la pathogenèse de l'hypertension et son traitement.

Références

- (1) Nussbaum R L, McInnes R R, Willard H F. *Thompson & Thompson Genetics in Medicine*. 6 ed. United States of America: 2004.
- (2) Rimoin D L, Connor J M, Pyeritz R E, Korf B R. Nature and frequency of genetic disease. In: Rimoin D L, Connor J M, Pyeritz R E, Korf B R, eds. *Emery and Rimoin's Principles and Practice of Medical Genetics*. 4 ed. London: Churchill LivingStone; 2002. p. 55-9.
- (3) Hamet P. Genetic determinants of the dynamics and kinetics of alcohol as an environmental modifier of blood pressure. *J Hypertens* 2003 June 1;21(6):1077-8.
- (4) Hamet P. Environmentally-regulated genes of hypertension. *Clin Exp Hypertens* 1996 April 1;18(3-4):267-78.
- (5) Seda O, Tremblay J, Sedova L, Hamet P. Integrating genomics and transcriptomics with geo-ethnicity and the environment for the resolution of complex cardiovascular diseases. *Curr Opin Mol Ther* 2005 December 1;7(6):583-7.
- (6) Marieb E N. *Anatomie et physiologie humaines*. 4 ed. Paris: De Boeck université; 1999.
- (7) Moncada S, Palmer RM, Higgs EA. Nitric oxide: physiology, pathophysiology, and pharmacology. *Pharmacol Rev* 1991 June 1;43(2):109-42.
- (8) Garg UC, Hassid A. Nitric oxide-generating vasodilators and 8-bromo-cyclic guanosine monophosphate inhibit mitogenesis and proliferation of cultured rat vascular smooth muscle cells. *J Clin Invest* 1989;83(5):1774-7.
- (9) McKee M, Scavone C, Nathanson JA. Nitric Oxide, cGMP, and Hormone Regulation of Active Sodium Transport. *PNAS* 1994 December 6;91(25):12056-60.
- (10) Pfeifer A, Klatt P, Massberg S, Ny L, Sausbier M, Hirneiss C, Wang GX, Korh M, Aszodi A, Andersson KE, Krombach F, Mayerhofer A, Ruth P, Fassler R, Hofmann F. Defective smooth muscle regulation in cGMP kinase I-deficient mice. *EMBO J* 1998 June 1;17(11):3045-51.
- (11) Hofmann F, Ammendola A, Schlossmann J. Rising behind NO: cGMP-dependent protein kinases. *J Cell Sci* 2000 May 15;113(10):1671-6.
- (12) Sander M, Chavoshan B, Victor RG. A Large Blood Pressure-Raising Effect of Nitric Oxide Synthase Inhibition in Humans. *Hypertension* 1999 April 1;33(4):937-42.

- (13) Yagil Y, Yagil C. Hypothesis: ACE2 Modulates Blood Pressure in the Mammalian Organism. *Hypertension* 2003 April 1;41(4):871-3.
- (14) Sachetelli S, Liu Q, Zhang SL, Liu F, Hsieh TJ, Brezniceanu ML, Guo DF, Filep JG, Ingelfinger JR, Sigmund CD, Hamet P, Chan JS. RAS blockade decreases blood pressure and proteinuria in transgenic mice overexpressing rat angiotensinogen gene in the kidney. *Kidney Int* 2006 March 1;69(6):1016-23.
- (15) Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, Donovan M, Woolf B, Robison K, Jeyaseelan R, Breitbart RE, Acton S. A Novel Angiotensin-Converting Enzyme-Related Carboxypeptidase (ACE2) Converts Angiotensin I to Angiotensin 1-9. *Circ Res* 2000 September 1;87(5):1e-9.
- (16) Vickers C, Hales P, Kaushik V, Dick L, Gavin J, Tang J, Godbout K, Parsons T, Baronas E, Hsieh F, Acton S, Patane M, Nichols A, Tummino P. Hydrolysis of Biological Peptides by Human Angiotensin-converting Enzyme-related Carboxypeptidase. *J Biol Chem* 2002 April 19;277(17):14838-43.
- (17) Tipnis SR, Hooper NM, Hyde R, Karran E, Christie G, Turner AJ. A Human Homologue of Angiotensin-converting Enzyme. Cloning and functional expression as a captopril-sensitive carboxypeptidase. *J Biol Chem* 2000;275:33238.
- (18) Ferrario CM, Trask AJ, Jessup JA. Advances in biochemical and functional roles of angiotensin-converting enzyme 2 and angiotensin-(1-7) in regulation of cardiovascular function. *Am J Physiol Heart Circ Physiol* 2005 December 1;289(6):H2281-H2290.
- (19) Zisman LS. ACE and ACE2: a tale of two enzymes. *Eur Heart J* 2005 February 2;26(4):322-4.
- (20) Burrell LM, Johnston CI, Tikellis C, Cooper ME. ACE2, a new regulator of the renin-angiotensin system. *Trends Endocrinol Metab* 2004 May 1;15(4):166-9.
- (21) Crackower MA, Sarao R, Oudit GY, Yagil C, Kozieradzki I, Scanga SE, Oliveira dS, da Costa J, Zhang L, Pei Y, Scholey J, Ferrario CM, Manoukian AS, Chappell MC, Backx PH, Yagil Y, Penninger JM. Angiotensin-converting enzyme 2 is an essential regulator of heart function. *Nature* 2002 June 20;417(6891):822-8.
- (22) Yi L, Gu YH, Wang XL, An LZ, Xie XD, Shao W, Ma LY, Fang JR, An YD, Wang F, Zhang DL. Association of ACE, ACE2 and UTS2 polymorphisms with essential hypertension in Han and Dongxiang populations from north-western China. *J Int Med Res* 2006 May 1;34(3):272-83.
- (23) Fan X, Wang Y, Sun K, Zhang W, Yang X, Wang S, Zhen Y, Wang J, Li W, Han Y, Liu T, Wang X, Chen J, Wu H, Hui R. Polymorphisms of ACE2 gene are associated with essential hypertension and antihypertensive effects of Captopril in women. *Clin Pharmacol Ther* 2007 August 1;82(2):187-96.

- (24) Katovich MJ, Grobe JL, Huentelman M, Raizada MK. Angiotensin-converting enzyme 2 as a novel target for gene therapy for hypertension. *Exp Physiol* 2005 May 1;90(3):299-305.
- (25) Joffres MR, Hamet P, MacLean DR, L'italien GJ, Fodor G. Distribution of blood pressure and hypertension in Canada and the United States. *Am J Hypertens* 2001 November 1;14(11 Pt 1):1099-105.
- (26) Burt VL, Whelton P, Roccella EJ, Brown C, Cutler JA, Higgins M, Horan MJ, Labarthe D. Prevalence of Hypertension in the US Adult Population : Results From the Third National Health and Nutrition Examination Survey, 1988-1991. *Hypertension* 1995 March 1;25(3):305-13.
- (27) Hall WD. Resistant hypertension, secondary hypertension, and hypertensive crises. *Cardiol Clin* 2002 May 1;20(2):281-9.
- (28) Akpunonu BE, Mulrow PJ, Hoffman EA. Secondary hypertension: evaluation and treatment. *Dis Mon* 1996 October 1;42(10):609-722.
- (29) O'Rourke JE, Richardson WS. Evidence based management of hypertension: What to do when blood pressure is difficult to control. *BMJ* 2001 May 19;322(7296):1229-32.
- (30) Januszewicz W, Chodakowska J, Styczynski G. Secondary hypertension in the elderly. *J Hum Hypertens* 1998 September 1;12(9):603-6.
- (31) Onusko E. Diagnosing secondary hypertension. *Am Fam Physician* 2003 January 1;67(1):67-74.
- (32) Biron P, Mongeau JG, Bertrand D. Familial aggregation of blood pressure in 558 adopted children. *Can Med Assoc J* 1976 October 23;115(8):773-4.
- (33) Kearney PM, Whelton M, Reynolds K, Muntner P, Whelton PK, He J. Global burden of hypertension: analysis of worldwide data. *Lancet* 2005 January 15;365(9455):217-23.
- (34) Hamet P, Pausova Z, Adarichev V, Adaricheva K, Tremblay J. Hypertension: genes and environment. *J Hypertens* 1998 April 1;16(4):397-418.
- (35) Xu X, Yang J, Rogus J, Chen C, Schork N. Mapping of a blood pressure quantitative trait locus to chromosome 15q in a Chinese population. *Hum Mol Genet* 1999 December 1;8(13):2551-5.
- (36) Kanellis J, Nakagawa T, Herrera-Acosta J, Schreiner GF, Rodriguez-Iturbe B, Johnson RJ. A single pathway for the development of essential hypertension. *Cardiol Rev* 2003 July 1;11(4):180-96.
- (37) Cowley AW. The genetic dissection of essential hypertension. *Nat Rev Genet* 2006 November 1;7(11):829-40.

- (38) Hamet P, Merlo E, Seda O, Broeckel U, Tremblay J, Kaldunski M, Gaudet D, Bouchard G, Deslauriers B, Gagnon F, Antoniol G, Pausova Z, Labuda M, Jomphe M, Gossard F, Tremblay G, Kirova R, Tonellato P, Orlov SN, Pintos J, Platko J, Hudson TJ, Rioux JD, Kotchen TA, Cowley AW. Quantitative founder-effect analysis of French Canadian families identifies specific loci contributing to metabolic phenotypes of hypertension. *Am J Hum Genet* 2005 May 1;76(5):815-32.
- (39) Reaven GM, Lithell H, Landsberg L. Hypertension and Associated Metabolic Abnormalities -- The Role of Insulin Resistance and the Sympathoadrenal System. *N Engl J Med* 1996 February 8;334(6):374-82.
- (40) Oparil S, Zaman MA, Calhoun DA. Pathogenesis of Hypertension. *Annals of Internal Medicine* 2003 November 4;139(9):761-76.
- (41) Ezzati M, Lopez AD, Rodgers A, Vander Hoorn S, Murray CJ. Selected major risk factors and global and regional burden of disease. *Lancet* 2002 November 2;360(9343):1347-60.
- (42) Halperin RO, Sesso HD, Ma J, Buring JE, Stampfer MJ, Michael Gaziano J. Dyslipidemia and the Risk of Incident Hypertension in Men. *Hypertension* 2006 January 1;47(1):45-50.
- (43) Sesso HD, Buring JE, Chown MJ, Ridker PM, Gaziano JM. A Prospective Study of Plasma Lipid Levels and Hypertension in Women. *Arch Intern Med* 2005 November 14;165(20):2420-7.
- (44) Haffner SM, Miettinen H, Gaskill SP, Stern MP. Metabolic precursors of hypertension. The San Antonio Heart Study. *Arch Intern Med* 1996 September 23;156(17):1994-2001.
- (45) Seda O, Liska F, Krenova D, Kazdova L, Sedova L, Zima T, Peng J, Pelinkova K, Tremblay J, Hamet P, Kren V. Dynamic genetic architecture of metabolic syndrome attributes in the rat. *Physiol Genomics* 2005 April 14;21(2):243-52.
- (46) Hedley AA, Ogden CL, Johnson CL, Carroll MD, Curtin LR, Flegal KM. Prevalence of Overweight and Obesity Among US Children, Adolescents, and Adults, 1999-2002. *JAMA* 2004 June 16;291(23):2847-50.
- (47) Christakis NA, Fowler JH. The Spread of Obesity in a Large Social Network over 32 Years. *N Engl J Med* 2007 July 26;357(4):370-9.
- (48) Hall JE, Brands MW, Henegar JR. Mechanisms of Hypertension and Kidney Disease in Obesity. *Ann N Y Acad Sci* 1999 November 18;892(1):91-107.
- (49) Pausova Z, Deslauriers B, Gaudet D, Tremblay J, Kotchen TA, Larochelle P, Cowley AW, Hamet P. Role of Tumor Necrosis Factor- α Gene Locus in Obesity and Obesity-Associated Hypertension in French Canadians. *Hypertension* 2000 July 1;36(1):14-9.

- (50) Pausova Z, Jomphe M, Houde L, Vezina H, Orlov SN, Gossard F, Gaudet D, Tremblay J, Kotchen TA, Cowley AW, Bouchard G, Hamet P. A Genealogical Study of Essential Hypertension with and without Obesity in French Canadians. *Obes Res* 2002 June 1;10(6):463-70.
- (51) Khunti K, Davies M. Metabolic syndrome. *BMJ* 2005 November 19;331(7526):1153-4.
- (52) Safar ME, Thomas F, Blacher J, Nzietchueng R, Bureau JM, Pannier B, Benetos A. Metabolic Syndrome and Age-Related Progression of Aortic Stiffness. *J Am Coll Cardiol* 2006 January 3;47(1):72-5.
- (53) Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, Platzer M, Howell GR, Burrows C, Bird CP, Frankish A, Lovell FL, Howe KL, Ashurst JL, Fulton RS, Sudbrak R, Wen G, Jones MC, Hurles ME, Andrews TD, Scott CE, Searle S, Ramsay J, Whittaker A, Deadman R, Carter NP, Hunt SE, Chen R, Cree A, Gunaratne P, Havlak P, Hodgson A, Metzker ML, Richards S, Scott G, Steffen D, Sodergren E, Wheeler DA, Worley KC, Ainscough R, Ambrose KD, Ansari-Lari MA, Aradhya S, Ashwell RI, Babbage AK, Bagguley CL, Ballabio A, Banerjee R, Barker GE, Barlow KF, Barrett IP, Bates KN, Beare DM, Beasley H, Beasley O, Beck A, Bethel G, Blechschmidt K, Brady N, Bray-Allen S, Bridgeman AM, Brown AJ, Brown MJ, Bonnin D, Bruford EA, Buhay C, Burch P, Burford D, Burgess J, Burrill W, Burton J, Bye JM, Carder C, Carrel L, Chako J, Chapman JC, Chavez D, Chen E, Chen G, Chen Y, Chen Z, Chinault C, Ciccodicola A, Clark SY, Clarke G, Clee CM, Clegg S, Clerc-Blankenburg K, Clifford K, Cobley V, Cole CG, Conquer JS, Corby N, Connor RE, David R, Davies J, Davis C, Davis J, Delgado O, Deshazo D, Dhama P, Ding Y, Dinh H, Dodsworth S, Draper H, Dugan-Rocha S, Dunham A, Dunn M, Durbin KJ, Dutta I, Eades T, Ellwood M, Emery-Cohen A, Errington H, Evans KL, Faulkner L, Francis F, Frankland J, Fraser AE, Galgoczy P, Gilbert J, Gill R, Glockner G, Gregory SG, Gribble S, Griffiths C, Grocock R, Gu Y, Gwilliam R, Hamilton C, Hart EA, Hawes A, Heath PD, Heitmann K, Hennig S, Hernandez J, Hinzmann B, Ho S, Hoffs M, Howden PJ, Huckle EJ, Hume J, Hunt PJ, Hunt AR, Isherwood J, Jacob L, Johnson D, Jones S, de Jong PJ, Joseph SS, Keenan S, Kelly S, Kershaw JK, Khan Z, Kioschis P, Klages S, Knights AJ, Kosiura A, Kovar-Smith C, Laird GK, Langford C, Lawlor S, Leversha M, Lewis L, Liu W, Lloyd C, Lloyd DM, Loulseged H, Loveland JE, Lovell JD, Lozado R, Lu J, Lyne R, Ma J, Maheshwari M, Matthews LH, McDowall J, McLaren S, McMurray A, Meidl P, Meitinger T, Milne S, Miner G, Mistry SL, Morgan M, Morris S, Muller I, Mullikin JC, Nguyen N, Nordsiek G, Nyakatura G, O'Dell CN, Okwuonu G, Palmer S, Pandian R, Parker D, Parrish J, Pastemak S, Patel D, Pearce AV, Pearson DM, Pelan SE, Perez L, Porter KM, Ramsey Y, Reichwald K, Rhodes S, Ridler KA, Schlessinger D, Schueler MG, Sehra HK, Shaw-Smith C, Shen H, Sheridan EM, Shownkeen R, Skuce CD, Smith ML, Sotheran EC, Steingruber HE, Steward CA, Storey R, Swann RM, Swarbreck D, Tabor PE, Taudien S, Taylor T, Teague B, Thomas K, Thorpe A, Timms K, Tracey A, Trevanion S, Tromans AC, d'Urso M, Verduzco D, Villasana D, Waldron L, Wall M, Wang Q, Warren J, Warry GL, Wei X, West A, Whitehead SL, Whiteley MN, Wilkinson JE, Willey DL, Williams G, Williams L, Williamson A, Williamson H, Wilming

- L, Woodmansey RL, Wray PW, Yen J. The DNA sequence of the human X chromosome. *Nature* 2005 March 17;434(7031):325-37.
- (54) Broman KW, Sen S, Owens SE, Manichaikul A, Southard-Smith EM, Churchill GA. The X Chromosome in Quantitative Trait Locus Mapping. *Genetics* 2006 December 1;174(4):2151-8.
- (55) Plath K, Mlynarczyk-Evans S, Nusinow DA, Panning B. Xist RNA and the mechanism of X chromosome inactivation. *Annu Rev Genet* 2002 January 1;36:233-78.
- (56) Williams BR, Wu Ct. Does Random X-Inactivation in Mammals Reflect a Random Choice Between Two X Chromosomes? *Genetics* 2004 July 1;167(3):1525-8.
- (57) Rastan S. Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. *J Embryol Exp Morphol* 1983 December 1;78:1-22.
- (58) Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N. Requirement for Xist in X chromosome inactivation. *Nature* 1996 January 11;379(6561):131-7.
- (59) Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, Willard HF. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 1991 January 3;349(6304):38-44.
- (60) Chadwick BP, Willard HF. Barring gene expression after XIST: maintaining facultative heterochromatin on the inactive X. *Semin Cell Dev Biol* 2003 December 1;14(6):359-67.
- (61) Chadwick BP, Willard HF. Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. *PNAS* 2004 December 14;101(50):17450-5.
- (62) Rougeulle C, Chaumeil J, Sarma K, Allis CD, Reinberg D, Avner P, Heard E. Differential Histone H3 Lys-9 and Lys-27 Methylation Profiles on the X Chromosome. *Mol Cell Biol* 2004 June 15;24(12):5475-84.
- (63) Silva J, Mak W, Zvetkova I, Appanah R, Nesterova TB, Webster Z, Peters AH, Jenuwein T, Otte AP, Brockdorff N. Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev Cell* 2003 April 1;4(4):481-95.
- (64) Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* 2005 March 17;434(7031):400-4.
- (65) Spatz A, Borg C, Feunteun J. X-chromosome genetics and human cancer. *Nat Rev Cancer* 2004 August 1;4(8):617-29.

- (66) Hilbert P, Lindpaintner K, Beckmann JS, Serikawa T, Soubrier F, Dubay C, Cartwright P, De Gouyon B, Julier C, Takahasi S. Chromosomal mapping of two genetic loci associated with blood-pressure regulation in hereditary hypertensive rats. *Nature* 1991 October 10;353(6344):521-9.
- (67) Elsheikh M, Dunger DB, Conway GS, Wass JAH. Turner's Syndrome in Adulthood. *Endocr Rev* 2002 February 1;23(1):120-40.
- (68) Gravholt CH, Juul S, Naeraa RW, Hansen J. Morbidity in Turner syndrome. *J Clin Epidemiol* 1998 February 1;51(2):147-58.
- (69) Ratcliffe S. Long term outcome in children of sex chromosome abnormalities. *Arch Dis Child* 1999 February 1;80(2):192-5.
- (70) Okada H, Fujioka H, Tatsumi N, Kanzaki M, Okuda Y, Fujisawa M, Hazama M, Matsumoto O, Gohji K, Arakawa S, Kamidono S. Klinefelter's syndrome in the male infertility clinic. *Hum Reprod* 1999 April 1;14(4):946-52.
- (71) Scriver CR. Human genetics: lessons from Quebec populations. *Annu Rev Genomics Hum Genet* 2001 January 1;2:69-101.
- (72) De Braekeleer M, Lamarre V, Scriver CR, Larochelle J, Bouchard G. Fertility in couples heterozygous for the tyrosinemia gene in Saguenay Lac-St-Jean. *Genet Couns* 1990 January 1;1(3-4):259-64.
- (73) Bouchard G, Roy R, Casgrain B, Hubert M. [Population files and database management: the BALSAC database and the INGRES/INGRID system]. *Hist Mes* 1989 January 1;4(1-2):39-57.
- (74) Levy D, DeStefano AL, Larson MG, O'Donnell CJ, Lifton RP, Gavras H, Cupples LA, Myers RH. Evidence for a Gene Influencing Blood Pressure on Chromosome 17 : Genome Scan Linkage Results for Longitudinal Blood Pressure Phenotypes in Subjects From the Framingham Heart Study. *Hypertension* 2000 October 1;36(4):477-83.
- (75) Wacholder S, Chanock S, Garcia-Closas M, El ghormli L, Rothman N. Assessing the Probability That a Positive Report is False: An Approach for Molecular Epidemiology Studies. *J Natl Cancer Inst* 2004 March 17;96(6):434-42.
- (76) Colhoun HM, McKeigue PM, vey Smith G. Problems of reporting genetic associations with complex outcomes. *Lancet* 2003 March 8;361(9360):865-72.
- (77) Kwon JM, Goate AM. The candidate gene approach. *Alcohol Res Health* 2000 January 1;24(3):164-8.
- (78) van Ooijen JW. Accuracy of mapping quantitative trait loci in autogamous species. *Theor Appl Genet* 1992 October 3;84:803-11.

- (79) Silverman EK, Palmer LJ. Case-Control Association Studies for the Genetics of Complex Respiratory Diseases. *Am J Respir Cell Mol Biol* 2000 June 1;22(6):645-8.
- (80) Uimari P, Kontkanen O, Visscher PM, Pirskanen M, Fuentes R, Salonen JT. Genome-wide linkage disequilibrium from 100,000 SNPs in the East Finland founder population. *Twin Res Hum Genet* 2005 June 1;8(3):185-97.
- (81) Belanger H, Beaulieu P, Moreau C, Labuda D, Hudson TJ, Sinnett D. Functional promoter SNPs in cell cycle checkpoint genes. *Hum Mol Genet* 2005 September 15;14(18):2641-8.
- (82) Baessler A, Hasinoff MJ, Fischer M, Reinhard W, Sonnenberg GE, Olivier M, Erdmann J, Schunkert H, Doering A, Jacob HJ, Comuzzie AG, Kissebah AH, Kwitek AE. Genetic Linkage and Association of the Growth Hormone Secretagogue Receptor (Ghrelin Receptor) Gene in Human Obesity. *Diabetes* 2005 January 1;54(1):259-67.
- (83) Zhu X, Bouzekri N, Southam L, Cooper RS, Adeyemo A, McKenzie CA, Luke A, Chen G, Elston RC, Ward R. Linkage and association analysis of angiotensin I-converting enzyme (ACE)-gene polymorphisms with ACE concentration and blood pressure. *Am J Hum Genet* 2001 May 1;68(5):1139-48.
- (84) Prokunina L, Castillejo-Lopez C, Oberg F, Gunnarsson I, Berg L, Magnusson V, Brookes AJ, Tentler D, Kristjansdottir H, Grondal G, Bolstad AI, Svenungsson E, Lundberg I, Sturfelt G, Jonssen A, Truedsson L, Lima G, Alcocer-Varela J, Jonsson R, Gyllenstein UB, Harley JB, Alarcon-Segovia D, Steinsson K, Alarcon-Riquelme ME. A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nat Genet* 2002 December 1;32(4):666-9.
- (85) Edwards AO, Ritter R, III, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement Factor H Polymorphism and Age-Related Macular Degeneration. *Science* 2005 April 15;308(5720):421-4.
- (86) Xu J, Turner A, Little J, Bleecker ER, Meyers DA. Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? *Hum Genet* 2002 December 1;111(6):573-4.
- (87) Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer CJ, Chew A, Han JH, Duan J, Carr JL, Lee MS, Koshy B, Kumar AM, Zhang G, Newell WR, Windemuth A, Xu C, Kalbfleisch TS, Shaner SL, Arnold K, Schulz V, Drysdale CM, Nandabalan K, Judson RS, Ruano G, Vovis GF. Haplotype Variation and Linkage Disequilibrium in 313 Human Genes. *Science* 2001 July 20;293(5529):489-93.
- (88) Zhao H, Pfeiffer R, Gail MH. Haplotype analysis in population genetics and association studies. *Pharmacogenomics* 2003 March 1;4(2):171-8.

- (89) Liu PY, Zhang YY, Lu Y, Long JR, Shen H, Zhao LJ, Xu FH, Xiao P, Xiong DH, Liu YJ, Recker RR, Deng HW. A survey of haplotype variants at several disease candidate genes: the importance of rare variants for complex diseases. *J Med Genet* 2005 March 1;42(3):221-7.
- (90) Stoll M, Corneliussen B, Costello CM, Waetzig GH, Mellgard B, Koch WA, Rosenstiel P, Albrecht M, Croucher PJ, Seegert D, Nikolaus S, Hampe J, Lengauer T, Pierrou S, Foelsch UR, Mathew CG, Lagerstrom-Fermer M, Schreiber S. Genetic variation in DLG5 is associated with inflammatory bowel disease. *Nat Genet* 2004 May 1;36(5):476-80.
- (91) Bland JM, Altman DG. Statistics notes: Multiple significance tests: the Bonferroni method. *BMJ* 1995 January 21;310(6973):170.
- (92) Westfall PH, Young SS. *Resampling-based multiple testing*. Wiley, New York: 1993.
- (93) Lin DY. An efficient Monte Carlo approach to assessing statistical significance in genomic studies. *Bioinformatics* 2005 March 15;21(6):781-7.
- (94) Kotchen TA, Kotchen JM, Grim CE, George V, Kaldunski ML, Cowley AW, Hamet P, Chelius TH. Genetic Determinants of Hypertension : Identification of Candidate Phenotypes. *Hypertension* 2000 July 1;36(1):7-13.
- (95) Hamet P, Seda O, Gaudet D, Brunelle P-L, Gurau A, Merlo E, Pilote L, Kotchen T, Cowley A, Tremblay J. Systematic search for sex-specific genomic determinants of 608 cardiovascular traits in French Canadians reveals novel loci/haplotypes. Human Genome Meeting (HUGO), Montréal, Canada, p.111. 2007.
Ref Type: Abstract
- (96) Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The Structure of Haplotype Blocks in the Human Genome. *Science* 2002 June 21;296(5576):2225-9.
- (97) Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005 January 15;21(2):263-5.
- (98) Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002 January 1;30(1):97-101.
- (99) Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 2006 January 15;25(1):127-41.
- (100) Hanley JA, Negassa A, Edwardes MD, Forrester JE. Statistical Analysis of Correlated Data Using Generalized Estimating Equations: An Orientation. *Am J Epidemiol* 2003 February 15;157(4):364-75.

- (101) Burton P, Gurrin L, Sly P. Extending the simple linear regression model to account for correlated responses: an introduction to generalized estimating equations and multi-level mixed modelling. *Stat Med* 1998 June 15;17(11):1261-91.
- (102) Uzun A, Leslin CM, Abyzov A, Ilyin V. Structure SNP (StSNP): a web server for mapping and modeling nsSNPs on protein structures with linkage to metabolic pathways. *Nucleic Acids Res* 2007 July 13;35(suppl_2):W384-W392.
- (103) Bhatti P, Church DM, Rutter JL, Struewing JP, Sigurdson AJ. Candidate Single Nucleotide Polymorphism Selection using Publicly Available Tools: A Guide for Epidemiologists. *Am J Epidemiol* 2006 October 15;164(8):794-804.
- (104) Marinescu VD, Kohane IS, Riva A. MAPPER: a search engine for the computational identification of putative transcription factor binding sites in multiple genomes. *BMC Bioinformatics* 2005 January 1;6:79.
- (105) Li S, Ma L, Li H, Vang S, Hu Y, Bolund L, Wang J. Snap: an integrated SNP annotation platform. *Nucleic Acids Res* 2007 January 12;35(suppl_1):D707-D710.
- (106) Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler aD. The Human Genome Browser at UCSC. *Genome Res* 2002 June 1;12(6):996-1006.
- (107) Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M,

- Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ. Initial sequencing and analysis of the human genome. *Nature* 2001 February 15;409(6822):860-921.
- (108) Lee PH, Shatkay H. F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res* 2007 November 5;gkm904.
- (109) Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* 2005 July 1;21(13):2933-42.
- (110) Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* 1995 February 24;23(23):4878-84.
- (111) Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, Podkolodny NL, Kolchanov NA. Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucleic Acids Res* 1998 January 1;26(1):362-7.
- (112) Sandelin A, Wasserman WW, Lenhard B. ConSite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic Acids Res* 2004 July 1;32(suppl_2):W249-W252.
- (113) Mooney SD, Altman RB. MutDB: annotating human variation with functionally relevant data. *Bioinformatics* 2003 September 22;19(14):1858-60.
- (114) Ovcharenko I, Nobrega MA, Loots GG, Stubbs L. ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. *Nucleic Acids Res* 2004 July 1;32(suppl_2):W280-W286.
- (115) Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 2004 July 1;32(suppl_2):W273-W279.
- (116) Hubbard TJP, Aken BL, Beal K, Ballester B, Caccamo M, Chen Y, Clarke L, Coates G, Cunningham F, Cutts T, Down T, Dyer SC, Fitzgerald S, Fernandez-Banet J, Graf S, Haider S, Hammond M, Herrero J, Holland R, Howe K, Howe K,

- Johnson N, Kahari A, Keefe D, Kokocinski F, Kulesha E, Lawson D, Longden I, Melsopp C, Megy K, Meidl P, Ouverdin B, Parker A, Prlic A, Rice S, Rios D, Schuster M, Sealy I, Severin J, Slater G, Smedley D, Spudich G, Trevanion S, Vilella A, Vogel J, White S, Wood M, Cox T, Curwen V, Durbin R, Fernandez-Suarez XM, Flicek P, Kasprzyk A, Proctor G, Searle S, Smith J, Ureta-Vidal A, Birney E. Ensembl 2007. *Nucleic Acids Res* 2006 December 5;gkl996.
- (117) Conde L, Vaquerizas JM, Dopazo H, Arbiza L, Reumers J, Rousseau F, Schymkowitz J, Dopazo J. PupaSuite: finding functional single nucleotide polymorphisms for large-scale genotyping purposes. *Nucleic Acids Res* 2006 July 1;34(suppl_2):W621-W625.
- (118) Blanchette M, Kent WJ, Riemer C, Elnitski L, Smit AFA, Roskin KM, Baertsch R, Rosenbloom K, Clawson H, Green ED, Haussler D, Miller W. Aligning Multiple Genomic Sequences With the Threaded Blockset Aligner. *Genome Res* 2004 April 1;14(4):708-15.
- (119) Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, Weinstock GM, Wilson RK, Gibbs RA, Kent WJ, Miller W, Haussler D. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 2005 August 1;15(8):1034-50.
- (120) Gordon D, Heath SC, Liu X, Ott J. A transmission/disequilibrium test that allows for genotyping errors in the analysis of single-nucleotide polymorphism data. *Am J Hum Genet* 2001 August 1;69(2):371-80.
- (121) Horvath S, Laird NM, Knapp M. The transmission/disequilibrium test and parental-genotype reconstruction for X-chromosomal markers. *Am J Hum Genet* 2000 March 1;66(3):1161-7.
- (122) Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, Cheung VG. Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet* 2007 February 1;39(2):226-31.
- (123) Abecasis GR, Noguchi E, Heinzmann A, Traherne JA, Bhattacharyya S, Leaves NI, Anderson GG, Zhang Y, Lench NJ, Carey A, Cardon LR, Moffatt MF, Cookson WO. Extent and distribution of linkage disequilibrium in three genomic regions. *Am J Hum Genet* 2001 January 1;68(1):191-7.
- (124) Wu G, Morris SM. Arginine metabolism: nitric oxide and beyond. *Biochem J* 1998 November 15;336 (Pt 1):1-17.
- (125) Moncada S, Higgs A. The L-Arginine-Nitric Oxide Pathway. *N Engl J Med* 1993 December 30;329(27):2002-12.
- (126) Giugliano D, Marfella R, Verrazzo G, Acampora R, Coppola L, Cozzolino D, D'Onofrio F. The Vascular Effects of L-Arginine in Humans . The Role of Endogenous Insulin. *J Clin Invest* 1997 February 1;99(3):433-8.

- (127) Kitiyakara C, Chabrashvili T, Chen Y, Blau J, Karber A, Aslam S, Welch WJ, Wilcox CS. Salt Intake, Oxidative Stress, and Renal Expression of NADPH Oxidase and Superoxide Dismutase. *J Am Soc Nephrol* 2003 November 1;14(11):2775-82.
- (128) Modlinger PS, Wilcox CS, Aslam S. Nitric oxide, oxidative stress, and progression of chronic renal failure. *Semin Nephrol* 2004 July 1;24(4):354-65.
- (129) Rubanyi GM, Vanhoutte PM. Superoxide anions and hyperoxia inactivate endothelium-derived relaxing factor. *Am J Physiol Heart Circ Physiol* 1986 May 1;250(5):H822-H827.
- (130) Grunfeld S, Hamilton CA, Mesaros S, McClain SW, Dominiczak AF, Bohr DF, Malinski T. Role of Superoxide in the Depressed Nitric Oxide Production by the Endothelium of Genetically Hypertensive Rats. *Hypertension* 1995 December 1;26(6):854-7.
- (131) Cuzzocrea S, Mazzon E, Dugo L, Di Paola R, Caputi AP, Salvemini D. Superoxide: a key player in hypertension. *FASEB J* 2004;18:94.
- (132) Freedman B, Bernstein J. Current Concepts Review - Sample Size and Statistical Power in Clinical Orthopaedic Research. *J Bone Joint Surg Am* 1999;81:1454-60.
- (133) Sala C, Arrigo G, Torri G, Martinazzi F, Riva P, Larizza L, Philippe C, Jonveaux P, Sloan F, Labella T, Toniolo D. Eleven X chromosome breakpoints associated with premature ovarian failure (POF) map to a 15-Mb YAC contig spanning Xq21. *Genomics* 1997 February 15;40(1):123-31.
- (134) Braybrook C, Warry G, Howell G, Arnason A, Bjornsson A, Moore GE, Ross MT, Stanier P. Identification and characterization of KLHL4, a novel human homologue of the Drosophila Kelch gene that maps within the X-linked cleft palate and Ankyloglossia (CPX) critical region. *Genomics* 2001 March 1;72(2):128-36.
- (135) Bahi N, Friocourt G, Carrie A, Graham ME, Weiss JL, Chafey P, Fauchereau F, Burgoyne RD, Chelly J. IL1 receptor accessory protein like, a protein involved in X-linked mental retardation, interacts with Neuronal Calcium Sensor-1 and regulates exocytosis. *Hum Mol Genet* 2003 June 15;12(12):1415-25.
- (136) Carrie A, Jun L, Bienvenu T, Vinet MC, McDonell N, Couvert P, Zemni R, Cardona A, Van Buggenhout G, Frints S, Hamel B, Moraine C, Ropers HH, Strom T, Howell GR, Whittaker A, Ross MT, Kahn A, Fryns JP, Beldjord C, Marynen P, Chelly J. A new member of the IL-1 receptor family highly expressed in hippocampus and involved in X-linked mental retardation. *Nat Genet* 1999 September 1;23(1):25-31.
- (137) Born TL, Smith DE, Garka KE, Renshaw BR, Bertles JS, Sims JE. Identification and Characterization of Two Members of a Novel Class of the Interleukin-1 Receptor (IL-1R) Family. Delineation of a new class of IL-1R-related proteins based on signaling. *J Biol Chem* 2000;275:29946.

- (138) Khan JA, Brint EK, O'Neill LAJ, Tong L. Crystal Structure of the Toll/Interleukin-1 Receptor Domain of Human IL-1RAPL. *J Biol Chem* 2004 July 23;279(30):31664-70.
- (139) Khraibi AA. Association between disturbances in the immune system and hypertension. *Am J Hypertens* 1991 July 1;4(7 Pt 1):635-41.
- (140) Mills PJ, Farag NH, Hong S, Kennedy BP, Berry CC, Ziegler MG. Immune cell CD62L and CD11a expression in response to a psychological stressor in human hypertension. *Brain Behav Immun* 2003 August 1;17(4):260-7.
- (141) Mahmud A, Feely J. Arterial Stiffness Is Related to Systemic Inflammation in Essential Hypertension. *Hypertension* 2005 November 1;46(5):1118-22.
- (142) Bautista LE, Vera LM, Arenas IA, Gamarra G. Independent association between inflammatory markers (C-reactive protein, interleukin-6, and TNF-alpha) and essential hypertension. *J Hum Hypertens* 2005 February 1;19(2):149-54.
- (143) Dalekos GN, Elisaf MS, Papagalanis N, Tzallas C, Siamopoulos KC. Elevated interleukin-1 beta in the circulation of patients with essential hypertension before any drug therapy: a pilot study. *Eur J Clin Invest* 1996 October 1;26(10):936-9.
- (144) Cucca F, Goy JV, Kawaguchi Y, Esposito L, Merriman ME, Wilson AJ, Cordell HJ, Bain SC, Todd JA. A male-female bias in type 1 diabetes and linkage to chromosome Xp in MHC HLA-DR3-positive patients. *Nat Genet* 1998 July 1;19(3):301-2.
- (145) Wei BQ, Mikkelsen TS, McKinney MK, Lander ES, Cravatt BF. A Second Fatty Acid Amide Hydrolase with Variable Distribution among Placental Mammals. *J Biol Chem* 2006 December 1;281(48):36569-78.
- (146) Batkai S, Pacher P, Osei-Hyiaman D, Radaeva S, Liu J, Harvey-White J, Offertaler L, Mackie K, Rudd MA, Bukoski RD, Kunos G. Endocannabinoids Acting at Cannabinoid-1 Receptors Regulate Cardiovascular Function in Hypertension. *Circulation* 2004 October 5;110(14):1996-2002.
- (147) Fletcher BS, Dragstedt C, Notterpek L, Nolan GP. Functional cloning of SPIN-2, a nuclear anti-apoptotic protein with roles in cell cycle progression. *Leukemia* 2002 August 1;16(8):1507-18.
- (148) Yuan X, Yamada K, Ishiyama S, Koyama W, Nonaka K. Identification of polymorphic loci in the promoter region of the serotonin 5-HT_{2C} receptor gene and their association with obesity and type II diabetes. *Diabetologia* 2000 March 1;43(3):373-6.
- (149) Frishman WH, Huberfeld S, Okin S, Wang YH, Kumar A, Shareef B. Serotonin and serotonin antagonism in cardiovascular and non-cardiovascular disease. *J Clin Pharmacol* 1995 June 1;35(6):541-72.

- (150) Lavezzaro G, Ladetto PE, Valente M, Stramignoni D, Zanna C, Assogna G, Salvetti A. Ketanserin and captopril interaction in the treatment of essential hypertensives. *Cardiovasc Drugs Ther* 1990 January 1;4 Suppl 1:119-22.
- (151) Pajukanta P, Cargill M, Viitanen L, Nuotio I, Kareinen A, Perola M, Terwilliger JD, Kempas E, Daly M, Lilja H, Rioux JD, Bretin T, Viikari JS, Ronnema T, Laakso M, Lander ES, Peltonen L. Two loci on chromosomes 2 and X for premature coronary heart disease identified in early- and late-settlement populations of Finland. *Am J Hum Genet* 2000 December 1;67(6):1481-93.
- (152) Kalin MF, Zumoff B. Sex hormones and coronary disease: a review of the clinical studies. *Steroids* 1990 August 1;55(8):330-52.
- (153) Gordon T, Kannel WB, Hjortland MC, McNamara PM. Menopause and coronary heart disease: the Framingham Study. *Ann Intern Med* 89, 157-161. 1978.
Ref Type: Abstract
- (154) Lawlor DA, Ebrahim S, vey Smith G. Sex matters: secular and geographical trends in sex differences in coronary heart disease mortality. *BMJ* 2001 September 8;323(7312):541-5.
- (155) Ikeda S, Kong SW, Lu J, Bisping E, Zhang H, Allen PD, Golub TR, Pieske B, Pu WT. Altered microRNA expression in human heart disease. *Physiol Genomics* 2007 November 14;31(3):367-73.
- (156) Naraba H, Iwai N. Assessment of the microRNA system in salt-sensitive hypertension. *Hypertens Res* 2005 October 1;28(10):819-26.
- (157) Wang X, Zhang J. Rapid Evolution of Mammalian X-Linked Testis-Expressed Homeobox Genes. *Genetics* 2004 June 1;167(2):879-88.
- (158) Iwashima Y, Horio T, Kamide K, Rakugi H, Ogihara T, Kawano Y. Uric Acid, Left Ventricular Mass Index, and Risk of Cardiovascular Disease in Essential Hypertension. *Hypertension* 2006 February 1;47(2):195-202.
- (159) Gaudio C, Ferri FM, Giovannini M, Pannarale G, Puddu PE, Vittore A, Fera MS, Vizza CD, Fedele F. Comparative effects of irbesartan versus amlodipine on left ventricular mass index in hypertensive patients with left ventricular hypertrophy. *J Cardiovasc Pharmacol* 2003 November 1;42(5):622-8.
- (160) Lieb W, Graf J, Gotz A, Konig IR, Mayer B, Fischer M, Stritzke J, Hengstenberg C, Holmer SR, Doring A, Lowel H, Schunkert H, Erdmann J. Association of angiotensin-converting enzyme 2 (ACE2) gene polymorphisms with parameters of left ventricular hypertrophy in men. Results of the MONICA Augsburg echocardiographic substudy. *J Mol Med* 2006 January 1;84(1):88-96.
- (161) De Mota N, Reaux-Le Goazigo A, El Messari S, Chartrel N, Roesch D, Dujardin C, Kordon C, Vaudry H, Moos F, Llorens-Cortes C. Apelin, a potent diuretic neuropeptide counteracting vasopressin actions through inhibition of vasopressin neuron activity and vasopressin release. *PNAS* 2004 July 13;101(28):10464-9.

- (162) Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL. The human disease network. *PNAS* 2007 May 22;104(21):8685-90.
- (163) Barabasi AL. Network Medicine -- From Obesity to the "Diseasome". *N Engl J Med* 2007 July 26;357(4):404-7.
- (164) Aklillu E, Carrillo JA, Makonnen E, Hellman K, Pitarque M, Bertilsson L, Ingelman-Sundberg M. Genetic Polymorphism of CYP1A2 in Ethiopians Affecting Induction and Expression: Characterization of Novel Haplotypes with Single-Nucleotide Polymorphisms in Intron 1. *Mol Pharmacol* 2003 September 1;64(3):659-69.
- (165) Menzaghi C, Paroni G, De Bonis C, Soccio T, Marucci A, Bacci S, Trischitta V. The -318 C>G Single-Nucleotide Polymorphism in GNAI2 Gene Promoter Region Impairs Transcriptional Activity through Specific Binding of Sp1 Transcription Factor and Is Associated with High Blood Pressure in Caucasians from Italy. *J Am Soc Nephrol* 2006 April 1;17(4_suppl_2):S115-S119.
- (166) Jacobson EM, Concepcion E, Oashi T, Tomer Y. A Graves' Disease-Associated Kozak Sequence Single-Nucleotide Polymorphism Enhances the Efficiency of CD40 Gene Translation: A Case for Translational Pathophysiology. *Endocrinology* 2005 June 1;146(6):2684-91.
- (167) Yamaguchi S, Shinmura K, Saitoh T, Takenoshita S, Kuwano H, Yokota J. A single nucleotide polymorphism at the splice donor site of the human MYH base excision repair genes results in reduced translation efficiency of its transcripts. *Genes Cells* 2002 May 1;7(5):461-74.
- (168) Clump DA, II, Flynn DC. A SNP in AFAP is associated with ovarian cancer. *AACR Meeting Abstracts* 2006 April 1;2006(1):776.
- (169) Li C, Liu Z, Zhang Z, Strom SS, Gershenwald JE, Prieto VG, Lee JE, Ross MI, Mansfield PF, Cormier JN, Duvic M, Grimm EA, Wei Q. Genetic variants of the vitamin D receptor gene alter risk of cutaneous melanoma. *J Invest Dermatol* 2007 February 1;127(2):276-80.
- (170) Chen Q, Reis SE, Kammerer C, Craig WY, LaPierre SE, Zimmer EL, McNamara DM, Pauly DF, Sharaf B, Holubkov R, Bairey Merz CN, Sopko G, Bontempo F, Kamboh MI. Genetic Variation in Lectin-Like Oxidized Low-Density Lipoprotein Receptor 1 (LOX1) Gene and the Risk of Coronary Artery Disease. *Circulation* 2003 July 1;107(25):3146-51.
- (171) Huang JL, Gao PS, Mathias RA, Yao TC, Chen LC, Kuo ML, Hsu SC, Plunkett B, Trogias A, Barnes KC, Stellato C, Beaty TH, Huang SK. Sequence variants of the gene encoding chemoattractant receptor expressed on Th2 cells (CRTH2) are associated with asthma and differentially influence mRNA stability. *Hum Mol Genet* 2004 November 1;13(21):2691-7.

- (172) Wray GA, Hahn MW, Abouheif E, Ballhoff JP, Pizer M, Rockman MV, Romano LA. The Evolution of Transcriptional Regulation in Eukaryotes. *Mol Biol Evol* 2003 September 1;20(9):1377-419.
- (173) Frith MC, Fu Y, Yu L, Chen JF, Hansen U, Weng Z. Detection of functional DNA motifs via statistical over-representation. *Nucleic Acids Res* 2004 February 26;32(4):1372-81.
- (174) Gosso MF, de Geus EJC, Polderman TJC, Boomsma DI, Heutink P, Posthuma D. Common variants underlying cognitive ability: further evidence for association between the SNAP-25 gene and cognition using a family-based study in two independent Dutch cohorts. *Genes, Brain and Behavior* , 1-10. 2007.
Ref Type: Abstract
- (175) Blaschke RJ, Rappold G. The pseudoautosomal regions, SHOX and disease. *Curr Opin Genet Dev* 2006 June 1;16(3):233-9.
- (176) Poitras L, Ghanem N, Hatch G, Ekker M. The proneural determinant MASH1 regulates forebrain Dlx1/2 expression through the I12b intergenic enhancer. *Development* 2007 May 1;134(9):1755-65.
- (177) Kirschner R, Erturk D, Zeitz C, Sahin S, Ramser J, Cremers FP, Ropers HH, Berger W. DNA sequence comparison of human and mouse retinitis pigmentosa GTPase regulator (RPGR) identifies tissue-specific exons and putative regulatory elements. *Hum Genet* 2001 September 1;109(3):271-8.
- (178) Hayward CS, Webb CM, Collins P. Effect of sex hormones on cardiac mass. *Lancet* 2001 April 28;357(9265):1354-6.
- (179) Staessen JA, van der H, Safar ME, Den Hond E, Gasowski J, Fagard RH, Wang JG, Boudier HA, Van Bortel LM. Menopause and the characteristics of the large arteries in a population study. *J Hum Hypertens* 2001 August 1;15(8):511-8.
- (180) Safar ME, Smulyan H. Hypertension in women. *Am J Hypertens* 2004 January 1;17(1):82-7.
- (181) Smulyan H, Asmar RG, Rudnicki A, London GM, Safar ME. Comparative effects of aging in men and women on the properties of the arterial tree. *J Am Coll Cardiol* 2001 April 1;37(5):1374-80.
- (182) Marcus R, Krause L, Weder AB, Dominguez-Meja A, Schork NJ, Julius S. Sex-specific determinants of increased left ventricular mass in the Tecumseh Blood Pressure Study. *Circulation* 1994 August 1;90(2):928-36.
- (183) Lopez de Lara C, Dumais A, Rouleau G, Lesage A, Dumont M, Chawky N, Alda M, Benkelfat C, Turecki G. STin2 variant and family history of suicide as significant predictors of suicide completion in major depression. *Biol Psychiatry* 2006 January 15;59(2):114-20.

- (184) Stan S, Lambert M, Delvin E, Paradis G, O'Loughlin J, Hanley JA, Levy E. Intestinal fatty acid binding protein and microsomal triglyceride transfer protein polymorphisms in French-Canadian youth. *J Lipid Res* 2005 February 1;46(2):320-7.
- (185) Ng MC, Wang Y, So WY, Cheng S, Visvikis S, Zee RY, Fernandez-Cruz A, Lindpaintner K, Chan JC. Ethnic differences in the linkage disequilibrium and distribution of single-nucleotide polymorphisms in 35 candidate genes for cardiovascular diseases. *Genomics* 2004 April 1;83(4):559-65.
- (186) Zhu X, Yan D, Cooper RS, Luke A, Ikeda MA, Chang YP, Weder A, Chakravarti A. Linkage Disequilibrium and Haplotype Diversity in the Genes of the Renin-Angiotensin System: Findings From the Family Blood Pressure Program. *Genome Res* 2003 February 1;13(2):173-81.
- (187) Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R, Lander ES. Linkage disequilibrium in the human genome. *Nature* 2001 May 10;411(6834):199-204.