



HAL
open science

Application des stratégies combinées utilisant le séquençage d'exome dans les maladies vasculaires rares

Juliette Albuissou

► **To cite this version:**

Juliette Albuissou. Application des stratégies combinées utilisant le séquençage d'exome dans les maladies vasculaires rares. Médecine humaine et pathologie. Université Sorbonne Paris Cité, 2015. Français. NNT : 2015USPCB119 . tel-01541588

HAL Id: tel-01541588

<https://theses.hal.science/tel-01541588>

Submitted on 19 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THESE DE DOCTORAT DE L'UNIVERSITE PARIS DESCARTES

Pour l'obtention du grade de Docteur de l'Université Paris Descartes

ECOLE DOCTORALE DEVELOPPEMENT, GENETIQUE, NEUROLOGIE ET VIEILLISSEMENT

DISCIPLINE GENETIQUE

Présentée et soutenue publiquement par

Juliette ALBUISSON

APPLICATION DES STRATEGIES COMBINEES UTILISANT LE SEQUENÇAGE D'EXOME DANS LES MALADIES VASCULAIRES RARES

Le 23 Novembre 2015

Devant le jury composé de:

Pr Stanislas Lyonnet	Président
Pr Laurence Olivier-Faivre	Rapporteur
Dr Jean-Jacques Schott	Rapporteur
Dr Sophie Dupuis-Girod	Examineur
Pr Xavier Jeunemaitre	Directeur de thèse
Pr Elisabeth Tournier-Lasserre	Membre invité
Pr Serge Amselem	Membre invité

REMERCIEMENTS

Je tiens à remercier tout d'abord les membres de mon jury, pour qui j'ai beaucoup d'estime et devant qui je suis fière de présenter mon travail de thèse.

Je remercie mon directeur de Thèse et chef de Service, Xavier Jeunemaître, qui m'a tout d'abord fait confiance il y a maintenant quatre ans, sans m'avoir rencontrée... et auprès de qui j'ai appris que l'exigence intellectuelle est une pratique de chaque instant.

Je remercie chaleureusement les membres de mon équipe de recherche, pour leur soutien, nos discussions éclairées, et nos bons moments au soleil du parc André Citroën... Je remercie tout particulièrement Nabila Bouatia-Naji pour ses conseils avisés, et son aide dans la réalisation de la partie statistique de ce travail. Je remercie aussi Romuald Kiando qui m'a apporté son aide dans cette étape.

Merci à l'équipe de Génétique Médicale et Moléculaire de l'HEGP, médecins, techniciens, secrétaire, pour m'avoir acceptée dans leur équipe: ils sont aujourd'hui toujours présents et prêts à offrir aide et témoignages d'amitié.

Merci aux membres du Centre de Référence des Maladies Vasculaires Rares, véritable équipe au service du malade, auquel je suis fière d'appartenir. Je remercie particulièrement Pierre-François Plouin pour m'avoir aidée dans la compréhension de cette étrange maladie qu'est la DFM.

Merci à Patrick Bruneval pour sa disponibilité, et les moments passés au microscope, à m'apprendre, me faire comprendre, et détailler... tous les prélèvements artériels rares d'étranges maladies que j'ai pu récolter depuis 4 ans.

Merci à Paul (il se reconnaîtra), pour la grande aide logistique apportée les derniers mois de cette année qui m'ont permis de sauver de précieuses heures de sommeil! Merci à Jean-Michael Mazzella et à Anne Legrand pour leur aide logistique aussi, pour la longue liste des menues tâches que vous avez bien voulu réaliser pour moi. Merci à Marie-Laure Baranne, Cyrielle Treard et Mathilde Padilla pour leur aide dans le séquençage d'APOL3 qui n'a pas été une mince affaire!

Je remercie mes amis qui savent que depuis siii longtemps je travaille, en plus de mon métier de médecin, à un truc bizarre qui consiste à se poser plein de questions, à prononcer des phrases incompréhensibles, et surtout à travailler aux heures où les "gens normaux" dorment ou s'adonnent à leurs loisirs préférés... Merci d'avoir été un soutien malgré ces bizarreries.

Et surtout, merci à ma famille, pour l'aide apportée dans mon quotidien, pour leur soutien indéfectible. Mais surtout pour cet endroit où il fait bon être quelque soit la latitude ou la longitude, du moment qu'on est ensemble... Family is power!

A mes chevaux, qui m'ont portée de loin en loin,

A ma famille et mes amis qui sont le sel de ma vie,

Et surtout... à mes deux plus belles expériences de génétique Mendélienne, qui illuminent mon quotidien!

ABSTRACT

Identifying genes of Mendelian disorders has started within the eighties. The pace of new genes discovery has been dramatically accelerated by the availability of the human genome sequence in the 2000s, and the next-generation sequencing technologies in the 2010s. However, a majority of the elucidated conditions so far correspond to relatively simplified situations, where the prevalence and the penetrance of the condition are high and the genetic heterogeneity is low. Nowadays, geneticists meet more and more situations where gene identification in unknown disorders can be tricky. Heritable conditions that are very rare, heterogenous or with imperfect Mendelian transmission can only be elucidated using large cohorts of patients, with a very well-characterized phenotype. This requires clinical, financial and logistical efforts to be made by the research teams. Generally, using exome sequencing alone is not efficient enough to elucidate these types of conditions. The power of recently developed strategies comes from its association with other genetic analysis tools, that have been specifically developed in the context of rare, heterogenous, or polygenic disorders. I employed exome sequencing in the identification of cardiovascular genetic conditions, using three different strategies. In the first condition, called hereditary xerocytosis, using linkage analysis together with exome sequencing of distant relatives was successful in identifying the causative gene. This was made possible by the identification of a reliable endophenotype, and the relative genetic homogeneity of the disorder. The second condition I studied is the abdominal aortic aneurysm (AAA), a common disorder with a strong hereditary component and rare situations of fully penetrant, dominant inheritance. I combined exome sequencing in a family with dominant inheritance with rare variants analysis of the candidate gene in a large cohort of sporadic AAA. This analysis is more complex and can be hazardous in the context of a candidate gene approach. The third strategy was developed for the study of fibromuscular dysplasia (FMD) which is a very heterogenous condition with low penetrance and no specific endophenotype. I combined exome sequencing in a group of 30 cases and relatives with filtering strategies for any type of Mendelian inheritance. I also used available bioinformatics tools and databases for refining the candidate genes filtering. This strategy provided promising results, probably due to the genetic characteristics of this condition.

In each of these examples, I adapted the analysis strategy to the peculiarities of the disorder. The results presented here enable to evaluate the efficiency of combined approaches using exome sequencing. Their specificities, limits, and the optimization that need to be done to elucidate the remaining unsolved genetic conditions are discussed.

RESUME

L'identification des gènes de maladies Mendéliennes été rendue possible dans les années 1980. Le séquençage du génome humain dans les années 2000, et l'arrivée du séquençage haut débit dans les années 2010 ont permis une progression phénoménale du rythme de ces découvertes génétiques. Cependant, ces techniques ont permis d'élucider les maladies les plus accessibles, de par leur homogénéité génétique, leur forte pénétrance pour les maladies dominantes, et leur prévalence élevée. Aujourd'hui, la communauté génétique se heurte à des difficultés de plus en plus nombreuses pour identifier les maladies non monogéniques, hétérogènes, ou très rares : recruter des familles porteuses d'une maladie très rare et cliniquement bien caractérisée, ou de grandes cohortes de patients atteints de maladies communes à composante génétique, nécessite un effort clinique, logistique et financier important. De plus, le séquençage d'exome pris isolément ne permet généralement pas l'élucidation de ces pathologies. Cet outil bien que puissant trouve ses limites dans les modèles génétiques sus-cités. La réussite des approches récentes vient de son utilisation en association avec d'autres techniques, adaptées aux caractéristiques des maladies comme la rareté, la dimension polygénique, ou l'hétérogénéité génétique. J'ai utilisé le séquençage d'exome dans l'identification de gènes de maladies cardiovasculaires rares, par trois stratégies combinées différentes. La première pathologie appelée stomatocytose héréditaire, est rare, relativement homogène mais présente des difficultés de phénotypage. Elle a été caractérisée par séquençage d'exome en association avec une analyse de liaison traditionnelle et l'identification d'une endophénotype fiable. Cette approche a été appliquée avec succès dans ce modèle relativement peu complexe de maladie. La seconde pathologie étudiée, l'anévrisme de l'aorte abdominale (AAA), est une maladie commune à forte composante héréditaire, avec de rares formes dominantes à pénétrance élevée. J'ai associé séquençage d'exome en modèle dominant et recherche de variants rares dans une cohorte de cas sporadiques afin d'identifier un gène de susceptibilité à la pathologie. Cette approche s'est révélée plus complexe à mettre en place, et son efficacité peut être discutée dans le cas d'une étude d'association centrée sur un gène candidat. Enfin, la troisième partie de ce travail est consacrée à la dysplasie fibromusculaire (DFM), maladie très hétérogène sur le plan génétique, peu pénétrante, et d'endophénotype peu accessible. J'ai appliqué dans cette troisième étape le séquençage d'exome à plus grande échelle (30 individus), en association à des stratégies de filtres sophistiqués exploitant tous les types de transmission Mendélienne. J'y ai aussi associé l'utilisation des outils bioinformatiques et bases de données biologiques accessibles à la communauté scientifique. Les résultats obtenus par cette dernière approche sont prometteurs, probablement du fait des caractéristiques inhérentes de cette pathologie.

L'utilisation de ces trois stratégies très différentes, adaptées aux contraintes de chaque pathologie, permet d'évaluer la puissance et l'efficacité des approches combinées utilisant le séquençage d'exome. Leurs difficultés inhérentes, leur inadaptation à certaines situations génétiques, ainsi que les pistes d'amélioration nécessaires pour l'élucidation des maladies génétiques de cause inconnue sont aussi abordés.

LISTE DES ABREVIATIONS

AAA	Anévrisme de l'aorte abdominale
ADN	Acide désoxyribonucléique
AIC	Anévrisme intracrânien
ATS	syndrome de tortuosité artérielle
BAO	Bicuspidie aortique
CML	Cellules musculaires lisses
CMV	Cytomégalovirus
DFM	Dysplasie fibromusculaire
GWAS	Genome-Wide Association Study
HAVS	Hand-Arm Vibration-induced Syndrome
HSV	Herpes simplex virus
IC	Intervalle de confiance
IFN- γ	interféron gamma
IHC	Immunohistochimie
KD	Knock-down
KI	Knock-in
KO	Knock-out
LDS	Syndrome de Loeys-Dietz
LZ	Leucine zipper
MAD	Membrane adressing domain
MAF	Minor allele frequency
MEC	Matrice extracellulaire
MMP	Métalloprotéinases matricielles
NF1	Neurofibromatose de type 1
NO	Oxyde nitrique
nsTAAD	Anévrisme et dissection de l'aorte ascendante non syndromique
OR	Odds-ratio
ROS	dérivés réactifs de l'oxygène
SEDv	Syndrome d'Ehlers-Danlos vasculaire
SID	SRA interacting domain
SNP	Single nucleotide polymorphism
SRA	Serum resistance agent
TAA	Anévrisme de l'aorte ascendante
TAAD	Anévrisme et dissection de l'aorte ascendante
TABC	Tronc artériel brachiocéphalique
TIMP	Inhibiteurs de métalloprotéinases matricielles
VALT	Tissu lymphoïde associé aux vaisseaux
VIH	Virus de l'immunodéficience humaine

TABLE DES MATIERES

Chapitre 1 : Les stratégies d'identification de gènes dans les maladies génétiques p14

- 1 Les maladies Mendéliennes p15
 - 1.1 Combien de gènes, combien de maladies ? p15
 - 1.2 Gènes, types de mutations et maladies Mendéliennes p15
 - 1.3 Les maladies identifiées p17
- 2 Les maladies polygéniques ou communes p19
- 3 Les stratégies d'identification des maladies génétiques Mendéliennes et communes avant le séquençage d'exome p19
 - 3.1 Maladies Mendéliennes: le clonage positionnel p19
 - 3.1.1 L'analyse de liaison simple (paramétrique) p20
 - 3.1.2 La cartographie d'homozygotie p20
 - 3.1.3 Les autres techniques de clonage positionnel p20
 - 3.1.4 L'évolution et les limites du clonage positionnel p21
 - 3.2 Maladies communes: Les études d'association GWAS p22
 - 3.2.1 Echec de la liaison génétique p22
 - 3.2.2 Les études d'association et l'âge d'or du GWAS p22
 - 3.2.3 Les limites du GWAS p25
- 4 Le séquençage d'exome p30
 - 4.1 Le séquençage haut débit p30
 - 4.2 Pourquoi séquencer un exome p32
 - 4.3 Le séquençage d'exome dans les maladies mendéliennes p34

- 4.3.1 Résultats bruts d'une analyse d'exome et utilisation de filtres basiques p34
 - 4.3.1.1 Bases de données de variants et fréquence allélique p34
 - 4.3.1.2 Prédications de pathogénicité p35
- 4.3.2 Optimisation du design de l'étude selon le modèle génétique p38
 - 4.3.2.1 Stratégie de ségrégation familiale p38
 - 4.3.2.2 Stratégie d'homozygotie p40
 - 4.3.2.3 Stratégie du double-hit p40
 - 4.3.2.4 Stratégie de croisement ou de chevauchement p41
 - 4.3.2.5 Stratégie *de novo*, ou analyse de trios p42
 - 4.3.2.6 Stratégie par gène candidat p42
- 4.3.3 Succès et limites du séquençage d'exome p43
- 4.4 Les stratégies combinées au séquençage d'exome en génétique Mendélienne p45
 - 4.4.1 Séquençage d'exome et cartographie génétique p45
 - 4.4.2 Séquençage d'exome et tests d'association p47
 - 4.4.3 Séquençage d'exome et outils de priorisation des gènes candidats p47
 - 4.4.3.1 Outils basés sur la fonction, les propriétés intrinsèques des gènes, et leur similitude avec les gènes de maladies p48
 - 4.4.3.2 Outils basés sur les similitudes phénotypiques p50
- 4.5 Stratégies combinées en maladies communes : la recherche de variants rares p52
 - 4.5.1 GWAS et séquençage d'exome dans les phénotypes extrêmes p52
 - 4.5.2 GWAS et séquençage d'exome dans les isolats géographiques p54
 - 4.5.3 Séquençage d'exome dans les formes familiales rares de maladies communes p54
 - 4.5.4 Etudes d'association pour les variants rares p55

Objectifs de travail p58

Chapitre 2 : Etude génétique de la Stomatocytose Hériditaire Déshydratée p59

- 1 Clinique, épidémiologie et physiopathologie p60
- 2 Génétique de la DHS p61
- 3 Application du séquençage d'exome dans l'identification du gène causal p65
- 4 Premières conclusions p67

Chapitre 3 : Etude génétique de l'Anévrisme de l'aorte abdominale p69

- 1 Généralités sur la paroi artérielle p70
 - 1.1 Histologie p70
 - 1.1.1 L'endothélium p70
 - 1.1.2 La media p70
 - 1.1.3 L'adventice p74
 - 1.2 Métabolisme de la paroi artérielle p77
 - 1.2.1 Les MMP p77
 - 1.1.2 Les TIMP p79
- 2 L'aorte p79
 - 2.1 Anatomie descriptive p79
 - 2.2 Pathologie aortique anévrismale et occlusive p81
 - 2.2.1 Anévrisme de l'aorte ascendante p81
 - 2.2.2 Anévrisme de l'aorte descendante p82
3. L'anévrisme de l'aorte abdominale p82
 - 3.1 Définition, diagnostic et histoire naturelle p82

- 3.2 Épidémiologie et facteurs de risque de l'AAA p84
- 3.3 Etiologie p89
- 3.4 Composante héréditaire de l'AAA p89
- 3.5 Aspects histologiques et pathogénèse de l'AAA p90
- 3.6 Spécificités du processus inflammatoire dans l'AAA p92
- 3.7 L'immunité p93
- 3.8 L'hypothèse infectieuse p93
- 4. Génétique de l' AAA p95
 - 4.1 Analyses de liaison p95
 - 4.2 Etude d'association gènes candidats p95
 - 4.3 GWAS p96
- 5 Etude génétique de l'AAA dominant et sporadique p99
 - 5.1 Présentation de la famille princeps p99
 - 5.2 Analyse d'exome p102
 - 5.3 Identification de gènes candidats, analyse de liaison aux locus correspondants p104
 - 5.3.1 PLA2G3 p106
 - 5.3.1.1 Données de la littérature p106
 - 5.3.1.2 Séquençage de PLA2G3 dans 127 cas familiaux d'AAA p107
 - 5.3.2 APOL3 p107
 - 5.3.2.1 Données de la littérature p107
 - 5.3.2.1.1 La famille des ApoL p107
 - 5.3.2.1.2 Apol1 p113
 - 5.3.2.1.1.1 Apol1 et développement de la résistance au trypanosome p113
 - 5.3.2.1.1.2 Apol1 et cellules humaines : données de la littérature p115
 - 5.3.2.1.1.3 Apol1 et cellules humaines : données issues de collaborations p118
 - 5.3.2.1.1.4 Apol2, Apol4, et Apol6 p119

5.3.2.2 Etude d'ApoI3 dans l'AAA p119

5.3.2.2.1 Expression d'ApoI3 : Immunohistochimie, données issues de collaborations et de bases de données publiques p119

5.3.2.2.2 ApoI3 et cellules dendritiques p122

5.3.2.2.3 Modèle explicatif du rôle d'ApoI3 dans l'AAA p124

5.4 Etude d'association au locus ApoI3 et recherche de variants rares p126

5.4.1 Cohortes constitutives p126

5.4.2 Résultats du séquençage Sanger p126

5.4.3 Analyse d'association des variants rares p128

5.4.4 Tests d'association des variants très rares p132

5.4.5 Etudes fonctionnelles p132

6 Premières conclusions p133

Chapitre 4 : Etude génétique de la Dysplasie Fibromusculaire p135

1. Définition p136

2. Histoire de la DFM p136

3. Epidémiologie et facteurs de risque p138

4. Histologie et classification histo-angiographique p140

4.1 DFM intimale p142

4.2 DFM mediale p142

4.2.1 Fibroplasie mediale p142

4.2.2 Fibroplasie périmediale p144

4.2.3 Hyperplasie mediale p144

4.3 DFM adventitielle p144

5 Aspects cliniques et histoire naturelle de la DFM p145

5.1 DFM rénale p146

5.2 DFM des artères cervicales, encéphaliques et anévrismes intracrâniens p146

5.3 DFM des artères coronaires p147

6 Composante héréditaire de la DFM p147

- 7 Génétique de la DFM p148
- 8 Hypothèses physiopathologiques p151
 - P151
 - 8.1 L'hypothèse hormonale p151
 - 8.2 L'hypothèse ischémique p151
 - 8.3 L'hypothèse mécanique p152
- 9 Diagnostics différentiels et pathologies vasculaires approchantes p157
 - 9.1 Les maladies sténosantes p157
 - 9.1.1 Les maladies de la voie RAS p157
 - 9.1.1.1 La voie RAS p157
 - 9.1.1.2 Les RASopathies p159
 - 9.1.1.3 La sténose supra-avalvulaire aortique et l'haploinsuffisance du gène ELN p161
 - 9.1.2 Les maladies de la voie NOTCH p163
 - 9.1.2.1 La voie Notch p163
 - 9.1.2.2 Les mutations de NOTCH1 p163
 - 9.1.2.3 Les mutations de NOTCH2, JAG1, et le syndrome d'Alagille p163
 - 9.1.2.4 Les mutations NOTCH3 et le syndrome CADASIL p165
 - 9.1.3 La cascade du NO p167
 - 9.1.3.1 Les mutations GUCY1A3 et le syndrome de Moya-Moya p169
 - 9.1.4 La vasculopathie liée à ACTA2 p172
 - 9.2 Les maladies anévrysmales et disséquantales p173
 - 9.2.1 Les maladies de la matrice extracellulaire p173
 - 9.2.1.1 Atteintes structurelles de la MEC p173
 - 9.2.1.1.1 Le syndrome d'Ehlers-Danlos vasculaire p173
 - 9.2.1.1.2 le syndrome HANAC et la porencéphalie héréditaire p175
 - 9.2.1.2 Trouble métabolique de la MEC : les maladies de la voie du TGFβ p177
 - 9.2.1.2.1 La maladie de Marfan p177
 - 9.2.1.2.2 Les syndromes de Loeys-Dietz p179
 - 9.2.1.2.3 Le syndrome de tortuosité artérielle et le cutis laxa AR1b p180
 - 9.2.2 Les maladies de la contraction de la CML p181
 - 9.2.3 Les TAAD de physiopathologie inconnue p181
- 10 Etude génétique de la DFM p183
 - 10.1 Présentation des familles étudiées p183
 - 10.2 Sélection des individus faisant l'objet d'un séquençage d'exome p183
 - 10.3 Séquençage d'exome : données brutes et application des filtres de base p186
 - 10.4 Stratégie d'analyse des données p186

- 10.4.1 Recherche ciblée des gènes de maladies sténosantes, anévrysmales et disséquantes p189
 - 10.4.1.1 Identification d'une mutation perte de fonction de SMAD3 p189
 - 10.4.1.2 Identification d'une mutation pathogène de COL4A2 p189
 - 10.4.1.3 Identification d'une mutation COL4A2 de pathogénicité inconnue p191
- 10.4.2 Analyse interfamiliale p193
- 10.4.3 Analyse intrafamiliale p196
 - 10.4.3.1 Analyse de trios p196
 - 10.4.3.2 Recherche ciblée de gènes candidats p200
 - 10.4.3.2.1 Phénotypes murins p200
 - 10.4.3.2.2 Voie du HAVS et des diagnostics différentiels p200
 - 10.4.3.3 Analyse combinée et priorisation de gènes candidats p205
 - 10.4.3.3.1 Analyse DAVID p209
 - 10.4.3.3.2 Analyse STRING p209
 - 10.4.3.3.3 Analyse ENDEAVOUR p210
 - 10.4.3.3.4 Rôle des gènes identifiés par priorisation p212

11. Premières conclusions p221

Chapitre 5 : Discussion générale et perspectives p225

Bibliographie p238

Annexes p273

Chapitre 1

Les stratégies d'identification de gènes de maladies génétiqes

Objectifs de travail

1. Les maladies mendéliennes

1.1 Combien de gènes, combien de maladies?

Le génome humain, selon les calculs les plus récents, contiendrait environ 19 000 gènes codant des protéines (source Projet Génome Humain, Ezkurdia et al., 2014; Flicek et al., 2014).

Les maladies mendéliennes sont individuellement rares et touchent moins de 5% des individus mais correspondent à un très grand nombre d'entités individuelles. Selon les récentes statistiques d'OMIM (<http://www.omim.org/statistics/entry>) d'août 2015 (tableau 1), le nombre de maladies à transmission mendélienne certaine ou probable est respectivement d'environ 5 700 et 1 700, soit un total de plus de 7000 maladies. Cependant, ces chiffres correspondent à une estimation très basse, compte-tenu de l'absence d'exhaustivité des entrées dans OMIM. En effet, la majorité de la population mondiale (représentée par les continents africain et asiatique, dont la Chine et l'Inde) n'ont pas accès à une médecine de qualité et la caractérisation phénotypique dans ces populations est très incomplète, par manque d'infrastructures, de ressources, et d'expertise dans le domaine de la génétique médicale. Compte-tenu du nombre important d'unions consanguines à travers le monde (20 à 50%, Bittles, 2001) surreprésentées dans ces continents, aucune estimation satisfaisante du nombre de maladies génétiques existantes ne peut être conduite à partir du catalogue OMIM de maladies mendéliennes humaines.

Des méthodes alternatives de calcul du nombre de maladies potentielles ont été proposées : d'une part, une extrapolation à partir du taux de mutation dans le génome humain, et du nombre de gènes essentiels (vitaux) ont permis d'estimer 7 000 à 15 000 gènes de maladies (Cooper et al., 2010).

D'autre part, l'utilisation du génie génétique à grande échelle pour la génération de mutants perte de fonction chez la souris (The International Knockout Mouse Consortium) ont montré que 70% des gènes inactivés étaient associés à un phénotype mais avec viabilité foetale, et 30% des gènes étaient associés à une mortalité embryonnaire (Ayadi et al., 2012).

1.2 Gènes, types de mutations et maladies Mendéliennes

PRÉFIXE	TYPE D'ENTRÉE OMIM	AUTOSOMIQUE	LIÉ À L'X	LIÉ À L'Y	MITOCHONDRIAL	TOTAL
*	Description de gène	14221	699	48	35	15003
+	Gène et phénotype combinés	82	2	0	2	86
#	# Description Phénotypique, bases moléculaires connues	4200	294	4	29	4527
%	% Description Phénotypique, bases moléculaires non connues	1515	130	5	0	1650
	Autres, phénotypes avec suspicion de transmission mendélienne	1706	112	2	0	1820
	Total	21724	1237	59	66	23086

Tableau 1: Statistiques du nombre d'entrées OMIM , par type (Aout 2015).

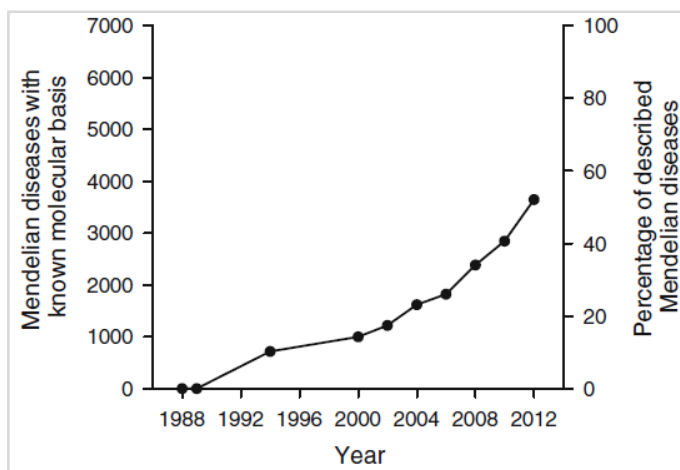


Figure 1: Nombre et pourcentages de maladies mendéliennes élucidées entre 1988 et 2012 (D'après Brunham et al, 2013).

Ce dernier résultat éclaire le nombre de maladies Mendéliennes auxquelles la communauté des généticiens devra faire face dans les décennies à venir, mais est peut être sous-estimé. En effet certaines maladies génétiques humaines ne trouvent pas d'équivalence exacte chez la souris en termes de sévérité du phénotype. Un exemple bien connu illustre cette problématique : dans la mucoviscidose, le modèle murin est très incomplet et modéré y compris après l'introduction de la mutation récurrente $\Delta F508$. Dans le cas du Syndrome d'Ehlers-Danlos vasculaire, le mutant KO hétérozygote pour COL3A1 ne présente aucun phénotype spontané (Faugeroux et al., 2013) contrairement à l'haploinsuffisance chez l'homme responsable d'un phénotype vasculaire retardé mais non atténué (Frank, Albuissou, Ranque et al., 2015; Pepin et al., 2014).

De plus, les phénotypes obtenus par le travail à grande échelle effectué chez la souris ne correspondent qu'à des pertes de fonction de gènes, et il reste à déterminer pour chaque gène potentiellement pathogène le rôle d'allèles hypomorphes, ou de mutations gains de fonction. A titre d'exemple, un quart des gènes associés à un phénotype dans OMIM (soit plus de 700 gènes) sont responsables d'au moins 2 phénotypes distincts (statistiques de février 2015, Selon Chong et al., 2015) ce qui n'est qu'un aperçu de l'effet pléiotrope des mutations attendues dans la population humaine pour les 19 000 gènes codant des protéines.

Ce dernier point soulève la question de l'élargissement du spectre phénotypique dans ces pathologies : dans certains cas bien connus d'effet pléiotrope, il s'agit bien de pathologies distinctes liées au même gène. C'est par exemple le cas du gène *PSEN1* appartenant au complexe gamma-sécrétase, dont les mutations inactivatrices sont responsables d'Acne Inversa, et les mutations gain de fonction sont responsables de maladie d'Alzheimer précoce (Chartier-Harlin et al., 1991). Cependant certaines maladies considérées comme "alléliques" telles que les syndromes craniosynostotiques de Pfeiffer et de Jackson-Weiss liés aux mutations de *FGFR1* ou *FGFR2* (Robin et al., 1994; Steinberger et al., 1996), illustrent bien la nécessité d'une approche modernisée des pathologies que nous étudions. Une vision intégrative et élargie du spectre phénotypique des pathologies, et une probable homogénéisation sur la base de catalogues de symptômes semblent nécessaires dans les années à venir. En l'absence d'évolution des pratiques dans ce domaine, nous risquons d'être noyés par un nombre erroné et surtout surestimé d'entités cliniques qui ne constitueraient pas de véritables entités sur le plan mendélien, induisant la multiplication d'informations non croisées auxquelles les scientifiques et généticiens font appel dans leurs travaux de recherche.

1.3 Les maladies identifiées

Parmi les maladies génétiques cataloguées dans OMIM, environ 70 % ont été élucidées sur le plan génétique (tableau 1). Cette proportion était de 50% il y a 3 ans (figure 1, Brunham et Hayden, 2013), et est aussi en augmentation constante, particulièrement depuis l'avènement du séquençage haut-débit et de la généralisation du séquençage d'exome dans les maladies orphelines. On estime entre 200 et 300 le nombre

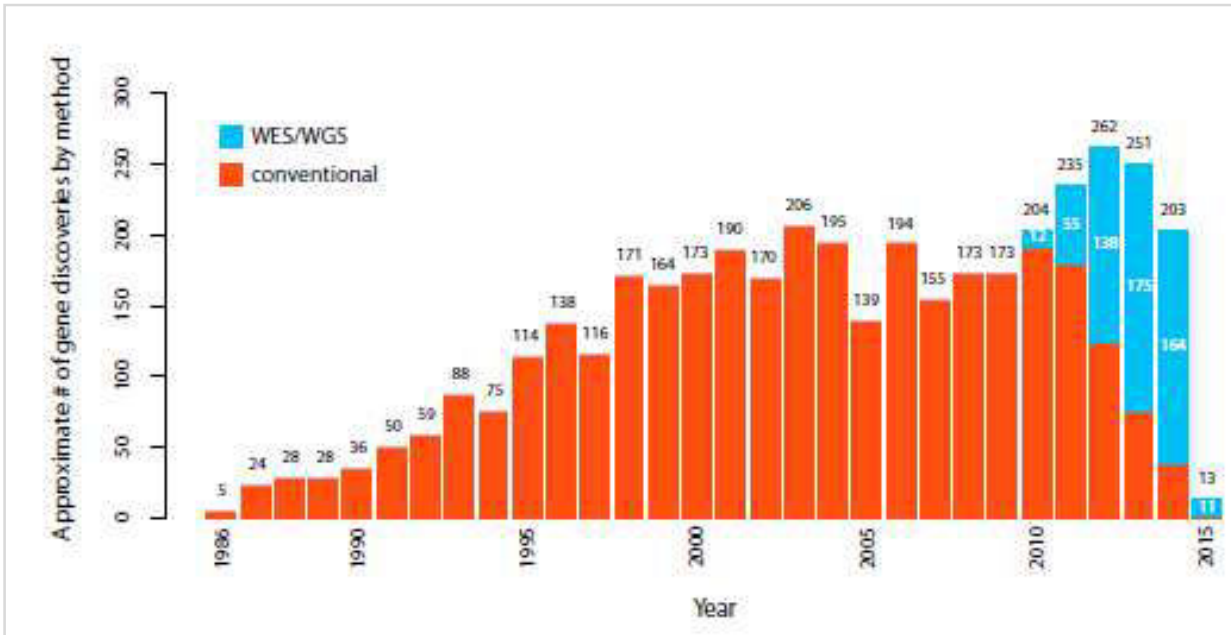


Figure 2: Nombre de gènes découverts par séquençage d'exome et de génome (en bleu) en comparaison des approches conventionnelles (en orange) depuis 2010 (d'après Chong et al, 2015).

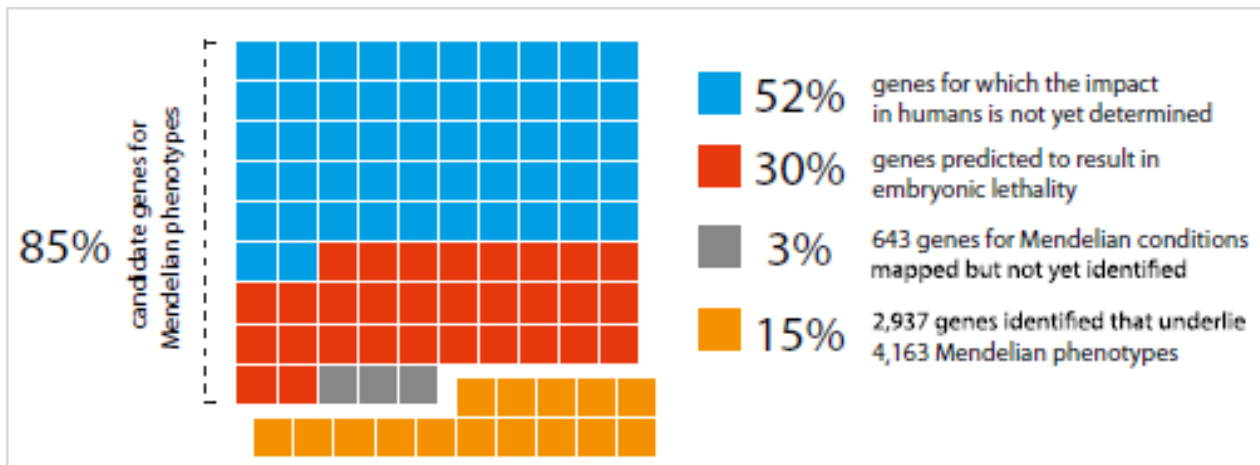


Figure 3: Lien entre gènes de protéines humains et phénotypes Mendéliens. Parmi les 19 000 gènes de protéine prédits dans le génome humain, environ 15% ($n \approx 2937$) ont été liés à des variants causaux (carrés oranges). Environ 643 ($\approx 3,38\%$) phénotypes Mendéliens ont été cartographiés (carrés gris) mais les gènes correspondants non identifiés. Sur la base des analyses de KO en modèle murin, les variations perte de fonction d'environ 30% des gènes ($n \approx 5960$) seraient responsables d'embryolétalité (carrés rouges). Un minimum de 52% de gènes ($n \approx 10330$) ont un impact phénotypique encore indéterminé (carrés bleus). Au total, environ 16 063 gènes pourraient être responsables de phénotypes Mendéliens (d'après Chong et al, 2015).

de gènes de maladies identifiées par an depuis 2010 (figure 2). L'ensemble de ces données est bien résumé dans la figure 3 issue des données de Chong et coll. (Chong et al., 2015).

2. Les maladies polygéniques ou communes

Les maladies polygéniques, aussi appelées à génétique complexe ou fréquentes, présentent une hérédité différente avec l'implication de nombreux facteurs génétiques et environnementaux. Une faible proportion des familles présente souvent des formes monogéniques de l'affection, mais la grande majorité des cas présente une récurrence familiale faible ou sont sporadiques. Le trait est alors multifactoriel, liée à l'interaction entre une composante génétique polygénique et des facteurs non génétiques. Dans la composante génétique, Les allèles impliqués ne sont, dans la très grande majorité des cas, pas délétères en eux-mêmes, mais confèrent seulement une susceptibilité accrue à la maladie. Ce mode héréditaire concerne une grande partie des maladies communes telles que les diabètes de type 1 et 2, les maladies cardiovasculaires, certains cancers fréquents, des maladies neurologiques et psychiatriques, des malformations congénitales, l'asthme et les maladies auto-immunes. Leur étude est complexe et comporte traditionnellement trois étapes : montrer que la maladie a une récurrence familiale, montrer que cette tendance familiale est due à des facteurs génétiques et, enfin, identifier les gènes impliqués.

3. Les stratégies d'identification des maladies génétiques Mendéliennes et communes avant le séquençage d'exome

3.1 Maladies Mendéliennes : le clonage positionnel

L'ère moderne de la génétique humaine a débuté dans les années 1980 grâce au clonage positionnel, rendu possible par l'identification de variants polymorphes naturels dans le génome humain (Botstein et al., 1980). Cette découverte a permis le développement de l'analyse de liaison, technique créée en 1913 par Alfred Sturtevant pour l'analyse génétique chez la drosophile.

3.1.1 L'analyse de liaison simple (paramétrique)

L'analyse de liaison est basée sur un principe simple de coségrégation au sein de familles d'un trait avec un ou plusieurs marqueur polymorphes parmi de nombreux marqueurs répartis le long du génome. Elle tire parti des recombinaisons méiotiques entre individus d'une même famille, d'autant plus nombreuses qu'ils sont distants et/ou nombreux, ce qui permet de restreindre la taille des régions incluant le/les marqueur(s) et associées au trait étudié. Le traitement statistique des données permet d'obtenir un "Z-score", score de vraisemblance (après transformation logarithmique) de la liaison entre un marqueur et le trait. Ce score est considéré comme significatif lorsque supérieur à 3 (la liaison du trait au marqueur est 1 000 fois plus probable que l'absence de liaison).

Outre son application dans les maladies mendéliennes homogènes sur le plan génétique, cette méthode s'est également avérée puissante pour l'étude de certains sous-types de maladies multifactorielles, comme par exemple les formes à âge de début précoce de la maladie d'Alzheimer, d'hypertension artérielle ou du cancer du sein (Chartier-Harlin et al., 1991; Miki et al., 1994; Shimkets et al., 1994).

3.1.2 La cartographie d'homozygotie

Elle représente une variante de l'analyse de liaison (Lander et Botstein, 1987) dans le cas des pathologies récessives avec consanguinité, faisant l'hypothèse d'une anomalie homozygote sous-jacente. L'analyse est focalisée sur les régions génomiques de grande taille contenant les marqueurs polymorphes homozygotes afin de diminuer le nombre de gènes à étudier par la suite. Un premier succès de cette technique a été apporté dans la génétique de l'alkaptonurie, à partir de 3 familles partageant un locus commun homozygote (Pollak et al., 1993).

3.1.3 Les autres techniques de clonage positionnel

Diverses techniques ont été développées dans les cas d'échec de l'analyse de liaison simple : analyse de liaison non paramétrique (méthode des paires de germains) permettant de s'affranchir de l'hypothèse du modèle génétique sous-jacent, et combinaison liaison-association (Transmission Disequilibrium test). Ces tests peuvent être aussi appliqués à l'étude génétique des maladies communes. Leur robustesse a pour contrepartie leur manque de puissance statistique, nécessitant donc un grand nombre de familles et de sujets atteints (Cox et al., 2001).

3.1.4 L'évolution et les limites du clonage positionnel

Dans la décennie qui a suivi la découverte des marqueurs polymorphes du génome humain, la cartographie d'environ 60 à 70 gènes de maladies a pu être réalisée : tout d'abord la maladie de Huntington en 1983 (Gusella et al., 1983), puis la mucoviscidose en 1985 (White et al., 1985).

Le nombre de gènes identifiés par analyse de liaison simple est allé croissant depuis lors (figure 2) pour atteindre 180 à 200 gènes par an dans les années 2000. Cette progression a été grandement facilitée par les premiers séquençages du génome humain en 2001 (Lander 2001, Venter 2001) : environ 1000 gènes de maladies ont été identifiés entre 1980 et 2000, contre 3000 entre 2000 et 2010 (figure 1).

Le séquençage du génome humain, et le traitement bioinformatique des données correspondantes ont permis de caractériser et de classer la majorité des gènes de protéines par des approches de phylogénie et de similitude de séquences et de motifs. Il en a résulté la possibilité de combiner les analyses de liaison avec une **approche gènes-candidats**, réduisant fortement le nombre de gènes à tester dans les intervalles de liaison identifiés par cartographie.

Cependant le rendement de ces techniques est nettement plus faible dans les maladies mendéliennes les plus difficiles à élucider. Les maladies de prévalence suffisamment importante, génétiquement homogènes, dont les bases physiopathologiques sont robustes, et de mode de transmission monogénique pur ont été les plus simples à caractériser par ces méthodes.

D'une façon générale, que ce soit pour l'étude de maladies monogéniques ou de formes rares de maladies multifactorielles, il est clair que l'analyse de liaison parvient facilement à détecter la corrélation très forte qui existe entre le gène et la maladie. En revanche, elle est nettement moins adaptée pour l'étude de maladies multifactorielles où l'hétérogénéité génétique (plusieurs gènes impliqués) est très vraisemblable et où les effets des facteurs génétiques impliqués sont faibles (faible corrélation trait-marqueur). A ce titre, le succès dans l'étude de sous-type de certaines maladies multifactorielles repose sur l'existence de grandes familles atteintes, même si globalement elles ne représentent qu'une très faible minorité des cas de la population générale. Les maladies plus complexes et/ou rares ont donc été le parent pauvre de cette période et ont fait l'objet d'une attention accrue lorsque le séquençage d'exome a été rendu possible (voir section 4. Le séquençage d'exome).

3.2 Maladies communes : Les études d'association GWAS

3.2.1 Echec de la liaison génétique

Les premières études génétiques des maladies communes ont été réalisées dans la continuité des succès du clonage positionnel. A partir de formes familiales rares de maladies communes, de nombreuses régions génomiques d'intérêt ont été publiées dans le diabète de type 2, ou la schizophrénie à titre d'exemple (figure 4). Cependant la majorité des résultats de ces études n'ont pas pu être répliqués, pour diverses raisons dont l'hétérogénéité génétique certaine de ces modèles familiaux rares, la présence de facteurs confondants (phénotopies, fréquentes *de facto* dans les maladies communes), ou le manque de puissance statistique.

3.2.2 Les études d'association et l'âge d'or du GWAS

Les études d'association centrées sur des gènes candidats étaient utilisées avec une efficacité relative avant l'ère du pangénomique dans ce domaine. Elles comportaient cependant une limite majeure : la nécessité de poser une hypothèse physiopathologique de départ, qui peut s'avérer très limitante. Cet aspect est très bien illustré par la discordance au sein des nombreuses études d'association réalisées dans le domaine des maladies cardiovasculaires : l'une des hypothèses implique le stress oxydatif, qui joue un rôle important dans le développement de l'athérosclérose. Plusieurs études ont incriminé des variants de la GST (Gluthatione S-transferase) dans la maladie coronaire (Cora et al., 2013; Hayes et Strange, 2000), mais d'autres études tendent à réfuter cette association (Miller et al., 2003; Yeh et al., 2013). Plusieurs meta-analyses ont vu le jour et permettent partiellement d'éclairer cette discordance : les variants non associés correspondraient à des effets protéiques peu marqués, seule une perte de fonction de *GSTM1* serait associée à la maladie coronaire, lorsque cumulée à l'intoxication tabagique (Attia et al., 2003; Zhou et al., 2014). Comme le démontre cet exemple, seule une approche de type méta-analyse a permis d'apporter une cohérence aux résultats pris séparément. Cela souligne les difficultés de conclure à partir d'études de faible ampleur avec une hypothèse physiopathologique de départ.

La recherche dans le domaine de la génétique des maladies communes s'est tournée ensuite vers les études d'association pangénomiques (GWAS), rendues possibles par les marqueurs à faible densité puis par la création de puces SNP à très haute densité (1 million de SNPs dans les années 2000). De façon concomitante a émergé le principe de "maladies communes-variants communs". Lander (Lander, 1996) explique que si on

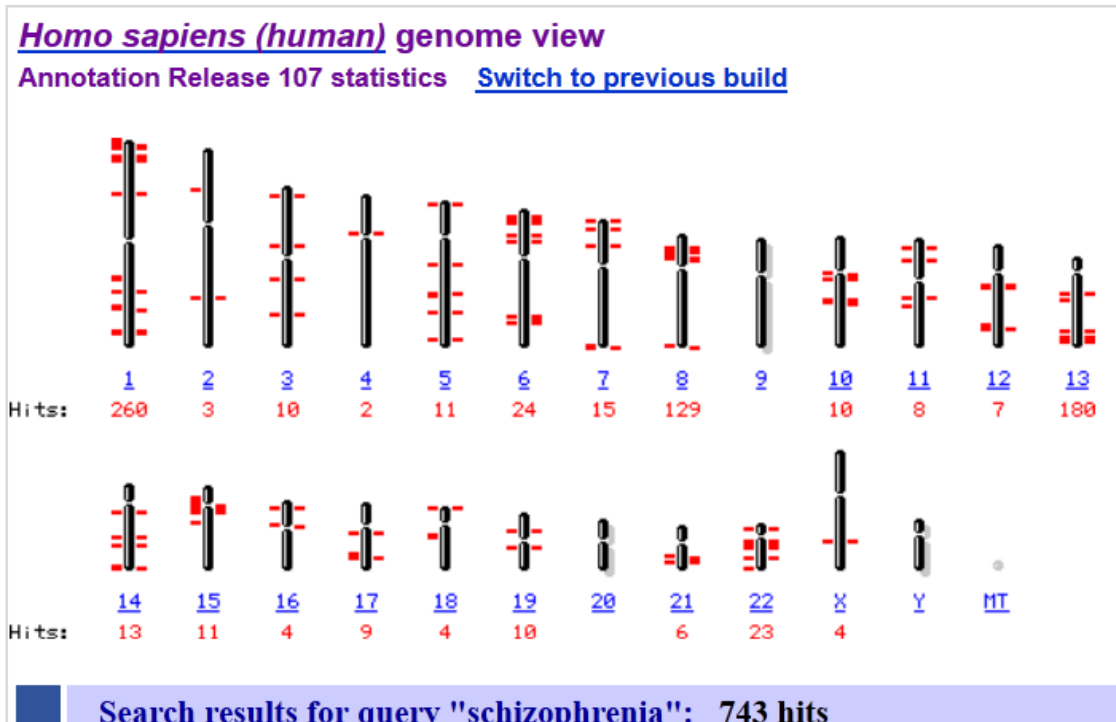


Figure 4: Résultats d’une requête sur le nombre de loci liés à la Schizophrénie à partir du portail NCBI Map Viewer (Aout 2015). De nombreux loci ont été identifiés dans une étude puis non reproduits dans les études ultérieures.

compare la séquence de deux individus, la plupart des différences sont situées dans les polymorphismes anciens et fréquents (>1% ou 5% de la population selon les références), qui devraient donc être responsables d'une partie du risque génétique de développer les maladies communes. Ces variations nucléotidiques deviennent fréquentes car elles sont neutres ou favorables à la survie. Néanmoins, quelques-unes pourraient entraîner des effets adverses modérés dans des conditions d'environnement modifiées (par exemple, préserver la graisse corporelle pendant l'ère glaciaire qui devient un facteur favorisant l'obésité dans notre ère actuelle). Il propose de les cataloguer et d'étudier leur association avec les différentes maladies dans de très grands échantillons, donnant naissance au principe du GWAS. L'hypothèse "maladies communes-variants communs" part du principe que plusieurs SNP fréquents confèrent un risque modeste aux maladies communes, mais qui peut être détecté par leur analyse extensive dans un très large échantillon d'individus. Le principe de cette méthode est donc basé sur une comparaison de fréquences des allèles (étude d'association) des marqueurs communs présents sur tout le génome, entre un groupe de cas et un groupe de témoins, sans hypothèse préalable sur les gènes d'intérêt.

Deux éléments principaux sont à prendre en compte : i) le nombre très important de tests statistiques effectués implique une correction sur tests multiples (une valeur de p inférieure à 5×10^{-8} étant généralement considérée comme significative) nécessitant des effectifs de plusieurs milliers de cas et de témoins pour obtenir une puissance statistique suffisante ; ii) des études de réplification indépendantes sont de plus nécessaires pour diminuer encore le risque de résultats positifs aléatoires, impliquant des projets multicentriques dans le cadre de consortia internationaux.

En 2005, une première étude permet l'identification de *CFH*, gène de susceptibilité dans la dégénérescence maculaire liée à l'âge, à partir de 96 cas, 50 contrôles et environ 120 000 marqueurs, ce qui représente aujourd'hui un succès très atypique compte tenu du peu de marqueurs et d'échantillons utilisés (Haines et al., 2005).

Dans les années qui ont suivi, des milliers d'études de ce type ont été réalisées et ont permis d'identifier sans a priori des centaines de régions et gènes de susceptibilité dans de nombreuses maladies communes comme les dyslipidémies, le diabète de type 2, le cancer, les maladies cardiovasculaires... L'un des plus grands succès d'études GWAS est publié en 2007 dans Nature (Burton et al, Wellcome Trust Case Control Consortium) utilisant 500 000 SNP, 2 000 cas et 3 000 contrôles. Parmi les 7 pathologies étudiées (trouble bipolaire, maladie coronaire, hypertension artérielle, maladie de Crohn, polyarthrite rhumatoïde, diabète de type 1 et 2), 24 marqueurs indépendants ont pu être associés à six de ces maladies ($p < 5 \cdot 10^{-7}$). Depuis, la plus grande densité des marqueurs SNPs disponibles sur puces à ADN, la possibilité d'inférer encore d'autres SNPs par les nombreuses études extensives effectuées sur plusieurs populations et surtout le génotypage d'un bien plus

grand nombre de cas et témoins ont permis l'identification d'un nombre supplémentaire de loci de susceptibilité pour ces différents traits ou maladies complexes.

3.2.3 Les limites du GWAS

Cependant comme l'illustre la figure 5 issue de Manolio et coll. (2009), La majorité des allèles identifiés par cette approche dans quasiment toutes les études menées confèrent un effet faible (Odds-Ratio <2) et n'ont aucune application clinique potentielle. En effet sur le plan théorique le GWAS se place dans l'hypothèse "maladie commune-variants communs" où la maladie commune est considérée comme la conséquence d'un nombre faible de variants de fréquence élevée (Halushka et al., 1999; Lander, 1996). Ces résultats décevants ont mené au concept d'héritabilité manquante (missing heritability) : A titre d'exemple, plus de 30 loci ont été identifiés dans la maladie de Crohn, qui sous l'hypothèse de l'effet additif n'expliquent que 20% de l'héritabilité totale estimée (Manolio et al., 2009). D'autres exemples sont illustrés dans le tableau 2, qui permet d'expliquer le succès des GWAS à petite échelle dans la dégénérescence maculaire liée à l'âge, où l'héritabilité est expliquée à 50% par des variants communs.

Cette héritabilité manquante peut être expliquée selon Manolio (Manolio et al., 2009) et Maher (Maher, 2008) par plusieurs facteurs :

Un plus grand nombre de variations avec un effet modéré non encore identifiées

Les SNP communément étudiés dans les GWAS, de par leur fréquence et leur ancienneté, représentent des blocs haplotypiques "verrouillés". Ainsi, par exemple, si deux individus ont le même génotype sur un SNP donné, ils peuvent tous deux être notés comme ayant la même version d'un gène lié à la pathologie sous-jacente dans cette région. Ceci est vrai même si l'une des deux personnes porte un variant relativement moins fréquent, inclus dans le même bloc haplotypique, ou non présent dans le pool de marqueurs utilisés, mais qui a un effet important sur la maladie. Les résultats de GWAS sur ce type de SNP pourrait permettre d'identifier un locus morbide, mais l'analyse à l'échelle de centaines d'individus pourrait donner l'impression que ses effets sont assez faibles : l'effet de la variation sous-jacente serait "dilué". Dans cet exemple, l'identification de l'effet de ce SNP implique d'utiliser des puces de plus haute densité avec des marqueurs de fréquence allélique un peu plus faible afin de distinguer les sous-groupes correspondants. Une autre option est de séquencer les gènes candidats et leurs régions avoisinantes chez des milliers de personnes pour mettre en évidence de nouvelles associations majeures avec le trait étudié. Les limites de ces

Estimates of heritability and number of loci for several complex traits

Disease	Number of loci	Proportion of heritability explained	Heritability measure
Age-related macular degeneration	5	50%	Sibling recurrence risk
Crohn's disease	32	20%	Genetic risk (liability)
Systemic lupus erythematosus	6	15%	Sibling recurrence risk
Type 2 diabetes	18	6%	Sibling recurrence risk
HDL cholesterol	7	5.2%	Residual* phenotypic variance
Height	40	5%	Phenotypic variance
Early onset myocardial infarction	9	2.8%	Phenotypic variance
Fasting glucose	4	1.5%	Phenotypic variance

*Residual is after adjustment for age, gender, diabetes.

Tableau 2: Estimation de l'héritabilité et du nombre de loci dans diverses maladies communes (d'après Manolio et al, 2009).

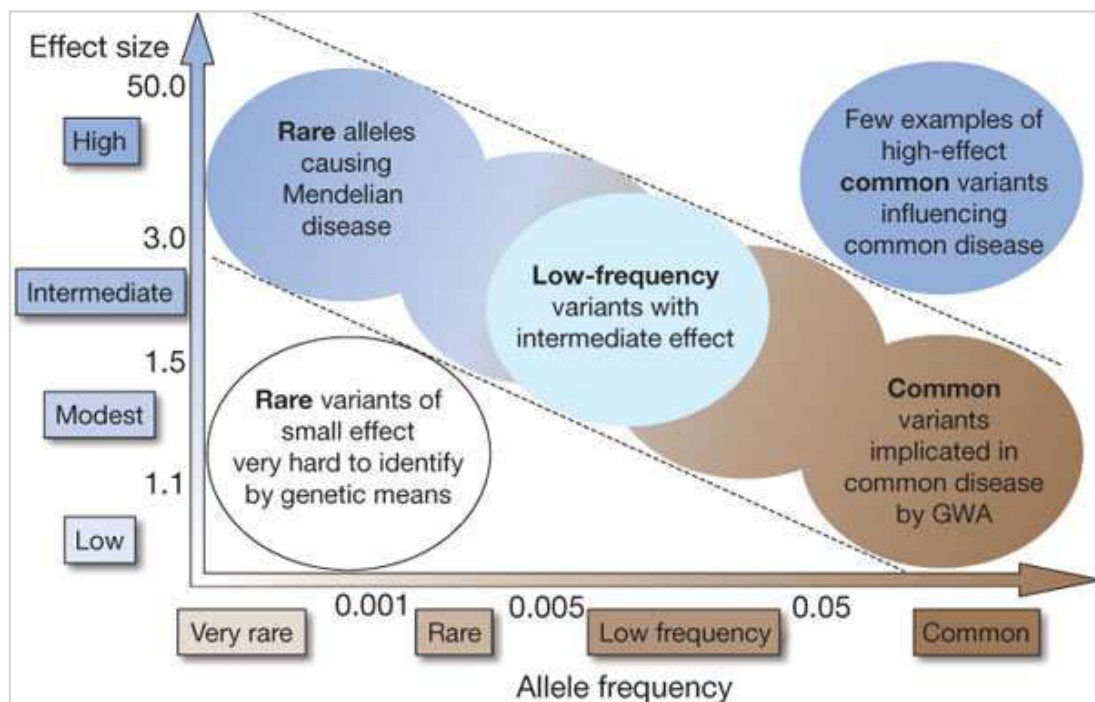


Figure 5: possibilités d'identification des variants selon la fréquence de l'allèle à risque et la force de l'effet génétique (odds-ratio). Jusqu'en 2009, les efforts étaient concentrés sur les situations représentées entre les lignes en pointillés (d'après Manolio et al, 2009).

approches sont liées au nombre importants de nouveaux cas à recruter et au coût encore élevé du séquençage à réaliser pour plusieurs milliers de cas.

Des variations rares (<0.5%) ou très rares (<0.05%) non présentes dans les systèmes de génotypage disponibles

Une grande partie de la spéculation sur l'héritabilité manquante dans les GWAS a porté sur la contribution possible des variations de fréquence allélique (ou MAF pour Minor Allelic Frequency) faible (< 0,5%). Ces variations sont généralement trop rares mais surtout bien trop nombreuses pour être capturées par les techniques de génotypage utilisées dans les GWAS actuels, et leur effet n'est pas suffisamment important pour être détecté par les analyses de liaison classiques (zone intermédiaire dans la figure 5. Récemment, les études d'exome à grande échelle ont révélé que plus de 80% des 500 000 variants nucléotidiques (ou SNV pour Single Nucleotide Variants) identifiés en moyenne chez un être humain sont des variants rares (Tennessen et al., 2012). Ceci suggère l'implication possible de ces variants dans des formes monogéniques ou oligogéniques de maladies communes, dont l'hétérogénéité génétique est majeure, menant à l'hypothèse « maladie commune-variants rares » : les variants à effet délétère majeur disparaissant *de facto* par sélection purificatrice, la fréquence de tels variants dans la population est donc très faible. Les variants rares sont par ailleurs 4 fois plus susceptibles d'être fonctionnels (basée sur des prédictions bioinformatiques) que les variants fréquents (Tennessen et al., 2012). La plupart des SNP utilisés dans les GWAS se composent de variants communs antérieurs à la migration de l'homme moderne hors d'Afrique et sont partagés entre populations des 5 continents, représentés sur HapMap (International HapMap Consortium, 2003). La plus grande part de la variabilité humaine représentée par les variants rares, est la conséquence de phénomènes mutationnels récents, qui ont eu lieu avec l'expansion rapide de la population au cours des 10 000 dernières années (Tennessen et al., 2012).

Les limites des GWAS spécifiques à cette part d'héritabilité manquante que sont les variants rares ont partiellement pu être repoussées grâce à la possibilité de séquencer l'exome de patients atteints. D'autre part, la mise sur le marché de puces SNP comportant une majorité de variants rares (et pas ou peu de variants communs), dites puces "Exome" contribue aux progrès dans ce domaine. Ces puces ont vu le jour secondairement à la génération des données issues du séquençage d'exomes dans de grandes cohortes d'individus issus de divers groupes ethniques. Elles sont donc très récentes.

Des variations structurelles mal détectées par les études actuelles : les CNV

Une autre hypothèse, permettant d'expliquer l'héritabilité non retrouvée par les études d'association, est basée sur l'observation d'une surabondance d'anomalies chromosomiques ou CNV (pour Copy Number Variations) chez les patients atteints de certaines pathologies comme la maladie de Crohn, l'autisme, le

Selected disease associations with rare CNVs and common CNPs						
Disease	Locus	Type of CNV	Size (kb)	Population frequency	Case frequency	Effect size (OR)
Rare CNVs						
Autism idiopathic MR	16p11.2	De novo deletion	600	1.10^{-4}	1%	100
Autism	16p11.2	Rare duplication	600	3.10^{-4}	0.50%	16
Schizophrenia	1q21.1	Rare deletion	1,4	2.10^{-4}	0.30%	15
IMR	1q21.1	Rare deletion	1,4	2.10^{-4}	0.47%	Not observed in 4,737 controls
Schizophrenia	15q13.3	Rare deletion	1,6	2.10^{-4}	0.20%	12
Epilepsy	15q13.3	Rare deletion	1,6	2.10^{-4}	1.0%	Not observed in 3,699 controls
IMR	15q13.3	Rare deletion	1,6	2.10^{-4}	0.30%	Not observed in 960 controls
Schizophrenia	22q11.2	Rare deletion	3	2.510^{-4}	1%	40
Common CNPs						
Crohn's disease	IRGM	Deletion polymorphism	20	7%	11%	1.5
Body mass index	NEGR1	Deletion polymorphism	45	65%	Quantitative trait	<1 kg
Psoriasis	LCE3C	Deletion polymorphism	30	55%	65%	1.3

Tableau 3: Quelques pathologies associées à des variants de nombre de copies génomiques ou CVN (pour Copy Number Variants). Certains sont des événements rares mais d'autres sont des polymorphismes fréquents, en rapport avec un odds-ratio faible (d'après Manolio et al, 2009).

psoriasis (tableau 3). Il a été démontré en 2009 que les CNV avaient un impact majeur sur l'expression de gènes, impliquant notamment leur épissage et leur activité transcriptionnelle dans, mais aussi à proximité, des réarrangements (Henrichsen et al., 2009). Des analyses plus sophistiquées réalisées à partir de puces de génotypage et une modernisation de ces puces permettent à ce jour d'obtenir les résultats de génotypage SNP et CNV à partir des mêmes données.

Une non-détection des marques et processus épigénétique et des interactions gène- environnement

La méthylation de l'ADN est un processus majeur de régulation de l'expression des gènes à l'échelle du génome chez les vertébrés. L'influence de l'environnement dans l'évolution de la méthylation (notamment des séquences promotrices) des gènes au cours de la vie a été démontrée dans de nombreux domaines. L'exemple le plus simple et le plus parlant en est probablement le spectre des malformations neuroectodermiques par déficit en acide folique pendant la croissance embryonnaire. Le phénomène sous-jacent est un défaut de méthylation des régions promotrices de gènes majeurs du développement. Les études de jumeaux sont un point de départ important de l'estimation de la part de méthylation du génome dans l'héritabilité manquante : Fraga et coll. (Fraga et al., 2005) ont démontré la dérive épigénétique entre jumeaux au cours du vieillissement en terme de méthylation d'ADN. Cette dérive est très évocatrice de l'influence d'un environnement non partagé sur l'épigénome. Compte-tenu du taux élevé de discordance phénotypique des jumeaux dans le domaine des maladies communes, il est probable que les marques épigénétiques acquises sous l'influence de l'environnement soient un facteur important d'héritabilité manquante, comme cela l'a été suggéré (Trerotola et al., 2015). Le profil de méthylation du génome (ou méthylome) fait cependant tout récemment l'objet d'études de type GWAS (Methylome-Wide Association Studies ou MWAS) et pourrait offrir à l'avenir un complément d'information prometteur : une étude menée en 2014 par McClay (McClay et al., 2014) chez plus de 700 individus de 25 à 92 ans a mis en évidence 11 loci différenciellement méthylés associés à l'âge ($p < 1.10^{-8}$), dans des familles de gènes de protocadherines, d'homeobox, de MAP Kinases et de récepteurs à la Ryanodine. Ces familles de gènes sont susceptibles d'être fortement impliquées dans les divers processus de vieillissement pathologique. Cependant les méthodes de MWAS sont très récentes et les données robustes issues de leur mise en œuvre sont quasiment inexistantes.

Dans le même ordre d'idées, l'origine parentale des allèles, sous-tendue par le phénomène d'empreinte, est un processus qui n'est jamais pris en compte dans les approches GWAS jusqu'à très récemment. Kong en 2009 a mis en évidence l'importance de l'origine parentale dans la survenue du diabète de type 2 (Kong et al., 2009) et estime à 10 à 15% l'héritabilité liée à cette origine parentale (Eichler et al., 2010).

D'une façon beaucoup plus large, les interactions gène-environnement ne sont pas réellement prises en compte dans les études de GWAS, simplement par manque de puissance statistique mais surtout d'outils puissants de modélisation des relations gène-environnement dans un organisme complexe comme l'homme.

Une faible puissance pour détecter les interactions gène-gène

L'interaction gène-gène, ou épistasie, se réfère dans le domaine des traits quantitatifs à la part de la variance génétique qui ne peut être expliquée ni par les effets additifs des allèles en présence, ni par les effets de dominance simple. Plus simplement elle définit l'existence d'un phénomène plus complexe que la somme des effets des gènes impliqués dans un phénotype. Des phénomènes d'épistasie ont été mis en évidence dans la susceptibilité au cancer (Fijneman et al., 1996) et dans les maladies auto-immunes (WandstratetWakeland, 2001) et des processus morphogénétiques (Leamy et al., 2002) mais reste complexe à évaluer par l'approche GWAS. A ce jour aucune étude à grande échelle en GWAS impliquant une analyse d'interaction gène-gène n'a été publiée.

4. Le séquençage d'exome

4.1 Le séquençage haut débit

C'est en 2005-2006 qu'est apparue sur le marché la première génération des appareils de "séquençage à haut débit" ou "séquençage massif parallèle".

Cette technologie utilise des procédés de synthèse chimique ou enzymatique de l'ADN en milieu solide ou liquide, générant dans le même temps sur un support dédié des milliards de séquences à partir de l'ADN déposé. La puissance de cette technique réside dans les capacités de miniaturisation et l'utilisation de supports microfluidiques permettant la séparation selon diverses modalités de ces milliards de séquence obtenues.

L'ADN qui correspond à la cible d'intérêt à séquencer, est obtenu par diverses techniques selon la taille de la cible recherchée : enrichissement (ou capture) par hybridation pour les cibles de grande taille comme l'exome, ou simple PCR multiplex pour les cibles de petite taille dans un contexte de "séquençage ciblé". Une grande variété de techniques d'enrichissement est disponible sur le marché correspondant à tous les types et toutes les tailles de cibles recherchés.

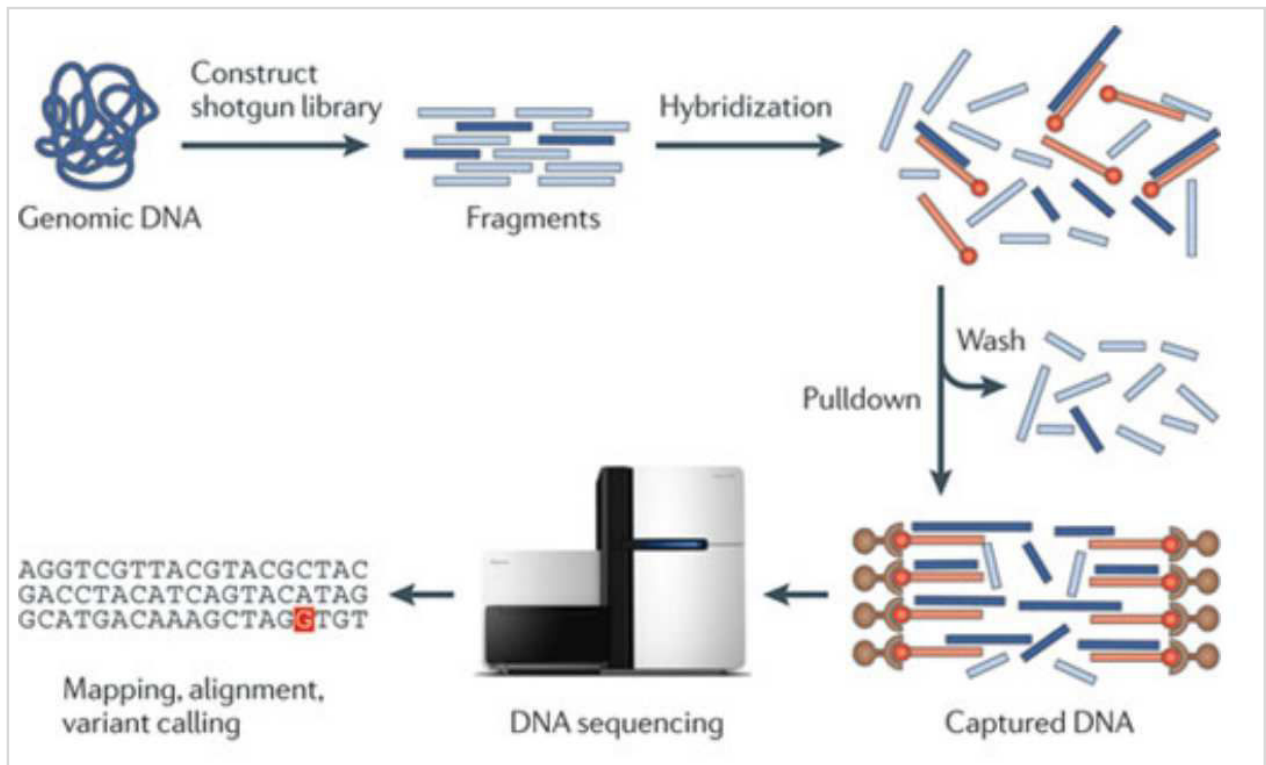


Figure 6: Principe et déroulement du séquençage à haut débit (d'après Bamshad et al, 2011).

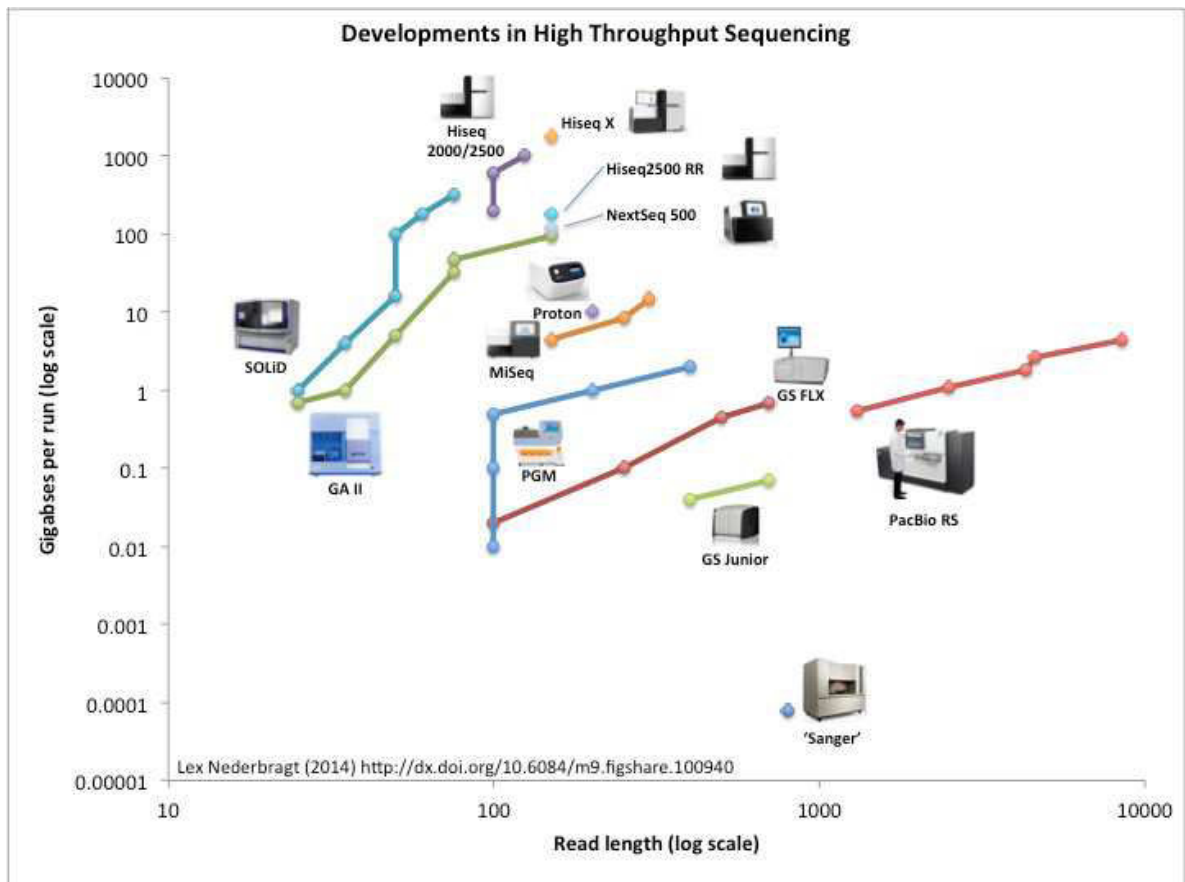


Figure 7: Echelle de séquençage des différents appareils. Le séquençage de type Sanger est représenté en bas par un séquenceur 96 capillaires Applied Biosystems (d'après Lex Nederbragt, Université d'Oslo).

La technique en elle-même comporte trois étapes distinctes : la préparation d'une librairie, l'hybridation, et le séquençage par synthèse d'ADN (figure 6).

Les étapes ultérieures à la génération des séquences font l'objet de divers traitements informatiques permettant l'obtention de séquences de qualité correcte. Ensuite sont mises en place des méthodologies de filtrage (discutées dans la section 4.3.1) dans les diverses applications qui en découlent.

Le séquençage haut débit représente un changement d'échelle par rapport au séquençage de type Sanger, allant d'un facteur 1 000 pour un appareil de type Roche Junior à 100 millions pour un appareil de type Illumina HiSeqX (figure 7). La baisse critique des coûts de séquençage par cette technologie d'année en année a permis un développement considérable de son utilisation à travers le monde.

Enfin le développement d'appareils de séquençage de plus en plus puissants (figure 8) a rendu possible le séquençage de génomes humains entiers, s'affranchissant des méthodes de préparation ou d'enrichissement de l'ADN et des biais de couverture qui y sont associés. Ces génomes ne sont pas des « chimères » constituées de séquences consensus tels que l'étaient les premiers génomes humains séquencés en 2001, et constituent donc une source réelle de toutes les informations génétiques d'un individu.

4.2 Pourquoi séquencer un exome

Le terme exome, néologisme créé en 2008 par l'équipe de Craig Venter (Ng et al, 2008), définit l'ensemble des exons codants d'un génome donné. Le séquençage d'exome représente jusqu'à présent le meilleur compromis entre "exhaustivité" en terme de mutations génétiques causales et effort de séquençage : d'une part l'exome correspond à environ 30-40 Mb de séquence codante soit seulement 1-3 % du génome humain, et d'autre part la très grande majorité des mutations identifiées dans des gènes par des approches de clonage positionnel siègent dans la séquence codante ou à proximité immédiate des gènes de protéines (Stenson et al., 2009).

Le premier séquençage d'exome a été réalisé en 2009 chez un individu sans maladie génétique connue (Ng et al., 2008), puis par la même équipe dans une maladie dominante de diagnostic connu, le syndrome de Freeman-Sheldon (Ng et al., 2009). Cette première étude apporte la preuve de l'identification de la mutation en cause par l'utilisation de filtres simples. Très rapidement cette équipe a appliqué avec succès cette approche dans une maladie récessive, le syndrome de Miller (Ng et al., 2010a). La même année, une équipe concurrente a identifié la cause génétique d'une maladie avec mutation *de novo*, le syndrome de Schinzel-Giedon (Hoischen et al., 2010).

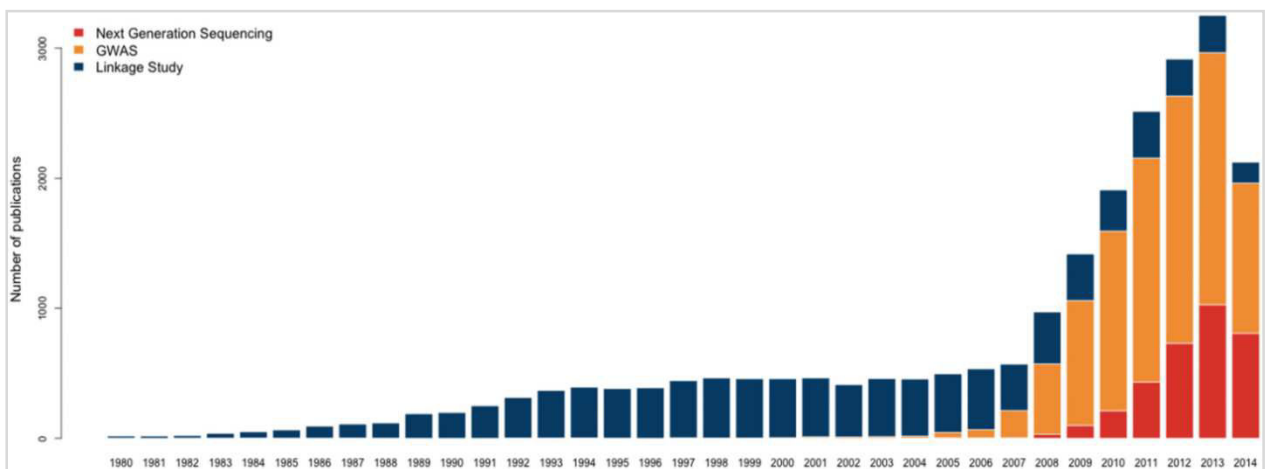


Figure 8: Nombre de publications portant sur la caractérisation de traits ou pathologies humains selon diverses techniques, entre 1980 et 2014. Seules les publications citées dans Medline sont incluses (d'après Wang et al, 2015).

Le nombre de publications ayant dès lors utilisé le séquençage d'exome est allé croissant depuis cette « preuve du concept », et a clairement supplanté la cartographie comme méthode d'identification de gènes dans le domaine des maladies mendéliennes (figure 2), mais aussi dans le domaine des maladies communes (figure 8). Comme l'illustre la figure 2, depuis 2013 trois fois plus de gènes ont été identifiés par des techniques de séquençage d'exome que par des approches « conventionnelles ». En fait, bon nombre d'études combinent une approche de cartographie de liaison et de séquençage d'exomes par l'utilisation de l'information génétique obtenue sur les exomes.

4.3 Le séquençage d'exome dans les maladies mendéliennes

4.3.1 Résultats bruts d'une analyse d'exome et utilisation de filtres basiques

4.3.1.1 Bases de données de variants et fréquence allélique

Les données issues des premiers séquençages d'exomes ont montré que chaque individu porte environ 20 000 variations nucléotidiques s'il est caucasien et jusqu'à 25 000 variations s'il est africain. Comme l'illustre le tableau 4, 95 à 98% de ces variations sont déjà connues dans les bases de données de polymorphismes comme dbSNP, 1000 génomes, et HapMap (Bamshad et al., 2011). Leur élimination constitue une bonne base de départ pour réduire le nombre de gènes à analyser. Cependant, ces bases doivent être utilisées avec précaution car elles sont « polluées » par des variations dont la pathogénicité est certaine, notamment dbSNP. Dans une étude menée en 2010 par Walsh sur la surdité non syndromique, il est apparu que l'une des mutations, confirmée par cette étude, était présente dans dbSNP, rapportée cependant une seule fois (Walsh et al., 2010). Une approche plus rationnelle de l'utilisation de ces bases, et largement utilisée à ce jour, consiste à n'en exclure que les variants fréquents (>1%). Ce seuil ne doit pas être considéré comme immuable, particulièrement dans le cas de maladies récessives fréquentes autosomales ou liées à l'X, où la fréquence des variants potentiels dans la population peut atteindre 1%.

Des bases de données d'exomes ont été récemment mises à disposition de la communauté scientifique : Exome Variant Server et EXAC, qui comportent respectivement les données de plus de 12 000 et 60 000 exomes d'individus de diverses origines ethniques. Un variant identifié à faible fréquence constitue parfois une difficulté importante d'interprétation des données d'exome selon la pathologie étudiée, dans la mesure où les cohortes composant ces grands agrégats d'exomes ne sont pas constituées par des individus sains. Des

Variant type	Mean number of variants (\pm sd) in African Americans	Mean number of variants (\pm sd) in European Americans
Novel variants		
Missense	303 (\pm 32)	192 (\pm 21)
Nonsense	5 (\pm 2)	5 (\pm 2)
Synonymous	209 (\pm 26)	109 (\pm 16)
Splice	2 (\pm 1)	2 (\pm 1)
Total	520 (\pm 53)	307 (\pm 33)
Non-novel variants		
Missense	10,828 (\pm 342)	9,319 (\pm 233)
Nonsense	98 (\pm 8)	89 (\pm 6)
Synonymous	12,567 (\pm 416)	10,536 (\pm 280)
Splice	36 (\pm 4)	32 (\pm 3)
Total	23,529 (\pm 751)	19,976 (\pm 505)
Total variants		
Missense	11,131 (\pm 364)	9,511 (\pm 244)
Nonsense	103 (\pm 8)	93 (\pm 6)
Synonymous	12,776 (\pm 434)	10,645 (\pm 286)
Splice	38 (\pm 5)	34 (\pm 4)
Total	24,049 (\pm 791)	20,283 (\pm 523)

Tableau 4: Nombre moyen de variants codants dans 2 populations. Données issues 100 échantillons AfroAméricains et 100 échantillons caucasiens. Les variants “non nouveaux” sont présents dans dbSNPv131 (d’après Bamshad et al, 2011).

Consortium/Cohort	Samples
1000 Genomes	1,851
Bulgarian Trios	461
GoT2D	2,502
Inflammatory Bowel Disease	1,675
Myocardial Infarction Genetics Consortium	14,622
NHLBI-GO Exome Sequencing Project (ESP)	3,936
National Institute of Mental Health (NIMH) Controls	364
SIGMA-T2D	3,845
Sequencing in Suomi (SISu)	948
Swedish Schizophrenia & Bipolar Studies	12,119
T2D-GENES	8,98
Schizophrenia Trios from Taiwan	1,505
The Cancer Genome Atlas (TCGA)	7,601
Tourette Syndrome Association International Consortium for Genomics (TSAICG)	297
Total	60,706

Tableau 5: cohortes constituant les 60 000 exomes de la base de données EXAC. Les cas et les témoins ont été inclus, les données phénotypiques ne sont pas disponibles en fonction des génotypes identifiés dans un ou des gène(s) donné(s) (Source EXAC).

pathologies comme les troubles psychiatriques, l'infarctus du myocarde, et les maladies inflammatoires de l'intestin y sont, entre autres, représentées (tableau 5).

4.3.1.2 Prédiction de pathogénicité

Cependant, même après élimination des variants silencieux et fréquents ($MAF > 1\%$), le séquençage d'exome chez un individu peut permettre d'identifier entre 100 et 500 variants rares_ et potentiellement causaux_ selon le design de l'étude, son niveau de consanguinité et son origine ethnique (Ng et al., 2010a; Pierce et al., 2010; Vissers et al., 2010). Ce résultat a mené à la nécessité immédiate d'utiliser un second niveau de filtres dans l'exploitation des résultats. La prédiction de pathogénicité des variants est ce second niveau de filtre communément utilisé.

On peut distinguer dans cette étape de filtrage deux types de pathogénicité prédite : d'une part les variants induisant une perte de fonction certaine ou très probable du gène (mutations stop, décalant le cadre de lecture, ou de sites canoniques d'épissage), que l'on appellera dans un souci de simplification « mutations perte de fonction » d'autre part les variants faux-sens dont la pathogénicité est prédite sur des bases plus complexes.

Considérer les variations perte de fonction peut paraître la solution la plus rationnelle dans une analyse de séquence d'exome. Cette approche a été couronnée de succès dans les études portant sur l'autisme, ou un excès significatif de mutations de ce type a été retrouvé chez les enfants atteints dans de multiples études familiales (Lossifov et al., 2014; Lim et al., 2013; Toma et al., 2014). Elle reste malgré tout d'une efficacité relative : il été démontré que chaque individu porte en moyenne 100 mutations perte de fonction, et 20 gènes totalement inactivés (Rivas et al., 2015), ce qui démontre le manque de spécificité d'un filtre simplement basé sur la perte de fonction de la protéine.

Les variants faux-sens sont globalement 30 à 60 fois plus nombreux que les précédents (tableau 4) et les filtres utilisés portent sur de multiples paramètres tels la conservation de la position et des propriétés physicochimiques de l'acide aminé sous-jacent ou la structure de la protéine ou du motif correspondant (tableau 6). Leur efficacité de prédiction, tous outils confondus serait de 75 à 80% (Hicks et al., 2011), mais ces données sont par définition biaisés par l'état des connaissances actuelles : la plupart en effet utilisent des jeux de données existantes de mutations comme HGMD (Human Mutation Database) pour valider la démarche de prédiction, données ayant été elles-mêmes souvent obtenues après sélection puis validation des variants sur des bases de prédiction de pathogénicité. En résumé, ces outils ne trouvent que ce qu'ils cherchent et ne cherchent que ce qu'ils savent déjà trouver. A titre d'exemple, lors de l'analyse historique du séquençage d'exome dans le syndrome de Miller, le gène responsable de la pathologie a été éliminé dans un premier temps car l'un des variants identifiés était prédit comme « bénin » par Polyphen (Ng et al., 2010a). Par ailleurs,

NAME	TYPE OF ANALYSIS	WEB LINK
AlignGVGD	Conservation of physico-chemical properties	agvgd.iarc.fr
Condel	Prediction method based on combining other methods	bg.upf.edu/condel
MAPP	Conservation of physico-chemical properties	mendel.stanford.edu/sidowlab/downloads/MAPP
MutationTaster	Bayes classifier over multiple sequence features and conservation	mutationtaster.org
PMut	Evolutionary and structural features combined using a machine learning method	mmb2.pcb.ub.es:8080/PMut
PolyPhen-2	Evolutionary and structural features combined using naïve Bayes classifier	genetics.bwh.harvard.edu/pph2
SIFT	Evolutionary method based on position-specific scoring matrix	sift-dna.org
SNAP	Several evolutionary and structural features combined using a neural network	www.rostlab.org/services/snap
SNPs3D	Combination of a phylogenetic and a structural method. Uses support vector machine	www.snps3d.org

Tableau 6: Quelques outils de prédiction de pathogénicité de variants faux-sens. Rares sont les outils qui effectuent des prédictions sur d'autres types de variations (indels par exemple) (d'après Sunyaev et al, 2012).

une faible concordance des prédictions bio-informatiques a été montrée à partir de données d'exome. Dans l'étude de Tennessen et coll. 50% des SNP étaient prédits délétères par 1 outil pris isolément, mais seuls 1% des variants étaient concordants pour l'ensemble des 7 outils évalués (Tennessen et al., 2012). Les méthodes sous-jacentes à ces outils bio-informatiques sont malgré tout relativement redondantes et aucune étude rétrospective de grande ampleur n'a été utilisée pour évaluer la meilleure combinaison; la plus utilisée dans les études actuelles associe les logiciels SIFT et Polyphen.

D'une façon générale, le filtrage constitue aujourd'hui l'étape limitante de toute analyse d'exome, et la dernière partie de mon travail de thèse (chapitre 4, Etude génétique de la Dysplasie Fibromusculaire) illustre bien les difficultés rencontrées pour limiter efficacement le nombre de gènes candidats. Devant le besoin croissant de restreindre le nombre de variants et de gènes après ces étapes de filtrage « basique », divers outils plus sophistiqués de filtres ont récemment été développés pour prioriser ces gènes candidats et sont décrits dans la section 4.4.3 Séquençage d'exome et outils de priorisation des gènes candidats.

4.3.2 Optimisation du design de l'étude selon le modèle génétique

La priorisation initiale des variants par les filtres sus-cités constitue donc une étape nécessaire mais non suffisante pour l'identification des mutations causales. Il s'agit alors d'exploiter au maximum le contexte et les données cliniques des patients afin de choisir la stratégie d'analyse la plus appropriée de séquençage d'exome. Cette stratégie est dépendante du mode de transmission supposé de la pathologie et s'appuie sur la ségrégation attendue des variants correspondants (figure 9).

4.3.2.1 Stratégie de ségrégation familiale

Dans le cas d'une maladie de transmission mendélienne, plusieurs membres atteints de la famille peuvent être séquencés afin d'identifier les variants génétiques partagés. Les membres de la famille non atteints peuvent être séquencés afin d'exclure les variants privés bénins. En sélectionnant les apparentés atteints les plus éloignés, le nombre de variants bénins partagés peut être considérablement réduit. Cette approche a été utilisée avec succès par Ng et coll. qui ont séquencé deux membres atteints d'une fratrie, réduisant le nombre de candidats récessifs à seulement 9 gènes (Ng et al., 2010a). Krawitz et coll. ont appliqué une version plus élaborée de cette approche en utilisant des données de séquençage de l'exome seulement. Les auteurs ont déterminé les haplotypes pour tous les variants partagés par trois membres atteints d'une fratrie, puis n'ont

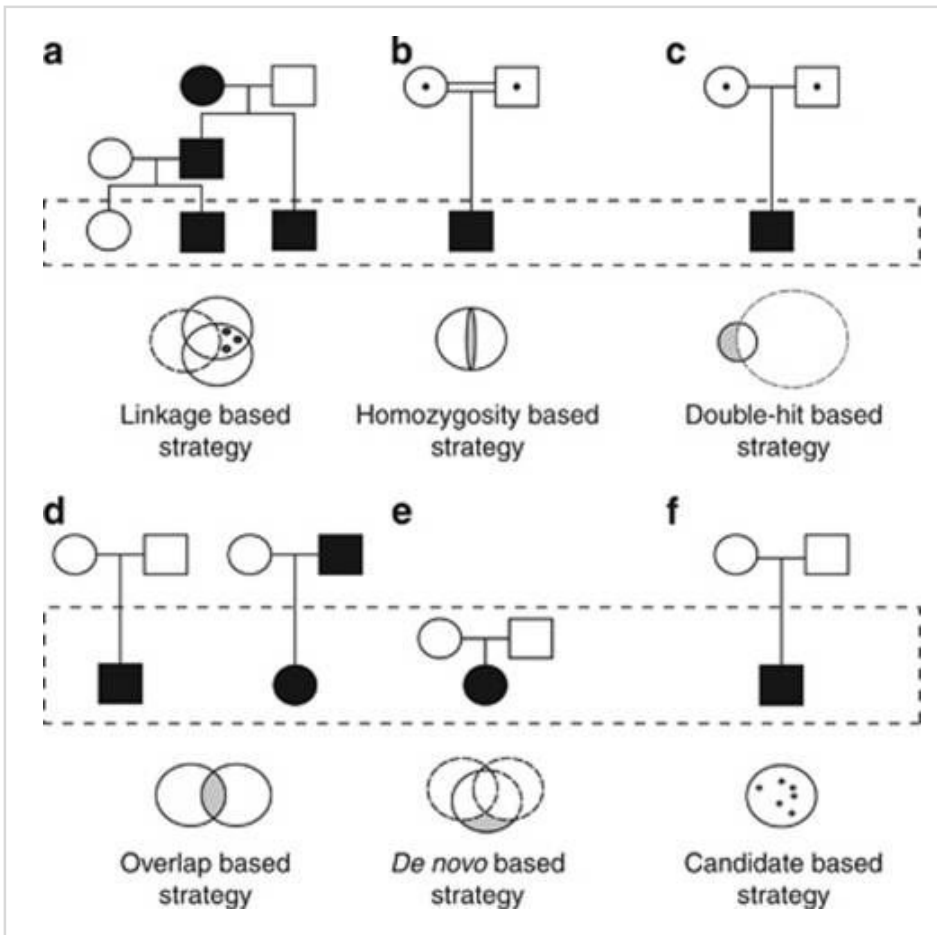
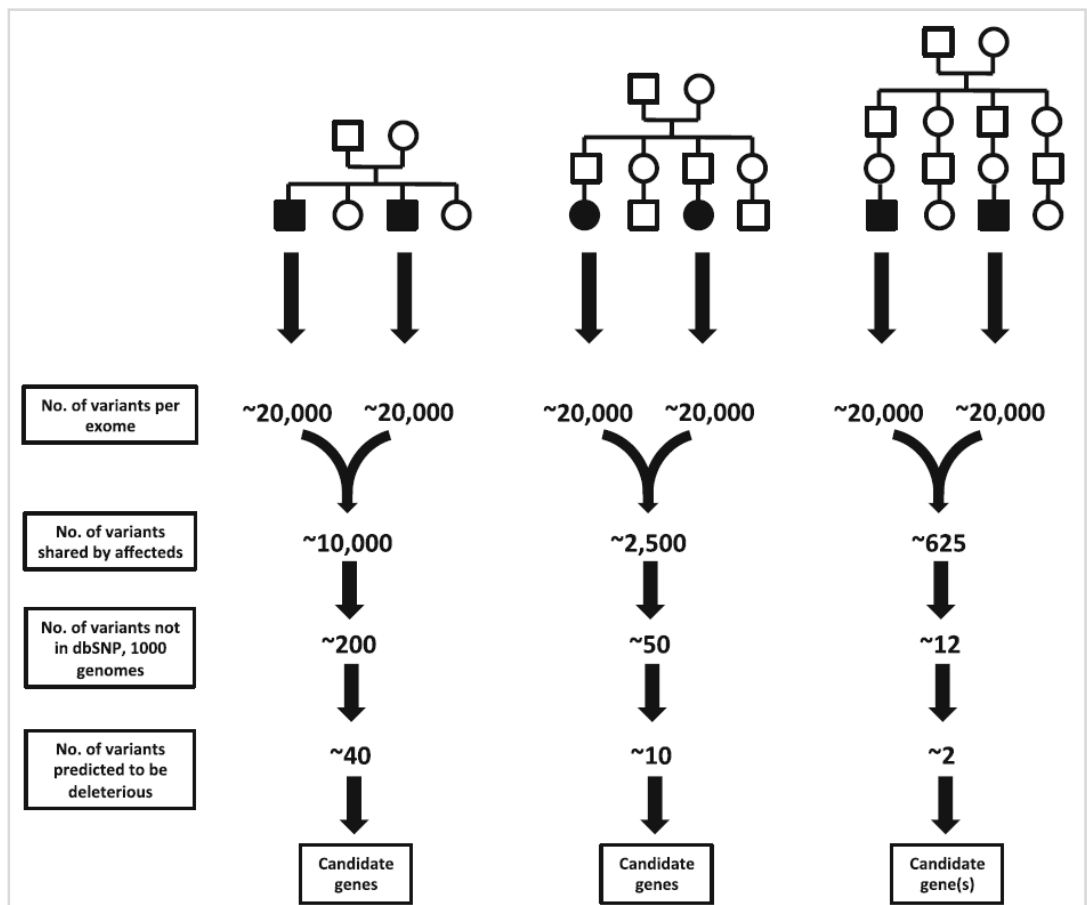


Figure 9: Stratégies de séquençage d'exome pour l'identification de gènes dans les maladies mendéliennes. Les encadrés en pointillé incluent les individus concernés par le séquençage d'exome. Les cercles sous chaque pedigree symbolisent les groupes de variants identifiés. Les zones grises représentent les variants potentiellement pathogènes, présents chez les atteints et absents chez les apparentés sains (d'après Gilissen et al, 2012).

Figure 10: Exemple théorique de résultats de séquençage d'exome pour diminuer le nombre de variants identifiés par analyse de ségrégation (cas d'une maladie dominante, d'après Brunham et al, 2013).



conservé que les variants rares et identiques par descendance. De cette façon, ils ont réduit le nombre de candidats de 14 à seulement 2 gènes (Krawitz et al., 2010). Cette stratégie est bien sûr d'autant plus efficace que les apparentés atteints sont éloignés : selon les estimations réalisés par Brunham et coll. (Figure 10), à partir du séquençage d'exome de 2 individus, le nombre de candidats peut être réduit d'un facteur 4 entre une analyse de fratrie et de cousins germains, et d'un facteur 20 si l'on analyse une paire de cousins issus de germains.

4.3.2.2 Stratégie d'homozygotie

Dans le cas d'une maladie récessive avec suspicion de consanguinité, l'hypothèse initiale est la même que dans la cartographie d'homozygotie : le variant homozygote doit siéger dans une région homozygote de grande taille, conservant essentiellement les variants identiques par descendance chez l'individu atteint. Ces régions peuvent être identifiées au préalable par génotypage SNP, mais Becker et coll. (Becker et al., 2011) ont récemment montré que les données de l'exome seul sont suffisamment informatives pour réaliser une cartographie d'homozygotie fiable. Dans cette étude, 17 des 318 variants non synonymes privés observés chez le cas-index étaient homozygotes, mais seulement trois étaient situés dans les grandes régions homozygotes et la mutation causale était située dans la plus grande de ces trois régions. Cette approche semble réduire suffisamment le nombre de variants pour permettre l'identification d'un gène de maladie chez un patient sans ajout d'apparentés. Elle ne peut toutefois pas détecter les loci homozygotes dans les régions avec peu de SNP informatifs.

4.3.2.3 Stratégie du double-hit

Quand un seul patient d'une famille est disponible et que le mode de transmission suspecté est récessif (sans suspicion de consanguinité), il est possible de séquencer seulement l'exome d'un patient et ne sélectionner que les gènes porteurs de variants homozygotes ou hétérozygotes composites (bien que dans ce dernier cas on ne puisse pas déterminer la phase). Gilissen et coll. ont utilisé cette stratégie pour réduire le nombre de gènes candidats chez deux patients atteints du syndrome de Sensenbrenner, respectivement de 139 et 158 à seulement 3 et 4 (Gilissen et al., 2010). Pierce et coll. ont utilisé la même approche pour un patient atteint du syndrome de Perrault, identifiant un seul gène parmi les 207 variants privés (Pierce et al., 2010). Les deux

études démontrent la puissance de cette approche qui permet d'identifier la cause génétique d'une maladie récessive à partir du séquençage d'un seul individu.

4.3.2.4 Stratégie de croisement ou de chevauchement

En l'absence d'hétérogénéité génétique, il est cohérent de rechercher les variants présents dans un seul gène chez plusieurs patients non apparentés présentant un phénotype similaire. Cette approche, dite de chevauchement, est particulièrement indiquée dans les pathologies dominantes, car les variants hétérozygotes simples sont beaucoup plus nombreux que les variants homozygotes ou hétérozygotes composites prédits pathogènes sur le même gène. Le nombre de gènes présentant des mutations chez plusieurs patients atteints diminuera d'autant plus rapidement que le nombre de patients non apparentés est important.

La première étude utilisant cette stratégie avec succès a porté sur quatre individus non apparentés porteurs du syndrome de Schinzel-Giedon (Hoischen et al., 2010). En 2012, Gilissen et coll. ont démontré l'applicabilité de cette stratégie à partir de seulement trois patients porteurs du syndrome de Bohring-Opitz. Chacun des patients était porteur de 130 à 222 variants privés, Le croisement des données des patients deux à deux permettait de réduire le nombre de gènes à une vingtaine, et à un seul après croisement avec le troisième patient (Gilissen et al., 2012). Cette stratégie n'est cependant applicable que dans les pathologies à forte homogénéité génétique comme nous le verrons dans le chapitre 4.

Par ailleurs la couverture incomplète des données générées par séquençage d'exome inhérentes aux étapes d'enrichissement est une autre limite de cette approche, mais aussi de toutes les approches de séquençage d'exome. Cette problématique est illustrée par l'étude de Ng et coll (Ng et al., 2010b) dans le syndrome de Kabuki : parmi les 10 cas-index séquencés dans cette étude, aucun gène candidat convaincant n'a été identifié chez l'ensemble des patients. Une analyse moins stringente a permis d'identifier des mutations dans le gène *MLL2* chez 7 de ces patients. Le Séquençage Sanger de la région codante de ce gène chez les trois cas négatifs a révélé que deux des patients étaient porteurs de mutations de type indels non détectées par le séquençage d'exome.

L'hétérogénéité génétique du syndrome de Kabuki a en outre été mise en évidence dans la cohorte de réplication utilisée dans l'étude de Hannibal et coll. (Hannibal et al., 2011) dans laquelle seulement 75% des cas étaient porteurs d'une mutation du gène *MLL2*.

Cette différence de proportions est évocatrice d'une sélection des cas les plus caractéristiques de la pathologie dans la première étude, soulignant l'importance de phénotypage clinique précis et cohérent pour la réussite d'une stratégie de chevauchement.

4.3.2.5 Stratégie *de novo*, ou analyse de trios

Dans les maladies sporadiques, et tout particulièrement dans le cas de maladies sévères responsables d'un phénotype fœtal ou infantile développemental, l'hypothèse d'une anomalie *de novo* peut être émise. Ces mutations *de novo* peuvent être identifiées de façon simple par l'étude de l'exome de trios parents-enfant ou parents-fœtus. Les variants candidats sont identifiés après élimination de tous les variants hérités. Cette stratégie très efficace aboutit à un nombre limité de gènes potentiels, un exome contenant en moyenne seulement 3 à 4 anomalies *de novo* (O'Roak et al., 2014; Roach et al., 2010; Vissers et al., 2010). Vissers et coll. ont utilisé cette stratégie en 2010 dans la déficience intellectuelle qui leur a permis d'identifier des anomalies *de novo* dans des gènes neuro-développementaux chez 7 des 10 patients séquencés. Des résultats similaires ont été obtenus dans l'autisme et la schizophrénie (O'Roak et al., 2014; Xu et al., 2011).

Cette stratégie a aussi été couronnée de succès dans les maladies cardiovasculaires : Dans une grande étude par séquençage d'exome de 41 familles atteintes du syndrome de Gordon (ou hypertension hypokaliémique), une première étape a permis d'établir la causalité du gène *KLHL3* dans des formes récessives et dominantes de la maladie. En seconde étape, la réanalyse des données de séquençage d'exome a permis d'identifier 4 mutations *de novo* du gène *CUL3* chez 4 cas sporadiques de cette cohorte. Le travail réalisé par cette équipe a permis, en combinant diverses approches et modèles génétiques, d'identifier deux nouveaux gènes en une seule stratégie d'analyse (Boyden et al., 2012).

L'ensemble de ces résultats souligne le rôle probablement majeur des anomalies *de novo* dans les maladies sporadiques, tout au moins dans les domaines cardiovasculaires, et neuropsychiatriques.

Cependant la mise en évidence d'une neomutation ne constitue pas en soi une preuve formelle de pathogénicité, et des études fonctionnelles ou de réplication restent nécessaires pour affirmer la pathogénicité des variations identifiées.

4.3.2.6 Stratégie par gène candidat

Dans le cas d'un seul individu atteint et en l'absence d'apparentés disponibles les options sont limitées, et les stratégies d'analyse reposent uniquement sur le filtrage comme vu précédemment, et la priorisation (voir section 4.4.3).

Ces différentes stratégies peuvent être combinées et elles le sont très généralement pour une optimisation des résultats : un des premiers exemples est l'étude de Munsurusu et coll. qui a ainsi associé recherche de double hit et ségrégation familiale et identifié un gène d'hypolipidémie familiale combinée (Musunuru et al., 2010a).

4.3.3 Succès et limites du séquençage d'exome

D'une façon générale, le séquençage d'exome a entraîné une accélération très importante de l'identification de gènes de maladies mendéliennes (figure 2), et s'est révélé d'une plus grande efficacité dans les maladies récessives que dominantes y compris dans les cas de mutations hétérozygotes composites (Bamshad et al., 2011).

Comme précédemment cité, sous certaines conditions le nombre d'individus minimum permettant d'identifier un gène causal peut être faible, jusqu'à 3 (Hoischen et al., 2011) en cas de stratégie de chevauchement ou jusqu'à un seul cas pour une stratégie de double hit. Les maladies récessives ou liées à une mutation *de novo* sont de toute évidence les plus faciles à identifier, et sont donc sur-représentées dans les exemples de réussite de la méthode. Les échecs de séquençage d'exome n'étant jamais (ou presque) publiés, il est difficile de connaître la puissance réelle de ces approches. A partir des données du Centre for Mendelian Genomics (University of Washington Center for Mendelian Genomics, Yale University Center for Mendelian Genomics, Baylor-John Hopkins Center for Mendelian Genomics) ayant réalisé un effort majeur de séquençage d'exome dans diverses maladies mendéliennes (18 800 échantillons, 8 800 familles, 1 050 maladies mendéliennes dont 580 de cause connue et 470 maladies de cause inconnue), Chong et coll. ont révélé que seuls 30% des cas ont pu être élucidés par cette approche. De façon intéressante, sur les plus de 300 gènes identifiés par cette grande étude, 40% correspondaient à un nouveau gène de maladie, 5% à un effet pléiotrope (gène identifié dans une maladie clairement distincte) et plus de 50% correspondait à une expansion d'un phénotype déjà lié à un gène et une maladie connus (tableau 7, Chong et al., 2015). Des résultats très similaires ont été obtenus dans 100 pathologies mendéliennes par la plate-forme FORGE (Finding of Rare diseases Genes Canada).

SUMMARY OF DISCOVERIES OF GENES UNDERLYING MENDELIAN PHENOTYPES				
Discovery Type		Evidence of Causality		
		Conservative	Suggestive	Total
Known	known gene; explained, known phenotype	320	19	339
	phenotype expansion	174	24	198
Novel	known gene; unexplained, known phenotype	4	0	4
	known gene; new phenotype	17	7	24
	novel gene; unexplained, known phenotype	25	27	52
	novel gene; new phenotype	107	232	339
Total novel		327	290	617
Total number of discoveries		647	309	956

Tableau 7: Distribution des résultats de séquençage d'exome réalisé dans le Centre for Mendelian Genomics (d'après Chong et al, 2015).

Ces résultats à grande échelle issus de grands programmes de séquençage d'exome soulignent nettement le taux relativement faible de réussite de la technique. Il est probable que les échecs de séquençage correspondent à des cas de maladies dominantes très rares ou très hétérogènes sur le plan génétique.

4.4 Les stratégies combinées au séquençage d'exome en génétique Mendélienne

L'efficacité du séquençage d'exome peut être fortement améliorée en utilisant des stratégies combinées.

Dans le cadre des maladies Mendéliennes, ces stratégies sont d'une part inspirées des méthodes existantes de cartographie (analyse de liaison simple, cartographie d'homozygotie). D'autre part, Elles utilisent des outils de plus en plus sophistiqués de priorisation des gènes candidats à partir des listes de candidats obtenues par les filtres simples.

4.4.1 Séquençage d'exome et cartographie génétique

Dans le cadre des maladies dominantes, après réalisation d'une analyse de liaison classique, le séquençage d'exome est un substitut rapide et efficace au séquençage Sanger. L'objectif est d'obtenir l'ensemble des séquences codantes des gènes inclus dans les intervalles de liaison définis par un Lod-Score supérieur à 3. L'approche la plus efficace est alors de réaliser le séquençage d'exome chez deux individus dont la taille de la ou des régions génomiques correspondant au chevauchement des haplotypes est la plus faible. Ceci permet de réduire au maximum le nombre de gènes et de variants candidats.

Cette stratégie a été appliquée dans les maladies dominantes et récessives liées à l'X, ou elle a démontré son efficacité (tableau 8). A titre d'exemple, Sobreira et coll. dans l'étude génétique de la métachondromatose ont réduit le nombre de candidats de 109 (par approche exome seul) à 1 seul gène présent dans l'intervalle de liaison (Sobreira et al., 2010). De même, dans une nouvelle forme de diarrhée dominante familiale, la réduction de la région génomique analysée à l'intervalle lié au trait (soit 2,9 Mb) a révélé une seule variation candidate hétérozygote dans le gène *GUCY2C*, confirmé par la suite par une étude fonctionnelle (Fiskerstrand et al., 2012). Dans les maladies récessives avec consanguinité, l'efficacité d'une stratégie combinée est encore plus importante : les intervalles peuvent être réduits de façon drastique par la conservation des régions homozygotes par descendance à partir de peu de cas, comme discuté dans la

Disorder	PMID**	Mode of inheritance	N*	Strategy	Gene(s)
Comparison of unrelated cases					
Kabuki	20711175	AD	10	10 cases / 10 kindreds	<i>MLL2</i>
Schinzel-Giedion	20436468	AD	4	4 cases / 4 kindreds	<i>SETBP1</i>
Hadju-Cheney	21378985	AD	3	3 cases / 3 kindreds	<i>NOTCH2</i>
Fowler	20518025	AR	2	2 cases / 2 kindreds	<i>FLVCR2</i>
Sensenbrenner	20817137	AR	2	2 cases / 2 kindreds	<i>WDR35</i>
Comparison of related cases					
Miller	19915526	AR	4	4 cases / 3 kindreds	<i>DHODH</i>
hyperphosphatasia-MR	20802478	AR	3	3 cases / 1 kindred	<i>PIGV</i>
hypolipidemia	20942659	AR	2	2 cases / 1 kindred	<i>ANGPTL3</i>
retinitis pigmentosa	21295283	AR	3	3 cases / 1 kindred	<i>DHDDS</i>
novel skeletal dysplasia	21455487	AR	4	2 cases + parents / 1 kindred	<i>POP1</i>
spinocerebellar ataxia	21106500	AD	4	linkage + 4 cases / 1 kindred	<i>TGM6</i>
familial ALS	21145000	AD	2	linkage + 2 cases / 1 kindred	<i>VCP</i>
dilated cardiomyopathy	21353195	AD		linkage + 4 cases / 1 kindred	<i>BAG3</i>
hidradenitis suppurativa	21430701	AD	3	linkage + 2 cases ⁹ / 1 kindred	<i>NCSTN</i>
spinocerebellar ataxia	21106500	AD	4	linkage + 4 cases / 1 kindred	<i>TGM6</i>
primary failure tooth eruption	21404329	AD	4	linkage + 4 cases ⁴ / 1 kindred	<i>PTH1R</i>
TARP ¹	20451169	XLR	2	linkage + 2 cases / 2 kindreds	<i>RBM10</i>
X-linked leukoencephalopathy	21415082	XLR	2	linkage + 1 case ⁹ / 1 kindred	<i>MCT8</i>
Homozygosity mapping					
DFNB82 (hearing loss)	20602914	AR	1	1 case / 1 kindred	<i>GPSM2</i>
CNS malformations	20729831	AR	2	2 cases / 1 kindred	<i>WDR62</i>
Seckel	21131973	AR	1	1 case / 1 kindred	<i>CEP152</i>
NPHP-related ciliopathy ²	20835237	AR	1	1 case / 1 kindred	<i>SDCCAG8</i>
autoimmune LPS	21109225	AR	1	1 case / 1 kindred	<i>FADD</i>
3MC	21035106	AR	1	1 case / 1 kindred	<i>MASP1</i>
complex I deficiency	21057504	AR	1	1 case / 1 kindred	<i>ACAD9</i>
non-syndromic MR	21212097	AR	2	2 obligate carrier parents	<i>TECR</i>
Ochoa	21450525	AR	1	1 case / 1 kindred	<i>HPSE2</i>
Identification of <i>de novo</i> mutations					
sporadic MR	21076407	complex	30	10 parent-child trios	multiple
autism		complex	60	20 parent-child trios	multiple

Tableau 8: Liste des pathologies identifiées par séquençage d'exome jusqu'en 2011. De nombreuses stratégies combinent séquençage d'exome et analyses traditionnelles (d'après Bamshad et al, 2011).

section 4.3.2.2 Stratégie d'homozygotie. Dans le syndrome de Malpuech-Michels-Mingarelli-Carnevale, la cartographie d'homozygotie sur 3 cas dans 2 familles a identifié une région de 1,81 Mb incluant 20 gènes. Le séquençage d'exome d'un seul de ces trois cas a révélé une mutation faux-sens dans le gène MASP1, prédite délétère par les outils traditionnels précédemment cités (Sirmaci et al., 2010).

4.4.2 Séquençage d'exome et tests d'association

Dans le cas de maladies dominantes dont l'hétérogénéité génétique est certaine ou suspectée, les tests d'association « two-samples » (Bamshad et al., 2011) entre exomes de cas-index et de témoins permettent de déterminer a posteriori la puissance attendue d'un projet de séquençage d'exome. L'objectif de ces tests est de déterminer pour les gènes candidats partagés par plusieurs cas-index si cette récurrence est significative ou non. En effet, identifier de façon récurrente un gène chez plus d'un cas-index dans ce contexte peut être lié à la grande taille du gène, à son faible niveau de conservation (inversement proportionnel à sa variabilité) ou bien à des artefacts de séquençage. Ce test développé par Bamshad et coll. peut être réalisé en utilisant comme contrôles des données publiques d'exome en respectant les règles de stratification pour l'origine ethnique.

4.4.3 Séquençage d'exome et outils de priorisation des gènes candidats

Le séquençage d'exome a augmenté massivement le nombre de gènes candidats pour un phénotype donné. Il est probable que le fort taux d'échecs rencontré dans cette approche provienne de l'incapacité à discriminer parmi la liste résiduelle des gènes, le ou les candidats en cause. Cette incapacité peut résulter d'une absence de données sur la fonction ou l'expression du gène, cas relativement courant puisque seuls 14 000 des 19 000 gènes de protéines environ sont décrits dans OMIM (tableau 1), mais pas uniquement.

L'information génétique disponible pour une majorité de gènes a fortement augmenté en raison (i) d'une sophistication et d'une systématisation des méthodes d'analyse par génie génétique (comme The International Knockout Mouse Consortium), (ii) de la centralisation de l'information génétique dans les bases de données génomiques publiques (telles que Ensembl, NCBI, et UCSC), les bases de données d'expression génique (telles que GEO) et les bases de données de variation humaines (telles que HapMap), (iii) des données portant sur les réseaux et les voies de régulation génétiques de plus en plus fournies (par exemple, KEGG et

Reactome). La littérature biomédicale est devenue par ailleurs une ressource trop massive pour être assimilée par les scientifiques (par exemple, plus de 22 millions de résumés sont répertoriés par PubMed en 2014, dont 15 millions traitant de l'homme). Face à cette avalanche de données, à la recherche d'arguments pour un gène candidat, Le défi consiste à utiliser de façon rationnelle et efficace l'information disponible, tâche impossible à entreprendre manuellement. De nombreuses méthodes d'évaluation de gènes candidats se sont développées, basées sur les caractéristiques des gènes sélectionnés (taille, structure, motifs), la similitude de séquence avec des gènes de maladies connues, et l'annotation fonctionnelle des gènes candidats. Une autre catégorie d'outils a récemment vu le jour dans ce domaine et consiste à utiliser l'information issue de bases de données de phénotypes (humains ou murins) dans la sélection des candidats.

4.4.3.1 Outils basés sur la fonction, les propriétés intrinsèques des gènes, et leur similitude avec les gènes de maladies

Par l'analyse des propriétés intrinsèques des gènes déjà associés aux maladies héréditaires (indépendamment du phénotype), des spécificités peuvent être trouvées dans les gènes responsables de maladies. Ces propriétés comprennent la structure du gène, sa taille ou la séquence de ses régions régulatrices, la conservation ou la présence de paralogues (source potentielle de redondance fonctionnelle). La méthode la plus utilisée est PROSPECTR (Adie et al., 2005) qui permet à partir d'une liste de gènes, d'éviter la sélection de gènes à faible potentiel délétère (un exemple simple est l'élimination de gènes de récepteurs olfactifs par cet outil). L'efficacité de ces outils devient aujourd'hui limitée avec l'augmentation des informations associées à un grand nombre de gènes.

Plusieurs méthodes proposent une priorisation sur la base d'une similitude de structure et de fonction avec des gènes déjà connus dans une maladie proche (ou la même maladie dans les cas d'hétérogénéité génétique). Cette approche logique ne peut parfois pas être réalisée manuellement devant la masse des informations sur les maladies connues, et son automatisation représente un gain de temps substantiel. L'usage de ces outils permet de plus un meilleur contrôle des requêtes effectuées car ils utilisent un vocabulaire formel, comme les termes MeSH ou GO (Gene Ontology). Des méthodes telles que G2D (Perez-Iratxeta et al., 2002), et ENDEAVOUR (Tranchevent et al., 2008) utilisent cette approche.

L'efficacité de cette approche est conditionnée par l'existence de maladies ou phénotypes similaires de cause génétique connue, ce qui n'est pas toujours le cas. Une autre limite de ces outils est le faible pourcentage de gènes humains associés à des données validées expérimentalement dans les bases de données interrogées : seuls 6% ont des liens vers la littérature non-génomique et plus de 70% des annotations correspondent à des

fonctions extrapolées à partir de domaines protéiques prédits (Perez-Iratxeta et al., 2002) ce qui augmente considérablement le risque d'erreur d'interprétation des résultats et diminue le rapport « signal-bruit ».

Certains outils comme Panther (Miet et Thomas, 2009), ou DAVID (Huang et al., 2009) permettent de déterminer, parmi une liste de gènes, un enrichissement dans les trois grandes classes d'éléments de la classification ontologique des gènes (ou GO) que sont les processus biologiques, la fonction moléculaire et les compartiments cellulaires. Ils ont été initialement développés pour l'analyse de données transcriptomiques ou d'immunoprécipitation de la chromatine à grande échelle (Chip-Chip ou Chip-Seq) mais sont parfois appliquées au séquençage d'exome.

Une approche complémentaire consiste à identifier les interactions protéiques et les voies ou cascades dans lesquelles sont impliqués les gènes candidats. L'hypothèse sous-jacente évidente est que plusieurs partenaires peuvent être impliqués dans une même maladie ou un même groupe de maladies. C'est une constatation universelle en génétique Mendélienne, qui se vérifie dans le domaine des maladies vasculaires rares, impliquant notamment la voie du TGF β pour la maladie de Marfan et les syndromes Marfanoides, ou la voie PI3K dans les malformations vasculaires acquises en mosaïque (syndrome de Protée et apparentés). STRING (Jensen et al., 2009) établit ainsi un réseau de gènes en effectuant des recherches d'interactions protéine-protéine, de co-occurrence dans des voies par l'utilisation de KEGG, de co-expression, et de co-citation dans la littérature scientifique. Il est possible avec cet outil d'ajouter des partenaires supplémentaires à partir du pool de gènes de la base interrogée. Cela permet d'identifier des gènes interagissant de façon indirecte par l'intermédiaire d'un partenaire commun. De nombreux autres outils que STRING existent, comme Ingenuity Pathway Analysis (outil commercial développé par Quiagen) ou GeneMania (Warde-Farley et al., 2010). Des méthodes telles qu'ENDEAVOUR, précédemment cité, et les versions récentes de G2D (Perez-Iratxeta et al., 2007) combinent cette recherche d'information avec une interrogation des termes GO.

Ces deux dernières approches (enrichissement ontologique et réseaux de gènes) s'avèrent intéressantes dans les analyses d'exomes de cohortes de taille suffisante, avec un nombre de gènes candidats importants dans des maladies hétérogènes ou oligogéniques voire polygéniques. Elles permettent en effet d'identifier dans une liste de gènes les partenaires potentiels (qu'une revue de la littérature ne suffit pas à identifier manuellement) constituant une hypothèse physiopathologique de départ.

4.4.3.2 Outils basés sur les similitudes phénotypiques

L'analyse phénotypique « *in silico* » permet d'évaluer la pertinence de chaque gène candidat en fonction du phénotype des patients. Ces outils utilisent l'ontologie phénotypique humaine présente dans la base de données HPO (Groza et al., 2015). Cette base est un catalogue des anomalies cliniques observées dans la maladie humaine. De même, les phénotypes murins sont interrogés par divers outils comme Exomizer, et sa version plus récente Phive (Robinson et al., 2014). La performance de ce dernier outil très prometteur a été évaluée par une stratégie de simulation : des mutations pathogènes issues la base de données Human Gene Mutation (contenant 28 516 mutations pathogènes et environ 1 000 gènes de maladies, (Stenson et al., 2009) ont été tirées au hasard et ajoutées aux données de 1 000 exomes issus du consortium 1000 génomes. Le jeu de données final comportait 100 000 exomes « mutés ». Après utilisation des filtres classiques de fréquence allélique et paramétrage du mode de transmission attendu, 83% et 66% des « bons gènes » étaient classés au somme de la liste de candidats pour les maladies récessives et dominantes respectivement. Un exemple de concordance phénotypique réalisée par Phive est présenté dans la figure 11.

Compte tenu de la taille grandissante des catalogues de phénotypes murins soutenus par divers consortiums internationaux, la recherche de concordance avec les phénotypes murins peut s'avérer intéressante. C'est particulièrement le cas lorsqu'aucune maladie humaine proche n'est élucidée sur le plan génétique et que la physiopathologie n'est pas connue : les analyses de concordance phénotypique utilisant le catalogue HPO, les recherches de similitudes avec des gènes de maladies proches, et les analyses de réseaux de gènes et de voies de signalisation ne peuvent pas être utilisés.

D'un point de vue général, les algorithmes nécessaires à l'intégration des ensembles de données phénotypiques et génétiques à grande échelle adaptés à l'analyse d'exome en sont encore à leurs débuts, et les preuves de leur efficacité restreintes : une étude récente menée dans l'autisme a identifié des variants dans des partenaires directs de gènes de l'autisme par enrichissement GO et analyse de voies de signalisation, mais le modèle génétique sous-jacent rend difficile la confirmation fonctionnelle de ces résultats (Toma et al., 2014). L'utilisation d'Endeavour à partir de données d'exome a permis d'identifier le gène causal d'une paraparésie spastique récessive (SPG30) à partir d'une liste restreinte de 15 gènes candidats (Erich et al., 2011) mais les exemples de réussite de l'utilisation de ces approches restent peu nombreux.

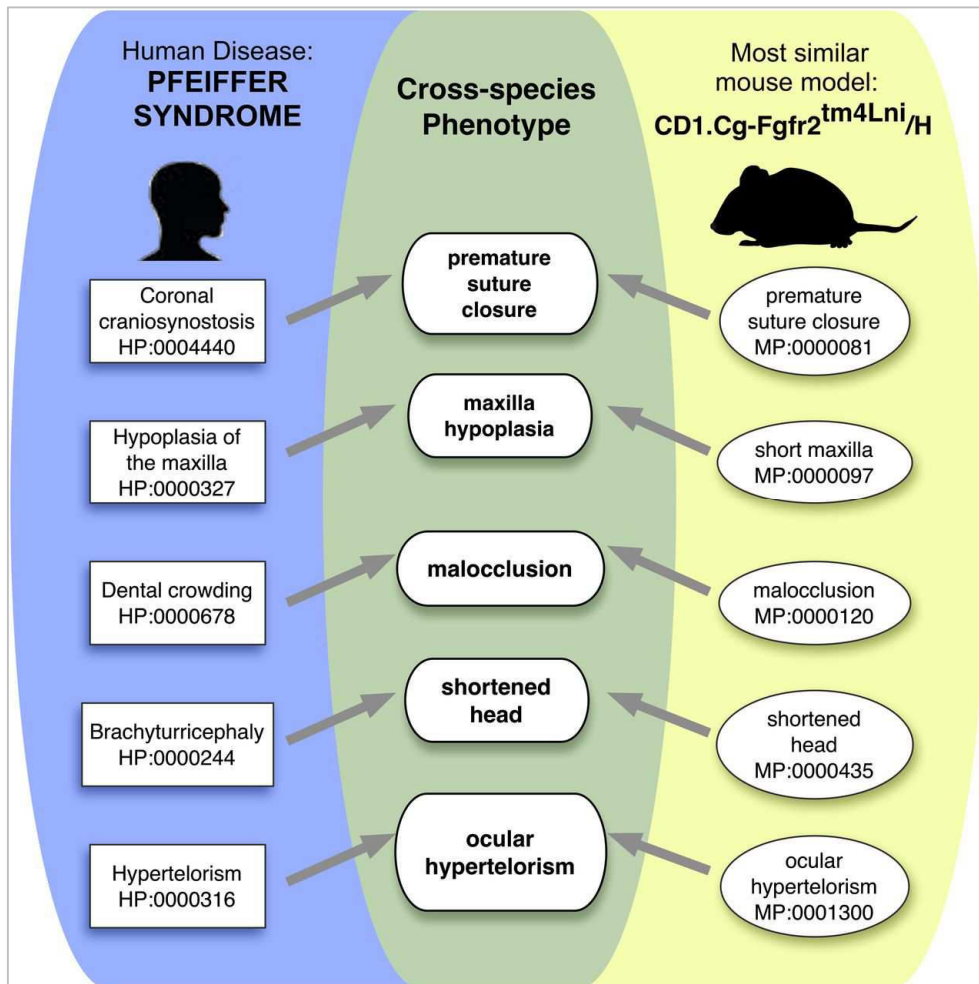


Figure 11: Exemple d'algorithme de similitudes phénotypiques dans la maladie de Pfeiffer (d'après Robinson et al, 2014).

4.5 Stratégies combinées en maladies communes : la recherche de variants rares

Comme vu dans la section 3.2.3, le GWAS a démontré ses limites en terme d'impact des SNP fréquents et des gènes qui y sont liés dans l'héritabilité des maladies communes, menant au concept d'héritabilité manquante sous-tendu par l'hypothèse « maladie commune-variant rare ».

Une idée simple serait de séquencer l'exome (ou le génome) de tous les cas et de tous les témoins, et de réaliser une étude d'association à partir de ces données. Cette approche sera probablement développée dans les années à venir, mais les coûts actuels sont encore prohibitifs pour de très grandes séries de sujets et les tests statistiques actuels ne sont pas adaptés au volume de données correspondantes : en effet dans une approche WGS (séquençage de Génome ou Whole Genome Sequencing), les corrections pour tests multiples nécessairement appliquées génèreraient une perte de puissance majeure, avec pour conséquence la nécessité d'inclure des dizaines de milliers de cas _ voire plus pour certaines maladies complexes _ pour atteindre les seuils de significativité requis. Compte-tenu des coûts de tels projets et de l'inadéquation des outils statistiques, ces approches ne sont pas encore utilisées en génétique multifactorielle.

De plus, les analyses statistiques portant sur les variants rares diffèrent de celles développées dans le champ des GWAS traditionnels (voir sous-section 4.5.4 Etudes d'association pour les variants rares), et leur développement récent nécessite probablement une évaluation rétrospective.

Comme dans la génétique des maladies mendéliennes, la combinaison de diverses stratégies s'avère l'approche la plus efficace. On peut dénombrer quatre types d'approches utilisées.

4.5.1 GWAS et séquençage d'exome dans les phénotypes extrêmes

Une stratégie complémentaire d'identification d'allèles rares à risque est le "séquençage des extrêmes" (figure 12). Le prototype de ce type d'étude a été réalisé par Cohen et coll. (Cohen et al., 2004) par le séquençage 3 gènes candidats dans l'hypoHDLémie (*ABCA1*, *APOA1* et *LCAT*) chez des individus dont le taux d'HDL était respectivement au 95^{ème} et 5^{ème} percentile de la population. Cette étude, réalisée en séquençage Sanger classique, a révélé un excès de variants rares chez les cas hypoHDLémiques (20 variants contre 3 chez les cas hyper HDLémiques), essentiellement retrouvés dans le gène *ABCA1*.

De façon similaire, le reséquençage de 4 gènes associés par GWAS à l'hypertriglycéridémie a démontré dans une cohorte de patients hypertriglycéridémiques (forme majeure) un excès de variants rares : 28% étaient porteurs de variants rares contre 15% des contrôles (Johansen et al., 2012).

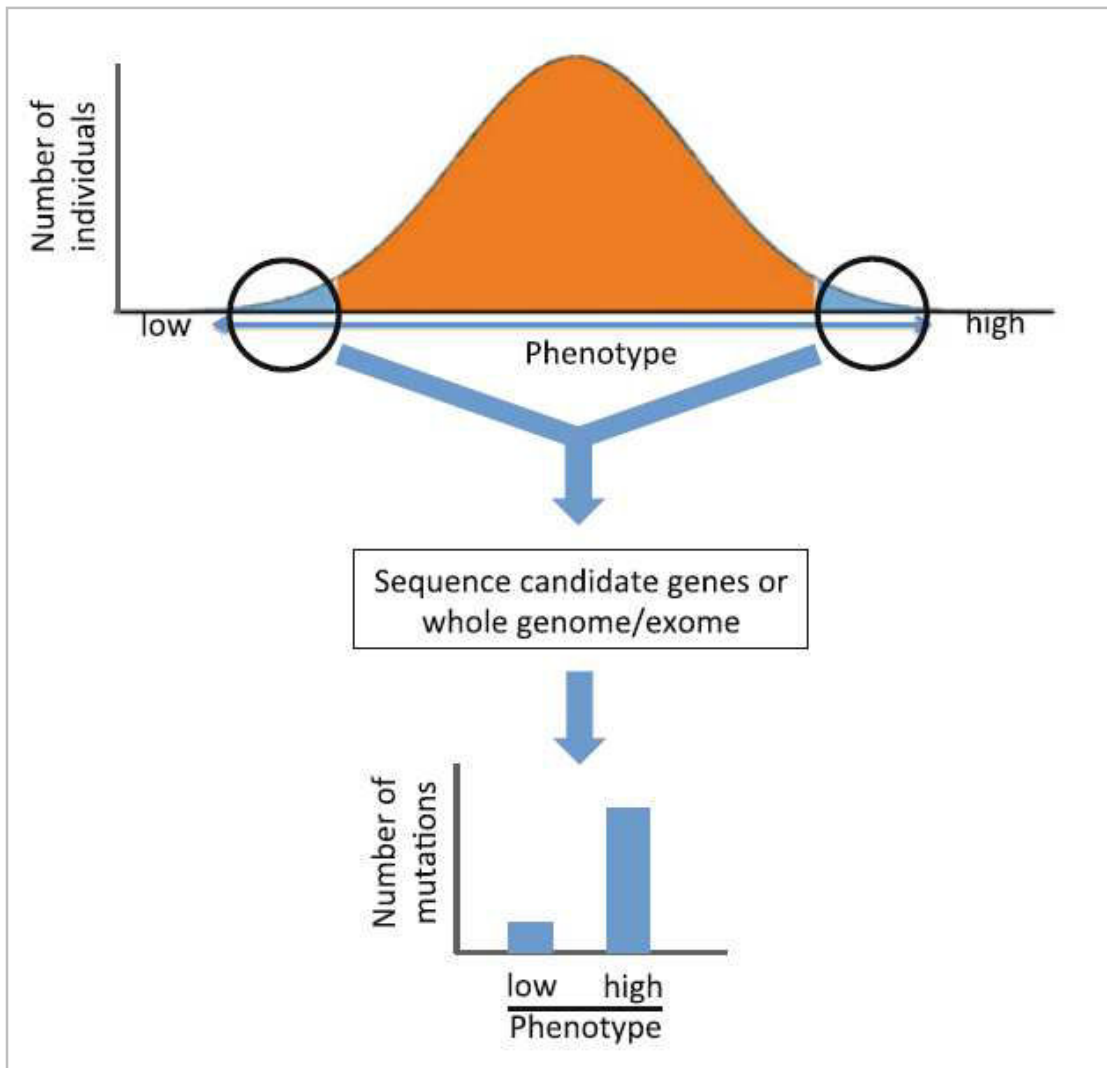


Figure 12: Séquencage des extrêmes. Un excès de variants rares dans un gène dans un des deux sous-groupes de phénotype extrême est un argument pour son implication dans la pathologie étudiée (d'après Brunham et al, 2013).

Cette dernière étude, ainsi que de nombreuses autres, démontre qu'une partie de l'héritabilité dans les maladies communes est liée à des variants rares présents dans les gènes identifiés précédemment par GWAS, ce qui a conforté l'idée générale que variants rares et communs influençant un même trait peuvent être identifiés dans les mêmes gènes.

Le séquençage des extrêmes a émergé comme un outil puissant démontrant l'implication de variants rares dans les gènes influençant les traits complexes, et cette approche a récemment été appliquée par séquençage d'exome à la recherche d'un facteur génétique de susceptibilité à l'infection par *Pseudomonas aeruginosa* dans la mucoviscidose (Emond et al., 2012). Cette susceptibilité est très hétérogène au sein des patients, et l'étude s'est portée sur 2 catégories de patients : ceux infectés dans le 10^{ème} percentile de la distribution d'âge à l'infection (atteinte précoce), et ceux non porteurs du germe après l'âge de 14 ans. Le séquençage de l'exome de ces individus a identifié un gène unique, *DCTN4*, statistiquement associé avec l'âge d'infection au germe. Cette étude princeps démontre la puissance et l'efficacité de ce type d'analyse dans un modèle proche des maladies communes, mais nécessite des approches statistiques dédiées afin d'asseoir la significativité des variants rares identifiés (développées dans la section 4.5.4).

4.5.2 GWAS et séquençage d'exome dans les isolats géographiques

Une étude portant sur le Sick Sinus Syndrome (SSS) dans la population islandaise illustre cette approche. L'analyse par GWAS chez environ 800 cas et 40 000 témoins a identifié un locus en 14q11 contenant 3 SNP associés au SSS. Le séquençage de génome réalisé chez 87 cas porteurs du SNP, a identifié un variant codant rare dans le gène *MYH6*, associé au SSS avec un odds-ratio de 12,5 (p -value $1,5 \cdot 10^{-29}$, (Holm et al., 2011). La fréquence allélique de ce variant codant dans cette population est relativement élevée (évaluée grâce à cette étude à 0,4%) du fait de la structure de la population islandaise, constituant un isolat géographique. Plus rare dans la population caucasienne non islandaise, il n'est pas représenté sur les puces traditionnellement utilisées pour les GWAS, malgré son odds-ratio élevé.

4.5.3 Séquençage d'exome dans les formes familiales rares de maladies communes

Le séquençage d'exome dans les formes familiales de maladies communes semble a priori une approche séduisante, s'apparentant aux méthodologies développées précédemment. Ce type d'analyse est cependant plus délicat que dans les maladies purement mendéliennes, les formes familiales de maladies communes étant caractérisées par une grande hétérogénéité génétique. A titre d'exemple dans le domaine des maladies cardiovasculaires, à ce jour 13 gènes et 4 loci ont été identifiés dans la fibrillation auriculaire familiale, 7 gènes

dans l'hypertension pulmonaire, et 7 gènes dans la maladie coronaire, (OMIM Aout 2015, Padmanabhan et al., 2010). L'identification des gènes responsables de formes rares est cependant un défi important sur le plan scientifique et médical, car elle permet de mieux comprendre les bases physiopathologies des maladies complexes et de dégager des pistes thérapeutiques : c'est le cas des maladies coronariennes où l'identification des gènes du métabolisme du cholestérol a permis de proposer les traitements par statines, inhibiteurs de l'HMG-CoA réductase, enzyme-clé de production du cholestérol.

L'application du séquençage d'exome dans ce domaine va probablement accélérer ce processus, et des exemples de réussite ont déjà été publiés : Le gène *KCNK3* a été identifié dans une famille par séquençage d'exome, puis confirmé chez 2 des 92 cas familiaux et 3 des 230 cas sporadiques analysés par séquençage Sanger (Ma et al., 2013). Ce succès souligne cependant la grande hétérogénéité génétique de ce type d'affection, et l'existence de probables facteurs modificateurs présents dans la maladie (du fait de l'identification de mutations dans un groupe de cas sporadiques). Ces données rendent compte des difficultés à venir dans l'identification des autres gènes impliqués dans cette maladie, et les maladies cardiovasculaires communes d'une façon générale.

Dans le domaine des maladies neurodéveloppementales, les études de patients atteints d'autisme révèlent l'hétérogénéité génétique extrême de cette affection, mais dans ce cas, le modèle émergent de la maladie semble être oligogénique : il impliquerait l'accumulation de mutations rares *de novo* et de mutations héréditaires perte de fonction qui contribuent au risque génétique globale chez un patient particulier (O'Roak et al., 2014; Toma et al., 2014, voir section 4.3.1.2 Prédiction de pathogénicité).

4.5.4 Etudes d'association pour les variants rares

L'identification de variants rares (<0,5%) ou très rares (<0,05%) dans de grandes cohortes de patients atteints d'une maladie commune, qu'elle soit réalisée par séquençage de génome (comme dans le SSS), d'exome (comme dans le phénotype sensible à *P. aeruginosa* dans la mucoviscidose) ou par simple séquençage Sanger doit faire l'objet d'un traitement statistique spécifique pour confirmer une association avec le trait étudié. Cette spécificité est inhérente à la rareté des événements traités, qui diminue la puissance statistique des tests effectués.

De nombreuses méthodes statistiques ont été proposées pour augmenter le rapport « signal-bruit » dans les tests d'association issus de données de séquençage ciblé. Elles sont en théorie applicables aux données de séquençage de génome ou d'exome entier.

Les tests unidirectionnels

Les premières méthodes connues sous le nom de « burden-test » comme par exemple, le test CAST (Morgenthaler et Thilly, 2007), consistent à agréger tous les variants rares (« collapsing methods ») dans une région génomique en une seule variable. Ces approches ne tiennent pas compte de l'effet potentiel de variants communs, ni généralement de la fréquence allélique de chacun des variants rares identifiés, à l'exception des tests de « somme pondérée » (WSTs, pour Weighted Sum Tests). Les approches type WST utilisent différents systèmes de pondération, mais en général, ils surestiment le poids des variants rares et inversement pour les variants fréquents. L'approche « seuil variable » (VT, Price et al., 2010) ne fixe pas un seuil de rareté, mais évalue une série de seuils et conserve le seuil maximisant la puissance du test.

Les tests bidirectionnels

Aucune de ces méthodes ne permet d'envisager des effets contradictoires des variants identifiés (protecteur versus délétère). L'approche de « sommation adaptative » (Asum) est la première méthode permettant cette distinction (HanetPan, 2010).

La flexibilité de ces tests a été améliorée par l'utilisation de méthodes basées sur l'apprentissage (ou « machine-learning ») comme par exemple le test SKAT (Sequence Kernel Association test, Wu et al., 2011). Comme pour les WST, SKAT intègre également un système de pondération, qui peut être modifié en fonction de la connaissance préalable de la cause de la maladie. L'importance et le sens des associations sont estimées à partir de données au lieu d'être fixés au préalable. Cette méthode a été appliquée sur des données de séquençage des gènes ANGPTL3, ANGPTL4, et ANGPTL5 chez 3 476 individus de la cohorte « Dallas Heart Study » (Lee et al., 2012b). Une association entre plusieurs des 97 variants de ces gènes et le taux de triglycérides a été identifiée de façon plus significative ($p < 10^{-5}$) en comparaison des tests unidirectionnels ($p < 10^{-4}$ à 10^{-3}).

Plusieurs modèles unifiant les méthodes existantes ont été proposés et permettent d'identifier la méthode la plus efficace, adaptée au cas étudié (Lee et al., 2012b; Tang et Lin, 2014). C'est notamment le cas de SKAT-O, modèle linéaire combinant SKAT et un test type Burden (Lee et al., 2012c). Cette méthode a été appliquée avec succès à l'identification de variants de susceptibilité à un second gène, *CAV2*, dans l'infection à *P.aeruginosa* chez les patients mucoviscidosiques (Emond et al., 2015) en appliquant la stratégie des phénotypes extrêmes : Les données issues de séquençage d'exome chez 85 cas d'infection précoce et 65 cas d'infection tardive, comparées aux données de 3 000 contrôles issus d'ESP (exome sequencing project) ont révélé le rôle protecteur du gène *CAV2*. L'analyse SKAT-O incluait 8 variants, et l'analyse ultérieure variant par variant a confirmé l'implication d'un seul SNP dans cet effet protecteur. Ces résultats ont ensuite été confirmés par l'ajout d'une population contrôle d'environ 700 individus génotypés pour ce SNP.

Enfin, une méthode d'agrégation, regroupant l'exploitation de variants rares et communs, appelée CMC (pour combine Multivariate et Collapsing Method, Li et Leal, 2008) propose une analyse multivariée après séparation en sous-groupe rares et fréquents des variants identifiés. Plus récemment, Ionita-Laza et coll. ont généralisé ce paramétrage à la méthode SKAT (Ionita-Laza et al., 2013).

La diversité de ces méthodes met en lumière leurs limites inhérentes et la nécessité d'une hypothèse préalable : dans le cadre des dyslipidémies, un modèle de type SKAT sera plus approprié qu'un test de charge dans la mesure où l'existence de variants protecteurs et délétères au sein de mêmes gènes est clairement établie.

Aujourd'hui les études de séquençage haut débit dans les maladies communes sont actuellement restreintes aux cas particuliers développés précédemment (phénotypes extrêmes, sous-groupes spécifiques, formes familiales). Les tests statistiques présentés ci-dessus nécessitant de grands jeux de données pour analyser correctement un gène ou un locus. Ils sont de fait encore peu appliqués et leur efficacité peut être compromise par divers paramètres, comme démontré dans une étude réalisée dans l'autisme (Liu et al., 2013).

Objectifs de travail

J'ai réalisé mon travail de thèse dans le laboratoire « Génétique des maladies artérielles rares ». Dans le cadre de ce travail, j'ai étudié trois maladies à composante cardiovasculaire dont les caractéristiques génétiques étaient très diverses. L'objectif de ma thèse était de comprendre les bases génétiques de ces trois pathologies en utilisant une approche de séquençage d'exome.

La première maladie a fait l'objet d'une étude génétique préalable dans mon laboratoire d'accueil, car les patients présentaient un phénotype cardiovasculaire (hypertension artérielle) associé à une maladie hématologique, dont la cause génétique n'était pas connue. Cette maladie appelée stomatocytose héréditaire déshydratée a donc fait l'objet d'un travail de cartographie génétique réalisé avant mon arrivée, à partir d'une grande famille préalablement publiée. Afin d'identifier le gène en cause, j'ai appliqué une méthodologie combinant analyse de liaison génétique et séquençage d'exome. Le travail réalisé dans cette première partie de ma thèse a permis d'identifier la cause génétique de cette affection et d'en comprendre les mécanismes moléculaires.

La seconde pathologie que j'ai étudiée est l'anévrisme de l'aorte abdominale. Cette maladie est commune, mais l'existence de formes mendéliennes rares et pénétrantes est connue depuis longtemps. Cependant aucun gène de forme Mendélienne n'a été identifié à ce jour malgré l'existence de plusieurs loci publiés. Afin de comprendre les bases génétiques de cette affection j'ai utilisé le séquençage d'exome dans une famille présentant une forme Mendélienne de la maladie. Après l'identification d'un gène candidat par cette technique, j'ai développé (i) une approche d'épidémiologie génétique centrée sur ce gène. J'ai analysé les variants fréquents, rares, et très rares, d'une grande cohorte de patients présentant une forme sporadique ou familiale de la maladie, et entrepris une étude d'association de ces variants ; (ii) une étude en modèle cellulaire basée sur la fonctionnalité prédite du gène et de la mutation correspondante.

La troisième partie de ma thèse est consacrée à la dysplasie fibromusculaire artérielle. Cette maladie rare, généralement sporadique, présente environ 10% de formes familiales. Son diagnostic est difficile, car (i) il existe de nombreuses formes infracliniques inaccessibles au diagnostic de certitude, (ii) et de nombreux diagnostics différentiels peuvent mimer la pathologie. J'ai utilisé le séquençage d'exome pour caractériser la génétique des formes familiales et sporadiques de la maladie. A cette fin j'ai analysé l'exome de 8 familles d'au moins deux cas, conjointement avec une étude de trios, et de cas sporadiques, soit 30 individus au total. J'ai donc associé analyse d'exome en formes familiales et sporadiques de la maladie. Cette analyse a été combinée à une exploitation des données et outils bioinformatiques les plus récents de priorisation de gènes candidats afin de dégager les meilleurs candidats de cette affection.

Chapitre 2

Etude génétique de la Stomatocytose Héritaire

Déshydratée

Clinique, épidémiologie et physiopathologie

La Stomatocytose Héritaire Déshydratée (DHS, ou Xérocytose Héritaire) est une maladie hématologique caractérisée par une triade clinique : anémie hémolytique, œdème périnatal, et pseudohyperkaliémie (Grootenboer 2000).

La DHS est transmise sur un mode autosomique dominant, elle est rare mais sa fréquence en population générale est inconnue.

Elle est caractérisée par une déshydratation érythrocytaire primitive : Les érythrocytes présentent une augmentation de la perméabilité membranaire au potassium et au sodium avec fuite majoritaire de potassium (Vives Corrons et al., 1995). Ce désordre ionique est associé à une concentration corpusculaire moyenne en hémoglobine élevée et une fragilité osmotique diminuée, reflet d'une déshydratation cellulaire (résumé par Zarychanski et al., 2012). Une petite proportion des globules rouges ont un aspect de stomatocytes au frottis sanguin (figure 1). Leur durée de vie est réduite *in vivo*, mais la présence d'un tableau hémolytique franc est très rare (voir article joint).

Malgré cette hémolyse modérée, les complications telles la splénomégalie et la lithiase biliaire, conséquences de la rétention splénique des globules rouges et de l'hyperbilirubinémie peuvent survenir, bien qu'elles soient inconstantes. De même, une surcharge en fer et une sidérose hépatique peuvent y être associés (éléments cliniques résumés et discutés dans l'article joint).

Les patients peuvent présenter inconstamment un œdème périnatal et/ou une ascite, sévère et parfois létale mais résolutive dans la première année de vie (Grootenboer et al., 1998; Rees et al., 2004). La physiopathologie de cet œdème n'est pas connue.

La pseudohyperkaliémie est secondaire à la perte de potassium des érythrocytes à température ambiante, phénomène qui ne s'observe pas ou peu à la température corporelle : la fuite de potassium est donc dépendante de la température (résumé par Iolascon et al., 1999), et associée à une résistance plus élevée à la pression osmotique. Cette caractéristique permet la réalisation d'un diagnostic sensible et spécifique par une ektacytomètre (viscosimètre à diffraction laser) qui mesure la déformabilité des globules rouges en fonction de l'osmolalité moyenne.

Les profils de déformabilité osmotique sont de spécificité variable selon la pathologie érythrocytaire étudiée, mais dans la DHS les xérococytes présentent un profil typique et donc spécifique (figure 2).

L'hémolyse peut être précipitée par certains exercices physiques impliquant l'impact répétitif des mains ou des pieds sur une surface rigide comme la marche, le jogging, les percussions, le karaté... (Platt et al., 1981). Ces auteurs ont démontré que les xérocytes ont une sensibilité accrue à l'hémolyse par contrainte de cisaillement, partiellement corrigée in vitro par réhydratation des xérocytes.

Le diagnostic de cette pathologie est important car contrairement aux autres anémies hémolytiques, notamment la stomatocytose hyperhydratée, la splénectomie est contre-indiquée. Elle peut être la cause de complications thromboemboliques étendues (article joint).

1. Génétique de la DHS

Cette maladie présente un degré modéré d'hétérogénéité génétique : Grootenboer et coll. (Grootenboer et al., 2000) ont réalisé une étude de liaison à partir de 10 familles européennes et identifié un seul locus en 16q23-24 avec un Z-score de 3,02. Ce locus était cohérent chez toutes les familles, et couvrait une région génomique de 7 Mb (figure 3).

Cependant un second locus en 2q35 a été identifié dans une grande famille de 20 individus atteints (Carella et al., 1998), plus récemment lié au gène ABCB6 (Andolfo et al., 2013).

Plusieurs familles présentant cette affection étaient étudiées depuis presque 10 ans par analyse de liaison dans mon laboratoire d'accueil. Les travaux réalisés précédemment dans l'équipe ont permis de restreindre l'intervalle génomique (Z-score 4,71) du locus identifié par Grootenboer à environ 5 Mb, contenant plus de 90 gènes (Beurain et al., 2007, figures 4 et 5). Ceci a été rendu possible grâce à la fusion de deux branches familiales dont l'une avait été analysée par Grootenboer en 2000 représentées sur la figure 6. C'est cette famille (dite famille « VA ») qui présentait un phénotype complexe d'hypertension artérielle associé à une anémie avec hémolyse. L'analyse a porté sur 19 membres dont 12 atteints et 6 conjoints non apparentés. Compte tenu de la variabilité d'expression du phénotype dans cette famille et dans la DHS d'une façon générale, le diagnostic a été affirmé par Ektacytométrie.

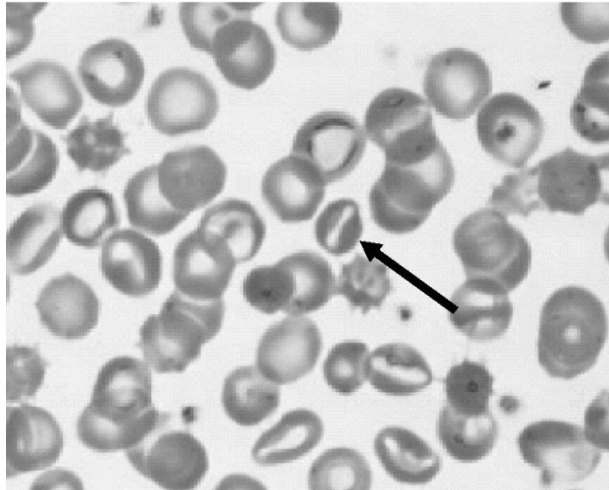


Figure 1: Stomatocytes sur frottis sanguin. d'après Martinaud et al, 2008

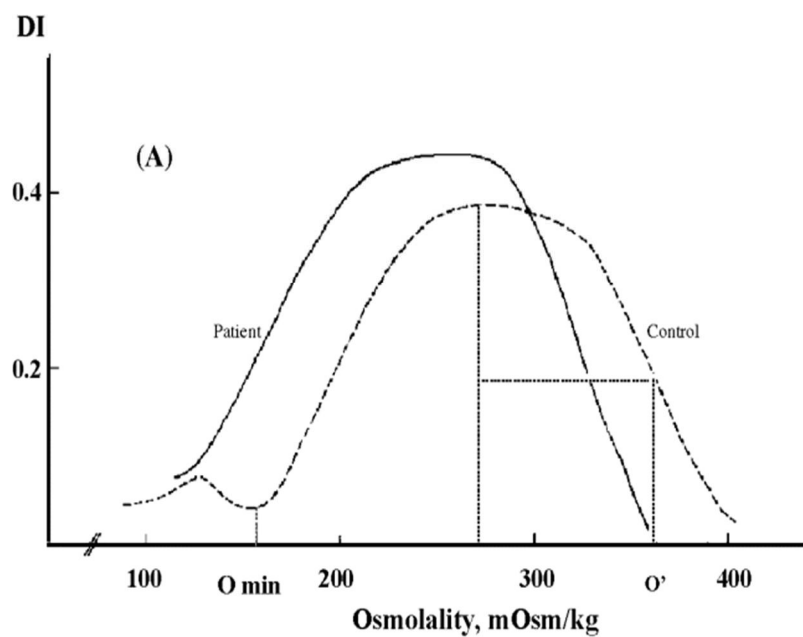


Figure 2: Profil ektacytométrique d'un patient atteint de DHS. En ordonnée la déformabilité érythrocytaire, en abscisse l'Osmolalité. Le déplacement vers la gauche de la courbe en cloche sans modification de son aspect est spécifique de la DHS. d'après Beaurain et al, 2007.

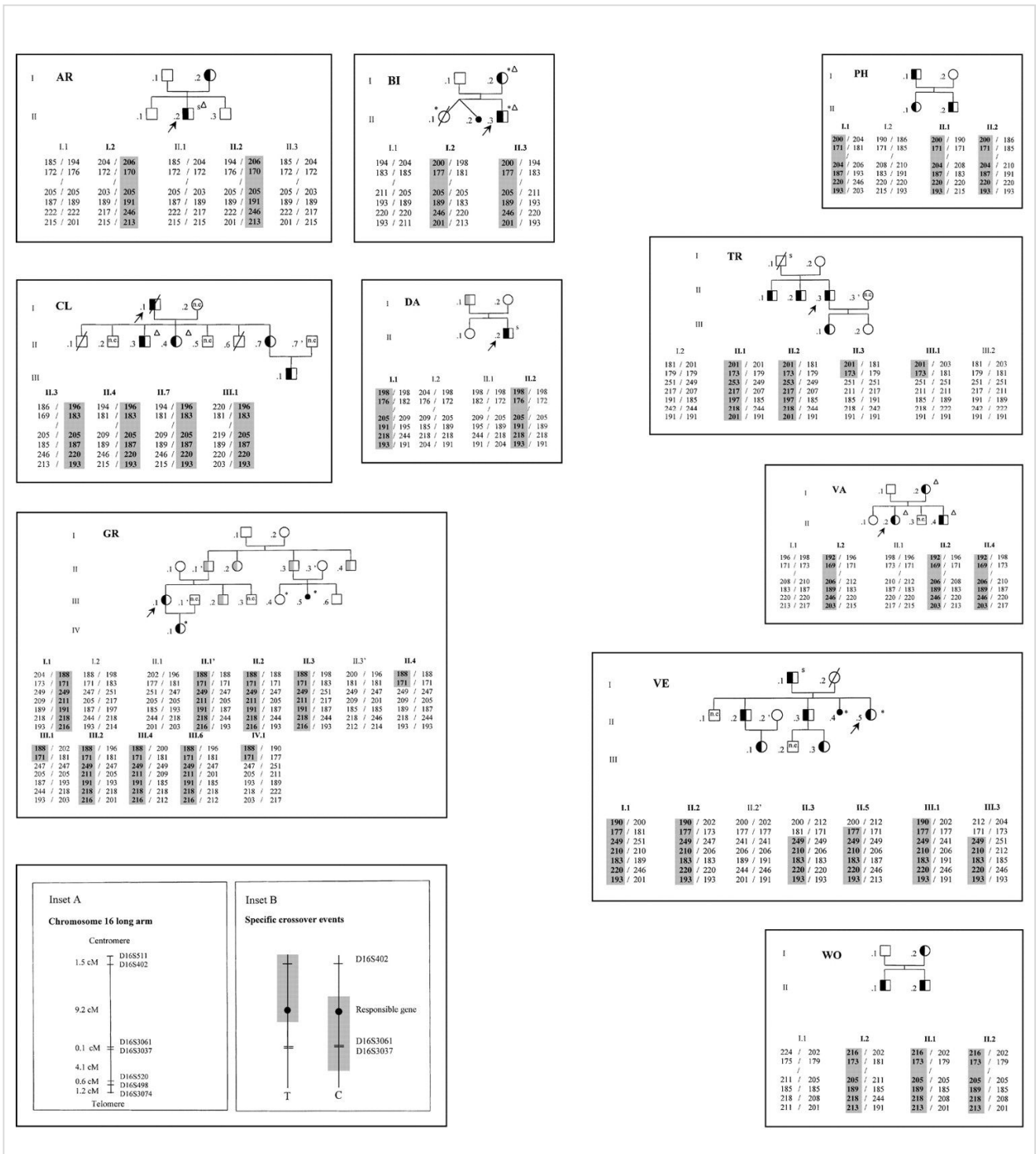


Figure 3: Liaison au locus 16q23-24 dans 10 familles porteuses de DHS. La région génomique fait environ 7 Mb. d'après Grootenboer et al, 2000.

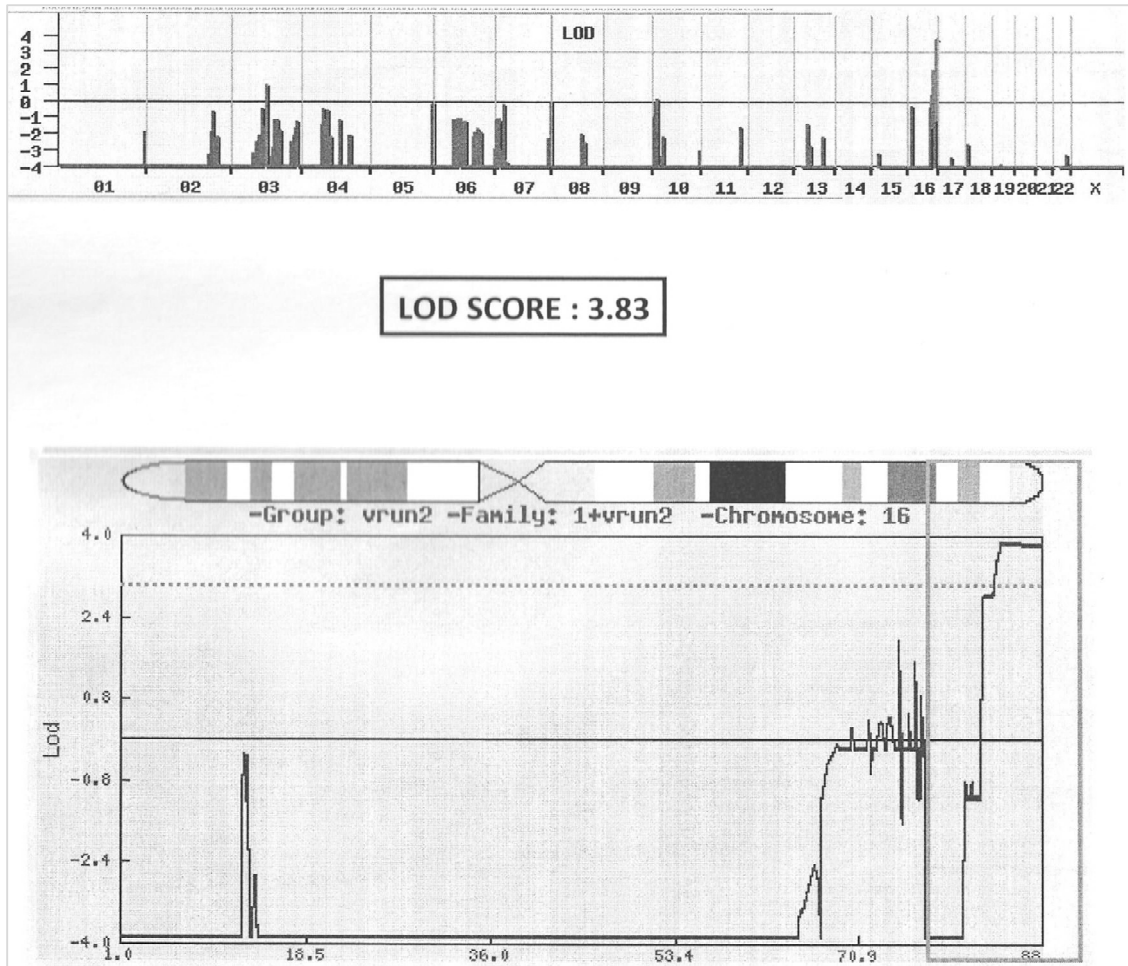


Figure 4: Résultat de l'analyse de liaison dans la famille VA.

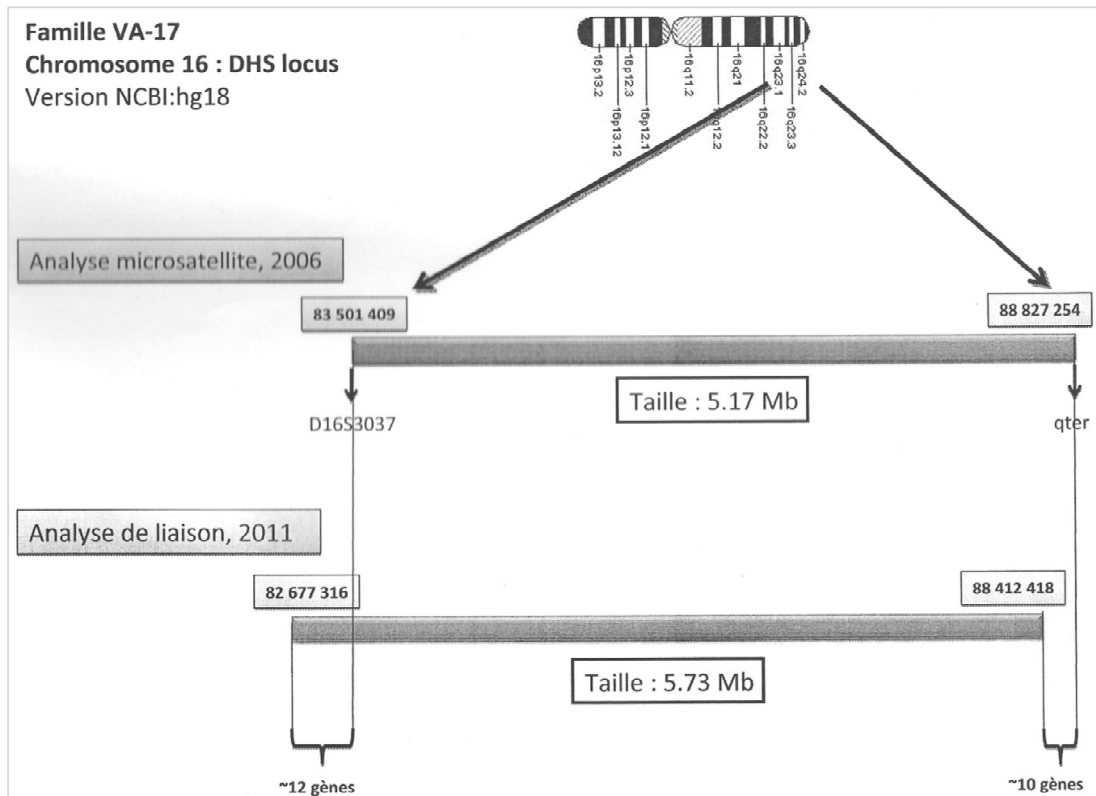


Figure 5: taille de la région génomique au locus 16q24 et liée à la DHS dans la famille VA.

En 2011, Houston et coll. (Houston et al., 2011) ont analysé 137 individus dont 18 atteints appartenant à une grande famille canadienne. L'analyse de liaison a réduit l'intervalle à 1,52 Mb, contenant 51 gènes confirmés ou prédits.

2. Application du séquençage d'exome dans l'identification du gène causal

Le séquençage d'exome des individus II-2 et III-6 de la famille VA a été réalisé (figure 6), l'intervalle de liaison génétique étant similaire entre apparentés plus distants.

L'analyse des résultats a montré qu'environ 4 000 variants étaient partagés entre les deux individus, Après élimination des variants de fréquence supérieure à 1% et des variants silencieux, seuls 5 candidats étaient situés dans l'intervalle de liaison défini par Houston qui comportait en réalité non pas 51 mais 30 gènes de protéines (tableau additionnel 1, article joint).

L'utilisation des outils de prédictions de pathogénicité a retenu uniquement deux candidats : *CDT1* et *PIEZO1*, alors appelé *FAM38A* tous deux exprimés dans les réticulocytes, d'après diverses bases de données en ligne (GEO Profiles, GTex portal, BIOGPS).

La protéine codée par *CDT1* est impliquée dans la formation du complexe de pré-replication de l'ADN, et est un régulateur de la transcription dans la phase G1/S de la mitose. Elle est impliquée dans le syndrome de Meier-Gorlin (nanisme avec microcéphalie de transmission autosomique récessive).

PIEZO1 s'est révélé un candidat idéal puisqu'en 2010 puis en 2012, Coste et coll. ont identifié son implication dans la génération d'un courant cationique, mécanodépendant (Coste et al., 2010, 2012). Cette équipe a caractérisé les protéines PIEZO1 et 2 comme une nouvelle famille de canaux ioniques impliqués dans la mécanotransduction, et fortement conservés au cours de l'évolution puisque leurs travaux ont porté sur la souris et la drosophile. Ce candidat correspondant parfaitement à la physiopathologie attendue concernant la fuite de cations et l'accoutumance au phénotype après stimulus mécanique important (exercice physique, voir section clinique, épidémiologie et physiopathologie).

Une seconde famille avait été identifiée par mon équipe d'accueil. Elle comportait 8 individus atteints et 5 sains. J'ai développé une collaboration avec le service d'Hématologie de l'hôpital Kremlin-Bicêtre (Dr Picard) disposant du seul éktytomètre de France, le service de Médecine Interne de l'Hôpital militaire Saint-Anne à Toulon (Pr de Jaureguiberry), et ai pu recruter une nouvelle famille de 2 cas et 11 cas isolés dont certains étaient publiés (Carli et al., 2007; Martinaud et al., 2008). Le séquençage Sanger des 51 exons codants de

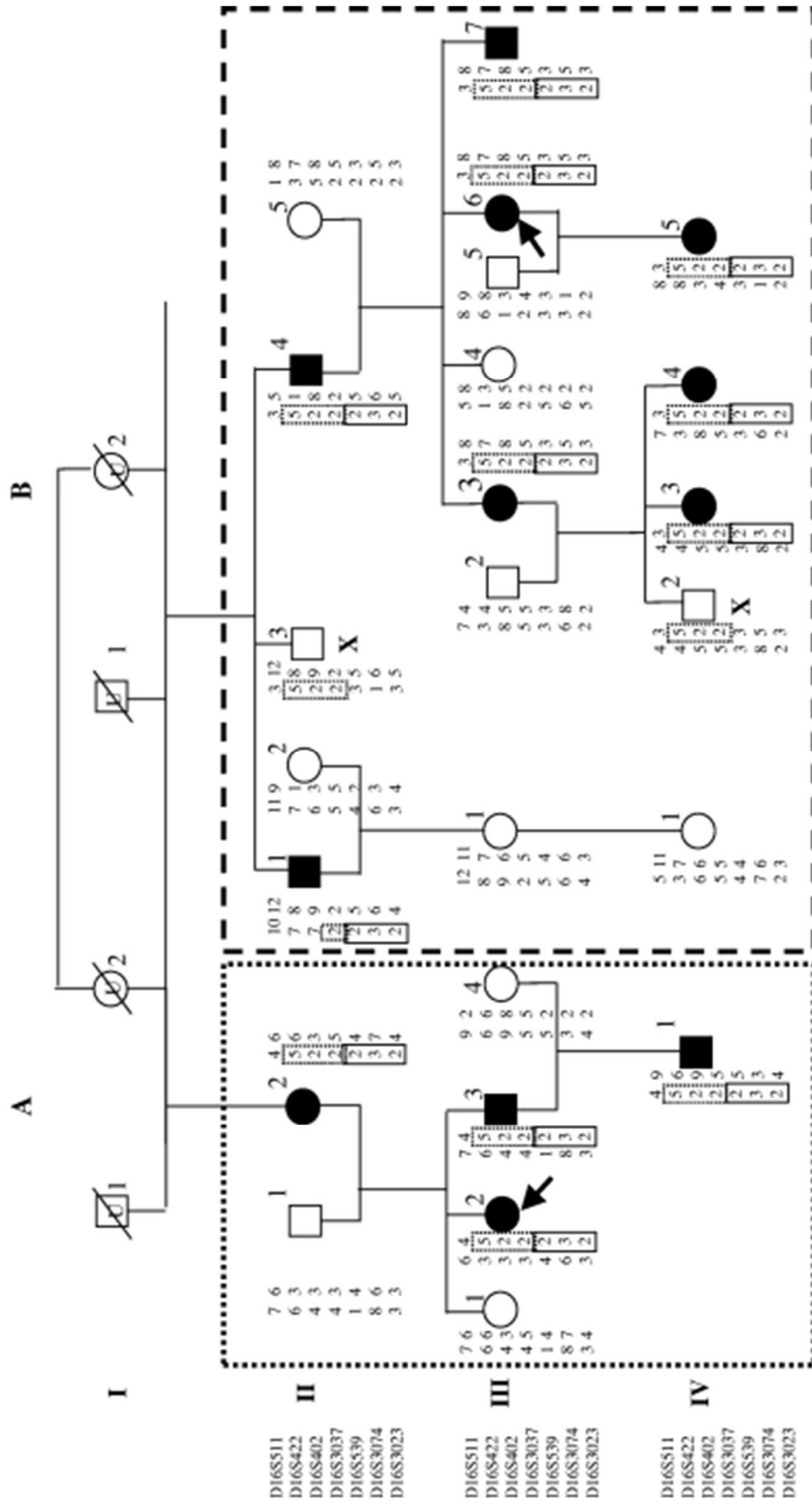


Figure 6: Haplotypes des membres de la famille VA au locus 16q24. La branche de gauche correspond à la famille VA de la figure 3 (publiée par Grootenboer et coll. en 2000).

PIEZO1 a permis d'identifier 3 nouvelles mutations chez les deux familles recrutées et chez 8 cas sporadiques (figure 1 et tableau 1 article joint). Trois de ces mutations étaient des faux-sens prédites délétères, la troisième une insertion en phase de 9 nucléotides.

Ce dernier variant est récurrent (p. E2496ELE) dans les deux familles supplémentaires ainsi que dans 6 cas sporadiques. L'analyse dans une cohorte de 600 contrôles a identifié deux individus porteurs de ce variant, dont l'un présentant une hyperkaliémie. Dans la base de données EVS, ce variant est présent à l'état homozygote ou hétérozygote chez un individu parmi 2 500 caucasiens et hétérozygote chez un individu parmi 1 400 africains.

Il s'agit donc d'un variant rare, et une analyse en comparaison de fréquence simple (test exact de Fisher) entre la population contrôle séquencée (MAF=0,0017) et la cohorte DHS (8/22 cas, MAF=0,36) était très significative ($p < 10^{-76}$) et n'a pas requis d'analyse plus sophistiquée de type Burden ou SKAT.

Nous avons ensuite mis en place une collaboration avec l'équipe américaine ayant publié des travaux d'électrophysiologie sur *PIEZO1* (Coste et al., 2010, 2012) et les quatre mutations identifiées ont été introduites par transfection dans des cellules humaines HEK293T. L'analyse de patch-clamp réalisée a révélé un gain de fonction pour les 4 mutations identifiées, c'est-à-dire une augmentation de l'activité du canal en réponse à un stimulus mécanique (figure 3 article joint), expliquant parfaitement la dysfonction de *PIEZO* dans l'organisme.

Deux mois avant la soumission de notre article à Nature Communications, Zarichansky et coll. (Zarychanski et al., 2012) ont publié 2 mutations de *PIEZO1* dans la famille canadienne précédemment publiée par Houston et dans une famille américaine (Miller et al., 1971) par séquençage d'exome. Le nombre de patients inclus dans l'analyse d'exome est inconnu, et les données fonctionnelles de ces deux variants faux-sens sont relativement pauvres. Nous avons donc inclus ces deux nouveaux variants dans l'étude électrophysiologique et caractérisé un gain de fonction tel qu'attendu par nos premiers résultats (figure 3 article joint).

3. Premières conclusions

Les résultats obtenus dans ce projet montrent bien la puissance de l'association liaison-exome, où le séquençage haut débit vient remplacer le séquençage Sanger traditionnel. A partir d'environ 30 gènes potentiellement impliqués par la liaison, une approche simple sans filtres sophistiqués a permis d'identifier un gène candidat non équivoque. Dans ce travail cependant, les efforts fournis pour l'analyse de liaison étaient

conséquents et le nombre d'individus génotypés importants. De plus, l'existence d'un endophénotype très sensible et spécifique a permis de classer les apparentés de façon formelle et fortement augmenté la puissance de l'analyse génétique. Cette situation « idéale » n'est pas toujours rendue possible selon la taille des familles et le type de phénotype étudié, comme nous le verrons dans le chapitre 4 consacré à la Dysplasie Fibromusculaire.

Enfin, depuis l'identification de *PIEZO1* dans la DHS, de nombreux éléments sont venus éclairer son rôle biologique dans les cellules sanguines et vasculaires. Très récemment, l'équipe de nos collaborateurs a publié une élégante étude topologique de cette grosse protéine à multiples domaines transmembranaires qu'est *PIEZO1*. Leur approche a associé alignement *in silico*, génération de protéines chimère souris/drosophile, spectrométrie de masse, insertion de tag et immunomarquage, et mutagenèse dirigée (Coste et al., 2015). Ce travail de caractérisation topologique a permis d'identifier le pore dans la région c-terminale de *PIEZO1*, zone où s'accumulent les mutations dans la DHS. Par mutagenèse dirigée ils ont identifié certains acides aminés impliqués dans la modulation de la balance cation/anion, et notamment dans l'importance du flux intracellulaire de calcium. Quelques mois plus tard une équipe concurrente a démontré l'implication de *PIEZO1* dans le relargage extracellulaire d'ATP par les érythrocytes (Cinar et al., 2015). Ces auteurs ont démontré que ce relargage était la conséquence du flux intracellulaire de calcium. Ce dernier est secondaire à la stimulation mécanique des hématies générée par le flux sanguin. L'ATP est impliquée dans l'homéostasie vasculaire, et notamment son relargage érythrocytaire, comme le souligne les auteurs, est responsable d'une vasodilatation microvasculaire par activation des récepteurs endothéliaux, puis activation de la voie de l'oxyde nitrique (NO, voir chapitre 4. Section 10.4.3.3.4 Rôle des gènes identifiés par priorisation). *PIEZO1* est donc, de par ses propriétés mécanosensibles, impliquée dans la régulation du signal ATP érythrocytaire.

Enfin, toujours en 2015, *PIEZO1* a été identifié comme un gène de la lymphogénèse : par séquençage d'exome dans une fratrie de patients atteints, Lukacs et coll. ont identifié deux mutations affectant fortement la fonction protéique dans une forme non syndromique de dysplasie lymphatique congénitale : les patients ont une mutation perte de fonction, associée à un faux-sens altérant fortement l'adressage de la protéine à la membrane (Lukacs et al., 2015). Ces mutations ont donc un effet opposé à celles identifiées dans la DHS, et leur phénotype est relativement inattendu puisque l'inactivation du gène est embryolétale chez la souris (Ranade et al., 2014). Il s'agit donc d'un exemple d'effet pléiotrope de *PIEZO1*, bien que l'existence d'un œdème périnatal transitoire chez les patients porteurs d'une DHS pose la question du rôle de ces mutations dans le système lymphatique en développement, qui est à rapprocher du phénotype récemment identifié par l'équipe de Lukacs.

ARTICLE

Received 11 Jan 2013 | Accepted 18 Apr 2013 | Published 21 May 2013

DOI: 10.1038/ncomms2899

Dehydrated hereditary stomatocytosis linked to gain-of-function mutations in mechanically activated PIEZO1 ion channels

Juliette Albuissou^{1,2,3,*}, Swetha E. Murthy^{4,*}, Michael Bandell⁵, Bertrand Coste^{4,†}, H el ene Louis-dit-Picard^{1,2}, Jayanti Mathur⁵, Madeleine F en ant-Thibault⁶, G erard Tertian^{7,8}, Jean-Pierre de Jaureguiberry⁹, Pierre-Yves Syfuss¹⁰, Stuart Cahalan⁴, Loic Gar on¹¹, Fabienne Toutain¹², Pierre Simon Rohrlich¹³, Jean Delaunay⁸, V eronique Picard^{7,8}, Xavier Jeunemaitre^{1,2,3} & Ardem Patapoutian^{4,5}

Dehydrated hereditary stomatocytosis is a genetic condition with defective red blood cell membrane properties that causes an imbalance in intracellular cation concentrations. Recently, two missense mutations in the mechanically activated *PIEZO1* (*FAM38A*) ion channel were associated with dehydrated hereditary stomatocytosis. However, it is not known how these mutations affect PIEZO1 function. Here, by combining linkage analysis and whole-exome sequencing in a large pedigree and Sanger sequencing in two additional kindreds and 11 unrelated dehydrated hereditary stomatocytosis cases, we identify three novel missense mutations and one recurrent duplication in *PIEZO1*, demonstrating that it is the major gene for dehydrated hereditary stomatocytosis. All the dehydrated hereditary stomatocytosis-associated mutations locate at C-terminal half of PIEZO1. Remarkably, we find that all *PIEZO1* mutations give rise to mechanically activated currents that inactivate more slowly than wild-type currents. This gain-of-function PIEZO1 phenotype provides insight that helps to explain the increased permeability of cations in red blood cells of dehydrated hereditary stomatocytosis patients. Our findings also suggest a new role for mechanotransduction in red blood cell biology and pathophysiology.

¹INSERM, UMRS-970, PARCC, Paris 75015, France. ²Universit  Paris Descartes, PRES Sorbonne Paris Cit , Paris 75015, France. ³AP-HP, D partement de G n tique, H pital Europ en Georges Pompidou, Paris 75015, France. ⁴Molecular and Cellular Neuroscience Department, Dorris Neuroscience Center, The Scripps Research Institute, La Jolla, California 92037, USA. ⁵Genomics Institute of the Novartis Research Foundation, San Diego, California 92121, USA. ⁶AP-HP, Service de Biochimie, H pital Bic tre, Le Kremlin Bic tre 94275, France. ⁷Univ Paris-Sud, EA 4531, Chatenay Malabry and Le Kremlin Bic tre 92296, France. ⁸AP-HP, Service d'H matologie et Immunologie, H pital Bic tre, Le Kremlin Bic tre 94275, France. ⁹Service de m decine interne, HIA Sainte-Anne, boulevard Sainte-Anne, 83800 Toulon Naval, France. ¹⁰Service de M decine Interne, Centre Hospitalier, Troyes 10000, France. ¹¹AP-HP, Service d'H matologie et Immunologie biologiques, H pital Saint Antoine, Paris 75012, France. ¹²Service d'H matologie Oncologie P diatrique, Universit  de Rennes-I, CHU de Rennes, Rennes 35033, France. ¹³Service d'H matologie-Oncologie P diatrique, CHU, H pital Jean Minjoz, Besan on 25030, France. [†] Present address: Ion Channels and Sensory Transduction Group, CRN2M, Marseille, France. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to X.J. (email: xavier.jeunemaitre@inserm.fr) or to A.P. (email: ardem@scripps.edu).

Dehydrated hereditary stomatocytosis, (DHS, Xerocytosis, MIM no. 194380) is a dominantly inherited pleiotropic syndrome characterized by increased permeability of red blood cell (RBC) membranes to cations (for review see Delaunay¹). It is typically associated with silent-to-mild haemolysis, pseudohyperkalaemia (that is, loss of K⁺ from red cells on storage at room temperature), and perinatal oedema². A minor proportion of the red cells appear as stomatocytes on blood films. As is frequently observed in chronic haemolytic diseases, complications such as splenomegaly, resulting from increased red cell trapping in the spleen, and cholelithiasis, resulting from elevated bilirubin levels, may occur. For unknown reasons, the course of DHS is frequently associated with iron overload that may lead to hemosiderosis³.

Two genetic loci have been associated with DHS. Linkage to locus 16q23-qter was demonstrated in 1998 (ref. 4), which we further refined to a 11.5-cM interval based on information available on a large three-generation pedigree⁵, an interval confirmed in another large Canadian kindred⁶. Recently, two missense mutations at the *FAM38A* gene lying in this interval and encoding the PIEZO1 protein (M2225R and R2456H) were associated to DHS in two large North-American pedigrees⁷. A second locus for DHS was identified on chromosome 2q35q36 for at least one family^{8,9}. Very recently the Iolascon's group identified two heterozygous missense gain-of-function mutations at a single residue (R375) in the *ABCB6* gene causing familial pseudohyperkalaemia, a related disorder, in three multigenerational families¹⁰. *ABCB6* is a porphyrin transporter present in erythrocyte membranes. Whether *ABCB6* and PIEZO1 act through a shared pathway is unknown.

PIEZOs are newly identified mechanically activated (MA) cation channels. These ion channels are proposed homotetramers, with each subunit encompassing 30–40 transmembrane domains. They are expressed in many cell types including human erythroid progenitor cells^{7,11,12}. Some insights into the roles of PIEZO1 have been obtained recently by knockdown experiments in mammalian epithelial cells from the kidney and the lung, and in lung cancer cells^{13,14}. In the context of kidney epithelial cells, PIEZO1 is proposed to act as a sensor of cell tension. McHugh *et al.* demonstrated that PIEZO1 inactivation promotes cell migration of normal lung epithelial cells, and leads to a switch towards β -integrin-independent amoeboid migration of small lung cancer cells. Evidence for PIEZO function *in vivo* is still limited. In *Drosophila*, Piezo is required for mechanical nociception¹⁵, and as mentioned above, two human *PIEZO1* mutations are associated with DHS patients described by Zacharykanski *et al.*⁷ However, the mechanism by which these mutations alter channel properties and in turn cause the observed phenotype is yet to be elucidated.

Here, we identify three additional missense mutations and one recurrent duplication mutation in *PIEZO1* of DHS patients. Functional characterization of recombinant PIEZO1 channels containing any one of these mutations indicates that they are all gain-of-function mutations that could cause a net increase in ion flux leading to DHS.

Results

Mutation detection. Our genetic analysis had begun when the gene causing DHS was not yet known. We first performed a whole-exome sequencing analysis using two distant affected subjects of our large previously described pedigree⁵ (Fig. 1). Predicted functional coding and unique variants were found at five genes contained in the linkage interval (Supplementary Table S1 and Supplementary Fig. S1). Only two missense variants were not shared by the unaffected subject, the *CDT1* and *FAM38A*—*PIEZO1* genes. Although *CDT1* (chromatin licensing and

DNA replication factor 1) belongs to a family of DNA replication proteins and is essential for DNA replication, *PIEZO1* appeared as a more attractive candidate because of its recent identification as MA ion channels involved in mammalian mechanotransduction¹².

Genotyping of the *PIEZO1* p.Ala2020Thr (A2020T) variant, identified by exome sequencing, in 14 affected and 5 unaffected members of the family 1 showed a complete co-segregation with the disease. Bioinformatic softwares (SIFT, Polyphen 2) showed a high likelihood of pathogenicity score (Supplementary Table S2). We then performed a systematic screening of the entire coding sequence of *PIEZO1* gene in a series of two other kindreds and 11 unrelated DHS cases. Index cases were referred to our centre for investigation of chronic Coombs test negative haemolysis, with few symptoms for most of them. DHS was diagnosed by ektacytometry, and showed the typical dehydrated profile in each case. As expected, the electrophoretic analysis of RBC membrane protein was normal. Other biological and clinical findings were in agreement with the DHS diagnosis (Table 1). Three additional mutations were found in 10 subjects (Table 1): two missense mutations, p.Arg1358Pro (R1358P) and p.Thr2127Met (T2127M), and a duplication of six nucleotides 5'-GGAGCT-3' at the very end of exon 51, about 70 nucleotides before the stop codon (Fig. 1 and Supplementary Fig. S2). This duplication (c. 7479_7484dup) occurs in a region of low complexity and rich in leucine (L) and glutamate (E) codons. The resulting staggered in-frame duplication of leucine 2,495 and glutamate 2,496 (p.Leu2495_Glu2496dup) will be addressed as E2496ELE.

We observed that E2496ELE was recurrently found in eight unrelated index cases (Table 1). Haplotype analysis using single nucleotide polymorphisms (SNPs) at the *PIEZO1* gene showed that this in-frame duplication was carried by at least four different haplotypes, thus excluding an ancestral allele (Supplementary Table S3). A majority of available bioinformatic softwares could not predict its pathogenicity, but Mutation Taster considered it as possibly polymorphic. However, it was absent from the 1000 Genomes, Exome Variant Server, or dbSNP v135 databases. The four mutations were found in *PIEZO1* motifs with a high level of conservation among vertebrates (Fig. 1a).

We further assessed the frequency of E2496ELE in 600 healthy French controls and found it twice among 1,200 alleles (minor allele frequency (MAF) = 0.0017), contrasting with a very highly significant proportion of mutated alleles in our DHS series (8/22, minor allele frequency = 0.36, $P < 10^{-76}$). Of note, one of the two positive healthy individuals was considered to have hyperkalaemia in one of his blood tests, with no additional information. Another argument for pathogenicity of this recurrent variant was provided by its co-segregation with DHS diagnosis in families 2 and 3 (Fig. 1b). In these two families, each positive subject had both pseudohyperkalaemia and abnormal haematologic parameters. The penetrance of the condition could therefore be considered as 100% in the tested individuals. No obvious/simple genotype–phenotype relationship appear from our data; however, we noticed that this restricted sample of eight index cases bearing the E2496ELE mutation had a tendency for lower age at diagnosis (20.5 versus 60.6 years) and a more severe haematological phenotype compared to the three index cases carrying a missense mutation (Table 2). This observation needs to be confirmed with larger number of subjects. Overall, four different mutations lying in highly conserved C-terminal regions of the PIEZO1 protein (Fig. 2) were found in 11 unrelated familial and sporadic DHS cases with strong arguments of pathogenicity for each of them.

Electrophysiological studies of DHS mutations in hPIEZO1 channels. We then set out to determine whether these mutations

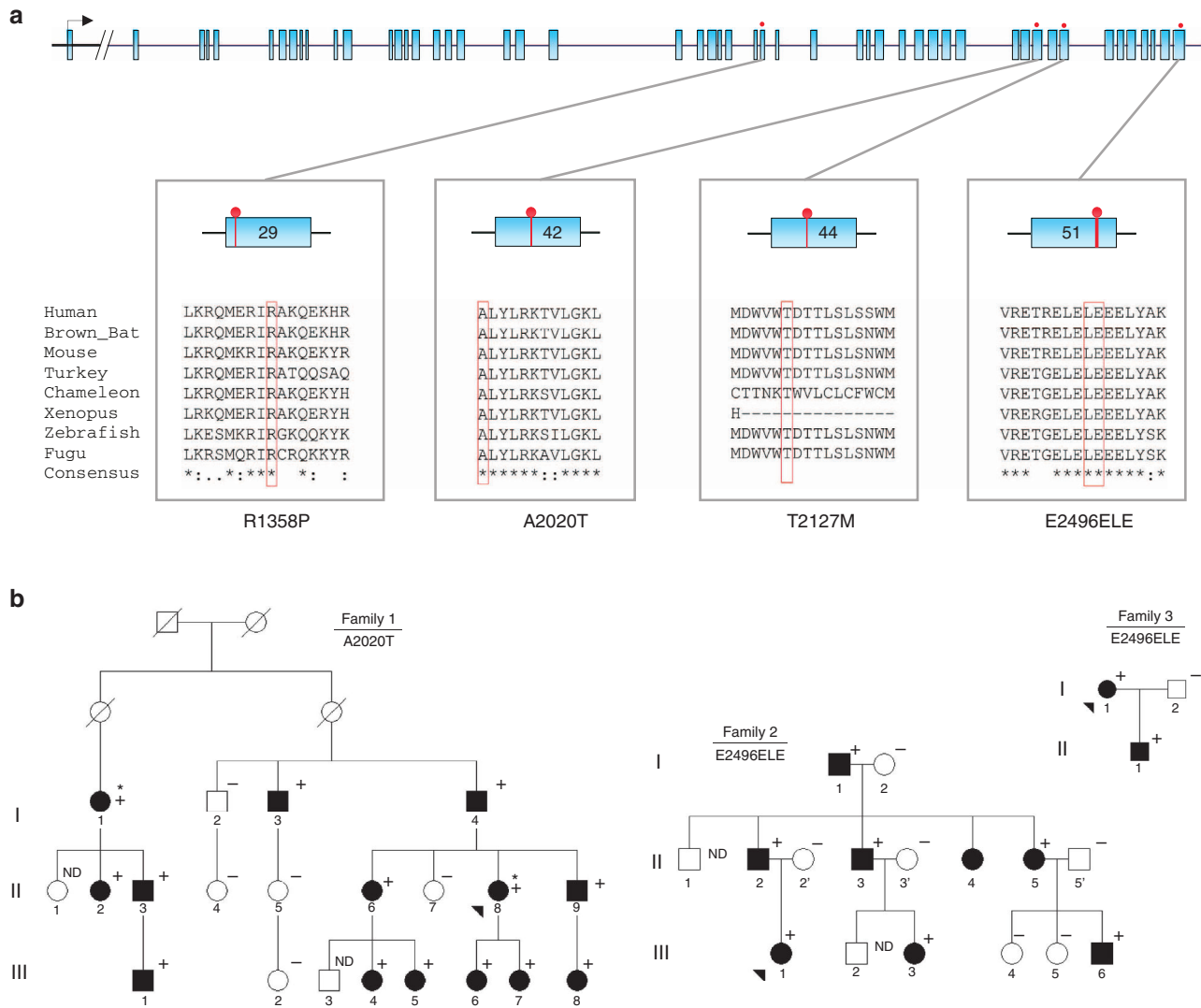


Figure 1 | Mutation screening results in *PIEZO1* among patients and relatives. (a) Graphical representation of the 51 exons of human *PIEZO1* gene, with evolutionary proteic conservation analysis of the mutants surrounding regions amongst vertebrates. Blue boxes: exons. Red lollipops: location of the identified mutations. Red boxes in aligned sequences: amino-acid residues concerned by the mutations. (b) Pedigrees of the three French DHS families with corresponding mutations. Black arrowheads: index cases; asterisks: affected patients selected for exome sequencing; + and -, respectively: presence or absence of the described mutation.

Table 1 | Clinical and haematological presentation of the 11 index cases.

Id	Piezo1 mutation	Age (years)*	Family history	PHK	RBC ($10^{12} l^{-1}$)	Hb ($g dl^{-1}$)	MCV (fl)	MCCH (%)	Retic. ($10^9 l^{-1}$)	PNE	Typical DHS ektacytometry (IDmax, Omin and O' (mOsm kg^{-1}))
F1	A2020T	48	Y	Y	5.03	15.6	88.9	34	259	N	Y (0.41, 112, 325)
K1	E2496ELE	16	N	N	3.09	10.3	102.8	32.5	136	N	Y (0.41, 100, 275)
K2	E2496ELE	15	ND	ND	4.05	14.4	98.5	36	202	ND	Y (0.39, 135, 286)
K3	E2496ELE	11	Y	ND	3.28	11.3	91.9	37.3	275	ND	Y (0.45, 119, 298)
K4	R1358P	69	N	N	4.47	15.4	98.1	35.2	367	ND	Y (0.41, 123, 301)
K5	E2496ELE	21	Y	ND	3.41	12.7	98.9	37.5	256	N	Y (0.45, 117, 288)
K6	E2496ELE	18	Y	N	3.38	12.5	102.5	36	378	N	Y (0.48, 99, 278)
K7	E2496ELE	42	Y	ND	3.19	12.6	108.6	36.5	290	ND	Y (0.50, 99, 283)
K8	T2127M	65	Y	Y	5.21	17.2	88.4	37.3	151	ND	Y (0.47, 128, 340)
F2	E2496ELE	30	Y	Y	ND	9.4	99.1	ND	182	Y	Y (0.42, 122, 303)
F3	E2496ELE	26	Y	Y	3.39	13.9	120	33.8	220	Y	Y (0.47, 119, 317)

DHS, dehydrated hereditary stomatocytosis; Hb, haemoglobin; MCCH, mean corpuscular concentration of haemoglobin; MCV, mean corpuscular volume; N, no; ND, not determined; PNE, perinatal oedema; PHK, pseudohyperkalaemia; RBC, red blood cells count; retic, reticulocytes count; Y, yes.

Typical DHS ektacytometry shows normal RBC deformability (IDmax) and decreased Omin and O' reflecting RBC dehydration. (Normal range, IDmax: 0.38-0.55, Omin: 135-155 mosm kg^{-1} , O': 340-375 mosm kg^{-1}).

*Age represents the age at diagnosis.

Table 2 | Clinical and haematological comparison between index patients with PIEZO1 mutations.

	E2496ELE (N = 8)	Others (N = 3)	P-value
Age (years)	20.5 (11–42)	60.6 (48–69)	0.01
Reticulocytes (n mm ⁻³)	256.2 (136–378)	259 (151–367)	0.9
Red blood cells count (10 ¹² l ⁻¹)	3.4 (3.09–4.05)	4.9 (4.47–5.21)	0.007
Haemoglobin (g l ⁻¹)	12.3 (9.4–14.4)	16.1 (15.6–17.2)	0.002
Mean cell volume (fl)	100.5 (98.5–120)	91.8 (88.4–98.1)	0.04
Mean corpuscular haemoglobin concentration (%)	36 (32.5–37.5)	35.5 (34–37.3)	0.9

Mean age at diagnosis and haematological parameters (including minimal and maximal values) between the eight index cases harbouring the recurrent mutation E2496ELE versus the three dehydrated hereditary stomatocytosis index cases with other missense PIEZO1 mutations (others).

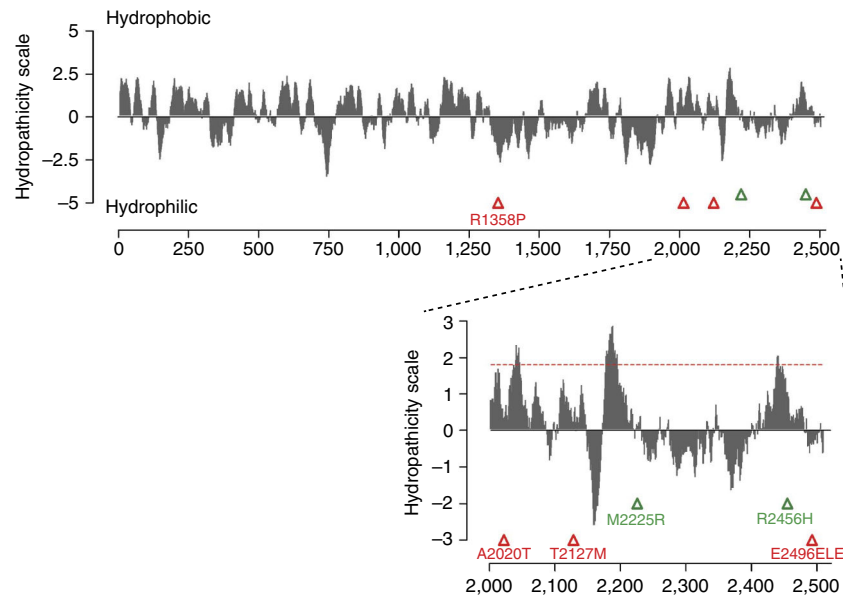


Figure 2 | Hydrophobicity plot of human PIEZO1 and position of reported mutations. Kyte-Doolittle hydrophobicity analysis (19 residues window) of human PIEZO1 prepared using the ProtScale program (Expasy). Triangles indicate the position of previously reported (green) and newly identified (red) mutations associated with DHS. Inset: magnification of C-terminal region. Peaks with scores greater than 1.8 (red dashed line) indicate possible transmembrane regions.

alter PIEZO1 channel activity. We introduced the four mutations identified from our study (R1358P, A2020T, T2127M and E2496ELE) and the two previously published (M2225R and R2456H) DHS-associated mutations⁷ in *hPIEZO1* full-length complementary DNA. We transfected HEK293T cells with wild-type (WT) or mutant *hPIEZO1* and analysed channel activity via whole-cell patch clamp configuration. MA inward currents were recorded by applying mechanical force on to the cell surface using a glass probe¹² (Fig. 3a). In cells expressing WT *hPIEZO1* such stimulation gives rise to rapid transient increase in current.

Sensitivity of a channel to mechanical stimulation, which can also be expressed as the distance of the mechanical stimulus, can be estimated by measuring ‘threshold’ values for each cell. We measure this by subtracting the distance at which the probe first touches the cell surface from the minimal distance at which MA currents are evoked. Threshold is therefore merely the distance of indentation, after making cell contact, required to activate the channel. We observed no apparent differences in threshold values for any of the *hPIEZO1* mutants relative to WT *hPIEZO1*, suggesting that the DHS mutations do not alter the channel’s sensitivity to mechanical stimulation (Table 3).

Similar to WT *hPIEZO1*, all the mutants had a linear current–voltage correlation with a reversal potential close to 0 mV (Supplementary Fig. S3). At -80 mV, the maximum current

(I_{\max}) recorded from cells expressing mutant *hPIEZO1* channels was found to be similar to that of WT *hPIEZO1* (Table 3). A potential exception to this was M2225R-induced currents, which showed overall larger I_{\max} currents. However, the average was not significantly different compared with WT. In addition, peak currents depend on expression levels in HEK293T cells, and might not represent endogenous expression.

All WT and mutant *hPIEZO1* MA inactivation currents could be fit with a mono-exponential function. Remarkably, we observed that all mutations caused considerable increase in the inactivation time constant relative to WT channel kinetics (Fig. 3b). The R2456H mutation had the maximum impact with a 2.6-fold slower time constant, while the remaining five mutations inactivate with 1.3- to 1.7-fold slower kinetics (Fig. 3c and Table 3). Therefore, all six mutations can be defined as gain-of-function, as they would give rise to an increased channel activity in response to a given mechanical stimulus.

Discussion

We present a comprehensive analysis of DHS-causing *PIEZO1* mutations. Four novel *PIEZO1* mutations were identified in 11 out of 14 index cases, demonstrating that this gene has a major role in DHS. DHS is a difficult diagnosis because of a highly

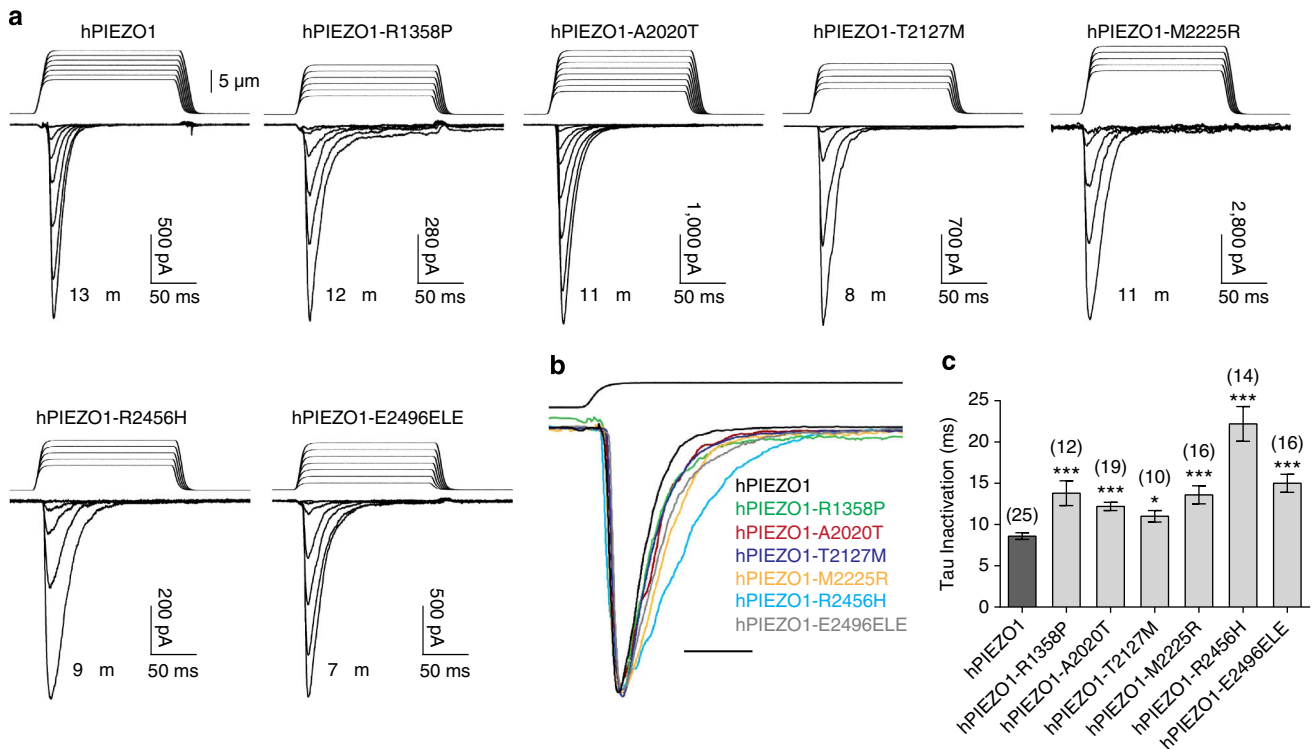


Figure 3 | Recombinant hPIEZO1 channels with DHS-associated mutations display slow inactivation kinetics. (a) Representative traces of mechanically activated inward currents recorded at -80 mV from HEK293T cells expressing either WT or indicated mutant hPIEZO1. Cells were stimulated by a series of mechanical steps (150 ms duration) in $1\ \mu\text{m}$ increments. For each representative trace, the channel response corresponds to the indentation distance of the glass probe and is shown as a series of responses starting from the step that generates the first mechanically activated current (the initial steps that did not cause channel activation are not shown) till the step that generates the maximum response for that cell. The distance of the probe for the maximum response for each construct is indicated. (b) Representative WT (black) and mutant (coloured) hPIEZO1 traces normalized to peak. Maximum currents from each construct shown in a are overlaid to highlight difference in inactivation kinetics. Scale bar; 25 ms. (c) Average of inactivation time constant (τ) for WT and mutant hPIEZO1 channels. Bars represent mean \pm s.e.m. and numbers in parenthesis indicate number of cells tested for each condition. * $P < 0.05$, *** $P < 0.001$; relative to WT hPIEZO1 (Student's t -test).

Table 3 | Inactivation kinetics of hPIEZO1 mutations.

Mutation	Inactivation kinetics			I_{max} at -80 mV	
	n	Threshold (μm)	Inactivation time constant, τ (ms)	n	nA
hPIEZO1	25	5.0 ± 0.3	8.6 ± 0.4	9	2.2 ± 0.4
hPIEZO1-R1358P	12	4.3 ± 0.5	$13.8 \pm 1.5^{***}$	12	1.4 ± 0.3
hPIEZO1-A2020T	19	4.7 ± 0.2	$12.2 \pm 0.5^{***}$	8	2 ± 0.4
hPIEZO1-T2127M	10	5.0 ± 0.5	$11 \pm 0.7^*$	10	2.2 ± 0.8
hPIEZO1-M2225R	16	4.4 ± 0.4	$13.6 \pm 1.1^{***}$	9	7.7 ± 3.7
hPIEZO1-R2456H	14	4.7 ± 0.4	$22.2 \pm 2.1^{***}$	7	1.6 ± 0.5
hPIEZO1-E2496ELE	16	4.2 ± 0.3	$15 \pm 1.1^{***}$	11	1.5 ± 0.2

All values measured from cells held at -80 mV membrane potential. Average values expressed as mean \pm s.e.m. * $P < 0.05$; *** $P < 0.001$ relative to wild-type hPIEZO1 (Student's t -test).

variable clinical expression, ranging from the absence of clinical symptoms to lethal perinatal oedema. Other features of DHS include severe iron overload leading to hepatic transplantation, or life-threatening thromboembolic disease after splenectomy. But the most frequent DHS condition is moderately symptomatic haemolysis. The only biological test for DHS is ektacytometry, which is available in a very limited number of laboratories. The disease may be overlooked for years or decades and is sometimes confused with spherocytosis. Our data, including the presence of a recurrent mutation, render a genetic diagnosis possible.

PIEZO1 is a large homotetrameric mechanosensitive ion channel with each subunit containing 2,521 amino acids and

predicted to encompass no less than 30 transmembrane domains¹¹. However, PIEZOs do not share homology with any known ion channels or proteins. Interestingly, all six DHS-associated PIEZO1 mutations are located in the second half of the coding sequence. Indeed, apart from R1358P, all are in the C terminal 1/5th of the protein, albeit distributed across several exons in the gene. It is difficult to conclude too much from this information, as we know virtually nothing about the topology of PIEZO1, or the role of different PIEZO regions in channel function. However, our electrophysiological studies implicate the C-terminal region in regulating gating kinetics.

Human *PIEZO1* is a very polymorphic gene. As an example, many coding variants of *PIEZO1* are present in the 1,000 genomes database, including variants predicted as deleterious. Interestingly, the distribution of these variants tends to cluster to the N-terminal end of the protein, unlike DHS mutations (Supplementary Fig. S1), suggesting an increasing selective pressure from the N terminal to the C-terminal end of the protein. One surprising finding is the recurrence of the E2496ELE mutation, observed in 8 out of the 11 DHS *PIEZO1*-related cases. It was also found twice in 1,200 healthy controls of French origin. We had no access to the medical record of these two *a priori* asymptomatic subjects, but one of these two heterozygous carriers had hyperkalaemia. As mentioned above, this in-frame insertion can neither be explained by a founder effect nor could be considered as a polymorphism taking into account its rarity in the general population, its associated clear phenotype and its electrophysiological consequences. The six-nucleotide-inserted sequence 'GGAGCT' represents a short intragenic in-frame insertion, embedded in a low complexity surrounding sequence between positions 2,491 and 2,499. Slippage mutagenesis, leading to the addition or removal of one copy of a short tandem repeat, is probably responsible for this recurrence. In the human genome, microdeletions and microinsertions of >1 to <20 nt were found to be strongly associated with surrounding short tandem repeats, similar to the sequence surrounding this duplication¹⁶.

Characterization of kinetic properties of mutant *PIEZO1* channels showed a striking effect: all six DHS-causing mutations resulted in a similar slowing of inactivation kinetics. We propose that these mutations could affect the inactivation mechanism by either destabilizing the inactivated state or stabilize the channel in the open state thereby biasing the open to inactivated state equilibrium towards open, resulting in an increased Na⁺ and Ca²⁺ ion influx. RBCs are unique in that they circulate in the vascular system for 120 days and undergo constant mechanical stress, specifically through the capillaries and splenic cords. Stabilization of the open conformation of *PIEZO1* channels in RBCs could also alter the transition between closed and open states of the channel leading to higher open probability compared with WT channels. This shift in kinetics could potentially also cause a slight increase in basal channel activity (independent of mechanical stimulation), which may also contribute to the DHS phenotype.

Detailed kinetic analysis of these channels in the future should tell us more about the gating mechanism and how these mutations in *PIEZO1* could lead to erythroid dysfunction. Our analysis here was restricted to heterologously expressed *PIEZO1* mutants. Recording of MA currents from control and affected RBCs in the future could more directly link *PIEZO1*-dependent current characteristics and pathophysiology of DHS.

It is possible that increased *PIEZO1* activity could directly contribute to RBC dehydration observed in DHS patients. Human RBCs have a resting membrane potential of about -10 mV, and one possibility is that the gain-of-function mutations cause increase in influx of Na⁺ and efflux of K⁺ through *PIEZO1* channels. Indirect mechanisms could also account for this phenotype. For example, *PIEZO1*-mediated rise in intracellular Ca²⁺ concentration could induce K⁺ efflux by the activation of Ca²⁺-activated K⁺ channels, the Gardos channel. Gardos channels have been shown to have a vital role in physiological and pathological states of erythrocytes. Additionally, it has been shown that Gardos channel activation causes a reduction in erythrocyte cell volume^{17,18}.

Multiple cell types along with erythroid progenitor cells express *PIEZO1* (ref. 7). The specific haematologic phenotype of DHS patients due to gain-of-function *PIEZO1* mutants is therefore

surprising. It is, however, important to note that DHS is a mostly tolerated disorder, and even the RBC phenotype is relatively subtle—as reflected by the modest changes of RBC constants such as mean cell volume and mean cell haemoglobin concentration or the weak (5–15%) number of stomatocytes^{1,19}. Therefore, variations in minor modulating factors could impact which cell types show phenotypes in DHS patients. For example, RBCs might be more sensitive to increased cation flux compared with other cell types, or *PIEZO1* channels expressed in RBCs might be modulated in a way that amplifies the consequences of the *PIEZO1* mutations. Another intriguing possibility is that RBC plasma membranes experience more mechanical stress compared with other cells because of their unique membrane properties and constant movement during blood circulation. It is also possible that other cell types and tissues are affected in DHS patients, but these have remained unknown or unexplained because of these modulating factors and the possible subtlety of phenotype. Perinatal oedema that is inconstantly observed in DHS patients is such an example. *PIEZO1* is expressed in endothelial cells (unpublished data) and this phenotype could be the consequence of a leak out of the vascular compartment at a specific development stage. *PIEZO1* has been implicated in cell–cell and cell–matrix interactions²⁰. In both *PIEZO1* knockdown experiments^{13,14}, the mechanosensitivity and gating properties of the protein were not assessed, but the influence of *PIEZO1* on cell migration/invasion and extrusion was described.

The function of a mechanosensitive *PIEZO1* channel in normal RBCs remains speculative. Erythrocytes are highly deformable cells, a property that allows them to pass through narrow capillaries of the microvasculature²¹. Mechanosensory feedback mechanisms could allow the cell to respond to its changing mechanical environment, possibly by regulating its volume and morphology. In this respect, it is interesting to comment on overhydrated hereditary stomatocytosis, another genetic disease that like DHS is associated with an imbalance in erythrocyte cation gradients. This disease is associated with mutations in Rhesus-associated glycoprotein, and is characterized by a dramatic reduction of stomatin from erythrocyte membranes^{1,22,23}. Strikingly, stomatin family members from mouse and worm have been shown to be required for mechanotransduction in sensory neurons^{24–26}. This tantalizing coincidence raises the possibility that the various genetic conditions that cause stomatocytosis are caused by disrupted mechanotransduction, and that perhaps *PIEZO1* and stomatin family members function in the same pathway. Yet another clue towards a role for mechanotransduction in erythrocytes comes from sickle cell disease, arguably the best-known erythrocyte disorder. In sickle RBCs, mutant haemoglobin molecules form large polymers under low oxygen conditions, which cause the deformation of the erythrocyte. This deformation coincides with the development of a cationic membrane current, termed pSickle, and subsequent erythrocyte dehydration¹⁷. Notably pSickle is blocked by a blocker of stretch-activated currents, the tarantula toxin Gsmtx4, which is also known to block *PIEZO1* (refs 27–29). Future studies will explore the relationship between *PIEZO1* and pSickle.

In conclusion, we identify four new *PIEZO1* mutations in DHS patients, demonstrating that *PIEZO1* is a major determinant of DHS. We also present the first functional analysis of DHS-causing *PIEZO1* mutations. Our studies provide a substantial step towards understanding physiopathology of DHS. Our findings will also be useful to establish DHS genotype/phenotype relationship and to refine the classification of RBC cation leak disorders on a molecular basis.

During the final revision of this manuscript, two related papers were published which also show that DHS mutations in *PIEZO1* alter channel activity^{30,31}.

Methods

Patients. Our study focused on three families and 11 independent index cases with a typical DHS clinical and biological phenotype. Patients were from all regions of France, and were diagnosed through specialized consultation by a haematologist or a clinical geneticist. Some of the patients have been published elsewhere: family 1, which was described by us and others (family VA in Grootenboer *et al.*² and Beaurain *et al.*³, presented with mild, uncomplicated haematological signs of the disease, pseudohyperkalaemia and no history of perinatal oedema. In family 2 (family VE in Grootenboer *et al.*) patients were affected by a mild-to-moderate haemolysis and anaemia, by a recurrence of perinatal oedema (three affected cases), and varying recurrence of pseudohyperkalaemia. Case 4 (ref. 3) and family 3 (refs 32,33) were case reports, these patients expressed a moderate form of anaemia and haemolysis without initial evidence of perinatal oedema or pseudohyperkalaemia. In these last three cases, diagnosis of DHS was done late after the onset of haematologic manifestations. The other cases were referred to the Haematology Department of the Hospital Kremlin-Bicêtre and their characteristics are summarized in Table 1.

Informed consents for genetic analysis were obtained for all patients, according to local institutional ethical board, Hôpital Européen Georges Pompidou Ethical Comité.

The research was conducted in the Hôpital Européen Georges Pompidou Genetics Department or the Haematology Department of Kremlin-Bicêtre University Hospital in Paris, France, and the Scripps Research Institute in San Diego, California.

Haematological tests. Blood electrolytes were measured in each center. RBC studies were performed in the same laboratory: cell counts and red cell constants were measured using an ADVIA 2120 (Siemens) instrument, blood smears were observed after May-Grünwald-Giemsa staining using standard methods, red cells membrane properties were assessed using osmolar gradient (ranging from 80 to 500 mosmol kg⁻¹) ektactometry using a Technicon ektactometer and erythrocyte membrane proteins analysed by SDS-polyacrylamide gel electrophoresis as previously described^{1,34,35}.

Molecular genetics analysis and targeted exome sequencing. Genomic DNA was extracted from EDTA blood samples using the QIAamp Midi Kit blood DNA extraction kit (Quiagen). Quality of the DNA was evaluated using nanodrop 2000 spectrophotometer (Thermo Scientific).

Library preparation, capture, sequencing, and variants detection and annotation were performed by IntegraGen, Evry, France. Exons of genomic DNA samples were captured using Agilent in-solution enrichment methodology with their biotinylated oligonucleotides probes library, followed by paired-end 75 bp massively parallel sequencing on Illumina HiSeq 2000. For detailed explanations of the process, see Gnirke *et al.*³⁶ publication in *Nature Biotechnology*. Sequence capture, enrichment and elution were performed according to the manufacturer's instruction and protocols (Human All exon kit V3, 50Mb, Agilent) without modification. Briefly, 3 µg of each genomic DNA were fragmented by sonication and purified to yield fragments of 150–200 bp. Paired-end adaptor oligonucleotides from Illumina were ligated on repaired A-tailed fragments, then purified and enriched by six PCR cycles. Six hundred nanograms of these purified libraries were hybridized to the SureSelect oligo probe capture library during 72 h. After hybridization, washing and elution, the eluted fraction was PCR amplified with 11 cycles, purified and quantified by qPCR to obtain sufficient DNA template for downstream applications. Each eluted-enriched DNA sample was then sequenced on an Illumina HiSeq 2000 as paired-end 75b reads. Image analysis and base calling were performed using Illumina Real Time Analysis Pipeline version 1.14 with default parameters.

Bioinformatics analysis of PIEZO1 for exome study. The bioinformatics analysis of deep sequencing data was based on the Illumina pipeline (CASAVA 1.8). CASAVA performs alignment of the reads to the human reference genome (hg19) with the alignment algorithm ELANDv2 (performs multiseed and gapped alignments), then calls the SNPs based on the allele calls and read depth, and detects variants (SNPs and Indels). Only the positions included in the bait coordinates were conserved. Genetic variations annotation was performed using IntegraGen in-house pipeline, which consists in gene annotation (RefSeq), detection of known polymorphisms (dbSNP v132, 1,000 genomes) followed by a mutation characterization (exonic, intronic, silent, nonsense and so on). For each position, the exomic frequencies (homo- and heterozygous) were determined from the IntegraGen Exome database, and the exome results provided by HapMap. Results were provided per sample and family upon tabulated text files, including coverage/depth statistical analysis per exome and per target.

Genomic sequences of all 51 exons of *PIEZO1* gene were amplified by PCR using either Red'y star reaction mix (Eurogentec) or Jumpstart reaction mix (Sigma). Twenty different primer pairs amplifying one or more exons each, targeting the unique reference transcript sequence of the gene (NM_001142864.2) were used. Sequencing reactions of the PCR products were performed using the Big Dye Terminator sequencing kit v3.3 (Life Technologies, Carlsbad, California, USA). Twenty-five additional primers were used to obtain the complete coding sequence

of the gene from both strands. Sequencing products were migrated on an ABI prism 3730 sequencer (Life Technologies, Carlsbad, California, USA) and analysed using the Sequencher 5.0 software (Genecodes, Ann Arbor, Michigan, USA).

Sequencing of *PIEZO* exon 51, for population frequency estimation of E2496ELE, was performed in 1,200 control alleles from a French healthy controls cohort. Primer pairs and PCR conditions are available upon request.

Sequence alignment and analysis were performed using Sequencher software 5.0 (Genecodes, Ann Arbor, Michigan, USA). Sequence variations were analysed using Alamut (Interactive Biosoftware, Rouen, France), and predicted pathogenicity of variants was assessed using Polyphen 2, SIFT and Mutation Taster online tools. Frequency of variants in published populations were assessed in 1,000 genomes, Exome Variant Server and dbSNP databases. Human to other vertebrates protein alignments of *PIEZO1* were performed using CLUSTAL Omega.

All web tools addresses are listed in the Supplementary Table S4.

Cloning and tissue culture. Halo-tagged *hPIEZO1* was purchased from Kazusa DNA Research Institute. This clone differs from the NCBI sequence with coding differences at I83T, V250A, V394L and R407G and non-coding differences at codons 1338, 1846, 2190, 2214 and 2353. We first confirmed the functionality of the clone through electrophysiological experiments. The cDNA was then cloned into pIRES2-eGFP expression vector, and the DHS mutations were generated either by overlapping PCR mutagenesis or by using Agilent's QuikChange II XL Site Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions. All clones were verified by full-length DNA sequencing.

Human embryonic kidney 293T (HEK293T) cells were grown in DMEM containing 4.5 mg ml⁻¹ glucose, 10% fetal bovine serum, 50 U ml⁻¹ penicillin and 50 mg ml⁻¹ streptomycin. Cells were plated onto poly-L-lysine-coated 12-mm round glass coverslips placed in 24-well plates and transfected using lipofectamine 2000 (Invitrogen) according to the manufacturer's instruction. Plasmid DNA (600 ng ml⁻¹) was transfected and green fluorescent protein-positive cells were recorded from, 12–48 h post transfection.

Electrophysiology and mechanical stimulation. Patch clamp experiments were performed in standard whole-cell recordings using an Axopatch 200B amplifier (Axon Instruments). Patch pipettes had a resistance of 1.5–3 MΩ when filled with intracellular solution consisting of (in mM) 133 CsCl, 5 EGTA, 1 CaCl₂, 1 MgCl₂, 10 HEPES, 4 MgATP and 0.4 Na₂GTP (pH adjusted to 7.3 with CsOH). The extracellular solution consisted of (in mM) 130 NaCl, 3 KCl, 1 MgCl₂, 2.5 CaCl₂, 10 HEPES and 10 glucose (pH adjusted to 7.3 with NaOH). All experiments were done at room temperature. Currents were sampled at 20 kHz and filtered at 2 kHz. Voltages were not corrected for a liquid junction potential. Leak currents before mechanical stimulations were subtracted off-line from the current traces.

Mechanical stimulation was achieved using a fire-polished glass pipette (tip diameter 3–4 µm) positioned at an angle of 80° to the recording cell. Downward movement of the probe towards the cell was driven by a Clampex controlled piezo-electric crystal microstage (E625 LVPZT Controller/Amplifier; Physik Instrumente). The probe had a velocity of 1 µm ms⁻¹ during the ramp segment of the command for forward motion and the stimulus was applied for 150 ms. To assess the mechanical sensitivity of a cell, the probe is first placed as close to the cell as possible (this distance could vary from cell to cell). Then, a series of mechanical steps in 1-µm increments was applied every 10 s, which allowed full recovery of mechanosensitive currents. Threshold was calculated by subtracting the distance at which the probe first touches the cell surface from the minimal distance at which MA currents are evoked. MA inward currents were recorded at a holding potential of –80 mV. For I–V relationship recordings, voltage steps were applied 0.7 s before the mechanical stimulation from a holding potential of –60 mV.

Statistical analysis. Frequency comparisons of variants were done using Fisher exact test. Clinical and haematological quantitative variables comparisons between subsets of patients were done using Student's *t*-tests. Channel kinetic properties between WT and mutant *PIEZO1* were compared using Student's *t*-test.

References

1. Delaunay, J. The hereditary stomatocytoses: genetic disorders of the red cell membrane permeability to monovalent cations. *Semin. Hematol.* **41**, 165–172 (2004).
2. Grootenboer, S. *et al.* Pleiotropic syndrome of dehydrated hereditary stomatocytosis, pseudohyperkalemia, and perinatal edema maps to 16q23-q24. *Blood* **96**, 2599–2605 (2000).
3. Syfuss, P. Y. *et al.* Mild dehydrated hereditary stomatocytosis revealed by marked hepatosiderosis. *Clin. Lab. Haematol.* **28**, 270–274 (2006).
4. Carella, M. *et al.* Genomewide search for dehydrated hereditary stomatocytosis (hereditary xerocytosis): mapping of locus to chromosome 16 (16q23-qter). *Am. J. Hum. Genet.* **63**, 810–816 (1998).
5. Beaurain, G. *et al.* Dehydrated hereditary stomatocytosis mimicking familial hyperkalaemic hypertension: clinical and genetic investigation. *Eur. J. Haematol.* **78**, 253–259 (2007).

6. Houston, B. L. *et al.* Refinement of the hereditary xerocytosis locus on chromosome 16q in a large Canadian kindred. *Blood Cells Mol. Dis.* **47**, 226–231 (2011).
7. Zarychanski, R. *et al.* Mutations in the mechanotransduction protein PIEZO1 are associated with hereditary xerocytosis. *Blood* **120**, 1908–1915 (2012).
8. Carella, M. *et al.* A second locus mapping to 2q35–36 for familial pseudohyperkalemia. *Eur. J. Hum. Genet.* **12**, 1073–1076 (2004).
9. Carella, M. *et al.* Genetic heterogeneity of hereditary stomatocytosis syndromes showing pseudohyperkalemia. *Haematologica* **84**, 862–863 (1999).
10. Andolfo, I. *et al.* Missense mutations in the ABCB6 transporter cause dominant familial pseudohyperkalemia. *Am. J. Hematol.* **88**, 66–72 (2013).
11. Coste, B. *et al.* Piezo proteins are pore-forming subunits of mechanically activated channels. *Nature* **483**, 176–181 (2012).
12. Coste, B. *et al.* Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. *Science* **330**, 55–60 (2010).
13. McHugh, B. J., Murdoch, A., Haslett, C. & Sethi, T. Loss of the integrin-activating transmembrane protein Fam38A (Piezo1) promotes a switch to a reduced integrin-dependent mode of cell migration. *PLoS ONE* **7**, e40346 (2012).
14. Eisenhoffer, G. T. *et al.* Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia. *Nature* **484**, 546–549 (2012).
15. Kim, S. E., Coste, B., Chadha, A., Cook, B. & Patapoutian, A. The role of *Drosophila* Piezo in mechanical nociception. *Nature* **483**, 209–212 (2012).
16. Ball, E. V. *et al.* Microdeletions and microinsertions causing human genetic disease: common mechanisms of mutagenesis and the role of local DNA sequence complexity. *Hum. Mutat.* **26**, 205–213 (2005).
17. Lew, V. L. & Bookchin, R. M. Ion transport pathology in the mechanism of sickle cell dehydration. *Physiol. Rev.* **85**, 179–200 (2005).
18. Maher, A. D. & Kuchel, P. W. The Gárdos channel: a review of the Ca²⁺-activated K⁺ channel in human erythrocytes. *Int. J. Biochem. Cell Biol.* **35**, 1182–1197 (2003).
19. Stewart, G. W. Hemolytic disease due to membrane ion channel disorders. *Curr. Opin. Hematol.* **11**, 244–250 (2004).
20. McHugh, B. J. *et al.* Integrin activation by Fam38A uses a novel mechanism of R-Ras targeting to the endoplasmic reticulum. *J. Cell Sci.* **123**, 51–61 (2010).
21. Mohandas, N. & Gallagher, P. G. Red cell membrane: past, present, and future. *Blood* **112**, 3939–3948 (2008).
22. Bruce, L. J. Hereditary stomatocytosis and cation leaky red cells — recent developments. *Blood Cells Mol. Dis.* **42**, 216–222 (2009).
23. Bruce, L. J. *et al.* The monovalent cation leak in overhydrated stomatocytic red blood cells results from amino acid substitutions in the Rh-associated glycoprotein. *Blood* **113**, 1350–1357 (2009).
24. Huang, M., Gu, G., Ferguson, E. L. & Chalfie, M. A stomatin-like protein necessary for mechanosensation in *C. elegans*. *Nature* **378**, 292–295 (1995).
25. Martinez-Salgado, C. *et al.* Stomatin and sensory neuron mechanotransduction. *J. Neurophysiol.* **98**, 3802–3808 (2007).
26. Wetzel, C. *et al.* A stomatin-domain protein essential for touch sensation in the mouse. *Nature* **445**, 206–209 (2007).
27. Vandorpe, D. H. *et al.* Hypoxia activates a Ca²⁺-permeable cation conductance sensitive to carbon monoxide and to GsMTx-4 in human and mouse sickle erythrocytes. *PLoS ONE* **5**, e8732 (2010).
28. Gottlieb, P. A., Bae, C. & Sachs, F. Gating the mechanical channel Piezo1: a comparison between whole-cell and patch recording. *Channels* **6**, 282–289 (2012).
29. Gottlieb, P. A. & Sachs, F. Piezo1: properties of a cation selective mechanical channel. *Channels* **6**, 214–219 (2012).
30. Bae, C., Gnanasambandam, R., Nicolai, C., Sachs, F. & Gottlieb, P. A. Xerocytosis is caused by mutations that alter the kinetics of the mechanosensitive channel PIEZO1. *Proc. Nat. Acad. Sci.* **110**, E1162–E1168 (2013).
31. Andolfo, I. *et al.* Multiple clinical forms of dehydrated hereditary stomatocytosis arise from mutations in PIEZO1. *Blood* doi:10.1182/blood-2013-02-482489 (2013).
32. Martinaud, C. *et al.* Antiphospholipid antibodies in a family with dehydrated hereditary stomatocytosis. *Thromb. Res.* **122**, 572–575 (2008).
33. Carli, P., Graffin, B., Gisserot, O., Landais, C. & De Jaureguiberry, J. P. Maladie thromboembolique récidivante après splénectomie pour stomatocytose héréditaire. *La Revue de Médecine Interne* **28**, 879–881 (2007).
34. Clark, M., Mohandas, N. & Shohet, S. Osmotic gradient ektacytometry: comprehensive characterization of red cell volume and surface maintenance. *Blood* **61**, 899–910 (1983).
35. Rocha, S. *et al.* Erythrocyte membrane protein destabilization versus clinical outcome in 160 Portuguese hereditary spherocytosis patients. *Br. J. Haematol.* **149**, 785–794 (2010).
36. Gnirke, A. *et al.* Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotech.* **27**, 182–189 (2009).

Acknowledgements

We thank the families for their interest and support in our research; Geneviève Beaurain for her help in the genetic analysis of family 1; Patrick Nitschké and Jean-Marc Plaza at the Foundation Imagine, Université Paris Descartes, for their help in the exome analysis and linkage analysis of family 1; and the technical staff of the biochemistry and haematology lab of Hôpital Bicêtre for technical skill and collaborating clinicians for referring patients to us. This work was supported by grants from INSERM, ANR, Fondation pour la Recherche Médicale to X.J. and National Institute of Health (NS083174) to A.P.

Author contributions

G.T., J.D., J.P.J., P.Y.S., L.G., F.T., P.S.R. and V.P. recruited patients, recorded clinical data and discussed data. J.A. performed the genetic analysis, H.L.D.P. interpreted the linkage and the exome analyses in family F1. V.P. and M.F.T. analysed biochemical and haematological RBC-related data and made DHS diagnoses. S.E.M. and B.C. recorded and analysed electrophysiological data. M.B., J.M. and S.C. performed mutagenesis and cloning. J.A., V.P. and X.J. designed, performed and interpreted genetic experiments. S.E.M., B.C. and A.P. designed electrophysiology experiments. J.A., S.E.M., B.C., M.B., V.P., X.J. and A.P. discussed results and wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Albuissou, J. *et al.* Dehydrated Hereditary stomatocytosis linked to gain-of-function mutations in mechanically activated PIEZO1 ion channels. *Nat. Commun.* **4**:1884 doi: 10.1038/ncomms2899 (2013).

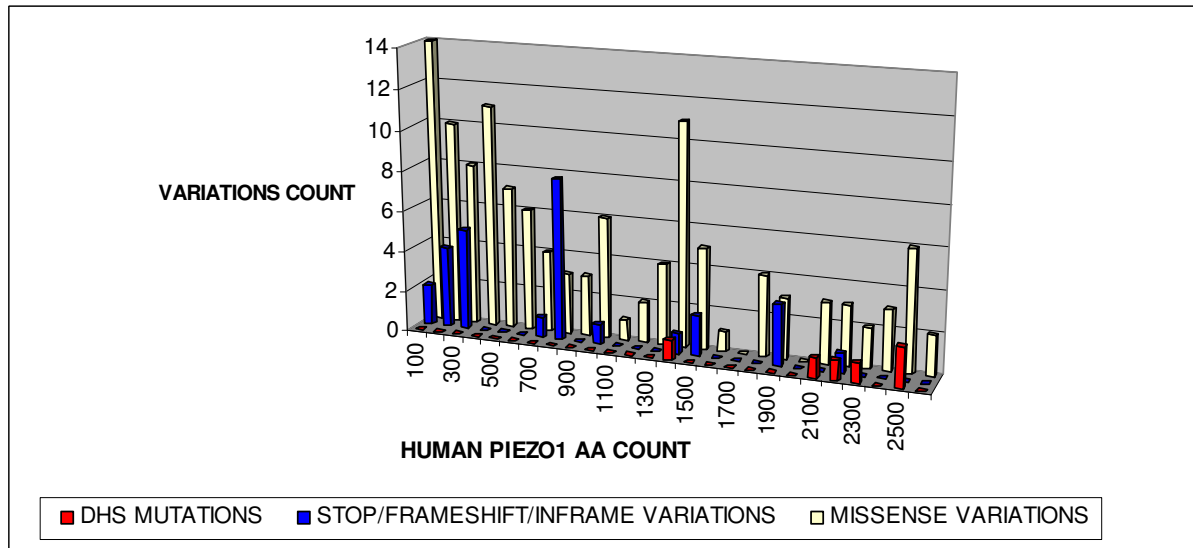
Corrigendum: Dehydrated hereditary stomatocytosis linked to gain-of-function mutations in mechanically activated PIEZO1 ion channels

Juliette Albuissou, Swetha E. Murthy, Michael Bandell, Bertrand Coste, H el ene Louis-dit-Picard, Jayanti Mathur, Madeleine F en eant-Thibault, G erard Tertian, Jean-Pierre de Jaureguiberry, Pierre-Yves Syfuss, Stuart Cahalan, Loic Gar on, Fabienne Toutain, Pierre Simon Rohrllich, Jean Delaunay, V eronique Picard, Xavier Jeunemaitre & Ardem Patapoutian

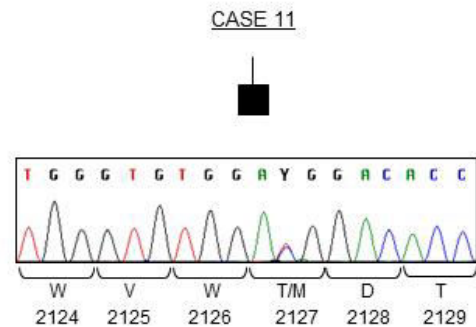
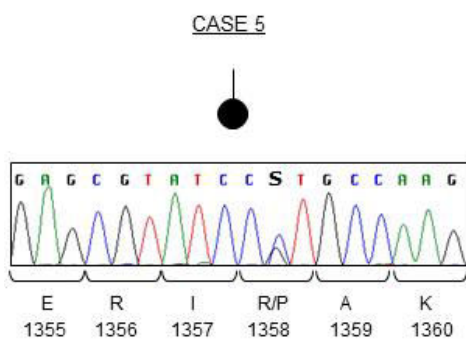
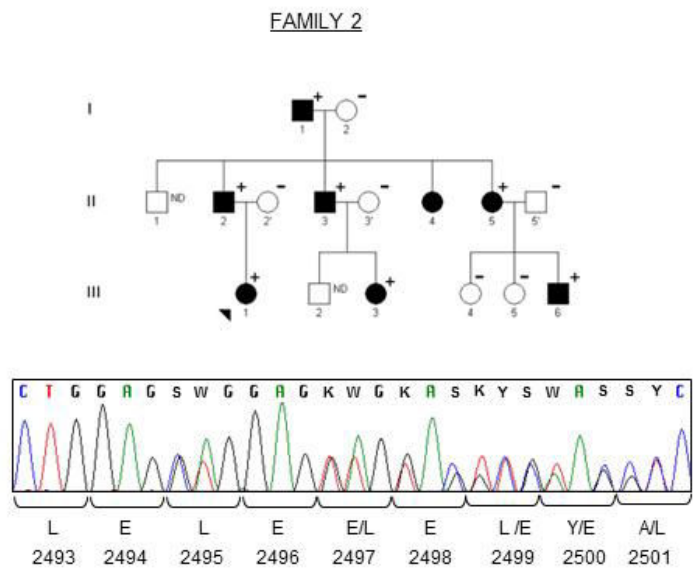
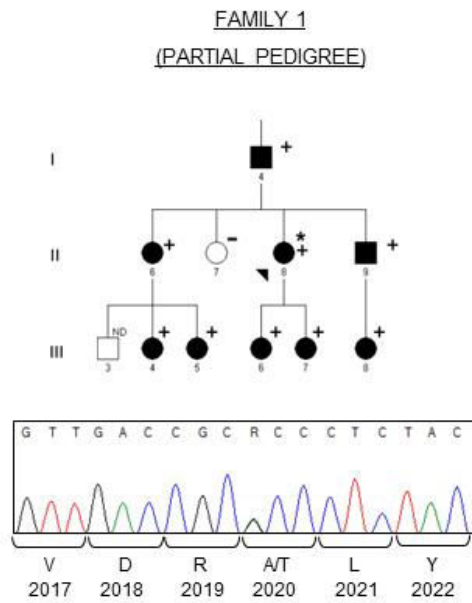
Nature Communications 4:1884 doi: 10.1038/ncomms2899 (2013); Published 21 May 2013; Updated 23 Sep 2013

In the first paragraph of the Methods section in this Article, case 8 (K8 in Table 1) was incorrectly referred to as case 4. The following is the correct sentence: ‘Case 8 (ref. 3) and family 3 (refs 32, 33) were case reports, these patients expressed a moderate form of anaemia and haemolysis without initial evidence of perinatal oedema or pseudohyperkalaemia.’

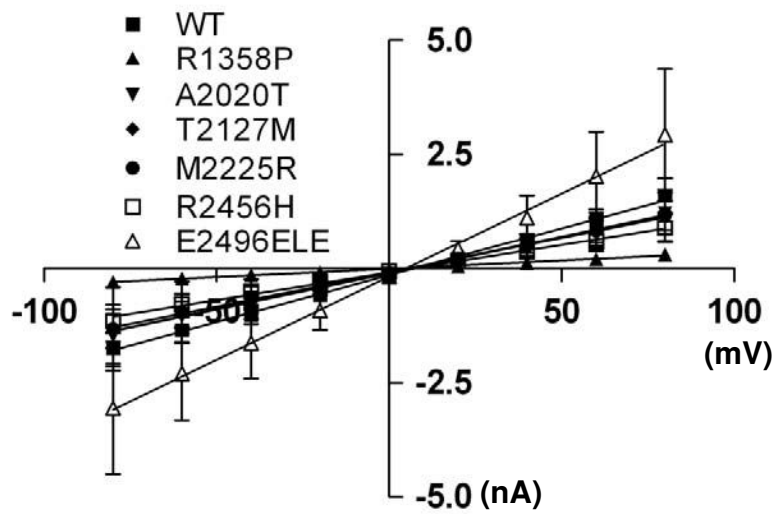
Supplementary Figures



Supplementary Figure S1. Distribution of coding variants in the general population compared to DHS mutations along PIEZO1 AA sequence. The human PIEZO1 coding variants have been extracted from the 1000 genomes variations table (gathering all variants from Ensembl entry ENSG00000103335). Only those predicted as probably deleterious/damaging by Polyphen or SIFT are presented here. Each bar represents the variations count (according to their type) per 100 Amino Acids block along the protein.



Supplementary Figure S2. Electrophoretic proof of the four different *PIEZO1* mutations found in DHS patients. Mutations have been confirmed (Family 1) or detected by conventional Sanger Sequencing. Alleles are represented by their symbols (R, S, or Y) at point mutation positions. Amino acid symbols and their positions are represented under the electropherograms, under the corresponding codons.



Supplementary Figure S3. I-V plot for DHS mutations in hPIEZO1 channels. Current-Voltage correlation of MA currents recorded from HEK293T cells transfected with wild-type or mutant hPIEZO1. Upon mechanical stimulation, currents were recorded in the whole-cell patch clamp configuration at holding potentials ranging from -80 mV to +80 mV. E_{rev} for WT: +7.2 mV (n = 8); R1358P: +3.0 mV (n = 5); A2020T: +6.8 mV (n = 6); T2127M: +6.0 mV (n = 10); M2225R: +5.4 mV (n = 4); R2456H: +7.3 mV (n = 6); E2496ELE: +5.0 mV (n = 4).

Supplementary Table S1. Exome analysis of the Family 1 pedigree

Variants	Affected 1	Affected 2	Shared
Coding (exons, exon boundaries, splicing)	9,815	9,515	4,375
Novel non-synonymous	2,407	2,346	231
Linkage location chr16 (hg19 : 84119815-89590825)	13	15	5
Probably Damaging (Polyphen2, SIFT, MutPred Server)	5	5	<i>PIEZO1</i> <i>CDT1</i>

Supplementary Table S2. Bioinformatics predictions of pathogenicity for the four different *PIEZO1* mutations identified

	R1358P	A2020T	T2127M	L2495_E2496dup
SIFT	Deleterious (score: 0.00)	Deleterious (score: 0.00)	Deleterious (score: 0.03)	NA
Polyphen	Probably damaging (score: 1.00)	Possibly damaging (score: 0.94)	Probably damaging (score: 1.00)	NA
Mutation Taster	disease causing (p=0.99)	disease causing (p=0.98)	disease causing (p=0.89)	Polymorphism (p=0.999)

SIFT and Polyphen only predict simple substitutions pathogenicity.

Supplementary Table S3. Genotypes of 10 SNPs at the *PIEZO1* gene for the 8 independent patients harbouring the E2496ELE mutation

LOCATION	intron 26		intron 27	exon 28	intron 29	exon 30	exon 39			intron 50	Exon 51
	rs2247261	rs79092320	rs4782429	rs4782430	rs78217493	rs11645197	rs4238686	rs35159887	rs75373715	rs3214887	E2496ELE
MAF	0.3219	0.057	0.115	0.1145	0.0618	0.163	0.3036	0.1644	unknown	unknown	-
WT GENOTYPE	GG	GG	CC	CC	CC	CC	TT	CC	-/-	-/-	-/-
CASE1	GA	GG	CC	CC	CC	CT	TC	CT	-/TCC	-/G	-/+
CASE2	GA	GG	CC	CC	CT	CT	TC	CT	-/TCC	-/G	-/+
CASE4	GG	GG	CC	CC	CC	TT	TT	TT	-/-	-/-	-/+
CASE7	GG	GG	CC	CC	CC	TT	TT	TT	-/-	-/-	-/+
CASE9	GG	GC	CC	CC	CC	CT	TC	CT	-/TCC	-/G	-/+
CASE10	GG	GG	TT	TT	TT	CC	TT	CC	-/-	-/-	-/+
FAMILY2	GG	GG	TT	CC	CT	CC	CC	CC	-/-	-/-	-/+
FAMILY3	GG	GG	TT	CC	TT	TT	TT	TT	-/-	-/-	-/+

Location of the selected SNP is given on top of the table. MAF: Minor Allele Frequency (Source dbSNP v137).

H0: GGCCCCTC - - -
H1: GGCCCTCT TCC G Dup
H2: GGCCCTTT - - Dup
H3: GGTTTCTC - - Dup
H4: GGTCTTTT - - Dup

Supplementary Table S4. Web tools addresses for bioinformatics sequence analyses

POLYPHEN	http://genetics.bwh.harvard.edu/pph2/
SIFT	http://sift.jcvi.org/
MUTATION TASTER	http://www.mutationtaster.org/
1000 Genomes	http://www.1000genomes.org/
Exome Variant Server	http://evs.gs.washington.edu/EVS/
Clustal Omega	http://www.ebi.ac.uk/Tools/msa/clustalo/
dbSNP	http://www.ncbi.nlm.nih.gov/projects/SNP/

Chapitre 3

Etude génétique de l'Anévrisme de l'Aorte Abdominale

1. Généralités sur la paroi artérielle

1.1 Histologie

La paroi de tous les vaisseaux sanguins à l'exception des capillaires se compose de trois couches principales (figure 1) : L'intima regroupe le revêtement pavimenteux endothélial et la couche sous-endothéliale composée de tissu conjonctif. La média est composée de la limitante élastique interne, d'une couche fibromusculaire d'épaisseur variable contenant du collagène, de l'élastine et des cellules musculaires lisses (ou CML), et de la limitante élastique externe. L'adventice, tissu conjonctif lâche en relation avec le tissu adipeux environnant, se compose essentiellement de faisceaux de fibres de collagène, de fibroblastes, de vaisseaux (artérioles, capillaires, veinules) incluant les vaisseaux lymphatiques, des terminaisons nerveuses et sympathiques, et des leucocytes résidents. Enfin, la substance fondamentale composée de glycosaminoglycanes et protéoglycanes plus ou moins hydratés est également présente dans la paroi artérielle, son rôle est encore peu connu (Dobrin, 1978).

1.1.1 L'endothélium

L'endothélium est un composant quantitativement mineur de la paroi, mais sur le plan fonctionnel il est impliqué dans de nombreux processus biochimiques, métaboliques et immunologiques. Il participe à l'angiogénèse et à la prolifération cellulaire par sa production de facteurs de croissance, module le calibre vasculaire en sécrétant, entre autres, l'endothéline vasoconstrictrice et le NO vasodilatateur. Il régule d'autre part la balance procoagulation-anticoagulation et dégrade les lipoprotéines par sa lipase membranaire. En réponse à différents stimuli particulièrement oxydants, il exprime les récepteurs de surface nécessaires à la diapédèse leucocytaire et régule la perméabilité vasculaire.

1.1.2 La média

Les trois composants principaux de la média sont l'élastine, le collagène fibrillaire, et les cellules musculaires lisses.

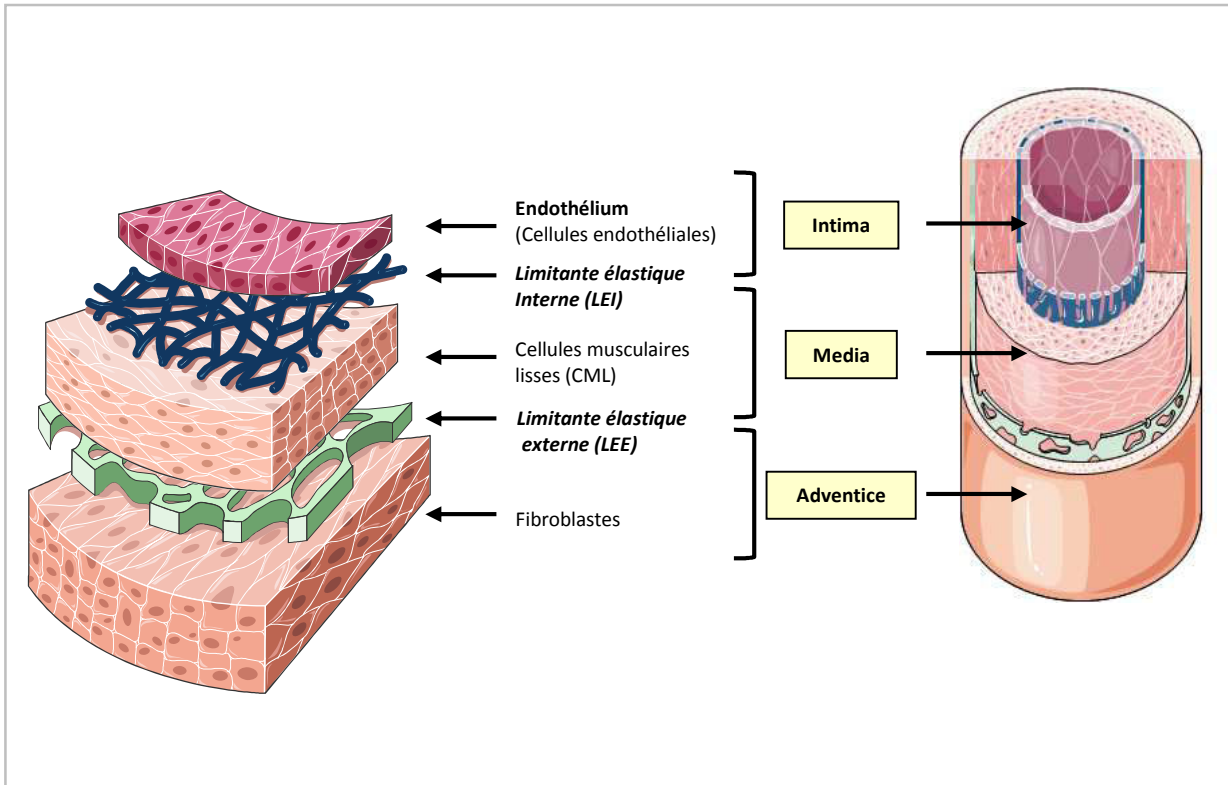


Figure 1: structure de la paroi artérielle.

Les fibres élastiques sont composées d'élastine, matériel hydrophobe insoluble enrobant des microfibrilles de fibrilline et de fibuline, synthétisée par les cellules musculaires lisses. Le précurseur de l'élastine ou tropoélastine se polymérise puis se complexifie en structure hélicoïdale comparable à un ressort que l'on peut étirer. Les complexes se déposent le long des microfibrilles qui représentent le squelette de la fibre élastique. L'élastine est synthétisée dans la vie fœtale et dans la première année de vie puis son expression diminue fortement chez l'adulte. Sa demi-vie très longue (70 ans, Mesh *et al.*, 1992; Powell *et al.*, 1992) permet le maintien des propriétés élastiques des artères élastiques jusqu'à un âge avancé. Sa protéolyse n'est jamais spontanée, mais implique l'activation de protéases élastolytiques. Les fibres élastiques ont une organisation variable selon le type d'artère, élastique ou musculaire (cf infra).

Le collagène fibrillaire (de type I et III) est arrangé en fibres d'orientation circonférentielle (Wolinsky et Glagov, 1964), avec une proportion de 2 pour 1 entre les quantités de collagène de type I et III respectivement. Les fibres collagènes sont composées d'une unité de base, le tropo-collagène, molécule insoluble composée de trois chaînes d'acides aminés agencés en triple hélice. Les unités de tropo-collagène s'alignent parallèlement les unes aux autres de façon répétitive pour former des fibrilles puis d'épais faisceaux extrêmement résistants à la traction. Contrairement à l'élastine, le collagène est constamment renouvelé : sa demi-vie est de deux à trois jours et sa synthèse est stimulée par les contraintes mécaniques s'exerçant sur le tissu conjonctif.

Ainsi, fibres élastiques et fibres de collagène confèrent élasticité et résistance à la paroi artérielle, mais les modalités de leur organisation respective permettent l'obtention de propriétés très distinctes selon le type d'artère, élastique ou musculaire (cf infra).

La composante cellulaire de la paroi est représentée par les cellules musculaires lisses (CML). Dans la plupart des artères, ces cellules sont majoritairement orientées de façon circonférentielle, conférant le rôle vasomoteur à la média (changement de calibre, résistances périphériques).

La média est une couche avasculaire en condition physiologique, excepté le tiers externe de la média aortique. Elle est donc un site immun privilégié, inaccessible aux leucocytes de par l'absence de microcirculation en conditions physiologiques.

La morphologie et la fonction de cette couche intermédiaire permet la distinction entre deux sous-types d'artères.

Les **artères de type élastique** sont les artères de gros calibre (artères de conductance) telles que l'aorte, le tronc brachio-céphalique, les artères carotides communes, ainsi que la partie proximale des artères sous-clavières, vertébrales et iliaques communes. Leur aspect macroscopique jaunâtre atteste de leur contenu important en fibres élastiques (35% du poids sec, Campa *et al.*, 1987).

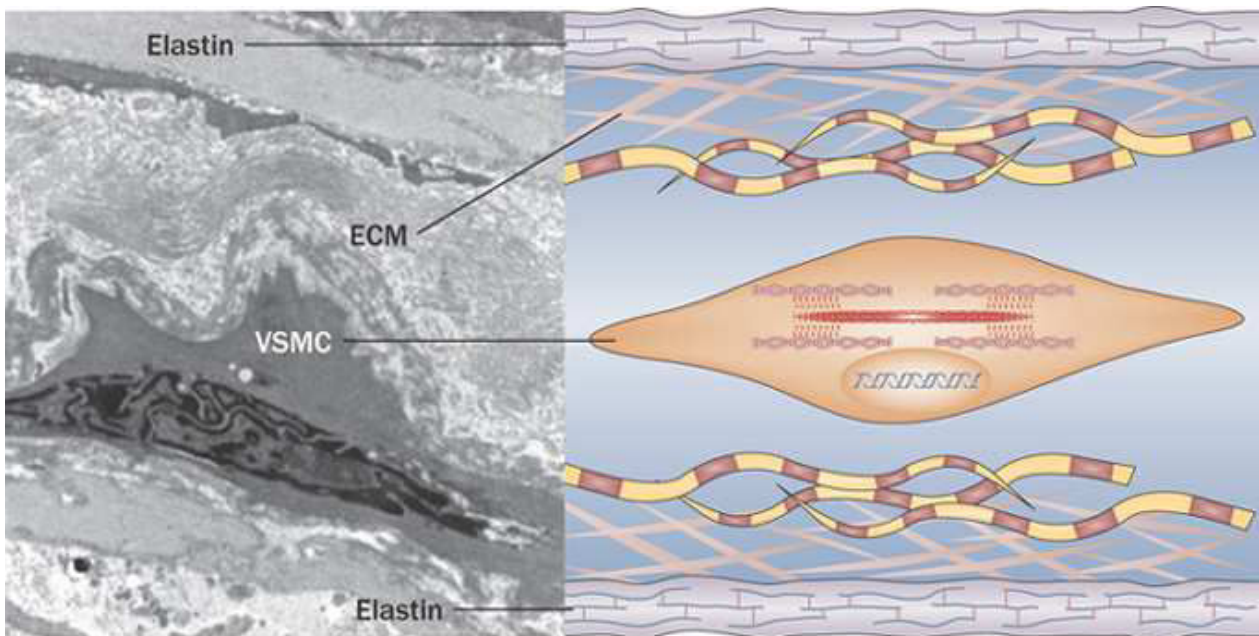


Figure 2: Aspect en microscopie électronique d'une unité lamellaire de la paroi aortique. Une cellule musculaire lisse est enchâssée entre eux couches d'élastine et environnée de protéines matricielles contenant des microfibrilles et des protéoglycanes . Entre les unités lamellaires s'intercalent de larges faisceaux de collagène. ECM: Matrice extracellulaire. vSCM: CML vasculaire. D'après El Hamamsy et al, 2009.

En microscopie, on observe une média riche en fibres élastiques arrangées en lames élastiques fenêtrées concentriques, d'autant plus denses qu'elles sont proches de l'intima. Le nombre de lames élastiques dans la média des artères de distribution étant proportionnel au diamètre et à la tension circonférentielle moyenne du vaisseau, quelque soit l'espèce animale étudiée pour ces paramètres, le concept d'unité lamellaire s'est imposé en tant qu'unité fonctionnelle et structurelle de la média des artères de conductance (figure 2). Les CML y sont peu nombreuses, orientées de façon concentrique ou légèrement hélicoïdale, et sont accolées les unes aux autres le long du grand axe du vaisseau pour former des bandelettes uni- ou bi-cellulaires plates maintenues par une lame basale commune et un fin réticule composé de fibrilles de collagène. Au-dessus et en-dessous de cette couche cellulaire, des fibres élastiques intimement associées entre elles, d'orientation également circonférentielle, tapissent étroitement les groupes cellulaires. La principale fonction physiologique des CML dans les artères de conductance est la synthèse de matrice extracellulaire. Elle assure à elle seule la fonction de contention du sang dans l'espace circulatoire.

Cette alternance de structures « élastine-CML-élastine » est considérée comme la structure de base de la paroi des artères élastiques. Ces couches successives sont appelés fascicules musculo-élastiques (ou FME, figure 3). Elles sont séparées les unes des autres par une zone acellulaire étroite composée d'épais faisceaux de collagène ondulés et orientés circonférentiellement.

Les artères de type musculaire sont les artères de distribution. De calibre moindre que les précédentes, elles présentent une média riche en CML arrangées en épais anneaux concentriques (figure 4), et sont pauvres en fibres élastiques et collagènes. Contrairement aux artères élastiques, les anneaux musculaires sont continus et d'une épaisseur de plusieurs cellules accolées les unes aux autres, il n'y a donc pas d'organisation en unités lamellaires ou en fascicules musculo-élastiques. Les artères de distribution et les artérioles ont comme principale fonction d'assurer et de moduler la résistance à l'écoulement du sang (fonction vasomotrice).

Les artères élastiques ont une grande capacité de distension, longitudinale (élongation) et circonférentielle (dilatation) en réponse à l'onde de pouls générée par les battements cardiaques. Les artères musculaires ont quant à elles une grande capacité de contraction et de réduction du diamètre artériel, en réponse à divers stimuli.

1.1.3 L'adventice

L'adventice est une couche formée d'un tissu conjonctif lâche comprenant principalement des fibroblastes, quelques CML, des protéoglycanes, et des fibres de collagène présentant une orientation oblique spiralée. Elle entoure et protège l'artère, et la rend solidaire des organes environnants. La synthèse des éléments de la matrice, et particulièrement des collagènes et de l'élastine, est assurée par les fibroblastes adventitiels.

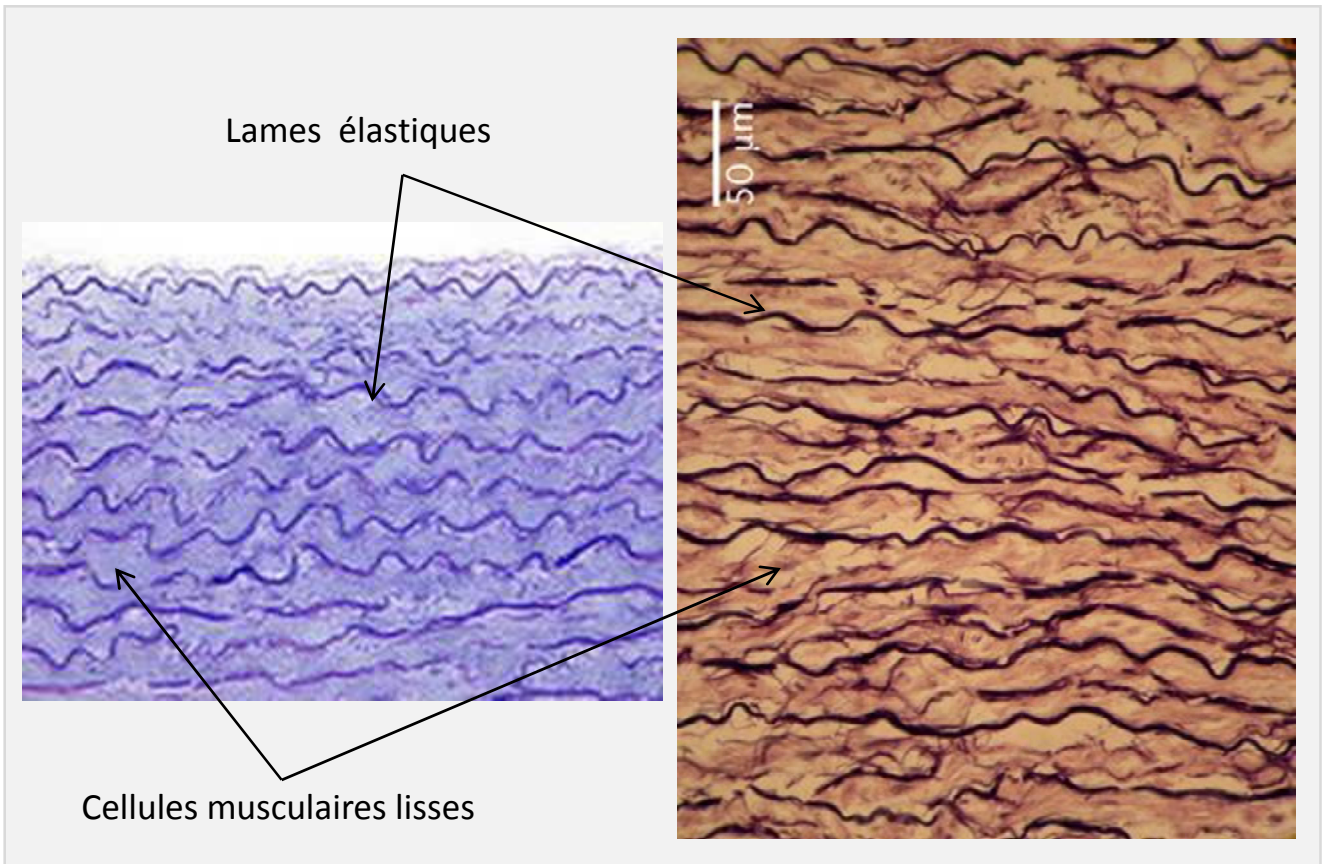


Figure 3: Aspect de la media aortique après coloration spécifique des fibres élastiques à l'Orcéine. On observe l'alternance des couches définissant les fascicules musculoélastiques caractéristiques de la media d'une artère élastique.

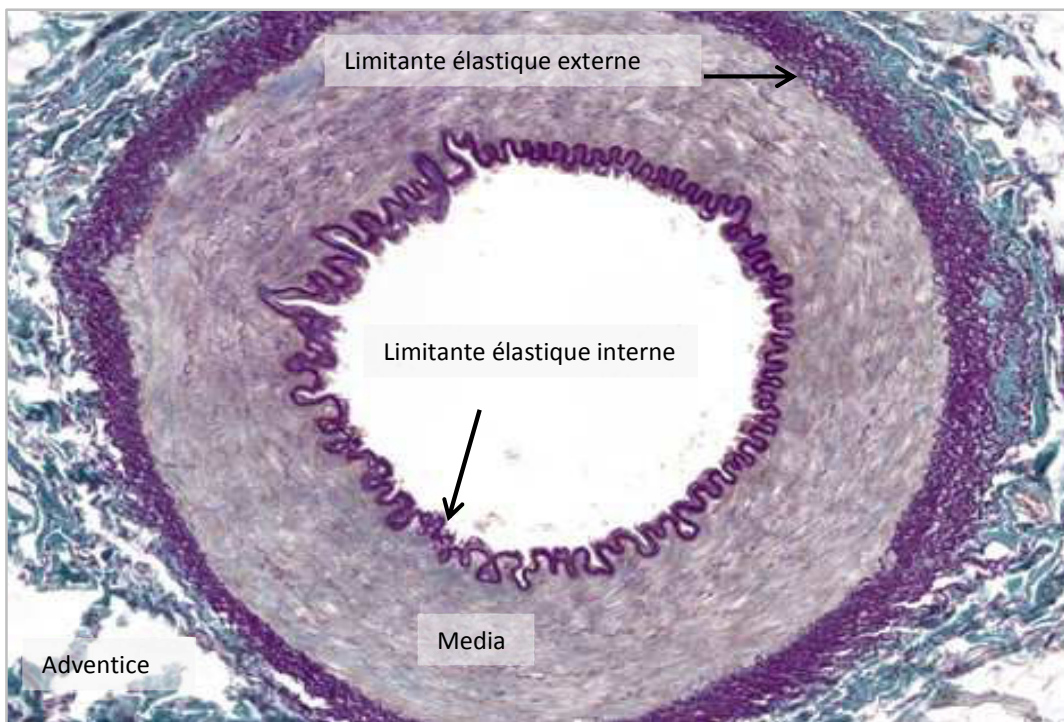


Figure 4: Section d'une artère musculaire (coloration au trichrome vert). L'élastine est colorée en rose foncé, les fibres de collagène en vert, et les cellules musculaires lisses apparaissent en rose pâle. La média est caractérisée par son épaisseur, contenant essentiellement des CML. L'adventice est riche en fibres de collagène.

Dans les artères de gros calibre, l'adventice contient également les *vasa vasorum*, petits vaisseaux drainant les produits de la média (sécrétion médiale ou transformation médiale des produits circulants).

1.1 Métabolisme de la paroi artérielle

La matrice extracellulaire est donc un réseau dynamique de macromolécules qui subit un remodelage constant, résultat d'un équilibre entre synthèse de protéines matricielles par les fibroblastes et CML, et dégradation. Cette dernière est physiologique et s'observe dans divers processus tels le développement, la réparation des tissus et l'angiogenèse, mais peut être associée à un processus pathologique comme le cancer, l'athérosclérose, l'arthrite (Fingleton, 2007).

Plusieurs types de protéinases sont impliqués dans la dégradation de la matrice extracellulaire, mais les enzymes principales appartiennent à une famille d'endopeptidases dépendantes du zinc, les métalloprotéinases de la matrice (Visse et Nagase, 2003).

1.2.1 Les MMP

Le groupe des métalloprotéinases matricielles (MMP) est constitué chez l'homme de 23 membres, capables de dégrader conjointement l'ensemble des protéines de la matrice extracellulaire. Elles sont en général peu exprimées dans les tissus sains, mais leur expression augmente lors des processus de remodelage tissulaire physiologique ou pathologique sous l'effet de cytokines, de facteurs de croissances ou d'hormones capables de moduler, par transcription, l'expression de leur gènes (Nagase et Woessner, 1999). La fonction principale des MMP est donc la dégradation et le renouvellement permanent du tissu. Cependant, il est désormais reconnu que les modifications induites par les MMP altèrent également les interactions cellules-matrice et cellules-cellules et permettent la libération de facteurs de croissance, les rendant ainsi disponibles pour leurs récepteurs. Les MMP peuvent être classées en six classes distinctes, sur la base de leur structure primaire et de leur spécificité vis-à-vis de certains substrats protéiques (tableau 1). Elles sont synthétisées le plus souvent sous forme inactive de zymogènes (pro-MMP) par de nombreuses cellules mésenchymateuses et immunitaires comme les macrophages. Leur activation requiert le clivage par les sérine protéases du pro-domaine qui maintient la protéine dans un état inactif.

Les MMPs ont été impliquées dans les maladies cardio-vasculaires telles que les pathologies anévrysmales, l'athérosclérose, l'hypertension artérielle pulmonaire, des maladies digestives tumorales ou chroniques inflammatoires (Maladie de Crohn). Des études menées *ex vivo* et en modèle murin ont montré que les

CLASSE	NUMÉROTATION	PRINCIPAUX SUBSTRATS MATRICIELS
Matrilysines	MMP7, MMP26	Fibronectine, gélatine, élastine, collagène de type IV, laminine
Collagénases interstitielles	MMP1, MMP8, MMP13	Fibronectine, gélatines, collagènes de types I,II,III, IV, VI, VII, X
Stromélysines	MMP3, MMP10, MMP11	Glycoprotéines , protéoglycanes, fibronectine
Gélatinases	MMP2, MMP9	Gélatine , élastine, fibronectine collagène de type I, IV, V, VII, XI
Métallo-protéinases de type membranaire (MT-MMPS)	MMP14, MMP15, MMP16, MMP17, MMP24, MMP25	Gélatine , fibronectine, collagènes de type I, II, III, IV
Autres (groupe hétérogène)	MMP12, MMP18, MMP19, MMP20, MMP21, MMP22, MMP23, MMP27, MMP28	Laminine, élastine, collagène de type IV, fibronectine, gélatine

Tableau 1: Les différentes classes de métalloprotéinases et leurs substrats

MMP-2 et MMP-9 sont impliqués dans les ruptures du ventricule gauche post-infarctus du myocarde, et dans le développement d'anévrismes de l'aorte abdominale (Hayashidani et al., 2003, Longo et al., 2002). Le rôle clé de ces enzymes dans le remodelage matriciel, et dans la libération de facteurs de croissance dépendant des cytokines nécessite une régulation fine qui s'effectue au niveau transcriptionnel, par activation protéolytique, ou par l'induction d'inhibiteurs endogènes spécifiques : les inhibiteurs tissulaires des MMP ou TIMPs (Hayakawa et al., 1994).

1.2.2 Les TIMP

Les inhibiteurs endogènes des MMPs forment une famille de quatre protéines TIMP (TIMP-1 à 4). Leur domaine N-terminal contient une unité indépendante pouvant inhiber l'activité des MMPs. En dehors de leurs propriétés inhibitrices, les TIMPs peuvent également agir comme des facteurs de croissance et stimuler, *in vitro*, la croissance de toutes les cellules de la paroi artérielle (Hayakawa et al., 1994). Elles participent donc à la stabilisation de la matrice mais également à la cicatrisation tissulaire.

2. L'aorte

L'aorte est l'artère principale de l'organisme assurant l'ensemble de la vascularisation artérielle systémique. Elle possède des propriétés biomécaniques qui dépendent des caractéristiques structurelles de sa paroi (artère élastique). Elle naît de la chambre de chasse du ventricule gauche, au niveau de la valve aortique. Après un court trajet ascendant, elle transite du thorax vers l'abdomen en donnant naissance à l'ensemble des collatérales artérielles du corps humain. Elle se termine, dans l'espace retro-péritonéal, en donnant naissance à deux collatérales terminales principales au niveau de la bifurcation iliaque.

2.1 Anatomie descriptive

Quatre segments sont habituellement décrits et diffèrent du point de vue anatomique et des pathologies associées (figure 5).

Le segment 1 ou aorte thoracique ascendante, va de la valve aortique cardiaque jusqu'au pied du Tronc Artériel Brachio-Céphalique (TABC) sur 7 à 10 cm en moyenne. Le segment II ou crosse aortique, va d'avant en arrière dans le médiastin supérieur et est à l'origine de trois artères collatérales : le TABC, l'artère carotide primitive gauche et l'artère sous-clavière gauche. Le segment III ou aorte thoracique descendante, va de l'artère sous-clavière gauche jusqu'au hiatus aortique diaphragmatique où elle entre dans la cavité abdominale. Le segment IV ou aorte abdominale, va dans la cavité abdominale après son passage trans-diaphragmatique. L'aorte abdominale donne de multiples artères collatérales telles que le tronc cœliaque, les artères mésentériques, rénales et iliaques primitives droite et gauche (collatérales terminales).

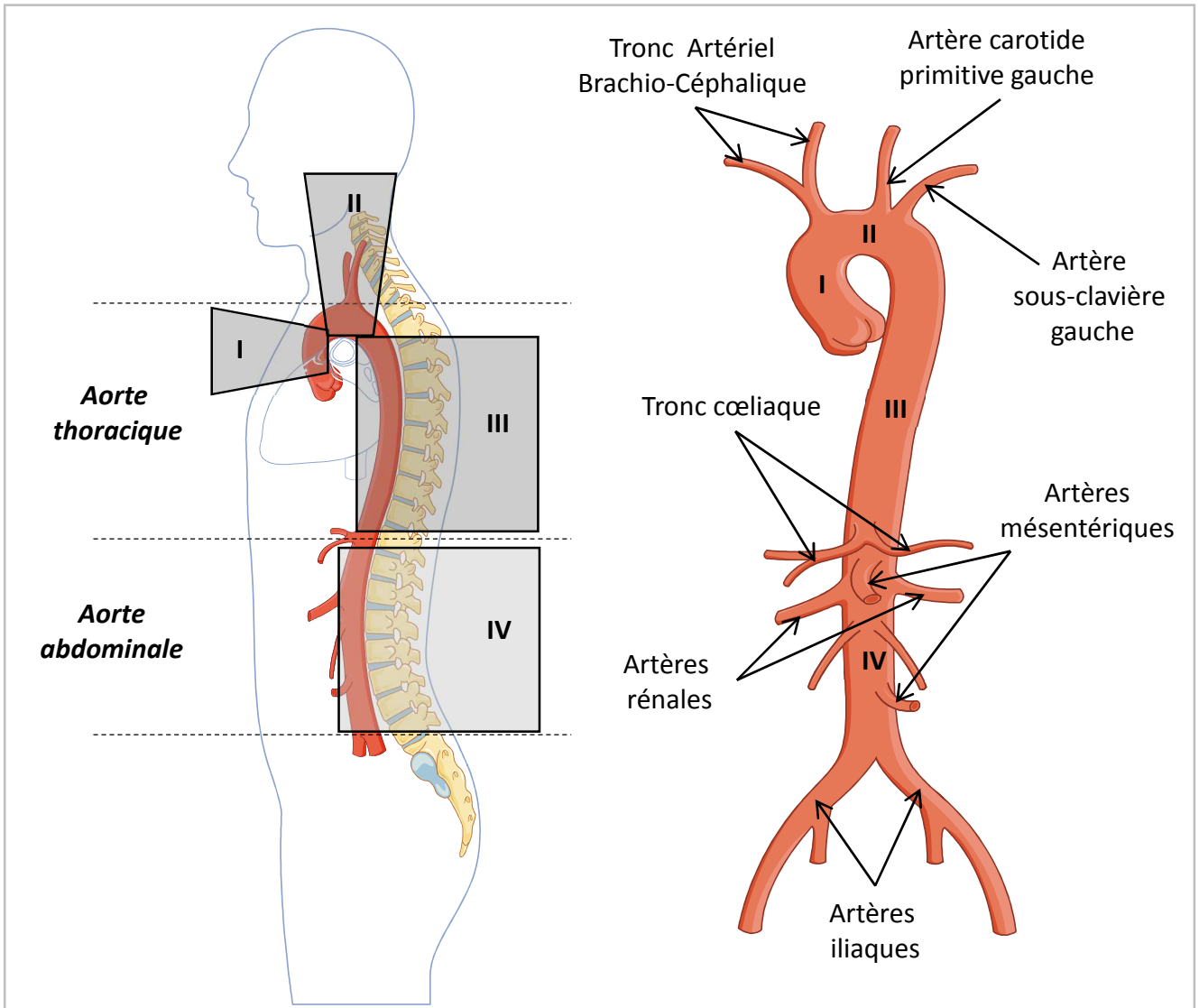


Figure 5: Les quatre segments de l'aorte et leurs collatérales.

2.2 Pathologie aortique anévrysmale et occlusive

La pathologie aortique courante est dominée par deux groupes d'affections aux mécanismes divergents, avec d'une part une diminution du diamètre de la lumière artérielle (pouvant aller jusqu'à l'occlusion) dont la cause principale est l'athérome, et d'autre part, une dilatation de la lumière artérielle représentée par l'anévrysme. Ces deux mécanismes ont en commun un déséquilibre de l'homéostasie de la paroi artérielle.

Les anévrysmes, définis comme des augmentations localisées du diamètre vasculaire, sont associées à la perte permanente du parallélisme des parois vasculaires et mettent en jeu les trois couches du vaisseau. Lors de ce processus, la paroi se fragilise et la composition de sa média est considérablement remaniée et présente une réduction et une fragmentation des fibres élastiques et de collagène. Ces modifications altèrent les propriétés mécaniques de l'artère qui perd progressivement sa capacité à lutter contre la distension. Lorsque la force d'étirement exercée par la pression artérielle devient supérieure à la force de cohésion des fibres élastiques et de collagène de la paroi, l'anévrysme se rompt.

D'un point de vue anatomique on décrit deux types d'anévrysmes.

2.2.1 Anévrysme de l'aorte ascendante

Les anévrysmes de l'aorte ascendante (ou TAA pour Thoracic Aortic Aneurysms) sont rares et peuvent être héréditaires. Dans cette dernière catégorie de pathologies, on distingue classiquement les formes syndromiques et non syndromiques.

Les formes syndromiques sont fréquemment liées à une altération d'un gène impliqué dans le métabolisme ou la structure de la matrice extracellulaire. Les pathologies les plus représentatives en sont la maladie de Marfan liées aux mutations de la fibrilline 1, les syndromes de Loeys-Dietz (gènes *TGF β 2* et *3*, *TGF β R1* et *2*, *SMAD2* et *3*) impliquant une altération de la voie du TGF β , et le syndrome d'Ehlers-Danlos vasculaire (gène *COL3A1*), conséquence d'un défaut structurel du collagène de type III. Le phénotype dans ces pathologies est souvent étendu à d'autres structures impliquant une composante matricielle importante comme les articulations ou la peau (Cook et al., 2015; Van Laer et al., 2014).

Les formes non syndromiques sont majoritairement en rapport avec une dysfonction de l'appareil cellulaire contractile de l'artère : TAA liés aux mutations du gène de l'actine (*ACTA2*) ou de la myosine (*MYH11*)

spécifiques des CML, ou de protéines intervenant dans le métabolisme de ces protéines contractiles (*MYLK*, *PRKG1*, Guo et al., 2013, 2007; Wang et al., 2010; Zhu et al., 2006). Cependant les limites entre formes syndromiques et non syndromiques ont tendance à s'estomper et de nouvelles formes syndromiques d'atteinte artérielle liée à *ACTA2* (Munot et al., 2012), ou des formes purement aortiques de maladies liées à *SMAD3* ou *TGF β 2* ont récemment été décrits (Gago-Díaz et al., 2014; Regalado et al., 2011).

2.2.2 Anévrisme de l'aorte descendante

Les anévrysmes de l'aorte descendante sont largement dominés par les anévrysmes de l'aorte abdominale (AAA) le plus souvent localisés entre les artères rénales et la bifurcation iliaque. On parle alors d'anévrisme infrarénal. L'épidémiologie et les processus physiopathologiques en jeu rendent cette affection entièrement distincte de l'anévrisme de l'aorte ascendante. Ils sont détaillés dans la section suivante.

3. L'anévrisme de l'aorte abdominale

3.1 Définition, diagnostic et histoire naturelle

Un anévrisme est défini comme une dilatation permanente, localisée, segmentaire, avec perte du parallélisme des bords, d'une artère dont le diamètre est augmenté d'au moins 50 % par rapport à son diamètre normal (Frydman et al., 2003; Verloes et al., 1995).

Dans la plupart des études épidémiologiques l'AAA est défini comme une dilatation localisée de l'aorte abdominale d'un diamètre antéro-postérieur supérieur à 30 mm, correspondant au critère d'augmentation de 50% de son diamètre (figure 6). L'AAA est couramment diagnostiqué par mesure du diamètre de l'aorte *via* une échographie abdominale.

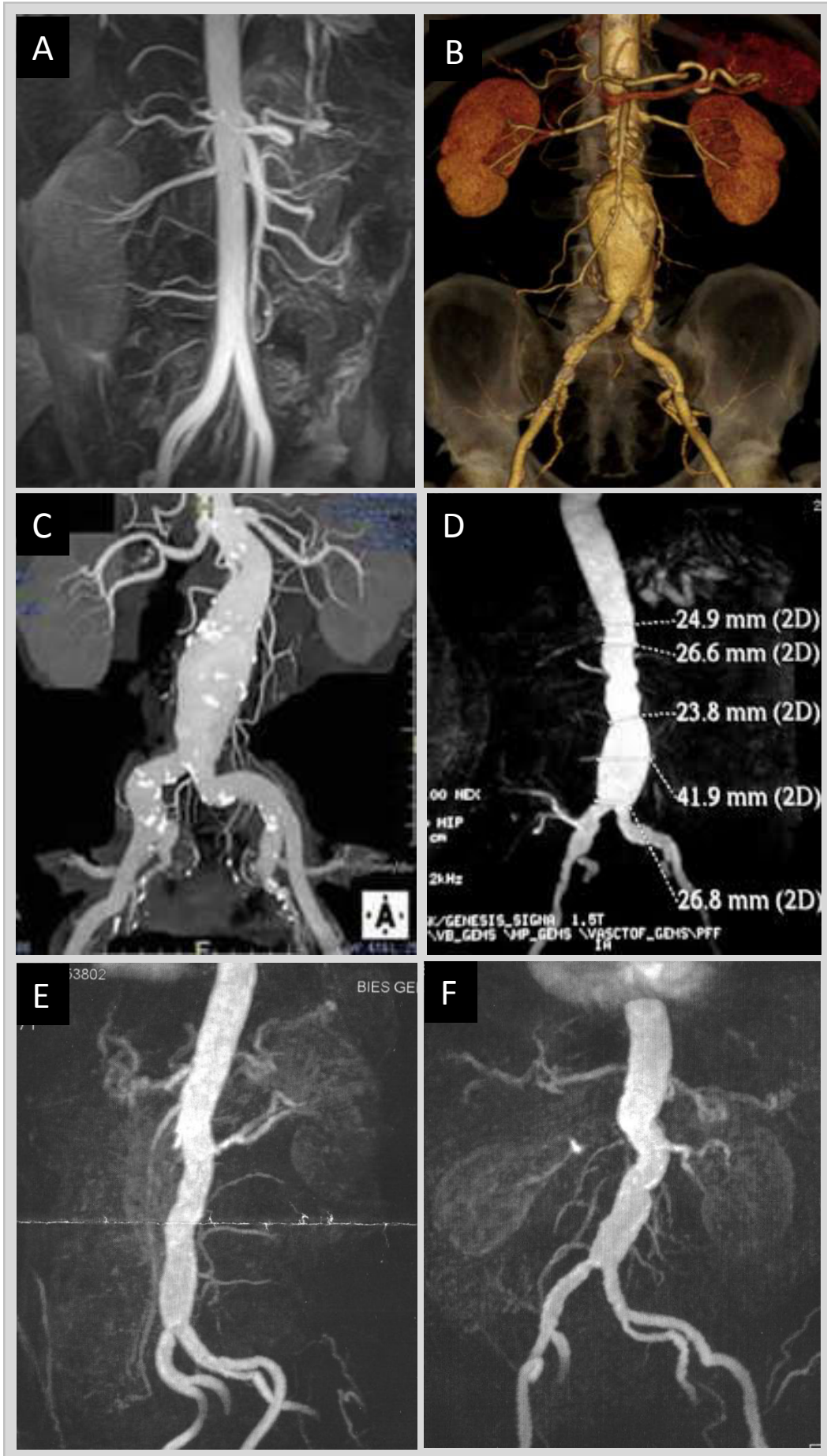


Figure 6: Aspect normal et divers types de dilatation de l'aorte abdominale en imagerie. A: aorte normale en angioIRM. B: anévrisme localisé de l'aorte infrarénale, par angioscanner avec reconstruction. C,D: aspects de méga-dolicho-aorte présentant des lésions diffuses. E,F: aspect en angioIRM de l'individu II-5 de la famille étudiée.

On distingue l'anévrisme de l'artériomégalie qui consiste en un élargissement artériel diffus concernant plusieurs segments artériels (figure 6, Band et al., 1972). On parle de dolicho-méga-aorte ou de méga-dolicho-aorte lorsque l'élargissement siège sur l'aorte. Elle est cependant fortement corrélée à l'AAA familial (Lawrence et al., 1998) : parmi les apparentés au premier degré d'un patient avec mégadolicho-aorte, 36% sont porteurs d'un anévrisme aortique, indiquant une probable parenté entre ces deux pathologies, l'une focale et l'autre plus diffuse.

Le diagnostic précoce de l'AAA est critique, la rupture d'anévrisme étant mortelle dans 80-90% des cas, contre 5% en cas d'intervention chirurgicale programmée (Karkos et al., 2000). La rupture d'anévrisme de l'aorte abdominale est la 13^{ème} cause de décès aux USA, représentant 1 à 2% de tous les décès chez les hommes de plus de 65 ans (Wilmink et Quick, 1998). Des chiffres similaires sont retrouvés en Angleterre (Law, 1998). Le risque de rupture est évalué entre 2,5 et 10% des AAA (Glimåker et al., 1991; Perko et al., 1993). L'histoire naturelle de l'AAA est marquée d'une part par un lien significatif entre sa taille et sa vitesse de croissance (la taille augmente de façon exponentielle avec le temps, (Stonebridge et al., 1996), et par l'absence de symptôme associé pour les anévrysmes de moins de 6 cm, taille à laquelle le risque de rupture est majeur. Cette pathologie constitue donc un véritable problème de santé publique. Dans de nombreux pays des stratégies de dépistage systématique ont été mises en place, notamment chez les hommes fumeurs de plus de 65 ans (Scandinavie, USA et Angleterre), et pourraient permettre de réduire la mortalité de 50% dans ce groupe (Fleming et al., 2005; U.S. Preventive Services Task Force, 2005).

3.2 Épidémiologie et facteurs de risque de l'AAA

La prévalence de l'AAA, estimée par dépistage échographique varie entre 1,3% et 11,4% chez l'homme et entre 0,6% et 2,1% chez la femme. Cette prévalence dépend fortement de l'âge des patients. Elle est faible avant 55 ans puis augmente de façon substantielle à partir de cet âge (figure 7, Bickerstaff et al., 1984). Si la prévalence de l'AAA est comparable dans la plupart des pays occidentaux, elle diffère fortement de celle mesurée au Japon qui semble très faible (Takei et al., 1995).

L'incidence de l'AAA est de 11,3 à 117,2 pour 100 000 années-personnes chez les hommes et de 2,2 à 33,9 pour 100 000 années-personnes chez les femmes (Castleden et Mercer, 1985). Ces chiffres très imprécis témoignent de la variabilité de l'usage des critères diagnostics de l'AAA et donc de sa définition, et de la population étudiée (figure 8, Bloomer et al., 2012). Diverses études transversales, prospectives et autopsiques ont été réalisées en populations caucasienne, afro-américaine et asiatique (Jackson et Rutter, 2001; Salem et al., 2009; Yii, 2003) indiquant que 1,9% à 18,5% des hommes de plus de 60 ans ont un

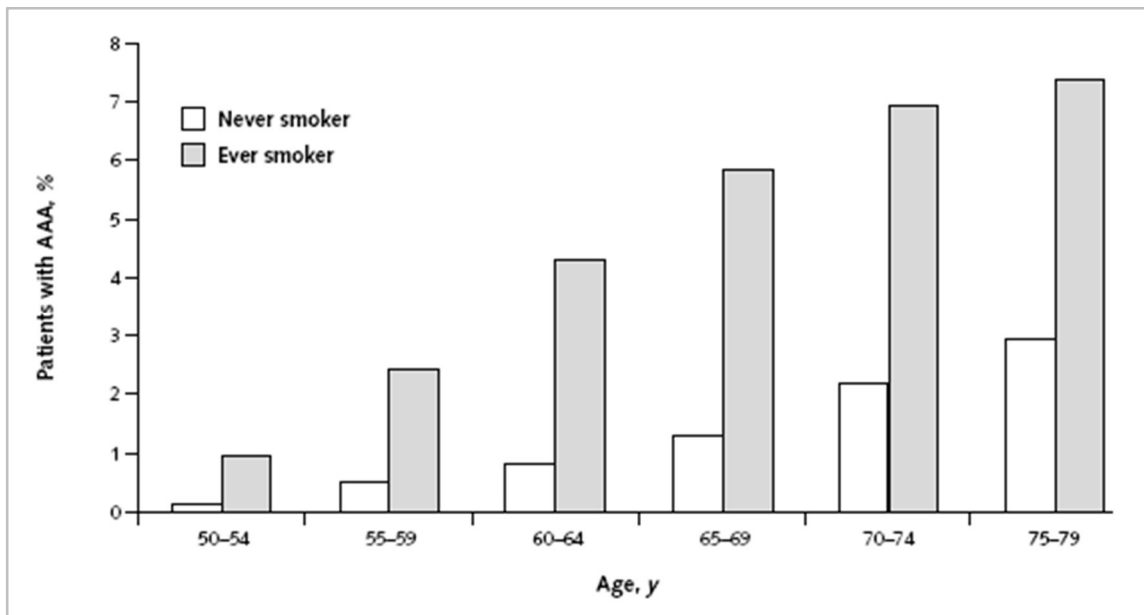


Figure 7: Prévalence de l'AAA entre 50 et 80 ans dans la population américaine, en fonction de l'exposition au tabac. D'après Fleming et al, 2005.

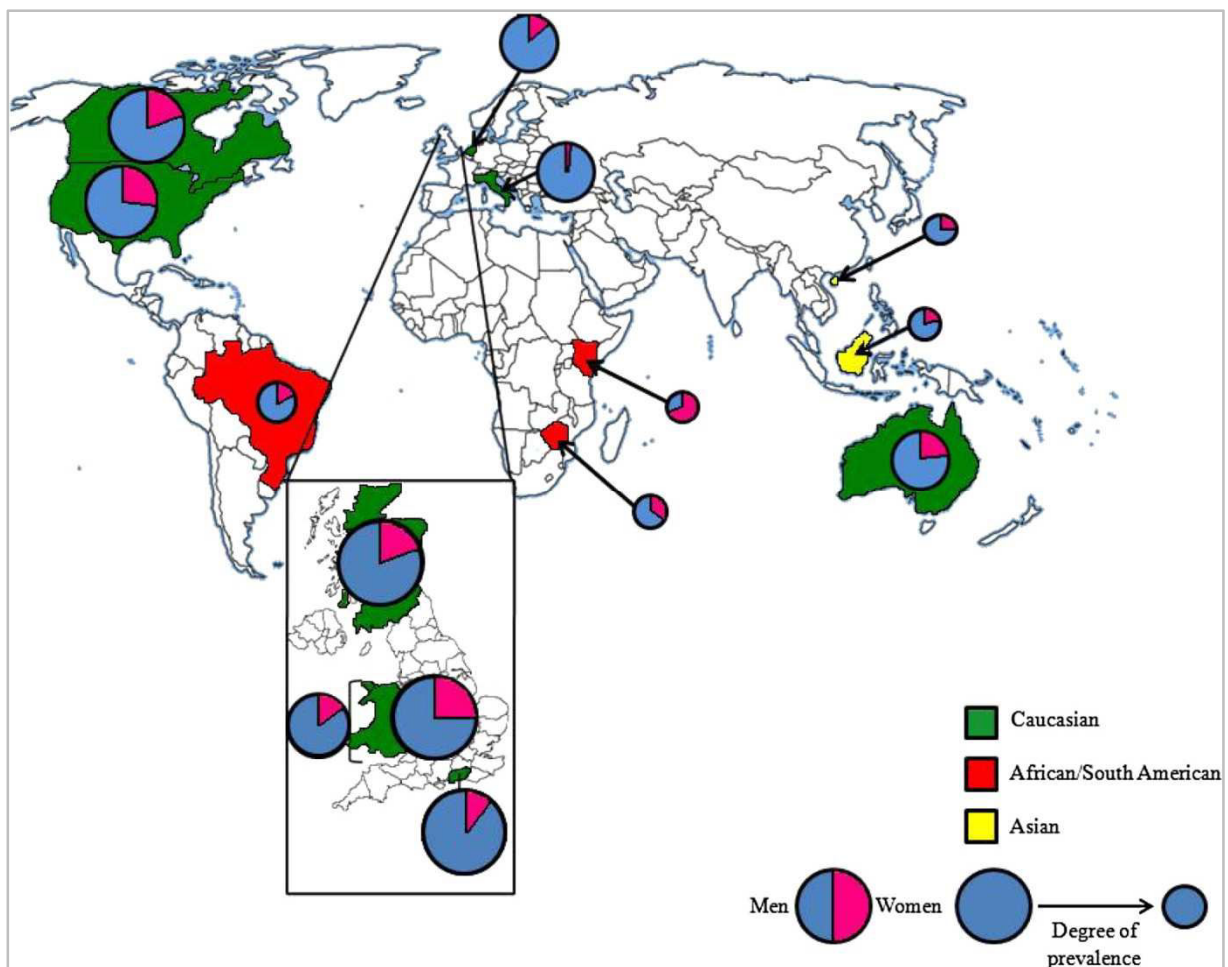


Figure 8: Répartition géographique de l'incidence rapportée de l'AAA. La prévalence de l'AAA dans chaque pays est indiquée par la taille du cercle. Le sex-ratio est indiqué par le contenu du cercle (bleu = hommes et rose = femmes).

diamètre de l'aorte abdominale de plus de 29 mm dans les pays industrialisés et que l'AAA est beaucoup plus fréquent chez l'homme que la femme (Ogata et al., 2005a). Les facteurs de risque de l'AAA sont très semblables à ceux des maladies cardiovasculaires, athéromateuses : ils comprennent l'âge, l'existence d'une maladie vasculaire périphérique (majoritairement artérite des membres inférieurs), le tabagisme, l'hyperlipidémie, l'hypertension, le sexe masculin, les antécédents familiaux (voir section 3.4 Composante héréditaire de l'AAA), l'origine caucasienne (Harthun, 2008; Karkos et al., 2000; van Vlijmen-van Keulen et al., 2002). A l'inverse, le diabète, qui est un facteur de risque connu dans la maladie coronarienne, est inversement corrélé à l'AAA (Naydeck et al., 1999). L'origine asiatique semble être un facteur protecteur de la maladie (van Vlijmen-van Keulen et al., 2002).

L'athérosclérose : hypothèse controversée

Historiquement, l'AAA était considérée comme une maladie cardiovasculaire liée à l'athérome au même titre que la coronaropathie (Madaric et al., 2005). Aujourd'hui, la pathogénie de ces deux maladies apparaît, au moins en partie, différente (voir section 3.5 Aspects histologiques et pathogénèse de l'AAA), bien que la co-existence de la maladie coronarienne et de l'AAA est relativement élevée : 15 à 50% des individus semblent porteurs des deux pathologies, (Madaric et al., 2005; Yusuf et al., 2004). Le risque relatif d'être porteur d'un AAA est estimé à 2,2 en cas de maladie coronarienne sous-jacente (Wilmink et Quick, 1998). Cette association peut s'expliquer partiellement par le chevauchement des facteurs de risque cardiovasculaires classiques entre maladie coronarienne et AAA.

Cependant l'AAA semble être un cas particulier en ce sens que l'athérome pourrait représenter une condition nécessaire mais non suffisante au développement anévrysmal : la maladie athéromateuse symptomatique (sténose ou occlusion) est de fréquence similaire à l'AAA (2 à 3% chez les hommes et 1 à 2 % chez les femmes de plus de 60 ans), mais elle est largement plus élevée si l'on tient compte des formes asymptomatiques (24% des patients de plus de 60 ans exempts d'anévrysme et d'AVC ischémique, Amarenco et al., 1994). Seuls 9% à 16% des patients atteints d'athérosclérose de l'aorte abdominale vont développer un AAA, qui peut cependant survenir secondairement à une atteinte occlusive de l'aorte abdominale (Koch et al., 1990). Ceci pourrait être expliqué par les caractéristiques hémodynamiques différentes de l'aorte sous-rénale, des artères coronaires, des carotides et des artères des membres inférieurs.

Le tabac

Le tabagisme est un facteur de risque majeur et bien documenté dans le développement de l'AAA : l'étude de Singh et coll. révèle que le risque est d'autant plus élevé que l'intoxication est ancienne, et ce de façon

indépendante du nombre de paquets-année consommés. Pour une exposition de 40 ans et plus, le risque relatif est de 8 ($p < 0,001$, IC : 5-12,6) contre 1,4 pour une exposition de moins de 20 ans ($p < 0,001$, IC : 0,8-2,4, Singh et al., 2001).

Les dyslipidémies

Les taux élevés de cholestérol LDL et des triglycérides (TG) sont des facteurs de risque connus (Yusuf et al., 2004), ainsi que le faible taux de cholestérol HDL : le risque relatif évolue de façon linéaire avec la baisse du taux de HDL, et l'élévation du taux de LDL et de TG. Plus précisément, un taux de HDL cholestérol inférieur à 1,20 mmol/L confère un risque relatif de 2 ($p < 0,05$, IC : 0,22-0,51), et une élévation du taux de LDL, par tranche de 0.1 mmol/L augmente le risque d'AAA de 3% (Hobbs et al., 2003; Singh et al., 2001).

L'hypertension artérielle

L'hypertension artérielle (HTA) est un facteur de risque modéré de l'AAA : l'HTA systolique (PAS > 150 mm Hg) est associée à une augmentation de 1,5 fois du risque d'AAA chez les femmes selon Singh et coll. Un risque relatif du même ordre est retrouvé avec une élévation de la PA diastolique (PAD > 90 mmHg, Blanchard et al., 2000). L'hypertension est aussi impliquée dans l'accélération de la croissance anévrysmale et le risque de rupture (Karkos et al., 2000).

Le sexe masculin

Le ratio homme/femme est très élevé dans cette pathologie (6 pour 1) et surpasse de loin celui de la maladie coronarienne (qui est seulement de 2 pour 1). Le sexe masculin a été identifié comme l'un des facteurs de risque les plus importants de l'AAA à travers différents groupes ethniques dans de nombreuses enquêtes à travers le monde. Les causes possibles de ce dimorphisme sexuel sont nombreuses (différences anatomiques, hormonales, de mode de vie impliquant notamment les facteurs de risque cardiovasculaires, la régulation du système immunitaire, ...).

Le diabète

Le diabète a été identifié comme un facteur protecteur de l'AAA dans plusieurs études, avec un odds-ratio cependant modéré (0,68 à 0,78, Jamrozik et al., 2000; LaMorte et al., 1995; Lederle et al., 1997; Mattes et al., 1997). Chez les patients diabétiques, la paroi de l'aorte est le siège d'une synthèse accrue de protéines de la matrice et d'une baisse associée de son catabolisme. L'association négative entre diabète et AAA pourrait être selon divers auteurs la conséquence d'un renforcement structurel de la paroi de l'aorte, notamment par glycation et pontage des protéines matricielles les rendant résistantes à l'action des protéases (Norman et al., 2007, figure 9) et, de façon plus hypothétique, par modification des interactions entre cellules de l'immunité et protéines matricielles (Golledge et al., 2008).

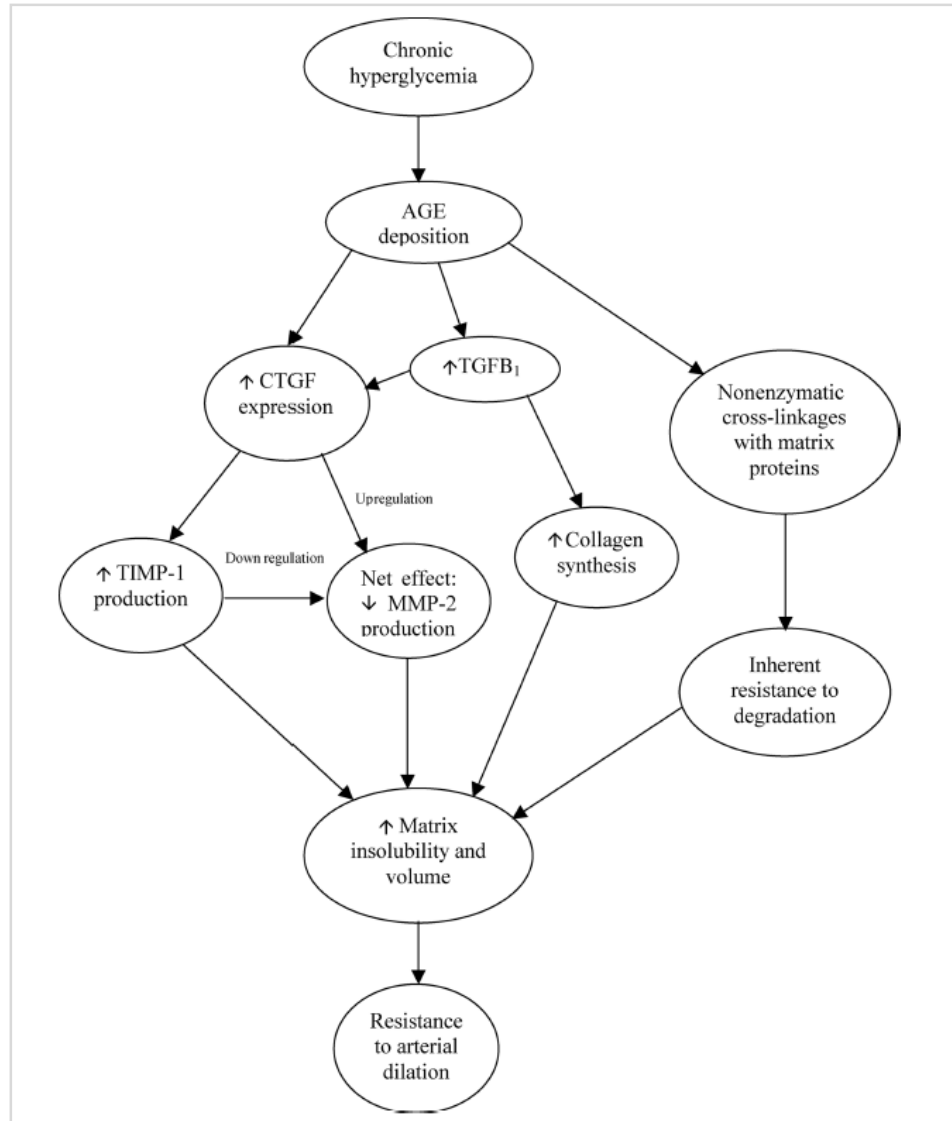


Figure 9: Mécanismes moléculaires potentiellement impliqués dans la résistance à l'AAA chez le diabétique. CTGF: Connective Tissue Growth Factor; AGE: advanced glycation end-products. D'après Norman et al, 2007.

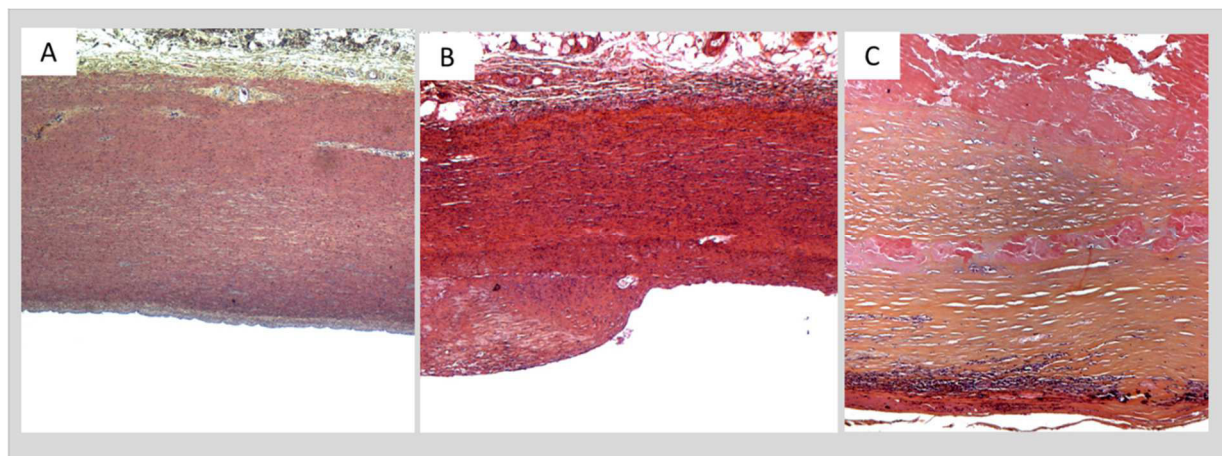


Figure 10: Aspect histologique en trichrome vert d'une aorte normale (A), d'un athérome aortique sans anévrisme (B) et d'une anévrisme de l'aorte abdominale (C). Dans l'athérome « simple » la média et l'adventice conservent un aspect normal, contrairement à l'AAA.

3.3 Etiologie

Rarement, l'AAA peut être la conséquence d'une infection systémique envahissant la paroi artérielle. On parle d'anévrisme mycotique, même dans le cas d'infections bactériennes. La cause la plus connue d'anévrisme mycotique est l'infection syphilitique, généralement responsable d'anévrismes sacciformes le long de la paroi aortique. Cette étiologie n'est pas abordée dans ce manuscrit.

Dans la majorité des cas l'AAA est donc une maladie fréquente, sporadique, soumise aux facteurs de risque cardio-vasculaires cités précédemment. Elle est considérée comme une maladie commune ou complexe telle que définie dans l'introduction du manuscrit.

3.4 Composante héréditaire de l'AAA

Les études épidémiologiques ont démontré que l'antécédent familial d'AAA était un facteur de susceptibilité important d'AAA (Nischán et al., 2010) : Environ 20 % des patients présentent à l'interrogatoire un antécédent d'AAA chez un apparenté au premier degré (Nischán et al., 2010), La prévalence de l'AAA chez un apparenté du premier degré d'un individu atteint est estimée à 8 fois supérieure à celle de la population générale (Ogata et al., 2005b).

Dans une étude récente réalisée en Suède portant sur 265 paires de jumeaux, les odds-ratios respectifs d'AAA était de 71 et 8 pour un vrai et faux jumeau d'atteint. Sur la base de ces données l'héritabilité estimée de cette pathologie est de 0,7 (Wahlgren et al., 2010). Cela n'exclut pas qu'un grand nombre de gènes puissent être impliqués, comme c'est le cas dans le déterminisme génétique de la taille adulte (Hemani et al., 2013). Cette hérédité peut être masquée par l'âge tardif d'apparition de la pathologie (après 55 voire 65 ans) et son histoire naturelle caractérisée par l'absence de symptômes avant rupture (Kuivaniemi et al., 2008). L'existence de formes monogéniques dominantes de fréquence sous-estimée peut même être évoquée dans ce contexte. Une étude réalisée dans une autre pathologie cardiovasculaire qu'est le rétrécissement aortique calcifié (RAC), a mis en lumière les difficultés d'identification des formes familiales en cas de développement tardif de la pathologie et de sa dimension pauci-symptomatique dans la population âgée. Une approche épidémiologique et généalogique en isolat géographique a été utilisée pour caractériser cette maladie considérée initialement comme non héréditaire. Elle a permis d'identifier de nombreux individus atteints apparentés mettant en évidence un mode de transmission dominant de l'affection, ignoré dans les études épidémiologiques précédentes (Probst et al., 2006).

3.5 Aspects histologiques et pathogénèse de l'AAA

Les trois caractéristiques histologiques dominantes de l'AAA comprennent (i) une dégénérescence de la média qui apparaît amincie, associée à (ii) un épaissement intimal siège d'une plaque athéromateuse (figure 10) et (iii) un processus inflammatoire chronique localisé majoritairement dans l'adventice (voir section 3.6 Spécificités du processus inflammatoire dans l'AAA).

L'intima est le siège d'un thrombus intraluminal développé à partir d'une plaque d'athérome, compliquée d'ulcération. La plaque d'athérome et le thrombus qui y est associé forment une structure en remodelage permanent, siège d'un infiltrat inflammatoire riche en polynucléaires et cellules mononucléées d'aspect particulier appelées cellules spumeuses.

La média est caractérisée par une perte par apoptose des cellules musculaires lisse (figure 11, X), et une destruction de la MEC. L'élastine est fragmentée (Henderson et al., 1999; Thompson et Parks, 1996) et il existe une perte des fibres collagènes. La phase précoce de l'AAA est associée à une synthèse accrue de collagène suggérant un processus de réparation (Ormezzano et al., 2010). Toutefois à un stade plus avancé la dégradation du collagène est supérieure à sa synthèse (figure 11, IX et XI). Cette destruction de la média, conséquence d'un déséquilibre entre dégradation et synthèse, est consécutive à une dérégulation des MMP. Ces dernières peuvent être sécrétées par tous les types cellulaires présents dans la paroi et dans le thrombus mural : cellules endothéliales, CML, fibroblastes, macrophages, polynucléaires neutrophiles, et lymphocytes (figure 11, IV et VII, Liu et al., 2006; Schneiderman et al., 1995; Shi et al., 1999; Shimizu et al., 2005).

La collagénase (MMP-1,8 et 13), l'élastase (MMP-12) et de nombreuses autres métalloprotéases (MMP-2, 3, 9, et 12), sont impliquées dans les anomalies de la paroi anévrysmale (Curci et al., 1998; Irizarry et al., 1993; Mao et al., 1999; McMillan et al., 1995; Newman et al., 1994; Senior et al., 1984). La gélatinase B (MMP-9) a été largement étudiée dans les AAA humains (Thompson et al., 1995), mais de nombreux autres MMP et les TIMP ont également été décrites. Des modèles animaux d'anévrysme de l'aorte confirment que les MMP produites par des cellules responsables d'une inflammation chronique sont des médiateurs de la dégradation de la paroi (Daugherty et al., 2000; Longo et al., 2002; Pyo et al., 2000); L'inhibition de certaines MMP induit un maintien de l'intégrité de la paroi artérielle dans certains modèles expérimentaux d'anévrysme, constituant une piste thérapeutique prometteuse (Mosorin et al., 2001).

D'autres sérine protéases comme la plasmine et l'activateur tissulaire du plasminogène, l'élastase leucocytaire des neutrophiles, ainsi que les cathepsine (cystéine protéases à l'activité élastolytique puissante) peuvent également contribuer à la protéolyse matricielle (Cohen et al., 1991; Dollery et al., 2003; Gacko et Glowinski, 1998; Jean-Claude et al., 1994; Reilly et al., 1994; Schneiderman et al., 1995; Shi et al., 1999; Shimizu et al., 2005; Sukhova et al., 1998); elles sont présentes dans toute l'épaisseur de la paroi

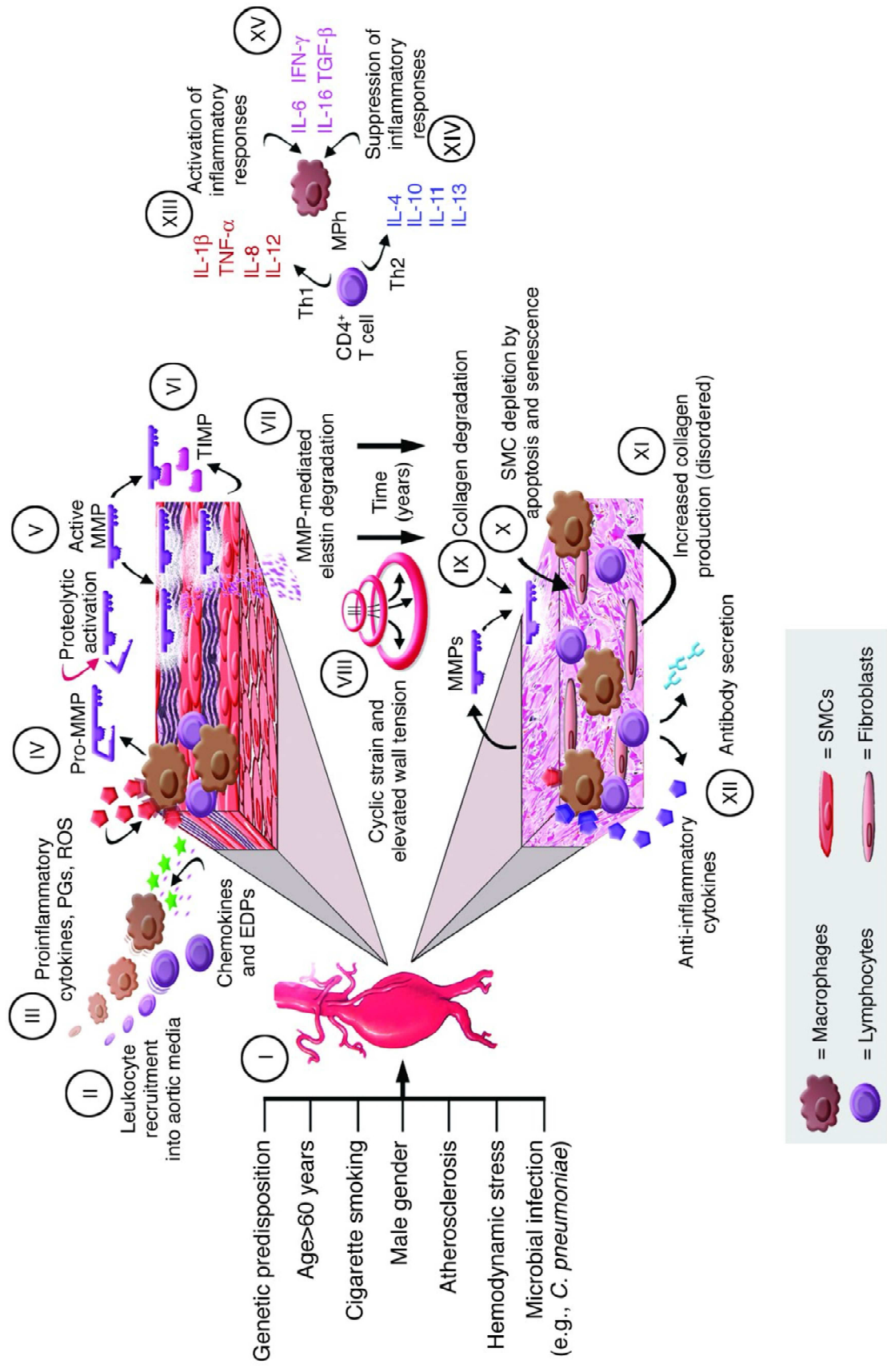


Figure 11: Physiopathologie de l'AAA. Les aspects inflammatoires, immunitaires, mécaniques, et métaboliques ainsi que les facteurs de risque sont représentés. D'après Curci et al, 2014.

anévrismale à des concentrations supérieures aux artères athéroscléreuses non anévrysmales. *A contrario*, la cystatine C, inhibiteur majeur des élastases de type cathepsines K et S, est sous-exprimée dans l'AAA (Shi et al., 1999); les souris haploinsuffisantes pour la cystatine C montrent une fragmentation accrue des fibres élastiques et un anévrisme aortique (Sukhova et al., 2005).

L'AAA est un processus pathologique chronique et silencieux, et les observations rapportées ci-dessus ne font pas état des phénomènes initiateurs de ce processus, dont on sait seulement qu'ils sont sous-tendus par les facteurs de risque cardiovasculaires d'une part, et une composante héréditaire d'autre part (figure 11, I).

Le recrutement des leucocytes dans la média et l'adventice semble être cependant un événement précoce et probablement sous l'influence de chémokines (Koch et al., 1993). Divers peptides dérivés de l'élastine peuvent également constituer un signal chémoattractant des neutrophiles, des fibroblastes, et des monocytes-macrophages (Hance et al., 2002; Senior et al., 1980, 1984).

3.6 Spécificités du processus inflammatoire dans l'AAA

L'AAA est le siège d'un infiltrat inflammatoire par les macrophages, mais il contient aussi un grand nombre de cellules T, B, de plasmocytes, et de cellules dendritiques (cellules présentatrices de l'antigène) localisés majoritairement dans la partie externe de la média et l'adventice (Bobryshev et Lord, 2001; Koch et al., 1990). Cet infiltrat cellulaire sécrète différents facteurs inflammatoires (cytokines, chemokines, leucotriènes, dérivés réactifs de l'oxygène (ROS), et immunoglobulines : figure 11, II et III). Il correspond à une réponse phagocytaire aux produits de dégradation du thrombus intraluminal, et lymphocytaire à des néoantigènes produits par cette dégradation.

Le recrutement des leucocytes et l'expression de cytokines pro-inflammatoires sont aussi une caractéristique de l'athérogénèse. L'AAA se développant généralement à partir d'un athérome aortique sévère, (Reed et al., 1992), il est communément appelé « anévrisme athéroscléreux » (Rose et Dent, 1981; Schneiderman et al., 1995). Cependant, comme nous l'avons vu dans la section 3.2 Epidémiologie et facteurs de risque de l'AAA, la majorité des patients souffrant d'athérosclérose coronarienne ou aortique abdominale ne développent pas d'AAA. L'athérosclérose en soi peut être la source d'un phénomène inflammatoire localisé, susceptible de déclencher sous certaines conditions spécifiques le processus anévrysmal. En effet dans l'athérosclérose « pure », les cellules immunocompétentes (essentiellement composé de lymphocytes T et les macrophages) s'accumulent dans l'intima, siège de la plaque d'athérome (Gown et al., 1986; Hansson et al., 1988; Jonasson et al., 1986; Ross, 1999). Il s'agit donc d'un processus inflammatoire relativement confiné à l'intima ou la portion interne de la média. En comparaison, l'infiltrat adventiciel de l'AAA s'avère bien plus riche et complexe (Koch et al., 1990; Parums, 1990; Ramshaw et Parums, 1990). Cet infiltrat est particulièrement important dans les formes dites « inflammatoires » d'AAA, ou périaortites, s'étendant dans

la graisse périaortique et pouvant résulter dans les formes les plus sévères en une fibrose rétropéritonéale (Mitchinson, 1986).

3.7 L'immunité

L'existence d'un tissu lymphoïde associé aux vaisseaux (VALT), constitué de cellules immunocompétentes (lymphocytes T et B) et de cellules dendritiques, a été récemment reconnu (Waltner-Romen et al., 1998; Wick et al., 1997). Dans la paroi artérielle normale, quelques-unes de ces cellules sont réparties dans la couche sous-endothéliale de l'intima et autour des *vasa vasorum* de l'adventice (Bobryshev et Lord, 1995; Bobryshev et al., 1997; Lord et Bobryshev, 1999; Waltner-Romen et al., 1998; Wick et al., 1997). Par analogie au tissu lymphoïde associé aux muqueuses (connu sous le nom de MALT) des voies respiratoires et gastro-intestinales, le VALT pourrait constituer une protection de la paroi contre divers agents infectieux (Wick et al., 1997). Les cytokines pro-inflammatoires sécrétées par les lymphocytes T helper Th1, notamment l'interféron gamma (IFN- γ), sont élevées dans le sang et le tissu aortique de patients avec AAA, le taux sérique d'IFN- γ étant corrélé à l'expansion de l'anévrisme (Juvonen et al., 1997; Szekanecz et al., 1994). L'adventice de l'AAA est également le siège d'un infiltrat Th2 sécrétant d'autres types de cytokines (figure 11, XIII à XV, Davis et al., 2001).

Le tissu aortique anévrysmal contient également une grande quantité d'IgG (figure 11, XII), suggérant qu'une réponse immunitaire humorale est un phénomène constitutif de la pathologie. Des travaux récents ont conduit à l'identification de plusieurs auto-antigènes putatifs, qui pourraient correspondre à des produits de dégradation de la matrice extracellulaire (Hirose et al., 1997). La recherche d'auto-anticorps dans l'AAA a révélé l'existence d'IgG contre un peptide humain exprimé dans le tissu aortique normal, la protéine AAAP40 (Gregory et al., 1996; Xia et al., 1996). Ces résultats semblent indiquer une réponse auto-immune liée à l'AAA. Cependant, des études plus sophistiquées ont révélé que les lymphocytes B dérivés de lésions d'AAA présentent un répertoire non clonal de chaînes lourdes d'immunoglobulines, de même que les cellules T (TCR) présentent une diversité importante de récepteurs, reflet d'une réponse polyclonale (Walton et al., 1997; Yen et al., 1997). Dans une analyse récente et complète, Ocana et coll. ont démontré que les lymphocytes infiltrant l'anévrisme sont constitués de cellules mémoire activées exprimant des molécules costimulatrices. Plus précisément, les cellules T CD4+ prédominantes dans l'AAA présentent un profil de marqueurs de surface similaire à celui retrouvé dans d'autres atteintes auto-immunes ou inflammatoires chroniques, soutenant l'hypothèse que l'AAA est une maladie inflammatoire chronique (Ocana, 2003).

3.8 L'hypothèse infectieuse

Les déclencheurs de l'immunité acquise dans l'AAA sont inconnus, et l'une des hypothèses prédominantes est l'existence d'une infection virale ou bactérienne passée ou présente.

L'infection par *Chlamydia pneumoniae* (*C.p*) semble être retrouvée dans 35% des AAA de plus de 5 cm (Petersen et al., 2002). De même, une étude allemande portant sur les tissus anévrismaux de patients a révélé la présence de cytomégalovirus (CMV) dans 21 aortes parmi 22. L'analyse par PCR du virus de l'herpès (HSV) et du CMV a démontré une prévalence plus élevée d'infection dans les anévrysmes « inflammatoires » que dans les formes « athéroscléreuses », ce qui plaide pour le rôle proinflammatoire de ces agents infectieux dans le processus anévrysmal (Gredmark-Russ et al., 2009).

Une étude portant sur *Borrelia burgdoferi* (*B.b*), agent de la maladie de Lyme, a démontré que plus d'un tiers des patients porteurs d'AAA ont des anticorps contre *B.b*, contre environ 10% en population contrôle (Hinterseher et al., 2012). Dans cette étude, l'absence d'identification de *B.b* dans la paroi anévrysmale par des techniques sensibles a fait suggérer aux auteurs un mécanisme de mimétisme moléculaire entre antigènes de *B.b* et constituants de la paroi aortique. De même, le peptide AAAP40 présente, selon ses découvreurs, des similitudes de séquence avec des protéines du CMV et de *Treponema pallidum*, agent de la syphilis.

Plus récemment, les pathogènes dits « faibles » buccodentaires, tels que *Porphyromonas gingivalis* (*P.g*) ont été impliqués dans la physiopathologie de l'anévrysmes de l'aorte abdominale. Ce germe intracellulaire facultatif est un constituant de la flore endobuccale dans la population générale. Dans un contexte particulier impliquant des phénomènes inflammatoires et environnementaux notamment, *P.g* est associé à la maladie parodontale, maladie locale du tissu de soutien de la dent. Une association a été mise en évidence entre un taux élevé d'anticorps sériques anti- *P.g*, la présence d'ADN bactérien dans la paroi anévrysmale, et la prévalence, mais aussi la sévérité, de l'AAA (Delbosc et al., 2011). *P.g*, par la sécrétion de certaines protéines, est responsable d'une modification phénotypique *ex vivo* des cellules musculaires lisses d'anévrysmes de patients, confirmé en modèle murin. Ce phénotype est retrouvé dans les AAA agressifs, et d'autant plus fréquemment chez les patients porteurs de maladie parodontale (Hokamura et al., 2010; Nakano et al., 2011).

Il est intéressant de noter que les agents infectieux identifiés par ces diverses études présentent un point commun : ils sont responsables d'infections chroniques voire latentes, et non de formes aiguës avec tableau infectieux bruyant au premier plan. Cette constatation est cohérente avec un processus inflammatoire tissulaire chronique à bas bruit responsable de l'AAA.

4. Génétique de l'AAA

4.1 Analyses de liaison

Les loci AAA1 et AAA2 (OMIM)

Shibamura et coll. (Shibamura et al., 2004) ont identifié par analyse de liaison chez 36 cas-index et 75 apparentés porteurs d'AAA deux régions génomiques sur le chromosome 19q13 (locus AAA1) et 4q31 (locus AAA2) avec des Z-scores de 4,64 et 3,73 respectivement. Le locus 19q13 a été confirmé dans une étude portant sur trois familles avec AAA (Z-score de 3,95) mais l'extension de l'analyse à 55 familles supplémentaires n'a pas permis de confirmer cette liaison (Shibamura et al., 2004; van Vlijmen-van Keulen et al., 2002). Ce résultat rend compte de la grande hétérogénéité génétique de l'AAA et des difficultés inhérentes à l'analyse de liaison dans cette pathologie.

Le locus 4q31 n'a pas fait l'objet d'investigations spécifiques. Le locus 19q13, qui contient un grand nombre de gènes connus, a été exploré et les études ultérieures se sont focalisées sur le gène *KLK1*.

Le gène KLK1

KLK1 code la Kallicroïne, sérine protéase qui convertit le kininogène en bradykinine, peptide proinflammatoire et vasoactif agissant par liaison à ses récepteurs B1 et B2. La bradykinine a en outre des propriétés procoagulantes, antibactériennes et antifongiques. Il a été montré que la dysfonction du récepteur B1 est responsable d'un risque accru de développement d'un AAA dans un modèle de souris (Merino et al., 2009). Une étude subséquente d'association menée au locus *KLK1* portant sur 1 629 patients (Biros et al., 2011a) a identifié un SNP impliqué spécifiquement dans les AAA de grande taille (> 5 cm). Ce SNP est responsable d'un épissage alternatif de *KLK1*, et il a été identifié une prépondérance de transcrits courts de *KLK1* dans un échantillon de 79 AAA de grande taille.

4.2 Etude d'association gènes candidats

Plus d'une centaine d'études d'association gènes-candidats ont été menées dans cette pathologie. La MEC, le système immunitaire, les facteurs de risque cardiovasculaire et les voies de signalisation inflammatoires et protéolytiques ont été étudiés. La plupart de ces études sont cependant de faible puissance (moins de 150 patients en moyenne, Hinterseher et al., 2011). Leurs résultats ne sont généralement pas confirmés par les études ultérieures, voire se révèlent contradictoires. Comme c'est le cas général dans les maladies communes

(décrit dans le chapitre 1), les odds-ratio caractérisés à ces divers loci ne rendent pas compte de l'héritabilité de la maladie (Saratzis et al., 2011), qui est comme vu précédemment très élevée pour une maladie commune (section 3.4 Composante héréditaire de l'AAA).

Plusieurs méta-analyses ont été réalisées à partir de ces nombreuses études, et semblent identifier des associations avec divers gènes codant des protéines de la MEC notamment : MMP-3, MMP-2, MMP-13, l'élastine, TGF β R1 et 2 mais aussi la CRP (Biros et al., 2011b; Saracini et al., 2012; Saratzis et al., 2014). Ces résultats se heurtent cependant aux limites et biais des méta-analyses (Attia et al., 2003).

4.3 GWAS

Le locus 9p21.3 et le gène ANRIL

Un premier locus, identifié à l'origine dans un GWAS pour la maladie coronaire, puis par la suite dans de nombreuses autres maladies dont les anévrysmes intracrâniens, la maladie d'Alzheimer et certains cancers (Helgadóttir et al., 2008), a été confirmé dans l'AAA par diverses études (OR=1,31 $p=1,2 \times 10^{-12}$, Biros et al., 2010; Bown et al., 2008; Helgadóttir et al., 2008; Thompson et al., 2009).

Le SNP identifié est situé dans un désert de gènes. Le meilleur gène candidat de la région est un gène d'ARN non codant appelé *CDKN2B-AS* ou *ANRIL*. Cet ARN, antisens de *CDKN2B*, réprime son expression par processus épigénétique (Kotake et al., 2011). Chez la souris, une délétion de 70 Kb englobant le SNP associé et une partie du gène *ANRIL*, mais pas les deux gènes les plus proches *CDKN2A* et *CDKN2B* (protéines checkpoints du cycle cellulaire) réduit l'expression vasculaire et cardiaque de ces deux gènes (Visel et al., 2010). Les souris KO *CDKN2B* soumises à une perfusion d'élastase développent un anévrysme aortique infrarénal plus massif que les souris témoins (Leeper et al., 2013). Ces souris présentent également une augmentation de la prolifération intinale en réponse à une lésion de la paroi vasculaire, associée à une apoptose des CML dans la paroi aortique (Leeper et al., 2013). Dans un modèle de culture cellulaire, le knock-down de *CDKN2B* entraîne une augmentation de l'expression de p53 et p21, molécules inductrices de l'apoptose. L'expression de *CDKN2B* est réduite dans les cellules musculaires lisses issues d'AAA humain.

Ces divers éléments confortent l'hypothèse actuelle que les individus porteurs de l'allèle à risque en 9p21 ont une expression affaiblie de *CDKN2B*, responsable d'un phénomène d'apoptose des CML dépendant de p53. Cette perte des CML pourrait constituer un élément de sensibilisation au processus anévrysmal.

GENETIC LOCUS	STUDY DESIGN	OR/LOD	REFERENCE
AAA-1, 19q13	Whole genome analysis using affected-relative-pair linkage analysis; 109 AAA-affected families included	LOD 4.64	Shibamura et al (2004)
AAA-2, 4q31	Whole genome analysis using affected-relative-pair linkage analysis; 109 AAA-affected families included	LOD 3.73	Shibamura et al (2004)
DAB2IP	Genome-wide association study; 1292 AAAs and 30 503 controls	OR 1.21	Gretarsdottir et al (2010)
LRP-1	Genome-wide association study; 1866 AAAs and 5435—further validated in 1491 AAAs and 11 060 controls	OR 1.15	Bown et al (2011)
LDLR	Genome-wide association study; 1830 AAAs and 5435 controls—further validated in 1292 AAAs and 30 503 controls and then further validated in three geographically distinct populations	OR 1.32	Bradley et al (2013)
SORT-1	Genome-wide association study; 7048 AAAs and 75 976 controls	OR 1.23	Jones et al (2013)

Tableau 2: Loci et gènes identifiés par études « génome entier » réalisées dans l'AAA. D'après Saratzis et al, 2014.

Le gène DAB2IP

Une étude en GWAS portant sur 1 292 patients, 30 503 contrôles et 30 000 SNP (Gretarsdottir et al., 2010) a identifié, outre le locus 9p21 précédemment cité, un SNP dans le gène *DAB2IP*, confirmé par étude de réplication (3 267 cas et 7 451 contrôles, $p=4,6.10^{-10}$). Ce SNP porté par l'intron 1 du gène est connu dans la maladie coronarienne, l'embolie pulmonaire et l'artérite des membres inférieurs. *DAB2IP* code une protéine d'échafaudage impliquée dans une grande variété de processus cellulaires et moléculaires : réponse immunitaire innée, inflammation et inhibition de la croissance cellulaire, apoptose et survie cellulaire, angiogenèse, et migration cellulaire. Elle joue un rôle de checkpoint du cycle cellulaire en réduisant les niveaux de cycline G1 résultant en un arrêt du cycle en phase G0. Elle est impliquée dans la transduction du signal inflammatoire *via* le TNF et l'interféron ou les lipopolysaccharides. Elle module l'équilibre entre survie induite par la voie PI3K-AKT et apoptose médiée par la voie MAP Kinases-JNK. Ce gène est donc un très bon candidat de l'AAA pour de multiples raisons.

Le locus SORT1

Jones et coll. (Jones et al., 2013) ont analysé 44 loci associés aux taux de lipides sanguins et à la maladie coronarienne dans une cohorte Néo-Zélandaise constituée de 11 sous-cohortes. Ils ont identifié à nouveau le locus 9p21 et un nouveau locus en 1p13.3 dont le gène candidat est *SORT1*. Ce locus n'a jamais été identifié dans la maladie coronarienne et semble donc spécifique de l'AAA. Le gène *SORT1* code la sortiline, petite protéine cargo présente à la surface cellulaire et dans les membranes de l'appareil de Golgi. Elle est impliquée dans la régulation du trafic intracellulaire par le réseau membranaire trans-golgien, et l'adressage de certaines protéines dans le lysosome et divers autres compartiments vésiculaires. Musurunu et coll. ont démontré qu'au locus *SORT1*, un snp fréquent (MAF = 30% chez les caucasiens) non codant créait un site de fixation du facteur de transcription C/EBP, responsable d'une augmentation de l'expression hépatique de *SORT1*, et d'une baisse du LDL cholestérol (Musunuru *et al.*, 2010b). Le mécanisme de régulation du taux de cholestérol par *SORT1* est inconnu.

Le gène LRP1

Un autre locus spécifique de l'AAA a récemment été identifié dans une cohorte de 2 000 cas et 5 000 contrôles par Bown et coll. (Bown et al., 2011). Sur neuf loci candidats ($p < 1 \times 10^{-5}$), la réplication (2 871 cas et 32 687 contrôles) a confirmé l'association à un SNP présent dans l'intron 2 du gène *LRP1* (OR=1,15 $p=4,5.10^{-10}$). Ce variant modifie l'expression transcriptionnelle de *LRP1* (Bown et al., 2011). Les souris inactivées pour *LRP1* dans les CML vasculaires présentent une susceptibilité accrue aux anévrysmes (Boucher et al., 2003). *LRP1* est

un récepteur d'endocytose impliqué dans la phagocytose des cellules apoptotiques. Il est aussi impliqué dans l'homéostasie lipidique cellulaire, la clairance plasmatique de résidus de chylomicrons et l' α 2 macroglobuline, et le métabolisme des complexes activateurs/inhibiteurs du plasminogène.

Le gène LDLR

Enfin cette même équipe (Bradley et al., 2013), à partir de la même cohorte utilisée dans l'identification de *LRP1*, a pu par le biais de répliques identifier le gène *LDLR* codant le récepteur aux LDL (OR=1,32 p=2.10⁻¹⁰), précédemment connu dans la maladie coronarienne.

Les données issues de ces diverses études sont résumées dans le tableau 2.

L'ensemble des études génétiques réalisées dans l'AAA apportent aujourd'hui des informations sur les processus sous-jacents à la pathologie : des protéines structurales (élastine) et du métabolisme de la MEC (récepteurs au TGF β , métalloprotéases) mais aussi des acteurs impliqués dans le cycle cellulaire (*ANRIL*, *DAB2IP*), l'apoptose (*LRP1*), le tri cellulaire et l'adressage au lysosome (sortiline), le métabolisme des lipides (récepteur au LDL), l'inflammation (*CRP*, *DAB2IP*, kalllicréine) et la réponse anti-infectieuse (kalllicréine) donnent un aperçu de la variété des processus impliqués dans l'AAA. Ces résultats sont cependant décevants car ils ne permettent pas, encore une fois, d'expliquer l'héritabilité importante de la pathologie. Comme précédemment discuté dans la section 3.4 Composante héréditaire de l'AAA, dans le cas particulier de cette pathologie l'héritabilité manquante peut être la conséquence d'un défaut diagnostique lié à son histoire naturelle et ses caractéristiques épidémiologiques.

5. Etude génétique de l'AAA dominant et sporadique

5.1 Présentation de la famille princeps

La famille « SE » a été identifiée par le professeur Pierre Julia, chirurgien vasculaire à l'hôpital Européen Georges Pompidou, à partir de trois frères qu'il a opérés d'AAA (individus II-1, II-9 et II-11). Ils appartiennent à une grande fratrie de 13 personnes. A cette génération, parmi 6 frères et 7 sœurs, on note 3 hommes et 3 femmes atteints (figure 12). L'anévrisme de l'aorte abdominale dans cette famille était hérité du père (I-1). La transmission dominante de l'affection dans la famille est donc très probable (6 atteints parmi 13), avec

HTA avérée = oui
 AAA = oui
 HTA avérée = non

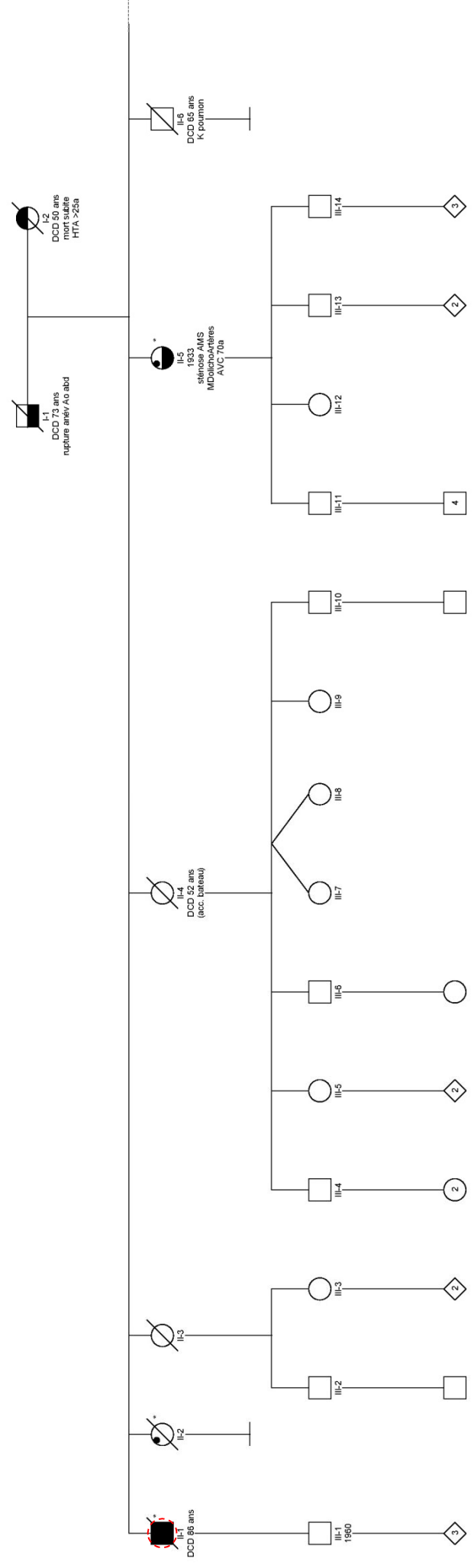


Figure 12: Pedigree de la famille SE. La partie supérieure du symbole est noircie en cas d'hypertension. La partie inférieure du symbole est noircie en cas d'hypercholestérolémie. Les patients analysés par séquençage d'exome sont entourés d'un cercle rouge. Les patients génotypés pour les divers gènes candidats sont marqués d'une astérisque.

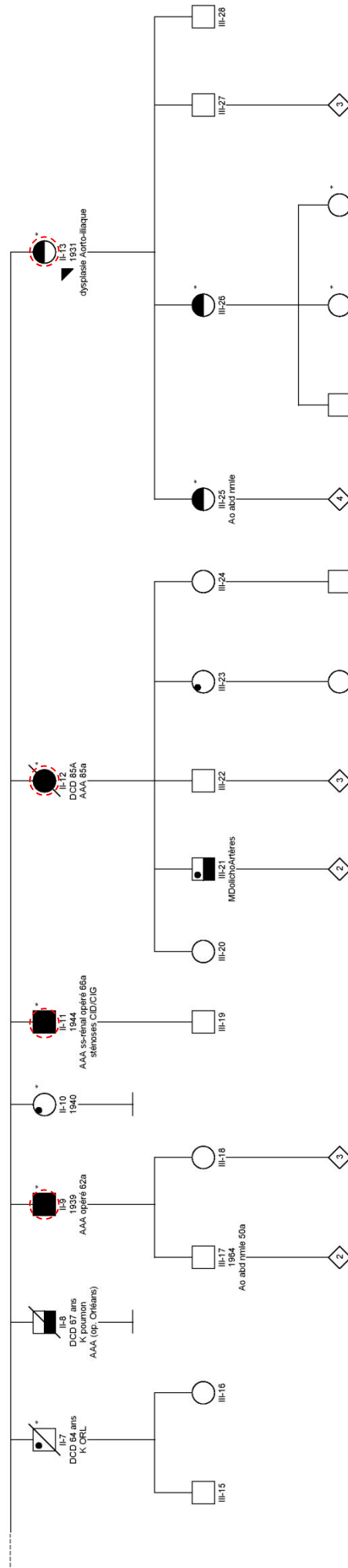


Figure 12 (suite): Pedigree de la famille SE. La partie supérieure du symbole est noircie en cas d'hypertension. La partie inférieure est noircie chez les patients atteints d'AAA ou de mega-dolicho-aorte. Les patients analysés par séquençage d'exome sont entourés d'un cercle rouge. Les patients génotypés pour les divers gènes candidats sont marqués d'une astérisque.

une pénétrance qui pourrait être proche de 1 à partir de 65 ans. Elle est cependant impossible à estimer de façon fiable sans mutation causale ni phénotype de certitude pour la génération suivante. La génération II étant relativement âgée (de 89 à 71 ans), 6 membres de la fratrie sont actuellement décédés. La troisième génération est beaucoup plus jeune (de 61 à 43 ans) et la pénétrance de l'AAA y est encore incomplète.

Dans cette famille, l'atteinte aortique chez les femmes semble atténuée et de découverte plus tardive (tableau 3), ce qui semble concordant avec les données épidémiologiques démontrant le risque plus élevé des hommes pour cette affection (voir section 3.2 Epidémiologie et facteurs de risque) : Parmi les trois femmes considérées atteintes, seule II-12 présentait un AAA, les femmes II-5 et II-13 présentent une dolicho-mega-aorte (figure 6, E et F). A la génération III, un homme de 63 ans présente aussi des mega-artères (III-21).

Comme nous l'avons vu dans la section 3.1 Définition, il existe une agrégation familiale connue entre mega-aorte et AAA. Les patients II-5, II-13 et III-21 ont été considérés comme porteurs d'un équivalent modéré d'AAA et portaient donc l'allèle à risque.

Le phénotype des patients inclus dans l'analyse génétique a été confirmé par échographie-doppler, angioIRM ou angioscanner. Les facteurs de risque de l'AAA que sont le tabac, l'hypertension et la dyslipidémie sont présents chez de nombreux membres de la famille (tableau 3). On y retrouve aussi des facteurs protecteurs que sont bien sûr le sexe féminin, et le diabète (patiente III-26).

5.2 Analyse d'exome

Je disposais de l'ADN de 12 patients des générations II et III : 6 individus atteints, 1 sain et 2 de statut indéterminé à la première génération, et un atteint et deux sains à la seconde génération (figure 12).

Au moment du choix des patients à inclure dans le séquençage d'exome, le statut clinique de II-5 et III-21 n'était pas connu. En conséquence, uniquement l'ADN des patients atteints appartenant à la génération II a été séquencé. Il s'agit des sujets II-1, II-9, II-11, II-12 et III-13. Deux des trois femmes sont porteuses d'une dolicho-mega-aorte.

J'ai analysé les données de séquençage grâce à un outil nommé Polyweb, développé par l'équipe la plateforme de bioinformatique de la Fondation Imagine (Dr Patrick Nitschké).

Les données brutes de séquençage ainsi que le croisement des résultats sont réunis dans le tableau 4 : le croisement des données pour les variants faux-sens prédits délétères par SIFT ou Polyphen, et les variants

Nom	Individu	Athérome	Phénotype aortique (âge de diagnostic)	HTA (chiffres tensionnels)	Tabac (nombres de PA)	Dyslipidémie	Commentaire
Se Pie	II1	+	AAA (64 ans)	+	+ (35 PA)	-	
Cho Col	II2	?	?	-	-	-	
Bon Mon	II3	?	?				décédée 81 ans (Alzheimer)
Cha Mar	II4	?	?				Décédée à 52 ans (accident)
Sch Pas	II5	+	Mega-dolicho-aorte et artères iliaques (80 ans)	-	+ (5 PA)	+ (tahor)	
Se Jac	II6	?	?				Décédé à 65 ans (K du poumon)
Se Jea	II7		?	-	+ (35PA)	+ (5,37mmol/l)	Décédé à 64 ans (K ORL)
Se Mar	II8		?	?	?	?	Décédé à 67 ans (K du poumon)
Se Jos	II9	+	AAA (62 ans)	+ traitée	+ (30PA)	-	
Se Mar	II10		normal	-	+ (15PA)	-	
Se Pat	II11	+	AAA (66 ans)	+ traitée (?)	+ (45 PA)	+ traitée	
Giv Odi	II12	+	AAA (85 ans)	+ (180/110)	-	-	décédée
Lev Don	II13	+	Mega-dolicho-aorte et artères iliaques (78 ans)	+(?)	-	-	
Giv Jea	III21	-	Mega-dolicho-aorte et artères iliaques (55 ans)	+ traitée (140/90)	+ (20PA)	-	
Pie Chr	III25	-	normal	+ (?)	-	-	
Vie Ann	III26	-	normal	+ traitée (140/100)	-	? (tahor)	diabète type 2

Tableau 3: Phénotype artériel et facteurs de risque de l'AAA des patients de la famille SE.

Individus	Variants				
	Total	MAF<1%	Faux-sens, épissage, et tronquants	Prédits pathogènes par Polyphen OU SIFT (pour les faux-sens uniquement)	Communs (modèle autosomique dominant)
II1	24006	5459	493	256	3
II9	24230	5469	519	238	
II11	23790	5388	481	234	
II12	24350	5654	529	253	
II13	24313	5567	487	243	

Tableau 4: résultats bruts du séquençage d'exome des 5 membres analysés de la famille SE.

stops, d'épissage ou décalant le cadre de lecture permet d'obtenir seulement 3 candidats à partir d'une liste d'environ 250 par individu.

5.3 Identification de gènes candidats, ségrégation et liaison aux locus correspondants

Les 3 candidats identifiés sont les gènes *PLA2G3*, porteur d'une mutation stop de l'exon 4 (p.Arg303*), *APOL3* porteur d'une variation faux-sens de l'exon 3 (p.Arg377Trp), et *NFKBIZ*, porteur d'une mutation faux-sens sur l'exon 1 (p.Gly80Cys).

Une lecture rapide de la littérature montre que ces trois gènes pouvaient être de bons candidats puisque *PLA2G3* a été impliqué dans le métabolisme endothélial de cholestérol LDL (Parthasarathy et al., 1985), *APOL3* a été identifié dans l'endothélium d'artères athéromateuses (Horrevoets et al., 1999), et *NFKBIZ* est un effecteur de la réponse immunitaire aux agents bactériens par l'intermédiaire notamment de l'INF γ (Hanihara-Tatsuzawa et al., 2014; Ishiguro-Oonuma et al., 2015).

Les trois variants identifiés ont été séquencés chez les autres membres de la famille SE. Les résultats indiquent que *NFKBIZ* ne ségrège pas avec le phénotype, contrairement aux deux autres gènes (tableau 6) : les trois individus sains seraient non pénétrants pour ce variant (en jaune) mais surtout deux individus atteints ne portent pas le variant (en rouge). Ce gène a été exclu des candidats potentiel.

Dans le cadre d'une méta-analyse sur les données GWAS de 30 000 cas réalisée par Matt Bown (communication personnelle) aucune association n'a été identifiée pour les deux candidats restants : une extraction des données de 4 000 SNP à proximité de ces deux gènes a montré une p-value de 0,043 pour le marqueur rs138784 situé à environ 1 Mb d'*APOL3*, et une p-value de 0,00057 pour le marqueur rs2286442 situé à environ 750 Kb de *PLA2G3*. Ces p-values sont très supérieures à la significativité attendue pour un marqueur associé dans une étude GWAS.

La coségrégation de *PLA2G3* et *APOL3* est attendue puisqu'ils sont portés tous deux par le chromosome 22 et sont distants d'environ 5 Mb. Le résultat de l'analyse de ségrégation est donc identique pour ces deux gènes : seul l'individu II-26 serait non pénétrant pour l'un ou l'autre variant. Fait intéressant, cette patiente présente un diabète sévère insulino-requérant (facteur protecteur de l'AAA) qui s'est révélé vers l'âge de 50 ans, qui pourrait expliquer la non-pénétrance de l'un ou l'autre de ces variants (ou les deux).

Une analyse de liaison au locus portant sur les 10 patients génotypés en utilisant uniquement les deux variants de *PLA2G3* et *APOL3* comme marqueurs a été effectuée avec l'outil Merlin. Le lod-score obtenu est

gene	description	Type	cDNA	protéine	identifiant dbSNP	Polyphen	sift	Mutation taster	AlignGVGD	Fréquence allélique (EXAC)	Fréquence allélique Caucasiens (EXAC)
PLA2G3	phospholipase A2, group III	Stop	c.907C>T	p.Arg303*	rs151265075	-	-	-	-	1/1000	1,7/1000
APOL3	apolipoprotein L3	Faux-sens	c.1129C>T	p.Arg377trp	rs145559179	0,999 (probably damaging)	1,01 (deleterious)	Polymorphism (p-value: 1)	less likely	8/100 000	5/100 000
NFKBIZ	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	Faux-sens	c.238G>T	p.Gly80Cys	-	0,997 (probably damaging)	0,08 (tolerated)	Disease causing (p-value =0,56)	less likely	0	0

Tableau 5: caractéristiques des trois variants identifiés après les étapes de filtrage classiques.

Nom	Individu	Statut (A: atteint, S: sain)	Présence du variant		
			PLA2G3	APOL3	NFKBIZ
Se Pie	II1	A	+	+	+
Sch Pas	II5	A	+	+	-
Se Jos	II9	A	+	+	+
Se Mar	II10	S	-	-	+
Se Pat	II11	A	+	+	+
Giv Odi	II12	A	+	+	+
Lev Don	II13	A	+	+	+
Giv Jea	III21	A	+	+	-
Pie Chr	III25	S	-	-	+
Vie Ann	III26	S	+	+	+

Tableau 6: Analyse de ségrégation des trois variants candidats chez les membres de la famille SE dont le phénotype est certain. Le variant identifié dans NFKBIZ ne ségrège pas avec le phénotype, tandis que les variants PLA2G3 et APOL3 ségrègent sauf pour III26 qui serait non pénétrante.

de 4,14 ($p=2.10^{-5}$). Avant l'analyse de ségrégation le lod-score pour les 5 patients analysés par séquençage d'exome était de 0,25 ($p=0,4$). Une analyse approfondie la littérature portant sur ces deux gènes a permis d'apporter des éléments de choix pour la suite de l'étude.

5.3.1 *PLA2G3*

5.3.1.1 Données de la littérature

PLA2G3 est un gène de 7 exons qui n'a qu'un seul transcrit. Il code la phospholipase A2 du groupe 3. Les phospholipases A2, dont *PLA2G3* sont présentes à la surface membranaire de nombreux types cellulaires dont les cellules endothéliales. Elles hydrolysent la phosphatidylcholine (phospholipide membranaire ou présent dans les particules de HDL et LDL) en lysophosphatidylcholine.

En 2007, Bostrom et coll. ont étudié le phénotype de souris *LDLR^{-/-}* (modèle murin d'athérosclérose) surexprimant ou sous-exprimant *PLA2G3* par mutagenèse de *PLA2G3* en cellules de moelle osseuse. Sous régime athérogène, Les souris surexprimant *PLA2G3* présentaient des plaques d'athérome aortique 2 à 3 fois plus développées que les souris non mutantes, et les souris sous-exprimant *PLA2G3* présentaient des plaques moins développées d'environ 30% (Bostrom et al., 2007). Sato et coll. ont généré une souris transgénique surexprimant *PLA2G3* humain, et montré que cette surexpression augmentait l'hydrolyse de phosphatidylcholine en lysophosphatidylcholine, facilitant la formation des cellules spumeuses ex vivo (Sato et al., 2008). L'accumulation de *PLA2G3* a été détectée dans les lésions athérosclérotiques de souris *apoE^{-/-}*, autre modèle bien connu d'athérosclérose. Ces souris nourries par régime athérogène développent comme les souris *LDLR^{-/-}* des lésions athéroscléroseuses aortiques plus sévères que les souris contrôle.

La mutation identifiée dans la famille SE est une mutation stop, présente chez un peu plus d'1 pour 1 000 individus caucasiens. La fréquence cumulée des variants identifiés dans la base EXAC, induisant une haploinsuffisance (mutations stop, frameshift et de la méthionine initiatrice) montre qu'environ 5,6 % individus ont une haploinsuffisance de *PLA2G3*, un variant récurrent (p.Ser70ValfsTer3) étant présent chez 4‰.

Cette fréquence relativement élevée associée aux données de la littérature sus-citées ne semblent pas pouvoir impliquer de mutations inactivatrices de *PLA2G3* dans la physiopathologie de l'AAA. Il a cependant été séquencé dans des formes familiales d'AAA afin de déterminer le type de variants identifiés portés par ces patients.

5.3.1.2 Séquençage de *PLA2G3* dans 127 cas familiaux d'AAA

Dans le cadre de deux collaborations, l'une avec le Pr Elisabeth Tournier-Lasserre (Génétique des maladies vasculaires, INSERM U740) et l'autre avec le Pr Danielle Majoor (Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands) j'ai pu obtenir l'ADN de patients de 2 cohortes de cas-index de familles avec AAA. Nous avons séquençé par technique Sanger 137 cas familiaux d'AAA et identifié trois variations : l'anomalie frameshift récurrente de la population (p.Ser70ValfsTer3), et deux variations faux-sens (p.His89Glu et p.Ala109Pro). Les deux premières variations sont à la même fréquence qu'en population générale (soit retrouvées 2 fois dans la cohorte). La dernière variation, retrouvée une fois chez un patient hollandais, et prédite bénigne par SIFT et Polyphen, n'est pas présente dans les bases de données EXAC et EVS. Elle se situe dans un domaine de fonction inconnue.

Les données de séquençage n'ayant pas apporté d'information supplémentaire convaincante concernant le rôle de *PLA2G3* dans l'AAA l'étude de ce gène n'a pas été poursuivie dans les analyses ultérieures.

5.3 2 *APOL3*

Ce gène a été initialement retenu comme un bon candidat sur des données de la littérature : une étude ancienne réalisée par differential display a démontré son expression dans l'aorte et les artères iliaques normales et athéromateuses. Cette expression est importante dans les cellules endothéliales de l'intima et des *vasa vasorum*, modérée dans les cellules musculaires lisses et absente des macrophages de la plaque d'athérome. Le traitement par le TNF α de cellules endothéliales de type HUVEC induit une forte surexpression (x20) de l'ARN d'*APOL3* (figure 13, Horrevoets et al., 1999). Le TNF α est une cytokine proinflammatoire impliquée dans la pathogénèse de l'AAA, surexprimé dans un modèle murin de la pathologie, et dont l'inactivation inhibe la formation de l'anévrisme dans ce modèle (Xiong et al., 2009).

Apol 3 a donc été retenu comme un candidat potentiel dans la pathogénèse de l'AAA.

5.3.2.1 Données de la littérature

5.3.2.1.1 La famille des Apol

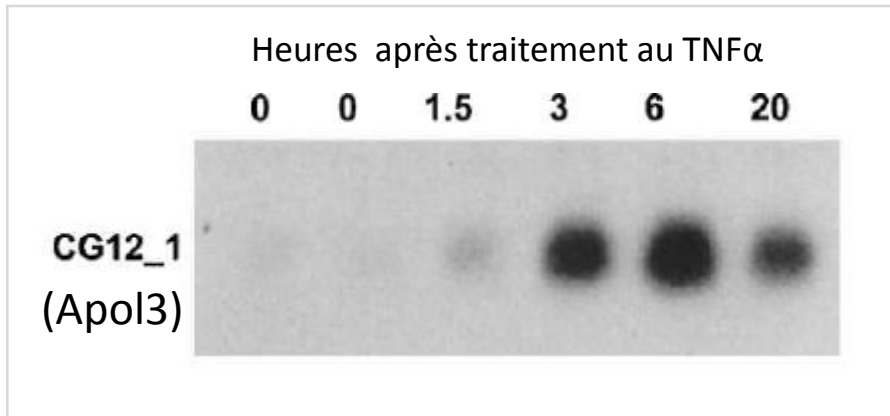


Figure 13: analyse en Northern-Blot de l' Induction de l'ARN *APOL3* après traitement par le TNF α dans les cellules endothéliales de type HUVEC. D'après Horrevoets et al, 1998.

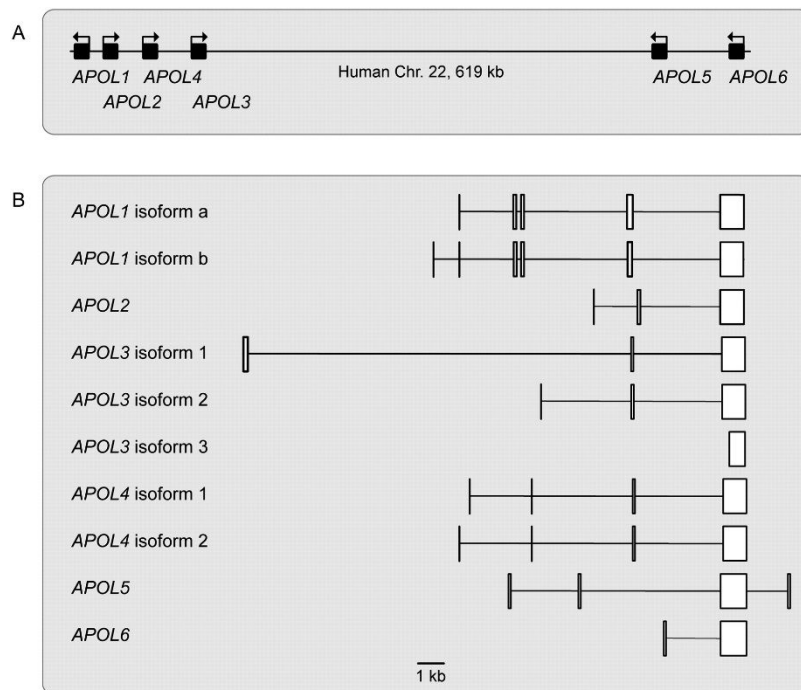


Figure 14: organisation génomique (A) et exonique (B) des gènes de la famille *APOL* humaine. D'après Smith et Malik, 2009.

Cette protéine appartient à une famille de 6 protéines humaines (APOL 1 à 6), dont les gènes sont organisés en un cluster d'environ 600 kb sur le chromosome 22 (figure 14). Cette famille de protéines a été appelée apolipoprotéines secondairement à l'identification d'APOL1 dans certaines particules de cholestérol HDL.

Les apolipoprotéines L ont évolué rapidement chez les primates (Smith et Malik, 2009), *APOL3* étant le paralogue ayant le moins rapidement évolué de cette famille. *APOL3* et *APOL4* en sont les membres les plus proches du point de vue phylogénétique (figure 15).

Les APOL appartiennent à la grande famille des protéines BCL2, qui partagent un ou plusieurs motifs « BH » (pour BCL2-homology), et sont impliquées dans la régulation de l'apoptose par perméabilisation de la membrane externe de la mitochondrie. Il existe 4 motifs BH (BH1 à 4), et les protéines de cette famille présentent généralement une combinaison particulière des différents motifs BH (Strasser, 2005, figure 16). Les membres de la famille BCL2 peuvent avoir une action anti- ou pro-apoptotique en fonction de la combinaison de ces motifs BH : toutes les protéines anti-apoptotiques portent les motifs BH1 et BH2, parfois associés à un motif BH4. *A contrario*, toutes les protéines pro-apoptotiques contiennent un motif BH3, impliqué dans l'activité cytolytique par hétérodimérisation avec des protéines anti-apoptotiques endogènes ou virales. Cette dichotomie entre domaines anti-et pro- apoptose est imparfaite puisque BH3 est parfois présent dans certaines protéines anti-apoptotiques comme BCL2.

Au sein de la grande famille des protéines BCL2, les protéines "BH3-only" ne contiennent qu'un domaine BH3 et sont donc très généralement pro-apoptotiques (comme les protéines Bim, Bid, BAD). On y retrouve toutes les APOL (Lomonosova et Chinnadurai, 2008, figure 17), et le rôle de BH3 dans l'induction de l'apoptose a été confirmé chez APOL1 et APOL6 (cf infra).

Les APOL partagent trois autres domaines conservés (figure 18) : le pore (ou Pore Forming Domain), région de grande taille (environ 180 aa) présentant des similitudes avec une protéine bactérienne appelée colicine, le domaine MAD (pour Membrane Addressing Domain), d'environ 55 aa, et le domaine SID (pour SRA Interacting Domain), d'environ 60 aa. Ce dernier domaine contient un motif spécifique en c-terminal de type « Leucine-Zipper » (LZ) d'environ 20 aa. Cette structure en hélice α fréquemment retrouvée dans les facteurs de transcription est un motif de dimérisation. Sa structure tridimensionnelle est secondaire à la répétition de leucines tous les 7 AA de ce motif, couvrant 8 tours d'hélice. Un motif LZ contient donc une hélice α , et se dimérise avec une seconde hélice en miroir, les leucines de la première hélice interagissant avec celles de la seconde hélice ce qui facilite la dimérisation.

En dehors des données anciennes de Horrevoets et coll. (1998) et des données de phylogénie sur le cluster des APOL, très peu d'informations sur le rôle d' *APOL3* sont disponibles dans la littérature. Une revue

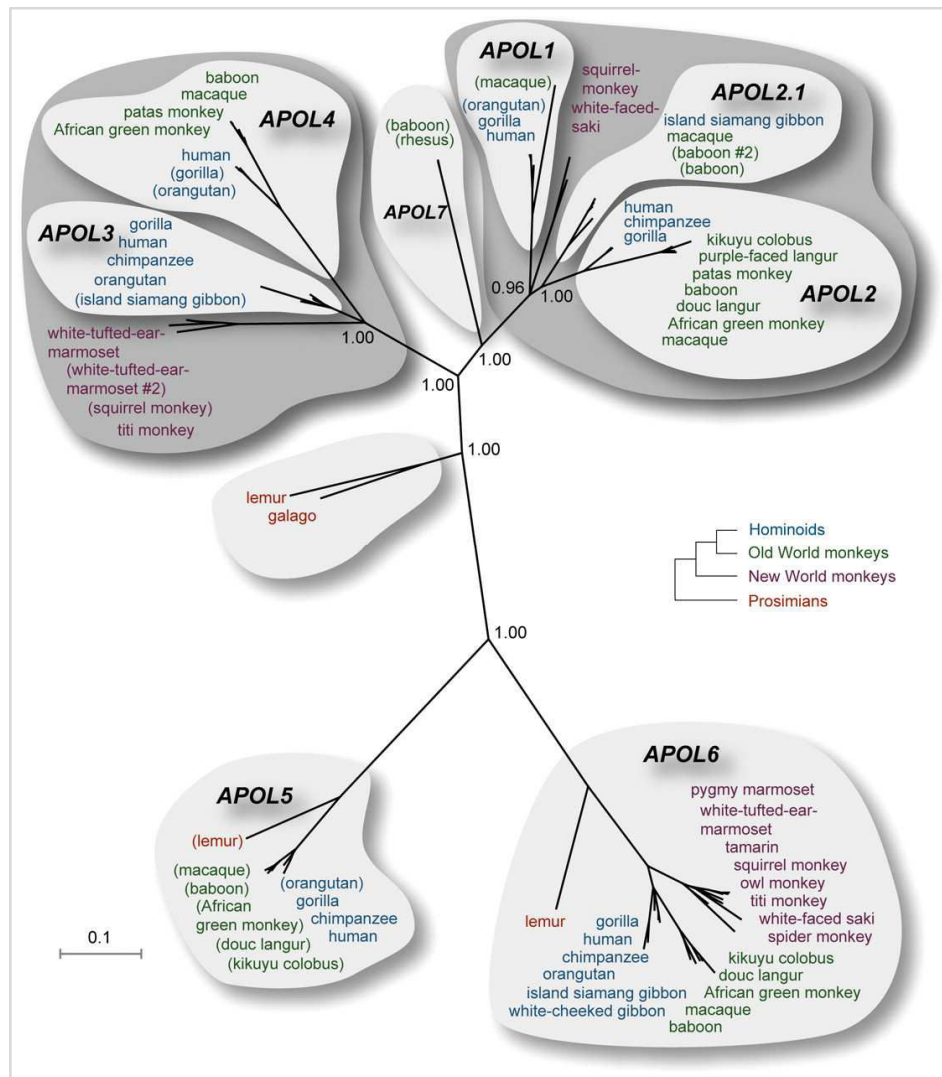


Figure 15: Phylogénie de la famille des gènes *APOL* chez les primates. D'après Smith & Malik, 2009.

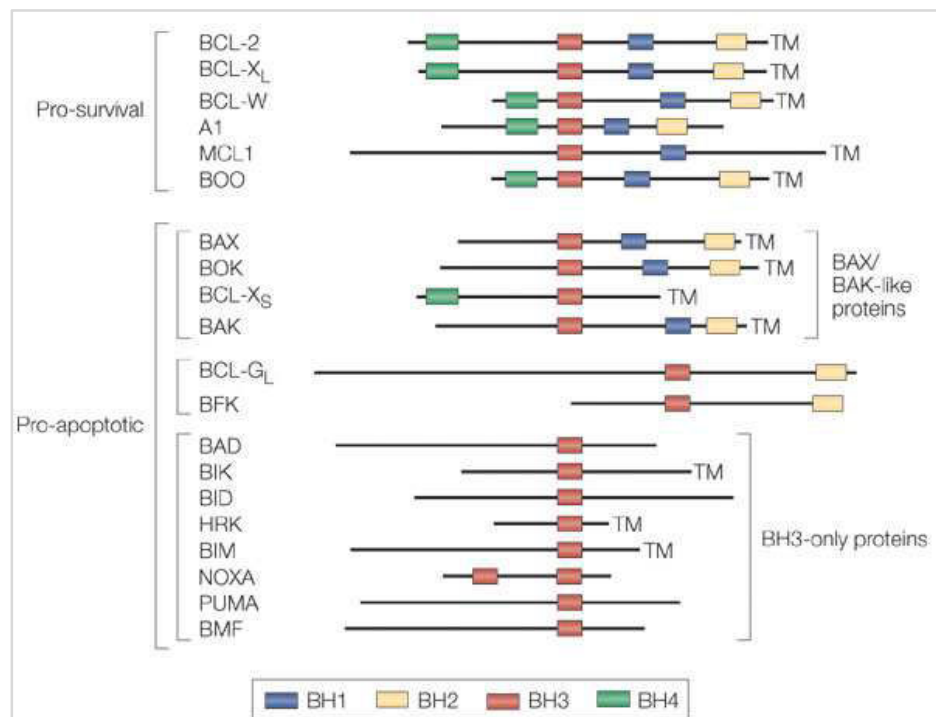


Figure 16: Membres de la famille de protéines BCL-2 chez les mammifères. D'après Strasser, 2005.

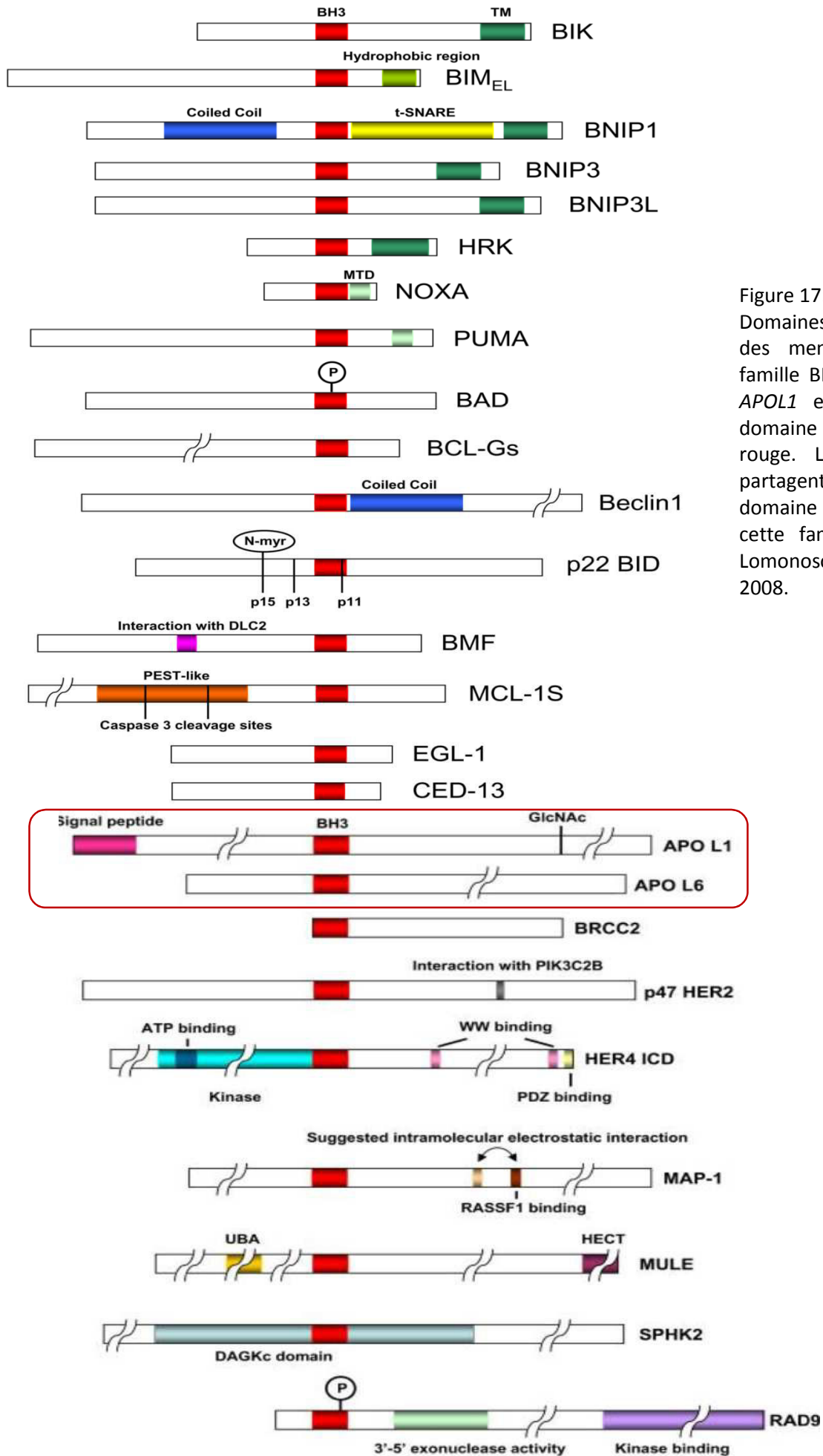


Figure 17: Domaines protéiques des membres de la famille BH3-only, dont *APOL1* et *APOL6*. Le domaine BH3 est en rouge. Les ApoL ne partagent aucun autre domaine connu de cette famille. D'après Lomonosova et al, 2008.

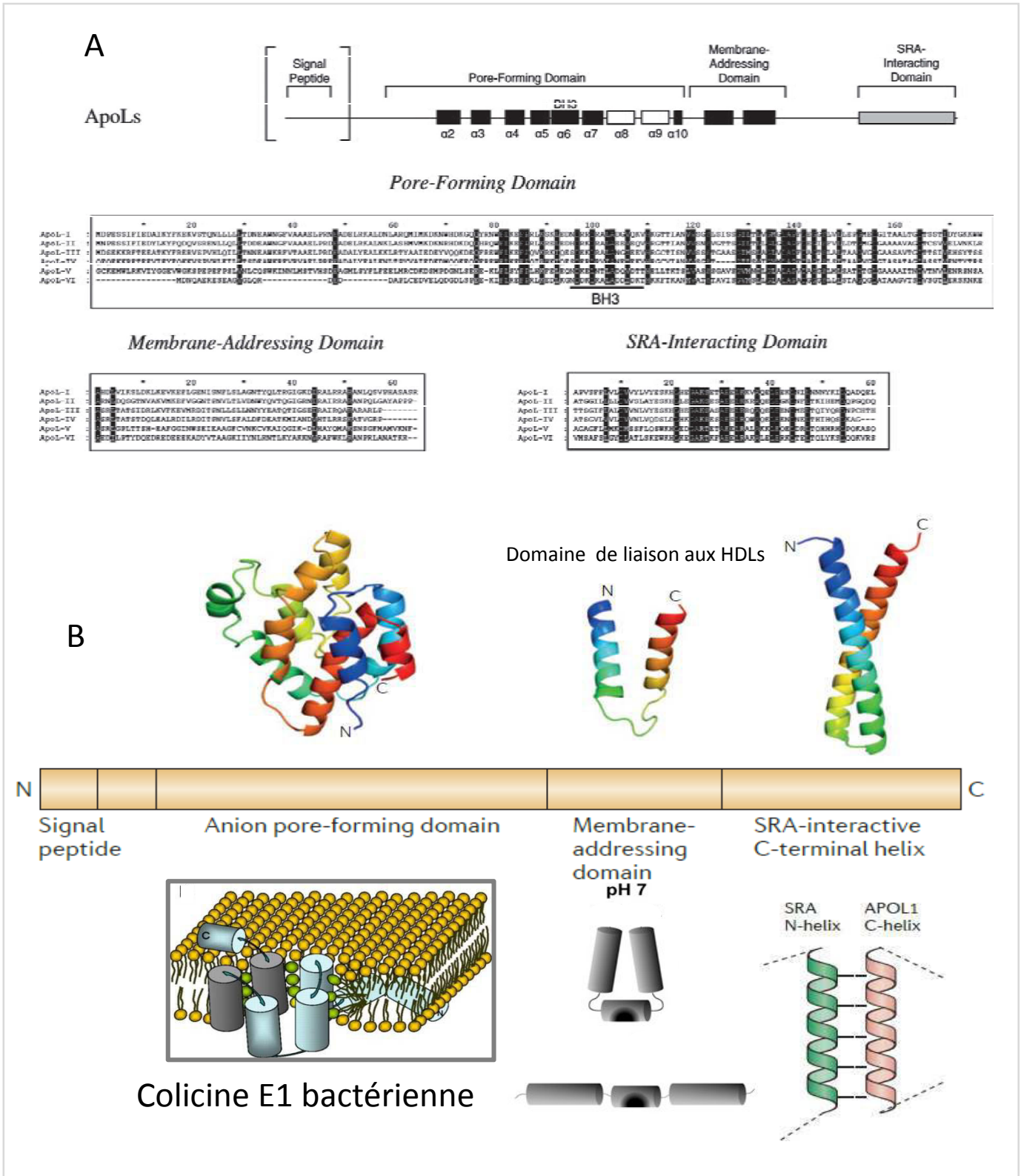


Figure 18: Domaines protéiques des membres de la famille APOL. Les trois domaines que sont le pore, le MAD et le SID sont conservés entre paralogues (A). (B) structure tridimensionnelle des domaines d'APOL1 (haut) et leur conformation présumée selon divers modèles (similitude, propriétés physicochimiques). D'après Vanhollebeke et Pays, 2014, Smith et Malik 2009.

exhaustive des données existantes sur les autres membres de la famille était donc nécessaire pour éclairer le rôle potentiel d'APOL3 dans l'AAA.

5.3.2.1.2 APOL1

Cette protéine a fait l'objet d'une caractérisation poussée dans le domaine de la parasitologie en raison de son rôle dans la résistance au trypanosome.

5.3.2.1.1.1 APOL1 et développement de la résistance aux trypanosomiasés

On distingue dans les trypanosomiasés, responsables de la maladie du sommeil, trois sous-espèces : *T. b. gambiense* responsable d'une forme chronique de développement lent de la maladie, *T. b. rhodesiense* responsable d'une forme aiguë voire fulminante, et *T. b. brucei*, responsable de zoonose mais qui n'infecte quasiment jamais l'homme. Cette résistance à *T.b.brucei* est liée à la fonction de la forme circulante d'APOL1, anciennement connue sous le nom facteur trypanolytique (TLF-1).

Contrairement aux autres APOL, APOL1 contient un peptide signal généré par épissage alternatif, et peut être sécrétée par diverses cellules (cellules endothéliales, macrophages, cellules T notamment) dans la circulation sanguine. Sa forme sérique est associée aux particules de cholestérol HDL par son domaine MAD (Vanhollebeke et Pays, 2006). Lors d'un contact avec *T.b.brucei* et son passage dans la circulation, APOL1 est phagocytée avec les particules de HDL par le trypanosome qui s'en « nourrit ». Elle est ensuite transportée par voie endosomale vers le lysosome, probablement ancrée à la membrane par son domaine MAD. A l'arrivée dans le lysosome, le pH acide modifie la conformation du domaine MAD qui se déploie, permettant au pore de s'intégrer à la membrane lysosomale et de laisser pénétrer les anions du cytosol dans le lysosome. Ce dernier se dilate puis libère ses protéases, induisant la mort cellulaire de *T.b.brucei* (Vanhollebeke et Pays, 2010, figure 19).

Par la génération de mutants ne contenant pas le pore, le MAD ou le SID, il a été démontré que chacun de ces trois domaines est nécessaire à l'action trypanolytique d'APOL1 (Lecordier et al., 2009). Le rôle du SID dans cette action n'est pas clairement identifié, bien que par définition son domaine LZ interagisse avec un partenaire protéique présent dans *T.b.brucei*.

Apol1 et trypanosomiase humaine aiguë

La forme sauvage d' APOL1 ne confère aucune résistance à *T. b. rhodesiense*, responsable de la forme aiguë de maladie du sommeil : Après contact et passage dans la circulation, APOL1 est phagocytée avec les particules de HDL par le trypanosome qui s'en « nourrit » de la même façon. Elle est transportée par voie endosomale vers le lysosome, où le pH acide permet le déploiement du domaine MAD. Une protéine des membranes internes de *T. b. rhodesiense* appelée SRA (pour Serum Resistance Agent, d'où le nom SID pour

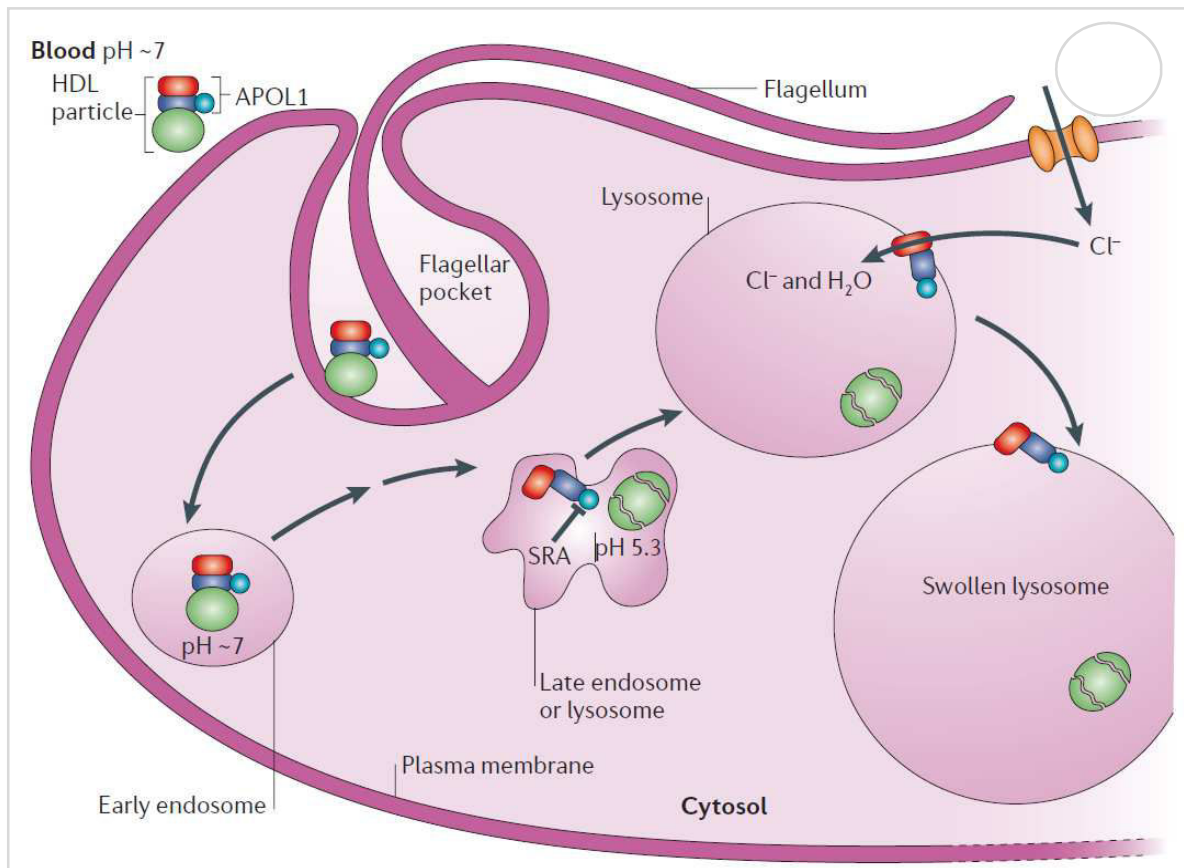


Figure 19: mécanisme moléculaire de l'effet trypanolytique d'APOL1 chez *T.b.brucei*. Le pore est représenté en rouge, en bleu foncé le domaine MAD, en turquoise le domaine SID, et en vert les particules de HDL. Un fois la particule de HDL associée à APOL1 phagocytée, elle est distribuée par trafic endosomal vers le lysosome et peut intégrer le pore à la membrane, permettant l'entrée des ions chlore et d'eau dans le lysosome qui gonfle, et libère ses enzymes dans le cytosol. D'après Vanhollebeke & Pays, 2010.

SRA-Interacting Protein), interagit avec le domaine LZ du SID. Cette interaction inhibe la fonction du pore par un mécanisme inconnu. APOL1 est ensuite dégradée dans le lysosome, et l'infection s'installe (figure 20).

Des variants naturels fréquents, localisés dans le domaine LZ interagissant avec le SRA, ont été identifiés en population africaine. Ces variants appelés G1 et G2 ont une fréquence cumulée très élevée de 22 et 13% respectivement, en comparaison des populations caucasiennes (1%, Genovese et al., 2010). Ils confèrent, in vivo une résistance à *T. b. rhodesiense*. Lecordier et coll. ont démontré in vitro ce phénomène de résistance : les mutations du SID altèrent l'interaction avec le SRA : la fonction du pore est restaurée et le trypanosome meurt (Lecordier et al., 2009). Très récemment, cette équipe a démontré qu'APOL1 était aussi transportée, via l'endolysosome, dans la mitochondrie de *T.b.brucei* et y induisait la perméabilisation de la membrane mitochondriale de façon dépendante de BH3. Ce phénomène correspond exactement à la transposition dans *T. b. rhodesiense* d'un processus apoptotique endogène (Vanwalleghem et al., 2015). Les différentes actions d'APOL1 sur les sous-espèces de trypanosome ainsi que la position des variants G1 et G2 dans le domaine LZ sont récapitulées dans la figure 21 (Pays et al., 2014).

La fréquence élevée des allèles G1 et G2 est clairement expliquée par leur sélection positive dans la résistance à cette maladie mortelle. Cependant, ils sont fortement associés dans la population africaine à une forme de glomérulosclérose focale, uniquement chez les individus homozygotes ou hétérozygotes composites (OR=10) mais pas les hétérozygotes simples. De façon inattendue, l'infection par le VIH des individus portant ces allèles de susceptibilité fait grimper l'odds-ratio à 30, témoignant d'un lien méconnu entre le VIH et ces variants (Genovese et al., 2013).

Ces derniers éléments ne sont pas sans rappeler le lien entre paludisme et drépanocytose, ou les hétérozygotes et homozygotes pour l'HbS sont résistants à *P.falciparum*, seuls les homozygotes étant atteints de drépanocytose.

5.3.2.1.1.2 APOL1 et cellules humaines : données de la littérature

Dans les podocytes, cellules épithéliales responsables de la filtration du plasma sanguin dans le glomérule rénal (figure 22), APOL1 peut être captée (forme sérique) ou induite par divers stimuli (notamment le TNF α). Dans ces cellules, la protéine APOL1 porteuse des variants G1 et G2 du SID induit le relargage du contenu lysosomal dans le cytosol, le gonflement, puis la nécrose des podocytes (figure 23). Ce processus est très fortement potentialisé par le VIH et le TNF α (Lan et al., 2014). Cette nécrose peut être inhibée par un traitement bloquant le trafic endosomal (comme la chloroquine) ou les canaux chlore.

Le rôle cytotoxique des variants G1 et G2 a été affirmé récemment par une étude de transgénèse dans des lignées de cellules 293T (cellules embryonnaires de rein) : l'introduction de ces variants induit un

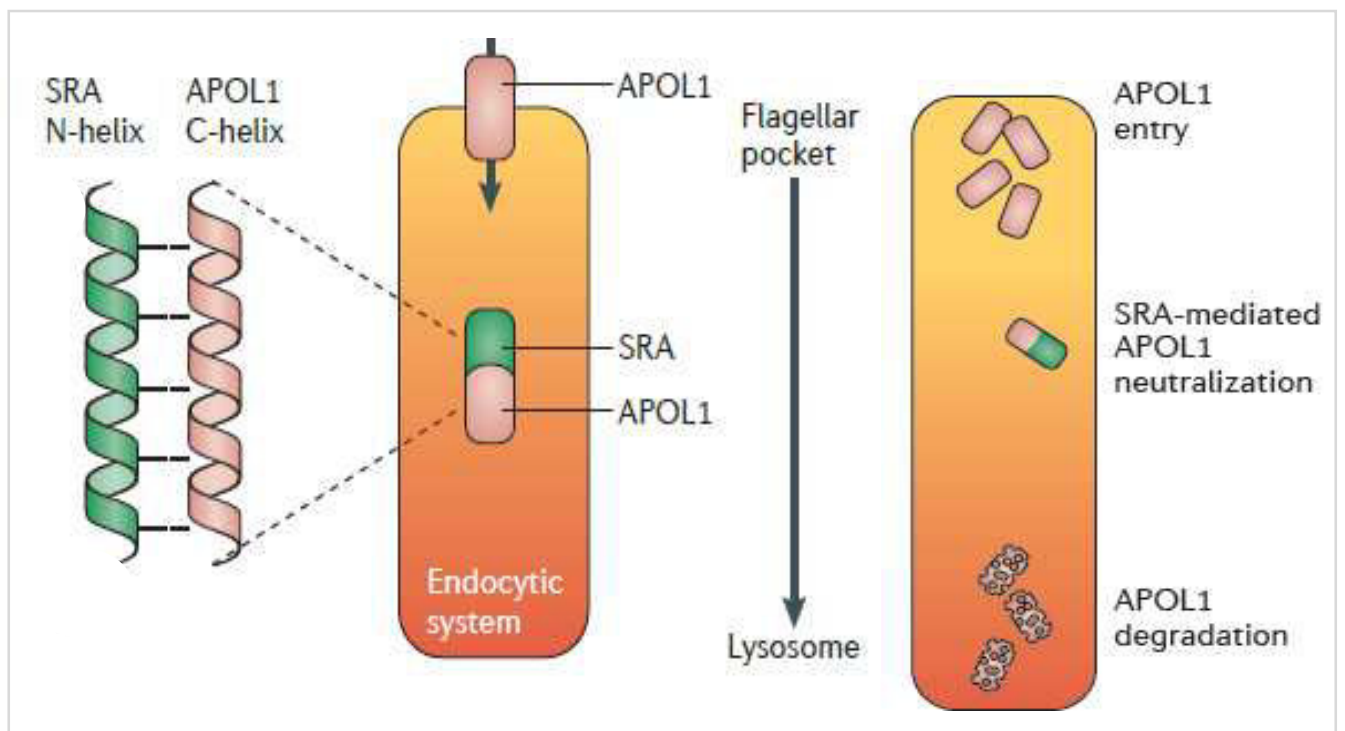


Figure 20: mécanisme moléculaire de la résistance à APOL1 chez *T.b.rhodensiense*. L'interaction SID-SRA prévient l'action du pore qui ne peut pas inactiver la fonction du lysosome. APOL1 est donc dégradée par voie lysosomale. D'après Vanhonnebeke & Pays, 2010.

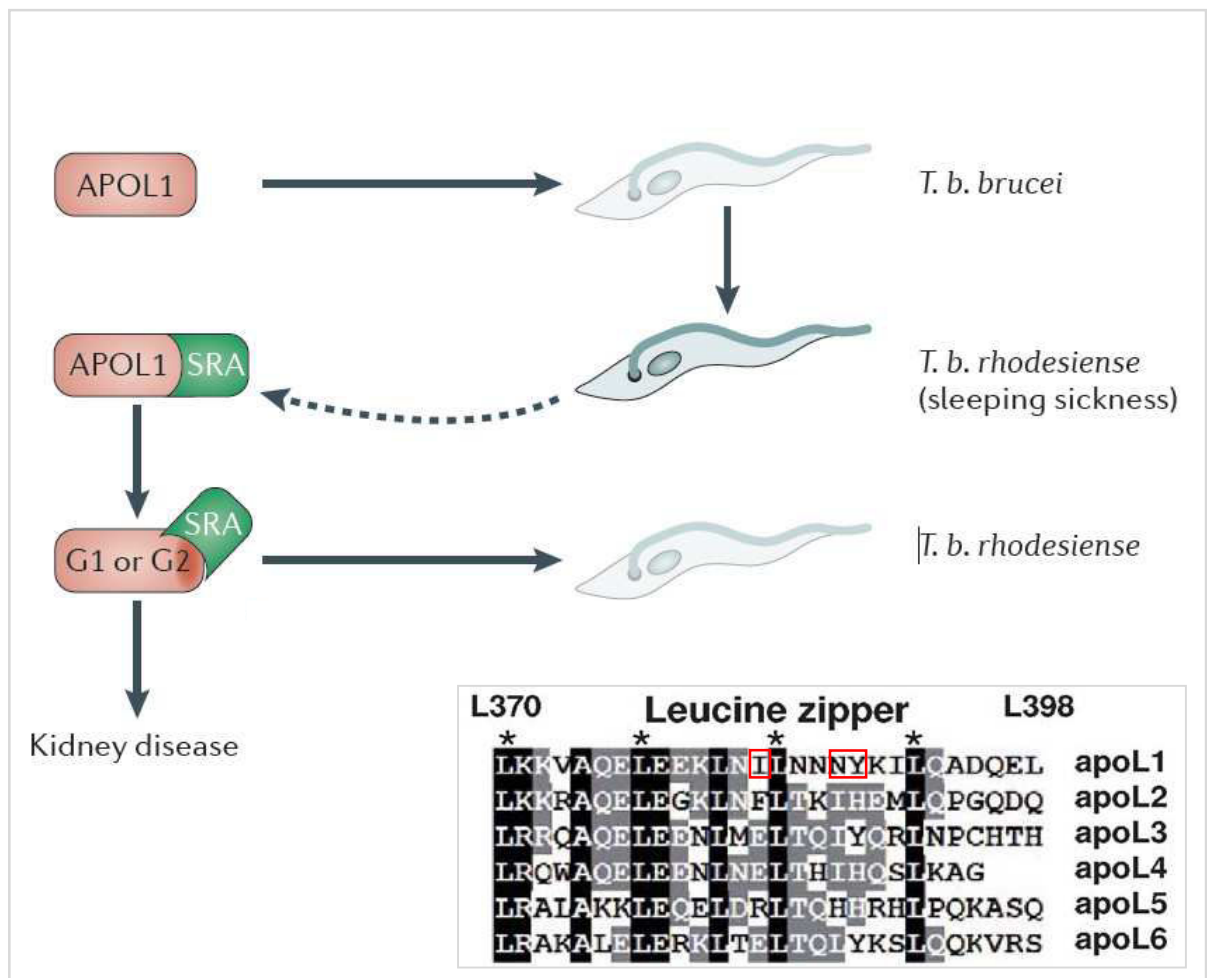


Figure 21: Différentes actions d'APOL1 selon la sous-espèce de trypanosome et le génotype correspondant. D'après Lecordier et al, 2009 et Pays et al, 2014. Cartouche: Les positions des Leucines du LZ sont représentées par des astérisques, les acides aminés concernés par les mutants G1 et G2 sont encadrés en rouge.

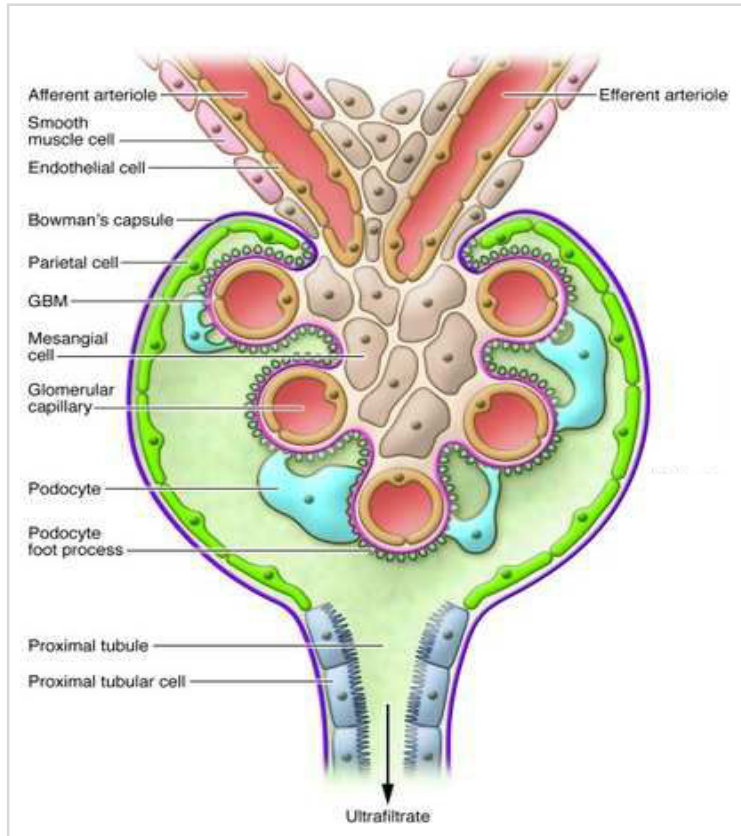


Figure 22: Structure du glomérule rénal. Les podocytes (en turquoise) entourent la pelote capillaire au centre de la structure. D'après Fogo et al, 2011.

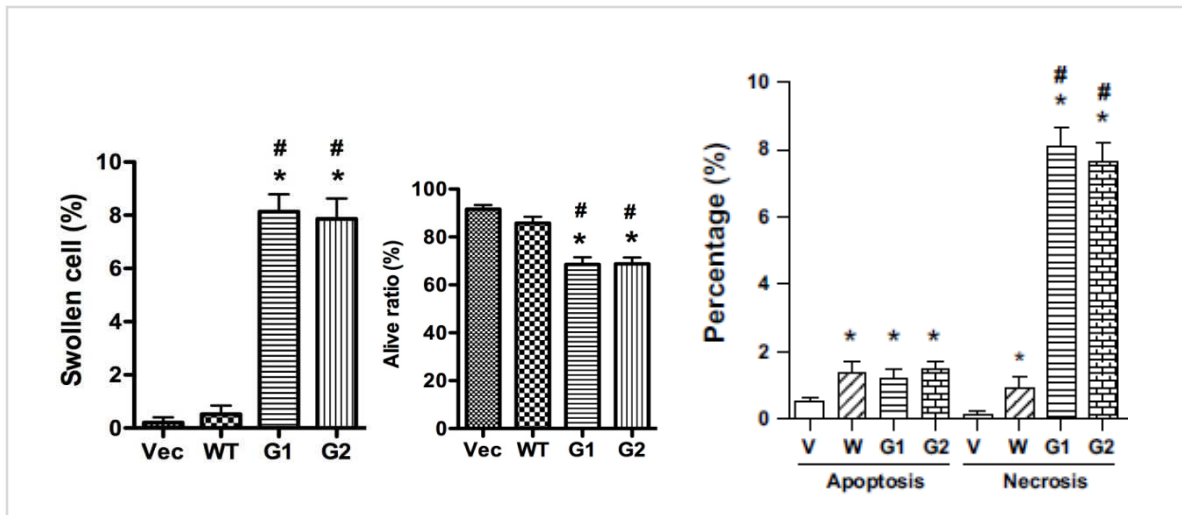


Figure 23: Effet des mutants G1/G2 sur le phénotype et la survie des podocytes à l'état de base. Les cellules gonflent puis meurent, le processus de nécrose semble impliqué dans cette perte cellulaire. Les cellules ont été transfectées avec *APOL1* humain sauvage, G1 ou G2 par vecteur viral. L'apoptose et la nécrose ont été évaluées par immunofluorescence utilisant des marqueurs spécifiques. D'après Lan et al, 2013.

gonflement et une mort des cellules, qui ne peut être contrebalancée par l'ajout de la forme sauvage D'APOL1 (Lan et al., 2015).

Dans les monocytes infectés par le VIH, APOL1 est induite par l'INF γ et limite l'infection virale de multiples façons (stimulation de l'endocytose et de la biogénèse lysosomale, induction de la dégradation de la protéine gag du VIH par voie endosomale, baisse de la réplication virale, Taylor et al., 2014).

De plus, il a été démontré dans de *multiples lignées cellulaires* qu'APOL1 induisait la mort cellulaire par la voie de l'autophagie, impliquant le lysosome et l'autophagosome, de façon clairement dépendante du domaine BH3 (Wan et al., 2008).

5.3.2.1.1.3 APOL1, APOL3 et cellules humaines : données issues de collaborations

Dans le cadre d'une collaboration sur *APOL3* avec l'équipe d'Etienne Pays (Genovese et al., 2010; Lecordier et al., 2009; Vanwalleghem et al., 2015), nous avons partagé des données **non publiées** sur la fonction des *APOL1* et 3.

Il apparaît des travaux réalisés par cette équipe, qu'*APOL1* est préférentiellement exprimé dans les podocytes, et dans une moindre mesure dans les cellules dendritiques, où il contrôle la durée de vie de ces cellules après stimulation virale :

Dans les podocytes, il est fortement induit par le VIH, ainsi que par stimulation du récepteur TLR3, récepteur de motifs ARN double-brin par le poly I:C (agent mimant l'ARN double brin). La voie TLR3 est donc une voie de la réponse antivirale. Cette induction a pour conséquence une baisse de la durée de vie des podocytes. L'inactivation par KO ou KD (par RNAi) d'*APOL1* restaure la survie de ces cellules. Au total, l'activation dans ces cellules par stimulus viral diminue la durée de vie des cellules, d'autant plus chez les mutants du SID, mais pas chez le KO. Ces résultats, associés aux données de coexpression en lignées cellulaires rénales des mutants G1, G2 et sauvage mènent à penser que les mutants G1/G2 agiraient par effet dominant négatif, dérégulant l'apoptose dans ces cellules (Lan et al., 2015).

L'activation (immunitaire) des cellules dendritiques par le poly I:C est associée à une induction d'*APOL1* qui est responsable d'une diminution de la survie des cellules (restaurée par inactivation transitoire en RNAi).

Enfin, dans la population indienne ont été identifiés des individus portant une forme totalement inactivée d'*APOL1*, sans phénotype particulier.

5.3.1.1.4 APOL2, APOL4, APOL5 et APOL6

Seul un article porte sur la fonction d'*APOL2*, mais il révèle que dans les cellules épithéliales bronchiques, *APOL2* est exprimé à l'état basal, et a une activité anti-apoptotique fortement induite par l'INF γ et le virus de la grippe (Liao et al., 2011). Il n'existe pas de données publiées concernant *APOL4* et 5. *APOL6* est surexprimé dans les cellules musculaires lisses vasculaires issues de lésions athérosclérotiques, où il induit la mort cellulaire par apoptose (Zhaorigetu et al., 2011).

5.3.2.2 Etude d'*APOL3* dans l'AAA

APOL3 est une protéine de 402 acides aminés qui contient les quatre domaines précédemment cités : le pore, contenant le domaine BH3, le domaine MAD et le domaine SID. Par épissage alternatif, une version plus courte d'*Apol3* (isoforme 2) existe, tronquant la partie N-terminale de la protéine de fonction inconnue, mais conservant les 4 domaines sus-cités.

La mutation identifiée dans la famille SE, substituant à l'Arginine en un Tryptophane (R377W) se situe dans le Leucine-Zipper du SID, à approximativement 10 acides aminés des équivalents G1 et G2 d'*APOL1* (figure 24). Ce changement d'acide aminé est prédit pour induire une perte de la structure tertiaire en « Zip » (figure 25). Par similitude avec *APOL1*, cette mutation pourrait induire une dérégulation de la protéine *via* une perte de l'interaction avec son partenaire, aujourd'hui encore inconnu.

5.3.2.2.1 Expression d' *APOL3* : Immunohistochimie, données issues de collaborations et de bases de données publiques

L'expression en immunohistochimie (IHC) d'*APOL3* en tissu artériel sain a été confirmée sur des prélèvements d'aorte, d'artère mammaire interne et poplitée (figure 26, A à C). L'expression d'*APOL3* (R377W hétérozygote) dans l'aorte thoracique de l'individu II-12 opéré d'une insuffisance aortique ne s'est pas révélée différente de l'expression attendue d'une homozygote « sauvage » (figure 26, D).

Les analyses réalisées sur *APOL3* par l'équipe d'Etienne Pays révèlent qu'il est exprimé fortement dans les cellules endothéliales aortiques et les cellules dendritiques, ou son induction est majeure par stimulus inflammatoire (TNF α) mais surtout viral par le poly I:C (figure 27). Ces données évoquent un rôle d'*APOL3* dans la défense antivirale, dans les cellules endothéliales et dendritiques.

L370	Leucine zipper										L398															
*	*	*	*	*	*	*	*	*	*	*																
LKKVAQELEEKLN	I	LN	NY	K	I	L	Q	A	D	Q	E	apoL1														
LKKFAQELEGKLN	F	L	T	K	I	H	E	M	L	Q	P	apoL2														
LR	Q	A	Q	E	L	E	N	I	M	E	L	T	Q	I	Y	Q	R	T	N	P	C	H	T	H	apoL3	
LR	Q	A	Q	E	L	E	N	I	N	E	L	T	H	I	H	Q	S	T	K	A	G			apoL4		
LR	A	L	A	K	K	L	E	Q	E	L	D	R	L	T	Q	H	R	H	P	Q	K	A	S	Q	apoL5	
LR	A	K	A	L	E	L	E	R	K	L	T	E	L	T	Q	L	Y	K	S	L	Q	Q	K	V	R	apoL6

Figure 24: Positions des mutations G1 et G2 d'APOL1 et de la mutation p.Arg377Trp d'APOL3. La mutation de la famille SE est au tout début du LZ. D'après Lecordier et al, 2009.

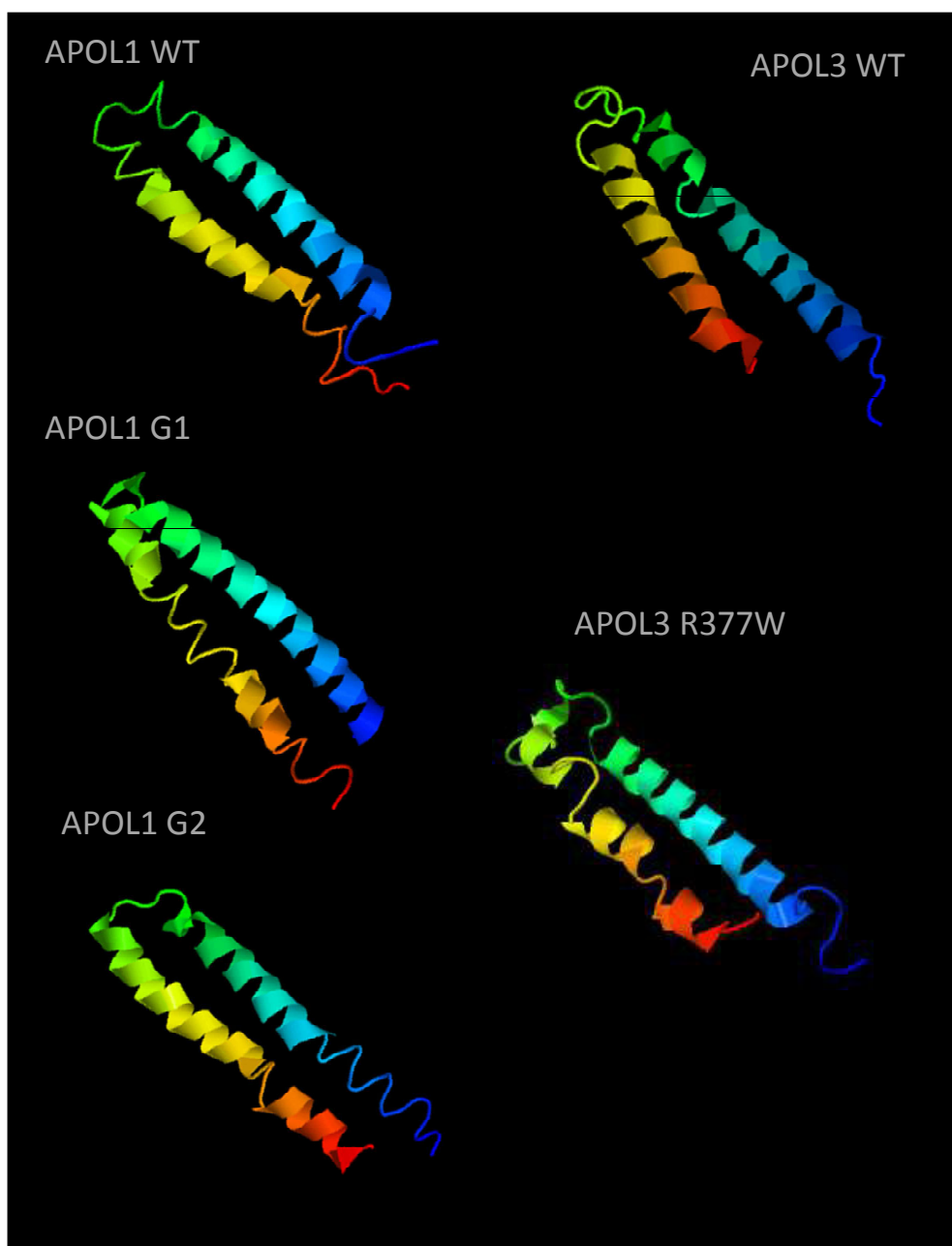


Figure 25: Prédiction des structures d'APOL1 et APOL3 sauvages, et mutants (G1 et G2 d'APOL1, p.Arg377W d'APOL3) avec l'outil I-Tasser.

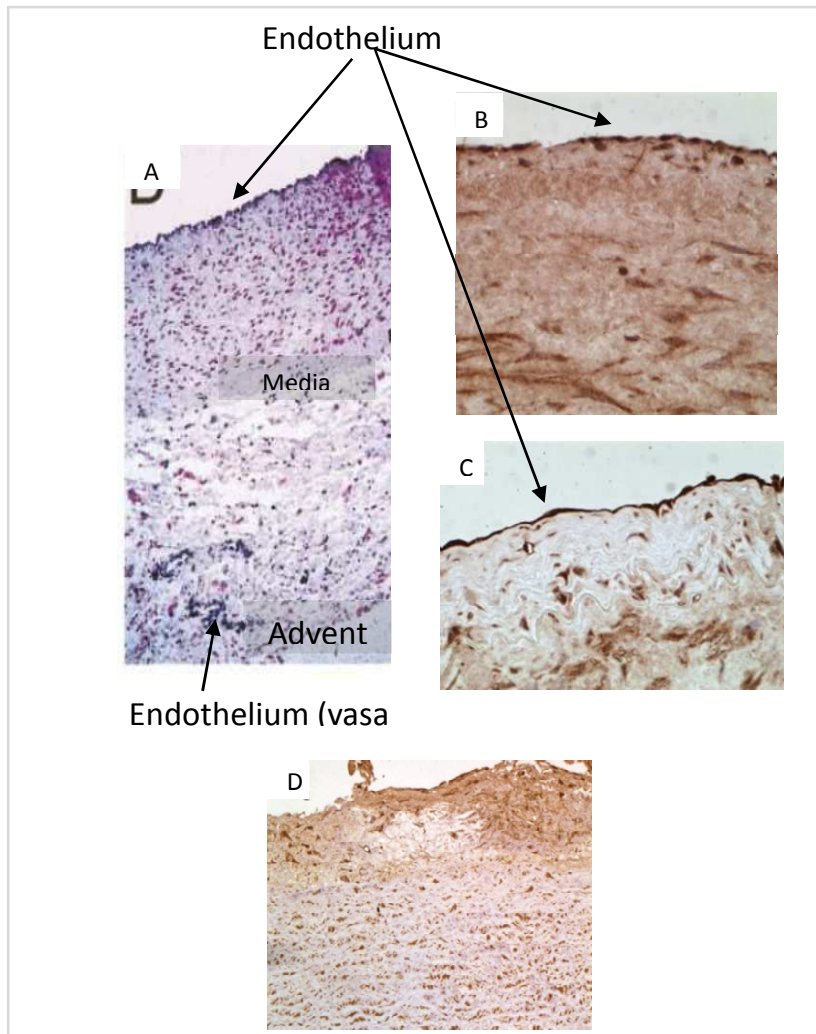


Figure 26: Expression protéique d'*APOL3* par immunohistochimie dans diverses artères normales et dans l'aorte ascendante de la patiente II-13. A: paroi aortique. B: Artère mammaire interne, C: artère poplitée. D: paroi aortique de II-12.

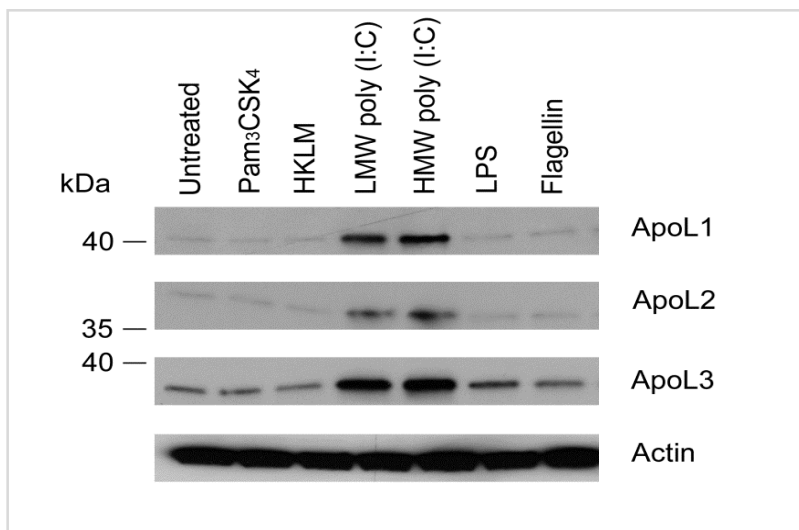


Figure 27: Effet de l'induction par diverses substances stimulant spécifiquement les voies des TLR sur l'expression d'*APOL1*, *APOL2* et *APOL3* en cellules dendritiques. Pam₃CSK₄ = lipoprotéine triacylée, stimulant des TLR 1 et 2. HKLM = *Listeria Monocytogenes* inactivé par la chaleur, stimulant du TLR2. LMW : Poly I:C de bas poids moléculaire (stimule TLR3). HMW= Poly I:C de haut poids moléculaire (stimule TLR3). LPS : Lipopolysaccharide (stimule TLR4). Flagelline: stimule TLR5. *APOL3* est majoritaire, et il est très fortement induit par stimulation de la voie TLR3, spécifique de l'infection virale. Collaboration E.Pays.

L'exploration de bases de données d'expression comme Expression Atlas est un bon moyen d'avoir une vision globale de l'expression d'un gène dans divers tissus et sous diverses conditions.

Cette analyse a permis d'identifier qu'*APOL3* peut être induit dans les cellules dendritiques par le lipopolysaccharide, molécule d'origine bactérienne se liant au récepteur TLR4 (récepteur de l'immunité innée) et par *Coxiella Burnetii* (Gorvel et al., 2014); dans les hépatocytes par le VHC ; dans les cellules endothéliales de cordon ombilical (HUVEC) par diverses souches de la grippe (Schmolke et al., 2009); dans la peau, les cellules endothéliales, les monocytes M1 (spécialisés dans la réponse antivirale), et l'épithélium nasal par l'INF γ (Indraccolo et al., 2007; Martinez et al., 2006). C'est un activateur de la voie NF κ B (Matsuda et al., 2003). Les données extraites de quatre études transcriptomiques distinctes semblent indiquer qu'il n'est pas différentiellement exprimé dans l'AAA (Choke et al., 2009; Lenk et al., 2006; Moxon et al., 2010; Zheng et al., 2002).

5.3.2.2 *APOL3* et cellules dendritiques

Les cellules dendritiques (ou DC) sont les cellules présentatrices de l'antigène les plus puissantes de l'organisme. Elles revêtent des rôles spécifiques en fonction notamment des marqueurs de surface qu'elles expriment.

Une extraction de données transcriptomiques issues d'un article traitant des sous-populations de cellules dendritiques a été menée (Robbins et al., 2008). L'analyse a mis en évidence qu'*APOL3* est fortement exprimé dans ces cellules, en comparaison des autres *APOL*. De plus, cette expression était relativement restreinte à une sous-population très minoritaire de DC appelée BDCA3+ (figure 28).

Ces dernières, comme toutes les cellules dendritiques, ont la capacité de pouvoir phagocytter des cellules nécrotiques ou apoptotiques infectées par des virus sans être infectées elles-mêmes, puis stimuler de façon atypique une réponse de type cellulaire cytotoxique (médiée par les lymphocytes TCD8+), généralement secondaire à une infection virale de la cellule présentant l'antigène. Ce phénomène s'appelle la présentation croisée. Les cellules BDCA3+ sont spécialisées dans ce type de présentation d'antigène lorsqu'elles reçoivent des **signaux de nécrose** des cellules environnantes (van der Aa et al., 2015, figure 29).

Elles expriment de plus **spécifiquement l'INF γ** , puissante cytokine antivirale par inhibition directe de la réplication de nombreux virus (grippe, VHC, VIH, rotavirus, HSV, CMV... Van der AA et al, 2015). L'INF γ induit également une réponse de type Th1 des lymphocytes T helpers, favorisant l'immunité cellulaire, la stimulation des lymphocytes TCD8 cytotoxiques, en réponse à un pathogène intracellulaire favorisant la disparition des cellules infectées par des virus (van der Aa et al., 2015, figure 29).

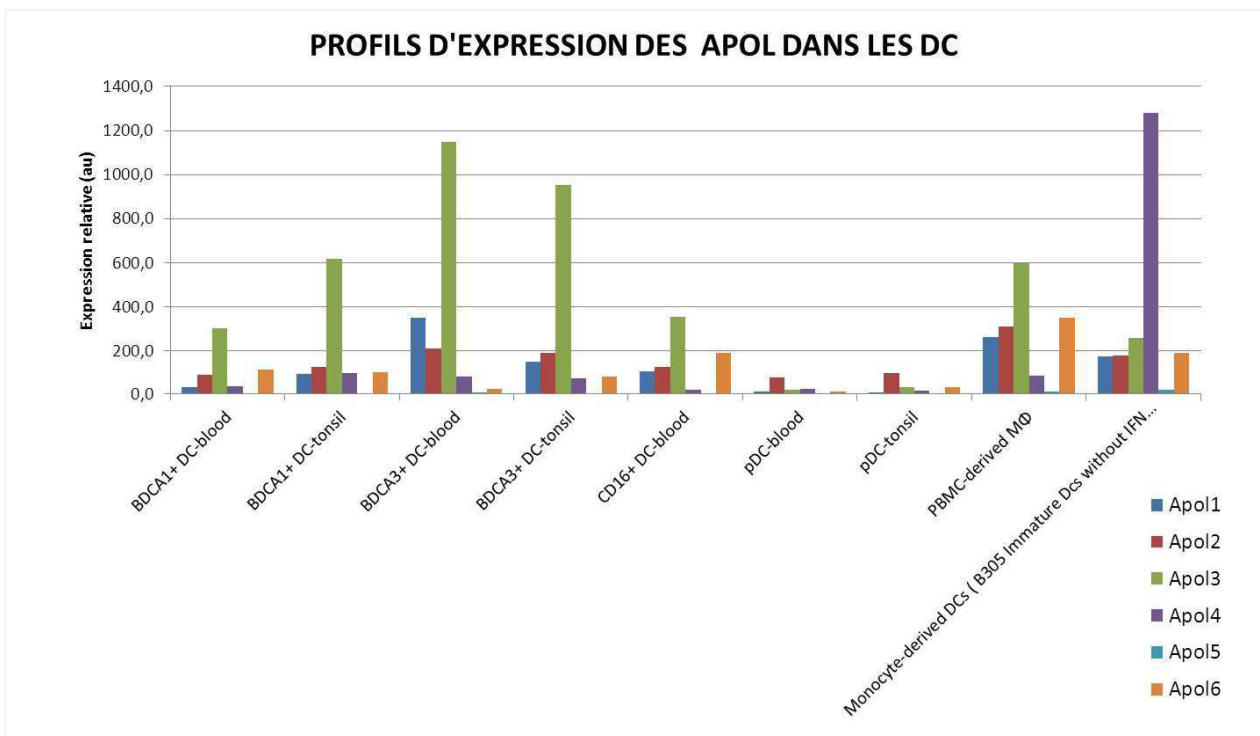
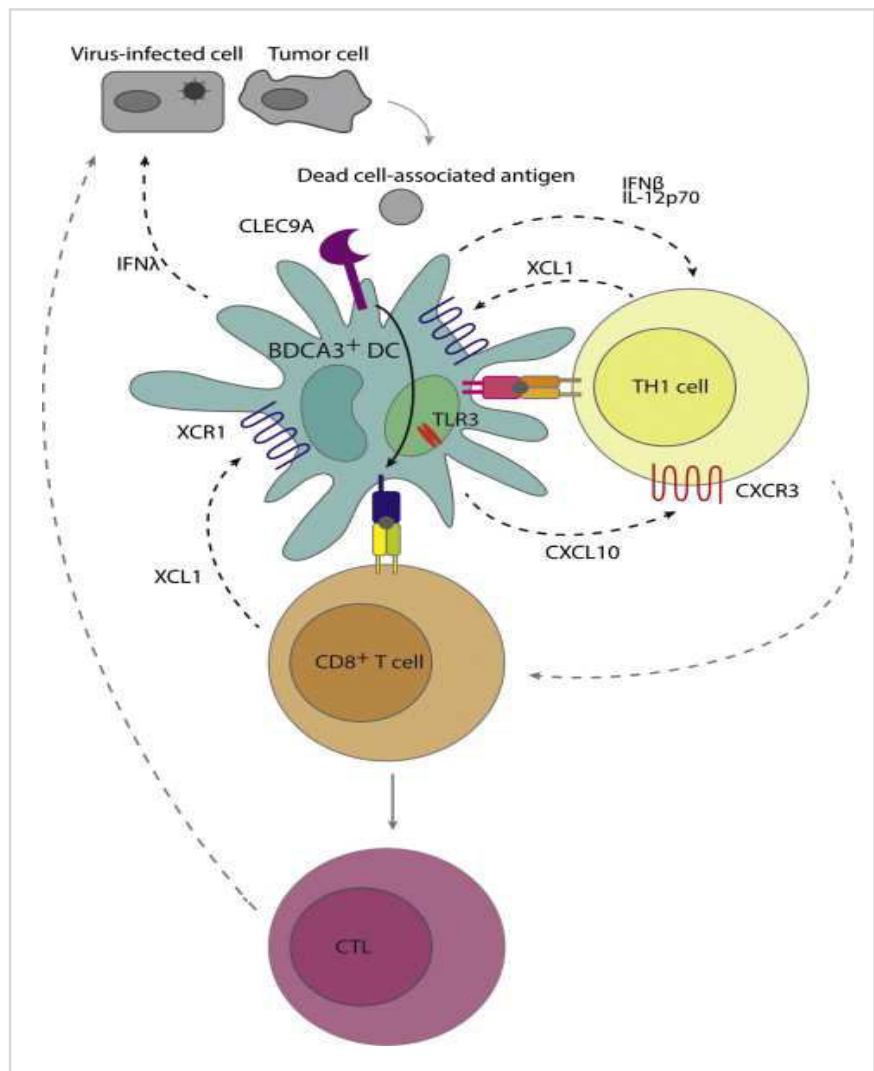


Figure 28: profils d'expression des Apol 1 à 6 dans les sous-populations de cellules dendritiques. Données extraites d'une étude transcriptomique (système Affymetrix) de Robbins et al, 2008. On note la prédominance d'APOL3 et sa relative spécificité pour les cellules BDCA3+.

Figure 29: La spécialisation fonctionnelle des cellules BDCA3+ suggère un rôle dans la défense antivirale. Le récepteur CLEC9A, spécifique de cette sous-population, spécialisé dans les signaux de nécrose. TLR3, récepteur non spécifique reçoit un signal d'infection virale. Le récepteur XCR1 au ligand des cellules CD8+ est spécifique des BDCA3+ . D'après Van der Aa et al, 2015.



5.3.2.2.3 Modèle explicatif du rôle d'Apol3 dans l'AAA

Tous ces éléments rassemblés permettent de dégager les points communs et divergents de cette famille de protéines : les APOLs sont des protéines intervenant dans la régulation de la mort cellulaire, notamment de l'apoptose mais pas uniquement (Lan et al, 2003). Cette régulation se fait par l'intermédiaire du domaine BH3, mais aussi du domaine SID, qui interagit dans les cellules endogènes avec un partenaire encore inconnu, possiblement d'origine virale. Les domaines MAD et le pore, par extrapolation de la fonction trypanolytique d'APOL1 (Vanwalleghem et al., 2015), pourraient servir à la fonction de pore lysosomal et/ou mitochondrial induisant l'apoptose par perméabilisation des membranes correspondantes.

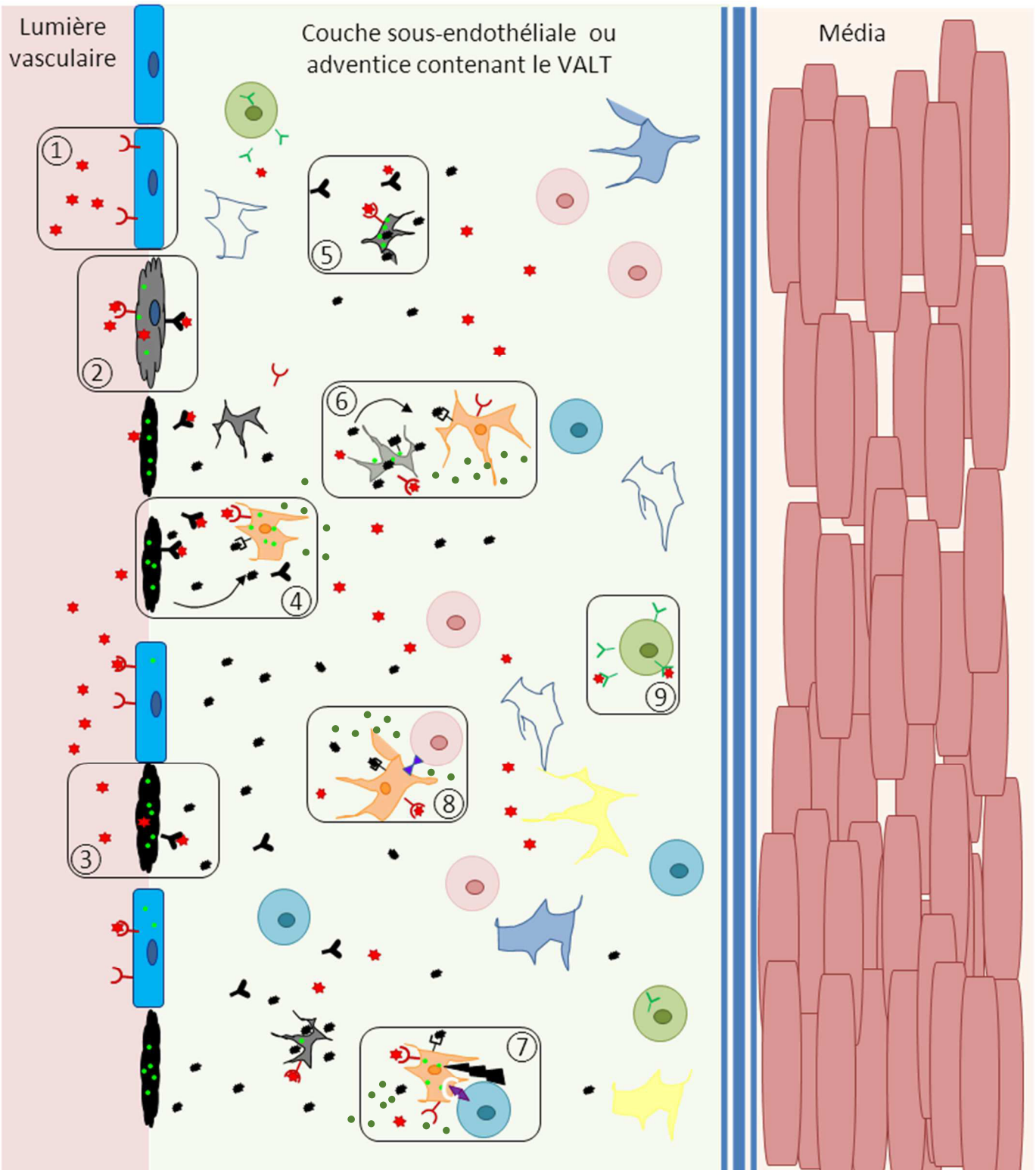
Ainsi, *APOL1* et *3* sont exprimées dans différents types cellulaires qui ont plusieurs points communs : elles sont toutes immunocompétentes car **ce sont des cellules présentatrices de l'antigène**, dites semi-professionnelles pour les cellules endothéliales et professionnelles pour les podocytes (Goldwich et al., 2013; Razakandrainibe et al., 2012). Ces cellules **répondent toutes à un stimulus viral par une surexpression de l'APOL1 ou 3 selon leur type**, induisant une **diminution de la survie de ces cellules**. Concernant APOL1, il a été prouvé que cette mort cellulaire accélérée pouvait être autophagique, apoptotique ou équivalent, et nécrotique dans le cas de mutants « dérégulés » du domaine SID (Lan et al., 2014; Vanwalleghem et al., 2015; Wan et al., 2008).

APOL3 est de même probablement impliquée dans des processus d'apoptose ou de nécro-apoptose comme les autres membres de sa famille mais ceci n'a encore jamais été confirmé expérimentalement.

De l'ensemble de ces éléments, il est possible de dégager un modèle putatif du rôle d'APOL3 dans l'AAA (figure 30, points ① à ⑨) : l'expression de la version dérégulée par la mutation R377W, induite initialement par une infection virale localisée (par exemple le CMV, ①) serait responsable d'une baisse de la survie des cellules endothéliales (② puis ③) par induction de l'apoptose et/ou de la nécrose par analogie avec les mutants G1/G2 dans les podocytes (Lan et al., 2014). Cette mort cellulaire siègerait dans l'intima mais aussi dans les *vasa vasorum* de l'adventice, stimulant le tissu lymphoïde associé aux vaisseaux, ou VALT, présent dans la couche sous-endothéliale et l'adventice de l'aorte (voir section 3.7 L'immunité). Cette perte cellulaire associée à la présence d'antigènes viraux stimulerait la réponse antivirale, médiée par les cellules dendritiques et particulièrement la sous-population BDCA3+ (④) spécialisée dans les stimuli viraux associés à la mort cellulaire. Sa stimulation par le signal de mort de l'endothélium induirait fortement et spécifiquement l'Apol3 y compris dans sa version dérégulée, responsable d'un phénomène d'amplification de la mort cellulaire (par analogie aux mutants G1/G2 du SID dans les podocytes) incluant les DC (⑤ et ⑥). Cette amplification serait responsable d'une activation cytotoxique majeure (cellules TCD8+ ⑦) directe, ou par stimulation par l'INF γ de la cascade immune antivirale cellulaire (cellules TH1 ⑧), suivie d'une réaction

Endothelium

Limitante élastique



- | | | | | | | | |
|--|---------------------------------------|--|--|--|---|--|--|
| | Cellule BDCA3+ | | Cellule endothéliale | | XCL10/XCL1 (communication entre BDCA3+ et Th1) | | Lymphocytes B |
| | Autres types de cellules dendritiques | | Cellule endothéliale nécrotique ou apoptotique | | XCLR1/XCL1 (communication entre BDCA3+ et Th1) | | Ly. T CD8+ cytotoxiques |
| | | | Apol3 (R377W) | | Récepteur CLEC9A aux signaux de mort cellulaire | | Mort cellulaire induite par les T CD8+ |
| | | | | | Anticorps | | Particule virale |
| | | | | | TLR3 | | |
| | | | | | CMH1 | | |
| | | | | | INFγ | | |

Figure 30: Modèle physiopathologique de l'atteinte des cellules de la paroi aortique par dérégulation d'APOL3.

humorale (©). De façon intéressante, ce modèle impliquerait un rôle délétère du mutant du SID dans deux populations de cellules en interaction.

5.4 Etude d'association au locus *ApoL3* et recherche de variants rares

5.4.1 Cohortes constitutives

Une collaboration a été mise en place avec l'équipe du Pr Nilesh Samani et du Dr Matt Bown (Départements de Cardiologie et de Chirurgie, Université de Leicester, UK). Cette équipe a identifié par des approches de GWAS deux gènes impliqués dans l'AAA (les gènes *LDLR* et *LRP1*, section 4.3 GWAS) à partir d'une cohorte d'environ 1 800 cas d'AAA dont 220 cas familiaux non apparentés.

5.4.2 Résultats du séquençage Sanger

Le séquençage Sanger d'*APOL3* a été réalisé chez les 1 800 cas de cette cohorte ainsi que les 137 cas issus des deux cohortes précédentes, soit presque 2 000 patients et environ 18% de cas familiaux.

J'ai identifié **6 autres variants très rares** dont 2 dans le domaine LZ du SID (figure 31). Ces variants sont tous des faux-sens, identifiés dans des cas sporadiques à l'exception d'un seul (p.Ser315Leu). L'utilisation de SIFT et de Polyphen est concordante pour quatre d'entre eux : l'un est prédit délétère (p.Ser230Cys), trois sont prédits bénins (p.Ile326Thr, p.Arg395His, p.Cys399Gly). Les résultats sont discordants pour 2 variants (p.Ala303Asn et p.Ser315Leu, tableau 7).

Les deux variants du domaine LZ sont prédits bénins par ces outils.

Par ailleurs, **7 variants rares** ont été identifiés (MAF entre 1% et 1‰ chez plusieurs cas sporadiques de cette cohorte). La fréquence allélique des variants fréquents a été extraite à partir des données de séquençage. L'ensemble des variants identifiés dans la cohorte ainsi que les fréquences alléliques correspondantes sont listés dans le tableau 8.

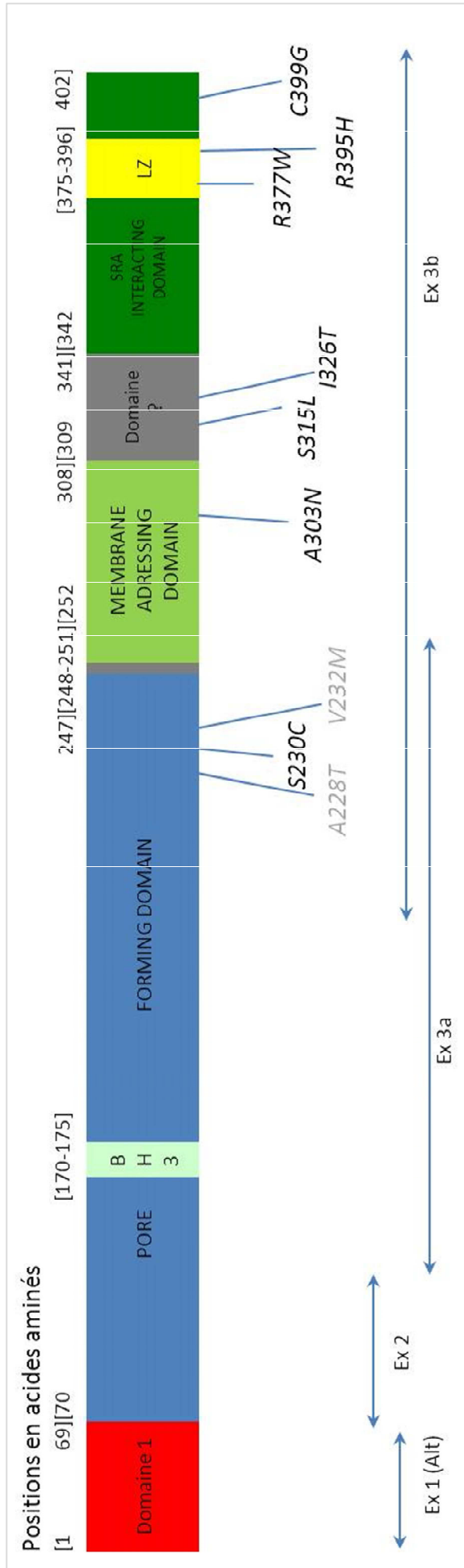


Figure 31: Protéine APOL3: localisation des 7 variants très rares (en noir) et des deux variants rares et fréquents associés à l'AAA (en gris).

5.4.3 Tests d'association des variants rares

Une comparaison de fréquence des variants rares et fréquents a été menée, entre les 1 800 cas de la cohorte et une population contrôle de 690 témoins de l'étude SUVIMAX (Hercberg et al, 2005). Les données issues de génotypage de ce groupe témoin par exome-chip (Infinium Human Exome-12 v1.2 BeadChip d'Illumina), étaient disponibles dans notre laboratoire. Au locus *APOL3*, 15 SNP étaient identifiés dans ces données. Parmi eux, 6 variants rares et 3 variants fréquents ont pu être utilisés pour la comparaison. L'outil SKAT-O a été utilisé, avec les options SKAT (variants rares uniquement) et SKAT CR (incluant les variants fréquents). Différents modèles ont été testés (cumulé, dominant, récessif) en distinguant les cas sporadiques des cas dominants. L'analyse a donc inclus 9 variants dont 3 fréquents, et mis en évidence des résultats significatifs (tableau 9).

L'analyse ultérieure variant par variant a confirmé l'association significative de deux variations, l'une rare (p.Ala228Thr, MAF 1,03‰ et 5,1‰ pour les cas et les contrôles respectivement) et l'autre fréquente (p.Val232Met, MAF 5,4% et 7% pour les cas et les contrôles respectivement) dans cet effet protecteur (tableau 10).

Le variant p.Ala228Thr présente un effet protecteur similaire en modèle additif et dominant, modèles non distinguables lors d'une étude centrée sur un seul gène. Cet effet semble d'autant plus significatif que l'ensemble des cas familiaux et sporadiques est utilisé, phénomène lié à l'effet de taille des groupes (Odds-ratio=0,20, intervalle de confiance= 0,051 - 0,78, p=0,021 ; tableau 11). Pour le variant fréquent p.Val232Met, l'association négative avec l'AAA semble la plus significative avec les cas sporadiques (pour les mêmes raisons), dans un modèle additif (Odds-ratio=0,74, intervalle de confiance= 0,56 - 0,96, p=0,026 ; tableau 11).

Ces deux variants sont proches dans la séquence protéique d'Apol3. Ils sont tous deux situés dans le pore, à distance du domaine BH3. Ils sont aussi très proches du variant très rare identifié p.Ser230Cys prédit délétère par Sift et Polyphen, et identifié une fois dans la cohorte de cas sporadiques du Royaume-Uni. Ces trois variants ne correspondent à aucun site connu de modification post-traductionnelle (voir Annexe 1), ni dans une région soumise à sélection positive (Smith et Malik, 2009).

Patient	variant dbSNP	c.	Prot	Domaine protéique	Fréquence dans EXAC	SIFT	Polyphen	Phenotype du patient
AAA1278	-	c.689C>G	p.Ser230Cys	Pore	0	Deleterious	probably pathogenic	Diagnostic fortuit à 80 ans, AAA 5,5cm. Ancien fumeur. Opéré en 2009 (endovasculaire). ATCD familial non connu
AAA1102	-	c.907_908GC>AA	p.Ala303Asn	MAD	0	Tolerated	Possibly damaging	Diagnostic fortuit à 77 ans, AAA 4,7 cm. Hypertendu, ancien fumeur. Opéré en 2009 (endovasculaire). ATCD familial non connu
AAA0493	rs201390959	c.944C>T	p.Ser315Leu	transition entre MAD et SID	8/1 million	Tolerated	Possibly damaging	AAA rompu à 63 ans. Non fumeur. Mère atteinte d'AAA. Opéré par voie endovasculaire puis réopéré par voie externe. ATCD familial non connu
AAA1502	rs143477399	c.977T>C	p.Ile326Thr	transition entre MAD et SID	5/100 000	Tolerated	Benign	Diagnostic fortuit à 80 ans, AAA 3 cm. Hypertendu, ancien fumeur. Opéré pour AAA à 5,8 cm. ATCD familial non connu
SE	rs145559179	1129C>T	p.Arg377W	LZ	8/100 000	Damaging	probably pathogenic	Forme dominante, famille SE
AAA227	rs145559179	c.1184G>A	p.Arg395His	LZ	8/100 000	Tolerated	Benign	Diagnostic à 66 ans, AAA 3,1 cm. Hypertendu, dyslipidémique, ancien fumeur. ATCD familial non connu
AAA0112	-	c.1195T>G	p.Cys399Gly	SID (3 AA après LZ)	0	Tolerated	Benign	Diagnostic fortuit à 78 ans, AAA 5,1 cm. Dyslipidémique, ancien fumeur, ATCD d'Infarctus du myocarde. Pas d'ATCD familial d'AAA. Opéré par voie externe

Tableau 7: Liste des variants très rares id'APOL3 identifiés dans la cohorte de cas sporadiques et familiaux d'AAA, incluant le variant de la famille SE. Leur localisation protéique, leur fréquence, les prédictions selon SIFT et Polyphen ainsi que les phénotypes correspondants sont détaillés.

identifiant dbSNP	c.	p.	Fréquence allélique					Disponibilité Exome-Chip Illumina
			cas sporadiques	cas familiaux	Fréquence cumulée	caucasiens cohorte ESP (~17200 allèles)	caucasiens cohorte EXAC (~132800 allèles)	
rs916334	c.-25G>A	p.?	3,91E-01	3,68E-01	3,89E-01	3,83E-01	4,01E-01	N
rs148712729	c.38A>G	p.Ser20Gly	0	0	0	0	0	O
rs150975698	c.75G>C	p.Val25Val	3,44E-03	0	3,16E-03	1,16E-03	8,42E-04	N
rs148885226	c.104G>T	p.Gly35Val	0	0	0	1,05E-03	1,35E-03	O
rs132653	c.117C>A	p.Ser39Arg	8,80E-01	8,24E-01	8,76E-01	8,26E-01	8,31E-01	O
rs11089781	c.172C>T	p.Gln58*	0	0	0	5,81E-04	2,85E-04	O
rs79419411	c.211A>G	p.Arg71Gly	0	3,88E-03	3,76E-04	2,33E-04	1,95E-04	O
rs45577431	c.223+19G>A	p.?	6,37E-02	4,40E-02	6,21E-02	5,86E-02	5,76E-02	N
rs132628	c.224-178G>A	?	NA	NA	NA	?	?	N
rs201432134	c.304A>G	p.Asn102Asp	0	0	0	0	0	O
rs141517147	c.335C>T	p.Ala112Val	0	0	0	3,49E-04	4,50E-05	O
rs373967437	c.350+31G>A	p.?	5,05E-04	0	4,62E-04	1,16E-04	1,06E-04	N
rs35695019	c.350+59del	p.?	3,33E-02	2,40E-02	3,20E-02	?	?	N
rs9510409	c.350+73C>G	p.?	1,21E-02	1,61E-02	1,25E-02	?	?	N
rs80575	c.351-1648C>A	p.?	NA	NA	NA	?	?	N
rs6000152	c.404C>T	p.Ala135Val	0	0	0	1,05E-03	1,05E-03	O
rs145787867	c.505A>C	p.Ile169Gln	0	0	0	1,16E-04	3,00E-05	O
rs3827346	c.564T>C	p.Asn188Asn	1,77E-01	1,54E-01	1,76E-01	1,88E-01	1,84E-01	N
rs140252051	c.586G>A	p.Ala195Thr	4,17E-03	0	3,77E-03	5,12E-03	4,36E-03	O
rs116147257	c.659C>T	p.Thr220Ile	0	0	0	3,49E-04	2,11E-04	O
rs142057520	c.682G>A	p.Ala228Thr	1,27E-03	0	1,13E-03	2,21E-03	3,81E-03	O
rs61741884	c.694G>A	p.Val232Met	5,80E-02	5,81E-02	5,80E-02	6,31E-02	5,55E-02	O
rs148518852	c.508G>A	p.Glu241Lys	0	0	0	0	0	O
rs61731692	c.732C>T	p.Tyr244Tyr	1,57E-01	1,01E-01	1,52E-01	1,54E-01	1,61E-01	O
rs132619	c.786A>C	p.Arg252Arg	4,37E-03	0	3,96E-03	5,93E-03	5,52E-03	N
rs132618	c.957T>A	p.Gly319Gly	4,72E-01	3,58E-01	4,68E-01	4,85E-01	4,87E-01	N
rs143477399	c.977T>C	p.Ile326Thr	4,17E-04	0	3,76E-04	1,16E-04	1,05E-04	O
rs144755172	c.812C>T	p.Ala342Val	0	0	0	1,16E-04	1,20E-04	O
rs145559179	c.1130G>A	p.Arg377Gln	1,47E-03	0	1,13E-03	2,56E-03	2,32E-03	O
rs150178861	c.*26G>C	p.?	3,76E-04	0	3,76E-04	0	1,54E-05	N

Tableau 8: Ensemble des variants du locus *APOL3* selon EXAC. Les variants identifiés par séquençage Sanger dans la cohorte sont en couleurs: variants fréquents (vert) et rares (jaune). Les variants non codants sont en gris.

TEST	Fréquence des variants testés	Echantillons	p-value	Nb total de marqueurs	Nb de marqueurs testés	Nb de marqueurs rares	Nb de marqueurs communs
SKAT	RARE (< 0.05)	Tous	0,048	6	6	6	NA
		Familiaux	0,047	5	5	5	NA
		Sporadiques	0,047	5	5	5	NA
SKAT-O	RARE (< 0.05)	Tous	0,027	6	6	6	NA
		Familiaux	0,017	5	5	5	NA
		Sporadiques	0,017	5	5	5	NA
SKAT-CR	ALL	Tous	0,038	9	9	6	3
		Familiaux	0,037	8	8	5	3
		Sporadiques	0,037	8	8	5	3

Tableau 9: Résultats des tests SKAT, SKAT-CR et SKAT-O sur les variants rares (et fréquents pour SKAT-CR) d'APOL3. Après correction pour tests multiples, l'association est significative.

identifiant dbSNP	c.	p.	Echantillons	Fréquence allélique cas	Fréquence allélique contrôles	χ^2	p-value	odds-ratio	Erreur standard	Intervalle de confiance
rs61741884	c.694G>A	p.VAI232Met	Tous	0,0545	0,06967	3,872	0,049	0,77	0,13	0,59-1
			sporadiques	0,05201	0,06967	5,204	0,023	0,73	0,14	0,56-0,95
			familiaux	0,08015	0,06967	0,365	0,546	1,16	0,25	0,71-1,90
rs142057520	c.682G>A	p.Ala228Thr	Tous	0,001033	0,005087	6,583	0,010	0,20	0,69	0,05-0,78
			sporadiques	0,001135	0,005087	5,698	0,017	0,22	0,69	0,05-0,86
			familiaux	?	?	?	?	?	?	?

Tableau 10: Effet protecteur des deux variants d'APOL3 associés à l'AAA. On note que l'odds-ratio de p.Ala228Thr est de 0,2 tous patients confondus ce qui confère un effet relativement élevé.

identifiant dbSNP	c.	p.	Echantillons	Modèle	Odds-ratio	Erreur Standard	Intervalle de confiance	p-value
rs61741884	c.694G>A	p.VAI232Met	Tous	Additif	0,78	0,13	0,59 - 1	0,053
			Familiaux		1,16	0,25	0,71 - 1,89	0,549
			Sporadiques		0,74	0,13	0,56 - 0,96	0,026
			Tous	Dominant	0,77	0,14	0,58 - 1,01	0,064
			sporadiques		1,26	0,26	0,74 - 2,1	0,389
			familiaux		0,73	0,14	0,54 - 0,96	0,027
Tous	Récessif	0,56	0,61	0,16 - 1,8	0,337			
rs142057520	c.682G>A	p.Ala228Thr	Tous	Additif	0,20	0,69	0,051 - 0,78	0,021
			sporadiques		0,22	0,69	0,057 - 0,85	0,029
			Tous	Dominant	0,20	0,69	0,051 - 0,78	0,021
			sporadiques		0,22	0,69	0,057 - 0,85	0,029

Tableau 11: Effet des deux variants rares d'APOL3 associés à l'AAA selon le modèle utilisé.

L'interprétation de ces premiers résultats est donc délicate. D'une part la proximité de ces deux variants, et l'association négative avec l'AAA suggère que seul le variant fréquent pourrait être un facteur protecteur de l'AAA, les deux variants étant probablement en déséquilibre de liaison. L'analyse des blocs de liaison et du coefficient de corrélation (r^2) n'est pas contributive dans ce cas car la différence de fréquence des variants est trop importante pour pouvoir conclure. D'autre part la comparaison des données issues de séquençage Sanger d'un côté et de génotypage de l'autre pose des problèmes de biais méthodologiques concernant le risque d'erreurs de génotypes. Une comparaison dans des conditions plus robustes est en cours de mise en place puisque le génotypage des 1 800 patients par Exome Chip a été réalisé par nos collaborateurs. Enfin, SKAT-O comme les autres tests d'association ne peut pas tester les variants de fréquence nulle, ce qui explique que seuls 9 des 15 variants identifiés dans la cohorte aient pu être exploités.

5.4.4 Tests d'association des variants très rares

Dans le cadre de cette dernière collaboration, le génotypage des 7 variants très rares identifiés par séquençage Sanger, incluant celui identifié dans la famille SE va être réalisé très prochainement. La population contrôle est une cohorte de 3 500 caucasiens issus du programme national de dépistage des AAA au Royaume-Uni ; ce sont donc essentiellement des hommes âgés de plus de 65 ans et indemnes d'AAA. Ils sont appariés pour l'âge et le sexe aux patients de la cohorte anglaise. L'association de ces variants très rares par la même approche que précédemment, c'est-à-dire par test combiné SKAT-O, sera ensuite réalisée.

5.4.5 Etudes fonctionnelles

Bien que les données d'expression et de génétique épidémiologiques indiquent qu'*APOL3* est un candidat très prometteur, il est maintenant nécessaire d'affirmer l'effet délétère du variant identifié dans la famille SE, voire des variants sélectionnées par étude d'association s'ils sont confirmés par réplique indépendante.

Dans le cadre de la collaboration avec l'équipe d'E. Pays, l'expression d'*APOL3* à l'état basal et après stimulation par le poly I:C a été testée dans les cellules dendritiques et les cellules endothéliales aortiques, et confirme l'induction d'*APOL3* après stimulus viral (figure 28, voir section 5.3.2.2.1 Expression d'*APOL3*). La génération du modèle mutant p.Arg377Trp dans ce dernier type cellulaire par méthode CRISPR est en cours dans notre laboratoire, et l'effet de ce variant sur la survie des cellules avec et sans induction sera testé très prochainement. Cette étude de survie sera complétée par un marquage de type TUNEL recherchant un phénomène d'apoptose sous-jacent.

Divers traitements pourront être appliqués au mutant cellulaire p.Arg377W, comme le blocage du trafic endosomal par la choroquine. Ce test a permis à Lan et coll. (2013) de lever l'effet cytotoxique des mutants G1/G2 d'*APOL1* dans les podocytes infectés par le VIH.

Une étude en immunofluorescence d'*APOL3* et du lysosome utilisant la cathepsine L (comme réalisé avec *APOL1* par Lan et coll.) permettrait de plus de déterminer s'il colocalise avec les membranes lysosomales comme c'est le cas pour *APOL1*. Cette approche pourra aussi être utilisée à la recherche d'une colocalisation avec la mitochondrie par des marqueurs spécifiques comme les anticorps anti-HSP60.

Enfin, une étude de la survie des cellules dendritiques de patients de la famille SE, portant particulièrement sur la sous-population BDCA3+ sera effectuée. Réalisée après tri cellulaire (FACS), elle peut se révéler techniquement difficile compte-tenu de la très faible représentation de ces cellules au sein des cellules dendritiques sanguines humaines. Une baisse de leur survie sera recherchée à l'état de base et après induction par le poly I:C.

6 Premières conclusions

Ce travail met en lumière les difficultés de l'analyse génétique dans une pathologie comme l'AAA : un projet global d'étude génétique de maladie complexe est relativement long, sa mise en œuvre n'est pas simple, notamment car elle requiert de multiples compétences, et se révèle relativement coûteuse. En effet, dans le cas le plus fréquent la présence de deux modèles distinct (multifactoriel et dominant) implique la nécessité de combiner de multiples approches de génétique Mendélienne (analyse de liaison par exemple) et épidémiologique (tests d'association). Ces approches combinées nécessitent le recrutement de nombreux cas sporadiques et familiaux, et impliquent nécessairement de multiples collaborations pour pouvoir être concluantes. De plus, elles requièrent des outils de séquençage et de génotypage puissants (notamment séquençage d'exome et exome-chip) qui ne sont disponibles que depuis récemment et impliquent encore un coût conséquent si l'on veut obtenir la puissance statistique suffisante (c'est-à-dire un nombre de patients suffisant). Les tests statistiques récemment développés comme le test SKAT-O semblent pouvoir tenir leurs promesses comme le démontrent les résultats de cette analyse. Cependant le manque de recul, notamment en termes de répllication, ne permet pas de considérer ces résultats comme totalement fiables à mon sens.

Le séquençage d'exome dans l'étude menée sur l'AAA s'est révélé très puissante puisque seuls trois variants candidats ont été obtenus après croisement des données de la fratrie étudiée. Cette approche est efficace si un certain nombre de conditions sont réunies, notamment l'obtention d'un phénotype ou endophénotype

fiable, et l'origine caucasienne de la famille étudiée. Nous verrons dans le chapitre suivant que l'analyse dans une famille d'origine ethnique différente, avec pourtant le même nombre d'individus atteints séquencés, est beaucoup plus délicate. Ces différences sont inhérentes à la variabilité naturelle plus importante dans d'autres populations, et l'inadéquation des bases de données de type EXAC et ESP dans certains groupes ethniques.

Enfin, quelque soit le niveau de sophistication des études combinées développées dans ce chapitre, on ne peut s'affranchir de tests fonctionnels : ils sont la confirmation ultime de l'implication du gène identifié dans le processus pathologique.

Chapitre 4

Etude génétique de la Dysplasie Fibromusculaire

1. Définition

La dysplasie fibromusculaire (DFM) est une maladie vasculaire non inflammatoire nonathéroscléreuse qui touche la paroi des artères de moyen calibre. Localement, elle se manifeste par une sténose artérielle qui peut se compliquer d'occlusion, d'anévrisme, ou de dissection (Olin et Sealove, 2011; Persu et al., 2012; Plouin et al., 2007). La dysplasie fibromusculaire peut toucher tout type d'artère de moyen calibre de la circulation cérébrale et systémique (figure 1), mais elle affecte dans 70 à 90% des cas les artères rénales, carotides et vertébrales (Kincaid et Davis, 1961; McCormack et al., 1966; Olin et al., 2012; Plouin et al., 2007). Dans sa forme la plus classique, elle est définie par son aspect angiographique en « collier de perles » ou en « pile d'assiettes » (figure 2).

2. Histoire de la DFM

La première description de la DFM est attribuée à Leadbetter et Burkland (Burkland et al., 1950) chez un garçon de 5½ ans avec hypertension artérielle sévère et sténose majeure unilatérale d'une artère rénale, traitée par néphrectomie. Le terme « hyperplasie fibromusculaire » a été introduit en 1958 par McCormack et coll. à partir de 3 patients atteints d'hypertension artérielle avec sténose de l'artère rénale (McCormack et al., 1958). La dysplasie fibromusculaire des artères cervicales a été identifiée par angiographie en 1964 par Palubinskas et Ripley puis confirmée histologiquement en 1965 chez une femme de 34 ans ayant développé un AVC sur thrombose de l'artère carotide (Palubinskas et Ripley, 1964, Connett et Lansche, 1965). En 1971, Harrison et McCormack ont proposé une classification histopathologique détaillée et corrélée aux données angiographiques de la DFM de l'artère rénale en 3 types distincts, selon la tunique artérielle touchée par le processus dysplasique (Harrison et McCormack, 1971, voir section 4. Histologie et classification histo-angiographique). En 1975, Stanley et coll. ont proposé une nouvelle classification histologique des DFM (Stanley et al., 1975).

En 2011, une réunion de consensus européenne, impliquant notamment le Centre de Référence des Maladies Vasculaires Rares de l'HEGP (CRMVR) a proposé des recommandations diagnostiques et thérapeutiques (Persu et al., 2012) suivies en 2014 par un consensus de l'AHA (American Heart Association)

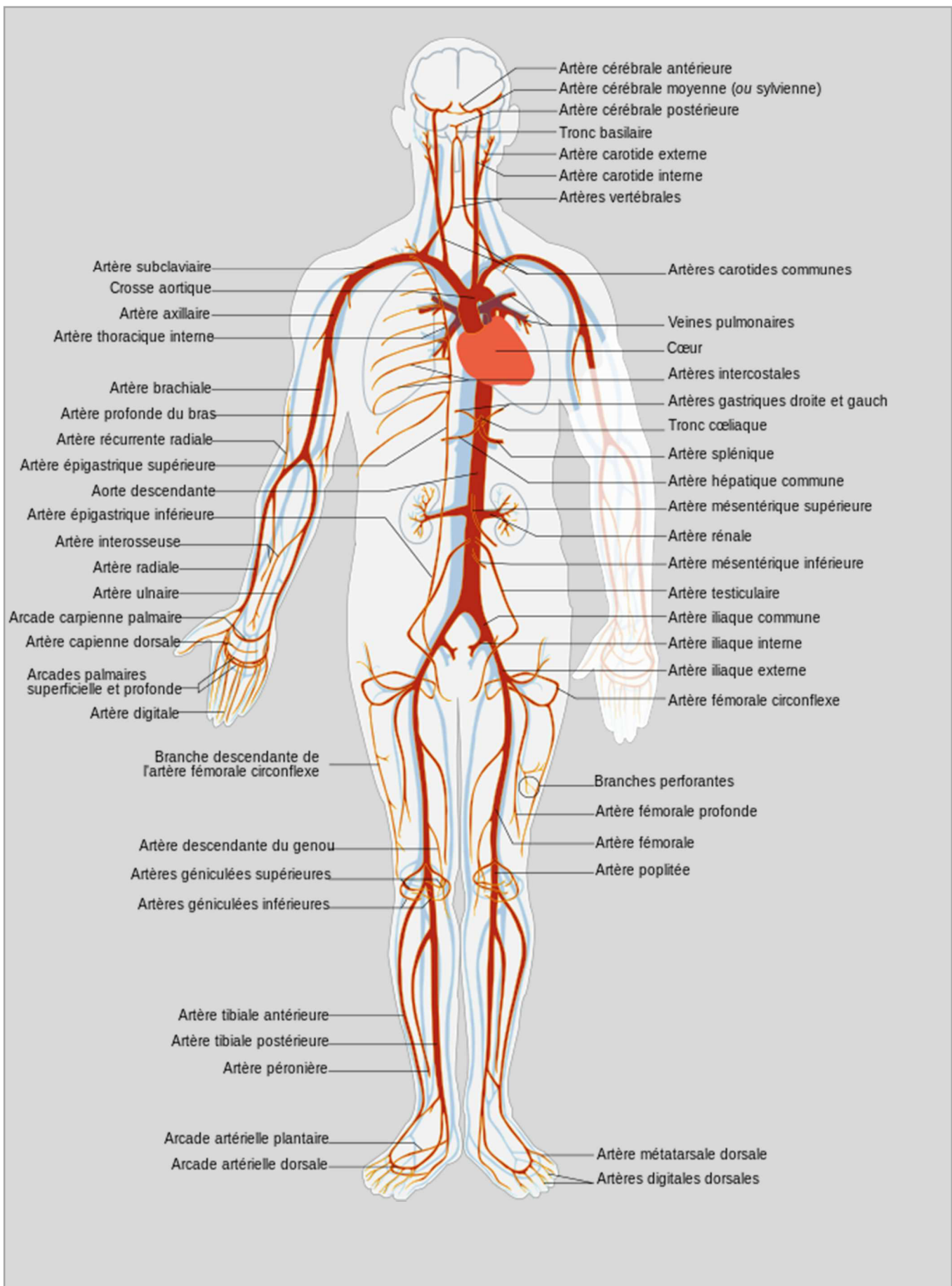


Figure 1: Artères de gros et moyen calibre de la circulation systémique et cérébrale. D'après Sobotta, Atlas d'anatomie humaine, vol1 et 2.

dirigé par J. Olin et coll. (Olin et Sealove, 2011) à partir de données issues de plus de 400 patients inscrits au registre américain de la dysplasie fibromusculaire.

3. Epidémiologie et facteurs de risque

La DFM est une maladie rare, sporadique dans environ 90% des cas (voir section 6. Composante héréditaire de la DFM). Elle est particulièrement fréquente chez les femmes, avec un sex-ratio moyen de 8:1 (Olin et al., 2012) tous types confondus.

Sa prévalence exacte n'est pas connue, et les estimations divergent, du fait du type de données utilisées à partir de patients symptomatiques (registres nationaux) ou non (données angiographiques de donneurs de rein ou données autopsiques). Il semble en effet que la DFM non symptomatique puisse être fréquente en comparaison des formes symptomatiques collectées dans les registres français et américain. La prévalence estimée à partir de données angiographiques de donneurs de rein est de l'ordre de 3,8 à 6,6% (données issues de deux séries d'environ 1800 et 500 patients respectivement (Cragg et al., 1989; Neymark et al., 2000). Notre équipe référente dirigée par le Pr Pierre-Francois Plouin au sein du CRMVR, a confirmé ces résultats par méta-analyse de 4 études angiographiques distincts impliquant 3 181 sujets asymptomatiques potentiels donneurs de rein. Cette étude relève que 139 sujets (4,4%) avaient une DFM selon les critères angiographiques définis par le consensus européen (Plouin et al., 2007). Au cours des 2 à 8 ans de suivi, environ un quart des individus non donneurs ont développé une hypertension artérielle (Andreoni et al., 2002; Cragg et al., 1989; Plouin et al., 2007). Ces données doivent être cependant prises avec précaution, les donneurs de rein potentiels de ces études étant des apparentés de patients atteints de diverses pathologies rénales chroniques dont la dimension familiale n'est pas connue.

Une grande étude autopsique réalisée à partir de 819 sujets a révélé une prévalence d'environ 1% de DFM des artères rénales (Heffelfinger et al, 1970) et pose la question de la sensibilité relative de l'examen autopsique en comparaison de l'angiographie. Les détails de l'étude autopsique sus-citée ne sont cependant pas connus et ne permettent pas d'évaluer la pertinence des critères utilisés.

La répartition des lésions dysplasiques dans l'arbre artériel est inhomogène : à partir de données issues de plus de 1 000 patients français et américains (Cragg et al., 1989; Mettinger et Ericson, 1982; Plouin et al., 2007) environ 60 à 75% des cas impliquaient l'artère rénale, 32% l'artère carotide ou vertébrale, et 10% d'autres lits artériels (artères iliaque ou intracrânienne). Il y a peu d'études traitant de la prévalence de la DFM des artères cervicales et intracrâniennes, mais ces dernières semblent moins souvent touchées (Olin et

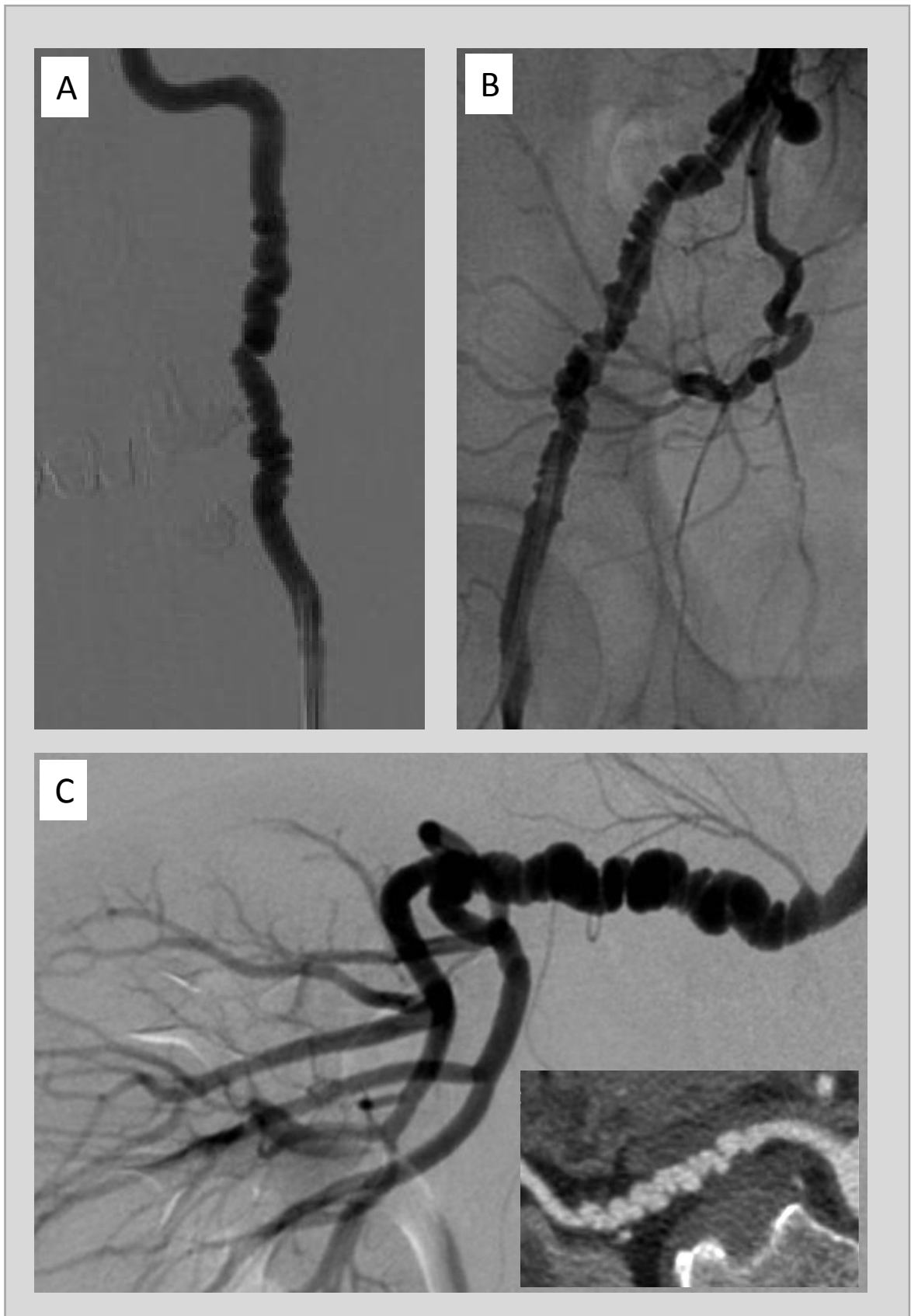


Figure 2: aspect angiographique de lésions en collier de perles d'une artère carotide interne (A), iliaque externe (B), et rénale (C). En médaillon l'aspect angioscannographique de la même lésion de l'artère rénale. D'après Olin et al, 2011, Savard et al, 2012.

al., 2012; Touzé et al., 2010). Une étude angiographique des atteintes carotidiennes et vertébrales fait état d'une prévalence de 0,3% à 3,2% (Touzé et al., 2010) mais ces chiffres sont probablement surestimés car les patients inclus dans cette série ne sont pas issus de la population générale mais porteurs de symptômes cervicaux ou cérébraux. La prévalence de la DFM cervicale sur étude autopsique est nettement inférieure aux chiffres sus-cités : parmi une série de 20 244 cas d'autopsie consécutives, seulement 4 avaient une DFM cervicale (Schievink et Björnsson, 1996).

La prévalence de la DFM dans les dissections spontanées des artères cervicales est relativement élevée, de l'ordre de 15% à 20% des cas (Debette et Leys, 2009; Schievink, 2001).

Les facteurs de risque majeurs de la DFM sont le sexe féminin et le tabagisme : dans les registres américain et français de la pathologie, 91 et 85% des patients respectivement sont de sexe féminin, le sex-ratio est donc environ de 8 à 9:1. Le tabagisme est un facteur de risque plus modéré, et particulièrement associé à la forme dite « unifocale » de la maladie (voir section 4. Histologie et classification histo-angiographique). Dans le registre français, environ 50% des patients porteurs de forme unifocale contre 25% dans les formes multifocales sont tabagiques (Savard et al., 2012). Toutes formes confondues, 30% des patients atteints de DFM sont fumeurs contre 18% parmi les témoins apparié pour l'âge, le sexe, la pression artérielle systolique, et le nombre d'antihypertenseurs pris par les patients ($p < 0,001$; odds-ratio 2,5; IC 1,06-3,09).

4. Histologie et classification histo-angiographique

La première classification consensuelle, basée sur la tunique artérielle touchée par la dysplasie date de 1971 (Harrison et McCormack, 1971), suivie en 1975 par la classification de Stanley. Cette dernière est simplifiée et définit 5 sous-types de fréquence variable, relativement corrélés avec l'aspect angiographique des lésions.

4.1 DFM intimale

La DFM ou fibroplasie intimale est le second sous-type le plus fréquent de DFM : elle est rapportée dans environ 5 à 10% des cas selon les séries (Harrison et McCormack, 1971; Stanley et al., 1975). L'intima est le siège d'une atteinte dysplasique caractérisée par une accumulation de fibres collagènes, des anomalies de la limitante élastique interne qui apparaît fragmentée et dupliquée, et la présence de cellules musculaires lisses peu abondantes et non orientées. La lumière vasculaire est amoindrie et la média est généralement normale, parfois amincie (figure 4A). Ce sous-type se présente à l'angiographie comme une sténose focale (constriction très localisée) ou tubulaire (figure 3, Olin 2014).

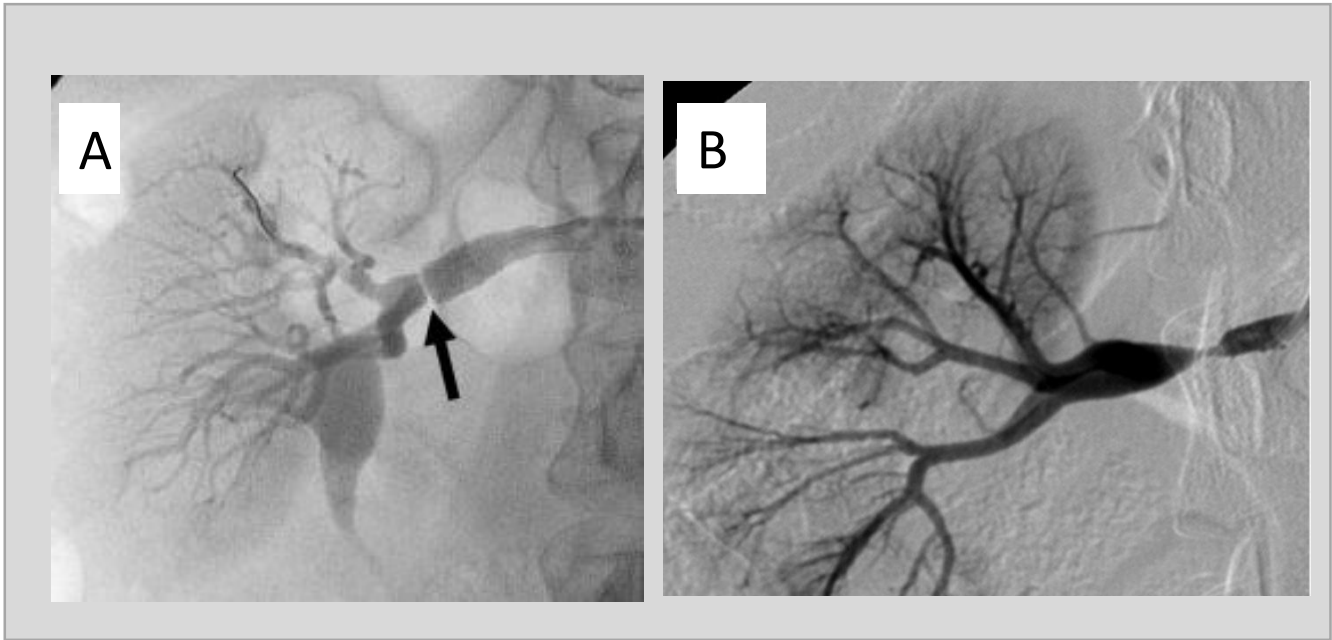


Figure 3: aspect angiographique des sténoses focales. A Sténose focale courte d'une artère rénale droite. B: sténose tubulaire d'une artère rénale droite.

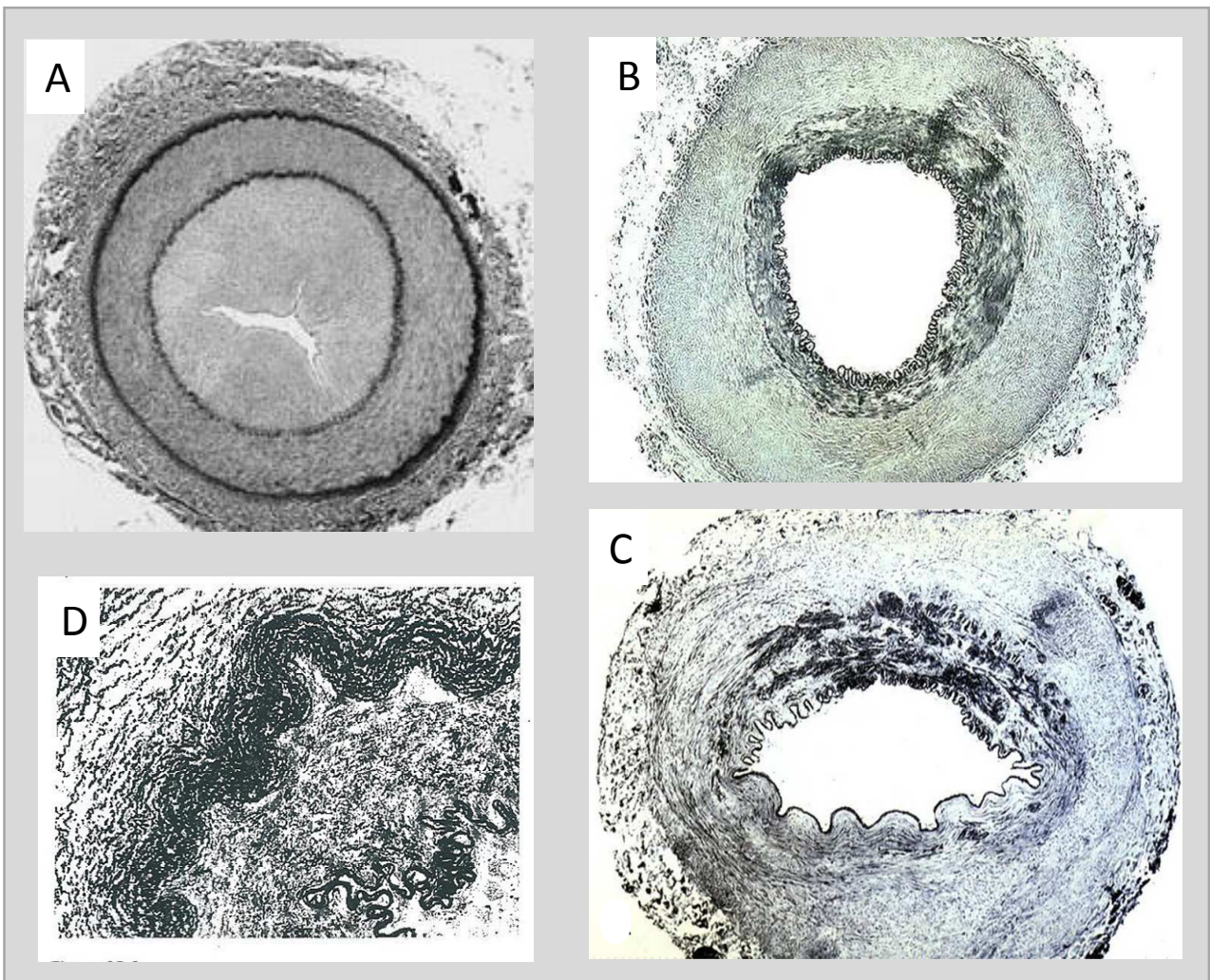


Figure 4: Formes rares de DFM.

A: Fibroplasie intimale. B: Fibroplasie périmédiale. C: Hyperplasie périmédiale. D: fibroplasie adventitielle. D'après McCormack et al, 1971, et Camilleri, Les maladies de la paroi artérielle, 1987, Ed Flammarion.

4.2 DFM médiale

La média est le siège de trois types lésionnels distincts sur le plan histologique : les fibroplasies médiale et péri-médiale, qui peuvent être associées sur un même segment artériel et corrélées à l'aspect angiographique en collier de perles, et l'hyperplasie médiale, d'aspect angiographique similaire à la fibroplasie intimale.

4.2.1 Fibroplasie médiale

Elle est la plus courante des dysplasies artérielles et représente entre 70 et 90% des cas selon les séries (Olin et al., 2012; Stanley et al., 1975). Macroscopiquement, l'aspect en collier de perles peut être visible (figure 5A, médaillon), et correspond à des sténoses multiples concentriques ou légèrement hélicoidales alternant avec des zones de dilatation post-sténotiques (la lumière artérielle au plus grand diamètre des « perles » étant de calibre supérieur à l'artère rénale saine). Cependant dans la majorité des cas la fibroplasie médiale est associée à un aspect macroscopique (et angiographique) moins typique de sténoses multiples visualisables sous la forme d'irrégularités pariétales non concentriques qui peuvent être associées à des zones anévrysmales. Histologiquement, la fibroplasie médiale est définie par quatre critères (figure 5) :

La dysplasie de la média, caractérisée par des anomalies des cellules musculaires lisses (d'où le nom de dysplasie) : on retrouve des zones de prolifération des CML, qui sont orientées en tous sens et non de façon concentrique ou hélicoidale. Des zones de perte en CML coexistent au sein d'un même fragment artériel. Ainsi l'épaisseur de la média apparaît irrégulière le long du segment atteint, alternant zones d'épaississement d'aspect bourgeonnant, et zones d'amincissement pouvant aller jusqu'à la disparition complète de la média dans les anévrysmes.

Une fibrose médiale dans les zones de perte en CML, caractérisée par des dépôts de collagène, de protéoglycanes et de glycosaminoglycanes, colorés en bleu par la coloration Trichrome vert.

Une disparition de la limitante élastique interne, tout particulièrement dans les zones hypoplasiques et notamment anévrysmales.

Une hyperplasie myointimale constituée de cellules musculaires lisses intimales. On la trouve en regard des zones d'amincissement de la média où la perte des CML est majeure, mais de façon inconstante. L'épaisseur totale de la paroi artérielle peut donc être conservée en certains endroits du fait de cette hyperplasie.

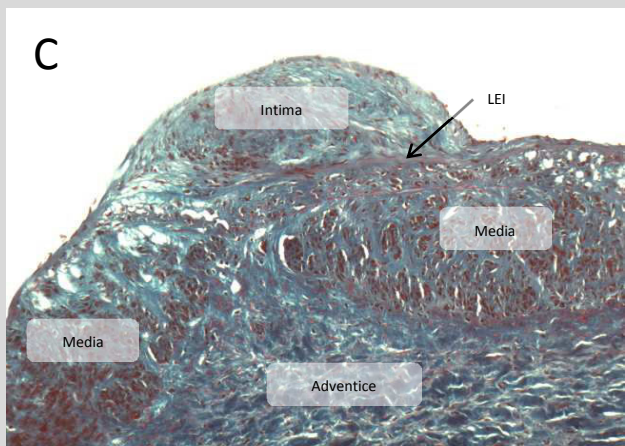
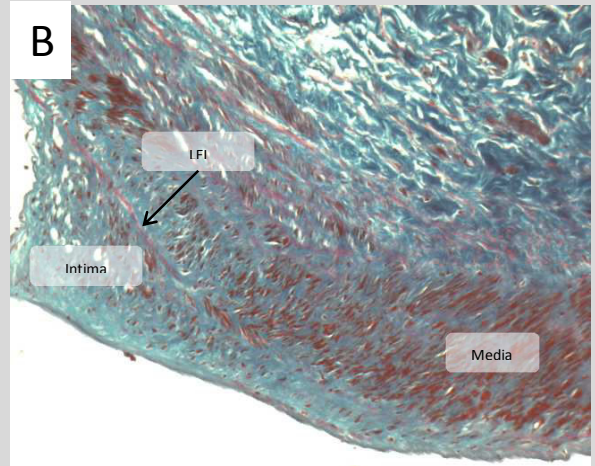
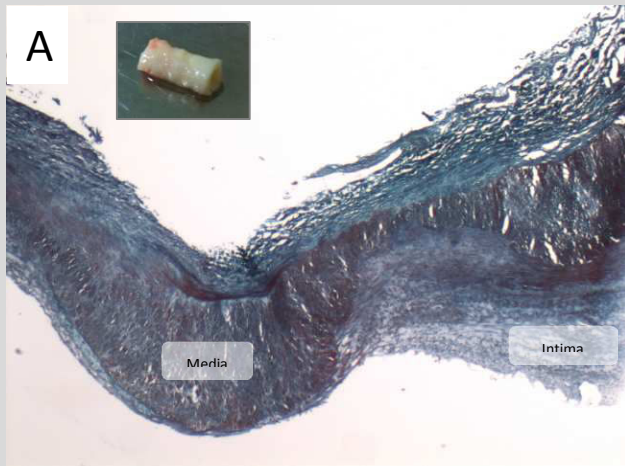


Figure 5: aspect histologique de différentes artères présentant une fibroplasie médiale. A: L'aspect général est typique d'une fibroplasie médiale, reproduisant les colliers de perles (zones d'hypertrophie mediale alternant avec des zones d'amincissement compensés par une hypertrophie myointimale). B: la limitante élastique interne (LEI) disparaît, et les CML de la media sont orientées en tous sens. C : la média s'amincit peu à peu, remplacée par une hyperplasie myointimale. Données anatomopathologiques de patients opérés à l'HEGP. Coloration trichrome vert.

4.2.2 Fibroplasie périmédiale

La fibroplasie périmédiale représente une atteinte rare voire très rare de DFM, actuellement estimée à 1% des cas (Olin et al., 2012). Elle se caractérise par une dysplasie localisée dans la moitié externe de la média, immédiatement en zone sous-adventitielle (figure 4B). L'aspect est proche de la fibroplasie médiale, mais les lésions respectent la zone interne de la média et la limitante élastique interne. L'analyse histologique poussée d'artères rénales montre que les sous-types médial et péri-médial coexistent généralement chez un même patient et au sein d'un même segment artériel, ce qui n'est pas le cas des autres sous-types. En certains endroits ces aspects forment un continuum lésionnel, suggérant qu'ils pourraient représenter des étapes différentes d'un même processus pathologique (Camilleri, 1987).

4.2.3 Hyperplasie médiale

L'hyperplasie médiale (figure 4C) est très rare (<1% des cas, Stanley et al., 1975) et est caractérisée par une prolifération hyperplasique cellulaire marquée et homogène au sein de la zone lésionnelle. Histologiquement, l'orientation de ces CML, l'aspect de la limitante élastique interne, de la matrice extracellulaire (fibres collagènes) et de l'intima sont conservés.

4.3 DFM adventitielle

La fibroplasie adventitielle (figure 4D) est très rare (<1%, Stanley et al., 1975) et caractérisée par un dépôt de collagène dans l'adventice et le tissu périartériel, associée à une infiltration focale de lymphocytes. Angiographiquement, elle apparaîtrait similaire à la DFM intimale et l'hyperplasie médiale, sous la forme d'une sténose focale (Weiner et al., 2009).

L'usage de la classification histologique est aujourd'hui assez limité dans la mesure où les sténoses dysplasiques nécessitant un geste sont aujourd'hui diagnostiquées par imagerie (angiographie ou angioscanner) et traitées par angioplastie transluminale. Dans les registres français et américain, moins de 5% des patients ont une analyse histologique, secondaire à un geste chirurgical à ciel ouvert (Olin et al., 2012; Savard et al., 2012).

Angiographiquement, la classification proposée par le consensus européen distingue les dysplasies « unifocale », « multifocale », et « tubulaire » (lorsque la sténose mesure plus d'1 cm, figure 3B). Les dysplasies focales, quelque soit leur longueur (sténoses focales ou tubulaires longues) sont définies par une sténose isolée sur un seul lit artériel. La DFM multifocale est définie par des sténoses multiples sur un ou plusieurs lits artériels.

Comme vu précédemment, les sténoses tubulaires correspondent à une DFM intimale, parfois adventitielle. La DFM unifocale avec sténose courte peut être la conséquence d'une dysplasie intimale, adventitielle ou d'une hyperplasie médiale. Enfin, la DFM multifocale correspond à la fibroplasie médiale et est donc de loin la plus couramment rencontrée. L'aspect en collier de perles, bien que relativement fréquent (75% des cas de la cohorte française, Savard et al., 2012) est inconstamment retrouvé dans ce sous-type.

Sur le plan clinique et épidémiologique, une étude menée par notre équipe a démontré la pertinence de la classification angiographique en 2 sous-types (DFM focale et multifocale) dans la discrimination de formes cliniques distinctes de patients. En effet les patients porteurs d'une atteinte focale (20% de cette cohorte de plus de 300 cas) sont plus jeunes au moment du diagnostic (âge médian 30 ans), ont un sex-ratio beaucoup moins favorable aux femmes (2 :1), une HTA plus sévère et présentent plus fréquemment une asymétrie rénale et une intoxication tabagique, en comparaison des cas atteints de DFM multifocale. Cette forme clinico-angiographique est donc plus sévère. Elle représente probablement une entité distincte de la forme multifocale bien qu'un chevauchement phénotypique ne puisse pas être exclu en l'absence de marqueur plus spécifique de chaque sous-type.

La forme multifocale est donc la plus couramment rencontrée et les patients présentent des caractéristiques cliniques et épidémiologiques propres : les femmes sont touchées 8 fois plus que les hommes, l'hypertension est plus modérée (+20 mm Hg systolique et diastolique au-delà des normes supérieures). L'âge médian au diagnostic est de 50 ans, révélant une maladie peu sévère ou silencieuse dans les premières décennies de vie. Savart et coll. ont par ailleurs démontré que l'aspect typique en collier de perles chez les patients atteints de DFM multifocale ne représentait pas un sous-type particulier de patients selon les critères développés ci-dessus. Ainsi, la DFM multifocale avec ou sans aspect de collier de perles est une entité peu sévère en comparaison de la forme focale, touchant très majoritairement les femmes de 30 à 50 ans.

5. Aspects cliniques et histoire naturelle de la DFM

Les manifestations cliniques de la DFM sont très variables et dépendent du lit artériel atteint et du degré de sténose, ainsi que de l'existence de lésions vasculaires associées (dissection artérielle, anévrisme). Théoriquement tous les lits artériels peuvent être touchés s'il s'agit d'artères musculaires. L'aorte, la naissance du tronc coeliaque et le TABC sont donc épargnés. Cependant l'atteinte des artères rénales et cervicales représente plus de 90% des formes, suivies par les lésions coronaires, puis tout autre type d'artère à destinée viscérale ou de membre. Les artères des membres supérieurs sont très rarement touchées (Olin et al., 2014).

La progression des lésions sténosantes dans la DFM n'est pas certaine. Elle fait l'objet d'une évaluation prospective dans le cadre du protocole français PROFILE étudiant par imagerie les lésions artérielles rénales et cervicales de patients sur une période de 3 ans.

5.1 DFM Rénale

La manifestation la plus commune de la DFM de l'artère rénale est l'hypertension, de gravité et d'âge d'apparition variables. Un âge de révélation de l'HTA précoce (avant 35 ans) ou une pharmacorésistance doit faire suspecter une sténose dysplasique de l'artère rénale. Cependant, l'âge médian de découverte de l'HTA dans les registres américain et français est plus tardif, notamment dans la DFM multifocale (entre 40 et 45 ans, Olin et al., 2012; Savard et al., 2012). Cet âge tardif et chevauchant avec l'âge médian de découverte de l'HTA essentielle rend le diagnostic étiologique parfois long, avec un délai diagnostique évalué de 4 à 9 ans (Olin et al., 2012; Savard et al., 2012).

La dissection de l'artère rénale est une complication rare (3 à 7%, Olin et al., 2012; Savard et al., 2012) se manifestant par une douleur lombaire, parfois une insuffisance rénale aiguë en cas d'occlusion complète. Une insuffisance rénale constituée est présente dans moins de 2% des cas (Olin et al., 2012; Savard et al., 2012).

Les anévrysmes des artères rénales sont plus fréquents que les dissections, retrouvés dans un peu plus de 10% des cas dans la cohorte française (Savard et al., 2012) et jusqu'à 30% dans la cohorte américaine (Olin et al., 2012).

5.2 DFM des artères cervicales, encéphaliques et anévrysmes intracrâniens

Comme vu dans la section 3. Epidémiologie et facteurs de risque, la localisation cervicale est plus rare que l'atteinte des artères rénales, mais les évaluations divergent sur le sujet : entre 9% (Stanley et al., 1975) et 28% (Lüscher et al., 1986) des patients atteints de DFM rénale sont porteurs de lésions dysplasiques carotidiennes ou vertébrales. Les manifestations cliniques de la DFM des artères encéphaliques sont très variables et souvent non spécifiques. La moitié des patients présente des céphalées très fréquentes voire quotidiennes, ou de véritables migraines dans 50% des cas (Olin et al., 2012). Les acouphènes sont rapportés par plus d'un quart (27,5%) des patients du registre américain. Des malaises avec sensation de vertige, de plénitude dans la tête ou les oreilles, sont fréquemment rapportés (Olin et al., 2012; Touzé et al., 2010). Les complications les plus sévères sont l'AIT (5-50%) ou l'AVC (8-33%) par mécanisme thromboembolique ou par bas débit cérébral, l'hémorragie méningée (3-40%) par rupture d'anévrysme intracrânien, et la dissection artérielle (Touzé et al., 2010). La présence d'anévrysmes intracrâniens (AIC) dans la DFM cervicale n'est pas

fortuite, la rupture d'anévrisme étant la première cause d'hémorragie méningée dans la pathologie. Dans le registre américain, environ 7% des patients sont porteurs d'un AIC (Olin et al., 2012), chiffres concordants avec l'unique étude de cohorte précédemment publiée (Cloft et al., 1998).

L'association entre DFM et dissection des artères cervicales est connue depuis de nombreuses années (Debette et Leys, 2009; Mokri et al., 1986; Schievink, 2001). Une DFM cervicale est retrouvée chez 15-20% des patients avec dissection carotidienne ou vertébrale spontanée. Inversement, environ 12% de patients dysplasiques ont présenté une dissection d'artère cervicale, responsables de douleurs cervicoencéphaliques, d'un syndrome de Claude-Bernard-Horner, ou d'un AVC. La récurrence des dissections cervicales n'est pas rare, le taux cumulé de dissections récurrentes étant de 3,5%, 5,0%, et de 12% à 2, 5 et 10 ans, respectivement (Schievink, 2001).

5.3 DFM des artères coronaires

Les manifestations coronaires de la DFM ont été revues récemment par Michelis et coll. et sont la conséquence d'une dissection d'un tronc artériel coronaire ou plus rarement d'une branche de division. L'aspect en collier de perles n'est généralement pas retrouvé et le diagnostic est souvent porté après la dissection (Michelis et al., 2014). Une sténose focale peut être identifiée par coronarographie ou coroscanner.

Cliniquement, la DFM coronaire se présente comme un syndrome coronarien aigu, un infarctus constitué, un décès subit, ou une dysfonction ventriculaire gauche aiguë chez un patient dont le diagnostic de DFM sur un autre lit artériel a nécessairement été porté (Pate et al., 2005; Saw et al., 2014). La DFM coronaire semble relativement rare bien que mal estimée aujourd'hui : moins de 5% des patients du registre américain présentent une maladie coronaire, DFM et athérome confondus.

6. Composante héréditaire de la DFM

L'existence d'un facteur héréditaire a été démontrée dès 1965 par plusieurs auteurs rapportant des cas de DFM dans des fratries ou chez des cousins germains (Halpern et al., 1965; Hansen et al., 1965; Major et al., 1977; Ouchi et al., 1989). Dans les années 1980, deux équipes (Gladstien et al., 1980; Mettinger et Ericson, 1982; Rushton, 1980) ont analysé cette composante héréditaire de façon plus systématisée : ces études réalisées respectivement avec 20 familles d'au moins deux cas et 37 cas de DFM cérébrovasculaire ont conclu à une transmission autosomique dominante à pénétrance incomplète, variable selon les familles. Cependant ces études ont été réalisées à partir d'évaluations cliniques des apparentés et non par approche

angiographique, ce qui rend leur interprétation incertaine. Plusieurs études subséquentes réalisées par angiographie et menées par notre équipe, ont démontré qu'environ 10% des cas de DFM rénale présentaient une récurrence familiale. Ces études montrent que cette récurrence est compatible avec une transmission autosomique dominante (Grimbert et al., 1998; Pannier-Moreau et al., 1997). La proportion de cas familiaux est du même ordre chez les patients du registre américain (7% d'antécédants familiaux de DFM). Enfin, l'analyse par échotracking carotidien à haute résolution a été réalisée dans ces familles françaises et semble mettre en évidence un endophénotype intéressant : un aspect échographique en « triple signal » et l'établissement d'un score échographique semi-quantitatif permettrait d'identifier les individus porteurs du trait génétique en l'absence de manifestations cliniques et de données angiographiques selon un modèle autosomique dominant (Perdu et al., 2007). Ce score a été utilisé pour identifier des apparentés potentiellement porteurs (figure 6). Aucune estimation de la sensibilité et de la spécificité de ce score n'a cependant été proposée, notamment au vu du faible nombre de cas atteints disponibles. La spécificité est limitée par l'âge de l'individu et l'existence d'un athérome sous-jacent notamment, et la sensibilité semble imparfaite au vu du score obtenu chez certains cas atteints (par exemple dans la famille FMD125, figure 6). Il est donc aujourd'hui difficile de se passer d'un examen angiographique ou angioscannographique pour déterminer le statut des apparentés et affirmer une forme familiale de la pathologie.

7. Génétique de la DFM

La DFM est aujourd'hui considérée comme une maladie multifactorielle avec forte composante héréditaire (comme vu dans la section précédente, environ 10% des patients présentant un antécédant familial de DFM), la composante environnementale connue étant liée au sexe et à l'intoxication tabagique. Dans les formes familiales, la pénétrance est variable et dépendante du sexe dans notre expérience : L'identification de formes familiales au sein du CRMVR démontre qu'au sein des 10 familles que nous suivons (comportant 2 à 5 atteints chacune) parmi les 27 cas cumulés, 21 sont des femmes, soit une proportion d'environ 75%. En première estimation, la pénétrance de la DFM semble donc deux fois plus importante chez les femmes que les hommes (voir section 10. Etude génétique de la DFM). Ces formes familiales sont compatibles avec une transmission autosomique dominante (figure 6), impliquant que de forts facteurs modificateurs liés au sexe soient à l'œuvre.

Aucune étude génétique de liaison ou d'association génome entier n'a été publiée jusqu'à très récemment. Un diagnostic histologique de DFM ayant été porté dans quelques cas de déficit en α 1-antitrypsine (Randers et al., 1998; Schievink et al., 1998a, 1998b; Sölder et al., 1997), notre équipe a testé l'association de 3

polymorphismes au locus AAT avec la DFM, dans une étude comportant 166 cas avec angiographie, et 3 groupes contrôles (incluant respectivement 353 sujets hypertendus, 288 et 444 sujets normotendus). Aucune association n'a pu être détectée avec cette approche (Perdu et al., 2006).

D'autres recherches ciblées sur des candidats comme le gène de l'actine spécifique du muscle lisse (*ACTA2*), le gène de l'élastine (*ELN*), et les gènes du syndrome d'Ehlers-Danlos (*COL3A1*), de la maladie de Marfan (*FBN1*) et des syndromes de Loeys-Dietz (*TGFBR1,2*, et *SMAD3*) ont démontré que dans une faible proportion de cas, la DFM peut être confondue avec l'un de ces syndromes : Poloskey et coll. ont en effet rapporté 2 cas de syndrome de Loeys-Dietz (mutations *TGFBR1*) avec dilatation de l'aorte ascendante et tortuosité artérielle inclus dans le registre américain de la DFM. Comme nous le verrons dans la section 10.4.1, la distinction diagnostique entre DFM et certaines présentations atypiques de ces syndromes peut être subtile. L'analyse des gènes correspondants au titre de l'exclusion se justifie particulièrement en cas d'anévrisme(s) artériel(s) associé(s), notamment aortique (Poloskey et al., 2012). A ce titre, il est surprenant de constater qu'environ 3% des patients inscrits au registre américain sont porteurs d'une dilatation de l'aorte ascendante, ce qui pose la question de la certitude diagnostique selon les critères utilisés pour la constitution de cette cohorte.

Notre équipe a plus récemment publié une étude de séquençage d'exome à partir de 16 cas familiaux. Elle comporte également une analyse d'association ciblant les variants rares identifiés chez les sujets familiaux en cohorte de cas sporadiques. Les résultats des analyses familiales démontrent la grande hétérogénéité génétique de la pathologie : parmi les 7 familles étudiées, aucun gène commun n'a pu être identifié par les approches classiques de sélection et de filtre de variants (Kiando et al., 2015). L'analyse d'association a été réalisée par exome-chip, utilisant la puce Illumina HumanExome-12v1 en comparant les variants rares de 309 cas non apparentés et 699 contrôles issus de l'étude SUVIMAX (voir chapitre 3 section 5.4.3). L'analyse par SKAT-O a permis d'identifier une association significative entre DFM multifocale et les gènes *MYLK* ($p=0,01$), impliqué dans le TAA (voir section 9.2.2), *OBSCN* codant l'obscurine, protéine sarcomérique ($p=0,003$), *DYNC2H1* codant la chaîne lourde cytoplasmique de la dynéine ($p=0,02$) et *RNF213* précédemment associée à la maladie de Moya-Moya ($p= 0,01$). Ces résultats d'association ne sont pas robustes et l'étude de cohortes plus importantes est nécessaire afin de confirmer leur implication dans la DFM.

8. Hypothèses physiopathologiques

Il n'existe pas d'étude physiopathologique importante publiée dans cette affection, et aucune hypothèse explicative robuste n'a été proposée. Les bases fondamentales du processus sténotique, et de l'association avec des anévrysmes à distance comme les AIC ne sont donc pas connus. J'ai souhaité dans un premier temps restreindre l'analyse génétique de la DFM en me basant sur des hypothèses physiopathologiques claires. Une revue bibliographique portant d'une part sur les théories explicatives proposées dans la littérature et d'autre part sur les maladies héréditaires anévrysmales, disséquantes, et sténosantes et leur physiopathologie, s'avérait nécessaire.

Je n'ai pas traité des diagnostics différentiels dans le cadre des formes sporadiques tels l'athérome, les vascularites inflammatoires ou la médiolyse segmentaire et focale. Ces maladies ne semblaient pas pertinentes pour dégager un cadre physiopathologique de la DFM, soit par définition (la DFM est non athéromateuse et non inflammatoire) soit par absence de connaissances sur les bases de ces maladies (cas de la médiolyse segmentaire, Slavin, 2009).

8.1 L'hypothèse hormonale

Des facteurs hormonaux (imprégnation oestrogénique) ont été proposés sur la base de la distorsion majeure du sex-ratio, sans élément de preuve supplémentaire. Une étude réalisée par Sang et coll. portant sur les facteurs de risque de la DFM a exclu la gestité ou l'utilisation de contraceptifs oraux comme facteurs associés (Sang et al., 1989). Une étude française récente aurait mis en évidence l'existence de récepteurs à la progestérone dans des artères rénales dysplasiques en comparaison d'artères saines. Cette étude porte sur 6 cas et 3 contrôles (Silhol et al., 2015). Nous avons réalisé un marquage immunohistochimique de ces récepteurs et n'avons pas confirmé ces résultats sur une seconde série de 4 cas distincts (données non publiées).

8.2 L'hypothèse ischémique

Elle a été proposée par Stanley en 1975 devant la constatation que les segments artériels rénaux et cervicaux les plus touchés étaient les moins riches en *vasa vasorum*. D'autre part, ces auteurs constatent que la DFM médiale dans ces segments affecte préférentiellement la portion externe de cette tunique. Cette hypothèse n'a pas fait l'objet d'études plus approfondies.

8.3 L'hypothèse mécanique

Le rein droit étant plus mobile que le gauche, cette hypothèse a été proposée en 1987 par Luscher et coll. devant la constatation que la DFM des artères rénales touchait plus fréquemment l'artère rénale droite que la gauche (Lüscher et al., 1986). Ce caractère a été confirmé dans la cohorte française où le ratio droite/gauche est de 5:1 et 3:1 pour la DFM multifocale et focale respectivement (Savard et al., 2012). Cependant une étude cas-contrôles réalisée deux ans plus tard a réfuté l'association entre mobilité rénale et DFM (Sang et al., 1989). De façon similaire, le segment le plus mobile des artères cervicales est la portion distale de la carotide interne, site le plus souvent touché par la DFM cervicale.

D'autres éléments viennent alimenter cette hypothèse : certains syndromes responsables de sténoses vasculaires impliquant un facteur mécanique présentent des similitudes histologiques avec la fibroplasie intimale ou médiale. C'est notamment le cas des **hémoglobinopathies** (drépanocytose, sphérocytose et thalassémies, figure 7) où la présence de sténoses carotidiennes est corrélée à la sévérité de l'anémie (Merkel et al., 1978; Tsomi et al., 2001). On retrouve généralement une hyperplasie myointimale, associée à une dysplasie médiale, et une fragmentation de la limitante élastique interne. Ces deux derniers aspects sont cependant inconstants, les lésions vasculaires dans ces affections ne sont donc pas exactement similaires celles de la fibroplasie médiale. La physiopathologie de l'atteinte vasculaire dans ces maladies de l'hémoglobine est mal connue, mais les hypothèses actuelles (figure 8) privilégient l'activation endothéliale *via* les interactions entre érythrocytes et endothélium, impliquant une sécrétion d'endothéline, et un piégeage du NO. Ces deux phénomènes sont responsables d'une vasoconstriction, et secondairement une prolifération myointimale par un mécanisme encore incertain (voir section 9.1.3).

Deux autres pathologies à composante mécanique présentent des similitudes avec la DFM. La thrombose veineuse de la veine iliofémorale gauche par compression de l'artère iliaque droite est appelée **syndrome de May-Thurner** (Ehrich et Krumbhaar, 1943; May et Thurner, 1957). Ce syndrome est caractérisé par trois phénomènes combinés appelés triade de Wirschov : une anomalie locale du flux veineux causée par la compression artérielle, une coagulopathie sous-jacente, et une lésion intrapariétale veineuse (hyperplasie intimale et dysplasie médiale). L'hypothèse physiopathologique, non démontrée, est que cette lésion pariétale serait induite par les vibrations secondaires aux pulsations artérielles iliaques.

Enfin, le **Hand-Arm Vibration-induced Syndrome (HAVS)** est une maladie professionnelle touchant les travailleurs manuels dont les mains sont exposées à des vibrations (marteau-piqueur, tronçonneuse). Cette maladie est responsable d'un phénomène de Raynaud acquis, une atrophie des phalanges, et de troubles de la sensibilité. La réversibilité et la sévérité de ces symptômes sont corrélées à la durée d'exposition aux vibrations. Sur le plan physiopathologique et histologique, sa composante est double, neurologique et

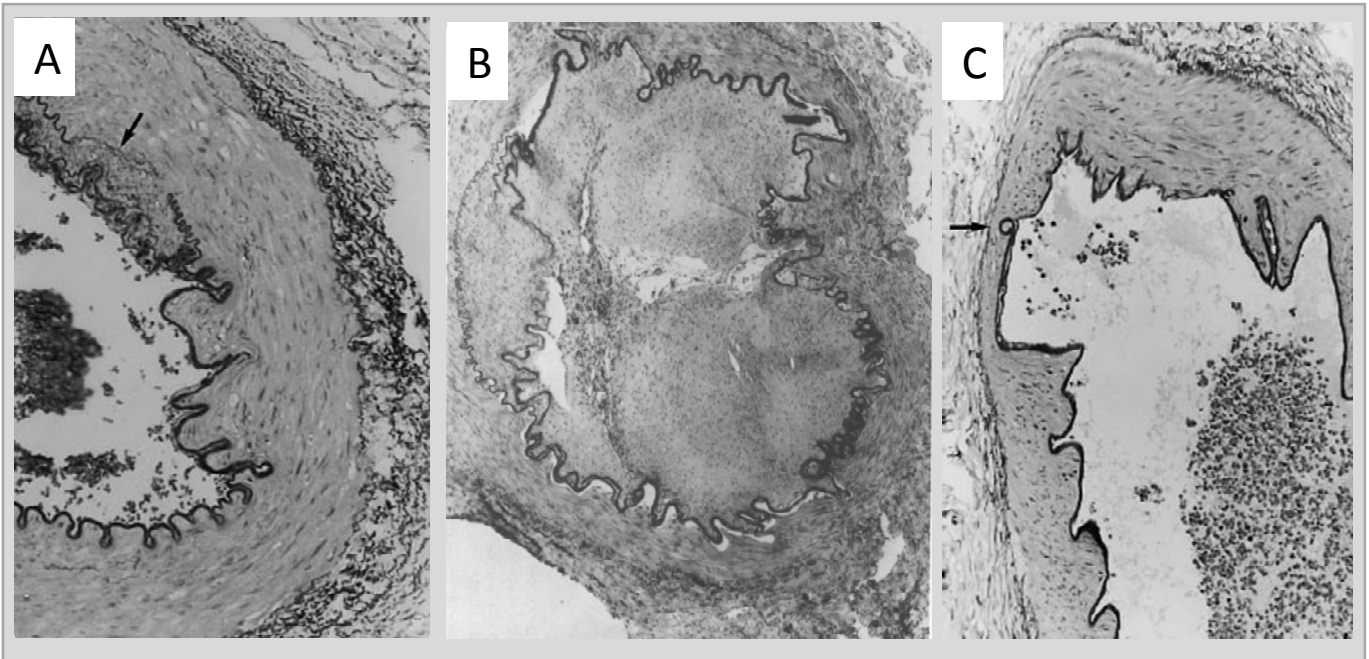


Figure 7: aspect histologique artériel de patients présentant une hémoglobinopathie. A et B: lésions d'une artère splénique de patient porteur de sphérocytose héréditaire. C lésion carotidienne interne d'un patient porteur d'une drépanocytose. D'après Merckel et al, 1978, Tsomi et al, 2001.

vasculaire, et présente des caractéristiques proches de la DFM (figure 9) : Une étude histologique portant sur 30 cas et 5 contrôles biospiés révèle une altération importante des nerfs périphériques et des vaisseaux digitaux (Takeuchi et al., 1986). Au sein des nerfs périphériques les axones sont peu nombreux et de faible calibre, et présentent une démyélinisation et un excès de collagène dans les couches endo- et perineurales. Au niveau vasculaire, il existe un épaissement des trois couches artérielles : 30% des vaisseaux biospiés présentent une fibrose intimale, 100% une fibrose adventitielle. Une hypertrophie et une hyperplasie continue de la média sont présentes sur tous les vaisseaux biopsiés. Sur le plan physiopathologique, ces vaisseaux sont caractérisés par une hypersensibilité à la stimulation α_2 adrénergique (responsable d'une vasoconstriction) et d'une insensibilité relative à la stimulation cholinergique (vasodilatatrice). Les vaisseaux étudiés dans un modèle murin d'exposition aux vibrations produisent moins de NO et plus de ROS (Krajnak, 2005; Krajnak et al., 2009) conséquence d'après les auteurs de l'hypersensibilité adrénergique.

L'endothéline a été identifiée comme un marqueur sérique et salivaire positivement corrélé à l'exposition aux vibrations dans cette pathologie professionnelle (Bovenzi et al., 2008; Kohout et al., 1995). Ce peptide, vasoconstricteur puissant sécrété par l'endothélium vasculaire, est responsable d'une contraction des CML par activation de la voie RAS (voir section 9.1.1.1) puis la cascade des MAP kinases, et de RhoA, protéine régulatrice de l'actine. Cette activation se produit par association de l'endothéline sur son récepteur (EDNR), couplé à diverses protéines G (Gaq, Ga11, Ga12/13), responsables de la transduction de ce signal (figure 19). La cascade MAPK a elle aussi été impliquée comme une voie activée dans l'endothélium soumis à un phénomène vibratoire, par phosphorylation de ERK1/2 (White et al., 2004). Stonyeva et coll. en 2003 ont récapitulé les mécanismes physiopathologiques à l'œuvre dans la dysfonction vasculaire dans cette pathologie impliquant de nombreux vasoconstricteurs et vasodilatateurs, ainsi qu'une dysfonction des terminaisons nerveuses périvasculaires (figure 10, Stonyeva, 2003).

Sur le plan histologique, toutes les pathologies développées dans l'hypothèse mécanique présentent un chevauchement incomplet avec la DFM. Dans ces exemples de pathologies, le stimulus mécanique est responsable d'une fibrose par dépôt de collagène, et d'une prolifération des cellules musculaires lisses se localisant dans la média (dysplasie et hyperplasie médiale) ou migrant vers l'intima (hyperplasie myointimale). Le stimulus mécanique est probablement responsable d'une turbulence du flux sanguin, générant de fortes contraintes de cisaillement (shear stress), et possiblement d'un étirement des cellules musculaires lisses (dans le cas du HAVS).

Ces deux types de stress mécanique ont été longuement étudiés et sont responsables de modifications structurelles de la paroi artérielle, par hyperplasie et hypertrophie ces CML et synthèse accrue de protéines de la MEC comme le collagène (Resnick et al., 2003; Wentzel et al., 2003). L'expression de nombreux gènes liés au remodelage vasculaire incluant la prolifération cellulaire, l'apoptose, et la migration, sont fortement

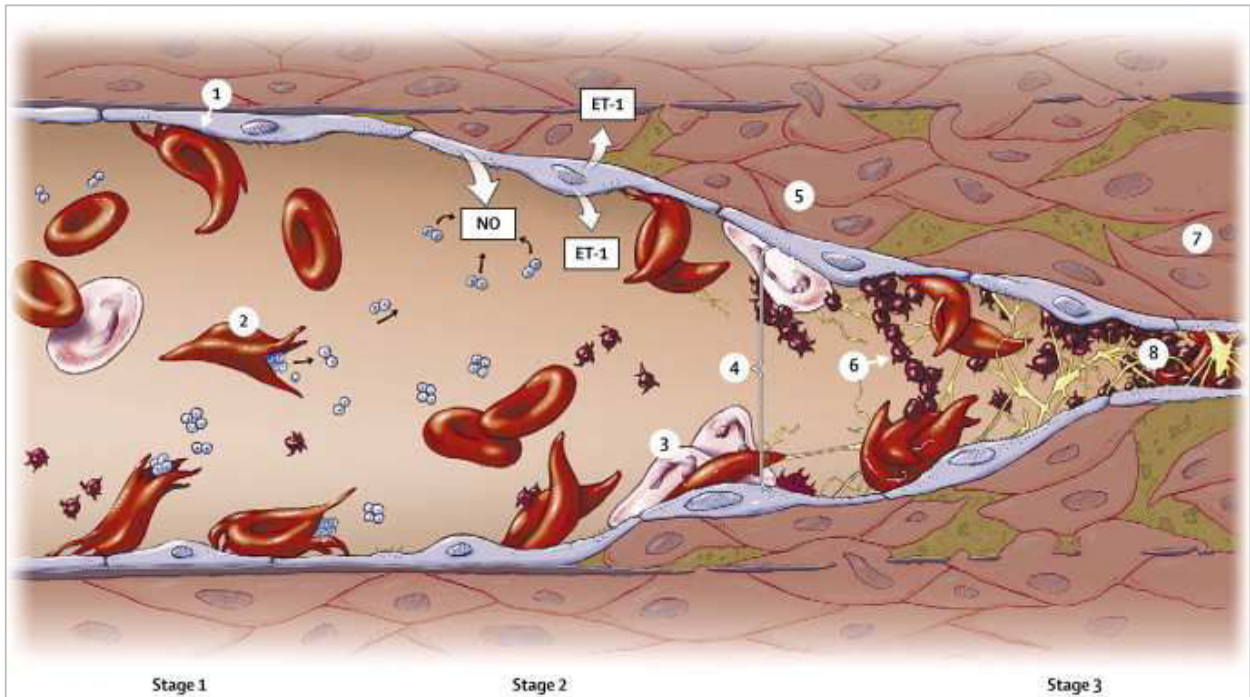


Figure 8: mécanismes impliqués dans la maladie vasculaire drépanocytaire. Adhérence anormale entre érythrocytes et endothélium (1) , hémolyse (2). Ces modifications induisent un état proinflammatoire, responsable d’une adhésion des leucocytes (3) , une aggrégation plaquettaire (6). Une sécrétion accrue d’endothéline et un piégeage du NO par les dimères d’hémoglobine libre augmente le tonus vasculaire (4). Le rétrécissement de la lumière survient ensuite par prolifération des CML et des fibroblastes qui migrent dans l’intima (5), responsables d’une vasculopathie (7) et de phénomènes occlusifs (8). D’après Switzer et al, 2006.

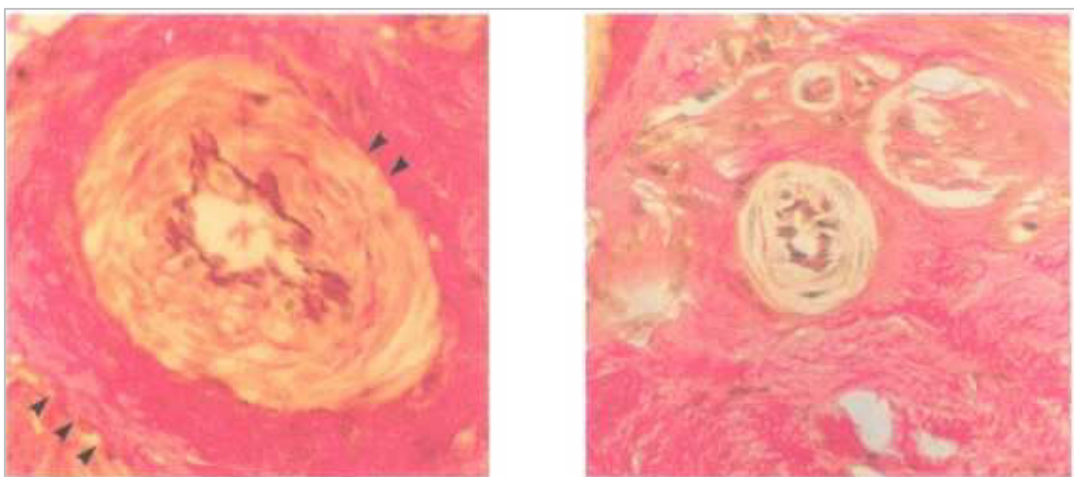


Figure 9: Aspect histologique des lésions artérielles des patients atteints de HAVS. La media est le siège d’une hyperplasie des CML, associée à une fibrose périvasculaire. (flèches). D’après Takeuchi et al, 1986.

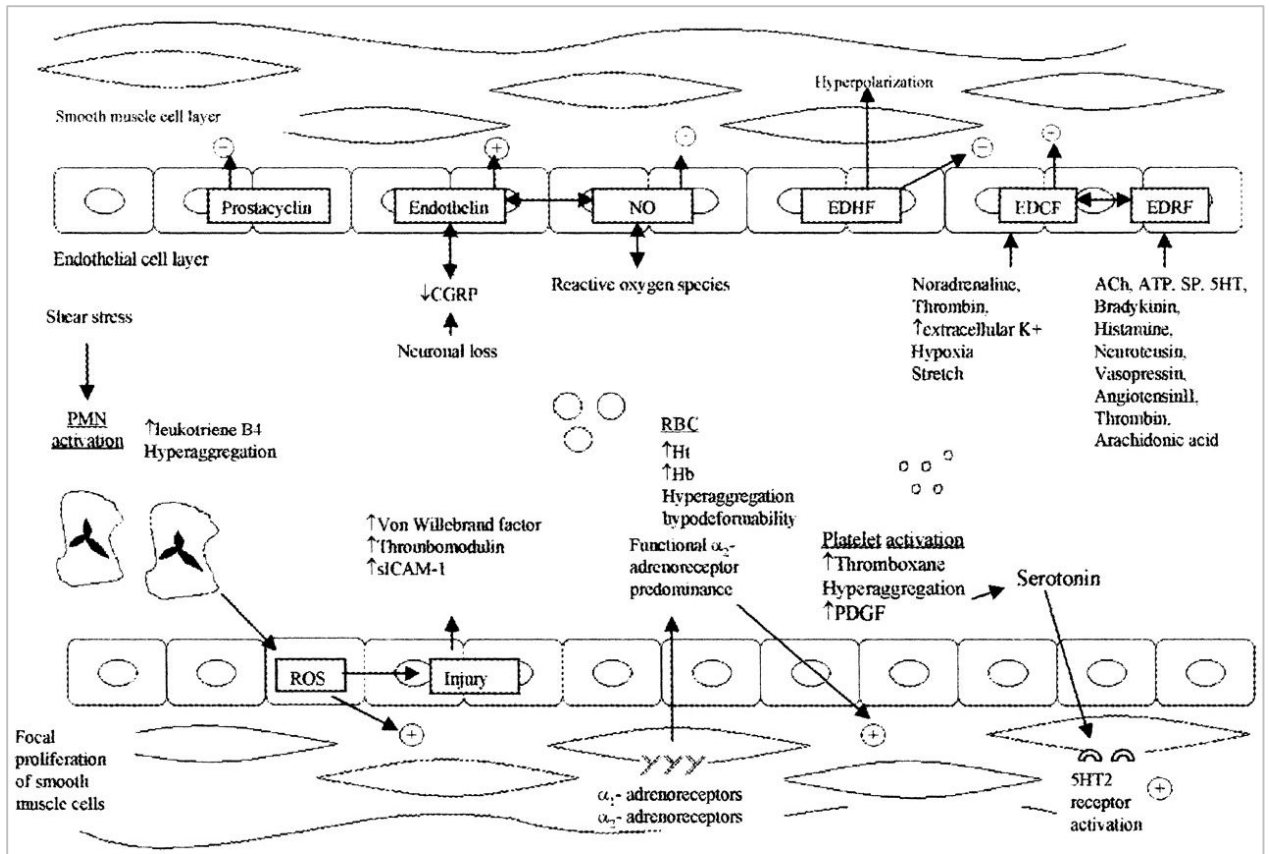


Figure 10: Modèle physiopathologique local de l'atteinte vasculaire dans le HAVS. Les + et - signifient respectivement une élévation ou une diminution de la concentration de la molécule correspondante ou de son effet. Ach: acétylcholine; ATP: adenosine triphosphate; CGRP: calcitonin gene-related peptide; EDCF, endothelial-derived constricting factor; EDHR: endothelium-derived hyperpolarizing factor; EDRF: endothelial-derived relaxing factors; 5HT2: sérotonine; Hb, hémoglobine; Ht, hémocrit; NO: oxyde nitrique; PDGF: platelet-derived growth factor; PMN: polynucléaire; RBC: érythrocyte; ROS: dérivés réactifs de l'oxygène; sICAM-1: soluble intercellular adhesion molecule-1; SP: substance P. D'après Stoyneva et al, 2003.

impactés par ce type de stimulus. Ce remodelage des CML conduit de façon localisée à une augmentation de l'épaisseur de paroi et une diminution du diamètre luminal, responsable d'un phénomène de turbulence accru qui entretient voire accélère le processus sténotique local.

9. Diagnostics différentiels et pathologies vasculaires approchantes : une classification physiopathologique

9.1 Les maladies sténosantes

La DFM est une maladie sténosante en premier lieu, qui se complique secondairement d'anévrismes et de dissections.

Un certain nombre de maladies héréditaires, syndromiques ou isolées, sont responsables de sténoses uniques ou multiples et présentent des caractéristiques histologiques partiellement chevauchantes avec la DFM. On peut distinguer plusieurs groupes selon les cascades biologiques impliquées : les maladies de la voie RAS, les maladies de la voie NOTCH, les maladies de la cascade du NO, et les maladies de l'appareil contractile de la CML (liées à ACTA2).

9.1.1 Les maladies de la voie RAS

9.1.1.1 La voie RAS

La voie Ras / MAP kinases (MAPK) joue un rôle essentiel dans le développement et est activée par un stimulus extracellulaire (Figure 11). La famille des gènes RAS est composée de HRAS, NRAS et KRAS. Les protéines Ras, véritables plaques tournantes de nombreux processus cellulaires, sont des petites GTPases liant la guanosine. Elles sont activées par des facteurs de croissance extracellulaires interagissant avec divers types de récepteurs : (1) les RTK ou récepteurs à tyrosine kinase (car ils contiennent tous un domaine tyrosine kinase) couplés aux protéines G, (ii) les récepteurs de cytokines (iii) et des récepteurs de certaines protéines de la matrice extracellulaire. La protéine RAS est présente sous deux formes, l'une active liée au GTP, et l'autre inactive liée au GDP. L'activation se produit *via* un stimulus de type facteur de croissance se liant au RTK, ce dernier s'autophosphoryle et interagit avec une protéine adaptatrice appelée GRB2, responsable du recrutement à la membrane plasmique d'une protéine nommée SOS. Cette dernière est un échangeur de guanosine di- et tri-phosphate qui augmente la proportion de forme activée de RAS (RAS GTP).

La voie MAPK est une cascade majeure de signalisation en aval de Ras. Ras activé conduit à l'activation de Raf (aRAF, bRAF, cRAF), la première MAPK de la voie (c'est donc une MAPKKK). Raf phosphoryle et active

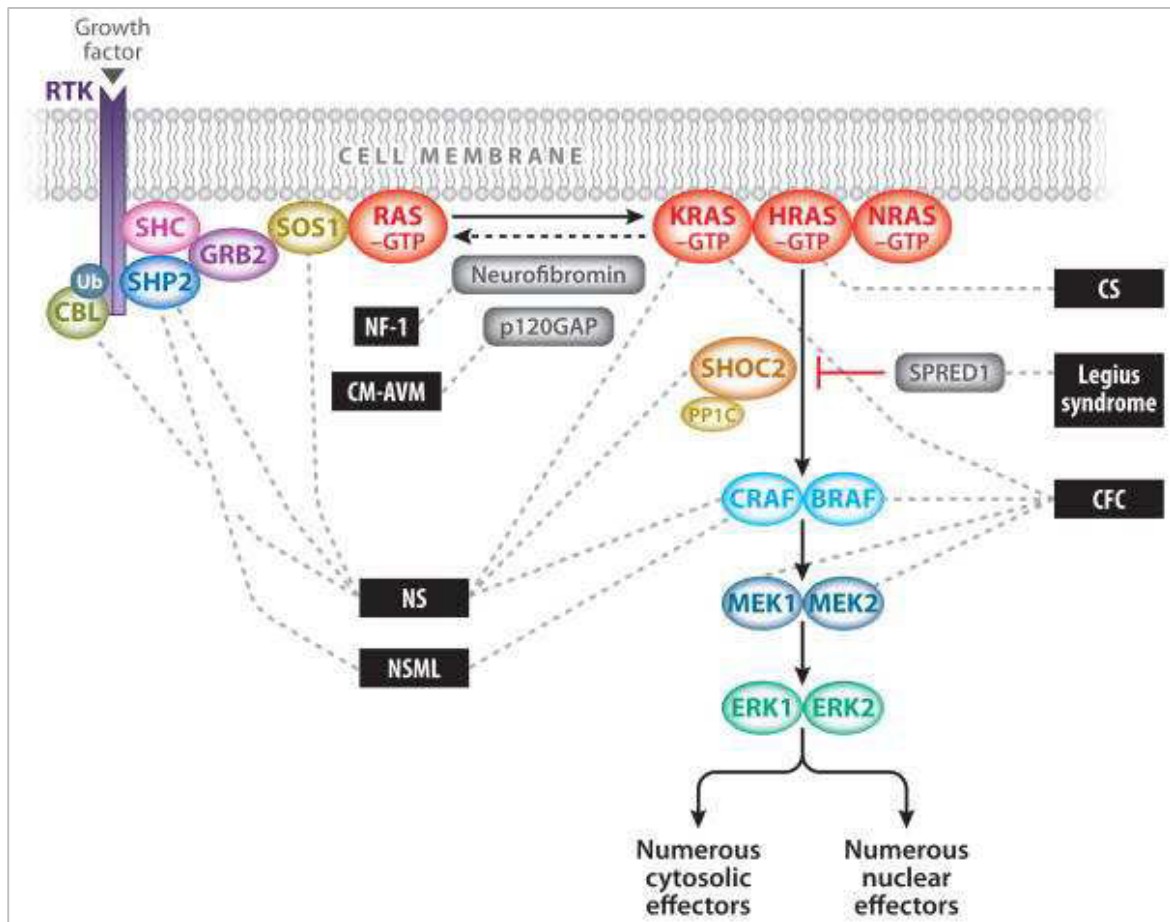


Figure 11: la voie RAS et les pathologies associées. NF-1: neurofibromatose de type 1. NS: Syndrome de Noonan. CS: syndrome de Costello. CFC: syndrome Cardio-Facio-Cutané. CM-AVM: syndrome Malformations capillaires et artérioveineuses. NSML: syndrome de Noonan avec lentilles multiples. D'après Rauen, 2013.

MEK1 et / ou MEK2 (MAPKK), qui à leur tour, phosphorylent et activent ERK1 et/ou ERK2 (MAPK). ERK1 et ERK2 sont les effecteurs finaux et interagissent avec un grand nombre de partenaires en aval, à la fois nucléaires et cytosoliques : facteurs de transcription, protéines membranaires et protéines kinases du cycle cellulaire, de la différenciation et la croissance cellulaire, et de la sénescence (Rauen, 2013).

La voie Ras / MAPK a été largement étudiée en oncologie à cause de son implication dans de nombreux cancers. Ras et BRAF sont mutés dans environ 30% des tumeurs malignes (Rauen, 2013). Pour cette raison, les maladies constitutionnelles de la voie RAS ou « RASopathies » sont classées comme des syndromes cancéreux, la majorité des mutations associées correspondant à une activation de la voie RAS. Cette activation induite par les mutations germinales est toutefois moindre que dans les mutations somatiques, ces dernières étant probablement embryolétales.

9.1.1.2 Les RASopathies

Les RASopathies sont causées par des mutations germinales des gènes de la voie Ras / MAPK. Chaque RASopathie présente un phénotype unique, mais en raison des mécanismes communs induits par l'hyperactivation de la voie RAS, elles partagent de nombreuses caractéristiques communes (dysmorphie faciale, malformations cardiaques ...) et notamment des atteintes vasculaires à type de sténoses :

La neurofibromatose de type 1 (NF1) a été le premier syndrome identifié impliquant la voie Ras / MAPK (Cawthon et al., 1990). Les vasculopathies ne sont pas rares dans la pathologie (de l'ordre de 10%) et sont multifformes et multisites : les anévrismes sont les plus fréquents (environ 60%), suivi de sténoses artérielles (30%), puis les malformations artérioveineuses et les anomalies valvulaires (10%). L'aorte, les artères rénales et mésentériques, cervicales sont impliquées avec une fréquence cumulée de 85%, suivies des atteintes des artères intracrâniennes qui sont plus rares (Oderich et al., 2007). Les atteintes vasculaires ne sont pas congénitales mais apparaissent généralement entre la première et la seconde décennie. Elles ont la deuxième cause de décès chez les patients atteints de NF1 (Oderich et al., 2007), secondaires à une hypertension rénovasculaire, des AVC, plus rarement une rupture d'anévrisme. Les atteintes des artères rénales sont les plus fréquentes, et revêtent les caractéristiques angiographiques des DFM focales (sténoses longues ou courtes) et parfois multifocales avec présence de multiples anévrismes et sténoses. Histologiquement, les lésions des artères rénales dans la NF1 sont associées à une prolifération des cellules musculaires lisses des artères de moyen calibre, et parfois de l'aorte. Elles sont considérées comme des lésions dysplasiques, généralement de type fibroplasie intimale (Oderich et al., 2007; Salyer et Salyer, 1974). Cependant nous avons

observé des lésions de type DFM médiale dans des cas de maladies artérielles rénales de patients NF1 (données non publiées).

Sur le plan physiopathologique, il a été démontré qu'une inactivation de *NF1* (modèle murin ou cellulaire *Nf1^{+/-}*) était responsable d'une activation de la voie RAS et MAPK dans les cellules endothéliales et les macrophages. Cette activation induit secondairement une activation de la voie AKT *via* une molécule appelée mTOR, voie aussi impliquée dans la régulation de la croissance, la survie, et le cycle cellulaire, la transcription et la traduction de nombreuses protéines (Stansfield et al., 2014).

Dans ce modèle murin, lorsqu'on lèse l'artère carotide, elle est le siège d'une prolifération neointimale et d'une diminution de la lumière artérielle. Cette observation est à rapprocher de l'hypothèse mécanique, en ce sens que la vasculopathie de la NF1 serait liée à un processus de réparation inadapté à une agression endovasculaire. L'atteinte vasculaire pourrait survenir à la suite de l'incapacité de l'endothélium à supprimer la migration et la prolifération excessive des CML en réponse aux agressions endovasculaires (Hamilton et Friedman, 2000).

Ce phénomène a été démontré dans les lésions cutanées de la NF1 : la neurofibromine est surexprimée dans le processus de réparation de plaies cutanées (Atit et al., 1999) ; Chez la souris *Nf1^{+/-}*, il existe une prolifération excessive des fibroblastes en réponse à une plaie cutanée, liée à la dérégulation de la voie RAS/MAPK.

Cette hypothèse d'un facteur mécanique pourrait expliquer la diversité des localisations des lésions artérielles rencontrées dans la pathologie

Les **syndromes de Noonan** (et syndrome Léopard) sont causés par des mutations activatrices dans de nombreuses protéines de la voie RAS : *SHP2 (PTPN11)*, *SOS1*, *RAF1*, *KRAS*, *NRAS*, *SHOC2* et *CBL* (Rauen, 2013). Le syndrome cardio-facio-cutané (CFC) est causé par l'activation de la voie MAPK par des mutations dans *BRAF*, *MEK1 (MAP2K1)* ou *MEK2 (MAP2K2)* (Rauen, 2013). L'anomalie vasculaire la plus fréquemment rencontrée dans ces deux pathologies est la sténose valvulaire pulmonaire, mais dans le syndrome de Noonan il n'est pas rare d'observer des sténoses ou dilatations supra-avalvulaires pulmonaires, une coarctation de l'aorte ou une dilatation de l'aorte ascendante, et un syndrome de Moya-Moya (voir section 9.1.3, Cornwall et al., 2014; Fujimoto et al., 2014; Schuster et Roberts, 1999). Les atteintes artérielles sténosantes dans ces pathologies sont caractérisées par une prolifération myointimale et une rupture de la limite élastique interne (Jimenez et al., 1999).

Les autres RASopathies, syndrome de Legius, causé par des mutations d'inactivation dans *SPRED1* (Rauen, 2013), et le syndrome CM-AVM (malformations capillaires et artérioveineuses) causé par des mutations de

RASA1, ne sont pas connues pour donner des sténoses ou des anévrismes artériels, bien que ce dernier soit une maladie malformative purement vasculaire.

9.1.1.3 La sténose supra-avalvulaire aortique et l'haploinsuffisance du gène *ELN*

La sténose supra-avalvulaire aortique, le cutis laxa autosomique dominant et le syndrome de Williams partagent en commun une haploinsuffisance pour le gène *ELN* codant l'élastine. Cette haploinsuffisance est responsable de sténoses aortiques, et plus rarement de dilatations de l'aorte ascendante voire de maladie artérielle plus diffuse (Graul-Neumann et al., 2008; van Son et al., 1994; Szabo et al., 2006). L'analyse histopathologique révèle une atteinte médiale, avec hypertrophie et désorganisation des cellules musculaires lisses (figure 12), des dépôts de collagène et une raréfaction des fibres élastiques qui sont fragmentées et interrompues (Stamm et al., 2001). Des aspects de fibroplasie médiale sont identifiés dans les lésions des artères de gros et moyen calibre, notamment les artères coronaires et pulmonaires (van Son et al., 1994). Les lésions histologiques rencontrées dans cette pathologie sont donc très proches de la DFM médiale. Sur le plan physiopathologique, il a été démontré que les produits de dégradation de l'élastine sont des molécules stimulatrices de la prolifération des cellules musculaires lisses (Hinek et Wilson, 2000). Cette prolifération est la conséquence d'une activation de la voie RAS/MAPK par les récepteurs membranaires à l'élastine, couplés aux protéines G (Mochizuki, 2002). La transduction du signal par ce récepteur induit aussi un flux de calcium intracellulaire, responsable d'une plus grande contractilité des CML (figure 13). Il a été démontré qu'une activation de mTOR (molécule activatrice de la voie AKT sus-citée) était présente dans les aortes de souris déficientes en élastine. L'administration de rapamycine, inhibiteur naturel de mTOR, réprime chez ces souris la prolifération des cellules musculaires lisses de la média (Li et al., 2013). Enfin, dans un modèle de fibroblastes, une activation de la voie MAPK *via* la voie AKT/PI3K, indépendamment de la voie RAS, a été décrite (Duca et al., 2005).

Ainsi, les pathologies liées au déficit en élastine ne sont pas à proprement parler des RASopathies, mais elles induisent, par stimulation du récepteur à l'élastine une activation relativement directe de la voie RAS ayant pour conséquence une prolifération des CML. Dans ce groupe de pathologies cependant, le phénotype histologique semble plus proche de la DFM médiale que dans les RASopathies pures.

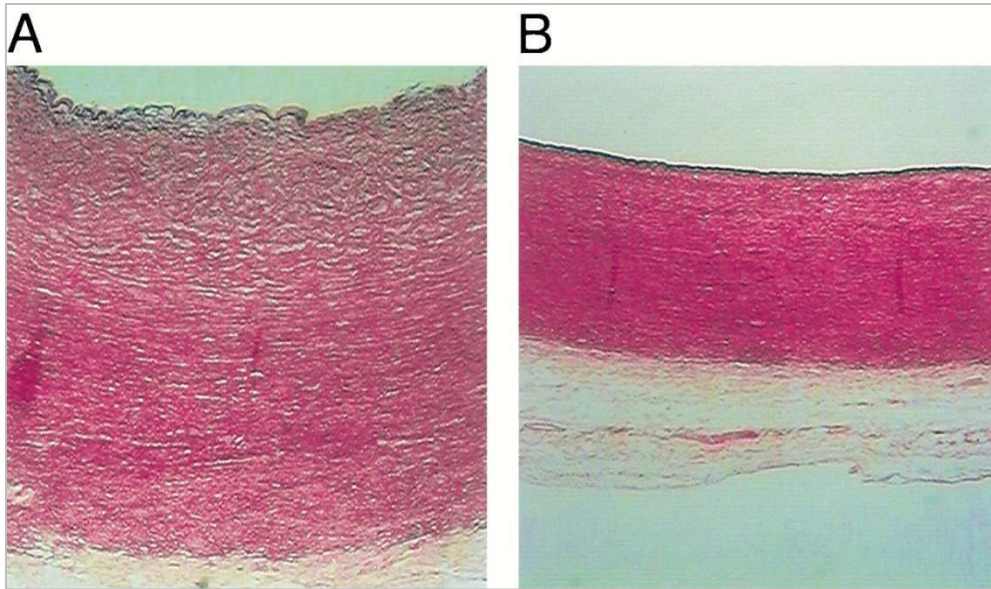


Figure 12: atteinte aortique dans la sténose supra-valvulaire aortique et le cutis laxa autosomique dominant (A) en comparaison d'une aorte normale (B). Coloration HES. D'après Stamm et al, 2001.

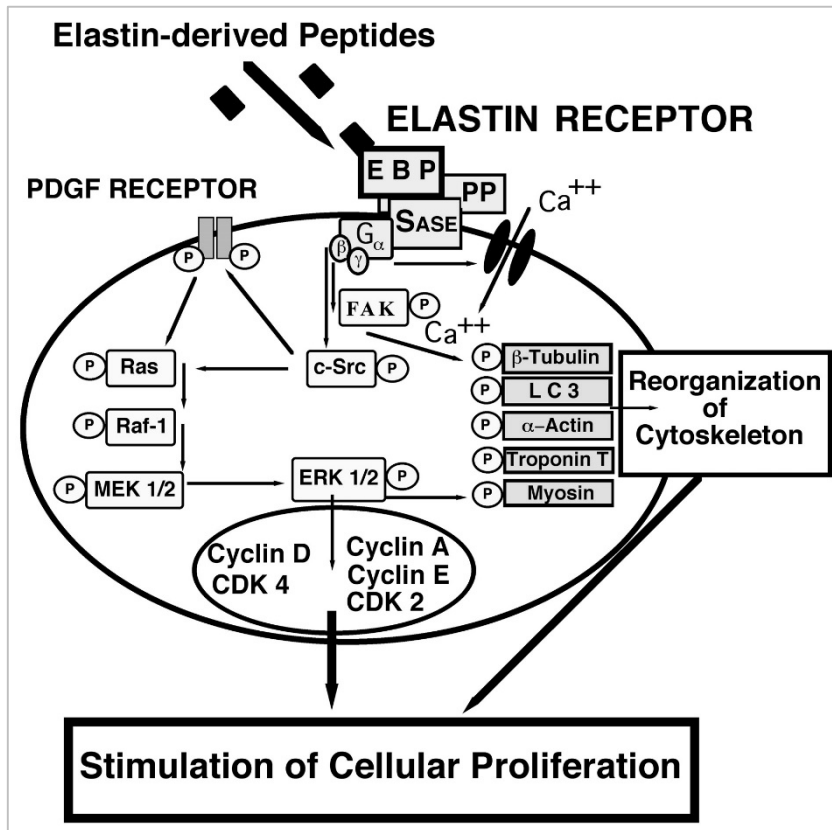


Figure 13: Modélisation des cascades de signalisation induites par activation du récepteur ER par les peptides dérivés de la dégradation de l'élastine. D'après Mochizuki et al, 2002.

9.1.2 Les maladies de la voie NOTCH

9.1.2.1 La voie Notch

La voie Notch est une voie de signalisation extrêmement conservée qui joue un rôle majeur dans de nombreux processus développementaux impliquant des phénomènes de différenciation et d'identité cellulaire, notamment cardiovasculaires (High et Epstein, 2008). Chez les mammifères, il existe 4 protéines NOTCH (NOTCH1 à 4) qui sont des récepteurs à un seul domaine transmembranaire. Le récepteur NOTCH (figure 14) est composé d'une partie extracellulaire interagissant avec son ligand (les protéines JAGGED et DELTA-LIKE), une région transmembranaire et une région intracellulaire appelée NICD (NOTCH IntraCellular Domain). L'interaction avec le ligand conduit à deux clivages protéolytiques successifs, l'un extracellulaire par une protéine ADAM qui libère le complexe ligand-domaine extracellulaire de NOTCH, l'autre par le complexe γ -sécrétase qui libère le NICD. Ce dernier est un régulateur transcriptionnel qui se transloque dans le noyau où il forme un complexe avec un facteur de transcription appelé CSL. Les cibles majoritaires de cette activation transcriptionnelle sont les gènes des familles *HES* et *HEY*, répresseurs transcriptionnels impliqués dans la différenciation et la prolifération cellulaire.

9.1.2.2 Les mutations de *NOTCH1*

La bicuspidie aortique est une atteinte malformative de la valve aortique qui touche 1 à 3% de la population (Tzemos et al., 2008). En 2005, *NOTCH1* a été identifié dans des formes dominantes de bicuspidie aortique non syndromique, en association cependant avec des calcifications valvulaires aortiques, et des atteintes ventriculaires gauches (Garg et al., 2005). Il a été démontré par la suite une association non fortuite entre coarctation de l'aorte et bicuspidie aortique, y compris chez les patients porteurs de mutations *NOTCH1* (McBride et al., 2005). J'ai séquencé *NOTCH1* dans une série de 60 cas de BAO et confirmé l'association entre ces deux anomalies (données non publiées).

9.1.2.3 Les mutations de *NOTCH2*, *JAG1*, et le syndrome d'Alagille

Le syndrome d'Alagille est une maladie syndromique associant une hypoplasie des canaux biliaires intrahépatiques avec cholestase, des anomalies cardiovasculaires, vertébrales, oculaires, et une dysmorphie caractéristique. Des mutations de *JAG1* ont été identifiées dans 60% à 75% des cas, *NOTCH2* étant plus rarement impliqué (McDaniell et al., 2006; Spinner et al., 2001). De multiples anomalies cardiovasculaires peuvent être associées à cette maladie, en premier lieu les sténoses des artères ou de la valve pulmonaires, la tétralogie de Fallot, les CIV et CIA (McElhinney et al., 2002). Les patients peuvent plus rarement avoir une coarctation de l'aorte, un syndrome de Moya-Moya, des sténoses des artères rénales responsables

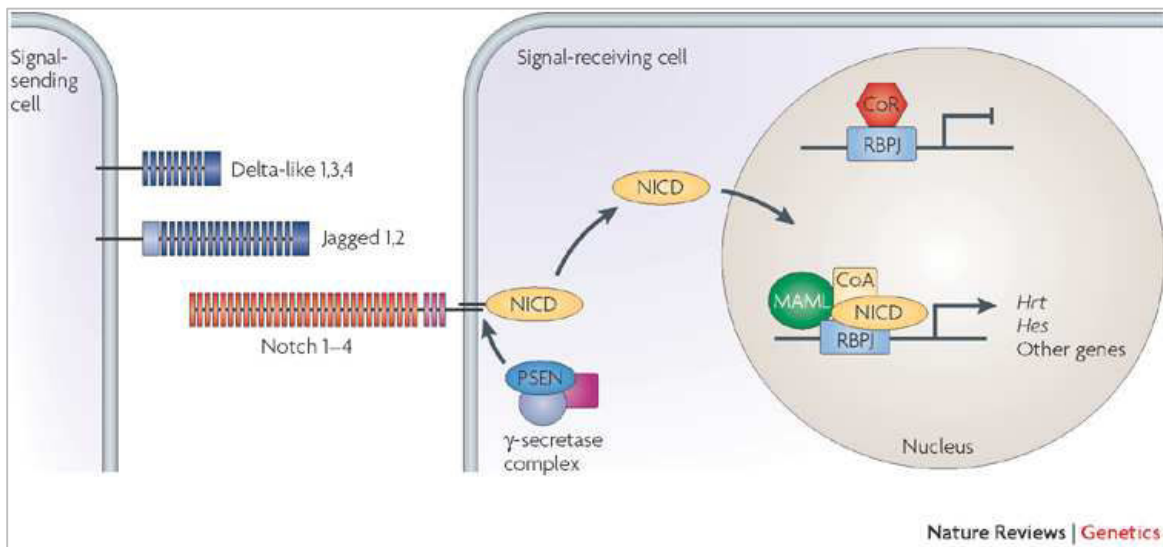


Figure 14: la voie NOTCH. Le récepteur est transmembranaire, et après activation par l'un de ses ligands subit une double protéolyse libérant un peptide extracellulaire (NECD) et un peptide intracellulaire (NICD), qui transloqué au noyau agit comme facteur de transcription sur ses gènes cibles, dont HEY, HES, HRT. D'après High et al, 2008.

d'hypertension rénovasculaire (Bérard et al., 1998; Kamath et al., 2004; Salem et al., 2012), et des anévrismes systémiques ou intracrâniens. Dans une analyse IRM de 26 patients sans symptômes neurologiques, environ 40% étaient porteurs d'une atteinte sténosante ou anévrismale des artères intracrâniennes (Emerick et al., 2005). Les lésions sténosantes, notamment artérielles, se présentent à l'imagerie comme une dysplasie focale courte ou tubulaire (Figure 15a).

L'analyse histologique chez ces patients révèle que la média est le siège d'une prolifération concentrique de cellules musculaires lisses qui sont responsables de la diminution du diamètre intraluminal (figure 15b).

9.1.2.4 Les mutations de *NOTCH3* et le syndrome CADASIL

Le syndrome CADASIL (Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy), est une maladie progressive qui associe des accidents vasculaires cérébraux, ischémiques et des migraines. Le début des symptômes se situe vers 40 ans. L'évolution est marquée par une récurrence des symptômes vasculaires, une atteinte cognitive et psychiatrique pouvant évoluer vers la démence. Cette maladie exclusive des artères cérébrales de petit calibre est la conséquence de mutations probablement gain de fonction de *NOTCH3*, secondaires à un défaut de repliement de la protéine qui devient résistante au clivage protéolytique. *NOTCH3* est maintenue à la membrane plasmique et séquestre son ligand. Il s'agit donc d'un processus différent de l'haploinsuffisance rencontrée dans les syndromes liés à *NOTCH1* et *2*.

Histologiquement, les artères cérébrales de petit calibre ont une paroi épaissie, pauvre en cellules musculaires lisses, et fibreuse, donnant une forme arrondie aux artérioles. L'adventice est aussi le siège d'une fibrose importante. Ce processus peut entraîner une oblitération complète des artérioles. L'immunohistochimie (IHC) révèle la présence de dépôts granulaires de peptide extracellulaire de *NOTCH3* dans la média.

Il n'existe pas de maladie humaine connue liée à *NOTCH4*, bien qu'un modèle murin présente des malformations artérioveineuses cérébrales majeures (Murphy et al., 2008). Dans le syndrome d'Alagille, l'atteinte artérielle des artères rénales peut mimer une DFM « idiopathique » et en partage certains aspects histopathologiques. Les atteintes vasculaires liées à *NOTCH1* et *NOTCH3* semblent cependant moins correspondre à un équivalent de DFM.

Sur le plan physiopathologique, il a été démontré que l'activation spécifique de la voie *JAG1/NOTCH2* (mais pas *NOTCH1* et *3*) dans les cellules endothéliales et musculaires lisses est responsable d'une inhibition de la prolifération des cellules musculaires lisses, notamment en cas d'agression endovasculaire. Dans ce modèle



Figure 15: lésions vasculaires dans le syndrome d'Alagille. A: aspect angiographique d'une sténose tubulaire courte d'une artère rénale. B: Aspect histologique d'une lésion sténosante artérielle ulnaire. La media est très hyperplasique et le siège d'une fibrose importante. D'après Bérard et al, 1998, Salem et al, 2012.

murin KO *Jag1*, la ligature carotidienne induit une prolifération intense des CML dans l'intima et la média qui est hyperplasique (figure 16). Le modèle proposé par ces auteurs est celui d'une dérégulation de la protéine p27 (Cdkn1b), responsable d'une mitogénèse intense (figure 17). Ces dernières études révèlent l'implication de la voie NOTCH dans le remodelage du phénotype des CML dans la paroi vasculaire mature.

9.1.3 La cascade du NO

L'oxyde nitrique ou NO, est synthétisé par trois enzymes appelées NO synthases : la NO synthase endothéliale (eNOS ou NOS3), neuronale (nNOS ou NOS1), et inductible (iNOS ou NOS2), chacune ayant des fonctions distinctes. nNOS et eNOS sont des enzymes dépendantes du calcium et produisent de faibles quantités de NO, qui a une action essentiellement locale. Dans le cerveau et le système nerveux périphérique, le NO agit comme un neurotransmetteur, impliqué dans la neurotoxicité associée à l'AVC et les maladies neurodégénératives notamment. Dans les macrophages, iNOS produit du NO en quantité importante de façon indépendante du calcium, en réponse à des stimuli comme le lipopolysaccharide. Elle a une action tumoricide et bactéricide. Cette réponse est impliquée dans les phénomènes de choc septique (pour une revue générale voir Lundberg et al., 2008).

L'endothélium vasculaire produit du NO (parfois appelé EDRF pour Endothelium-Derived Relaxing Factor), en réponse à de multiples stimuli comme le stress mécanique (forces de cisaillement), des facteurs plaquettaires, l'acétylcholine, et diverses cytokines dont les intégrines de la famille $\beta 1$ (figure 18). L'activation de eNOS est un processus en plusieurs étapes dont la première est dépendante du calcium (Quillon et al., 2015) : l'entrée de calcium dans la cellule permet l'interaction d'eNOS avec la calmoduline, induisant son relargage membranaire à partir des caveolae ou radeaux lipidiques où elle est séquestrée. Elle s'associe à HSP90 qui rend possible sa dimérisation, et recrute la PI3K (enzyme activatrice de la voie AKT). Cette dernière en la phosphorylant permet une synthèse accrue de NO, tout particulièrement en cas de stress mécanique important (Dimmeler et al., 1999). La protéine kinase A possède aussi la propriété de phosphoryler et d'activer eNOS après stimulation mécanique (Balligand et al., 2009). De plus, la cascade MAPK est activée *via* la voie AKT en cas de stimulation répétée, induisant l'activation transcriptionnelle du gène *NOS3*. Ceci souligne la complexité et la multiplicité des moyens d'activation d'eNOS en contexte physiologique.

Le NO a des propriétés vasodilatatrices puissantes et induit donc une augmentation du débit sanguin, permettant notamment une normalisation des contraintes de cisaillement. Il a une demi-vie très courte (quelques secondes) et une action essentiellement paracrine et autocrine. Il active la guanylate cyclase soluble (cytoplasmique) qui génère du GMP cyclique (cGMP) à partir de GTP. Ce dernier est le second

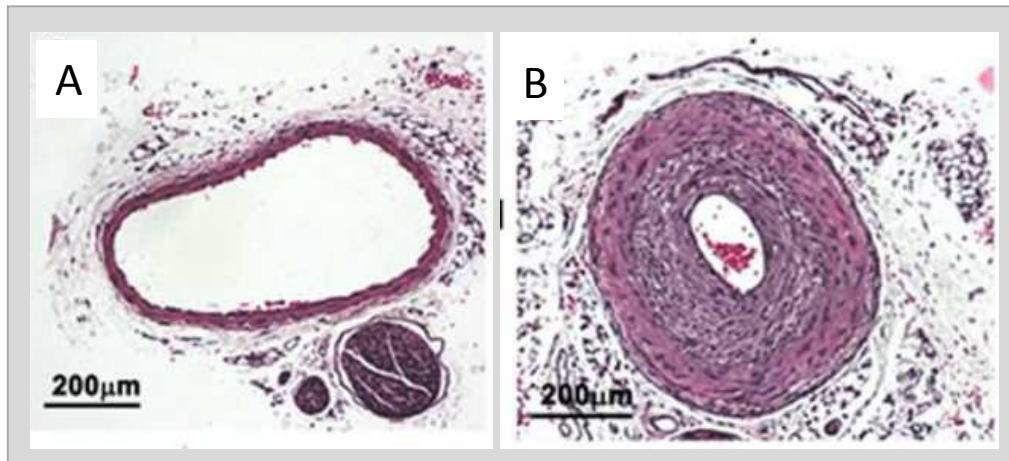


Figure 16: aspect des lésions vasculaires induites par blessure carotidienne chez le rat $JAG1^{-/-}$ en coloration HES. A: artère carotide non lésée. B: aspect hyperplasique de la media et l'intima, une semaine après lésion carotidienne. D'après Boucher et al, 2013.

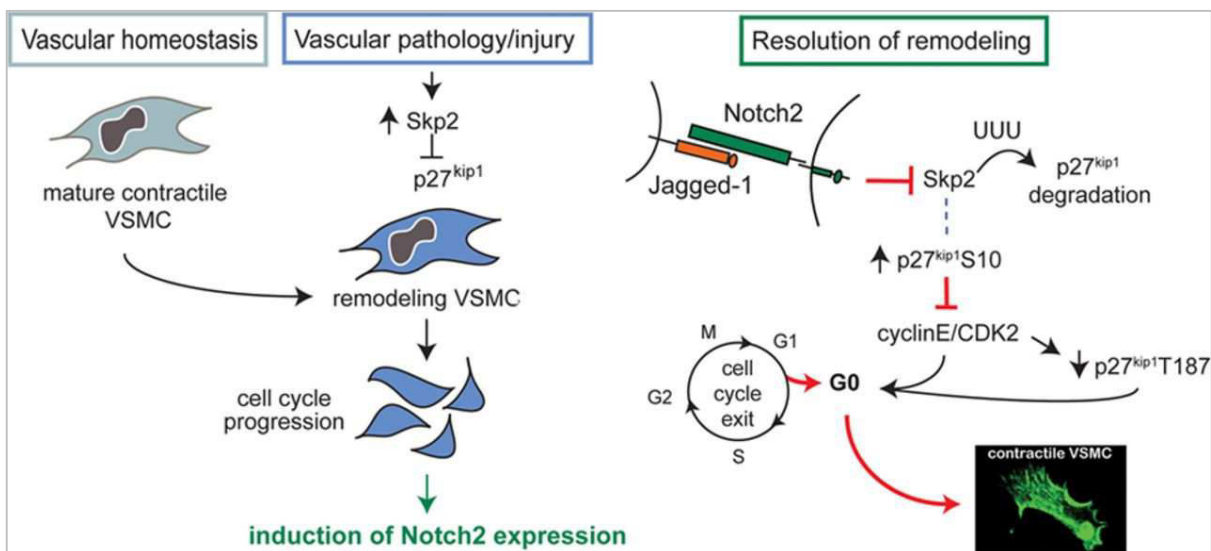


Figure 17: Modélisation de la régulation de la prolifération des CML lors du remodelage vasculaire lié à NOTCH2. D'après Boucher et al, 2013.

messenger majeur de la relaxation vasculaire. Il agit en activant la protéine kinase G qui induit la relaxation des muscles lisses par déphosphorylation des chaînes de myosine (figure 19).

9.1.3.1 Les mutations *GUCY1A3* et la maladie de Moya-Moya

La maladie de Moya-Moya est une atteinte cérébrovasculaire de cause inconnue jusqu'à très récemment, caractérisée par une sténose progressive de la partie terminale des artères carotides internes et leurs branches proximales. Cette maladie est associée au développement de vaisseaux collatéraux fins et fragiles à la base du cerveau (en aspect de bouffée de fumée ou "Moya-Moya" en japonais). La diminution du débit sanguin cérébral et la rupture des collatérales sont responsables d'AVC hémorragiques et ischémiques, chez des patients d'âge généralement infantile ou adulte jeune. Sur le plan histologique, les sténoses des portions terminales des carotides sont associées à une prolifération myointimale, une fibrose intimale, des phénomènes de thrombose luminale, et un amincissement médial (Yamashita et al., 1983).

Cette maladie est à distinguer du syndrome de Moya-Moya, que l'on peut retrouver comme précédemment cité dans la NF1, le syndrome d'Alagille, ou certaines mutations du gène *ACTA2* (voir section suivante). La prévalence de la maladie est très dépendante de l'origine ethnique puisqu'elle est 10 fois plus fréquente au Japon que dans les pays caucasiens, avec une forte composante familiale (de l'ordre de 5-10% au Japon). Divers modes de transmission ont été signalés, et plusieurs études de liaison ont été réalisées sans identification du gène jusqu'en 2014 (Roder et al., 2010). Par une approche GWAS, le gène *RNF213* a été identifié dans deux études indépendantes (Kamada et al., 2011; Liu et al., 2011b). Il code une protéine de fonction inconnue, contenant un domaine RING finger et un domaine ATPase. En 2014, l'équipe du CERVCO (Centre de Référence des Maladies Vasculaires Rares du Cerveau et de l'Oeil) a identifié par cartographie d'Homozygotie et séquençage d'exome des mutations pertes de fonction dans le gène *GUCY1A3* dans un syndrome récessif associant achalasie et maladie de Moya-Moya (Hervé et al., 2014). Ce gène code la sous-unité $\alpha 1$ de la guanylate cyclase soluble (sGC), enzyme ubiquitaire et principal récepteur du NO (figure 19).

L'hypothèse avancée par l'équipe est qu'une altération de cette voie serait responsable d'un remodelage inapproprié dans certaines localisations vasculaires où le flux laminaire est interrompu. La portion terminale de la carotide interne, de par sa conformation anatomique, est particulièrement concernée par le risque de turbulences du flux sanguin. Face à un stimulus mécanique comme une forte contrainte de cisaillement, un défaut de relaxation lié à l'inactivation de la cascade du NO pourrait être responsable du maintien de ces contraintes, et induire une prolifération des cellules musculaires lisses comme vu dans la section 8.3 L'hypothèse mécanique. Inversement, il a été récemment identifié un variant faux-sens de *GUCY1A3* présent chez les habitants des hauts-plateaux du Kirghizstan. Ce variant est responsable d'une plus grande sensibilité

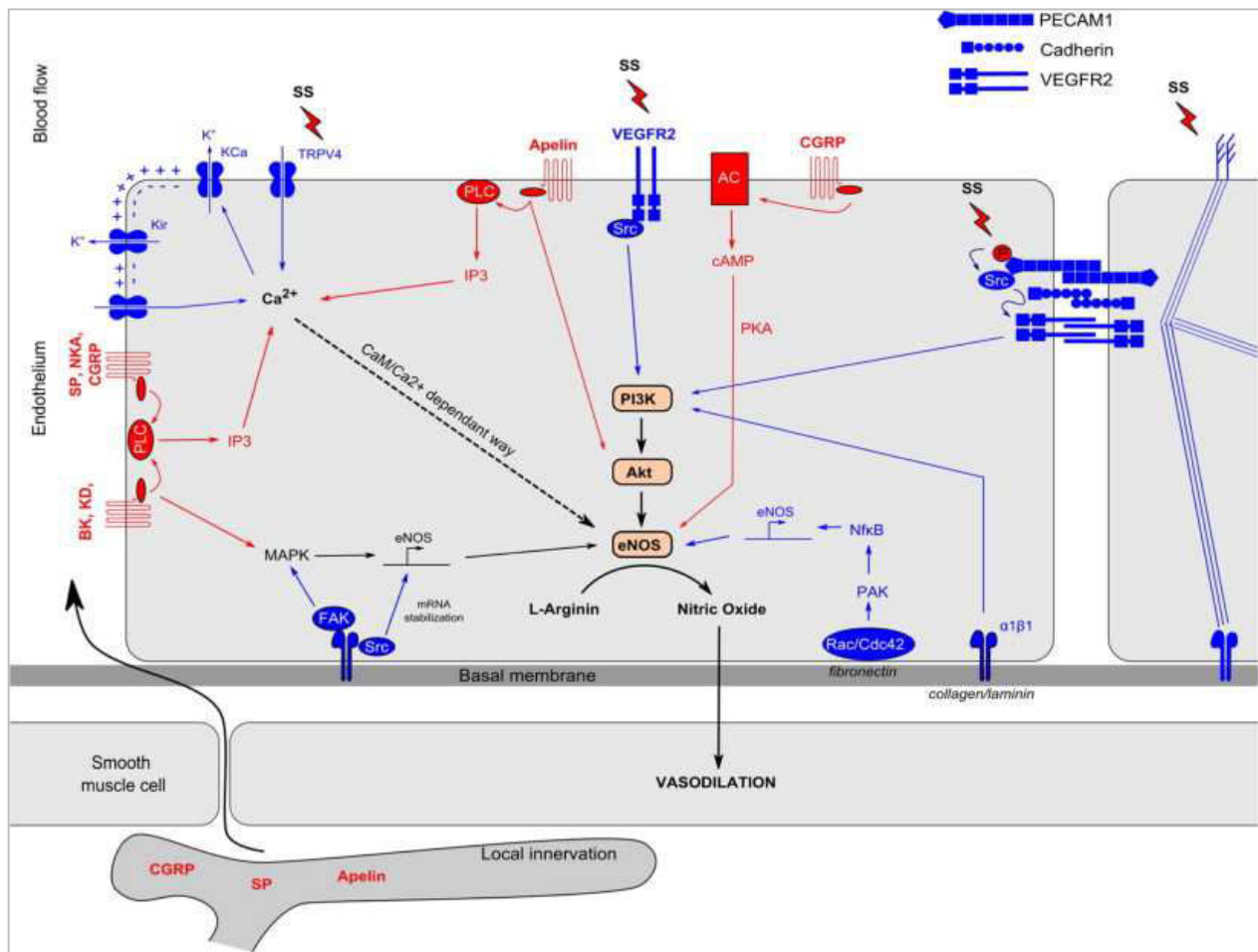


Figure 18: Modélisation de la stimulation de eNOS par perturbation du flux sanguin (shear stress, SS) ou par stimulation par les terminaisons nerveuses efférentes.

L'activation des canaux ioniques (récepteurs vanilloïdes à potentiel transitoire 4 (TRPV4) augmente la perméabilité au calcium, qui entraîne l'ouverture de canaux potassiques calcium-dépendants responsables d'une hyperpolarisation qui active eNOS par la voie Calcium-calmoduline. Le récepteur du facteur de croissance de l'endothélium vasculaire 2 (VEGFR2) active les voies PI3K et MAP kinases (MAPK). L'interaction entre divers composants de la MEC (collagène, laminine, fibronectine entre autres) et les intégrines, entraîne une adhésion focale et à l'activation de la voie MAPK, du facteur nucléaire kappa B (NfκB) et l'expression du transcrite de eNOS. Les neuropeptides d'origine humorale ou neuronale peuvent stimuler eNOS via leurs récepteurs couplés à la protéine G. La bradykinine (BK), la kallidine (KD), l'apéline, la substance P (SP), la neurokinine A (NKA) et le peptide lié à la calcitonine (CGRP) sont capables d'activer la phospholipase C (PLC) qui augmentent la concentration de calcium intracellulaire activant eNOS. De nombreuses interrelations existent entre ces différentes voies (BK et KD peuvent activer la voie MAP, l'apéline induit une phosphorylation d'Akt et le CGRP, via l'adénylate cyclase (AC), augmente la protéine kinase A (PKA) et induit secondairement l'activité d'eNOS).

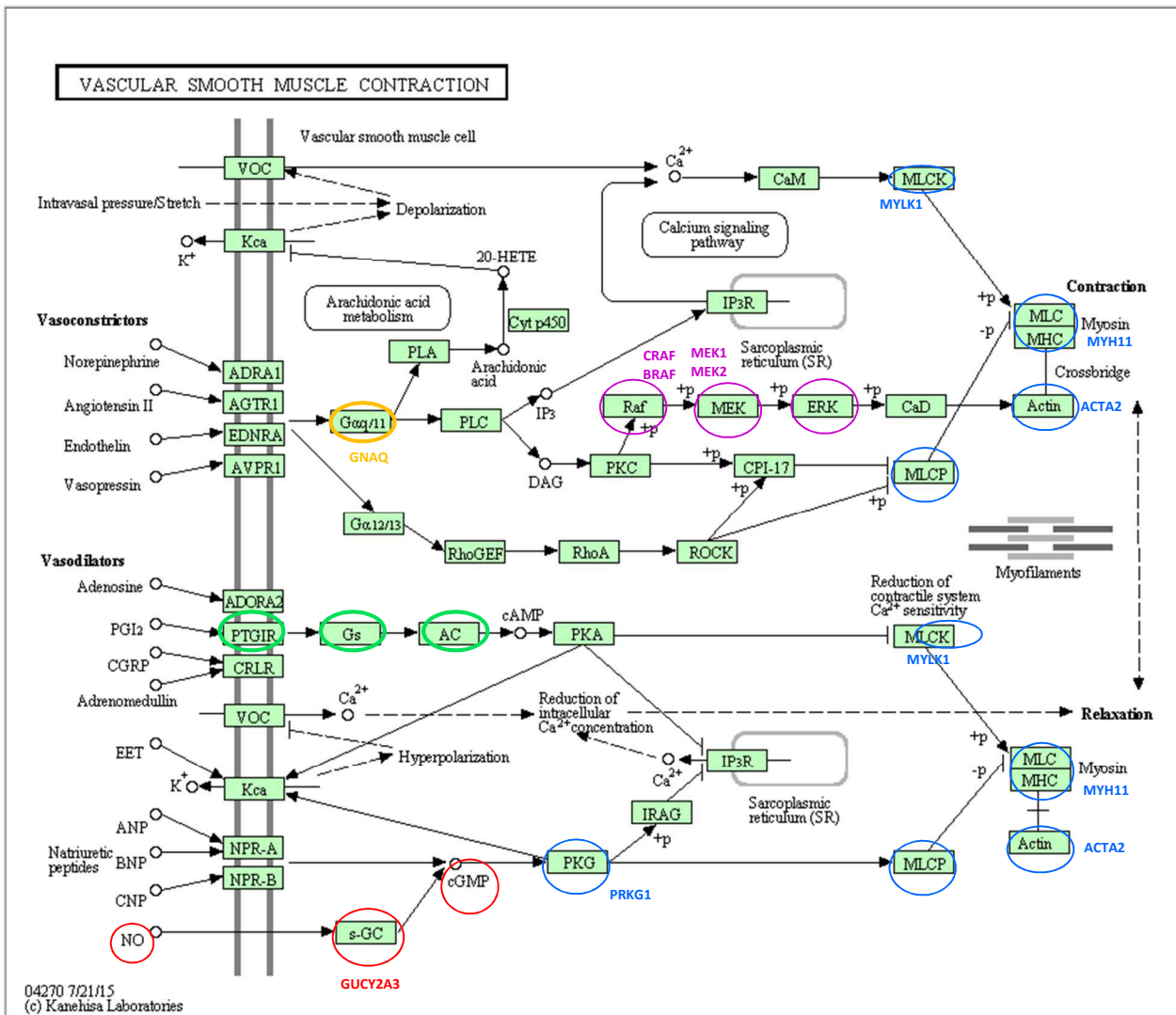


Figure 19: Modélisation de la contraction de la cellule musculaire lisse (source KEGG). Les intermédiaires de la voie du NO sont entourés en rouge, la voie RAS/MAPK (représentée en partie) est entourée en rose, les acteurs directs de la fonction contractile sont entourés en bleu. Les gènes ajoutés en couleur correspondent aux pathologies connues à composante vasculaires développées dans le manuscrit (Rouge: Moya-Moya, Rose: RASopathies, Bleu:TAAD). Une pathologie isolée dans ce diagramme (en orange) est le syndrome de Sturge-Weber, angiomatose en mosaïque proche des maladies protéiformes, liée au gène *GNAQ*.

La voie des prostacyclines a été entourée en vert, discutée en tant que gène candidat dans la dernière partie de ce chapitre.

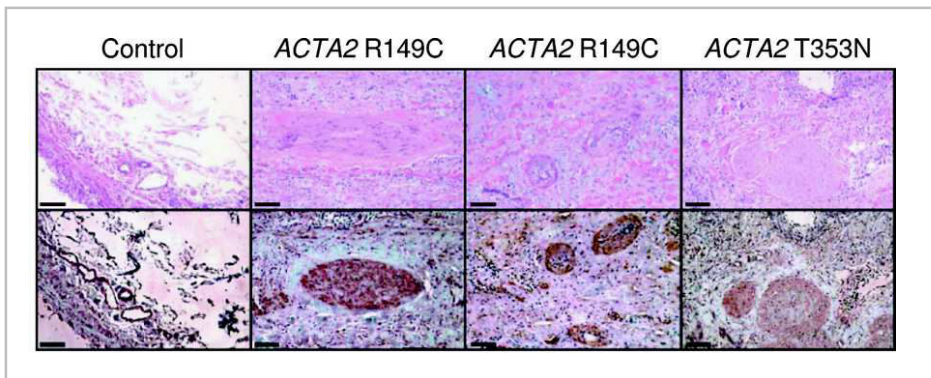


Figure 20: Hypertrophie des vasa vasorum aortiques des patients porteurs de certaines mutations d'*ACTA2*. Ligne du haut: coloration HES. Ligne du bas: immunohistochimie α -Actine spécifique du muscle lisse. D'après Milewicz et al, 2010.

de la guanylate cyclase au NO, et protège les individus qui le portent de l'hypertension pulmonaire liée à l'altitude (Wilkins et al., 2014).

9.1.4 La vasculopathie liée à ACTA2

Des mutations dans le gène *ACTA2*, codant l'actine spécifique du muscle lisse, sont responsables d'environ 10% des dissections et anévrismes de l'aorte ascendante non syndromiques (nsTAAD) (Guo et al., 2007). Sur le plan histologique, l'atteinte aortique est non spécifique, associant dégénérescence kystique de la média, perte focale des cellules musculaires lisses, et fragmentation des fibres élastiques. Ces lésions sont similaires à toutes celles identifiées dans les TAAD syndromiques et non syndromiques (voir section 9.2). Plus récemment, des mutations d'*ACTA2* ont été impliquées dans un phénotype vasculaires occlusif précoce associant atteinte coronaire, AVC et syndrome de Moya-Moya (Guo et al., 2009). Les patients de cette cohorte présentaient plus tardivement une dilatation aortique avec lésions dégénératives non spécifiques, comprenant cependant des zones focales d'hyperplasie des cellules musculaires lisses, et un épaissement de la média des artères musculaires comme les *vasa vasorum* (figure 20, Guo et al., 2007). *Ex vivo*, les cellules musculaires lisses ou les myofibroblastes issus de ces patients présentent une prolifération accrue en comparaison de CML non mutés. Le mécanisme impliqué dans cette induction de la prolifération cellulaire n'est pas connu.

9.2 Les maladies anévrysmales et disséquantes

On dénombre une vingtaine de gènes aujourd'hui responsables de maladies anévrysmales et disséquantes. Ces gènes peuvent être responsables de formes syndromiques ou non syndromiques d'anévrysmes et de dissections artériels. Certaines de ces maladies ne touchent que l'aorte alors que d'autres sont responsables d'une atteinte diffuse de l'arbre artériel systémique, et parfois cérébral. Une classification (non exhaustive) basée sur la physiopathologie permet de distinguer les maladies de la matrice extracellulaire d'une part, et les maladies de la contractilité de la CML d'autre part.

9.2.1 Les maladies de la matrice extracellulaire

9.2.1.1 Atteinte structurelle de la MEC

9.2.1.1.1 Le syndrome d'Ehlers-Danlos vasculaire

Le syndrome d'Ehlers-Danlos vasculaire (SEDv) est une pathologie dominante rare (prévalence estimée 1/200 000), caractérisé par une fragilité excessive des artères de gros et moyen calibre, du tube digestif, et de l'utérus gravide. Le SEDv est classé comme une maladie artérielle syndromique compte-tenu d'un phénotype cutané, squelettique et morphotypique reconnaissable (Beighton et al., 1998).

Les patients présentent dès la seconde décennie des anévrysmes, dissections et ruptures artériels, et des ruptures digestives et utérines, responsables du pronostic sombre de la maladie, dont la médiane de survie est d'environ 50 ans (Pepin et al., 2000). Les localisations artérielles de ces lésions sont multiples (figure 21, Oderich et al., 2005).

Cette maladie est secondaire à un défaut du collagène de type III, collagène fibrillaire présent dans la paroi des artères et des organes creux. Les fibres de collagène III sont composées de fibrilles, elles-mêmes constituées d'un assemblage de triples hélices de monomères de chaînes α . Les mutations du gène *COL3A1*, codant la chaîne α du collagène de type III (il n'existe pas de chaîne β), ont un effet dominant négatif : elles sont responsables d'un défaut d'enroulement des triples hélices et génèrent des anomalies quantitatives et qualitatives des fibres de collagène. La structure hélicoidale de ce collagène fibrillaire est liée à la succession de triplets d'acides aminés débutant par une glycine. Les mutations des résidus glycine sont responsables de la forme la plus couramment rencontrée dans la pathologie. J'ai récemment caractérisé plusieurs classes de mutations responsables de formes cliniques de SEDv de sévérité variable (Frank, Albuissou et al., 2015) : les mutations haploinsuffisantes se révèlent plus tardivement que les formes liées aux mutations glycine, et ne sont pas responsables de complications digestives ou utérines. Elles sont cependant fréquemment responsables de complications aortiques, qui ne sont pas typiques de la pathologie. Une classe de mutations particulièrement sévères, liées à une perte d'un segment de collagène fibrillaire (saut d'exon) est

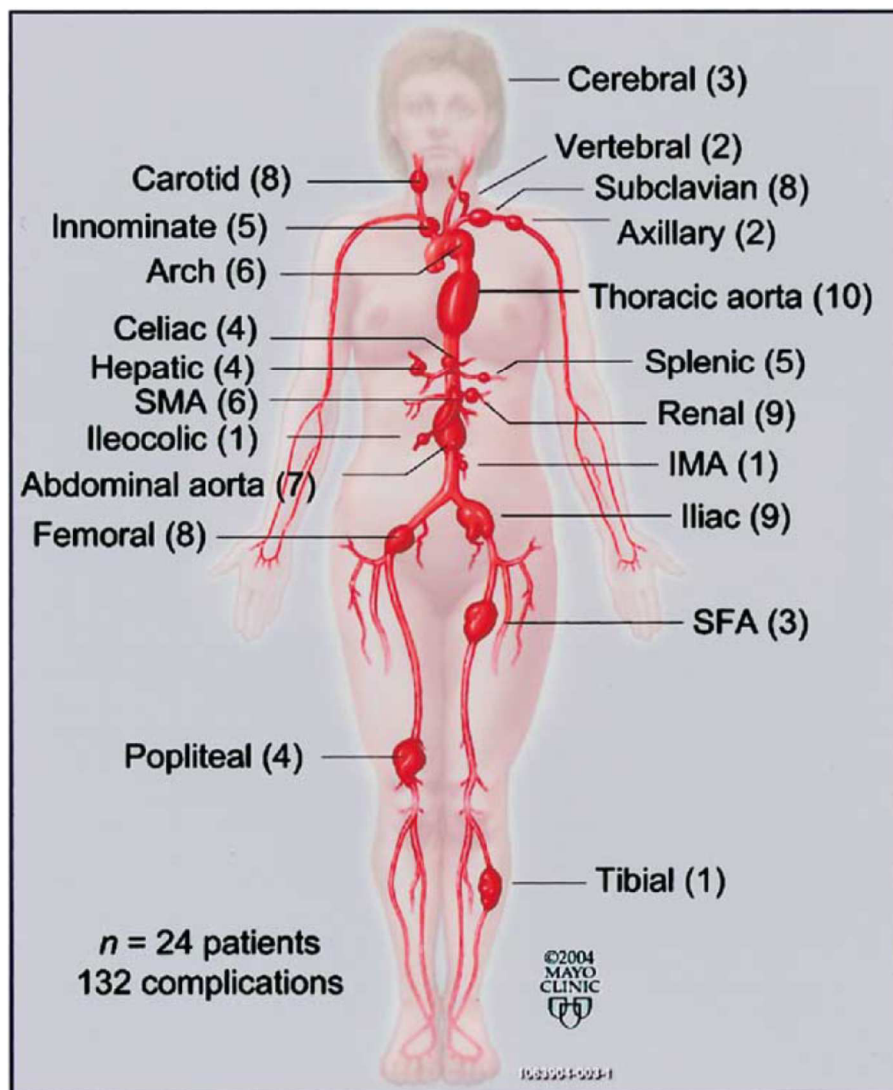


Figure 21: Topographie des lésions dans le SED vasculaire. Données issues d'une série de 132 cas, les chiffres entre parenthèses représentent le nombre d'occurrences dans cette série. D'après Oderich et al, 2005.

responsable de complications précoces survenant avant l'âge de 25 ans. L'article correspondant à ce travail clinico-génétique est joint dans l'annexe 2.

L'excessive fragilité artérielle des patients porteurs d'un SEDv contre-indique généralement tout geste interventionnel artériel. Cependant certains patients ont bénéficié d'une chirurgie à ciel ouvert, et parfois d'une étude autopsique extensive. Les prélèvements artériels identifiés dans la cohorte de 300 patients suivis par le CRMVR ont été collectés en une collection d'environ 10 cas et de sources variables (aorte, artère fémorale, rénale, tronc coeliaque). L'analyse histologique en collaboration avec le Pr Bruneval (Service d'Anatomo-pathologie de l'HEGP) montre que plusieurs des fragments artériels ont une atteinte de type fibroplasie médiale, généralement à proximité de sites de dissection artérielle (figure 22). Les atteintes aortiques sont *a contrario* aspécifiques, telles que décrites dans toutes les aortopathies dégénératives (voir section 9.2.2). Une étude plus complète du phénotype artériel, associant notamment une analyse en immunohistochimie de certains marqueurs de la voie du TGF β , est en cours et fait l'objet d'une publication en préparation.

9.2.1.1.2 Le syndrome HANAC et la porencéphalie héréditaire

Le syndrome HANAC, maladie très rare (prévalence estimée à 1/ 1 million) présente des similitudes physiopathologiques avec le SEDv : cette maladie syndromique associe anévrysmes intracrâniens (localisés majoritairement sur le siphon carotidien), néphropathie, crampes et angiopathie rétinienne. Il est causé par des mutations du gène *COL4A1*, codant la chaîne α du collagène de type 4. Ce collagène non fibrillaire est présent dans les membranes basales de nombreux tissus, et notamment des artères. Le collagène de type IV est une triple hélice constituée de 2 chaînes $\alpha 1$ et d'une chaîne $\alpha 2$. Les mutations de *COL4A1* sont aussi très généralement des glycines induisant un défaut de formation de la triple hélice, qui est séquestrée dans le reticulum endoplasmique des cellules qui le synthétisent (Weng et al., 2012).

La porencéphalie héréditaire est une maladie artérielle caractérisée par la présence de cavités intracérébrales circonscrites de taille variable, correspondant aux zones de perfusion des artères cérébrales. Cliniquement cette maladie se manifeste de diverses façons selon la localisation des zones cavitaires : une paralysie cérébrale hémiplegique, un déficit intellectuel et une épilepsie sont fréquemment observés. Cette maladie est la conséquence de mutations majoritairement situées sur les codons glycine des gènes *COL4A1* et *COL4A2* qui codent les chaînes $\alpha 1$ et $\alpha 2$ du collagène IV. Une étude récente sur un modèle murin KI a démontré que ces mutations induisent des anomalies de l'angiogénèse, et que la sévérité des hémorragies intracérébrales est corrélée à l'accumulation intracellulaire de collagène IV muté, suggérant que cette accumulation intracellulaire est un mécanisme moléculaire clé dans la fragilité artérielle (Jeanne et al., 2015).

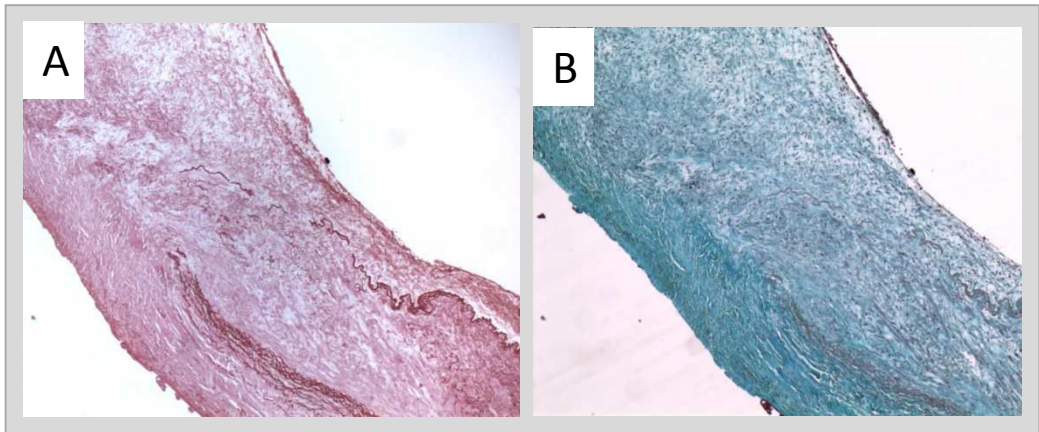


Figure 22: Lésion vasculaire du tronc coeliaque d'une patiente atteinte de SED vasculaire. A: coloration à l'orcéine montrant la perte de la LEI. B: coloration au trichrome vert montrant la désorganisation des CML de la media.

9.2.1.2 Trouble métabolique de la MEC : les maladies de la voie du TGFβ

La voie du TGFβ est une voie de signalisation impliquée dans de nombreux processus cellulaires comme la migration, la croissance et la différenciation cellulaire et l'apoptose. Elle joue un rôle majeur au cours du développement et dans l'homéostasie de nombreux tissus ; Le TGFβ est un immunorégulateur et un suppresseur de tumeurs impliqué dans de nombreux cancers (Massagué, 2008). Dans les tissus interstitiels, elle a un rôle profibrotique, induisant une synthèse accrue de protéines de la matrice extracellulaire comme le collagène de type I notamment.

Les ligands TGFβ sont une superfamille de ligands comportant notamment les BMP, les TGFβ, et les activines. Dans la voie dite « canonique » du TGFβ (figure 23), ils se lient à un récepteur de type II (TGFβR2), qui recrute et phosphoryle un récepteur de type I (TGFβR1). Ce dernier phosphoryle de nombreuses protéines SMAD (les R-Smads : SMAD1, SMAD2, SMAD3, SMAD5 et SMAD8). Une fois phosphorylé(s), le(s) R-Smad(s) s'associe(nt) à SMAD4. Ce complexe se transloque dans le noyau où il joue un rôle de cofacteur transcriptionnel. Il existe une voie dite « non canonique », *via* le récepteur à l'angiotensine II que le ligand TGFβ a la capacité d'activer, activant de nombreuses voies intracellulaires impliquant les SMADs, les voies RAS et MAPK entre autres (figure 23).

9.2.1.2.1 La maladie de Marfan

Le syndrome de Marfan est la première maladie artérielle identifiée impliquant la voie du TGFβ : cette maladie connue depuis 1972 (McCusick et al, 1972) associe des manifestations cardiovasculaires, musculo-squelettiques, ophtalmologiques et pulmonaires. Elle est la plus fréquente des maladies artérielles anévrysmales et disséquantes (prévalence 1/5 000). L'atteinte artérielle touche quasi-exclusivement l'aorte ascendante, sous la forme d'une dégénérescence de la média telle que décrite précédemment, associée à des anévrysmes, dissections et ruptures aortiques.

La maladie est liée à un défaut du gène *FBN1*, codant la fibrilline 1, protéine riche en cystéines constitutive de la MEC. Les mutations ponctuelles faux-sens représentent 2/3 des mutations du gène de la fibrilline-1 et correspondent généralement à une délétion ou substitution d'un résidu cystéine. Elles induisent ainsi une conformation incorrecte de la protéine et sa rétention dans le réticulum endoplasmique (Whiteman et Handford, 2003). Dans 1/3 des cas, les mutations sont responsables d'une haploinsuffisance, induisant chez certains patients un défaut purement quantitatif de la fibrilline-1. Certaines mutations responsables d'une anomalie de conformation de monomères de fibrilline induisent un effet dominant négatif (Milewicz et al., 1992) de façon relativement similaire au mécanisme dominant négatif dans le SEDv.

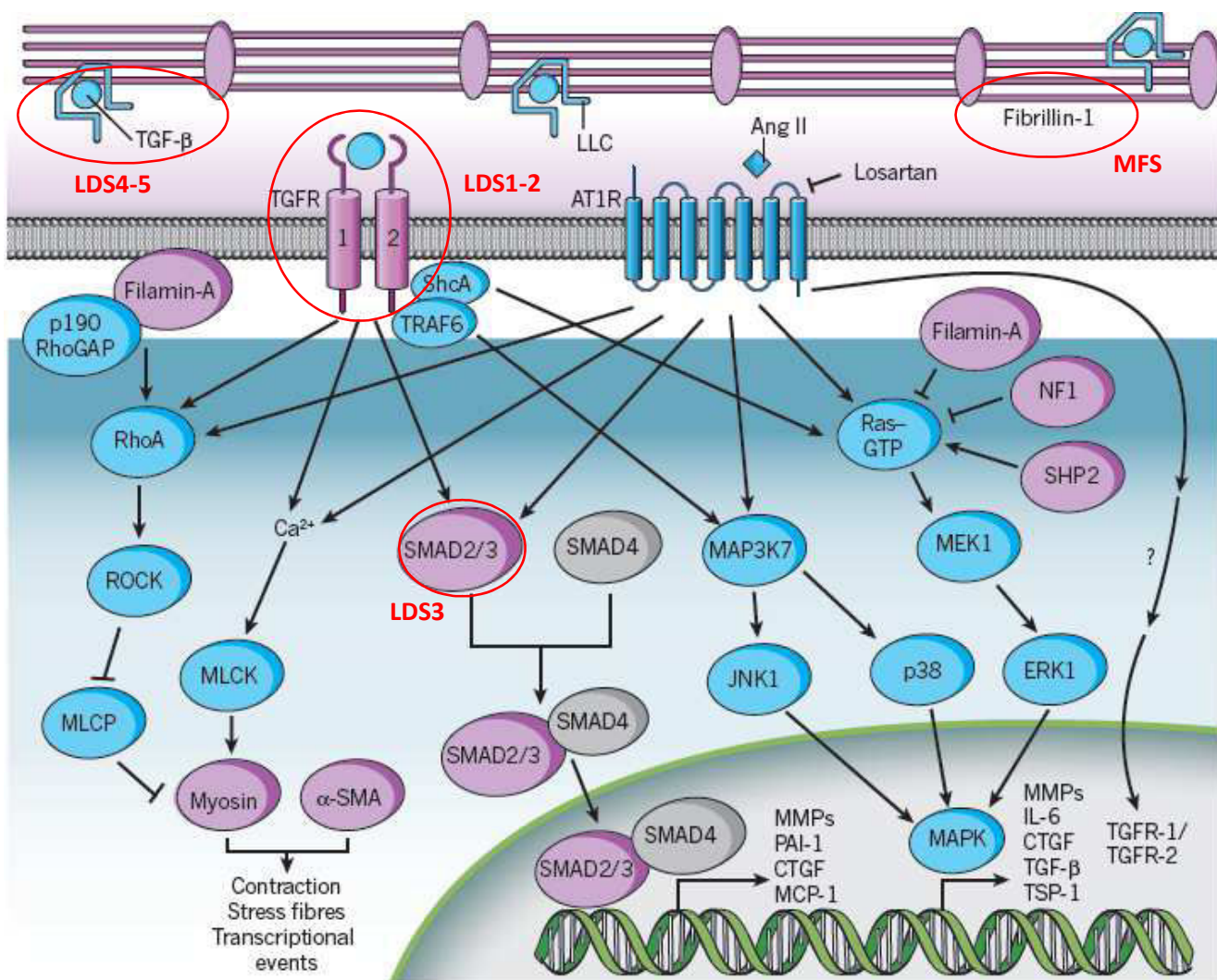


Figure 23: Voie du TGFβ dans la paroi artérielle. Les gènes /protéines entourés en rouge sont impliqués dans les pathologies du spectre de la maladie de Marfan. LDS: syndrome de Loeys-Dietz. MFS: maladie de Marfan. D'après Lindsay et Dietz, 2011.

Cependant les études récentes ont mis en lumière l'importance du dysfonctionnement de la voie du TGF β dans cette maladie (Kaartinen et Warburton, 2003; Neptune et al., 2003) : la fibrilline 1 constitue une réserve de TGF β , séquestrée sous forme d'un complexe protéique inactif appelé complexe latent. La fibrilline se lie au TGF β (sous forme latente) *via* un domaine appelé TB riche en cystéines. Ces dernières sont des sites fréquents de mutations dans la maladie, induisant un défaut d'interaction entre ces deux molécules, et une augmentation de la forme active du TGF β . Cette dérégulation de la voie du TGF β est considérée comme une activation, notamment à cause de l'effet bénéfique des traitements par sartans (bloqueurs de la voie non canonique du TGF β) chez la souris.

9.2.1.2.2 Les syndromes de Loeys-Dietz

Depuis 2004, cinq maladies artérielles anévrysmales et disséquantes impliquant la voie du TGF β ont été découvertes : les syndromes de Loeys- (LDS) de type 1 à 5 sont des maladies dites marfanoides car l'ensemble des signes (squelettiques, pulmonaires, vasculaires, cutanés...) présents dans la maladie de Marfan peuvent être retrouvés chez un patient porteur de syndrome de Loeys-Dietz, à l'exception de la subluxation du cristallin (MacCarrick et al., 2014).

On y retrouve cependant une atteinte vasculaire plus étendue impliquant, en plus de l'aorte, les artères moyennes de la circulation systémique et cérébrale, responsable d'une tortuosité plus ou moins marquée, d'anévrysmes, de dissections et de ruptures. Les patients présentent de plus une atteinte inconstante du système immunitaire, dépendante du gène muté dans cette voie.

Les LDS de type 1 et 2 sont liés à des mutations inactivatrices des récepteurs TGF β R1 et 2 (Loeys et al., 2005, 2006). Les patients présentent une forme syndromique avec dysmorphie faciale, lèvre bifide, et parfois une crâniosténose. L'atteinte aortique est souvent très agressive, responsable de dissections et de ruptures aortiques à un niveau de dilatation modérée. Sur le plan immunitaire ces patients présentent une allergie de type IV fréquente (asthme, eczéma, rhinites) pouvant parfois mimer un syndrome Hyper-IgE (Felgentreff et al., 2014).

le LDS de type 3 est lié aux mutations inactivatrices de *SMAD3* (van de Laar et al., 2011), facteur de transcription appartenant à la cascade intracellulaire de signalisation. Dans cette pathologie aussi appelée syndrome Anévrysmes-Ostéoarthrite, la composante immunitaire est parfois au premier plan, les patients pouvant présenter une ostéoarthrite précoce ou une polyarthrite rhumatoïde, une maladie de Charcot-Marie-Tooth, une thyroïdite de Hashimoto (Aubart et al., 2014; Berthet et al., 2015), une anémie de Biermer, un diabète auto-immun (données non publiées issues de patients suivis au CRMVR).

Le syndrome de Loey-Dietz de type 4 est secondaire à des mutations inactivatrices du gène *TGFβ2* (Boileau et al., 2012; Lindsay et al., 2012). Cette pathologie associe atteinte aortique et des artères intracrâniennes. Elle est considérée comme moins sévère que les types 1 à 3, notamment car l'atteinte aortique est peu évolutive (Ritelli et al., 2014). J'ai caractérisé au sein du CRMVR les phénotypes vasculaires de patients d'une grande famille avec forme sévère de maladie aortique associée à des lésions des artères systémiques de moyen calibre. Cette caractérisation étend le spectre phénotypique du LDS4 qui doit être considéré comme aussi sévère que les types 1 à 3 (Mazzella et al, soumis dans l'EJHG).

Enfin en 2015, le syndrome de Loey-Dietz de type 5 lié au gène *TGFβ3* a récemment été identifié (Bertoli-Avella et al., 2015). L'atteinte aortique est associée de façon inconstante à une fente palatine ou une lèvre bifide, un pied-bot, et une croissance squelettique excessive.

Tous ces syndromes ont la particularité d'être liés à des mutations inactivatrices des gènes impliqués (avec une fréquence élevée de mutations d'haploinsuffisances notamment), cependant toutes les études s'y rapportant montrent une hyperactivation paradoxale de la voie du TGFβ. Ce phénomène est appelé paradoxe des vasculopathies liées au TGFβ (Akhurst, 2012) : dans l'aorte de patients atteints de LDS sont observés des taux élevés de SMAD3 et SMAD2 sous forme phosphorylée, c'est-à-dire active (Boileau et al., 2012; Lindsay et al., 2012). Cette activation paradoxale est aussi observée in vivo chez la souris *Tgfβ2*^{-/-}, bien que d'apparition tardive (après 4 mois, Lindsay et al., 2012).

9.2.1.2.3 Le syndrome de tortuosité artérielle et le cutis laxa

autosomique récessif de type 1b

Le syndrome de tortuosité artérielle (ATS) est une maladie récessive liée aux mutations inactivatrices du gène *GLUT10* (ou *SLC2A10*) transporteur non exclusif du glucose (Coucke et al., 2006). Cette maladie de découverte généralement néonatale est caractérisée par une atteinte des grosses et moyennes artères systémiques, associant sténoses, anévrysmes et tortuosité majeure. Des hernies pariétales et diaphragmatiques, un cutis laxa, une dysmorphie faciale et squelettiques sont aussi présents (Callewaert et al., 2008). Au sein du CRMVR j'ai formalisé la prise en charge génétique de cette affection et participé à la caractérisation d'un cas atypique avec tortuosité veineuse associée (Albuisson et al., 2015; Moceri et al., 2013, articles en annexe)

La fonction de GLUT10 dans le métabolisme de la MEC est inconnue, mais les caractéristiques cliniques de ces patients permettent de considérer cette affection comme appartenant aux maladies de la matrice extracellulaire.

Le cutis laxa autosomique récessif de type 1b est une maladie très proche de l'ATS, bien que généralement plus sévère et précoce (Huchtagowder et al., 2006). Il est lié aux mutations du gène *EFEMP2* codant la fibuline 4, protéine matricielle impliquée dans la maturation des fibres élastiques par le recrutement d'une enzyme permettant leur polymérisation (Horiguchi et al., 2009).

Dans ces deux maladies, les sténoses artérielles siègent majoritairement sur les artères pulmonaires et sont marquées par un épaississement médial majeur et une prolifération des cellules musculaires lisses objectivées par la présence en immunohistochimie d'un marqueur de prolifération appelé ki-67 (données non publiées). Une activation de la voie du TGF β a aussi été suggérée dans ces deux cas (Coucke et al., 2006; Renard et al., 2010).

D'une façon générale, on peut observer que les **maladies de la matrice extracellulaires** responsables d'une atteinte artérielles sont liées à un **défaut structurel fragilisant la paroi de l'artère** (comme les collagénopathies) ou à une **perturbation du métabolisme de la MEC** caractérisées par une dérégulation de la voie du TGF β par un mécanisme encore incompris.

9.2.2 Les maladies de la contraction de la CML

Ce groupe de maladies est caractérisé par une aortopathie non syndromique au premier plan, associant anévrysmes et dissections de l'aorte ascendante (TAAD), et caractéristiques histopathologiques communes (dégénérescence kystique de la média, figure 24). Elles ont toutes en commun une altération (i) d'une protéine de l'appareil contractile de la CML, actine et myosine spécifiques du muscle lisse codées par *ACTA2* et *MYH11* (Guo et al., 2007; Zhu et al., 2006) (ii) ou d'une enzyme impliquée dans la phosphorylation de ces protéines : *MYLK*, *PRKG1* (figure 19, Guo et al., 2013; Wang et al., 2010). Dans certains cas l'atteinte aortique peut être accompagnée d'anomalies cardiovasculaires comme une persistance du canal artériel (mutations de *MYH11*), ou une CIV, et un livero racemosa (mutations d'*ACTA2*). La physiopathologie de ces affections est encore inconnue, aucun lien n'ayant été établie entre dysfonctionnement de l'appareil contractile des CML et mort cellulaire.

9.2.3 Les TAAD de physiopathologie inconnue

Enfin très récemment le gène *MAT2A* a été identifié dans une famille avec forme dominante de TAA associé à une prévalence élevée de bicuspidie aortique (Guo et al., 2015). La pénétrance de la mutation identifiée pour le TAA semble faible (environ 40% chez l'adulte), et très faible pour la BAO (environ 20%). *MAT2A* code la methionine adenosyltransferase II alpha (MAT II α). Bien qu'une étude en modèle animal ait confirmé l'implication de ce gène dans le développement cardiovasculaire du poisson-zèbre, son rôle dans la physiologie et le développement vasculaire est inconnu.

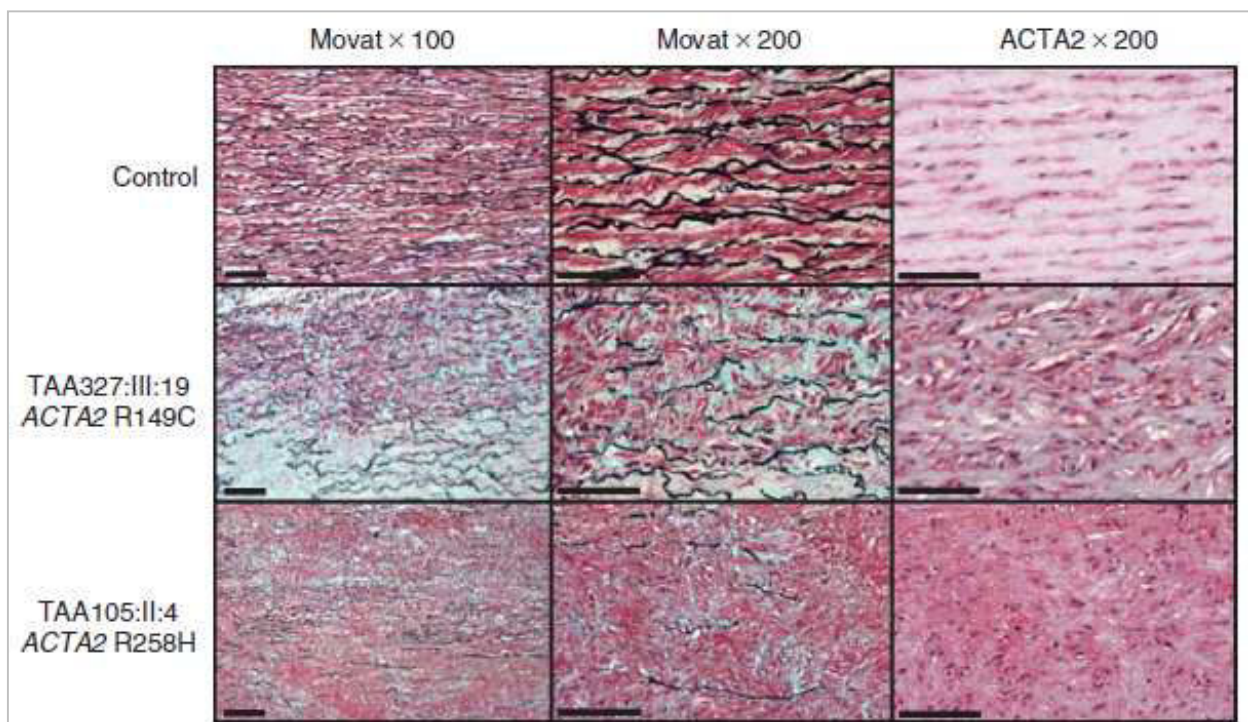


Figure 24: Dégénérescence kystique typique des aortopathies dégénératives: la média est le siège d'une perte en CML remplacée par du tissu hyalin (protéoglycanes), les fibres élastiques sont peu nombreuses et fragmentées. Coloration HES. D'après Guo et al, 2007.

10. Etude génétique de la DFM

10.1 Présentation des familles étudiées

Au sein du CRMVR, nous avons recruté 10 familles d'au moins deux cas atteints de DFM, confirmés par une imagerie de type angioscanner ou angiographie. Certaines de ces familles ont été précédemment publiées et ont fait l'objet d'une étude en échotracking, et d'une première analyse d'exome (Kiando et al., 2015; Perdu et al., 2007) ne permettant pas de dégager de gène muté en commun. Dans la majorité des cas, les patients sont atteints d'une DFM multifocale des artères rénales, plus rarement des artères cervicales. Les arbres généalogiques de ces familles sont représentés dans la figure 25. Plusieurs cas de DFM sévère d'apparition précoce, et sporadiques, ont été par ailleurs identifiés et ont fait l'objet d'investigations poussées au niveau imagerie (figure 26).

10.2 Sélection des individus faisant l'objet d'un séquençage d'exome

Un séquençage d'exome a été réalisé chez un total de 30 individus dont 26 atteints et 2 couples de parents non atteints, par un prestataire externe (Integragen). Une majorité d'atteints ont été séquençés dans la mesure où la pénétrance estimée de la maladie semble faible et que le statut génétique des apparentés non symptomatiques est de fait incertain. De plus, la réalisation d'un angioscanner systématique chez les apparentés sains de patients se justifie difficilement au vu du coût et de l'irradiation induite par cet examen. Les apparentés sains n'ont donc pas bénéficié d'explorations radiologiques fiables pour déterminer leur statut vasculaire. Comme précisé dans la section 7. Génétique de la DFM, une majorité de cas atteints sont des femmes (environ 75%), impliquant une plus faible pénétrance de la mutation causale chez les apparentés de sexe masculin dans ces familles.

Le séquençage d'exome a donc porté sur 8 des 10 familles suivies au sein du CRMVR (figure 25). Les deux familles restantes sont actuellement en cours de séquençage et feront l'objet d'une analyse conjointe ultérieure. Le séquençage a donc porté sur 5 familles de 2 cas (FDP11, FDP18, FDP43, FDP59, FDP143), 2 familles de 3 cas (FDP8, FDP125), 1 famille de 5 cas d'origine malgache (SED294), trois cas sporadiques seuls (FDP37, FDP412, FDP530), et deux cas sporadiques analysés en trio avec les parents non atteints (FDP200 et FDP249).

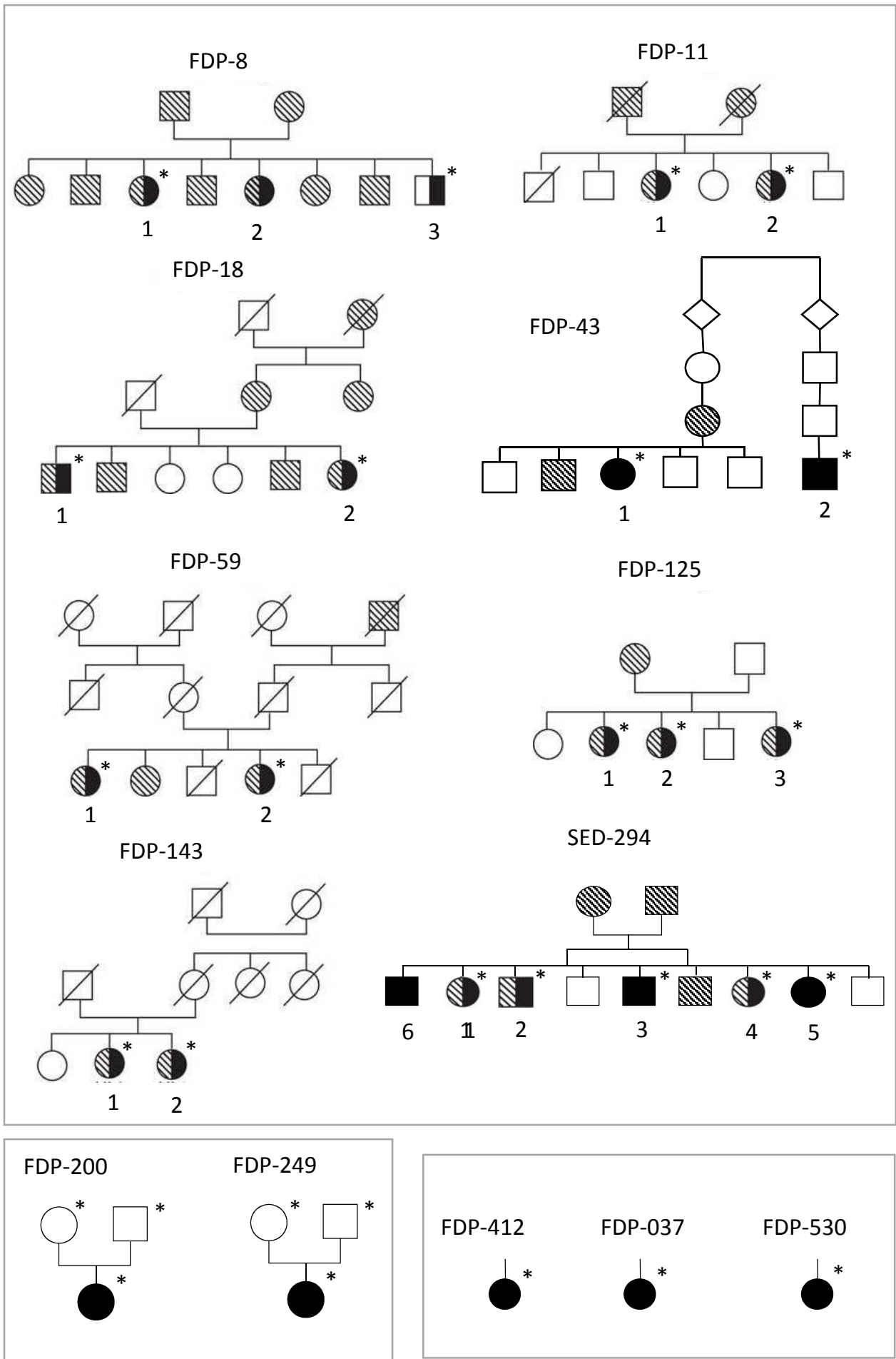


Figure 25: Structure des 8 familles et des 5 cas sporadiques atteints de DFM analysés par séquençage d'exome. Les astérisques désignent les individus séquencés. Noir: DFM. Hachuré: HTA.

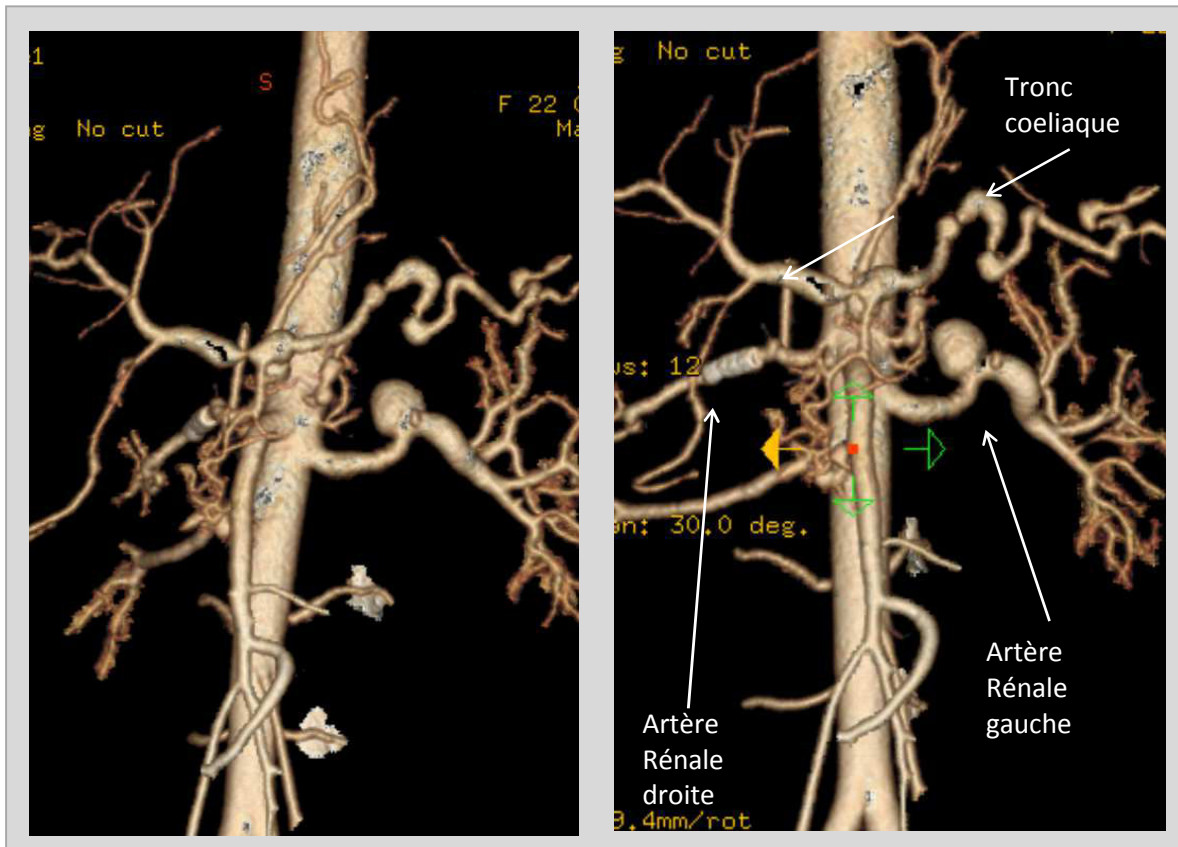


Figure 26: Lésions artérielles rénales et du tronc coeliaque du cas FDP249001. on note des lésions sténosantes et anévrysmales, une forme très diffuse de DFM multifocale ne se présentant pas en pile d'assiettes.

10.3 Séquençage d'exome : données brutes et application des filtres de base

L'analyse d'exome a été réalisée avec le même outil bioinformatique que pour l'étude génétique de l'AAA : le logiciel en ligne Polyweb, développé par l'équipe de P. Nitschké à l'institut Imagine (Université Paris Descartes). Cet outil permet l'application des filtres de base comme l'exclusion de variants fréquents, ou prédits bénins par les outils SIFT et Polyphen. Il permet de plus d'appliquer un modèle génétique (dominant, de novo, récessif homozygote ou hétérozygote composite) et de croiser les données familiales afin de déterminer si des mutations sont présentes dans un gène commun à plusieurs familles. En premier lieu, les variants fréquents, et prédits bénins par SIFT et par Polyphen ont été éliminés en conservant les variants dont les prédictions sont divergentes.

Les données brutes de séquençage ainsi que le croisement des résultats sont réunis dans le tableau 1.

L'analyse de ces données permet de distinguer au sein de ces cas un groupe d'individus caucasiens et un groupe d'individus d'origine non caucasienne, en partie africaine (patients FDP037, FDP43, SED294, et FDP53). Un plus grand nombre de variants est retrouvé dans ce dernier groupe, ce qui est concordant avec les données présentées dans le chapitre 1, section 4.3.1 Résultats bruts d'une analyse d'exome.

10.4 Stratégie d'analyse des données

L'analyse d'un grand nombre de cas issus de multiples familles dans le cas d'une maladie génétiquement hétérogène est complexe, et nécessite le développement d'une stratégie robuste établie à l'avance. L'analyse préliminaire des données de séquençage d'exome de 6 des 8 familles du CRMVR, réalisée par notre équipe de recherche, a démontré la grande hétérogénéité de cette affection (Kiando et al., 2015) et la nécessité de recourir à des stratégies alternatives de traitement de données multifamiliales.

L'analyse de ces 30 exomes a été réalisée en étapes successives, en terminant par la recherche de voies biologiques communes à plusieurs familles (figure 27) :

Etape 1 : recherche ciblée de mutations dans les gènes de maladies sténosantes, anévrysmales et disséquantes afin d'identifier un éventuel diagnostic différentiel.

Etape 2 : recherche interfamiliale (gènes communs à deux familles ou plus) en modèles dominant et récessif.

Etape 3 : recherche famille par famille de gènes candidats de protéines impliquées dans les voies développées dans les sections 8. et 9. Hypothèses physiopathologiques et Diagnostics différentiels.

Etape 4 : recherche de voies communes par l'utilisation des outils développés dans le chapitre 1, section outils de priorisation de gènes candidats.

Individus	Nombre de variants						gènes présentant des variations compatibles avec le mode de transmission familial	gènes présentant des variations perte de fonction compatibles avec le mode de transmission familial
	Total	MAF<1%	Faux-sens, épissage, et perte de fonction	Faux-sens prédits pathogènes par Polyphen OU SIFT (pour les faux-sens uniquement)	variants perte de fonction			
FDP8-1	20457	4660	561	242	6	14	2	
FDP8-2	20122	4697	574	256	4			
FDP8-4	20383	4970	612	235	7			
FDP11-1	20017	4050	474	249	8	48	7	
FDP11-2	20366	4130	412	252	10			
FDP18-1	20664	4136	552	256	9	34	4	
FDP18-4	20753	4081	487	245	6			
FDP37-1	28597	6230	702	307	14	NA	NA	
FDP43-1	20437	6025	680	292	11	21	1	
FDP43-2	20565	6450	665	307	9			
FDP59-1	20766	4350	465	251	9	47	4	
FDP59-2	20811	4373	543	246	7			
FDP125-1	20698	4814	617	249	4	15	1	
FDP125-2	20580	4884	628	227	6			
FDP125-3	20657	4710	607	2	5			
FDP143-1	20469	4345	538	229	3	44	3	
FDP143-2	20608	4449	545	231	5			
FDP200-1	22183	4201	486	225	7	1	0	
FDP200-10	21571	4472	530	237	4			
FDP200-11	21822	4390	526	241	6			
FDP249-1	21654	4841	562	252	5	1	0	
FDP249-10	21763	4729	549	234	8			
FDP249-11	21478	4654	519	248	3			
FDP412-1	21903	4012	567	230	6	NA	NA	
FDP530-1	22704	6040	705	301	7	NA	NA	
SED294-1	22643	6234	712	305	10	34	5	
SED294-2	23538	6054	694	287	9			
SED294-3	24161	6128	731	279	7			
SED294-4	23965	6321	729	298	8			
SED294-5	23592	5932	687	304	6			

Tableau 1: Résultats bruts et après application des filtres de base des données de séquençage d'exome des 30 individus, cas et apparentés.



Loeys-Dietz, Marfan, EDS, HANAC, Moya-Moya...

FDP412:
SMAD3

ANALYSE INTERFAMILIALE

GENES COMMUNS ?

OUI

THADA : FDP59+SED294
DOK7: FDP59+FDP11

PERTINENCE BIOLOGIQUE?

NON

ANALYSE INTRAFAMILIALE (EXCLUSION DES CAS ISOLEES)

TRIOS
SED200 SED249

VOIES KEGG
RAS
MAPK
NOTCH
HVS
TGFB...

PHENOTYPES MURINS

MUTATIONS PERTE DE FONCTION

NEO-MUTATION?
NON

RECESSIF?
HMZ?
NON

HTZ
COMPOSITE?

OUI

Aucun candidat

27 candidats (~3 par famille)

PRIORISATION

FDP200: *OBSCN*
FDP249: *AHNAK1*

>40 candidats biologiquement pertinents

0 à 3 candidats biologiquement pertinents par famille

Figure 27: Stratégie d'analyse des données des 30 séquençages d'exomes pour la DFM.

Afin de déterminer si un gène pouvait être candidat lorsque sa fonction vasculaire m'était inconnue, une recherche bibliographique combinée (utilisant les opérateurs booléens) dans Pubmed a été réalisée de la façon suivante : (endothel* OR vascul* OR arter* OR « smooth muscle cell » OR « extracellular matrix ») et # nom du gène #.

L'ensemble des variants identifiés par ces approches et cité ci-dessous ont fait l'objet d'une confirmation par séquençage Sanger.

10.4.1 Recherche ciblée des gènes de maladies sténosantes, anévrysmales et disséquantes

L'ensemble des gènes de maladies développées dans la section 9 ont été recherchés (selon un modèle dominant ou récessif selon le mode de transmission connu de la maladie). Cette analyse a été réalisée famille par famille et individu par individu dans les cas isolés.

10.4.1.1 Identification d'une mutation perte de fonction de *SMAD3*

Cette analyse a permis d'identifier une mutation perte de fonction non décrite (et absente des bases de données en population) de *SMAD3* chez le cas FDP412 (c.359_360insT p.Thr120*). *A posteriori*, l'analyse clinico-radiologique de cette patiente révèle qu'elle est porteuse d'une maladie diffuse, plus anévrysmale que sténosante (figure 28), qui est concordante avec ce résultat. Cependant, l'étude des cas décrits dans la littérature révèle des formes moins diffuses que celle présentée par la patiente, bien que ressemblants par l'aspect monoliforme des lésions artérielles. Il est de ce fait probable que d'autres cas classés DFM, notamment dans les cohortes françaises et américaines, présentent en réalité un syndrome de Loey-Dietz atypique, diffus, et associant sténoses et anévrysmes multiples. Cette patiente présente de plus un trouble bipolaire et une gastrite antrale, symptômes difficiles à interpréter dans ce contexte. Compte tenu de la richesse des symptômes associés aux mutations *SMAD3*, il n'est pas exclu que ces signes fassent partie du spectre de la maladie. Les données d'exome de cet échantillon ont été exclues des analyses ultérieures.

10.4.1.2 Identification d'une mutation pathogène de *COL4A2*

Cette analyse a par ailleurs permis d'identifier chez 2 cas parmi 3 de la famille FDP125 une mutation glycine de *COL4A2* (c.1312G>A p.Gly438Arg). Cette anomalie est non décrite, absente des bases de données en population, et n'a jamais été identifiée au laboratoire de génétique du CERVCO réalisant le diagnostic génétique des maladies liées à *COL4A2*. Elle siège dans la zone codant la triple hélice, et correspond sur le plan physiopathologique aux mutations décrites dans la porencéphalie héréditaire. Elle est donc considérée comme pathogène. Elle est portée par les individus 1 et 3 de la famille FDP125 : ces patients ne présentent pas de signes de porencéphalie ni du syndrome HANAC à ma connaissance, et sont atteints d'une dysplasie localisée sur les artères rénales. L'individu non porteur de cette mutation (FDP125-2) a bénéficié d'une

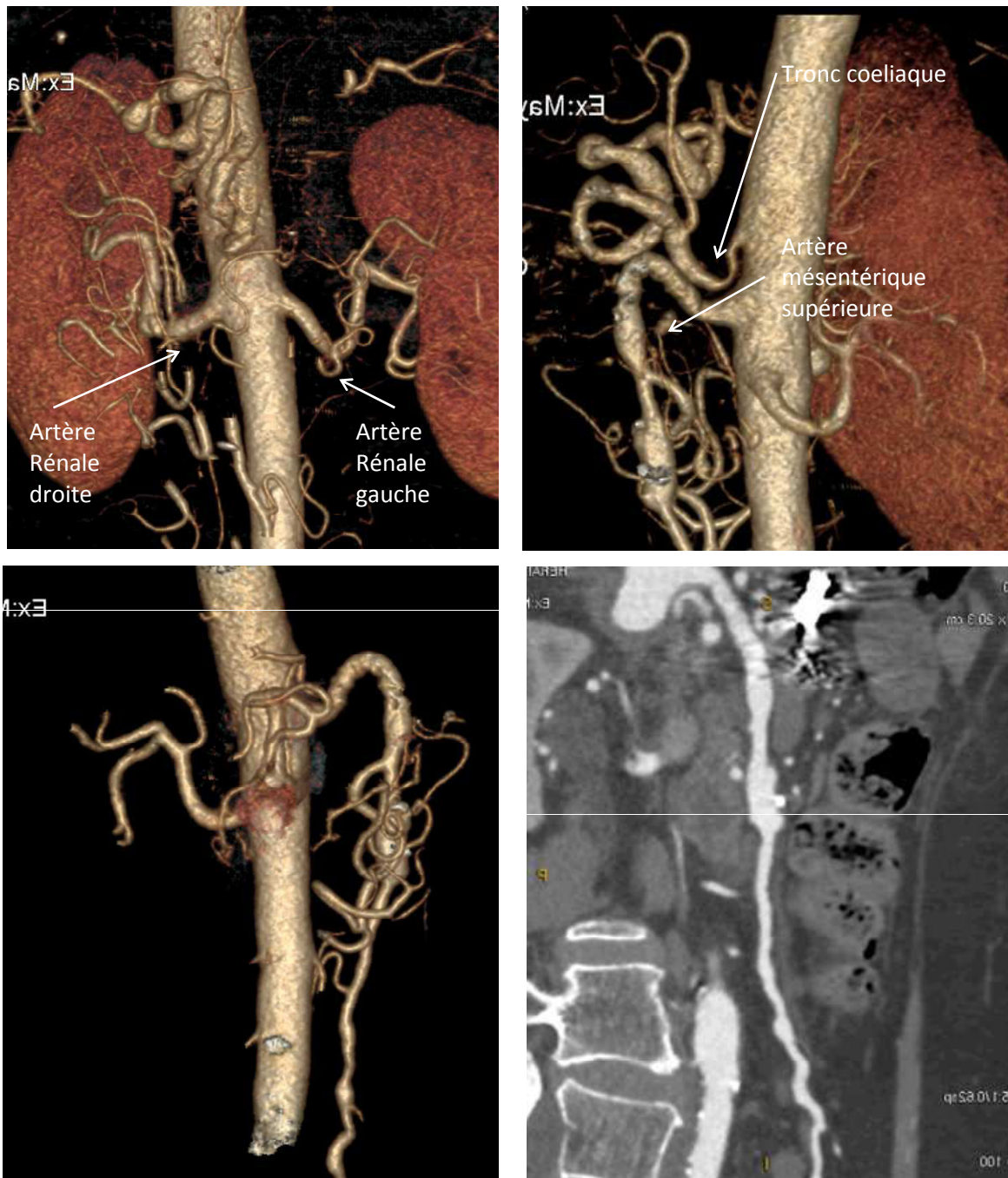


Figure 28: analyse par angioscanner avec et sans reconstruction, des artères digestives de la patiente FDP412. On Note une sténose importante de la naissance du tronc coeliaque, une tortuosité et un aspect monoliforme de l'artère mésentérique supérieure pouvant évoquer une dysplasie artérielle, bien que l'aspect général soit plus en faveur d'une maladie anévrysmale.

imagerie vasculaire cérébrale et n'est pas porteur d'anévrismes intracrâniens. Enfin, un individu de cette famille, porteur d'une sténose d'une artère rénale de cause inconnue, n'est pas non plus porteur de l'anomalie de *COL4A2*. Il est donc probable que cette découverte soit fortuite, et une exploration des cas de cette famille à la recherche d'une atteinte rénale (par échographie) et d'une tortuosité rétinienne (par examen du fond d'œil) va être réalisée très prochainement.

10.4.1.3 Identification d'une mutation de *COL4A2* de pathogénicité inconnue

J'ai identifié une mutation siégeant sur une glycine hors de la zone codant la triple hélice (c.4981G>A p.Gly1661Ser) dans la famille SED294, famille d'origine malgache porteuse d'une forme relativement diffuse de la maladie (DFM multisites). Cette variation est à l'état homozygote chez quatre des 5 cas séquencés, et hétérozygote chez le cas SED294-5. Dans cette famille, certains cas atteints de DFM présentent aussi un phénotype de fragilité aortique associé à une HTA sévère : les individus 1 et 5 ont présenté une dissection aortique, ainsi qu'un sixième individu non disponible pour une analyse génétique.

Cette variation siège dans le domaine NC1 de *COL4A2* (figure 29, Kalluri, 2003), domaine globulaire qui a deux fonctions distinctes : il participe à l'initiation de l'enroulement de la triple hélice (hétérotrimère constitué de 2 chaîne $\alpha 1$ codées par *COL4A1* et d'une chaîne $\alpha 2$ codée par *COL4A2*). Il est d'autre part libéré par clivage protéolytique, et porte alors le nom de canstatine, molécule aux propriétés anti-angiogéniques et anti-tumorales. Il inhibe la prolifération et la migration des cellules endothéliales, et a des propriétés proapoptiques par réduction du potentiel de membrane mitochondriale (Kamphaus et al., 2000).

Le statut homozygote de 4 des 5 membres de cette famille est probablement lié à la structure de la population malgache, qui correspond sur le plan génétique à un isolat géographique. Pour vérifier s'il existait une consanguinité méconnue dans cette famille, une analyse de cartographie d'homozygotie à partir des données d'exome par l'outil Homozygosity Mapper (Seelow et al., 2009) a été réalisée : La région d'homozygotie commune incluant *COL4A2* fait environ 2 Mb. Il s'agit de la plus grande région homozygote dans cette famille, correspondant à un profil de population isolée et non à une consanguinité vraie (Kirin et al., 2010). De plus, compte-tenu de l'origine géographique des patients, la fréquence de ce variant est difficile à estimer. La population malgache est issue de deux origines distinctes : océanienne et africaine (Poetsch et al., 2013). Les données issues des bases comme EXAC et EVS ne sont pas représentatives de cette population, et il n'est pas exclu que ce variant soit fréquent dans cet isolat. Il est donc nécessaire dans un premier temps de déterminer sa fréquence dans une population contrôle correctement appariée. Pour ce faire j'ai contacté plusieurs équipes ayant travaillé à la caractérisation de cette population (Hewitt et al., 1996; Poetsch et al., 2013; Tofanelli et al., 2009) qui disposent d'un échantillon cumulé d'environ 200 ADNs d'individus malgaches. Le variant p.Gly1661Ser va être prochainement recherché dans cette population par

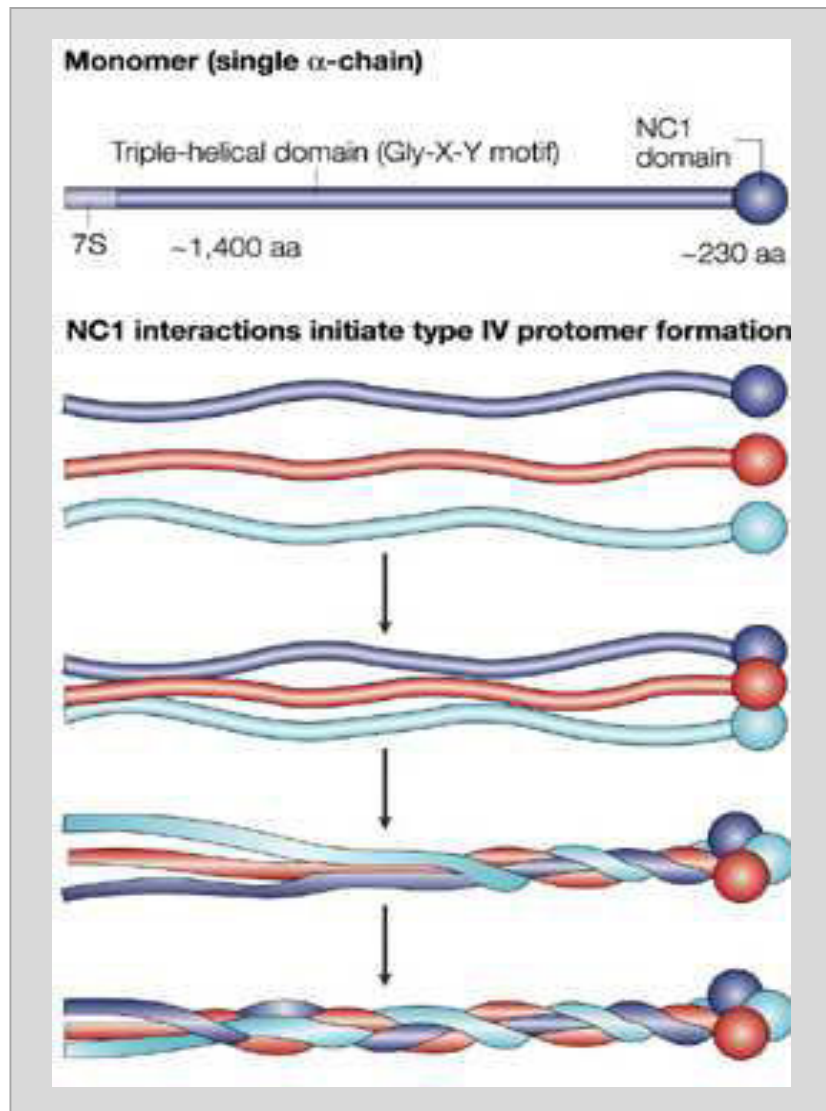


Figure 29: Structure des collagènes IV et rôle de leur domaine globulaire NC1 dans l'enroulement des hétérotrimères. D'après Kalluri et al, 2003.

séquençage direct dans le cadre de collaborations avec ces équipes. Cependant l'échantillon correspondant reste de faible taille et l'absence du variant parmi 200 contrôles ne permettrait pas, dans l'absolu, d'apporter d'argument majeur quant à sa causalité dans la pathologie familiale.

Enfin, dans le cadre de la prise en charge de cette famille dans le CRMVR, les descendants de ces 5 cas sont en cours d'exploration échographique et d'inclusion dans un protocole de recherche génétique : la ségrégation du variant avec une éventuelle DFM à la génération suivante est en cours d'analyse. Cependant, le statut génétique (homozygote et hétérozygote) du variant dans la famille ne permet pas d'interpréter le mode de transmission éventuel de la DFM si elle y était associée, et rend difficile l'analyse de résultats à la génération suivante.

Ces deux dernières familles dans lesquelles des anomalies de *COL4A2* ont été identifiées n'ont pas été exclues de l'analyse subséquente pour les raisons développées ci-dessus.

10.4.2 Analyse interfamiliale

Par stratégie de croisement telle que définie dans le chapitre 1, section 4.3.2.4. la recherche d'un gène commun à au moins 2 familles ou cas isolés (en modèle dominant, récessif, ou mixte) a été effectuée.

Comme l'indique le tableau 2, aucun gène n'a été identifié dans plus de 2 cas indépendants. Deux gènes, partagés par deux familles, portent des variations faux-sens hétérozygotes :

Le gène *THADA*

Il code la protéine associée à l'adénome thyroïdien. Ce gène est partagé par les familles FDP59 (portant la variation c.320C>G p.Pro107Arg inconnue des bases de données et dont les prédictions de pathogénicité divergent), et SED294 (portant la variation c.2998A>G p.Thr1000Ala, aussi inconnue des bases de données et dont les prédictions divergent). Cette protéine est de fonction inconnue, et les seules données disponibles semblent l'impliquer dans le syndrome des ovaires polykystiques par étude GWAS (Ha et al., 2015). Il n'existe pas de modèle murin d'inactivation de ce gène, mais il semble exprimé dans les artères (source GTEX portal). Devant le peu de données disponibles concernant ce gène, ce candidat n'a pas été retenu dans un premier temps.

Le gène *DOK7*

Ce gène code la protéine docking 7, et est partagé entre la famille FDP059 (portant le variant c.988G>A p.ValV330Ile dont les prédictions divergent mais absent des bases de données) et la famille FDP11 (portant le variant c.1487G>C p.Gly496Ala, prédit délétère et présent à une fréquence de 5/100 000). La protéine

symbole	nom	expression vasculaire	Bibliographie appareil vasculaire	PHENOTYPE murin	EXAC	FAMILLE 1	c.	p.	Scores de prédiction		FAMILLE 2	c.	p.	Scores de prédiction		EXAC (africains)
									POLYPH EN	SIFT				POLYPH EN	SIFT	
THADA	thyroid adenoma associated [Source:HGNC Symbol;Acc:19217]	+	non	pas de phénotype CV	1/100000	FDP059(2)	c.320C>G	p.Pro107Arg	0,878	0,15	SED294(5)	c.2998A>G	p.Thr1000Ala	0,619	0,1	0
DOK7	docking protein 7	+	non	pas de phénotype CV	1/100000	FDP059(2)	c.988G>A	p.ValN330Ile	0,99	0,82	FDP011(2)	c.1487G>C	p.Gly496Ala	0,996	0	0
MTMR7	myotubularin related protein 7	+	non	NA	1/100000	SED294(5)	c.571C>T	p.LeuL191Phe	0,844	-	FDP037(1)	c.1925C>T	p.Pro642Leu	0,048	0	1/100000 (1/100000)
ARAP1	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1	+	oui	NA	5/1000000	FDP011(2)	c.2600G>A	p.Arg867His	0,978	0,02	FDP037(1)	c.1124C>T	p.AlaA375Val	0,278	0	5/100000 (5%)
FGD5	FYVE, RhoGEF and PH domain containing 5	+	oui	NA	5/1000000	FDP018(2)	c.3233C>T	T1078I	0,885	0,01	FDP037(1)	c.3247G>A	p.Asp1083Asn	0,406	0,1	5/100000 (6%)
ADAMTS14	ADAM metalloproteinase with thrombospondin type 1 motif, 14	+	oui	NA	5/1000000	FDP008(3)	c.1136delT	p.Leu379Argfs*17			FDP037(1)	c.976G>A	p.Gly326Arg	0,996	-	1/100000 (1%)
HYDIN	HYDIN, axonemal central pair apparatus protein	-	non	phénotype CV: hémorragie	5/1000000	FDP018(2)	c.9548G>A	C3183Y	0,919	1	FDP037(1)	c.866A>G	p.TyrY289Cys	0,994	0	0

Tableau 2: Données issues d'une recherche croisée interfamiliale. Deux gènes sont communs à 2 familles chacun, mais ne sont pas des candidats connus. 5 gènes sont communs au cas FDP37 et à 1 parmi les 8 familles analysées. Parmi eux, trois gènes sont des candidats potentiels.

DOK7 est impliquée dans la synaptogénèse neuromusculaire : elle active le récepteur à tyrosine kinase spécifique du muscle strié (appelé MUSK), impliqué dans la différenciation post-synaptique, et le regroupement de récepteur de l'acétylcholine dans les myotubes nécessaire à la maturation de la synapse cholinergique. Des mutations inactivatrices de *DOK7* sont responsables de la myasthénie congénitale de type 10, maladie récessive. Aucune donnée de la littérature n'existe sur son implication vasculaire. *DOK7* n'a pas été retenu comme candidat potentiel.

Le tableau 2 révèle par ailleurs la présence de 3 gènes communs entre une famille et 1 cas sporadique, toujours le même : l'individu FDP037. Bien que son origine familiale ne le laisse pas supposer, cette patiente présente 1,5 fois plus de variants que les individus caucasiens (tableau 1), laissant suggérer qu'elle ait une origine africaine. L'identification dans son exome de variants africains rares et fréquents confirme cette hypothèse.

Ces trois variants correspondent à des gènes candidats intéressants au vu de leur fonction biologique, mais on peut noter que la fréquence en population appariée (d'origine africaine) est environ 10 fois plus élevée que la fréquence cumulée dans EXAC (tableau 2). L'analyse bibliographique montre un rôle vasculaire potentiel pour chacun d'entre eux.

Le gène *ARAP1*

ARAP1 (pour ArfGAP Avec domaines RhoGAP, Ankyrin et PH1) est un intermédiaire entre la voie RAS et la protéine RhoA. C'est une GTPase impliquée dans le remodelage du cytosquelette d'actine par régulation des membres de la famille ARF et RHO. Cette protéine est fortement exprimée dans le système vasculaire rénal ; Elle y est associée au récepteur de type 1 à l'angiotensine (*AGTR1*), présent sur les cellules musculaires lisses vasculaires et responsable d'une vasoconstriction après stimulation par l'angiotensine II. *ARAP1* pourrait avoir un rôle de recyclage local des récepteurs de type AT1 (Guo et al., 2003), et donc de modulation la sensibilité à l'angiotensine dans le système vasculaire, notamment rénal (Doblinger et al., 2012). Un modèle murin KO existe mais dont le phénotype n'est pas décrit.

Le gène *FGD5*

Ce gène code un facteur d'échange GTP/GDP ciblant les protéines de la famille Rho. C'est donc un régulateur de Rho GTPases, qui est exprimé dans les cellules endothéliales où il régule l'action pro-angiogénique du VEGF, et notamment la communication intercellulaire, la perméabilité, la migration, et la prolifération endothéliale. Le modèle murin KO de *Fgd5* est embryolétal (Gazit et al., 2014).

Le gène *ADAMTS14*

Ce gène appartient à la famille des métalloprotéases responsables de la maturation du propeptide de collagène de type 1. Ce candidat est intéressant bien que son implication dans la physiopathologie

cardiovasculaire soit peu connue : il a été identifié dans la plaque d'athérome au même titre que plusieurs autres procollagénases (Lee et al., 2012a).

D'une façon générale, l'analyse réalisée par croisement montre que l'ajout d'un cas isolé augmente fortement le nombre de candidats à analyser. La difficulté est encore augmentée dans ce cas par l'origine ethnique de la patiente, qui ajoute une masse critique supplémentaire de variants dont certains passent nécessairement les filtres de base.

Dans les analyses intrafamiliales développées ci-après, l'inclusion de cas sporadiques a généré un nombre de gènes candidats trop important si l'on sélectionne les variants rares, prédits pathogènes (tableau 1). Ils ont donc été écartés des analyses ultérieures dans un souci de simplification stratégique.

10.4.3 Analyse intrafamiliale

Les gènes candidats ont été analysés famille par famille, en débutant par l'analyse des trios.

10.4.3.1 Analyse de trios

L'analyse des données des familles FDP200 et FDP249 a été menée selon deux modèles : *de novo* et récessif. L'analyse en modèle *de novo* et en modèle récessif avec mutations homozygotes ne révèle aucune variation codante ou d'épissage. La recherche de variants hétérozygotes composites a identifié dans chacune de ces deux familles un unique gène compatible avec une transmission récessive :

Dans la famille FDP249, le cas-index porte 4 variants dans un seul gène appelé *AHNAK2* (tableau 3) : un variant est hérité du père et trois variants sont hérités de la mère. Ils ont tous une fréquence en population comprise entre 1/10 000 et 1‰, et parmi les variants hérités de la mère, l'un est prédit bénin et les autres ont des prédictions divergentes. La protéine *AHNAK2* est une composante du complexe de la dysferline (Huang et al., 2007), protéine dont le rôle dans la réparation des membranes plasmiques des cellules du muscle strié est bien connu. Cependant la dysferline est aussi exprimée dans l'endothélium, où son inactivation résulte en un défaut d'adhésion et de croissance cellulaire (Sharma et al., 2010). Selon la source GTEX, *AHNAK2* serait exprimé dans les tissus artériels, mais son rôle n'y est pas connu. Il n'existe pas de modèle murin muté pour ce gène. Les arguments pour son implication dans la DFM sont relativement faibles.

Dans la famille FDP200, deux variants faux-sens (c.5013A>C p.Lys1671Asn, et c.5029G>A p.Val1677Met) étaient hérités des parents dans un seul gène : le gène *OBSCN* codant l'obscurine (tableau 3). Ces variants sont présents dans les bases de données à une fréquence respective de 0, 3% et 0,02%, et sont très proches l'un de l'autre compte tenu de la taille de cette protéine. Les prédictions de pathogénicité divergent pour chacun d'entre eux, et ils sont situés dans un domaine de type immunoglobuline (cf infra).

OBSCN est l'un des 4 gènes associés à la DFM multifocale (Kiando et al., 2015), présentant la significativité la plus élevée ($p=0,003$, voir section 7. Génétique de la DFM). Les variants identifiés dans la famille FDP200, malgré leur présence dans les bases de données de population, ne sont pas présents sur la puce exome chip et nous n'avons donc pas d'information concernant leur fréquence dans la cohorte analysée.

L'obscurine, aussi appelée obscurin-myosin light chain kinase ou obscurin-RhoGEF, est une protéine sarcomérique géante des cellules musculaires striées, associée à la titine. Elle contient par ailleurs de multiples motifs structurels mais aussi catalytiques : au moins 68 domaines de type immunoglobuline, deux domaines fibronectine, un domaine de liaison calcium-calmoduline, un domaine RhoGEF, et deux domaines sérine-thréonine kinase de type Myosine Light Chain Kinase. C'est donc aussi un facteur d'échange GTP/GDP de Rho au même titre que FGD5 vu précédemment. De récentes études révèlent son rôle comme régulateur de RhoA dans un modèle cellulaire de cancer épithélial mammaire : l'obscurine active la contractilité cellulaire en induisant la phosphorylation de RhoA, de la chaîne légère de la myosine et de la myosine phosphatase (Perry et al., 2014). Son inactivation induit une baisse de la contractilité cellulaire (figure 30). Une récente étude menée par et Ackermann et coll. démontre l'existence de multiples isoformes (figure 31) dans de multiples tissus musculaires et non musculaires, dont certaines sont ubiquitaires et d'autres tissu spécifiques (Ackermann et al., 2014). Par ailleurs, selon le type cellulaire, la localisation de ces isoformes peut être nucléaire, cytosolique ou membranaire. Le modèle murin inactivé pour ce gène présente une forme tardive et modérée de myopathie avec lésions histologiques associées, mais pas d'atteinte artérielle connue (Lange et al., 2009). Dans cette étude, l'expression abondante dans les tissus vasculaires d'isoforme(s) contenant les domaines liant l'ankyrine, kinase, N- et C- terminaux a été démontrée par IHC.

Les variants

De par son identification par deux approches indépendantes, étude d'association de variants rares en grande cohorte d'une part et unique gène candidat dans une analyse de trio en modèle récessif d'autre part, *OBSCN* est un très bon candidat. De plus son rôle pro-contrastile relativement direct par activation des régulateurs de la myosine le rend particulièrement attrayant (figure 30).

Gène	description	p.	c.	POLYPHEN	SIFT	EXAC caucasiens
AHNAK2	AHNAK nucleoprotein 2	p.Lys3802Arg	c.11405A>G	0,981	0,03	0,03%
		p.His3050Gln	c.9150C>G	0,005	0,58	0,03%
		p.Gly2302Glu	c.6905G>A	0,988	0,2	0,03%
		p.Pro2225His	c.6674C>A	0,784	0,34	0,04%
OBSCN	obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF	p.Lys327Asn	c.981A>C	0,018	0,03	0,30%
		p.Val333Met	c.997G>A	0,605	0,1	0,02%

Tableau 3: Données issues de la recherche de variants hétérozygotes composites dans les trios FDP200 et FDP249

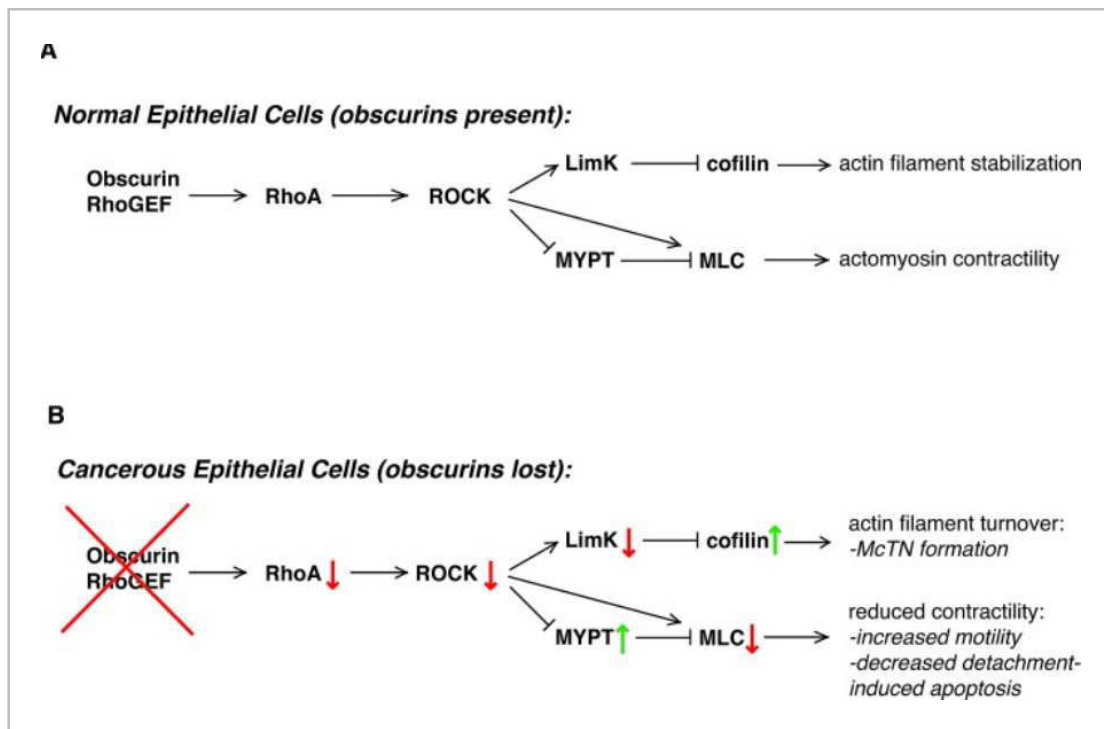


Figure 30: Modélisation du rôle de l'obscurine dans la contractilité des cellules épithéliales. D'après Perry et al, 2014.

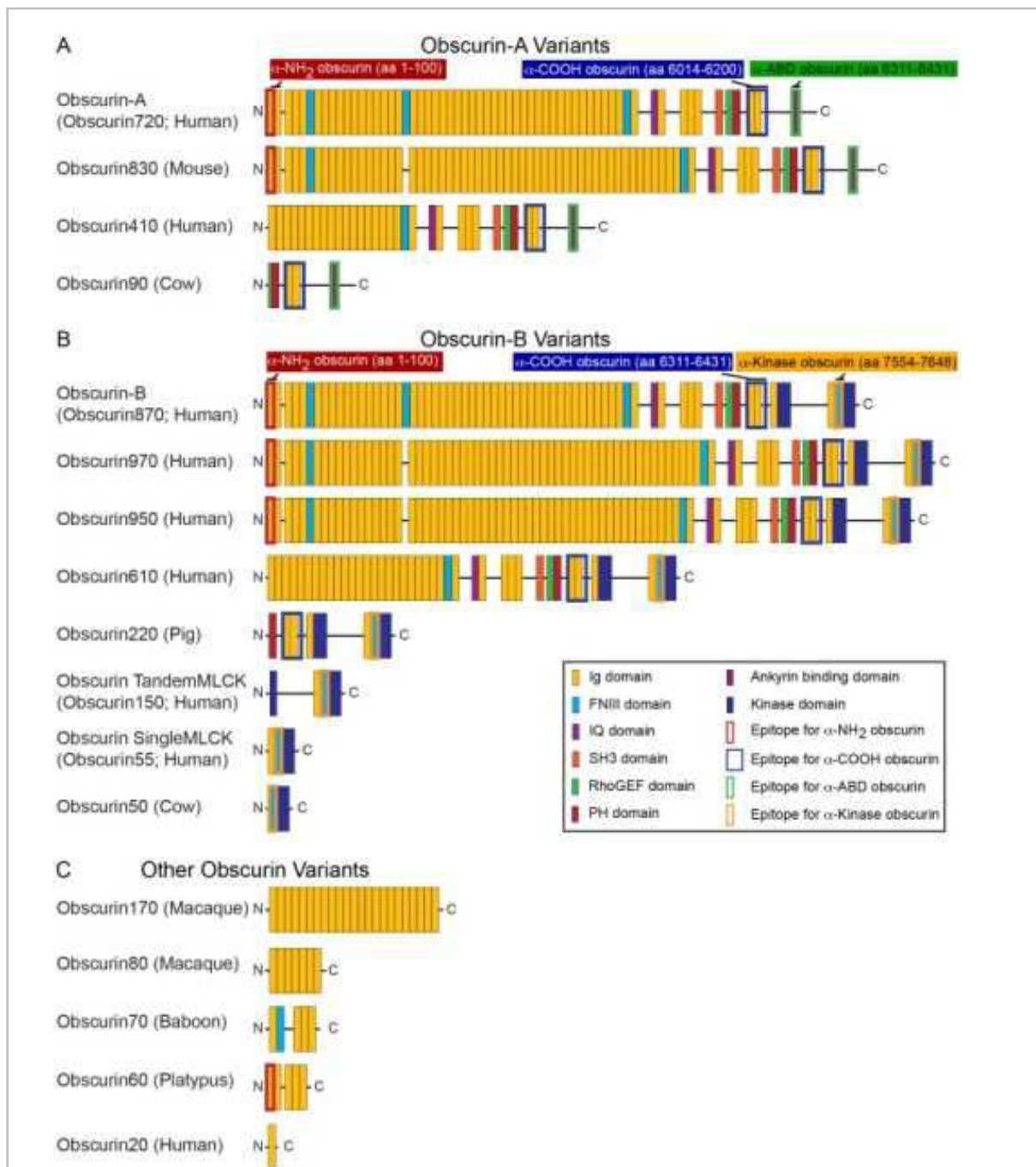


Figure 31: Isoformes de l'obscurine, selon Ackermann et al, 2014.

Terme	Nb genes
artery stenosis	47
abnormal kidney arterial blood vessel morphology	28
vascular smooth muscle cell hyperplasia	17
vascular smooth muscle cell hypoplasia	33
vascular smooth muscle hypertrophy	1
vascular smooth muscle hypotrophy	0
dilated vasculature	51
perivascular fibrosis	20

Tableau 4: Termes du MPO utilisés pour la recherche de gènes candidats, et nombre de gènes correspondants dans la base de données de modèles KO murins (source Jackson Laoratory).

10.4.3.2 Recherche ciblée de gènes candidats

A cette étape, une recherche systématisée de gènes candidats, famille par famille, a été menée, en excluant les trios analysés précédemment et les cas isolés.

Les gènes ont été choisis à partir de plusieurs sources, constituant une liste de plusieurs centaines de candidats.

10.4.3.2.1 Phénotypes murins

En première approche, les gènes connus pour donner un phénotype pertinent au niveau vasculaire chez la souris ont été choisis à partir de 8 entrées de la classification ontologique des phénotypes de mammifères (Mammalian Phenotype Ontology, tableau 4) soit 105 gènes candidats (tableau 5).

Cette approche n'a pas permis d'identifier de gènes de phénotype vasculaire murin chez les patients de ce projet.

10.4.3.2.2 Voies du HAVS et des diagnostics différentiels

La recherche systématique, famille par famille, à partir de listes de gènes potentiellement impliqués dans les pathologies approchantes a été effectuée : d'une part à partir de la liste des gènes de protéines de la cascade du HAVS en s'inspirant de l'article de Stoyneva et coll. récapitulant les voies mecano- et neuro-transmettrices impliquées dans cette maladie vasoconstrictrice (figure 10, Stoyneva, 2003). D'autre part, les données des voies impliquées dans les maladies Mendéliennes sténosantes, anévrysmales et disséquantes ont été utilisées, et non uniquement les gènes connus des maladies exclues en première analyse : il s'agit des voies RAS, NOTCH, PI3K/AKT, TGF β , de la contraction de la cellule musculaire lisse, et de la signalisation calcique (impliquée dans la biosynthèse du NO, la vasodilatation et la vasoconstriction notamment, figures 19, 32,33, 34, 35, 36). Le site KEGG (Kyoto Encyclopedia of Genes and Genomes) propose à partir de chaque voie développée sur son site, une interface interactive, dans laquelle les gènes connus impliqués sont cités. L'ensemble des gènes de chaque voie a été extrait, donnant une liste concaténée d'environ 250 candidats potentiels.

La recherche de variants de ces candidats dans les 8 familles étudiées met en évidence plus de 40 gènes potentiels. Dans cette approche, chaque gène est par définition un bon candidat, et les variants identifiés sont à plus de 95% des faux-sens dont les prédictions divergent. Il n'a pas été possible de restreindre encore plus cette liste à partir d'arguments supplémentaires (de ségrégation par exemple) et seule une confirmation par séquençage Sanger a été réalisée à titre systématique.

Symbole	Nom	Symbole	Nom
<i>Aars</i>	alanyl-tRNA synthetase	<i>lepr</i>	leptin receptor
<i>Abcc9</i>	ATP-binding cassette, sub-family C	<i>Lmna</i>	lamin A/C
<i>Ace</i>	angiotensin I converting enzyme	<i>Lox</i>	lysyl oxidase
<i>Ace2</i>	angiotensin I converting enzyme 2	<i>Lpar4</i>	lysophosphatidic acid receptor 4
<i>Adm</i>	adrenomedullin	<i>Mapk14</i>	mitogen-activated protein kinase 14
<i>agt</i>	angiotensinogen	<i>Mkl2</i>	MKL/myocardin-like 2
<i>agtr1a</i>	angiotensin II receptor, type 1a	<i>Mmp9</i>	matrix metalloproteinase 9
<i>apoe</i>	apolipoprotein E	<i>Msx1</i>	msh homeobox 1
<i>Arntl</i>	aryl hydrocarbon receptor nuclear translocator-like	<i>Msx2</i>	msh homeobox 2
<i>Atp2a2</i>	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2	<i>Mybpc3</i>	myosin binding protein C, cardiac
<i>bmp5</i>	bone morphogenetic protein 5	<i>Myh6</i>	myosin, heavy polypeptide 6, cardiac muscle, alpha
<i>Bmpr1a</i>	bone morphogenetic protein receptor, type 1A	<i>Nos1</i>	nitric oxide synthase 1, neuronal
<i>Bmpr2</i>	bone morphogenetic protein receptor, type II	<i>Nos2</i>	nitric oxide synthase 2, inducible
<i>Calcr1</i>	calcitonin receptor-like	<i>nos3</i>	nitric oxide synthase 3
<i>Cbs</i>	cystathionine-beta-synthase	<i>Notch2</i>	notch 2
<i>Ccm2</i>	cerebral cavernous malformation 2	<i>nphs2</i>	nephrosis 2, idiopathic, steroid-resistant
<i>CD59</i>	CD59 molecule, complement regulatory protein	<i>Npr1</i>	natriuretic peptide receptor 1
<i>Cd59a</i>	CD59a antigen	<i>Nrp1</i>	neuropilin 1
<i>Cdh5</i>	cadherin 5, type 2 (vascular endothelium)	<i>Pax3</i>	paired box 3
<i>Cited2</i>	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	<i>Pdgfra</i>	platelet-derived growth factor receptor, beta polypeptide
<i>colla2</i>	collagen, type I, alpha 2	<i>Pgr</i>	progesterone receptor
<i>cstc</i>		<i>Pitx2</i>	paired-like homeodomain transcription factor 2
<i>Cts3</i>	cathepsin 3	<i>Prkcd</i>	protein kinase C, delta
<i>Cxcr4</i>	chemokine (C-X-C motif) receptor 4	<i>Prox1</i>	protein 5
<i>Dll1</i>	delta-like 1	<i>Psen1</i>	presenilin 1
<i>Dll4</i>	delta-like 4	<i>Ptpn11</i>	protein tyrosine phosphatase, receptor type, J
<i>Dnaic1</i>	dynein, axonemal, intermediate chain 1	<i>Raf1</i>	Raf-1 proto-oncogene, serine/threonine kinase
<i>Efemp2</i>	EGF containing fibulin-like extracellular matrix protein 2	<i>Rapgef1</i>	Rap guanine nucleotide exchange factor (GEF) 1
<i>Eln</i>	elastin	<i>ren1</i>	renin 1 structural
<i>Fam204</i>	family with sequence similarity 20, member A	<i>Rgs2</i>	regulator of G-protein signaling 2
<i>fbn4</i>	fibulin 4	<i>Rras</i>	related RAS viral (r-ras) oncogene homolog
<i>Foxc1</i>	forkhead box C1	<i>Rxra</i>	retinoid X receptor, alpha
<i>Gata4</i>	GATA binding protein 4	<i>S1pr1</i>	sphingosine-1-phosphate receptor 1
<i>Gata5</i>	GATA binding protein 5	<i>Scn5a</i>	sodium channel, voltage gated, type V alpha subunit
<i>Gper1</i>	G protein-coupled estrogen receptor 1	<i>Serpind1</i>	serpin peptidase inhibitor, clade D (heparin cofactor), member 1
<i>Hand2</i>	heart and neural crest derivatives expressed transcript 2	<i>Sgcd</i>	sarcoglycan, delta (dystrophin-associated glycoprotein)
<i>hbb</i>	hemoglobin, beta	<i>Shc1</i>	SHC (Src homology 2 domain containing) transforming protein 1
<i>Hey2</i>	hairly/enhancer-of-split related with YRPW motif 2	<i>slc2a10</i>	solute carrier family 2 (facilitated glucose transporter), member 10
<i>Hhex</i>	hematopoietically expressed homeobox	<i>Smad7</i>	SMAD7
<i>Hmax1</i>	heme oxygenase 1	<i>Tagln</i>	Tagln
<i>Hoxa13</i>	homeobox A13	<i>Tbx1</i>	Tbx1
<i>Il1m</i>	interleukin 1 receptor antagonist	<i>Tek</i>	TEK tyrosine kinase, endothelial
<i>Itga7</i>	integrin alpha 7	<i>tgfb1</i>	tgfb1
<i>Jag1</i>	jagged 1	<i>Thbs1</i>	thrombospondin 1
<i>Kcnj6</i>	potassium inwardly-rectifying channel, subfamily J, member 6	<i>urehr1</i>	urea phenotype 1
<i>Kdr</i>	kinase insert domain receptor	<i>Vav2</i>	vav 2 guanine nucleotide exchange factor
<i>Klf15</i>	Kruppel-like factor 15	<i>vegfa</i>	vascular endothelial growth factor A
<i>Klf5</i>	Kruppel-like factor 5 (intestinal)	<i>Wnt1</i>	wingless-type MMTV integration site family, member 1
<i>Krit1</i>	KRIT1, ankyrin repeat containing	<i>wt1</i>	Wilms tumor 1
<i>Lama4</i>	laminin, alpha 4	<i>Xirp1</i>	xin actin-binding repeat containing 1
<i>lcat</i>	lecithin-cholesterol acyltransferase	<i>Zfand5</i>	Zfand5
<i>lcr</i>	hemoglobin, alpha 2 locus control region	<i>Zfpn2</i>	zinc finger protein, multitype 2

Tableau 5: Liste des candidats extraits du MPO (source Jackson Laboratory).

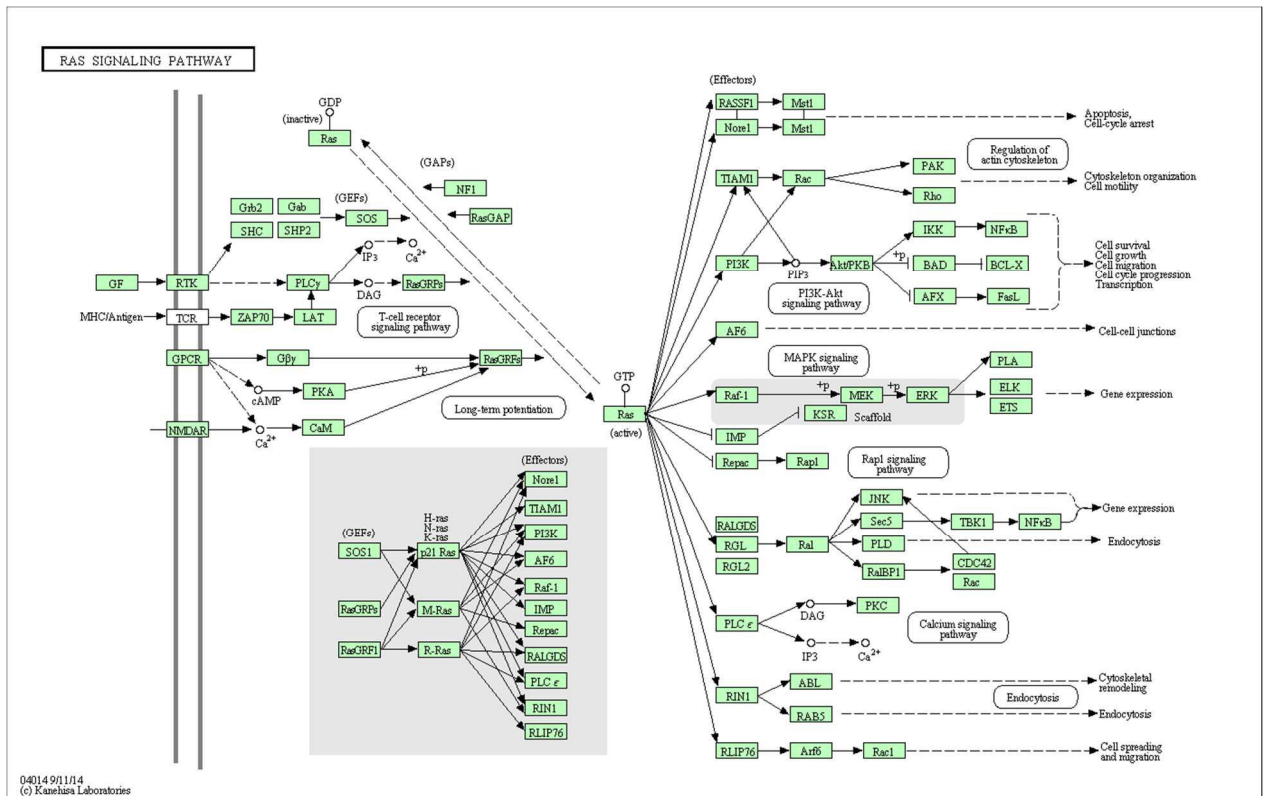


Figure 32 : La voie RAS (source KEGG).

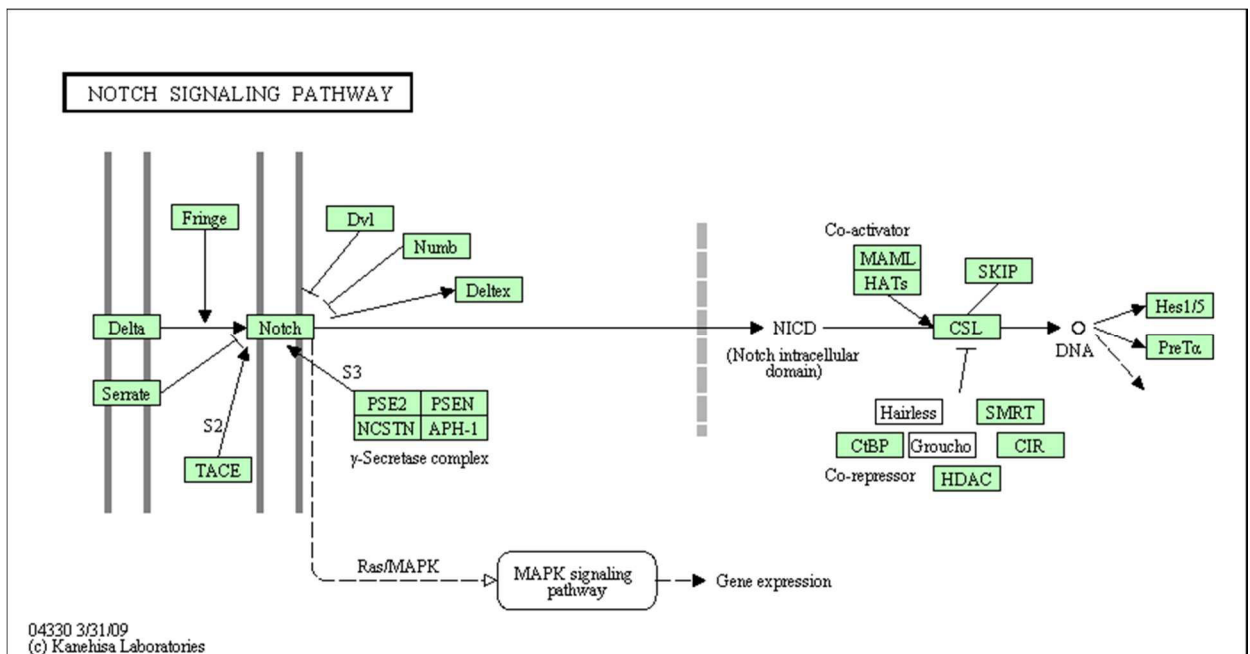


Figure 33 : La voie NOTCH (source KEGG).

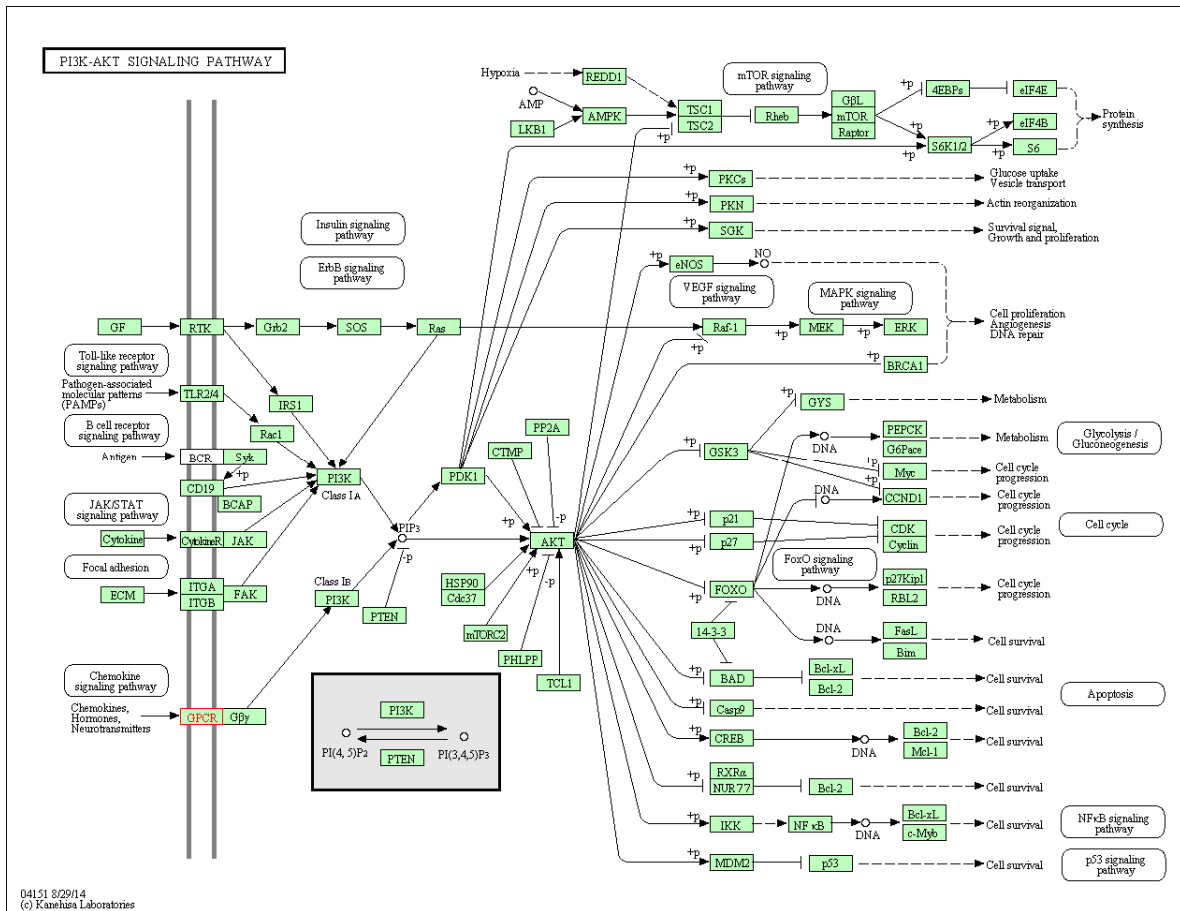


Figure 34 : La voie PI3K/AKT (source KEGG).

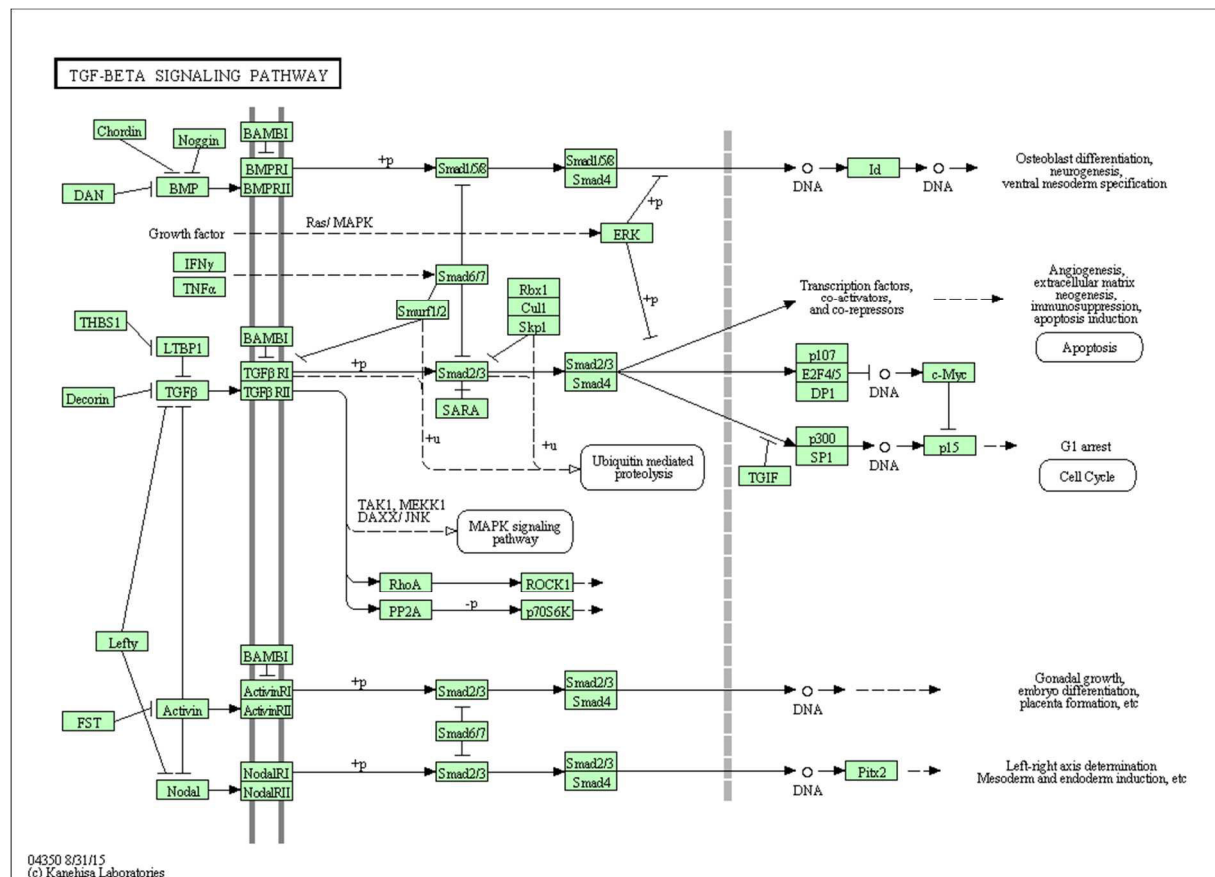


Figure 35 : La voie TGFβ (source KEGG).

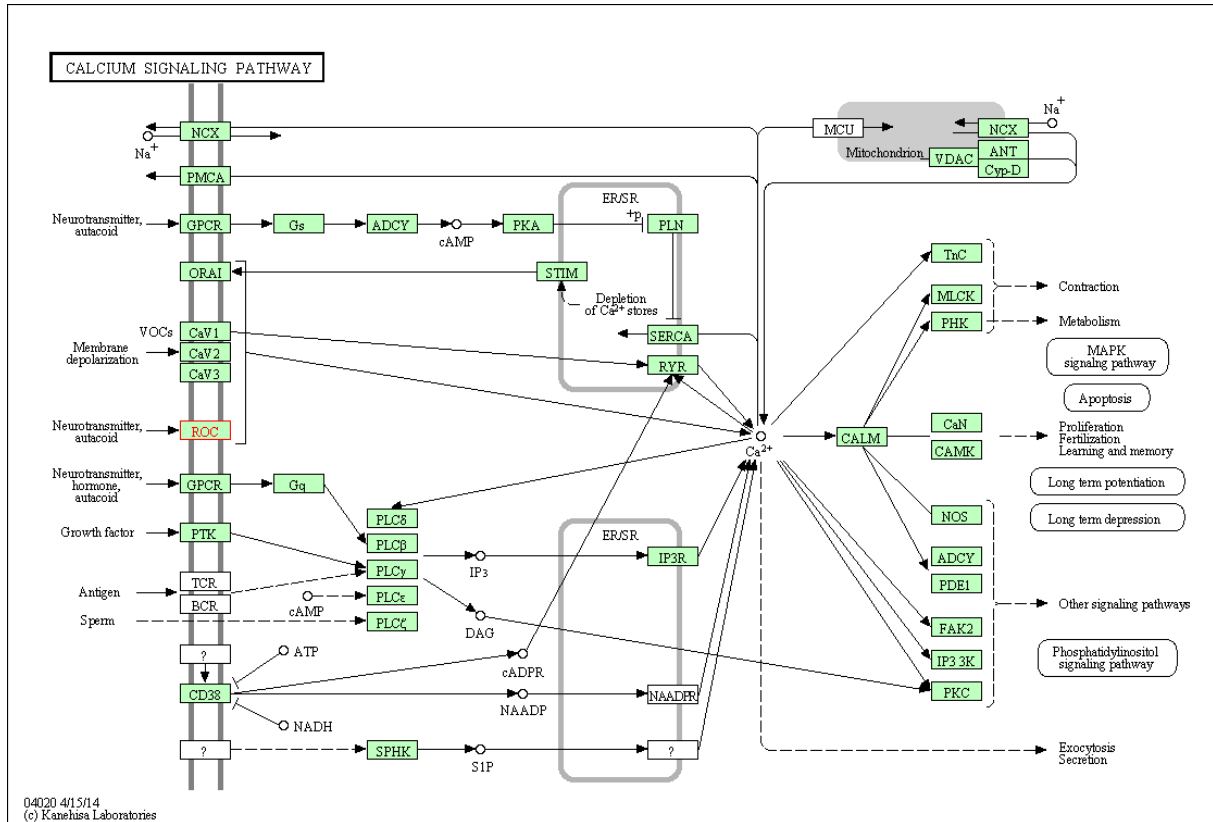


Figure 36 : La voie de signalisation calcique (source KEGG).

Quelques gènes sont d'intérêt particulier, comme *SMTN* codant la smootheline (protéine de l'appareil contractile de la cellule musculaire lisse, nécessaire à la réponse aux vasoconstricteurs, (Niessen et al., 2005) ou *ADAMTS2* (codant une enzyme de maturation du collagène de type 1 notamment, responsable du SED type dermatosparaxis), respectivement identifiés dans les familles FDP18 et SED294. Cette approche très chronophage, et faiblement sélective a été écartée au profit d'une approche plus stringente mais efficace.

10.4.3.3 Analyse combinée et priorisation des gènes candidats

L'analyse précédente a mis en évidence les difficultés d'analyse dans le cadre de variants rares avec des hypothèses physiopathologiques relativement floues. La liste de gènes sélectionnés par l'approche ci-dessus contient peut-être un, voire plusieurs gènes impliqués dans la DFM du fait de l'hétérogénéité de la maladie. Cependant le très grand nombre de candidats sélectionnés par cette approche la rend inefficace et non pertinente dans l'optique d'une validation ultérieure par des tests fonctionnels. Ces derniers ne peuvent pas être réalisés à haut débit dans la mesure où le phénotype cellulaire testé va dépendre en partie de la fonction de la protéine correspondante.

La somme des gènes et variations, concordants avec une transmission dominante dans les 8 familles séquencées était d'environ 260 (de 14 gènes pour la famille FDP-8 comportant 3 individus à 48 gènes pour la famille FDP11 comportant 2 individus, tableau 2). Un sous-groupe de variants dont la prédiction de pathogénicité est la plus forte, sans hypothèse physiopathologique préalable, a été sélectionné. Seules les mutations stop, frameshift, de la méthionine initiatrice, et de sites d'épissage canonique ont été analysés, correspondant à une liste de 27 gènes : 11 portaient des frameshifts, 8 des stops, 4 des mutations d'épissage canonique, et 4 des mutations de la méthionine initiatrice (tableau 6). Un traitement séquentiel des données de ces 27 gènes a permis de réduire encore la liste de ces candidats (figure 37) : 5 gènes de fonction inconnue ont été éliminés des analyses ultérieures. Puis un calcul de la fréquence cumulée dans EXAC de toutes les mutations perte de fonction de chacun des gènes permet d'obtenir une estimation réelle de la fréquence des haploinsuffisances dans la population caucasienne, et donc de la pression de sélection s'exerçant sur leur fonction. Les gènes *GANC* (50% d'haploinsuffisants chez les caucasiens), *CELA1* (22%), *CDH23* (3%) et *XIRP2* (2,5%) ont été éliminés à cette étape.

L'analyse bibliographique des 18 gènes restants montre que 8 d'entre eux n'ont aucune fonction connue dans la physiopathologie vasculaire, de la matrice extracellulaire, ou les voies de signalisation impliquées dans les maladies approchantes développées précédemment, ils ont été éliminés.

gène	Famille	consequence	description	exon	p.	c.
1	C3orf52	frameshift	chromosome 3 open reading frame 52 [Source:HGNC Symbol;Acc:26255]	4	p.Ala216Argfs*2	c.642delG
2	C9orf85	splicing,frameshift	chromosome 9 open reading frame 85 [Source:HGNC Symbol;Acc:28784]	2	p.Ile36Metfs*20	c.105_108delAATT
3	DCAF4	frameshift	DDB1 and CUL4 associated factor 4 [Source:HGNC Symbol;Acc:20229]	5	49	c.145_152delCTGC GTCT
4	P2RX3	frameshift	purinergic receptor P2X, ligand-gated ion channel, 3 [Source:HGNC Symbol;Acc:8534]	2	p.Tyr80Thrfs*2	c.237delIT
5	PTGIR	stop	prostaglandin I2 (prostaglyclin) receptor (IP) [Source:HGNC Symbol;Acc:9602]	2	p.Gln163*	c.487C>T
6	RALGDS	Méthionine initiatrice	ral guanine nucleotide dissociation stimulator [Source:HGNC Symbol;Acc:9842]	1	p.Met1?	c.2T>C
7	THAP9	frameshift	THAP domain containing 9 [Source:HGNC Symbol;Acc:23192]	5	p.Gln480Profs*4	c.1438dup
8	BIN3	Epissage canonique	bridging integrator 3 [Source:HGNC Symbol;Acc:1054]	i7		
9	CCDC183	frameshift	coiled-coil domain containing 183 [Source:HGNC Symbol;Acc:28236]	6	p.Gly185Aspfs*23	c.552delA
10	CREM	stop	cAMP responsive element modulator [Source:HGNC Symbol;Acc:2352]	1	p.Glu25*	c.73G>T
11	GANC	stop	glucosidase, alpha; neutral C [Source:HGNC Symbol;Acc:4139]	5	165	c.493C>T
12	CNDP1	Méthionine initiatrice	carnosine dipeptidase 1 (metallopeptidase M20 family) [Source:HGNC Symbol;Acc:20675]	1	p.Met1?	c.3G>A
13	ECH1	stop	enoyl CoA hydratase 1, peroxisomal [Source:HGNC Symbol;Acc:3149]	4	148	c.442C>T
14	ELOVL7	frameshift	ELOVL fatty acid elongase 7 [Source:HGNC Symbol;Acc:26292]	8	188	c.563delA
15	ZCWPW1	stop	zinc finger, CW type with PWWP domain 1 [Source:HGNC Symbol;Acc:23486]	14	p.Arg428*	c.1282C>T
16	CELA1	frameshift	chymotrypsin-like elastase family, member 1 [Source:HGNC Symbol;Acc:3308]	5	p.Ala140Valfs*47	c.419delC
17	ARFGEF1	stop	ADP-ribosylation factor guanine nucleotide-exchange factor 1 (brefeldin A-inhibited) [Source:HGNC Symbol;Acc:15772]	39	p.Gln1838*	c.5512C>T
18	CDH23	frameshift	cadherin-related 23 [Source:HGNC Symbol;Acc:13733]	41	p.Asn1818Thrfs*2	c.5453delC
19	LPAR6	stop	lysophosphatidic acid receptor 6 [Source:HGNC Symbol;Acc:15520]	5	p.Arg49*	c.145C>T
20	TMPRSS2	Méthionine initiatrice	transmembrane protease, serine 2 [Source:HGNC Symbol;Acc:11876]	1	p.Met1?	c.2T>G
21	ADAMTS14	frameshift	ADAM metallopeptidase with thrombospondin type 1 motif, 14 [Source:HGNC Symbol;Acc:14899]	7	p.Cys379Phefs*4	c.1136delIT
22	ZNF616	stop	zinc finger protein 616 [Source:HGNC Symbol;Acc:28062]	4	p.Gln204*	c.610C>T
23	DEFA5	Epissage canonique	defensin, alpha 5, Paneth cell-specific [Source:HGNC Symbol;Acc:2764]	i1	p.?	c.173-1G>A
24	LEPROTL1	Epissage canonique	leptin receptor overlapping transcript-like 1 [Source:HGNC Symbol;Acc:6555]	i3	p.?	c.280-2A>G
25	SLC28A3	Epissage canonique	solute carrier family 28 (concentrative nucleoside transporter), member 3 [Source:HGNC Symbol;Acc:16484]	i10	p.?	c.1023+1G>A
26	XIRP2	Méthionine initiatrice	xin actin-binding repeat containing 2 [Source:HGNC Symbol;Acc:14303]	2	p.Met1?	c.2T>C
27	ZIC5	frameshift	Zic family member 5 [Source:HGNC Symbol;Acc:20322]	1	p.Pro13Glnfs*2	c.38delC

Tableau 6: Liste des gènes portant une mutation perte de fonction .

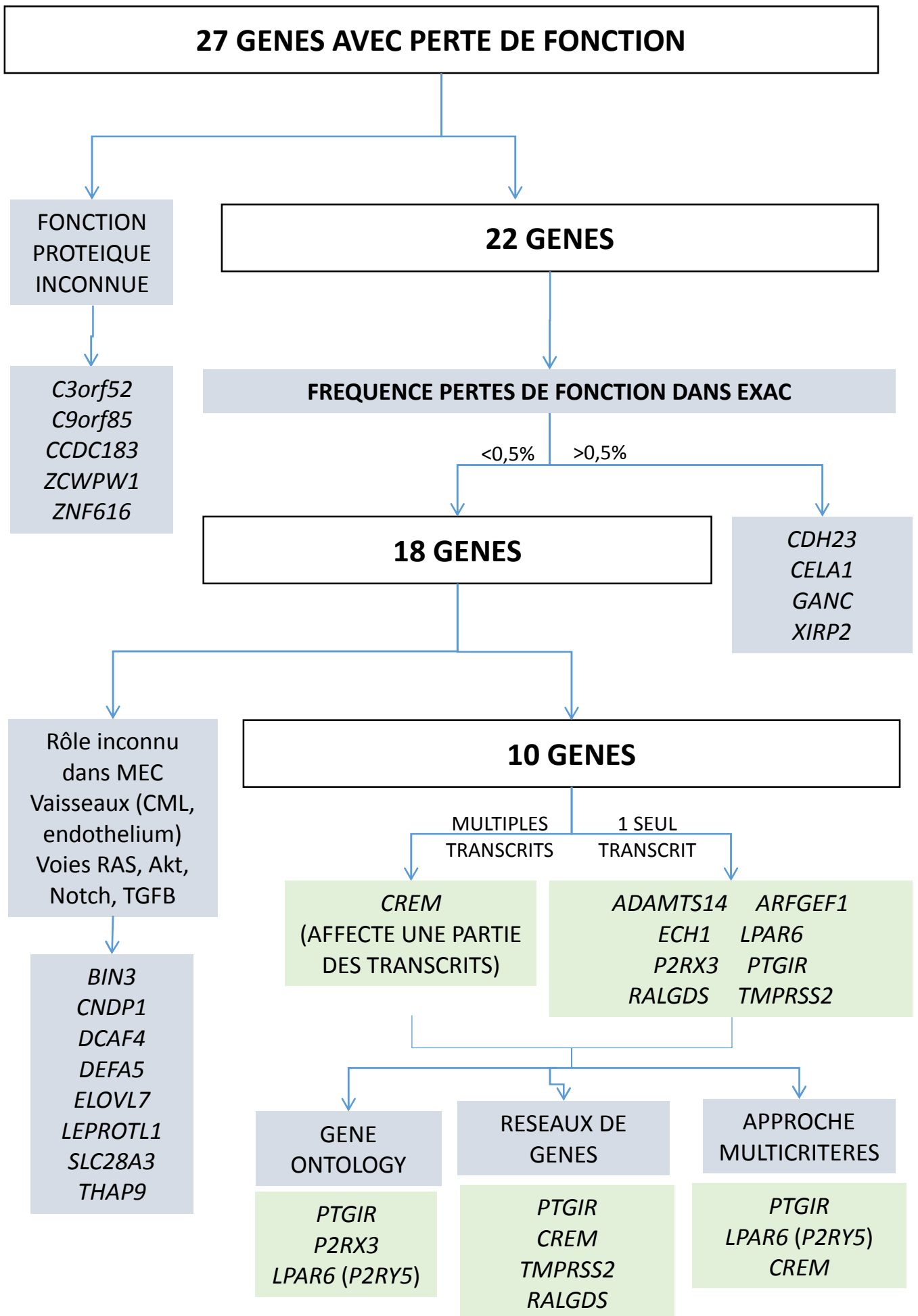


Figure 37 : Stratégie d'analyse des variations perte de fonction.

Parmi les 10 gènes restants, l'existence de transcrits alternatifs était un élément important pour déterminer l'impact des mutations sur l'ensemble des transcrits existants. La phase des exons correspondant aux mutations d'épissage permet de prédire l'effet d'un éventuel saut d'exon pouvant induire un décalage de lecture. Seuls deux gènes portent des transcrits alternatifs et sont donc concernés par cette analyse : (i) *CREM*, dont la fréquence cumulée des pertes de fonction est de 0,2%, mais concerne une partie des 22 transcrits alternatifs de ce gène très polymorphe. La mutation identifiée concerne uniquement trois isoformes parmi 22. Nous verrons plus loin la fonction de *CREM* et le rôle biologique de ces diverses isoformes. (ii) *SLC28A3*, portant une mutation d'épissage d'un exon en phase. Ce gène code deux isoformes transcriptionnelles toutes deux impactées par la mutation d'épissage.

De façon évidente certains de ces gènes comme *PTGIR*, codant le récepteur à la prostacycline étaient d'excellents candidats. Cependant certaines familles présentent deux candidats possibles par cette approche (tableau 6). La liste des 10 candidats a été soumise à divers outils informatiques (voir chapitre 1, section 4.4.3.1) afin de déterminer l'existence d'un réseau de relations entre certains d'entre eux : DAVID (outil de clustering fonctionnel de gènes) à la recherche d'un enrichissement dans certaines classes ontologiques de gènes, STRING à la recherche d'un lien fonctionnel ou prédit entre gènes, et enfin ENDEAVOUR à la recherche d'une concordance avec les gènes des maladies proches acquises et héréditaires.

10.4.3.3.1 Analyse DAVID

David compare une liste de gènes à l'ensemble des gènes humains codants, en termes de distribution dans les différentes classes ontologiques de gènes. Cette analyse montre un enrichissement pour des protéines impliquées dans l'interaction ligand neuroactif/récepteur (30% de gènes de la liste, $p_{\text{corr}}=0,043$, figure 38). Ce sont les trois gènes *PTGIR*, *P2RX3*, et *LPAR6* codant des récepteurs membranaires à des molécules neuroactives. Parmi eux, deux gènes codent des protéines appartenant à la famille des récepteurs aux bases puriques : *LPAR6* et *P2RX3*. Enfin, DAVID révèle qu'en plus de ces trois récepteurs, deux autres gènes codent des protéines membranaires : un récepteur à tyrosine kinase appelé *TMPRSS2*, et un transporteur de bases puriques (*SLC28A3*). Ce résultat n'est pas significatif ($p_{\text{corr}}=0,79$). Enfin, deux gènes codent des protéines d'échange GTP/GDP : les gènes *ARFGEF1* et *RALGDS*. Ce résultat n'est pas non plus différent de la répartition attendue dans les gènes humain ($p_{\text{corr}}=0,74$).

10.4.3.3.2 Analyse STRING

STRING analyse les liens fonctionnels entre gènes à partir de données expérimentales de coexpression et d'interaction, de données bibliographiques et de prédiction de fonction protéique. L'analyse des 10 candidats permet d'obtenir un réseau de gènes impliquant *PTGIR*, *RALGDS*, *CREM* et *TMPRSS2* (figure 39).

Functional Annotation Chart

[Help and Manual](#)

Current Gene List: 11-LOF-these
 Current Background: Homo sapiens
 11 DAVID IDs

Options

10 chart records

[Download File](#)

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	KEGG_PATHWAY	Neuroactive ligand-receptor interaction	RT		3	27,3	7,3E-3	4,3E-2
<input type="checkbox"/>	SP_PIR_KEYWORDS	alternative promoter usage	RT		2	18,2	2,4E-2	7,9E-1
<input type="checkbox"/>	GOTERM_MF_FAT	nucleotide receptor activity	RT		2	18,2	2,5E-2	8,5E-1
<input type="checkbox"/>	GOTERM_MF_FAT	purinergic nucleotide receptor activity	RT		2	18,2	2,5E-2	8,5E-1
<input type="checkbox"/>	UP_SEQ_FEATURE	topological domain:Extracellular	RT		5	45,5	4,2E-2	9,1E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	polymorphism	RT		10	90,9	4,7E-2	7,8E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	guanine-nucleotide releasing factor	RT		2	18,2	6,1E-2	7,4E-1
<input type="checkbox"/>	UP_SEQ_FEATURE	sequence variant	RT		10	90,9	6,6E-2	8,4E-1
<input type="checkbox"/>	UP_SEQ_FEATURE	topological domain:Cytoplasmic	RT		5	45,5	8,3E-2	8,0E-1
<input type="checkbox"/>	UP_SEQ_FEATURE	lipid moiety-binding region:S-palmitoyl cysteine	RT		2	18,2	9,1E-2	7,3E-1

Figure 38: résultats de l'analyse DAVID identifiant un enrichissement significatif en gènes codant des récepteurs de substances neuroactives.

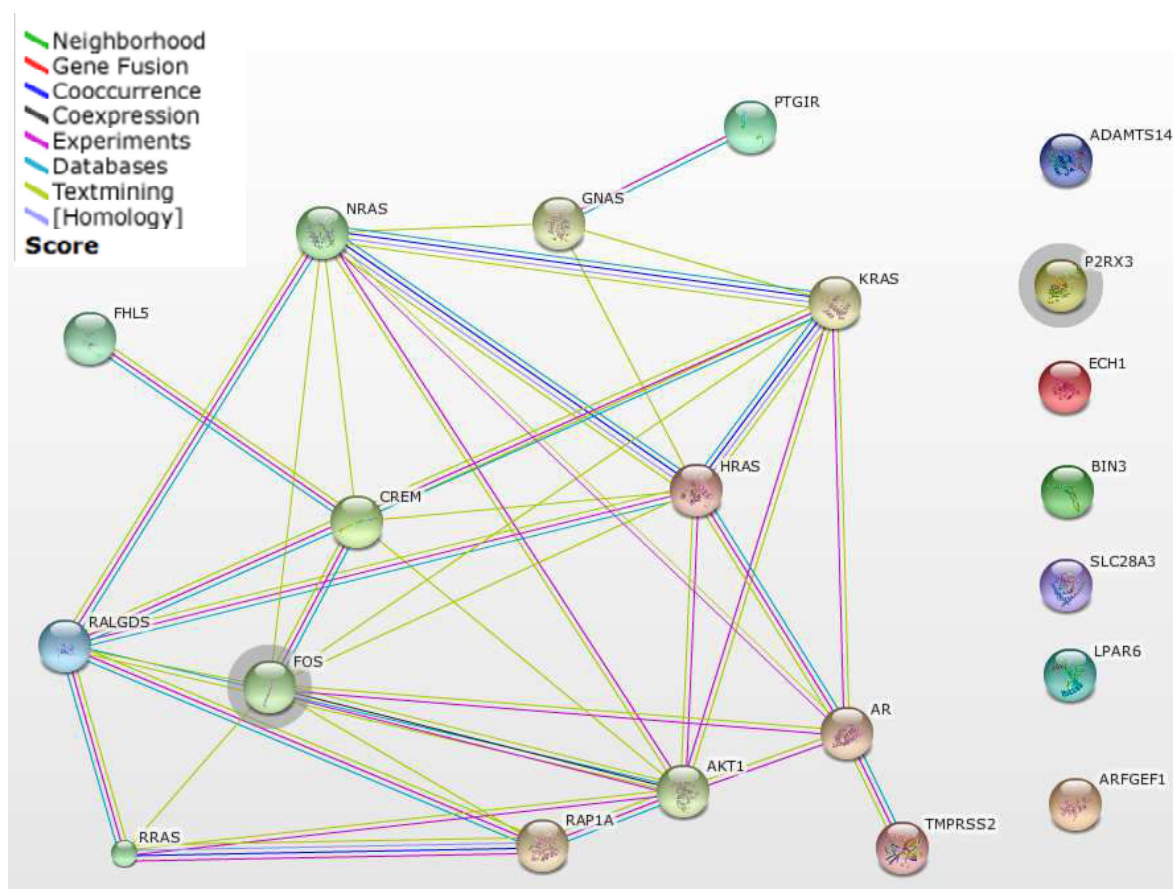


Figure 39: résultats de l'analyse STRING mettant en évidence des interactions connues ou prédites, indirectes, entre *PTGIR*, *CREM*, *RALGDS* et *TMPRSS2*. On note que la voie RAS est présente en interaction avec ces 4 protéines.

Ces gènes sont indirectement liés les uns aux autres à travers les voies RAS et AKT notamment. Les 6 autres gènes ne sont pas identifiés dans un réseau fonctionnel de gènes connu.

La caractérisation des liens de ces 4 gènes avec les voies citées est développée dans la section suivante.

10.4.3.3 Analyse ENDEAVOUR

L'analyse Endeavour utilise une liste de « gènes d'entraînement » pertinents pour la pathologie étudiée (gènes causaux de maladies proches, ou dans le cadre de maladies hétérogènes dont certains gènes sont déjà connus). Endeavour utilise des données d'annotation (issues de KEGG, Gene Ontology, SWISSPROT notamment) d'interaction, d'expression, de motifs protéiques afin de prioriser les gènes d'une liste de candidats. Une liste d'entraînement contenant les gènes de maladies sténosantes, anévrysmales et disséquantes, et de la voie biologique impliquée dans le HAVS a été proposée à cet outil (environ 75 gènes).

Les résultats de priorisation montrent que *PTGIR* est le candidat classé en premier parmi les 10 gènes soumis, suivi de *LPAR6* (appelé aussi *P2RY5*), et *CREM* (tableau 7).

Ces analyses distinctes montrent qu'en première lecture, il n'existe pas un seul réseau très clairement identifié commun à ces trois outils. Il est possible que cette liste ne soit pas particulièrement enrichie en gènes causaux. Il est aussi possible que le lien physiopathologique entre ces gènes puisse ne pas encore avoir été établi et n'apparaisse pas à travers ces outils interrogeant les données existantes. Quelque soit l'analyse cependant, le gène *PTGIR* semble systématiquement mis en avant et représente donc le candidat le plus robuste de cette liste.

Une analyse bibliographique de tous les gènes de ce set a permis de les classer selon la pertinence des données bibliographiques et physiopathologiques.

famille	symbole	score global	rang	ratio	Score GO	rang	ratio	score Swissprot	rang	ratio	Blast score	rang	ratio	score d'expression (5onEtA)	rang	ratio	score d'interactions protéiques	rang	ratio	score de motifs	rang	ratio
FDP11	PTGIR	1	0,0	0,0	1	0,0	0,1	1	0,0	0,1	1	0,0	0,1	3	0,7	0,2	4	-	0,5	1	0,1	0,0
FDP59	P2RY5	2	0,1	0,1	9	0,0	0,4	2	0,0	0,1	2	0,0	0,1	2	0,6	0,1	3	-	0,5	9	1,0	0,6
FDP143	CREM	3	0,2	0,2	4	0,0	0,3	8	0,0	0,5	7	436,0	0,5	1	0,6	0,1	2	-	0,5	5	0,9	0,4
FDP11	RALGDS	4	0,2	0,2	5	0,0	0,4	4	0,0	0,3	5	49,9	0,3	5	0,7	0,4	1	2,00	0,0	4	0,9	0,4
FDP59	TMPRSS2	5	0,2	0,2	6	0,0	0,5	7	0,0	0,4	3	22,6	0,2	7	0,8	0,6	7	-	0,5	3	0,9	0,4
FDP18	ECH1	6	0,4	0,3	8	0,0	0,7	5	0,0	0,4	8	551,0	0,7	4	0,7	0,3	5	-	0,5	2	0,9	0,3
FDP8	ADAMTS14	7	0,4	0,3	7	0,0	0,6	3	0,0	0,3	4	34,0	0,2	6	0,7	0,5	6	-	0,5	7	1,0	0,6
FDP59	ARFGEF1	8	0,7	0,6	2	0,0	0,2	7	0,0	0,5	6	120,0	0,4	8	0,8	0,8	8	-	0,5	6	1,0	0,5
FDP43	P2RX3	9	0,9	0,8	3	0,0	0,3	9	0,0	0,6	9	-	0,9	9	0,9	1,0	9	-	0,5	8	1,0	0,6

Tableau 7 : résultats de l'analyse ENDEAVOUR, classant comme premier candidat le gène *PTGIR* par ontologie de gènes, alignement de séquences, et recherche de motifs en utilisant une liste de 72 gènes d'entraînement, gènes de maladies Mendéliennes artérielles et gènes du HAVS.

10.4.3.3.4 Rôle des gènes identifiés par priorisation

Le gène *PTGIR*

Une mutation perte de fonction de ce gène a été identifiée dans la famille FDP11 (p.Gln163* dans l'exon 2, tableau 6). *PTGIR* ou *IP* code le récepteur à la prostacycline. Il appartient à la famille des protéines à sept domaines transmembranaires couplés aux protéines G. La prostacycline (PGI₂) est une molécule vasodilatatrice produite par les cellules endothéliales. Elle appartient à la famille des prostanoides, métabolites de l'acide arachidonique produits par les enzymes cyclooxygénases. Plus spécifiquement, PGI₂ est produit par la cyclooxygénase de type 2 (COX2).

Les prostanoides sont des molécules neuroactives paracrines, exprimés dans de multiples tissus, et d'actions diverses. PGI₂ est fortement exprimée dans les mégacaryocytes ou elle inhibe le processus d'activation plaquettaire. Dans le thymus et la rate, elle est exprimée par les thymocytes matures et lymphocytes spléniques, où son rôle est encore inconnu (Narumiya et al., 1999).

Dans le système vasculaire, PGI₂ est exprimée par les cellules endothéliales en réponse à un stimulus mécanique. Elle a une action paracrine sur les CML où elle induit une relaxation cellulaire (entouré en vert dans la figure 19), et inhibe la prolifération des CML. Son expression dans les artérioles efférentes des glomérules rénaux est majeure, où elle régule la filtration glomérulaire par vasodilatation en réponse à une baisse du flux artériel. La prostacycline, et son récepteur IP, sont aussi les principaux médiateurs de la voie des prostanoides dans la vasodilatation et la perméabilisation vasculaire liée à l'inflammation aiguë, en synergie avec d'autres médiateurs tels que l'histamine et la bradykinine (Murata et al., 1997).

De façon surprenante, les souris *Ip*^{-/-} ne présentent pas d'hypertension artérielle, bien qu'elles présentent de fait une insensibilité à PGI₂ (Murata et al., 1997). Ces souris ont un développement et un vieillissement normal, et aucun sur-risque d'accident artériel en conditions basales. Cependant, les données issues d'expérimentations murines chez la souris sauvage et *Ip*^{-/-} démontre qu'une expression élevée de *Pgi2* par surexpression de *Cox2*, est une réponse physiologique à une baisse du flux artériel (par exemple par ligature artérielle). Chez la souris mutante, l'insensibilité à *Pgi2* induit un profond remodelage artériel caractérisé par une prolifération neointimale et une diminution de la lumière artérielle. Ce phénomène de remodelage est reproduit par les agents pharmacologiques bloqueurs de COX2 comme le nimesulide (Kawabe et al., 2010; Rudic, 2005). Ces données sont corroborées par le surrisque d'IDM et d'AVC chez les patients traités par inhibiteurs spécifiques de COX2 comme le celecoxib ou le rofecoxib, ayant mené à un retrait du marché de ces médicaments (Rudic, 2005). Le phénomène impliqué dans ce surrisque est probablement double : une levée de l'inhibition plaquettaire d'une part, et une prolifération intimale dans les zones de flux turbulent d'autre part, chez les patients présentant une maladie vasculaire athéromateuse sous-jacente.

Cheng et coll. ont démontré le rôle exactement inverse de PGI2 et du thromboxane (TXA2), molécule vasoconstrictrice, dans ce phénomène de réponse à l'agression endothéliale : la prolifération néointimale exacerbée chez le souris *Ip^{-/-}* est abolie chez le double mutant *Ip^{-/-} Tp^{-/-}* (mutant du récepteur au thromboxane). Chez le mutant *Tp^{-/-}*, la prolifération est inhibée partiellement dans la réponse à l'agression endovasculaire (Cheng et al., 2002).

Enfin, il est intéressant de noter que *IP* est exprimé dans les neurones des ganglions spinaux, où il colocalise avec la substance P (neurotransmetteur impliqué dans la perception et l'intégration de la douleur au niveau du système nerveux central), indiquant que PGI2 puisse être impliquée dans la perception douloureuse. Dans certains neurones, *IP* et le récepteur à la prostaglandine (*EP*) sont coexprimés, où ils pourraient jouer un rôle complémentaire dans la transmission de la douleur. Diverses études menées notamment en modèle murin démontrent l'importance de la voie PGI2/*IP*, exprimée dans les terminaisons nerveuses des fibres afférentes nociceptives, dans la facilitation de la sensation de douleur (Bley et al., 1998). PGI2/*IP* ont été clairement impliqués dans la migraine : L'ARNm d'*IP* est un biomarqueur sanguin identifié dans la migraine (Kheirollahi et al., 2015), et la perfusion de PGI2 déclenche des crises migraineuses chez les patients migraineux chroniques (Wienecke et al., 2010). Compte-tenu de la prévalence de la maladie migraineuse chez les patients atteints de DFM, ces données semblent dresser un tableau cohérent avec l'expression de la DFM.

Le rôle de *PTGIR (IP)* et des autres gènes identifiés dans cette dernière approche est discuté dans la section conclusions préliminaires.

Le gène *CREM*

CREM a été identifié dans la famille FDP143 (p.Glu25*). Il code le modulateur du facteur de réponse à l'AMP cyclique (ou cAMP response element modulator). *CREM* et *CREB* (facteur de réponse à l'AMP) sont deux protéines structurellement proches qui forment un complexe protéique, et ont un rôle de facteurs transcriptionnels de réponse à l'AMPc et le GMPc notamment (Pilz et Casteel, 2003). Leurs gènes présentent tous deux une diversité fonctionnelle notable par l'existence de variants d'épissage et la formation d'homodimères ou hétérodimères liant la séquence CRE (élément de réponse à l'AMPc) présente dans les promoteurs de leurs gènes cibles (Mayr et Montminy, 2001).

L'AMPc favorise la quiescence et le maintien d'un phénotype contractile (par opposition au phénotype synthétique, prolifératif) dans les CML vasculaires, sous la dépendance des récepteurs β -adrénergiques et de la PKA (protéine kinase AMPc-dépendante). *CREB* est impliqué dans la régulation de la croissance et la

migration cellulaire, l'apoptose, et l'hypertrophie des CML (Funakoshi et al., 2002; Tokunou et al., 2003). L'élévation de l'AMPC inhibe, *via* CREB, la croissance cellulaire par blocage de la voie PDGF, par activation de PKA, et par inhibition de la voie MAPK notamment (Begum et al., 2011; Seidl et al., 2015). Plusieurs études ont démontré le rôle clé de CREB dans le développement de maladies vasculoprolifératives, où il est généralement sous-exprimé en réponse à une agression endothéliale (Ichiki, 2006; Reusch et Watson, 2004; Schauer et al., 2010).

L'inhibition de CREB en tant que facteur transcriptionnel est médiée par de nombreux variants de CREM, les petites isoformes appelées Icer et Sicer étant les plus puissantes inhibitrices de CREB (Molina et al., 1993). Ces répresseurs contiennent un domaine de liaison à la séquence CRE, mais ne contiennent pas de domaines de transactivation ou d'inductibilité par des kinases, contrairement à CREB. Ils agissent donc indirectement, possiblement par inhibition compétitive de CREB sur son site de liaison aux séquences CRE.

Il a été démontré que dans les CML artérielles, *CREM* était transcrit en réponse à une stimulation par le NO (Steinbicker et al., 2011), l'aldostérone (Leopold et al., 2007), et le beraprost (analogue de PGI₂). Dans ce dernier modèle d'induction, CREM inhibe la prolifération neointimale après agression endothéliale chez le rat (Ohtsubo et al., 2007). Dans cette étude, l'induction transcriptionnelle de *CREM* est bloquée par inhibition de PKA, suggérant que PKA est aussi un inducteur de *CREM*. Il existe un modèle murin KO de *Crem* présentant un phénotype complexe qui associe une stérilité chez les mâles et une baisse de la contractilité et de la relaxation myocardique, sans hypertrophie associée (Blendy et al., 1996). Ces souris présentent une insensibilité myocardique aux β -adrénergiques par perte des récepteurs β membranaires. Une analyse du phénotype vasculaire de ces souris réalisée par Seidl et coll. démontre qu'en conditions basales, l'aorte de souris *Crem*^{-/-} ne présente pas de phénotype prolifératif accru, et que le tonus aortique est normal. Après ligation carotidienne cependant, l'expression transcriptionnelle de *Crem* augmente chez la souris sauvage. Chez la souris KO, en l'absence de *Crem*, une prolifération des CML vasculaires est observée dans la média, avec présence d'un épaississement neointimal (Seidl et al., 2015).

Ces données sont étonnamment similaires à celle obtenues dans le modèle d'inactivation de la voie PGI₂, en terme d'induction du phénotype prolifératif secondairement à une agression vasculaire. Cependant dans ce modèle, toutes les isoformes de CREM sont inactivées, et non représentatives de la mutation portée par les patients FDP143. Cette dernière concerne les isoformes 15, 16 et 25, l'isoforme 16 étant d'action répressive bien que n'étant pas à ce jour rattachée au groupe des Icer.

Les gènes *P2RX3* et *LPAR6/P2RY5*

P2RX3 et *LPAR6*, aussi appelé *P2RY5*, ont respectivement été identifiés dans les familles FDP43 et 59 (p.Tyr80Thrfs*2 et p.Arg49* respectivement). Ils codent des récepteurs membranaires aux bases puriques (ATP essentiellement, et UTP). L'ATP, outre son rôle dans le métabolisme énergétique et dans de nombreuses cascades intracellulaires de signalisation, est une molécule neuroactive de l'espace extracellulaire libérée par de nombreux types cellulaires. Il régule de nombreux processus physiologiques dont le tonus vasculaire, la fonction cardiaque et l'hémostase.

L'ATP extracellulaire provient de 5 sources distinctes : dans les neurones sympathiques il est libéré en association avec la noradrénaline ; après activation plaquettaire, il est libéré dans la circulation sanguine par dégranulation ; les cellules chromaffines libèrent aussi l'ATP, ainsi que les CML et les cellules endothéliales vasculaires. La dernière source provient de la perméabilisation membranaire secondaire à la mort cellulaire.

Dans le système vasculaire, la stimulation par l'ATP induit des effets complexes, encore mal compris malgré le très grand nombre d'études portant sur le sujet. L'ATP, d'une façon générale, est impliqué dans le contrôle du tonus et du remodelage vasculaire (Burnstock et Ralevic, 2013). Certaines études démontrent qu'il induit la production de PGI₂ et de NO par l'endothélium. Cette source d'ATP peut provenir d'une agression artérielle, responsable du relargage de ces deux molécules vasodilatatrices dans un contexte de réaction endothéliale (Boeynaems et Galand, 1983; Kelm et al., 1988). Dans les cellules musculaires lisses, l'ATP libéré dans le cadre d'une stimulation par les terminaisons des neurones sympathiques est généralement responsable d'une vasoconstriction, bien que selon les modèles expérimentaux des réactions inverses aient été observées. L'ATP induit aussi dans ces cellules une prolifération cellulaire, notamment par l'activation de la voie MAPK (Kunapuli et Daniel, 1998). Enfin, il a été montré dans divers modèles *ex vivo* que l'agression vasculaire pouvait être responsable d'une vasoconstriction par relargage d'ATP. Ce résultat est en contradiction avec les données précédemment citées, mettant en évidence la complexité des effets de l'ATP selon le modèle utilisé.

Cette variabilité de la réponse vasculaire à l'ATP provient de la très grande diversité des récepteurs, présents à la membrane des cellules vasculaires. Selon le lit artériel concerné, plusieurs dizaines de profils différents combinant des sous-types de récepteurs, certains responsables d'une réponse vasodilatatrice, d'autres d'une réponse vasoconstrictrice. Il en est de même pour l'induction de la prolifération cellulaire (variable selon l'artère, et le(s) type(s) de récepteur(s) étudié(s)). La combinaison de récepteurs à la surface des cellules vasculaires varie selon le rôle physiologique de l'artère sous-jacente, et permet ainsi une grande spécificité de la réaction artérielle (Burnstock et Ralevic, 2013; Phillips et al., 1998). Burnstock et Ralevic ont récemment effectué une revue complète de l'effet de la signalisation purinergique dans la physiopathologie

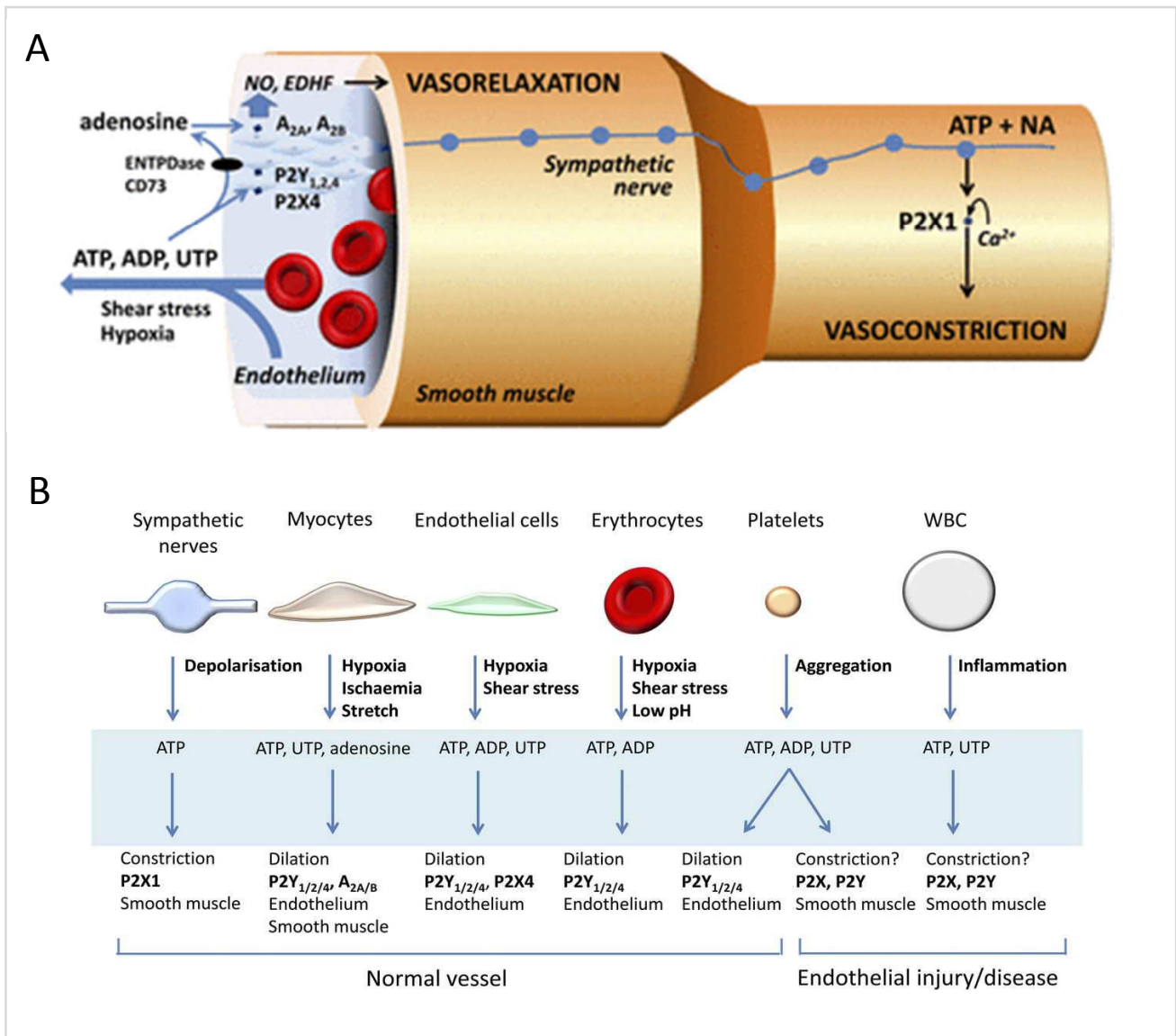


Figure 40: multiplicité des effets du relargage de l'ATP selon la source et l'état artériel. A: Endothélium versus terminaisons sympathiques: en conditions physiologiques, l'ATP active les récepteurs endothéliaux et génère un phénomène de vasodilatation. La stimulation par les nerfs sympathiques (efférents) génère une vasoconstriction.

B: différentes voies de stimulation des récepteurs P2X et P2Y selon l'état physiologique ou pathologique (agression endothéliale) vasculaire.

D'après Burnstock et Ralevic, 2013.

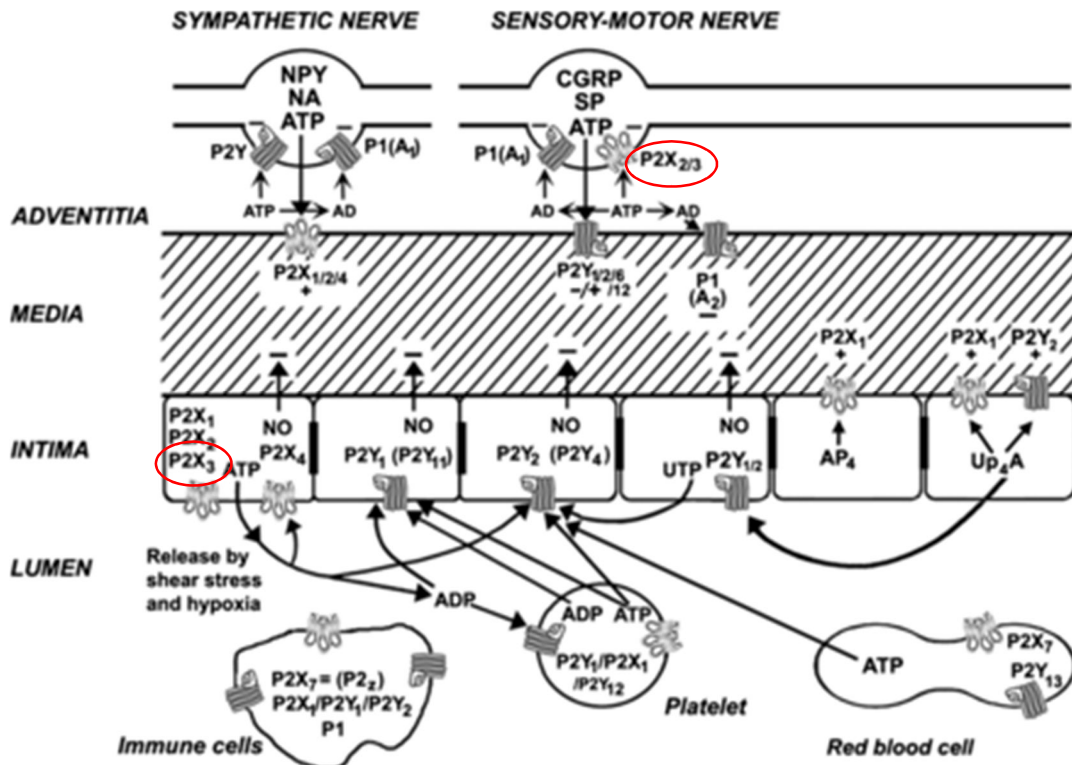


Figure 41: Distribution des récepteurs P2X et P2Y dans le tissu artériel: nerfs périphériques, cellules musculaires lisses, endothélium, et plaquettes. P2X3 est présent dans les terminaisons des nerfs afférents, et de l'endothélium. D'après Burnstock et Ralevic, 2013.

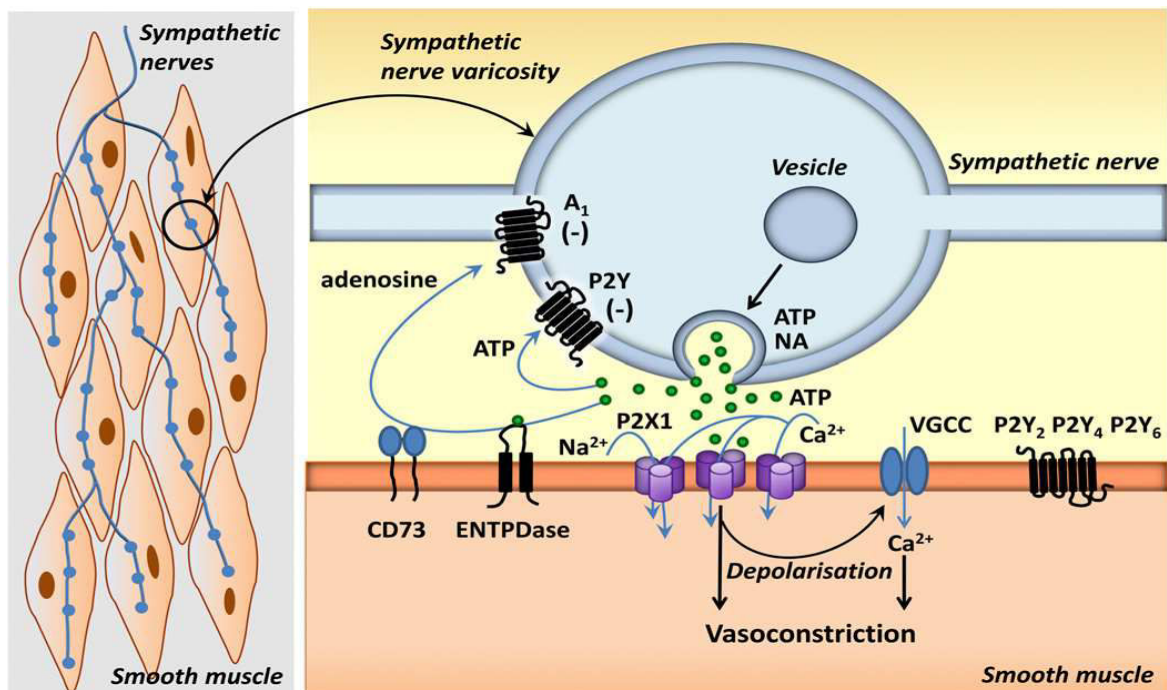


Figure 42: Effet vasoconstricteur des récepteurs P2X localisés sur la membrane des CML. D'après Burnstock et Ralevic, 2013.

vasculaire, et démontré que cette régulation du tonus vasculaire était dépendante de deux sources dont les actions sont contraires (Figures 40 et 41) : les cellules endothéliales et les terminaisons nerveuses périvasculaires, illustrant bien la diversité potentielle des effets vasculaires de l'ATP (Burnstock et Ralevic, 2013). Les récepteurs membranaires sont classés en canaux cationiques (P2X) d'une part, et récepteurs à 7 domaines transmembranaires couplés aux protéines G (P2Y) d'autre part.

Il existe 7 membres de la famille des récepteurs P2X (P2X1 à 7). Ce sont des canaux ioniques ATP dépendants, non sélectifs, homo- ou hétéro-dimériques, qui génèrent une entrée rapide de calcium en réponse à l'ATP, responsable d'une vasoconstriction dans la plupart des cas (figure 42).

En regard des récepteurs P2X1, P2X2 et P2X4, peu de données existent sur la fonction de P2X3 : c'est un canal cationique perméable au calcium et au sodium, fonctionnel sous la forme d'un homodimère P2X3/X3 ou hétérodimère P2X2/X3. Il a été identifié dans l'endothélium de plusieurs types artériels : à faible taux dans l'aorte et l'artère mésentérique (Phillips et al., 1998), et à un taux plus élevé dans l'artère rénale, pulmonaire, et le système vasculaire thymique. Un exemple illustre bien la complexité de la composition d'un lit artériel en récepteurs : dans l'artère mésentérique P2X1 et 4 sont fortement exprimés, P2X3 et 5 plus faiblement. P2X1 et dans une moindre mesure P2X3 sont responsables d'une vasoconstriction induite par la stimulation électrique qui génère une réponse adrénargique (Phillips et al., 1998). Dans l'artère pulmonaire, les récepteurs P2X3 sont particulièrement présents dans les terminaisons nerveuses périvasculaires. Ils sont activés par le relargage de l'ATP en réponse à l'étirement des artères pulmonaires, et induisent une vasoconstriction (Burnstock, 2007). De même P2X3 neuronal a un rôle vasoconstricteur dans l'artère fémorale : chez le rat, l'occlusion de l'artère fémorale induit, lors de l'exercice musculaire une élévation du tonus vasopresseur et sympathique local. Il a été démontré que cette réponse était secondaire à l'augmentation des récepteurs P2X3 dans les neurones afférents de l'artère rénale, induite par l'occlusion artérielle (Liu et al., 2011a).

Le rôle de P2X3 a été étudié dans l'artère rénale : les récepteurs P2X1 et 3, et P2Y1,2,6 y sont exprimés. Les récepteurs P2X sont essentiellement exprimés par les cellules musculaires lisses, et induisent une vasoconstriction secondaire à un relargage de l'ATP d'origine neuronale. Les récepteurs P2Y sont plutôt exprimés dans l'endothélium, où ils induisent une vasodilatation par relargage du NO (à l'exception de P2Y6, vasoconstricteur, Vonend et al., 2005).

Enfin, P2X3 a été impliqué dans la migraine et les phénomènes de sensibilisation à la douleur : il est exprimé dans les terminaisons nerveuses des artères cérébrales, où sa surexpression serait responsable d'une hypersensibilité douloureuse aux phénomènes de vasodilatation notamment (Burnstock, 2001; Fabbretti et al., 2006). Un modèle intéressant d'étude des phénomènes douloureux induits par des vibrations chez le rat a mis en évidence l'aspect biphasique de la réponse douloureuse : l'endothéline libérée par l'endothélium

soumis au stimulus mécanique induit un phénomène douloureux aigu. La réponse chronique, liée à une sensibilisation par l'endothéline est secondaire à la réponse de l'hétéromère P2X2/3 des terminaisons nerveuses périvasculaires au relargage de l'ATP endothélial (figure 43, Joseph et al., 2013). Il est surprenant de constater la proximité de ce phénomène avec le HAVS, comportant une dysplasie acquise des artères digitales. Il existe un modèle murin KO de *P2rx3* dont le phénotype artériel n'a pas été exploré.

Les récepteurs P2Y induisent une transduction du signal par liaison aux protéines G. Comme vu précédemment, ils présentent aussi une diversité de rôles, vasoconstricteurs ou dilatateurs, bien qu'ils soient globalement considérés comme des vasodilatateurs exprimés par l'endothélium (Burnstock et Ralevic, 2013). Nombre d'entre eux ont été impliqués dans des phénomènes de prolifération des CML (figure 44) impliquant les voies MAPK, RAS, et PI3K. Cependant, peu d'informations existent sur l'implication de P2Y5 dans la biologie vasculaire : c'est un récepteur à l'ATP et à l'acide lysophosphatidique (LPA), molécule à action mitogénique dérivée des phospholipides. P2Y5 est exprimé dans l'endothélium des cellules HUVEC, où il induit une contraction cellulaire en réponse à une stimulation par le LPA (Yanagida et al., 2009). Il a été identifié comme responsable de la perte de l'intégrité endothéliale de la microcirculation pulmonaire et pourrait être impliqué dans l'HTAP (Ren et al., 2013). Enfin, une perte de fonction complète du gène est responsable d'une forme d'hypotrichose congénitale (Khan et al., 2011). Dans les articles se rapportant à cette pathologie, aucun phénotype vasculaire n'est rapporté (Azeem et al., 2008; Nahum et al., 2011).

Le gène *RALGDS*

RALGDS est le second gène portant une mutation perte de fonction dans la famille FDP11, avec le récepteur à PGI2. C'est un effecteur direct de RAS, de type facteur d'échange GTP/GDP agissant sur la protéine RAL, impliquée dans les phénomènes d'exocytose. Il active AKT par recrutement d'une kinase spécifique (Hao et al., 2008), et est impliqué dans la dynamique du cytosquelette indépendamment de la voie RAS (Bhattacharya et al., 2002). Des mutations gain de fonction de *RALGDS* ont été impliquées dans l'hypertrophie cardiaque (Kawai et al., 2003), mais aucune donnée n'est disponible concernant son rôle potentiel dans la biologie vasculaire.

Le gène *ARFGEF1*

ARFGEF1, aussi appelé *BIG1*, a été identifié dans la famille FDP59, au même titre que la mutation de P2Y5. Il code un facteur d'échange GTP/GDP de ARF, protéine impliquée dans le trafic vésiculaire intracellulaire. C'est un activateur des voies AKT/PI3K et MAPK (Zhou et al., 2013) qui a été impliqué dans le développement neuritique. Aucune donnée de biologie vasculaire ni de modèle murin n'est disponible pour ce gène.

Le gène *TMPRSS2*

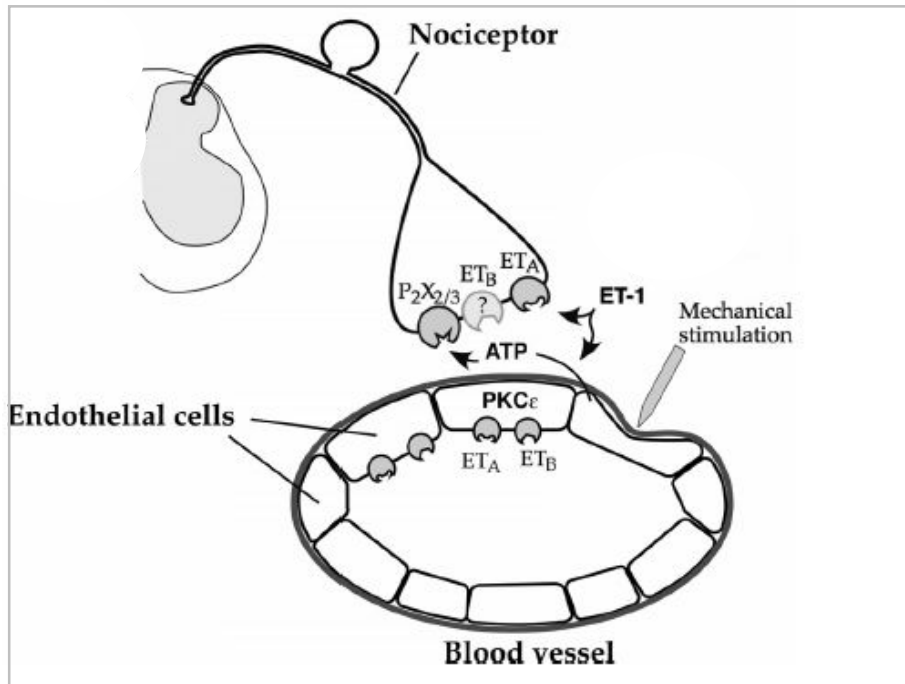


Figure 43: Effet d'hypersensibilisation à la douleur via P2X3. L'endothéline est libérée par l'intima lésée, et stimule par phénomène paracrine et autocrine les terminaisons nerveuses et les cellules endothéliales. Cette stimulation, associée au relargage d'ATP, active les récepteurs P2X2/3 qui développent un phénomène de potentialisation de la douleur. D'après Joseph et al, 2013.

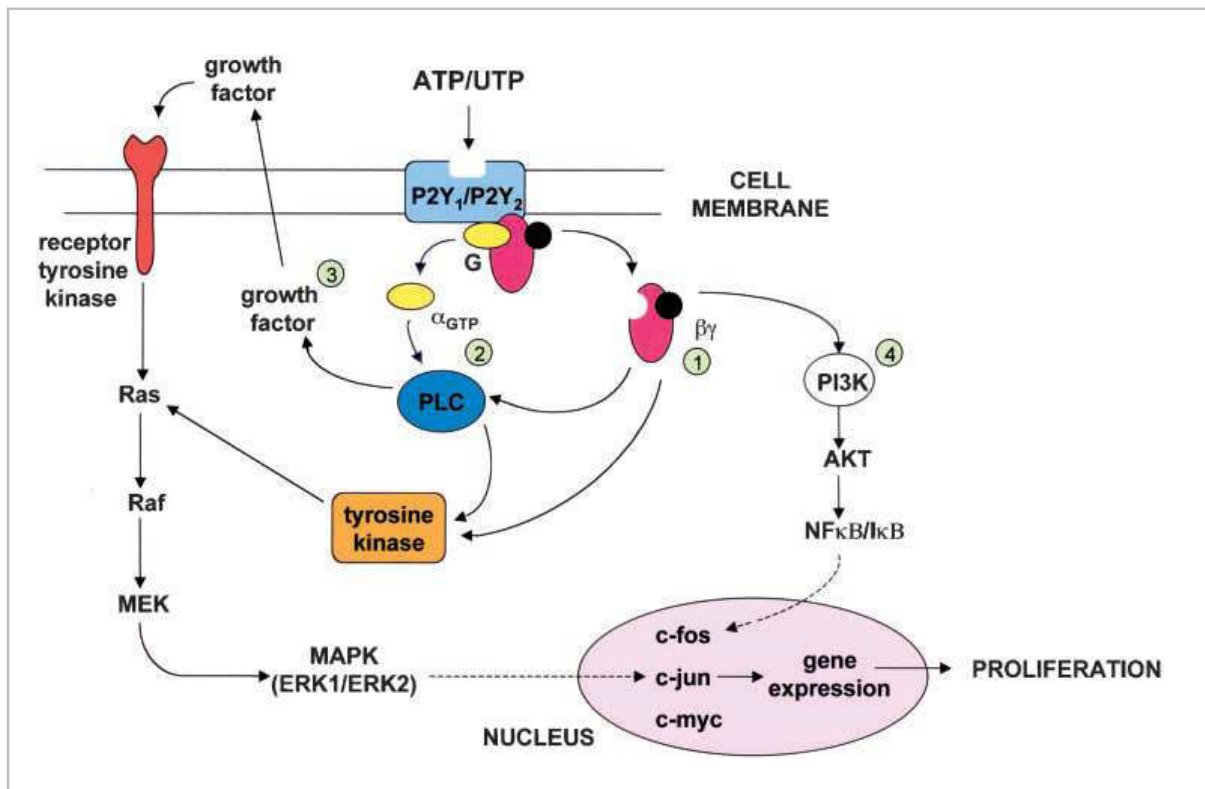


Figure 44: Effes des récepteurs à l'ATP de type P2Y sur la prolifération cellulaire. D'après Burnstock et al, 2002.

TMPRSS2 est le troisième gène impliqué dans la famille FDP59. Il code un récepteur de type tyrosine kinase (RTK), exprimé dans l'endothélium microvasculaire (Aimes et al., 2003). Son rôle n'est connu que dans le cancer de la prostate où il active la voie PI3K en réponse à un stimulus androgénique (Tilli et al., 2015). Aucune donnée de biologie vasculaire ni de modèle murin n'est non plus disponible pour ce gène.

Le gène *ECH1*

ECH1 a été identifié dans la famille FDP18. Il code l'Enoyl CoA Hydratase 1 peroxisomale, impliquée dans l'isomérisation de l'enoyle-CoA, et joue un rôle dans le métabolisme des Acyl-CoA. *ECH1* est exprimé dans les CML vasculaire où son expression est induite par la thrombine (Vendrov et al., 2006), dans un modèle cellulaire d'athérome. Un modèle murin existe, dont le phénotype n'est pas connu.

Le gène *ADAMTS14*

Comme vu précédemment, *ADAMTS14* est une métalloprotéase capable de cliver le procollagène 1 dans le cadre de la maturation des fibres collagènes. Il n'existe pas de modèle murin inactivé, et aucune autre donnée n'est disponible sur la fonction de ce gène.

11. Premières conclusions

Comme le montre ce travail réalisé dans le cadre de la DFM, maladie clairement très hétérogène et de physiopathologie peu connue, toutes les approches d'analyse d'exome afin de dégager un ou plusieurs gènes candidats ne se valent pas.

L'analyse réalisée à partir de l'ensemble des variants, incluant les faux-sens, dans ce type de projet, est longue et fastidieuse. De plus, elle se heurte rapidement à des difficultés d'interprétation et à des constats hasardeux, car elle ne permet pas d'extrapolation à partir des données bibliographiques, notamment d'inactivation de gènes ou d'inhibition pharmacologique ou transcriptionnelle. En effet, les variants faux-sens peuvent acquérir des fonctions très diverses (notamment toxiques, gain de fonction ou dominant négatif) qui ne sont généralement pas couvertes par les données de la littérature.

L'analyse des variants perte de fonction a été particulièrement développée, et à chaque étape, un travail a été réalisé sur la façon la plus pertinente d'enrichir au maximum le set final de gènes à soumettre à un test fonctionnel. La démarche correspondante ne se voulait pas exhaustive mais rationnelle. Cette analyse semblait très prometteuse, pour plusieurs raisons : elle peut se baser sur des données bibliographiques plus robustes, notamment par la manipulation d'outils de type ENDEAVOUR ou STRING qui font appel aux données

d'inactivation en modèle cellulaire ou murin. Elle peut par ailleurs exploiter de façon beaucoup plus puissante les données de fréquence de mutations perte de fonction puisque la fréquence cumulée peut être calculée et permettre d'en estimer l'impact en terme de sélection purificatrice. Cette sélection a permis d'écartier plusieurs gènes de la liste de candidats. Enfin, malgré la complexité des voies évoquées ci-dessous, elle permet une approche simplifiée au maximum de l'effet de la mutation (sauf dans le cas de protéines hétéromériques) comme un effet de dosage génique. Cependant on observe bien que dans certaines familles aucun candidat robuste ne peut être proposé par cette approche, indiquant que des mutations faux-sens seraient aussi responsables de DFM dominante. Dans les familles SED294, FDP8, FDP18, FDP125 le nombre de variations communes aux cas atteints est respectivement de 18, 14, 34, et 15, ce qui rend l'approche plus abordable famille par famille à cette échelle. Une étude plus approfondie dans ces familles des données bibliographiques de ces faux-sens, et l'analyse en modèle récessif, doit être réalisée.

Bien sûr, l'approche « variants perte de fonction » ne peut s'affranchir d'analyses fonctionnelles complémentaires, notamment dans des modèles KO hétérozygotes, ni d'une confirmation par identification d'autres patients porteurs de mutations similaires. L'existence de modèles murins KO de plusieurs des gènes identifiés, et de données cellulaires riches permettra cependant de simplifier au maximum les études fonctionnelles à mettre en place.

Cette approche est donc très séduisante, mais elle a été rendue possible uniquement par le grand nombre de cas inclus, ce qui n'est réalisable encore une fois que grâce au travail de fond réalisé pour la constitution de cette cohorte. Bien que le nombre de patients analysés en séquençage d'exome rende le travail d'analyse particulièrement complexe, il a permis d'obtenir un relativement grand nombre de mutations inactivatrices, et de dégager, en apparence, un voire plusieurs modèles physiopathologiques compatibles avec la DFM.

Un premier modèle correspondant à une forme sévère, diffuse, et probablement récessive de la maladie, impliquerait l'obscurine, protéine impliquée dans l'activation de l'appareil contractile de la CML. Étrangement, la pathologie présentée par la patiente n'est pas anévrysmale, mais sténosante. Elle ne correspond donc pas aux maladies anévrysmales de l'appareil contractile connues, à l'exception des atteintes sténosantes liées aux mutations dominantes d'ACTA2. Cependant l'atteinte anévrysmale des patients dans cette dernière pathologie est tardive en comparaison de l'atteinte sténosante qui est juvénile. La patiente FDP200-1 est âgée de moins de 30 ans, et il n'est pas impossible qu'elle développe tardivement une maladie anévrysmale. Des analyses fonctionnelles à partir de modèles KI des deux mutations identifiées en CML vasculaires seront nécessaires. Cependant, même si l'obscurine était impliquée de façon formelle par ces tests, elle n'a pas été identifiée dans la famille FDP249. Ce résultat peut être la conséquence soit d'une hétérogénéité génétique y compris dans les formes récessives de la maladie, soit d'un défaut de couverture du séquençage comme précédemment évoqué. Il peut s'agir aussi d'une confusion avec une maladie à composante inflammatoire

dans la seconde famille, certaines formes de maladie de Takayasu diffuse pouvant être impossibles à confirmer à distance du processus inflammatoire, et pouvant mimer une DFM.

Le second modèle, correspondant à un mode de transmission dominant, met en évidence une absence totale de gènes impliqués de près ou de loin dans la voie du TGF β , ce qui laisse à penser que la DFM n'est pas une maladie de cette voie, bien que des erreurs diagnostiques auraient pu le suggérer. Elle semble plus proche sur le plan physiopathologique des maladies sténosantes héréditaires comme le Moya-Moya lié à la voie du NO. Bien que cette dernière pathologie soit récessive, la similitude des lésions vasculaires et des processus physiopathologiques sous-jacents à un défaut de *GUCY1A3*, du récepteur IP et du facteur CREM (lui-même induit par la voie PGI₂/IP) chez la souris est frappante. Les cas atteints des deux familles portant ces variants hétérozygotes sont peu nombreux, et il est probable qu'il existe de nombreux porteurs sains apparentés. Ce dernier point soulève la question du modèle génétique le plus adapté à ces données : La faible proportion d'atteints pourrait s'expliquer par un phénomène de « second hit » somatique, voire épigénétique, chez certains individus, seuls à développer la maladie dans leur famille. Cependant, cette hypothèse n'est pas en accord avec les données épidémiologiques qui mettent en avant le sexe féminin et le tabac comme facteurs de risque majeurs. Aucun élément issu de l'analyse d'exome ni dans la littérature en rapport avec ces gènes n'a permis d'étayer cette hypothèse.

L'implication possible des récepteurs aux bases puriques semble intéressante, notamment par la présence de mutations perte de fonction dans 2 récepteurs appartenant à cette classe parmi 10 gènes sélectionnés, ce qui constitue un enrichissement ontologique significatif. Le modèle de pathologie acquise qu'est le HAVS semblait prometteur en terme de candidats. En réalité, aucun gène des récepteurs α adrénergiques n'a été impliqué, mais de façon indirecte ces deux pathologies pourraient être liées par les récepteurs P2X₃ voire P2Y₅. L'hypersensibilisation à la douleur induite par P2X₃, sa localisation au niveau des terminaisons nerveuses et des cellules musculaires lisses artérielles, son rôle vasoconstricteur, et probablement prolifératif (par extrapolation des autres P2X) révéleraient un lien très clair entre ces deux pathologies, confirmant l'hypothèse mécanique évoquée par de nombreux auteurs.

La confirmation du rôle des récepteurs aux bases puriques dans la DFM nécessitera, à mon sens, une analyse dépassant P2X₃ et P2Y₅, ne serait-ce qu'à cause de l'existence de formes hétéromériques des canaux P2X₃. De plus, si le rôle de P2X₃ et de P2Y₅ dans la DFM était confirmé, la combinatoire complexe de tous ces récepteurs en fonction du lit artériel rend clairement possible l'implication d'autres isoformes dans la pathologie.

Enfin, il est surprenant de constater que IP et P2X₃ sont tous deux impliqués dans la migraine, mais dans un sens inverse à celui attendu dans la DFM : ils sont tous deux surexprimés dans les vaisseaux et/ou les neurones afférents de patients migraineux, et sont considérés comme des « coupables » dans la physiopathologie de la

crise migraineuse. Le lien physiopathologique ne semble pas évident entre leur inactivation et le phénotype migraineux chez les patients.

L'utilisation des outils de priorisation dans la dernière étape de ce travail n'a pas clairement permis de proposer *ab initio* un processus biologique cohérent, mais a permis de cibler plus rapidement l'analyse bibliographique sur les candidats les plus évidents. Ces outils ne permettent pas de se passer d'une analyse bibliographique approfondie, mais les données obtenues sur chacun des candidats de la dernière section confirment bien l'impression générale que les 4 gènes *IP*, *CREM*, *P2RX3* et *P2RY5* sont les meilleurs candidats dans la pathologie.

Enfin, l'identification d'une patiente porteuse d'une mutation *SMAD3* n'est pas surprenante, et permet de réorienter le diagnostic. Ce résultat pose la question des critères diagnostics de la DFM, qui, bien qu'ils soient conservateurs pour la cohorte française, sont clairement insuffisants pour identifier l'entité clinique que représente la DFM multifocale « idiopathique ». La DFM doit donc rester un diagnostic d'élimination. Ce résultat nous a amené à réévaluer les cas de la cohorte française afin d'identifier un sous-groupe avec anévrysmes associés notamment, ou avec atteinte squelettique (critère plus difficile à extraire de la base de données de patients). Nous avons identifié une cinquantaine de patients porteurs d'anévrysmes dans notre cohorte, qui feront l'objet d'un séquençage systématique des gènes d'anévrysmes et de dissections artériels dans le cadre du diagnostic.

L'identification d'un cas familial de mutations *COL4A2*, ne ségrégeant pas avec le phénotype est plus probablement de l'ordre de la découverte fortuite (*unsolicited finding*), et pose la question des découvertes de ce type dans toute analyse d'exome. A ce titre, aucune recherche de mutation dans les gènes dits « actionnables », de prédisposition aux cancers (hors voie RAS et AKT) ou de maladie neurodégénérative n'ont été recherchés.

Chapitre 5

Discussion générale et perspectives

Les stratégies développées utilisant le séquençage d'exome dans mon travail de thèse sont diverses, car elles ont été adaptées à des contextes de maladies très différents :

L'analyse génétique de la DHS a été un succès pour plusieurs raisons :

Tout d'abord cette pathologie présente clairement une certaine homogénéité génétique : le locus *PIEZO1* (1-q23-24) était présent dans l'ensemble des 10 familles analysées par liaison génétique par Grootenboer (Grootenboer et al., 2000), locus aussi partagé par la famille canadienne étudiée par Houston (Houston et al., 2011). Un seul second locus a été identifié dans cette pathologie par Andolfo et coll. (Andolfo et al., 2013), correspondant au gène *ABCB6*.

Cette relative homogénéité génétique a rendu possible une confirmation en cohorte de taille modeste. Seuls 11 cas isolés et 2 familles supplémentaires étaient disponibles, mais le taux de détection était en rapport avec l'homogénéité sous-jacente puisque seuls 3 cas sporadiques ne portaient pas de mutation dans *PIEZO1*. Cette caractéristique de la DHS a permis d'accumuler rapidement des arguments en faveur de la pathogénicité de *PIEZO1*, et les tests fonctionnels réalisés n'ont été réellement nécessaires que pour affirmer le type d'effet mutationnel supposé (gain de fonction).

Ensuite, *PIEZO1* se révélait, lors de l'analyse des données d'exome, comme un bon candidat : c'est un canal cationique présentant des propriétés de mécanorécepteur parfaitement en accord avec la fuite de potassium et l'accutisation des symptômes lors de stress physiques, décrits dans la DHS. Cette situation est relativement confortable mais pas toujours rencontrée dans l'approche de type exome, rendant probablement compte d'une certaine proportion des échecs rencontrés. A titre d'exemple, Huebner et coll. ont identifié une mutation perte de fonction dans deux familles indépendantes atteintes de surdité postlinguale liée à l'X, dans le gène *SMPX*. Ce gène, bien qu'exprimé dans la cochlée, y a une fonction inconnue. Aucune donnée bibliographique ne permettait d'extrapoler son rôle dans l'oreille interne : il a été initialement cloné à partir du muscle squelettique, où il code une petite protéine de 88 acides aminés aux propriétés mécanosensibles, fortement induite *in vivo* par l'étirement passif musculaire. Cette protéine fait le lien entre sarcomère, sarcolemme, et matrice extracellulaire. De façon surprenante, les individus porteurs des mutations perte de fonction hétéro- ou hémizygotes ne présentent aucun phénotype musculaire, de même que le modèle murin correspondant, mais uniquement un phénotype sensoriel (Huebner et al., 2011). Dans cet exemple, l'affirmation de la causalité de ce gène vient essentiellement de l'identification d'une perte de fonction dans deux familles indépendantes, rendant cette association univoque malgré l'absence d'argument physiopathologique.

Enfin, le succès du clonage de *PIEZO1* vient aussi de l'accès à un endophénotype fiable, spécifique. La DHS étant une maladie à pénétrance incomplète, la numération sanguine et le ionogramme ne sont clairement pas

suffisants pour identifier les porteurs du gène de susceptibilité. Il n'est pas certain que l'intervalle familial ait pu être réduit de la sorte si tous les cas génotypés de la famille VA n'avaient pas pu être classés formellement en porteurs ou sains grâce à l'ektacytométrie. L'identification de *PIEZO1* dans la famille VA, ainsi que dans les familles étudiées par l'équipe de Zarichansky et coll. qui ont aussi utilisé l'ektacytométrie en aurait été retardée.

C'est toute la difficulté de l'analyse génétique dans la DFM, où aucun endophénotype fiable n'existe à ce jour, et où l'angioscanner, et *a fortiori* l'angiographie ne se justifient pas chez les apparentés sains. L'échotracking utilisé dans notre équipe constitue un endophénotype tellement peu spécifique et sensible, qu'il n'a pas été utilisé pour l'analyse de ségrégation des mutations identifiées. De fait, cette dernière n'a pas été réalisée et constitue un facteur très limitant dans l'étude de la DFM. Il est cependant envisageable d'identifier un endophénotype de type sérique, compte-tenu des hypothèses physiopathologiques dégagées par l'analyse d'exome. L'identification d'un taux élevé d'endothéline sérique ou salivaire dans l'HAVS, ou de transcrite circulant du récepteur à la prostacycline dans la migraine, constituent une piste intéressante pour rechercher un type de marqueur sérique associé à la DFM. L'identification de porteurs sains rendrait les analyses de ségrégation très puissantes compte-tenu de la faible pénétrance de la maladie, notamment chez les hommes.

Ces trois éléments, homogénéité génétique, identification d'un très bon candidat et existence d'un endophénotype exploitable constituent les bases de cette réussite.

L'analyse génétique de l'AAA s'est révélée plus difficile. On sait par définition que les formes Mendéliennes de maladies communes présentent une hétérogénéité génétique majeure, bien plus importante que la DHS. Dans l'AAA, les quatre loci identifiés (AAA1 à 4) ne correspondent pas au locus *APOL3*, et ceci est confirmé par les résultats du séquençage Sanger réalisé dans la cohorte de 350 cas familiaux et 1 600 cas sporadiques. Les résultats du séquençage montrent que parmi les 6 variants privés (très rares) identifiés, un seul l'a été dans un cas familial, les 5 autres correspondant à des cas sporadiques.

La présence de variants d'*APOL3* dans les cas sporadiques ne constitue pas un résultat incohérent en soi : dans l'hypertension artérielle pulmonaire (HTAP), les mutations du gène *BMP2* sont retrouvées dans 75% des formes familiales et 15-20% des formes sporadiques (Machado et al., 2009). Bien que dans ce cas *BMP2* soit un gène majeur d'HTAP, cet exemple démontre la faible pénétrance d'un grand nombre des variants identifiés dans les cas sporadiques, ce qui peut être le cas des variants *APOL3* identifiés dans cette cohorte.

Cependant dans la famille SE, la pénétrance de l'AAA semble complète. Ce résultat est peut-être biaisé par la grande taille de la fratrie analysée et la présence importante de facteurs de risque cardiovasculaires, précipitant le phénotype dans la famille. A ce titre il est intéressant de rappeler que le seul cas non pénétrant pour la mutation *APOL3* est une femme diabétique présentant donc 2 facteurs protecteurs de la maladie.

Trois éléments de cette étude sont en faveur de l'implication d'*APOL3* dans l'AAA : La forte liaison de la mutation avec le phénotype dans la famille SE (Z-score 4,14, $p=2.10^{-5}$), l'identification de variants très rares dans une grande cohorte de sujets atteints, et l'association d'un variant fréquent avec l'absence d'AAA identifiée par SKAT-O (OR=0,74; IC= 0,56-0,96; $p=0,026$).

Cependant le niveau de preuve apporté par ces résultats est modéré et doit faire l'objet d'une étude d'association sur les variants très rares d'une part, et d'une réplication des résultats sur les variants rares d'autre part. Le groupe de témoins proposé par nos collaborateurs (Dr Matt Bown) constitue un excellent contrôle pour ces études puisque les individus ne présentent pas d'AAA. Compte tenu de la prévalence élevée de la maladie chez l'homme âgé de plus de 65 ans (de l'ordre de 5-10% comme vu précédemment), ces témoins sont susceptibles de porter plus de variants protecteurs, et/ou moins de variants délétères que la cohorte séquencée. Cette analyse présente des similitudes avec l'approche « phénotypes extrêmes » développée dans l'introduction de ce manuscrit.

Dans cette étude, SKAT-O a été d'un apport intéressant malgré son utilisation dans des conditions méthodologiques non optimales. La comparaison de données issues de séquençage de type Sanger, en comparaison de données de génotypage SNP, constitue un biais certain, notamment en termes de taux de faux positifs et faux négatifs. Cet aspect va être corrigé par l'utilisation des données de génotypage de cette même cohorte mises à disposition par nos collaborateurs. L'utilisation de SKAT-O a constitué un avantage certain par rapport aux méthodes plus rudimentaires et unidirectionnelles comme CAST. Il n'est pas certain que l'effet protecteur du variant p.Val232Met, qui semble modéré, aurait été identifié par ces méthodes, notamment parce que le sens de l'effet est inattendu (effet protecteur). SKAT-O présente malgré tout, comme les autres tests statistiques d'association de variants, une limite majeure à mon sens : parmi les 15 variants rares et fréquents identifiés dans la cohorte par séquençage, seuls 9 d'entre eux ont pu être utilisés. Ceci est dû au fait que l'étude d'association ne peut pas tester un variant absent de l'un des deux groupes (cohorte AAA ou SUVIMAX dans notre cas). Ceci pose un problème majeur car si un variant présente un effet délétère ou protecteur fort, on s'attend à ce que sa fréquence tende vers 0 dans un des deux groupes, bien qu'on ne puisse pas tester cette hypothèse. A ce titre, les variants perte de fonction d'*APOL3* sont fréquents dans la population : 23% des africains dont plus de 2% d'homozygotes (correspondant à un seul allèle p.Gln58*), 0,5% des caucasiens, et 1% des asiatiques et sud-américains (source EXAC). Il est possible que la perte de fonction constitue un effet protecteur de l'AAA, dans la mesure où la mutation identifiée dans la famille SE semble liée à un effet de dérégulation, voire de gain de fonction. La fréquence allélique cumulée chez les caucasiens impliquerait d'identifier 4 allèles perte de fonction parmi les 1 800 individus séquencés, et aucun n'a été identifié. La significativité de ce résultat ne peut malheureusement pas être testée pour les raisons sus-citées : l'absence des allèles correspondants les exclut nécessairement d'un test d'association. L'analyse de certains

variants perte de fonction par Exome-Chip dans la population contrôle appariée permettra peut-être d'apporter plus d'informations sur ce point. Enfin, la fréquence très élevée des pertes de fonction d'*APOL3* en population africaine va à l'encontre d'un éventuel effet protecteur, dans la mesure où la prévalence de l'AAA est similaire entre africains et caucasiens (voir chapitre 3).

L'identification d'un variant codant d'*APOL3* à effet protecteur dans l'AAA n'est pas surprenant : il existe de façon intéressante une surreprésentation de variants à effet protecteur dans le champ des maladies cardiovasculaires. Ceci est particulièrement vrai pour les gènes ayant trait au métabolisme lipidique, comme *PCSK9*, *APOC3*, ou *ABCA1* (Cohen et al., 2005, 2004, 2006; TG et HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute et al., 2014). Ces données sont peut-être cependant simplement le reflet d'une analyse plus systématisée à la recherche de cet effet dans ce domaine.

L'identification du mutant p.Arg377Trp dans l'AAA doit maintenant être suivie de son analyse fonctionnelle. Les données physiopathologiques et bibliographiques obtenues sont en faveur d'un effet gain de fonction, et non perte de fonction de la mutation, comme vu précédemment. Cet aspect implique que les tests fonctionnels soient réalisés après mutagenèse dirigée, par exemple par système CRISPR, mais un certain nombre d'outils de biologie moléculaire et cellulaire comme l'inactivation (par RNAi, oligos antisens, composés pharmacologiques...), ou l'interruption du gène sont de fait d'un apport limité. La transgénèse, bien qu'elle présente des limites méthodologiques importantes, pourrait être une alternative possible. En effet l'induction de la mort cellulaire dans les modèles mutants *APOL1* semble liée à sa dérégulation (Lan et al., 2015), et l'ajout exogène d'un mutant *APOL3* pourrait suffire à induire une dérégulation de la fonction endogène de cette protéine. Par ailleurs, analyser séparément et conjointement ce variant avec le variant à effet protecteur, si ce dernier était confirmé, apporterait des éléments intéressants sur son rôle et le fonctionnement du pore d'*APOL3*.

Le type cellulaire dans lequel réaliser ces analyses fonctionnelles pose aussi question : l'analyse en cellules endothéliales est en cours. S'il s'avère cependant que l'effet cellulaire et tissulaire du mutant *APOL3* est majoritairement lié à son rôle dans les cellules dendritiques, le type d'analyse à mettre en place sera plus délicat. Une étude de l'effet du variant en cellules dendritiques dans le cadre d'une maladie artérielle chronique impliquant une communication entre de nombreux types cellulaires sera plus complexe. Elle nécessitera probablement d'analyser des tissus artériels frais de patients.

L'identification de ce variant, et la confirmation de son rôle dans l'AAA par tests fonctionnels pourrait permettre d'identifier de nouvelles pistes physiopathologiques et surtout thérapeutiques de la maladie, centrées sur les phénomènes d'immunité humorale et cellulaire antivirale, l'apoptose, et l'autoimmunité

correspondant aux hypothèses développée dans la littérature (Gredmark-Russ et al., 2009; Gregory et al., 1996; Hinterseher et al., 2012; Ocana, 2003; Xia et al., 1996).

Cependant, l'étude menée ici ne permet pas à ce jour, d'apporter des éléments d'amélioration sur l'héritabilité cachée dans la pathologie. Les résultats présentés ici montrent qu'*APOL3* n'est clairement pas LE gène majeur de la forme dominante d'AAA, ce qui en première analyse est assez frustrant compte tenu de l'effort de séquençage réalisé. Rappelons que l'héritabilité estimée de la pathologie est très élevée (Wahlgren et al., 2010), le risque relatif de développer un AAA d'un apparenté au premier degré étant de 8. Les résultats de cette étude tendent à alimenter l'hypothèse de l'existence nombreux variants (et gènes) privés impliqués dans les maladies communes (Tennessen et al., 2012).

Enfin, l'étude génétique de la DFM est très riche, et dans un premier temps, l'analyse systématisée telle que présentée dans le chapitre 4 n'a pas été réalisée, au profit d'une approche gènes candidats. Compte tenu du trop grand nombre de gènes candidats obtenus par cette approche, l'analyse systématique des variants perte de fonction a été envisagée secondairement. Cette approche sans hypothèse physiopathologique de départ a permis d'obtenir des résultats particulièrement intéressants, et semble pertinente dans les analyses de cohortes familiales de taille suffisante comme celle menée dans la DFM (8 familles, 26 atteints dont 21 cas familiaux).

A plus grande échelle, ce type d'approche a porté ses fruits notamment dans le domaine des maladies cardiovasculaires : Li et coll. ont utilisé des données de séquençage d'exome de plus de 8 500 individus. Ils ont croisé les résultats concernant les mutations perte de fonction avec divers types de données cliniques, portant notamment sur les facteurs de risque cardiovasculaires. Cette approche a permis de confirmer le rôle de *PCSK9* et d'*APOC3* déjà connus d'une part, et d'identifier 8 nouveaux loci associés entre autres au diabète de type 1 et à la magnésémie (Li et al., 2015). Une étude portant sur la population finlandaise (Lim et al., 2014) a identifié plusieurs dizaines de variants perte de fonction enrichis dans cette population en comparaison des européens non finlandais. Une étude d'association entre ces variants et une soixantaine de phénotypes parmi plus de 30 000 finlandais a révélé que 5 des gènes et pertes de fonction correspondants étaient associés à un phénotype. Parmi eux, *LPA* codant la lipoprotéine a, a été confirmé comme facteur impliqué dans les maladies cardiovasculaires. Ce type d'étude pourrait très clairement, par ailleurs, s'appliquer à l'AAA et permettre d'identifier de nouveaux gènes impliqués dans cette pathologie.

Ces études ont l'avantage d'être menées à partir de données génomiques, donc à grande échelle et sans hypothèse de départ. Dans le cadre d'une approche restreinte, ou impliquant une hypothèse physiopathologique, il existe un réel risque de mauvaise interprétation des données face à un mutant perte de fonction : Elsayed et coll. ont ainsi identifié par cartographie d'homozygotie dans une forme de surdité

recessive une mutation perte de fonction homozygote dans le gène *AHI1*. Des pertes de fonction situées dans la partie terminale de ce gène sont responsables d'un syndrome polymalformatif et neurodéveloppemental sévère, le syndrome de Joubert. Une analyse familiale plus poussée a démontré l'absence de ségrégation de la mutation avec la surdité familiale (Elsayed et al., 2015). Ainsi, une mutation plus en amont dans ce gène n'est responsable d'aucun phénotype identifiable, contrairement aux mutations situées en partie terminale du gène. Il est probable que dans le cas du syndrome de Joubert les mutations n'induisent pas de destruction des ARNm aberrants par le système du NMD (Nonsense-mediated mRNA decay). Cette situation illustre bien qu'à chaque étape d'une étude génétique, n'incluant par ailleurs pas nécessairement une analyse d'exome, nous tenons pour acquis des phénomènes biologiques fréquents comme le NMD. Nous ne les vérifions pas, mais cet exemple montre qu'accumuler des hypothèses non vérifiées et considérées comme des acquis peut avoir des conséquences majeures dans l'interprétation des résultats. D'une façon générale, cela démontre que toute mutation perte de fonction doit être confirmée par analyse de ségrégation et tests fonctionnels, notamment portant sur le NMD. Dans le cadre des maladies vasculaires rares j'ai rencontré ce type de situation en étudiant une famille consanguine présentant une maladie vasculaire rare anévrysmale et disséquante par séquençage d'exome : une mutation homozygote perte de fonction a été identifiée chez le cas atteint, absente chez les deux apparentés sains analysés. Le gène correspondant, *HTR7*, codant le récepteur de type 7 à la sérotonine était un très bon candidat. Il est exprimé dans les CML, et induit une vasodilatation secondaire à la stimulation par la sérotonine relarguée par les plaquettes. La perte de fonction s'avère par ailleurs très rare dans les populations d'EXAC. En dernier lieu, l'analyse familiale a révélé que deux membres non atteints de la fratrie étaient homozygotes pour le variant, le rendant très peu probablement causal.

Cependant la recherche de pertes de fonction porte souvent ses fruits, comme cela l'a été démontré dans l'autisme (O'Roak et al., 2014; Toma et al., 2014) révélant un modèle non Mendélien d'accumulation de mutations *de novo* et de pertes de fonction. Dans la DFM, un tel phénomène d'accumulation n'est pas exclu, puisque dans les familles FDP11 et FDP59 notamment, plusieurs gènes ont résisté à la sélection systématique réalisée sur les variations perte de fonction. Dans le cas de FDP11 par exemple, il est possible que l'inactivation du récepteur à la prostacycline et de *RALGDS*, effecteur de la voie RAS, soient responsables conjointement du phénotype. De même dans la famille FDP59, *ARFGEF1*, facteur d'échange GTP/GDP activant la voie AKT pourrait avoir un effet cumulé avec *P2RY5*, récepteur membranaire à l'ATP activant le système contractile. Dans ces familles présentant en apparence une transmission dominante de la maladie, un oligogénisme sous-jacent pourrait expliquer le faible nombre de cas atteints par famille. Seuls les porteurs de ces deux mutations seraient atteints, les individus porteurs d'une ou aucune mutation seraient sains. L'analyse de ségrégation réalisée dans ces deux familles sur les rares apparentés disponibles sans signes cliniques de DFM donne des résultats concordant avec ce modèle. Cependant cette analyse est très limitée dans la mesure où seul le

phénotype clinique a été pris en compte, pour les raisons évoquées précédemment. De plus, aucun résultat significatif ne peut être obtenu par cette approche compte tenu du faible nombre d'individus concernés.

Ce modèle oligogénique ne rend cependant pas compte de l'excès de femmes atteintes dans les familles. En modèle dominant simple ou oligogénique, les mutations responsables de la DFM devraient être théoriquement présentes chez un plus grand nombre d'hommes non atteints, puisque la pénétrance y est moindre. Aucune tendance de ce type n'a pu être dégagée à partir des mutations perte de fonction des 4 familles essentiellement concernées (FDP11, FDP43, FDP59, et FDP143), leur taille étant trop modeste pour en tirer des conclusions robustes.

D'une façon générale, l'approche centrée sur les pertes de fonction permet une mise en place **simplifiée des tests fonctionnels**, à partir de mutants souvent disponibles en lignées cellulaires ou murines. C'est le cas par exemple de *PTGIR* ou *P2RX3*.

Si l'on se projette dans l'analyse fonctionnelle des 3 meilleurs candidats de cette analyse (*PTGIR*, *CREM*, *P2RX3*), on observe cependant que le niveau de difficulté diffère :

Pour *PTGIR*, l'étape de génération des mutants sera épargnée par l'existence du modèle murin, et de nombreux tests *in vivo*, *ex vivo* ou *in vitro* pourront être réalisés à partir de souris hétérozygotes. Seules les conditions appliquées (stress mécanique ou autre) devront faire l'objet d'un design approprié.

Ce travail se complexifie pour l'analyse de *CREM* : la mutation identifiée ne concerne que trois des nombreux transcrits alternatifs de ce gène dont on sait que certains sont inhibiteurs et d'autres activateurs de la transcription. La mutation KO en modèle murin n'est donc pas un modèle reproduisant parfaitement la mutation humaine. Il semble dans ce cas impossible de se passer d'une mutagénèse dirigée en cellules humaines reproduisant le génotype des patients. L'interprétation en sera aussi plus délicate, puisqu'aucune analyse *in vivo* ou *ex vivo* ne pourra être réalisée. Cependant les tests cellulaires à mettre en place sont très proches de ceux à réaliser dans l'analyse de *PTGIR* puisque nous avons vu que *CREM* est induit pas la stimulation de la voie des prostacyclines (Ohtsubo et al., 2007).

L'analyse fonctionnelle de *P2RX3* pose un autre type de difficulté : dans l'artère, le récepteur *P2X3* semble être porté par l'endothélium et les terminaisons nerveuses, cellules largement minoritaires de la paroi artérielle. Une caractérisation plus approfondie de la localisation de ces récepteurs, notamment dans l'artère rénale, semble nécessaire avant de mettre en place une étude dans un type cellulaire donné. La seconde difficulté méthodologique d'une analyse de *P2X3* est liée à sa forme multimérique, homomère mais surtout hétéromère *P2X2/P2X3*. Il est difficile de se représenter l'effet de la perte d'un allèle de *P2RX3* sur la répartition des récepteurs membranaires, notamment hétéromériques, et il est possible qu'une analyse plus sophistiquée impliquant un mutant *P2RX2* soit nécessaire. Il existe des souris inactivées pour les gènes de ces

deux récepteurs. Elles ne présentent pas de phénotype vasculaire superposable à celui des souris inactivées pour *Creml* ou *Ptgir*, mais ce phénotype n'a pas spécifiquement été recherché. De plus, dans le cas des récepteurs P2X et P2Y, il semble que les autres récepteurs de cette famille doivent faire l'objet d'une analyse fonctionnelle et génétique : la recherche dans la base de données du Jackson Laboratory de phénotypes murins liés à l'inactivation des autres gènes de cette famille est contributive (Sumida et al., 2010; Yamamoto et al., 2006) : l'inactivation de *P2RY9* est responsable d'une hypoplasie des cellules musculaires lisses vasculaires, tandis que l'inactivation de *P2RX4* est responsable d'une hyperplasie médiale, associée à une absence de relargage du NO, de vasodilatation et de remodelage artériel secondaire à une perturbation du flux sanguin.

Cette constatation amène à l'idée qu'une analyse génétique à la recherche de nouvelles mutations dans les gènes *PTGIR*, *CREM*, *P2RX3* et *P2RY5* dans la cohorte française de DFM doit être nécessairement associée à l'analyse d'autres gènes de la famille des récepteurs de bases puriques, dont *P2RX4* et *P2RY9*.

Enfin, l'identification de l'obscure par deux approches indépendantes, et sans hypothèse physiopathologique de départ, est très prometteuse. L'utilisation de SKAT-O a révélé des gènes candidats pertinents, comme *RFN213* et *MYLK* ($p=0,01$ pour chacun) précédemment impliqués dans les maladies artérielles sténosantes et disséquantes respectivement. A ce titre, l'association la plus significative relevée par cet outil est avec *OBSCN* ($p=0,003$), totalement inconnu dans les maladies artérielles avant cette étude. Identifier ce gène par approche de trio dans une forme sporadique de la maladie, et par étude d'association de variants rares dans la cohorte française, est très évocateur de son implication selon un modèle récessif, mais peut-être aussi dominant de la maladie. Le séquençage systématique d'*OBSCN* dans la cohorte permettrait d'étayer cette hypothèse.

L'analyse fonctionnelle des variants faux-sens identifiés dans l'obscure sera forcément plus complexe que l'analyse de pertes de fonction. Dans le cas de mutations faux-sens, nous faisons souvent une hypothèse de départ quant au rôle joué par ces mutations (perte de fonction, gain de fonction, dérégulation, effet toxique ou structurel...). En ce sens dans le cas de l'obscure, deux éléments sont intéressants à relever : tout d'abord une analyse dans EXAC portant sur le transcrit majeur révèle que 4% de la population porte une mutation perte de fonction, et 0,6% porte 2 mutations, soit une probable inactivation complète de cet isoforme. L'obscure, comme vu précédemment, présente de nombreux transcrits, qui sont tissu- cellule- et compartiment- spécifiques, et l'interprétation des données d'EXAC est assez hasardeuse dans ce contexte. On peut cependant penser que les mutations identifiées dans la famille FDP200 dérégulent cette protéine très complexe, qui porte de multiples domaines fonctionnels et notamment myosine kinase-like. De plus, la recherche de sites de modification post-traductionnels révèle que l'une des deux mutations est portée par un site très conservé, qui est acétylé chez la souris (source PhosphoSitePlus). Pour rappel, les deux mutations ne

sont distantes que de 6 acides aminés dans une protéine qui en compte plus de 9 000 pour l'isoforme majoritaire. Ces éléments permettent de restreindre partiellement les hypothèses sur l'effet des mutations identifiées, afin de mettre en place les tests fonctionnels cellulaires. Une mutagénèse en système cellulaire introduisant les faux-sens identifiés semble ici encore indispensable pour la poursuite de l'étude du rôle d'*OBSCN* dans la DFM.

En dernier lieu, il semble que **les outils de priorisation** dans l'analyse des listes de variants issus du séquençage d'exome ne soient pas encore arrivés à maturité :

L'analyse de réseaux de gènes par STRING s'est révélée relativement décevante, bien qu'elle mette en évidence un lien entre plusieurs des gènes identifiés et la voie RAS. Etonnamment, STRING n'a pas permis de relier *P2RX3* et *P2RY5*, alors qu'ils codent tous deux des récepteurs à l'ATP. Cet outil utilise en théorie les liens familiaux entre gènes. *P2RX3* étant un canal ionique et *P2RY5* un récepteur à 7 domaines transmembranaires, cette divergence de structures peut expliquer ce résultat. Il n'existe pas de dimension statistique à ce type d'analyse, contrairement à la recherche ontologique. Cet outil n'est donc qu'un indicateur imparfait permettant de gagner un peu de temps dans la recherche de liens entre gènes candidats.

L'analyse ontologique (par DAVID) s'avère plus intéressante puisqu'elle met en évidence la surreprésentation de 3 récepteurs aux neuropeptides, de façon significative (*PTGIR*, *P2RX3*, *P2RY5*). Ce résultat est d'autant plus intéressant que les récepteurs aux neuropeptides sont impliqués dans de nombreux processus douloureux, dont la migraine qui est très fréquente dans la DFM (plus de 25% des patients, Olin et al., 2012). Il pointe du doigt un lien encore mal identifié entre les phénomènes douloureux importants présentés par ces patientes (après des événements de dissection mais pas uniquement), la forte prévalence de la migraine, et le processus dysplasique. D'autre part, ce résultat renforce encore les similitudes entre la DFM et la maladie de Moya-Moya, impliquant la voie du NO. Le NO, comme vu précédemment, est aussi un neurotransmetteur, notamment exprimé par les neurones des ganglions sphéno-palatins impliqués dans l'innervation du système artériel cérébral (Uddman et al., 1999). Il est relargué dans la paroi de ces artères suite à l'entrée massive de calcium dans les terminaisons nerveuses, qui activent la NO synthase neuronale. Il s'ensuit une vasodilatation de la circulation cérébrale. De fait, le NO est aussi un coupable idéal dans la migraine, et l'association entre *NOS3* et cette maladie a déjà été démontrée par meta-analyse. La similitude des processus associés à ces deux gènes, et l'intrication possible des rôles de *P2RX3* et du NO dans la communication entre cellules nerveuses et vasculaires du système artériel ouvre des champs d'investigation passionnants, notamment dans le domaine de la migraine mais aussi des phénomènes prolifératifs liés à la resténose.

L'outil ENDEAVOUR n'a pas été exploité de façon optimale, notamment pour des raisons d'interface et de formatage des listes de gènes à analyser. De plus, la soumission d'une liste d'entraînement implique un parti pris certain sur les hypothèses physiopathologiques de la maladie étudiée, contrairement aux autres outils utilisés.

Enfin, la connaissance *a posteriori* du rôle des candidats identifiés montre bien que l'exploitation des données du catalogue de phénotypes murins (Mouse Phenotype Ontology) était incomplète : une interrogation plus formelle du MPO aurait dû permettre d'identifier *PTGIR* puisque il est associé au terme « vascular restenosis ». L'utilisation manuelle, et forcément plus erratique de ce catalogue n'a pas permis de l'identifier car elle ne se basait que sur les phénotypes des cellules musculaires lisses, les anévrismes, et les atteintes de l'artère rénale. L'outil Phive présenté dans l'introduction de ce manuscrit n'a pas été utilisé, pour des raisons de formatage de fichiers et de manipulation de l'interface web, et il est possible que son utilisation aurait permis l'identification de *PTGIR*. Une évaluation plus poussée de cet outil dans un second temps sur les familles ne présentant pas de gène candidat robuste mérite d'être entreprise.

Cependant l'analyse *in silico* réalisée sur les nombreux gènes candidats de ce troisième projet a été fastidieuse, car faite manuellement. Il n'existe pas aujourd'hui d'outil publié permettant d'analyser, par exemple, un groupe de mutations perte de fonction issu d'une liste de gènes candidats. Le calcul de la fréquence cumulée des pertes de fonction issues d'EXAC, l'identification des transcrits concernés par les mutations, l'interrogation du catalogue de phénotypes murins, la recherche d'un enrichissement dans une classe ontologique de protéines, etc... sont des étapes qui pourraient clairement être automatisées et rassemblées au sein d'une même interface. Un outil intégré permettrait de gagner du temps de façon substantielle pour se consacrer à l'étude fonctionnelle des mutations identifiées. Dans l'avenir il est probable que ce type d'outil voie le jour. Il correspond en effet à un réel besoin dans les projets de séquençage d'exome actuels, qui de fait incluent de plus en plus de maladies génétiquement hétérogènes. Dans ce type d'approche, trouver rapidement du sens dans une liste de candidats potentiels est crucial pour la réussite du projet.

D'une façon générale, le travail réalisé sur ces trois pathologies met en évidence des **éléments majeurs permettant d'améliorer les chances de réussite d'un projet de séquençage d'exome** :

Tout d'abord le **design initial de l'étude**, c'est-à-dire le choix des individus à séquencer, est très important : dans la DFM, l'analyse en trio de la famille FDP200 en est un exemple très illustratif, puisqu'en modèle récessif seul le gène *OBSCN* a été identifié par cette approche. De même, l'analyse de membres très éloignés de la famille FDP43 (présentant 1/128^{ème} de leur génome en commun seulement) a permis de restreindre le nombre de candidats de façon drastique, notamment dans les gènes porteurs de mutations perte de fonction.

Ensuite, l'application d'une **stratégie d'analyse planifiée** est très importante dans les projets de séquençage d'exome d'un nombre important de patients. Une approche famille par famille à la recherche du gène idéal ne s'est pas révélée efficace dans l'analyse de la DFM. La connaissance de différentes approches possible, incluant notamment l'exploitation de données d'ontologie de gènes ou de phénotypes n'est pas à négliger pour dégager des groupes de gènes candidats.

Les difficultés rencontrées dans l'analyse des exomes DFM viennent entre autres de l'origine ethnique des patients analysés (africaine, maori, et malgache). Pour les ethnies rares, les données d'EXAC ne sont d'aucune aide. Ceci illustre le second élément majeur d'une réussite de ce type de projet : la nécessité de pouvoir utiliser des **données de population contrôlée adaptée**. C'est pourquoi l'identification d'un variant du domaine NC1 de *COL4A2* dans une famille malgache n'a aujourd'hui que peu de poids dans l'analyse globale. L'absence de ce type de données exclut aujourd'hui les isolats géographiques de bon nombre d'études génétiques, alors qu'ils sont, de par leur structure, plus clairement concernés par les maladies héréditaires que les populations panmictiques.

Un quatrième point majeur est la nécessité absolue **d'avoir une connaissance des bases physiopathologiques de la maladie** que l'on aborde par un séquençage d'exome : dans la première partie de mon travail consacré à la DHS, la connaissance préalable d'un comportement cellulaire lié aux phénomènes d'osmose (ayant mené à l'identification d'un profil ektacytométrique) et aux traumatismes mécaniques par divers auteurs ont été d'un apport majeur. Ces données, associées à la caractérisation récente du rôle de *PIEZO1* par Coste et coll. ont permis de rapidement faire le lien entre le gène et la maladie (Coste et al., 2012).

Dans l'AAA, seule une connaissance approfondie de la physiologie vasculaire et du processus pathologique, impliquant l'inflammation comme un facteur-clé, ont permis d'identifier *APOL3* comme un bon candidat. Dans la DFM, la connaissance physiopathologique des maladies vasculaires proches a été importante pour dégager un modèle de processus pathologique. Ceci permet d'apporter des arguments de poids pour l'implication des gènes *PTGIR*, *CREM* et *P2RX3* identifiés par l'approche perte de fonction, et de proposer des tests fonctionnels adaptés aux modèles proposés.

En conclusion, la multiplicité des approches d'analyse génétique, associé à une connaissance approfondie de la maladie sont des éléments nécessaires à l'identification de nouveaux gènes de maladie dans le contexte actuel. De nombreux gènes restent à découvrir, mais les difficultés rencontrées étant de plus en plus importantes, seule une évolution des pratiques et des outils disponibles permettra la réussite de tels projets. Associer l'étude de formes Mendéliennes et sporadiques de maladies est une approche efficace, grâce à l'existence de tests statistiques appropriés qui sont en constante évolution. La baisse des coûts de séquençage

d'exome, et la sophistication des outils d'analyse *in silico* permettront de même d'identifier à l'avenir des gènes et voies impliqués dans les maladies très hétérogènes.

BIBLIOGRAPHIE

- van der Aa, E., van Montfoort, N., and Woltman, A.M. (2015). BDCA3(+)/CLEC9A(+) human dendritic cell function and development. *Semin. Cell Dev. Biol.* *41*, 39–48.
- Ackermann, M.A., Shriver, M., Perry, N.A., Hu, L.-Y.R., and Kontrogianni-Konstantopoulos, A. (2014). Obscurins: Goliaths and Davids take over non-muscle tissues. *PLoS One* *9*, e88162.
- Adie, E.A., Adams, R.R., Evans, K.L., Porteous, D.J., and Pickard, B.S. (2005). Speeding disease gene discovery by sequence based candidate prioritization. *BMC Bioinformatics* *6*, 55.
- Aimes, R.T., Zijlstra, A., Hooper, J.D., Ogbourne, S.M., Sit, M.-L., Fuchs, S., Gotley, D.C., Quigley, J.P., and Antalis, T.M. (2003). Endothelial cell serine proteases expressed during vascular morphogenesis and angiogenesis. *Thromb. Haemost.* *89*, 561–572.
- Akhurst, R.J. (2012). The paradoxical TGF- β vasculopathies. *Nat. Genet.* *44*, 838–839.
- Albuisson, J., Mocerri, P., Flori, E., Belli, E., Gronier, C., and Jeunemaitre, X. (2015). Clinical utility gene card for: Arterial tortuosity syndrome. *Eur. J. Hum. Genet. EJHG*.
- Amarenco, P., Cohen, A., Tzourio, C., Bertrand, B., Hommel, M., Besson, G., Chauvel, C., Touboul, P.J., and Bousser, M.G. (1994). Atherosclerotic disease of the aortic arch and the risk of ischemic stroke. *N. Engl. J. Med.* *331*, 1474–1479.
- Andolfo, I., Alper, S.L., De Franceschi, L., Auriemma, C., Russo, R., De Falco, L., Vallefuoco, F., Esposito, M.R., Vandorpe, D.H., Shmukler, B.E., et al. (2013a). Multiple clinical forms of dehydrated hereditary stomatocytosis arise from mutations in PIEZO1. *Blood* *121*, 3925–3935, S1–S12.
- Andreoni, K.A., Weeks, S.M., Gerber, D.A., Fair, J.H., Mauro, M.A., McCoy, L., Scott, L., and Johnson, M.W. (2002). Incidence of donor renal fibromuscular dysplasia: does it justify routine angiography? *Transplantation* *73*, 1112–1116.
- Atit, R.P., Crowe, M.J., Greenhalgh, D.G., Wenstrup, R.J., and Ratner, N. (1999). The Nf1 tumor suppressor regulates mouse skin wound healing, fibroblast proliferation, and collagen deposited by fibroblasts. *J. Invest. Dermatol.* *112*, 835–842.
- Attia, J., Thakkinstian, A., and D’Este, C. (2003). Meta-analyses of molecular association studies: Methodologic lessons for genetic epidemiology. *J. Clin. Epidemiol.* *56*, 297–303.
- Aubart, M., Gobert, D., Aubart-Cohen, F., Detaint, D., Hanna, N., d’Indya, H., Lequintrec, J.-S., Renard, P., Vigneron, A.-M., Dieudé, P., et al. (2014). Early-onset osteoarthritis, Charcot-Marie-Tooth like neuropathy, autoimmune features, multiple arterial aneurysms and dissections: an unrecognized and life threatening condition. *PLoS One* *9*, e96387.
- Ayadi, A., Birling, M.-C., Bottomley, J., Bussell, J., Fuchs, H., Fray, M., Gailus-Durner, V., Greenaway, S., Houghton, R., Karp, N., et al. (2012). Mouse large-scale phenotyping initiatives: overview of the European Mouse Disease Clinic (EUMODIC) and of the Wellcome Trust Sanger Institute Mouse Genetics Project. *Mamm. Genome* *23*, 600–610.

- Azeem, Z., Jelani, M., Naz, G., Tariq, M., Wasif, N., Kamran-Ul-Hassan Naqvi, S., Ayub, M., Yasinzai, M., Amin-Ud-Din, M., Wali, A., et al. (2008). Novel mutations in G protein-coupled receptor gene (P2RY5) in families with autosomal recessive hypotrichosis (LAH3). *Hum. Genet.* *123*, 515–519.
- Balligand, J.-L., Feron, O., and Dessy, C. (2009). eNOS activation by physical forces: from short-term regulation of contraction to chronic remodeling of cardiovascular tissues. *Physiol. Rev.* *89*, 481–534.
- Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A., and Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nat. Rev. Genet.* *12*, 745–755.
- Band, W., Goedhard, W.J., and Knoop, A.A. (1972). Effects of aging on dynamic viscoelastic properties of the rat's thoracic aorta. *Pflüg. Arch. Eur. J. Physiol.* *331*, 357–364.
- Beaurain, G., Mathieu, F., Grootenboer, S., Fiquet, B., Cynober, T., Tchernia, G., Delaunay, J., and Jeunemaitre, X. (2007). Dehydrated hereditary stomatocytosis mimicking familial hyperkalaemic hypertension: clinical and genetic investigation. *Eur. J. Haematol.* *78*, 253–259.
- Becker, J., Semler, O., Gilissen, C., Li, Y., Bolz, H.J., Giunta, C., Bergmann, C., Rohrbach, M., Koerber, F., Zimmermann, K., et al. (2011). Exome sequencing identifies truncating mutations in human SERPINF1 in autosomal-recessive osteogenesis imperfecta. *Am. J. Hum. Genet.* *88*, 362–371.
- Begum, N., Hockman, S., and Manganiello, V.C. (2011). Phosphodiesterase 3A (PDE3A) deletion suppresses proliferation of cultured murine vascular smooth muscle cells (VSMCs) via inhibition of mitogen-activated protein kinase (MAPK) signaling and alterations in critical cell cycle regulatory proteins. *J. Biol. Chem.* *286*, 26238–26249.
- Beighton, P., De Paepe, A., Steinmann, B., Tsipouras, P., and Wenstrup, R.J. (1998). Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). *Am. J. Med. Genet.* *77*, 31–37.
- Bérard, E., Sarles, J., Triolo, V., Gagnadoux, M.F., Wernert, F., Hadchouel, M., and Niaudet, P. (1998). Renovascular hypertension and vascular anomalies in Alagille syndrome. *Pediatr. Nephrol. Berl. Ger.* *12*, 121–124.
- Berthet, E., Hanna, N., Giraud, C., and Soubrier, M. (2015). A case of rheumatoid arthritis associated with SMAD3 gene mutation: a new clinical entity? *J. Rheumatol.* *42*, 556.
- Bertoli-Avella, A.M., Gillis, E., Morisaki, H., Verhagen, J.M.A., de Graaf, B.M., van de Beek, G., Gallo, E., Kruithof, B.P.T., Venselaar, H., Myers, L.A., et al. (2015). Mutations in a TGF- β ligand, TGFB3, cause syndromic aortic aneurysms and dissections. *J. Am. Coll. Cardiol.* *65*, 1324–1336.
- Bhattacharya, M., Anborgh, P.H., Babwah, A.V., Dale, L.B., Dobransky, T., Benovic, J.L., Feldman, R.D., Verdi, J.M., Rylett, R.J., and Ferguson, S.S.G. (2002). Beta-arrestins regulate a Ral-GDS Ral effector pathway that mediates cytoskeletal reorganization. *Nat. Cell Biol.* *4*, 547–555.
- Bickerstaff, L.K., Hollier, L.H., Van Peenen, H.J., Melton, L.J., Pairolero, P.C., and Cherry, K.J. (1984). Abdominal aortic aneurysms: the changing natural history. *J. Vasc. Surg.* *1*, 6–12.
- Biros, E., Cooper, M., Palmer, L.J., Walker, P.J., Norman, P.E., and Golledge, J. (2010). Association of an allele on chromosome 9 and abdominal aortic aneurysm. *Atherosclerosis* *212*, 539–542.

- Biros, E., Norman, P.E., Walker, P.J., Nataatmadja, M., West, M., and Golledge, J. (2011a). A single nucleotide polymorphism in exon 3 of the kallikrein 1 gene is associated with large but not small abdominal aortic aneurysm. *Atherosclerosis* 217, 452–457.
- Biros, E., Norman, P.E., Jones, G.T., van Rij, A.M., Yu, G., Moxon, J.V., Blankensteijn, J.D., van Sterkenburg, S.M., Morris, D., Baas, A.F., et al. (2011b). Meta-analysis of the association between single nucleotide polymorphisms in TGF- β receptor genes and abdominal aortic aneurysm. *Atherosclerosis* 219, 218–223.
- Bittles, A. (2001). Consanguinity and its relevance to clinical genetics. *Clin. Genet.* 60, 89–98.
- Blanchard, J.F., Armenian, H.K., and Friesen, P.P. (2000). Risk factors for abdominal aortic aneurysm: results of a case-control study. *Am. J. Epidemiol.* 151, 575–583.
- Blendy, J.A., Kaestner, K.H., Weinbauer, G.F., Nieschlag, E., and Schütz, G. (1996). Severe impairment of spermatogenesis in mice lacking the CREM gene. *Nature* 380, 162–165.
- Bley, K.R., Hunter, J.C., Eglen, R.M., and Smith, J.A. (1998). The role of IP prostanoid receptors in inflammatory pain. *Trends Pharmacol. Sci.* 19, 141–147.
- Bloomer, L.D.S., Bown, M.J., and Tomaszewski, M. (2012). Sexual dimorphism of abdominal aortic aneurysms: a striking example of “male disadvantage” in cardiovascular disease. *Atherosclerosis* 225, 22–28.
- Bobryshev, Y.V., and Lord, R.S. (1995). Ultrastructural recognition of cells with dendritic cell morphology in human aortic intima. Contacting interactions of Vascular Dendritic Cells in athero-resistant and athero-prone areas of the normal aorta. *Arch. Histol. Cytol.* 58, 307–322.
- Bobryshev, Y.V., and Lord, R.S. (2001). Vascular-associated lymphoid tissue (VALT) involvement in aortic aneurysm. *Atherosclerosis* 154, 15–21.
- Bobryshev, Y.V., Ikezawa, T., and Watanabe, T. (1997). Formation of Birbeck granule-like structures in vascular dendritic cells in human atherosclerotic aorta. Lag-antibody to epidermal Langerhans cells recognizes cells in the aortic wall. *Atherosclerosis* 133, 193–202.
- Boeynaems, J.M., and Galand, N. (1983). Stimulation of vascular prostacyclin synthesis by extracellular ADP and ATP. *Biochem. Biophys. Res. Commun.* 112, 290–296.
- Boileau, C., Guo, D.-C., Hanna, N., Regalado, E.S., Detaint, D., Gong, L., Varret, M., Prakash, S.K., Li, A.H., d’Indy, H., et al. (2012). TGF β 2 mutations cause familial thoracic aortic aneurysms and dissections associated with mild systemic features of Marfan syndrome. *Nat. Genet.* 44, 916–921.
- Bostrom, M.A., Boyanovsky, B.B., Jordan, C.T., Wadsworth, M.P., Taatjes, D.J., de Beer, F.C., and Webb, N.R. (2007). Group v secretory phospholipase A2 promotes atherosclerosis: evidence from genetically altered mice. *Arterioscler. Thromb. Vasc. Biol.* 27, 600–606.
- Botstein, D., White, R.L., Skolnick, M., and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32, 314–331.
- Boucher, J.M., Harrington, A., Rostama, B., Lindner, V., and Liaw, L. (2013). A Receptor-Specific Function for Notch2 in Mediating Vascular Smooth Muscle Cell Growth Arrest Through Cyclin-dependent Kinase Inhibitor 1B. *Circ. Res.* 113, 975–985.
- Boucher, P., Gotthardt, M., Li, W.-P., Anderson, R.G.W., and Herz, J. (2003). LRP: role in vascular wall integrity and protection from atherosclerosis. *Science* 300, 329–332.

- Bovenzi, M., D'Agostin, F., Rui, F., Ambrosi, L., and Zefferino, R. (2008). Salivary endothelin and vascular disorders in vibration-exposed workers. *Scand. J. Work. Environ. Health* 34, 133–141.
- Bown, M.J., Braund, P.S., Thompson, J., London, N.J.M., Samani, N.J., and Sayers, R.D. (2008). Association between the coronary artery disease risk locus on chromosome 9p21.3 and abdominal aortic aneurysm. *Circ. Cardiovasc. Genet.* 1, 39–42.
- Bown, M.J., Jones, G.T., Harrison, S.C., Wright, B.J., Bumpstead, S., Baas, A.F., Gretarsdottir, S., Badger, S.A., Bradley, D.T., Burnand, K., et al. (2011a). Abdominal aortic aneurysm is associated with a variant in low-density lipoprotein receptor-related protein 1. *Am. J. Hum. Genet.* 89, 619–627.
- Bown, M.J., Jones, G.T., Harrison, S.C., Wright, B.J., Bumpstead, S., Baas, A.F., Gretarsdottir, S., Badger, S.A., Bradley, D.T., Burnand, K., et al. (2011b). Abdominal Aortic Aneurysm Is Associated with a Variant in Low-Density Lipoprotein Receptor-Related Protein 1. *Am. J. Hum. Genet.* 89, 619–627.
- Boyden, L.M., Choi, M., Choate, K.A., Nelson-Williams, C.J., Farhi, A., Toka, H.R., Tikhonova, I.R., Bjornson, R., Mane, S.M., Colussi, G., et al. (2012). Mutations in kelch-like 3 and cullin 3 cause hypertension and electrolyte abnormalities. *Nature* 482, 98–102.
- Bradley, D.T., Hughes, A.E., Badger, S.A., Jones, G.T., Harrison, S.C., Wright, B.J., Bumpstead, S., Baas, A.F., Gretarsdottir, S., Burnand, K., et al. (2013). A Variant in LDLR Is Associated With Abdominal Aortic Aneurysm. *Circ. Cardiovasc. Genet.* 6, 498–504.
- Brunham, L.R., and Hayden, M.R. (2013). Hunting human disease genes: lessons from the past, challenges for the future. *Hum. Genet.* 132, 603–617.
- Burkland, C.E., Goodwin, W.E., and Leadbetter, W.F. (1950). The cure of hypertension by nephrectomy; a ten-year follow-up of a case. *Surgery* 28, 67–70.
- Burnstock, G. (2001). Purine-mediated signalling in pain and visceral perception. *Trends Pharmacol. Sci.* 22, 182–188.
- Burnstock, G. (2007). Purine and pyrimidine receptors. *Cell. Mol. Life Sci.* 64, 1471–1483.
- Burnstock, G., and Ralevic, V. (2013). Purinergic Signaling and Blood Vessels in Health and Disease. *Pharmacol. Rev.* 66, 102–192.
- Callewaert, B.L., Willaert, A., Kerstjens-Frederikse, W.S., De Backer, J., Devriendt, K., Albrecht, B., Ramos-Arroyo, M.A., Doco-Fenzy, M., Hennekam, R.C.M., Pyeritz, R.E., et al. (2008). Arterial tortuosity syndrome: clinical and molecular findings in 12 newly identified families. *Hum. Mutat.* 29, 150–158.
- Camilleri Jean Pierre, Berry Collin I., Bariety Jean, Fiessinger Jean Noël: les maladies de la paroi artérielle, 1987, éd Flammarion.
- Campa, J.S., Greenhalgh, R.M., and Powell, J.T. (1987). Elastin degradation in abdominal aortic aneurysms. *Atherosclerosis* 65, 13–21.
- Carella, M., Stewart, G., Ajetunmobi, J.F., Perrotta, S., Grootenboer, S., Tchernia, G., Delaunay, J., Totaro, A., Zelante, L., Gasparini, P., et al. (1998). Genomewide search for dehydrated hereditary stomatocytosis (hereditary xerocytosis): mapping of locus to chromosome 16 (16q23-qter). *Am. J. Hum. Genet.* 63, 810–816.

- Carli, P., Graffin, B., Gisserot, O., Landais, C., and De Jaureguiberry, J.-P. (2007). [Recurrence of thromboembolic disease after splenectomy for hereditary xerocytosis]. *Rev. Médecine Interne Fondée Par Société Natl. Française Médecine Interne* 28, 879–881.
- Castleden, W.M., and Mercer, J.C. (1985). Abdominal aortic aneurysms in Western Australia: descriptive epidemiology and patterns of rupture. *Br. J. Surg.* 72, 109–112.
- Cawthon, R.M., O’Connell, P., Buchberg, A.M., Viskochil, D., Weiss, R.B., Culver, M., Stevens, J., Jenkins, N.A., Copeland, N.G., and White, R. (1990). Identification and characterization of transcripts from the neurofibromatosis 1 region: the sequence and genomic structure of EVI2 and mapping of other transcripts. *Genomics* 7, 555–565.
- Chartier-Harlin, M.C., Crawford, F., Houlden, H., Warren, A., Hughes, D., Fidani, L., Goate, A., Rossor, M., Roques, P., and Hardy, J. (1991). Early-onset Alzheimer’s disease caused by mutations at codon 717 of the beta-amyloid precursor protein gene. *Nature* 353, 844–846.
- Cheng, Y., Austin, S.C., Rocca, B., Koller, B.H., Coffman, T.M., Grosser, T., Lawson, J.A., and FitzGerald, G.A. (2002). Role of prostacyclin in the cardiovascular response to thromboxane A2. *Science* 296, 539–541.
- Choke, E., Cockerill, G.W., Laing, K., Dawson, J., Wilson, W.R.W., Loftus, I.M., and Thompson, M.M. (2009). Whole genome-expression profiling reveals a role for immune and inflammatory response in abdominal aortic aneurysm rupture. *Eur. J. Vasc. Endovasc. Surg. Off. J. Eur. Soc. Vasc. Surg.* 37, 305–310.
- Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D., Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T., et al. (2015). The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *Am. J. Hum. Genet.* 97, 199–215.
- Cinar, E., Zhou, S., DeCoursey, J., Wang, Y., Waugh, R.E., and Wan, J. (2015). Piezo1 regulates mechanotransductive release of ATP from human RBCs. *Proc. Natl. Acad. Sci. U. S. A.* 112, 11783–11788.
- Cloft, H.J., Kallmes, D.F., Kallmes, M.H., Goldstein, J.H., Jensen, M.E., and Dion, J.E. (1998). Prevalence of cerebral aneurysms in patients with fibromuscular dysplasia: a reassessment. *J. Neurosurg.* 88, 436–440.
- Cohen, J., Pertsemlidis, A., Kotowski, I.K., Graham, R., Garcia, C.K., and Hobbs, H.H. (2005). Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat. Genet.* 37, 161–165.
- Cohen, J.C., Kiss, R.S., Pertsemlidis, A., Marcel, Y.L., McPherson, R., and Hobbs, H.H. (2004). Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* 305, 869–872.
- Cohen, J.C., Boerwinkle, E., Mosley, T.H., and Hobbs, H.H. (2006). Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N. Engl. J. Med.* 354, 1264–1272.
- Cohen, J.R., Keegan, L., Sarfati, I., Danna, D., Ilardi, C., and Wise, L. (1991). Neutrophil chemotaxis and neutrophil elastase in the aortic wall in patients with abdominal aortic aneurysms. *J. Investig. Surg. Off. J. Acad. Surg. Res.* 4, 423–430.
- Connett, M.C., and Lansche, J.M. (1965). FIBROMUSCULAR HYPERPLASIA OF THE INTERNAL CAROTID ARTERY: REPORT OF A CASE. *Ann. Surg.* 162, 59–62.
- Cook, J.R., Carta, L., Galatioto, J., and Ramirez, F. (2015). Cardiovascular manifestations in Marfan syndrome and related diseases; multiple genes causing similar phenotypes. *Clin. Genet.* 87, 11–20.

- Cooper, D.N., Chen, J.-M., Ball, E.V., Howells, K., Mort, M., Phillips, A.D., Chuzhanova, N., Krawczak, M., Kehrer-Sawatzki, H., and Stenson, P.D. (2010). Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. *Hum. Mutat.* *31*, 631–655.
- Cora, T., Tokac, M., Acar, H., Soylu, A., and Inan, Z. (2013). Glutathione S-transferase M1 and T1 genotypes and myocardial infarction. *Mol. Biol. Rep.* *40*, 3263–3267.
- Cornwall, J.W., Green, R.S., Nielsen, J.C., and Gelb, B.D. (2014). Frequency of aortic dilation in Noonan syndrome. *Am. J. Cardiol.* *113*, 368–371.
- Coste, B., Mathur, J., Schmidt, M., Earley, T.J., Ranade, S., Petrus, M.J., Dubin, A.E., and Patapoutian, A. (2010). Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. *Science* *330*, 55–60.
- Coste, B., Xiao, B., Santos, J.S., Syeda, R., Grandl, J., Spencer, K.S., Kim, S.E., Schmidt, M., Mathur, J., Dubin, A.E., et al. (2012a). Piezo proteins are pore-forming subunits of mechanically activated channels. *Nature* *483*, 176–181.
- Coste, B., Murthy, S.E., Mathur, J., Schmidt, M., Mechioukhi, Y., Delmas, P., and Patapoutian, A. (2015). Piezo1 ion channel pore properties are dictated by C-terminal region. *Nat. Commun.* *6*, 7223.
- Coucke, P.J., Willaert, A., Wessels, M.W., Callewaert, B., Zoppi, N., De Backer, J., Fox, J.E., Mancini, G.M.S., Kambouris, M., Gardella, R., et al. (2006). Mutations in the facilitative glucose transporter GLUT10 alter angiogenesis and cause arterial tortuosity syndrome. *Nat. Genet.* *38*, 452–457.
- Cox, N.J., Wapelhorst, B., Morrison, V.A., Johnson, L., Pinchuk, L., Spielman, R.S., Todd, J.A., and Concannon, P. (2001). Seven Regions of the Genome Show Evidence of Linkage to Type 1 Diabetes in a Consensus Analysis of 767 Multiplex Families. *Am. J. Hum. Genet.* *69*, 820–830.
- Cragg, A.H., Smith, T.P., Thompson, B.H., Maroney, T.P., Stanson, A.W., Shaw, G.T., Hunter, D.W., and Cochran, S.T. (1989). Incidental fibromuscular dysplasia in potential renal donors: long-term clinical follow-up. *Radiology* *172*, 145–147.
- Curci, J.A., Liao, S., Huffman, M.D., Shapiro, S.D., and Thompson, R.W. (1998). Expression and localization of macrophage elastase (matrix metalloproteinase-12) in abdominal aortic aneurysms. *J. Clin. Invest.* *102*, 1900–1910.
- Daugherty, A., Manning, M.W., and Cassis, L.A. (2000). Angiotensin II promotes atherosclerotic lesions and aneurysms in apolipoprotein E-deficient mice. *J. Clin. Invest.* *105*, 1605–1612.
- Davis, V.A., Persidskaia, R.N., Baca-Regen, L.M., Fiotti, N., Halloran, B.G., and Baxter, B.T. (2001). Cytokine pattern in aneurysmal and occlusive disease of the aorta. *J. Surg. Res.* *101*, 152–156.
- DeBette, S., and Leys, D. (2009). Cervical-artery dissections: predisposing factors, diagnosis, and outcome. *Lancet Neurol.* *8*, 668–678.
- Delbosc, S., Alsac, J.-M., Journe, C., Louedec, L., Castier, Y., Bonnaure-Mallet, M., Ruimy, R., Rossignol, P., Bouchard, P., Michel, J.-B., et al. (2011). Porphyromonas gingivalis Participates in Pathogenesis of Human Abdominal Aortic Aneurysm by Neutrophil Activation. Proof of Concept in Rats. *PLoS ONE* *6*, e18679.
- Dimmeler, S., Fleming, I., Fisslthaler, B., Hermann, C., Busse, R., and Zeiher, A.M. (1999). Activation of nitric oxide synthase in endothelial cells by Akt-dependent phosphorylation. *Nature* *399*, 601–605.

Doblinger, E., Höcherl, K., Mederle, K., Kattler, V., Walter, S., Hansen, P.B., Jensen, B., and Castrop, H. (2012). Angiotensin AT1 receptor-associated protein Arap1 in the kidney vasculature is suppressed by angiotensin II. *Am. J. Physiol. Renal Physiol.* *302*, F1313–F1324.

Dobrin, P.B. (1978). Mechanical properties of arterises. *Physiol. Rev.* *58*, 397–460.

Dollery, C.M., Owen, C.A., Sukhova, G.K., Krettek, A., Shapiro, S.D., and Libby, P. (2003). Neutrophil Elastase in Human Atherosclerotic Plaques Production by Macrophages. *Circulation* *107*, 2829–2836.

Duca, L., Lambert, E., Debret, R., Rothhut, B., Blanchevove, C., Delacoux, F., Hornebeck, W., Martiny, L., and Debelle, L. (2005). Elastin peptides activate extracellular signal-regulated kinase 1/2 via a Ras-independent mechanism requiring both p110gamma/Raf-1 and protein kinase A/B-Raf signaling in human skin fibroblasts. *Mol. Pharmacol.* *67*, 1315–1324.

Ehrich, W.E., and Krumbhaar, E.B. (1943). A frequent obstructive anomaly of the mouth of the left common iliac vein. *Am. Heart J.* *26*, 737–750.

Eichler, E.E., Flint, J., Gibson, G., Kong, A., Leal, S.M., Moore, J.H., and Nadeau, J.H. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.* *11*, 446–450.

Elsayed, S.M., Phillips, J.B., Heller, R., Thoenes, M., Elsobky, E., Nürnberg, G., Nürnberg, P., Seland, S., Ebermann, I., Altmüller, J., et al. (2015). Non-manifesting AHI1 truncations indicate localized loss-of-function tolerance in a severe Mendelian disease gene. *Hum. Mol. Genet.* *24*, 2594–2603.

Emerick, K.M., Krantz, I.D., Kamath, B.M., Darling, C., Burrowes, D.M., Spinner, N.B., Whittington, P.F., and Piccoli, D.A. (2005). Intracranial vascular abnormalities in patients with Alagille syndrome. *J. Pediatr. Gastroenterol. Nutr.* *41*, 99–107.

Emond, M.J., Louie, T., Emerson, J., Zhao, W., Mathias, R.A., Knowles, M.R., Wright, F.A., Rieder, M.J., Tabor, H.K., Nickerson, D.A., et al. (2012). Exome sequencing of extreme phenotypes identifies DCTN4 as a modifier of chronic *Pseudomonas aeruginosa* infection in cystic fibrosis. *Nat. Genet.* *44*, 886–889.

Emond, M.J., Louie, T., Emerson, J., Chong, J.X., Mathias, R.A., Knowles, M.R., Rieder, M.J., Tabor, H.K., Nickerson, D.A., Barnes, K.C., et al. (2015). Exome Sequencing of Phenotypic Extremes Identifies CAV2 and TMC6 as Interacting Modifiers of Chronic *Pseudomonas aeruginosa* Infection in Cystic Fibrosis. *PLOS Genet.* *11*, e1005273.

Erlich, Y., Edvardson, S., Hodges, E., Zenvirt, S., Thekkat, P., Shaag, A., Dor, T., Hannon, G.J., and Elpeleg, O. (2011). Exome sequencing and disease-network analysis of a single family implicate a mutation in KIF1A in hereditary spastic paraparesis. *Genome Res.* *21*, 658–664.

Ezkurdia, I., Juan, D., Rodriguez, J.M., Frankish, A., Diekhans, M., Harrow, J., Vazquez, J., Valencia, A., and Tress, M.L. (2014). Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Hum. Mol. Genet.* *23*, 5866–5878.

Fabbretti, E., D'Arco, M., Fabbro, A., Simonetti, M., Nistri, A., and Giniatullin, R. (2006). Delayed Upregulation of ATP P2X3 Receptors of Trigeminal Sensory Neurons by Calcitonin Gene-Related Peptide. *J. Neurosci.* *26*, 6163–6171.

Faugeroux, J., Nematalla, H., Li, W., Clement, M., Robidel, E., Frank, M., Curis, E., Ait-Oufella, H., Caligiuri, G., Nicoletti, A., et al. (2013). Angiotensin II promotes thoracic aortic dissections and ruptures in Col3a1 haploinsufficient mice. *Hypertension* *62*, 203–208.

- Felgentreff, K., Siepe, M., Kotthoff, S., von Kodolitsch, Y., Schachtrup, K., Notarangelo, L.D., Walter, J.E., and Ehl, S. (2014). Severe eczema and Hyper-IgE in Loey's-Dietz-syndrome - contribution to new findings of immune dysregulation in connective tissue disorders. *Clin. Immunol. Orlando Fla* 150, 43–50.
- Fijneman, R.J., de Vries, S.S., Jansen, R.C., and Demant, P. (1996). Complex interactions of new quantitative trait loci, Sluc1, Sluc2, Sluc3, and Sluc4, that influence the susceptibility to lung cancer in the mouse. *Nat. Genet.* 14, 465–467.
- Fingleton, B. (2007). Matrix metalloproteinases as valid clinical targets. *Curr. Pharm. Des.* 13, 333–346.
- Fiskerstrand, T., Arshad, N., Haukanes, B.I., Tronstad, R.R., Pham, K.D.-C., Johansson, S., Håvik, B., Tønder, S.L., Levy, S.E., Brackman, D., et al. (2012). Familial diarrhea syndrome caused by an activating GUCY2C mutation. *N. Engl. J. Med.* 366, 1586–1595.
- Fleming, C., Whitlock, E.P., Beil, T.L., and Lederle, F.A. (2005). Screening for abdominal aortic aneurysm: a best-evidence systematic review for the U.S. Preventive Services Task Force. *Ann. Intern. Med.* 142, 203–211.
- Flicek, P., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., et al. (2014). Ensembl 2014. *Nucleic Acids Res.* 42, D749–D755.
- Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suñer, D., Cigudosa, J.C., Urioste, M., Benitez, J., et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. U. S. A.* 102, 10604–10609.
- Frank, M., Albuissou, J., Ranque, B., Golmard, L., Mazzella, J.-M., Bal-Theoleyre, L., Fauret, A.-L., Mirault, T., Denarié, N., Mousseaux, E., et al. (2015). The type of variants at the COL3A1 gene associates with the phenotype and severity of vascular Ehlers-Danlos syndrome. *Eur. J. Hum. Genet. EJHG.*
- Frydman, G., Walker, P.J., Summers, K., West, M., Xu, D., Lightfoot, T., Codd, C., Dique, T., and Nataatmadja, M. (2003). The value of screening in siblings of patients with abdominal aortic aneurysm. *Eur. J. Vasc. Endovasc. Surg. Off. J. Eur. Soc. Vasc. Surg.* 26, 396–400.
- Fujimoto, N., Nakajima, H., Sugiura, E., Dohi, K., Kanemitsu, S., Yamada, N., Aoki, Y., Nakatani, K., Shimpo, H., Nobori, T., et al. (2014). Bilateral giant coronary aneurysms in a 40-year-old male with Noonan syndrome caused by a KRAS germline mutation. *Int. J. Cardiol.* 173, e63–e66.
- Funakoshi, Y., Ichiki, T., Takeda, K., Tokuno, T., Iino, N., and Takeshita, A. (2002). Critical role of cAMP-response element-binding protein for angiotensin II-induced hypertrophy of vascular smooth muscle cells. *J. Biol. Chem.* 277, 18710–18717.
- Gacko, M., and Glowinski, S. (1998). Cathepsin D and cathepsin L activities in aortic aneurysm wall and parietal thrombus. *Clin. Chem. Lab. Med. CCLM FESCC* 36, 449–452.
- Gago-Díaz, M., Blanco-Verea, A., Teixidó-Turà, G., Valenzuela, I., Del Campo, M., Borregan, M., Sobrino, B., Amigo, J., García-Dorado, D., Evangelista, A., et al. (2014). Whole exome sequencing for the identification of a new mutation in TGFβ2 involved in a familial case of non-syndromic aortic disease. *Clin. Chim. Acta* 437, 88–92.
- Garg, V., Muth, A.N., Ransom, J.F., Schluterman, M.K., Barnes, R., King, I.N., Grossfeld, P.D., and Srivastava, D. (2005). Mutations in NOTCH1 cause aortic valve disease. *Nature* 437, 270–274.

- Gazit, R., Mandal, P.K., Ebina, W., Ben-Zvi, A., Nombela-Arrieta, C., Silberstein, L.E., and Rossi, D.J. (2014). Fgd5 identifies hematopoietic stem cells in the murine bone marrow. *J. Exp. Med.* *211*, 1315–1331.
- Genovese, G., Friedman, D.J., Ross, M.D., Lecordier, L., Uzureau, P., Freedman, B.I., Bowden, D.W., Langefeld, C.D., Oleksyk, T.K., Uscinski Knob, A.L., et al. (2010). Association of Trypanolytic ApoL1 Variants with Kidney Disease in African Americans. *Science* *329*, 841–845.
- Genovese, G., Friedman, D.J., and Pollak, M.R. (2013). APOL1 variants and kidney disease in people of recent African ancestry. *Nat. Rev. Nephrol.* *9*, 240–244.
- Gilissen, C., Arts, H.H., Hoischen, A., Spruijt, L., Mans, D.A., Arts, P., van Lier, B., Steehouwer, M., van Rieuwijk, J., Kant, S.G., et al. (2010). Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome. *Am. J. Hum. Genet.* *87*, 418–423.
- Gilissen, C., Hoischen, A., Brunner, H.G., and Veltman, J.A. (2012). Disease gene identification strategies for exome sequencing. *Eur. J. Hum. Genet. EJHG* *20*, 490–497.
- Gladstien, K., Rushton, A.R., and Kidd, K.K. (1980). Penetrance estimates and recurrence risks for fibromuscular dysplasia. *Clin. Genet.* *17*, 115–116.
- Glimåker, H., Holmberg, L., Elvin, A., Nybacka, O., Almgren, B., Björck, C.G., and Eriksson, I. (1991). Natural history of patients with abdominal aortic aneurysm. *Eur. J. Vasc. Surg.* *5*, 125–130.
- Goldwich, A., Burkard, M., Olke, M., Daniel, C., Amann, K., Hugo, C., Kurts, C., Steinkasserer, A., and Gessner, A. (2013). Podocytes Are Nonhematopoietic Professional Antigen-Presenting Cells. *J. Am. Soc. Nephrol.* *24*, 906–916.
- Golledge, J., Karan, M., Moran, C.S., Muller, J., Clancy, P., Dear, A.E., and Norman, P.E. (2008). Reduced expansion rate of abdominal aortic aneurysms in patients with diabetes may be related to aberrant monocyte-matrix interactions. *Eur. Heart J.* *29*, 665–672.
- Gorvel, L., Ben Amara, A., Ka, M.B., Textoris, J., Gorvel, J.-P., and Mege, J.-L. (2014). Myeloid decidual dendritic cells and immunoregulation of pregnancy: defective responsiveness to *Coxiella burnetii* and *Brucella abortus*. *Front. Cell. Infect. Microbiol.* *4*.
- Gown, A.M., Tsukada, T., and Ross, R. (1986). Human atherosclerosis. II. Immunocytochemical analysis of the cellular composition of human atherosclerotic lesions. *Am. J. Pathol.* *125*, 191–207.
- Graul-Neumann, L.M., Hausser, I., Essayie, M., Rauch, A., and Kraus, C. (2008). Highly variable cutis laxa resulting from a dominant splicing mutation of the elastin gene. *Am. J. Med. Genet. A.* *146A*, 977–983.
- Gredmark-Russ, S., Dzabic, M., Rahbar, A., Wanhainen, A., Björck, M., Larsson, E., Michel, J.-B., and Söderberg-Nauclér, C. (2009). Active cytomegalovirus infection in aortic smooth muscle cells from patients with abdominal aortic aneurysm. *J. Mol. Med. Berl. Ger.* *87*, 347–356.
- Gregory, A.K., Yin, N.X., Capella, J., Xia, S., Newman, K.M., and Tilson, M.D. (1996). Features of autoimmunity in the abdominal aortic aneurysm. *Arch. Surg. Chic. Ill* *1960* *131*, 85–88.
- Gretarsdottir, S., Baas, A.F., Thorleifsson, G., Holm, H., den Heijer, M., de Vries, J.-P.P.M., Kranendonk, S.E., Zeebregts, C.J.A.M., van Sterkenburg, S.M., Geelkerken, R.H., et al. (2010). Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. *Nat. Genet.* *42*, 692–697.

- Grimbert, P., Fiquier-Kempf, B., Coudol, P., Vuagnat, A., Pannier-Moreau, I., Corvol, P., Plouin, P.F., and Jeunemaître, X. (1998). [Genetic study of renal artery fibromuscular dysplasia]. *Arch. Mal. Coeur Vaiss.* *91*, 1069–1071.
- Grootenboer, S., Schischmanoff, P.O., Cynober, T., Rodrigue, J.C., Delaunay, J., Tchernia, G., and Dommergues, J.P. (1998). A genetic syndrome associating dehydrated hereditary stomatocytosis, pseudohyperkalemia and perinatal oedema. *Br. J. Haematol.* *103*, 383–386.
- Grootenboer, S., Schischmanoff, P.O., Laurendeau, I., Cynober, T., Tchernia, G., Dommergues, J.P., Dhermy, D., Bost, M., Varet, B., Snyder, M., et al. (2000a). Pleiotropic syndrome of dehydrated hereditary stomatocytosis, pseudohyperkalemia, and perinatal edema maps to 16q23-q24. *Blood* *96*, 2599–2605.
- Groza, T., Köhler, S., Moldenhauer, D., Vasilevsky, N., Baynam, G., Zemojtel, T., Schriml, L.M., Kibbe, W.A., Schofield, P.N., Beck, T., et al. (2015). The Human Phenotype Ontology: Semantic Unification of Common and Rare Disease. *Am. J. Hum. Genet.* *97*, 111–124.
- Guo, D., Regalado, E., Casteel, D.E., Santos-Cortez, R.L., Gong, L., Kim, J.J., Dyack, S., Horne, S.G., Chang, G., Jondeau, G., et al. (2013). Recurrent gain-of-function mutation in PRKG1 causes thoracic aortic aneurysms and acute aortic dissections. *Am. J. Hum. Genet.* *93*, 398–404.
- Guo, D., Gong, L., Regalado, E.S., Santos-Cortez, R.L., Zhao, R., Cai, B., Veeraraghavan, S., Prakash, S.K., Johnson, R.J., Muilenburg, A., et al. (2015). MAT2A Mutations Predispose Individuals to Thoracic Aortic Aneurysms. *Am. J. Hum. Genet.* *96*, 170–177.
- Guo, D.-C., Pannu, H., Tran-Fadulu, V., Papke, C.L., Yu, R.K., Avidan, N., Bourgeois, S., Estrera, A.L., Safi, H.J., Sparks, E., et al. (2007). Mutations in smooth muscle α -actin (ACTA2) lead to thoracic aortic aneurysms and dissections. *Nat. Genet.* *39*, 1488–1493.
- Guo, D.-C., Papke, C.L., Tran-Fadulu, V., Regalado, E.S., Avidan, N., Johnson, R.J., Kim, D.H., Pannu, H., Willing, M.C., Sparks, E., et al. (2009). Mutations in smooth muscle alpha-actin (ACTA2) cause coronary artery disease, stroke, and Moyamoya disease, along with thoracic aortic disease. *Am. J. Hum. Genet.* *84*, 617–627.
- Guo, D.-F., Chenier, I., Tardif, V., Orlov, S.N., and Inagami, T. (2003). Type 1 angiotensin II receptor-associated protein ARAP1 binds and recycles the receptor to the plasma membrane. *Biochem. Biophys. Res. Commun.* *310*, 1254–1265.
- Gusella, J.F., Wexler, N.S., Conneally, P.M., Naylor, S.L., Anderson, M.A., Tanzi, R.E., Watkins, P.C., Ottina, K., Wallace, M.R., and Sakaguchi, A.Y. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* *306*, 234–238.
- Ha, L., Shi, Y., Zhao, J., Li, T., and Chen, Z.-J. (2015). Association Study between Polycystic Ovarian Syndrome and the Susceptibility Genes Polymorphisms in Hui Chinese Women. *PloS One* *10*, e0126505.
- Haines, J.L., Hauser, M.A., Schmidt, S., Scott, W.K., Olson, L.M., Gallins, P., Spencer, K.L., Kwan, S.Y., Noureddine, M., Gilbert, J.R., et al. (2005). Complement factor H variant increases the risk of age-related macular degeneration. *Science* *308*, 419–421.
- Halpern, M.M., Sanford, H.S., and Viamonte, M. (1965). Renal-artery abnormalities in three hypertensive sisters. Probable familial fibromuscular hyperplasia. *JAMA* *194*, 512–513.
- Halushka, M.K., Fan, J.B., Bentley, K., Hsie, L., Shen, N., Weder, A., Cooper, R., Lipshutz, R., and Chakravarti, A. (1999). Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat. Genet.* *22*, 239–247.

- Hamilton, S.J., and Friedman, J.M. (2000). Insights into the pathogenesis of neurofibromatosis 1 vasculopathy. *Clin. Genet.* 58, 341–344.
- Han, F., and Pan, W. (2010). A data-adaptive sum test for disease association with multiple common or rare variants. *Hum. Hered.* 70, 42–54.
- Hance, K.A., Tataria, M., Ziporin, S.J., Lee, J.K., and Thompson, R.W. (2002). Monocyte chemotactic activity in human abdominal aortic aneurysms: role of elastin degradation peptides and the 67-kD cell surface elastin receptor. *J. Vasc. Surg.* 35, 254–261.
- Hanihara-Tatsuzawa, F., Miura, H., Kobayashi, S., Isagawa, T., Okuma, A., Manabe, I., and MaruYama, T. (2014). Control of Toll-like receptor-mediated T cell-independent type 1 antibody responses by the inducible nuclear protein I κ B- ζ . *J. Biol. Chem.* 289, 30925–30936.
- Hannibal, M.C., Buckingham, K.J., Ng, S.B., Ming, J.E., Beck, A.E., McMillin, M.J., Gildersleeve, H.I., Bigham, A.W., Tabor, H.K., Mefford, H.C., et al. (2011). Spectrum of MLL2 (ALR) mutations in 110 cases of Kabuki syndrome. *Am. J. Med. Genet. A.* 155A, 1511–1516.
- Hansen, J., Holten, C., and Thorborg, J.V. (1965). Hypertension in two sisters caused by so-called fibromuscular hyperplasia of the renal arteries. *Acta Med. Scand.* 178, 461–474.
- Hansson, G.K., Jonasson, L., Lojstved, B., Stemme, S., Kocher, O., and Gabbiani, G. (1988). Localization of T lymphocytes and macrophages in fibrous and complicated human atherosclerotic plaques. *Atherosclerosis* 72, 135–141.
- Hao, Y., Wong, R., and Feig, L.A. (2008). RalGDS couples growth factor signaling to Akt activation. *Mol. Cell. Biol.* 28, 2851–2859.
- Harrison, E.G., and McCormack, L.J. (1971). Pathologic classification of renal arterial disease in renovascular hypertension. *Mayo Clin. Proc.* 46, 161–167.
- Harthun, N.L. (2008). Current issues in the treatment of women with abdominal aortic aneurysm. *Gend. Med.* 5, 36–43.
- Hayakawa, T., Yamashita, K., Ohuchi, E., and Shinagawa, A. (1994). Cell growth-promoting activity of tissue inhibitor of metalloproteinases-2 (TIMP-2). *J. Cell Sci.* 107 (Pt 9), 2373–2379.
- Hayashidani, S., Tsutsui, H., Ikeuchi, M., Shiomi, T., Matsusaka, H., Kubota, T., Imanaka-Yoshida, K., Itoh, T., and Takeshita, A. (2003). Targeted deletion of MMP-2 attenuates early LV rupture and late remodeling after experimental myocardial infarction. *Am. J. Physiol. Heart Circ. Physiol.* 285, H1229–H1235.
- Hayes, J.D., and Strange, R.C. (2000). Glutathione S-transferase polymorphisms and their biological consequences. *Pharmacology* 61, 154–166.
- Helgadottir, A., Thorleifsson, G., Magnusson, K.P., Grétarsdottir, S., Steinthorsdottir, V., Manolescu, A., Jones, G.T., Rinkel, G.J.E., Blankensteijn, J.D., Ronkainen, A., et al. (2008). The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat. Genet.* 40, 217–224.
- Hemani, G., Yang, J., Vinkhuyzen, A., Powell, J.E., Willemssen, G., Hottenga, J.-J., Abdellaoui, A., Mangino, M., Valdes, A.M., Medland, S.E., et al. (2013). Inference of the genetic architecture underlying BMI and height with the use of 20,240 sibling pairs. *Am. J. Hum. Genet.* 93, 865–875.

- Henderson, E.L., Geng, Y.J., Sukhova, G.K., Whittemore, A.D., Knox, J., and Libby, P. (1999). Death of smooth muscle cells and expression of mediators of apoptosis by T lymphocytes in human abdominal aortic aneurysms. *Circulation* *99*, 96–104.
- Hervé, D., Philippi, A., Belbouab, R., Zerah, M., Chabrier, S., Collardeau-Frachon, S., Bergametti, F., Essongue, A., Berrou, E., Krivosic, V., et al. (2014). Loss of $\alpha 1\beta 1$ Soluble Guanylate Cyclase, the Major Nitric Oxide Receptor, Leads to Moyamoya and Achalasia. *Am. J. Hum. Genet.* *94*, 385–394.
- Hewitt, R., Krause, A., Goldman, A., Campbell, G., and Jenkins, T. (1996). Beta-globin haplotype analysis suggests that a major source of Malagasy ancestry is derived from Bantu-speaking Negroids. *Am. J. Hum. Genet.* *58*, 1303–1308.
- Hicks, S., Wheeler, D.A., Plon, S.E., and Kimmel, M. (2011). Prediction of missense mutation functionality depends on both the algorithm and sequence alignment employed. *Hum. Mutat.* *32*, 661–668.
- High, F.A., and Epstein, J.A. (2008). The multifaceted role of Notch in cardiac development and disease. *Nat. Rev. Genet.* *9*, 49–61.
- Hinek, A., and Wilson, S.E. (2000). Impaired elastogenesis in Hurler disease: dermatan sulfate accumulation linked to deficiency in elastin-binding protein and elastic fiber assembly. *Am. J. Pathol.* *156*, 925–938.
- Hinterseher, I., Tromp, G., and Kuivaniemi, H. (2011). Genes and Abdominal Aortic Aneurysm. *Ann. Vasc. Surg.* *25*, 388–412.
- Hinterseher, I., Gäbel, G., Corvinus, F., Lück, C., Saeger, H.D., Bergert, H., Tromp, G., and Kuivaniemi, H. (2012). Presence of *Borrelia burgdorferi* sensu lato antibodies in the serum of patients with abdominal aortic aneurysms. *Eur. J. Clin. Microbiol. Infect. Dis. Off. Publ. Eur. Soc. Clin. Microbiol.* *31*, 781–789.
- Hirose, H., Ozsvath, K.J., Xia, S., and Tilson, M.D. (1997). Molecular cloning of the complementary DNA for an additional member of the family of aortic aneurysm antigenic proteins. *J. Vasc. Surg.* *26*, 313–318.
- Hobbs, S.D., Claridge, M.W.C., Quick, C.R.G., Day, N.E., Bradbury, A.W., and Wilmink, A.B.M. (2003). LDL cholesterol is associated with small abdominal aortic aneurysms. *Eur. J. Vasc. Endovasc. Surg. Off. J. Eur. Soc. Vasc. Surg.* *26*, 618–622.
- Hoischen, A., van Bon, B.W.M., Gilissen, C., Arts, P., van Lier, B., Steehouwer, M., de Vries, P., de Reuver, R., Wieskamp, N., Mortier, G., et al. (2010). De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat. Genet.* *42*, 483–485.
- Hoischen, A., van Bon, B.W.M., Rodríguez-Santiago, B., Gilissen, C., Vissers, L.E.L.M., de Vries, P., Janssen, I., van Lier, B., Hastings, R., Smithson, S.F., et al. (2011). De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. *Nat. Genet.* *43*, 729–731.
- Hokamura, K., Inaba, H., Nakano, K., Nomura, R., Yoshioka, H., Taniguchi, K., Ooshima, T., Wada, K., Amano, A., and Umemura, K. (2010). Molecular analysis of aortic intimal hyperplasia caused by *Porphyromonas gingivalis* infection in mice with endothelial damage. *J. Periodontal Res.* *45*, 337–344.
- Holm, H., Gudbjartsson, D.F., Sulem, P., Masson, G., Helgadóttir, H.T., Zanon, C., Magnusson, O.T., Helgason, A., Saemundsdóttir, J., Gylfason, A., et al. (2011). A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat. Genet.* *43*, 316–320.

Horiguchi, M., Inoue, T., Ohbayashi, T., Hirai, M., Noda, K., Marmorstein, L.Y., Yabe, D., Takagi, K., Akama, T.O., Kita, T., et al. (2009). Fibulin-4 conducts proper elastogenesis via interaction with cross-linking enzyme lysyl oxidase. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 19029–19034.

Horrevoets, A.J., Fontijn, R.D., van Zonneveld, A.J., de Vries, C.J., ten Cate, J.W., and Pannekoek, H. (1999). Vascular endothelial genes that are responsive to tumor necrosis factor-alpha in vitro are expressed in atherosclerotic lesions, including inhibitor of apoptosis protein-1, stannin, and two novel genes. *Blood* *93*, 3418–3431.

Houston, B.L., Zelinski, T., Israels, S.J., Coghlan, G., Chodirker, B.N., Gallagher, P.G., Houston, D.S., and Zarychanski, R. (2011a). Refinement of the hereditary xerocytosis locus on chromosome 16q in a large Canadian kindred. *Blood Cells. Mol. Dis.* *47*, 226–231.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* *37*, 1–13.

Huang, Y., Laval, S.H., van Remoortere, A., Baudier, J., Benaud, C., Anderson, L.V.B., Straub, V., Deelder, A., Frants, R.R., den Dunnen, J.T., et al. (2007). AHNAK, a novel component of the dysferlin protein complex, redistributes to the cytoplasm with dysferlin during skeletal muscle regeneration. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *21*, 732–742.

Huchtagowder, V., Sausgruber, N., Kim, K.H., Angle, B., Marmorstein, L.Y., and Urban, Z. (2006). Fibulin-4: a novel gene for an autosomal recessive cutis laxa syndrome. *Am. J. Hum. Genet.* *78*, 1075–1080.

Huebner, A.K., Gandia, M., Frommolt, P., Maak, A., Wicklein, E.M., Thiele, H., Altmüller, J., Wagner, F., Viñuela, A., Aguirre, L.A., et al. (2011). Nonsense Mutations in SMPX, Encoding a Protein Responsive to Physical Force, Result in X-Chromosomal Hearing Loss. *Am. J. Hum. Genet.* *88*, 621–627.

Ichiki, T. (2006). Role of cAMP response element binding protein in cardiovascular remodeling: good, bad, or both? *Arterioscler. Thromb. Vasc. Biol.* *26*, 449–455.

Indraccolo, S., Pfeffer, U., Minuzzo, S., Esposito, G., Roni, V., Mandruzzato, S., Ferrari, N., Anfosso, L., Dell'Eva, R., Noonan, D.M., et al. (2007). Identification of Genes Selectively Regulated by IFNs in Endothelial Cells. *J. Immunol.* *178*, 1122–1135.

Iolascon, A., Stewart, G.W., Ajetunmobi, J.F., Perrotta, S., Delaunay, J., Carella, M., Zelante, L., and Gasparini, P. (1999). Familial pseudohyperkalemia maps to the same locus as dehydrated hereditary stomatocytosis (hereditary xerocytosis). *Blood* *93*, 3120–3123.

Ionita-Laza, I., Lee, S., Makarov, V., Buxbaum, J.D., and Lin, X. (2013). Sequence kernel association tests for the combined effect of rare and common variants. *Am. J. Hum. Genet.* *92*, 841–853.

Iossifov, I., O’Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature* *515*, 216–221.

Irizarry, E., Newman, K.M., Gandhi, R.H., Nackman, G.B., Halpern, V., Wishner, S., Scholes, J.V., and Tilson, M.D. (1993). Demonstration of interstitial collagenase in abdominal aortic aneurysm disease. *J. Surg. Res.* *54*, 571–574.

Ishiguro-Oonuma, T., Ochiai, K., Hashizume, K., and Morimatsu, M. (2015). The role of IFN- γ in regulating Nfkbiz expression in epidermal keratinocytes. *Biomed. Res. Tokyo Jpn.* *36*, 103–107.

- Jackson, W., and Rutter, P. (2001). Epidemiology of abdominal aortic aneurysms in the Asian community (Br J Surg 2001; 88: 382-4). Br. J. Surg. 88, 1017.
- Jamrozik, K., Norman, P.E., Spencer, C.A., Parsons, R.W., Tuohy, R., Lawrence-Brown, M.M., and Dickinson, J.A. (2000). Screening for abdominal aortic aneurysm: lessons from a population-based study. Med. J. Aust. 173, 345–350.
- Jean-Claude, J., Newman, K.M., Li, H., Gregory, A.K., and Tilson, M.D. (1994). Possible key role for plasmin in the pathogenesis of abdominal aortic aneurysms. Surgery 116, 472–478.
- Jeanne, M., Jorgensen, J., and Gould, D.B. (2015). Molecular and Genetic Analyses of Collagen Type IV Mutant Mouse Models of Spontaneous Intracerebral Hemorrhage Identify Mechanisms for Stroke Prevention. Circulation 131, 1555–1565.
- Jensen, L.J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., et al. (2009). STRING 8—a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res. 37, D412–D416.
- Jimenez, M., Daret, D., Choussat, A., and Bonnet, J. (1999). Immunohistological and ultrastructural analysis of the intimal thickening in coarctation of human aorta. Cardiovasc. Res. 41, 737–745.
- Johansen, C.T., Wang, J., McIntyre, A.D., Martins, R.A., Ban, M.R., Lanktree, M.B., Huff, M.W., Péterfy, M., Mehrabian, M., Lusis, A.J., et al. (2012). Excess of rare variants in non-genome-wide association study candidate genes in patients with hypertriglyceridemia. Circ. Cardiovasc. Genet. 5, 66–72.
- Jonasson, L., Holm, J., Skalli, O., Bondjers, G., and Hansson, G.K. (1986). Regional accumulations of T cells, macrophages, and smooth muscle cells in the human atherosclerotic plaque. Arterioscler. Dallas Tex 6, 131–138.
- Jones, G.T., Bown, M.J., Gretarsdottir, S., Romaine, S.P.R., Helgadottir, A., Yu, G., Tromp, G., Norman, P.E., Jin, C., Baas, A.F., et al. (2013). A sequence variant associated with sortilin-1 (SORT1) on 1p13.3 is independently associated with abdominal aortic aneurysm. Hum. Mol. Genet. 22, 2941–2947.
- Joseph, E.K., Green, P.G., Bogen, O., Alvarez, P., and Levine, J.D. (2013). Vascular Endothelial Cells Mediate Mechanical Stimulation-Induced Enhancement of Endothelin Hyperalgesia via Activation of P2X2/3 Receptors on Nociceptors. J. Neurosci. 33, 2849–2859.
- Juvonen, J., Surcel, H.M., Satta, J., Teppo, A.M., Bloigu, A., Syrjälä, H., Airaksinen, J., Leinonen, M., Saikku, P., and Juvonen, T. (1997). Elevated circulating levels of inflammatory cytokines in patients with abdominal aortic aneurysm. Arterioscler. Thromb. Vasc. Biol. 17, 2843–2847.
- Kaartinen, V., and Warburton, D. (2003). Fibrillin controls TGF- β activation. Nat. Genet. 33, 331–332.
- Kalluri, R. (2003). Angiogenesis: Basement membranes: structure, assembly and role in tumour angiogenesis. Nat. Rev. Cancer 3, 422–433.
- Kamada, F., Aoki, Y., Narisawa, A., Abe, Y., Komatsuzaki, S., Kikuchi, A., Kanno, J., Niihori, T., Ono, M., Ishii, N., et al. (2011). A genome-wide association study identifies RNF213 as the first Moyamoya disease gene. J. Hum. Genet. 56, 34–40.
- Kamath, B.M., Spinner, N.B., Emerick, K.M., Chudley, A.E., Booth, C., Piccoli, D.A., and Krantz, I.D. (2004). Vascular anomalies in Alagille syndrome: a significant cause of morbidity and mortality. Circulation 109, 1354–1358.

- Kamphaus, G.D., Colorado, P.C., Panka, D.J., Hopfer, H., Ramchandran, R., Torre, A., Maeshima, Y., Mier, J.W., Sukhatme, V.P., and Kalluri, R. (2000). Canstatin, a novel matrix-derived inhibitor of angiogenesis and tumor growth. *J. Biol. Chem.* *275*, 1209–1215.
- Karkos, C.D., Mukhopadhyay, U., Papakostas, I., Ghosh, J., Thomson, G.J., and Hughes, R. (2000). Abdominal aortic aneurysm: the role of clinical examination and opportunistic detection. *Eur. J. Vasc. Endovasc. Surg. Off. J. Eur. Soc. Vasc. Surg.* *19*, 299–303.
- Kawabe, J. -i., Yuhki, K. -i., Okada, M., Kanno, T., Yamauchi, A., Tashiro, N., Sasaki, T., Okumura, S., Nakagawa, N., Aburakawa, Y., et al. (2010). Prostaglandin I₂ Promotes Recruitment of Endothelial Progenitor Cells and Limits Vascular Remodeling. *Arterioscler. Thromb. Vasc. Biol.* *30*, 464–470.
- Kawai, M., Kawashima, S., Sakoda, T., Toh, R., Kikuchi, A., Yamauchi-Takahara, K., Kunisada, K., and Yokoyama, M. (2003). Ral GDP dissociation stimulator and Ral GTPase are involved in myocardial hypertrophy. *Hypertension* *41*, 956–962.
- Kelm, M., Feelisch, M., Spahr, R., Piper, H.M., Noack, E., and Schrader, J. (1988). Quantitative and kinetic characterization of nitric oxide and EDRF released from cultured endothelial cells. *Biochem. Biophys. Res. Commun.* *154*, 236–244.
- Khan, S., Habib, R., Mir, H., Umm-e-Kalsoom, null, Naz, G., Ayub, M., Shafique, S., Yamin, T., Ali, N., Basit, S., et al. (2011). Mutations in the LPAR6 and LIPH genes underlie autosomal recessive hypotrichosis/woolly hair in 17 consanguineous families from Pakistan. *Clin. Exp. Dermatol.* *36*, 652–654.
- Kheirollahi, M., Kazemi, M., Amini, G., Khorvash, F., Ahangari, F., Kolahdouz, M., and Koulivand, L. (2015). Expression of prostaglandin I₂ (prostacyclin) receptor in blood of migraine patients: A potential biomarker. *Adv. Biomed. Res.* *4*, 121.
- Kiando, S.R., Barlassina, C., Cusi, D., Galan, P., Lathrop, M., Plouin, P.-F., Jeunemaitre, X., and Bouatia-Naji, N. (2015). Exome sequencing in seven families and gene-based association studies indicate genetic heterogeneity and suggest possible candidates for fibromuscular dysplasia. *J. Hypertens.* *33*, 1802–1810.
- Kincaid, O.W., and Davis, G.D. (1961). Renal arteriography in hypertension. *Proc. Staff Meet. Mayo Clin.* *36*, 689–701.
- Kirin, M., McQuillan, R., Franklin, C.S., Campbell, H., McKeigue, P.M., and Wilson, J.F. (2010). Genomic Runs of Homozygosity Record Population History and Consanguinity. *PLoS ONE* *5*, e13996.
- Koch, A.E., Haines, G.K., Rizzo, R.J., Radosevich, J.A., Pope, R.M., Robinson, P.G., and Pearce, W.H. (1990). Human abdominal aortic aneurysms. Immunophenotypic analysis suggesting an immune-mediated response. *Am. J. Pathol.* *137*, 1199–1213.
- Koch, A.E., Kunkel, S.L., Pearce, W.H., Shah, M.R., Parikh, D., Evanoff, H.L., Haines, G.K., Burdick, M.D., and Strieter, R.M. (1993). Enhanced production of the chemotactic cytokines interleukin-8 and monocyte chemoattractant protein-1 in human abdominal aortic aneurysms. *Am. J. Pathol.* *142*, 1423–1431.
- Kohout, J., Topolcan, O., and Bejcková, H. (1995). The serum level of endothelin in patients with hand-arm vibration syndrome. *Cent. Eur. J. Public Health* *3 Suppl*, 43–44.
- Kong, A., Steinthorsdottir, V., Masson, G., Thorleifsson, G., Sulem, P., Besenbacher, S., Jonasdottir, A., Sigurdsson, A., Kristinsson, K.T., Jonasdottir, A., et al. (2009). Parental origin of sequence variants associated with complex diseases. *Nature* *462*, 868–874.

- Kotake, Y., Nakagawa, T., Kitagawa, K., Suzuki, S., Liu, N., Kitagawa, M., and Xiong, Y. (2011). Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene* 30, 1956–1962.
- Krajnak, K. (2005). Acute vibration increases 2C-adrenergic smooth muscle constriction and alters thermosensitivity of cutaneous arteries. *J. Appl. Physiol.* 100, 1230–1237.
- Krajnak, K., Waugh, S., Johnson, C., Miller, R., and Kiedrowski, M. (2009). Vibration disrupts vascular function in a model of metabolic syndrome. *Ind. Health* 47, 533–542.
- Krawitz, P.M., Schweiger, M.R., Rödelsperger, C., Marcelis, C., Kölsch, U., Meisel, C., Stephani, F., Kinoshita, T., Murakami, Y., Bauer, S., et al. (2010). Identity-by-descent filtering of exome sequence data identifies PIGV mutations in hyperphosphatasia mental retardation syndrome. *Nat. Genet.* 42, 827–829.
- Kunapuli, S.P., and Daniel, J.L. (1998). P2 receptor subtypes in the cardiovascular system. *Biochem. J.* 336 (Pt 3), 513–523.
- van de Laar, I.M.B.H., Oldenburg, R.A., Pals, G., Roos-Hesselink, J.W., de Graaf, B.M., Verhagen, J.M.A., Hoedemaekers, Y.M., Willemsen, R., Severijnen, L.-A., Venselaar, H., et al. (2011). Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. *Nat. Genet.* 43, 121–126.
- LaMorte, W.W., Scott, T.E., and Menzoian, J.O. (1995). Racial differences in the incidence of femoral bypass and abdominal aortic aneurysmectomy in Massachusetts: relationship to cardiovascular risk factors. *J. Vasc. Surg.* 21, 422–431.
- Lan, X., Jhaveri, A., Cheng, K., Wen, H., Saleem, M.A., Mathieson, P.W., Mikulak, J., Aviram, S., Malhotra, A., Skorecki, K., et al. (2014). APOL1 risk variants enhance podocyte necrosis through compromising lysosomal membrane permeability. *AJP Ren. Physiol.* 307, F326–F336.
- Lan, X., Wen, H., Lederman, R., Malhotra, A., Mikulak, J., Popik, W., Skorecki, K., and Singhal, P.C. (2015). Protein domains of APOL1 and its risk variants. *Exp. Mol. Pathol.* 99, 139–144.
- Lander, E.S. (1996). The new genomics: global views of biology. *Science* 274, 536–539.
- Lander, E.S., and Botstein, D. (1987). Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* 236, 1567–1570.
- Lange, S., Ouyang, K., Meyer, G., Cui, L., Cheng, H., Lieber, R.L., and Chen, J. (2009). Obscurin determines the architecture of the longitudinal sarcoplasmic reticulum. *J. Cell Sci.* 122, 2640–2650.
- Law, M. (1998). Screening for abdominal aortic aneurysms. *Br. Med. Bull.* 54, 903–913.
- Lawrence, P.F., Wallis, C., Dobrin, P.B., Bhirangi, K., Gugliuzza, N., Galt, S., and Kraiss, L. (1998). Peripheral aneurysms and arteriomegaly: is there a familial pattern? *J. Vasc. Surg.* 28, 599–605.
- Leamy, L.J., Routman, E.J., and Cheverud, J.M. (2002). An epistatic genetic basis for fluctuating asymmetry of mandible size in mice. *Evol. Int. J. Org. Evol.* 56, 642–653.
- Lecordier, L., Vanhollebeke, B., Poelvoorde, P., Tebabi, P., Paturiaux-Hanocq, F., Andris, F., Lins, L., and Pays, E. (2009). C-Terminal Mutants of Apolipoprotein L-I Efficiently Kill Both *Trypanosoma brucei brucei* and *Trypanosoma brucei rhodesiense*. *PLoS Pathog.* 5, e1000685.

- Lederle, F.A., Johnson, G.R., Wilson, S.E., Chute, E.P., Littooy, F.N., Bandyk, D., Krupski, W.C., Barone, G.W., Acher, C.W., and Ballard, D.J. (1997). Prevalence and associations of abdominal aortic aneurysm detected through screening. Aneurysm Detection and Management (ADAM) Veterans Affairs Cooperative Study Group. *Ann. Intern. Med.* *126*, 441–449.
- Lee, C.W., Hwang, I., Park, C.-S., Lee, H., Park, D.-W., Kang, S.-J., Lee, S.-W., Kim, Y.-H., Park, S.-W., and Park, S.-J. (2012a). Expression of ADAMTS-2, -3, -13, and -14 in culprit coronary lesions in patients with acute myocardial infarction or stable angina. *J. Thromb. Thrombolysis* *33*, 362–370.
- Lee, S., Wu, M.C., and Lin, X. (2012b). Optimal tests for rare variant effects in sequencing association studies. *Biostat. Oxf. Engl.* *13*, 762–775.
- Lee, S., Emond, M.J., Bamshad, M.J., Barnes, K.C., Rieder, M.J., Nickerson, D.A., Christiani, D.C., Wurfel, M.M., and Lin, X. (2012c). Optimal Unified Approach for Rare-Variant Association Testing with Application to Small-Sample Case-Control Whole-Exome Sequencing Studies. *Am. J. Hum. Genet.* *91*, 224–237.
- Leeper, N.J., Raiesdana, A., Kojima, Y., Kundu, R.K., Cheng, H., Maegdefessel, L., Toh, R., Ahn, G.-O., Ali, Z.A., Anderson, D.R., et al. (2013). Loss of CDKN2B promotes p53-dependent smooth muscle cell apoptosis and aneurysm formation. *Arterioscler. Thromb. Vasc. Biol.* *33*, e1–e10.
- Lenk, G.M., Tromp, G., Skunca, M., Gatalica, Z., Berguer, R., and Kuivaniemi, H. (2006). Global expression profiles in human normal and aneurysmal abdominal aorta based on two distinct whole genome microarray platforms. *Ann. N. Y. Acad. Sci.* *1085*, 360–362.
- Leopold, J.A., Dam, A., Maron, B.A., Scribner, A.W., Liao, R., Handy, D.E., Stanton, R.C., Pitt, B., and Loscalzo, J. (2007). Aldosterone impairs vascular reactivity by decreasing glucose-6-phosphate dehydrogenase activity. *Nat. Med.* *13*, 189–197.
- Li, B., and Leal, S.M. (2008). Methods for Detecting Associations with Rare Variants for Common Diseases: Application to Analysis of Sequence Data. *Am. J. Hum. Genet.* *83*, 311–321.
- Li, A.H., Morrison, A.C., Kovar, C., Cupples, L.A., Brody, J.A., Polfus, L.M., Yu, B., Metcalf, G., Muzny, D., Veeraraghavan, N., et al. (2015). Analysis of loss-of-function variants and 20 risk factor phenotypes in 8,554 individuals identifies loci influencing chronic disease. *Nat. Genet.* *47*, 640–642.
- Li, W., Li, Q., Qin, L., Ali, R., Qyang, Y., Tassabehji, M., Pober, B.R., Sessa, W.C., Giordano, F.J., and Tellides, G. (2013). Rapamycin inhibits smooth muscle cell proliferation and obstructive arteriopathy attributable to elastin deficiency. *Arterioscler. Thromb. Vasc. Biol.* *33*, 1028–1035.
- Liao, W., Goh, F.Y., Betts, R.J., Kemeny, D.M., Tam, J., Bay, B.-H., and Wong, W.S.F. (2011). A novel anti-apoptotic role for apolipoprotein L2 in IFN- γ -induced cytotoxicity in human bronchial epithelial cells. *J. Cell. Physiol.* *226*, 397–406.
- Lim, E.T., Raychaudhuri, S., Sanders, S.J., Stevens, C., Sabo, A., MacArthur, D.G., Neale, B.M., Kirby, A., Ruderfer, D.M., Fromer, M., et al. (2013). Rare complete knockouts in humans: population distribution and significant role in autism spectrum disorders. *Neuron* *77*, 235–242.
- Lim, E.T., Würtz, P., Havulinna, A.S., Palta, P., Tukiainen, T., Rehnström, K., Esko, T., Mägi, R., Inouye, M., Lappalainen, T., et al. (2014). Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet.* *10*, e1004494.

- Lindsay, M.E., Schepers, D., Bolar, N.A., Doyle, J.J., Gallo, E., Fert-Bober, J., Kempers, M.J.E., Fishman, E.K., Chen, Y., Myers, L., et al. (2012). Loss-of-function mutations in *TGFB2* cause a syndromic presentation of thoracic aortic aneurysm. *Nat. Genet.* *44*, 922–927.
- Liu, J., Sukhova, G.K., Yang, J.-T., Sun, J., Ma, L., Ren, A., Xu, W.-H., Fu, H., Dolganov, G.M., Hu, C., et al. (2006). Cathepsin L expression and regulation in human abdominal aortic aneurysm, atherosclerosis, and vascular cells. *Atherosclerosis* *184*, 302–311.
- Liu, J., Li, J.D., Lu, J., Xing, J., and Li, J. (2011a). Contribution of nerve growth factor to upregulation of P2X3 expression in DRG neurons of rats with femoral artery occlusion. *AJP Heart Circ. Physiol.* *301*, H1070–H1079.
- Liu, L., Sabo, A., Neale, B.M., Nagaswamy, U., Stevens, C., Lim, E., Bodea, C.A., Muzny, D., Reid, J.G., Banks, E., et al. (2013). Analysis of rare, exonic variation amongst subjects with autism spectrum disorders and population controls. *PLoS Genet.* *9*, e1003443.
- Liu, W., Morito, D., Takashima, S., Mineharu, Y., Kobayashi, H., Hitomi, T., Hashikata, H., Matsuura, N., Yamazaki, S., Toyoda, A., et al. (2011b). Identification of *RNF213* as a susceptibility gene for moyamoya disease and its possible role in vascular development. *PLoS One* *6*, e22542.
- Loeys, B.L., Chen, J., Neptune, E.R., Judge, D.P., Podowski, M., Holm, T., Meyers, J., Leitch, C.C., Katsanis, N., Sharifi, N., et al. (2005). A syndrome of altered cardiovascular, craniofacial, neurocognitive and skeletal development caused by mutations in *TGFBR1* or *TGFBR2*. *Nat. Genet.* *37*, 275–281.
- Loeys, B.L., Schwarze, U., Holm, T., Callewaert, B.L., Thomas, G.H., Pannu, H., De Backer, J.F., Oswald, G.L., Symoens, S., Manouvrier, S., et al. (2006). Aneurysm syndromes caused by mutations in the TGF-beta receptor. *N. Engl. J. Med.* *355*, 788–798.
- Lomonosova, E., and Chinnadurai, G. (2008). BH3-only proteins in apoptosis and beyond: an overview. *Oncogene* *27*, S2–S19.
- Longo, G.M., Xiong, W., Greiner, T.C., Zhao, Y., Fiotti, N., and Baxter, B.T. (2002). Matrix metalloproteinases 2 and 9 work in concert to produce aortic aneurysms. *J. Clin. Invest.* *110*, 625–632.
- Lord, R.S., and Bobryshev, Y.V. (1999). Clustering of dendritic cells in athero-prone areas of the aorta. *Atherosclerosis* *146*, 197–198.
- Lukacs, V., Mathur, J., Mao, R., Bayrak-Toydemir, P., Procter, M., Cahalan, S.M., Kim, H.J., Bandell, M., Longo, N., Day, R.W., et al. (2015). Impaired *PIEZO1* function in patients with a novel autosomal recessive congenital lymphatic dysplasia. *Nat. Commun.* *6*, 8329.
- Lundberg, J.O., Weitzberg, E., and Gladwin, M.T. (2008). The nitrate-nitrite-nitric oxide pathway in physiology and therapeutics. *Nat. Rev. Drug Discov.* *7*, 156–167.
- Lüscher, T.F., Keller, H.M., Imhof, H.G., Greminger, P., Kuhlmann, U., Largiadèr, F., Schneider, E., Schneider, J., and Vetter, W. (1986). Fibromuscular hyperplasia: extension of the disease and therapeutic outcome. Results of the University Hospital Zurich Cooperative Study on Fibromuscular Hyperplasia. *Nephron* *44 Suppl 1*, 109–114.
- Ma, L., Roman-Campos, D., Austin, E.D., Eyries, M., Sampson, K.S., Soubrier, F., Germain, M., Trégouët, D.-A., Borczuk, A., Rosenzweig, E.B., et al. (2013). A novel channelopathy in pulmonary arterial hypertension. *N. Engl. J. Med.* *369*, 351–361.

- MacCarrick, G., Iii, J.H.B., Bowdin, S., El-Hamamsy, I., Frischmeyer-Guerrero, P.A., Guerrero, A.L., Sponseller, P.D., Loeys, B., and Iii, H.C.D. (2014). Loeys-Dietz syndrome: a primer for diagnosis and management. *Genet. Med.* *16*, 576–587.
- Machado, R.D., Eickelberg, O., Elliott, C.G., Geraci, M.W., Hanaoka, M., Loyd, J.E., Newman, J.H., Phillips, J.A., Soubrier, F., Trembath, R.C., et al. (2009). Genetics and genomics of pulmonary arterial hypertension. *J. Am. Coll. Cardiol.* *54*, S32–S42.
- Madaric, J., Vulev, I., Bartunek, J., Mistrik, A., Verhamme, K., De Bruyne, B., and Riecanaky, I. (2005). Frequency of abdominal aortic aneurysm in patients >60 years of age with coronary artery disease. *Am. J. Cardiol.* *96*, 1214–1216.
- Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature* *456*, 18–21.
- Major, P., Genest, J., Cartier, P., and Kuchel, O. (1977). Hereditary fibromuscular dysplasia with renovascular hypertension. *Ann. Intern. Med.* *86*, 583.
- Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorf, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex diseases. *Nature* *461*, 747–753.
- Mao, D., Lee, J.K., VanVickle, S.J., and Thompson, R.W. (1999). Expression of collagenase-3 (MMP-13) in human abdominal aortic aneurysms and vascular smooth muscle cells in culture. *Biochem. Biophys. Res. Commun.* *261*, 904–910.
- Martinaud, C., Gisserot, O., Graffin, B., Gaillard, T., Brisou, P., Cynober, T., de Jaureguiberry, J.P., Delaunay, J., and Aguillon, P. (2008). Antiphospholipid antibodies in a family with dehydrated hereditary stomatocytosis. *Thromb. Res.* *122*, 572–575.
- Martinez, F.O., Gordon, S., Locati, M., and Mantovani, A. (2006). Transcriptional Profiling of the Human Monocyte-to-Macrophage Differentiation and Polarization: New Molecules and Patterns of Gene Expression. *J. Immunol.* *177*, 7303–7311.
- Massagué, J. (2008). TGFbeta in Cancer. *Cell* *134*, 215–230.
- Matsuda, A., Suzuki, Y., Honda, G., Muramatsu, S., Matsuzaki, O., Nagano, Y., Doi, T., Shimotohno, K., Harada, T., Nishida, E., et al. (2003). Large-scale identification and characterization of human genes that activate NF-kappaB and MAPK signaling pathways. *Oncogene* *22*, 3307–3318.
- Mattes, E., Davis, T.M.E., Yang, D.N., Ridley, D., Lund, H., and Norman, P.E. (1997). Prevalence of abdominal aortic aneurysms in men with diabetes. *Med. J. Aust.* *166*, 630–633.
- May, R., and Thurner, J. (1957). The Cause of the Predominantly Sinistral Occurrence of Thrombosis of the Pelvic Veins. *Angiology* *8*, 419–427.
- Mayr, B., and Montminy, M. (2001). Transcriptional regulation by the phosphorylation-dependent factor CREB. *Nat. Rev. Mol. Cell Biol.* *2*, 599–609.
- McBride, K.L., Pignatelli, R., Lewin, M., Ho, T., Fernbach, S., Menesses, A., Lam, W., Leal, S.M., Kaplan, N., Schliekelman, P., et al. (2005). Inheritance analysis of congenital left ventricular outflow tract obstruction malformations: Segregation, multiplex relative risk, and heritability. *Am. J. Med. Genet. A.* *134A*, 180–186.

- McClay, J.L., Aberg, K.A., Clark, S.L., Nerella, S., Kumar, G., Xie, L.Y., Hudson, A.D., Harada, A., Hultman, C.M., Magnusson, P.K.E., et al. (2014). A methylome-wide study of aging using massively parallel sequencing of the methyl-CpG-enriched genomic fraction from blood in over 700 subjects. *Hum. Mol. Genet.* *23*, 1175–1185.
- McCormack, L.J., Poutasse, E.F., Meaney, T.F., Noto, T.J., and Dustan, H.P. (1966). A pathologic-arteriographic correlation of renal arterial disease. *Am. Heart J.* *72*, 188–198.
- McCormack, L.J., Noto Jr., T.J., Meaney, T.F., Poutasse, E.F., and Dustan, H.P. (1967). Subadventitial fibroplasia of the renal artery, a disease of young women. *Am. Heart J.* *73*, 602–614.
- McDaniell, R., Warthen, D.M., Sanchez-Lara, P.A., Pai, A., Krantz, I.D., Piccoli, D.A., and Spinner, N.B. (2006). NOTCH2 mutations cause Alagille syndrome, a heterogeneous disorder of the notch signaling pathway. *Am. J. Hum. Genet.* *79*, 169–173.
- McElhinney, D.B., Krantz, I.D., Bason, L., Piccoli, D.A., Emerick, K.M., Spinner, N.B., and Goldmuntz, E. (2002). Analysis of cardiovascular phenotype and genotype-phenotype correlation in individuals with a JAG1 mutation and/or Alagille syndrome. *Circulation* *106*, 2567–2574.
- McMillan, W.D., Patterson, B.K., Keen, R.R., Shively, V.P., Cipollone, M., and Pearce, W.H. (1995). In situ localization and quantification of mRNA for 92-kD type IV collagenase and its inhibitor in aneurysmal, occlusive, and normal aorta. *Arterioscler. Thromb. Vasc. Biol.* *15*, 1139–1144.
- Merino, V.F., Todiras, M., Mori, M.A., Sales, V.M.T., Fonseca, R.G., Saul, V., Tenner, K., Bader, M., and Pesquero, J.B. (2009). Predisposition to atherosclerosis and aortic aneurysms in mice deficient in kinin B1 receptor and apolipoprotein E. *J. Mol. Med. Berl. Ger.* *87*, 953–963.
- Merkel, K.H., Ginsberg, P.L., Parker, J.C., and Post, M.J. (1978). Cerebrovascular disease in sickle cell anemia: a clinical, pathological and radiological correlation. *Stroke J. Cereb. Circ.* *9*, 45–52.
- Mesh, C.L., Baxter, B.T., Pearce, W.H., Chisholm, R.L., McGee, G.S., and Yao, J.S. (1992). Collagen and elastin gene expression in aortic aneurysms. *Surgery* *112*, 256–261; discussion 261–262.
- Mettinger, K.L., and Ericson, K. (1982). Fibromuscular dysplasia and the brain. I. Observations on angiographic, clinical and genetic characteristics. *Stroke J. Cereb. Circ.* *13*, 46–52.
- Mi, H., and Thomas, P. (2009). PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol. Biol. Clifton NJ* *563*, 123–140.
- Michelis, K.C., Olin, J.W., Kadian-Dodov, D., d'Escamard, V., and Kovacic, J.C. (2014). Coronary artery manifestations of fibromuscular dysplasia. *J. Am. Coll. Cardiol.* *64*, 1033–1046.
- Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P.A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L.M., and Ding, W. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* *266*, 66–71.
- Milewicz, D.M., Pyeritz, R.E., Crawford, E.S., and Byers, P.H. (1992). Marfan syndrome: defective synthesis, secretion, and extracellular matrix formation of fibrillin by cultured dermal fibroblasts. *J. Clin. Invest.* *89*, 79–86.
- Miller, D.R., Rickles, F.R., Lichtman, M.A., La Celle, P.L., Bates, J., and Weed, R.I. (1971). A new variant of hereditary hemolytic anemia with stomatocytosis and erythrocyte cation abnormality. *Blood* *38*, 184–204.

- Miller, E.A., Pankow, J.S., Millikan, R.C., Bray, M.S., Ballantyne, C.M., Bell, D.A., Heiss, G., and Li, R. (2003). Glutathione-S-transferase genotypes, smoking, and their association with markers of inflammation, hemostasis, and endothelial function: the atherosclerosis risk in communities (ARIC) study. *Atherosclerosis* 171, 265–272.
- Mitchinson, M.J. (1986). Retroperitoneal fibrosis revisited. *Arch. Pathol. Lab. Med.* 110, 784–786.
- Moceri, P., Albuissou, J., Saint-Faust, M., Casagrande, F., Giuliano, F., Devos, C., Benoit, P., Hugues, N., Ducreux, D., Cerboni, P., et al. (2013). Arterial tortuosity syndrome: early diagnosis and association with venous tortuosity. *J. Am. Coll. Cardiol.* 61, 783.
- Mochizuki, S. (2002). Signaling Pathways Transduced through the Elastin Receptor Facilitate Proliferation of Arterial Smooth Muscle Cells. *J. Biol. Chem.* 277, 44854–44863.
- Mokri, B., Sundt, T.M., Houser, O.W., and Piepgras, D.G. (1986). Spontaneous dissection of the cervical internal carotid artery. *Ann. Neurol.* 19, 126–138.
- Molina, C.A., Foulkes, N.S., Lalli, E., and Sassone-Corsi, P. (1993). Inducibility and negative autoregulation of CREM: an alternative promoter directs the expression of ICER, an early response repressor. *Cell* 75, 875–886.
- Morgenthaler, S., and Thilly, W.G. (2007). A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). *Mutat. Res.* 615, 28–56.
- Mosorin, M., Juvonen, J., Biancari, F., Satta, J., Surcel, H.M., Leinonen, M., Saikku, P., and Juvonen, T. (2001). Use of doxycycline to decrease the growth rate of abdominal aortic aneurysms: a randomized, double-blind, placebo-controlled pilot study. *J. Vasc. Surg.* 34, 606–610.
- Moxon, J.V., Parr, A., Emeto, T.I., Walker, P., Norman, P.E., and Golledge, J. (2010). Diagnosis and monitoring of abdominal aortic aneurysm: Current status and future prospects. *Curr. Probl. Cardiol.* 35, 512–548.
- Munot, P., Saunders, D.E., Milewicz, D.M., Regalado, E.S., Ostergaard, J.R., Braun, K.P., Kerr, T., Lichtenbelt, K.D., Philip, S., Rittey, C., et al. (2012). A novel distinctive cerebrovascular phenotype is associated with heterozygous Arg179 ACTA2 mutations. *Brain* 135, 2506–2514.
- Murata, T., Ushikubi, F., Matsuoka, T., Hirata, M., Yamasaki, A., Sugimoto, Y., Ichikawa, A., Aze, Y., Tanaka, T., Yoshida, N., et al. (1997). Altered pain perception and inflammatory response in mice lacking prostacyclin receptor. *Nature* 388, 678–682.
- Murphy, P.A., Lam, M.T.Y., Wu, X., Kim, T.N., Vartanian, S.M., Bollen, A.W., Carlson, T.R., and Wang, R.A. (2008). Endothelial Notch4 signaling induces hallmarks of brain arteriovenous malformations in mice. *Proc. Natl. Acad. Sci. U. S. A.* 105, 10901–10906.
- Musunuru, K., Pirruccello, J.P., Do, R., Peloso, G.M., Guiducci, C., Sougnez, C., Garimella, K.V., Fisher, S., Abreu, J., Barry, A.J., et al. (2010a). Exome sequencing, ANGPTL3 mutations, and familial combined hypolipidemia. *N. Engl. J. Med.* 363, 2220–2227.
- Musunuru, K., Strong, A., Frank-Kamenetsky, M., Lee, N.E., Ahfeldt, T., Sachs, K.V., Li, X., Li, H., Kuperwasser, N., Ruda, V.M., et al. (2010b). From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature* 466, 714–719.
- Nagase, H., and Woessner, J.F. (1999). Matrix Metalloproteinases. *J. Biol. Chem.* 274, 21491–21494.

- Nahum, S., Morice-Picard, F., Taieb, A., and Sprecher, E. (2011). A novel mutation in LPAR6 causes autosomal recessive hypotrichosis of the scalp. *Clin. Exp. Dermatol.* *36*, 188–194.
- Nakano, K., Wada, K., Nomura, R., Nemoto, H., Inaba, H., Kojima, A., Naka, S., Hokamura, K., Mukai, T., Nakajima, A., et al. (2011). Characterization of aortic aneurysms in cardiovascular disease patients harboring *Porphyromonas gingivalis*. *Oral Dis.* *17*, 370–378.
- Narumiya, S., Sugimoto, Y., and Ushikubi, F. (1999). Prostanoid receptors: structures, properties, and functions. *Physiol. Rev.* *79*, 1193–1226.
- Naydeck, B.L., Sutton-Tyrrell, K., Schiller, K.D., Newman, A.B., and Kuller, L.H. (1999). Prevalence and risk factors for abdominal aortic aneurysms in older adults with and without isolated systolic hypertension. *Am. J. Cardiol.* *83*, 759–764.
- Neptune, E.R., Frischmeyer, P.A., Arking, D.E., Myers, L., Bunton, T.E., Gayraud, B., Ramirez, F., Sakai, L.Y., and Dietz, H.C. (2003). Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. *Nat. Genet.* *33*, 407–411.
- Newman, K.M., Ogata, Y., Malon, A.M., Irizarry, E., Gandhi, R.H., Nagase, H., and Tilson, M.D. (1994). Identification of matrix metalloproteinases 3 (stromelysin-1) and 9 (gelatinase B) in abdominal aortic aneurysm. *Arterioscler. Thromb. J. Vasc. Biol. Am. Heart Assoc.* *14*, 1315–1320.
- Neymark, E., LaBerge, J.M., Hirose, R., Melzer, J.S., Kerlan, R.K., Wilson, M.W., and Gordon, R.L. (2000). Arteriographic detection of renovascular disease in potential renal donors: incidence and effect on donor surgery. *Radiology* *214*, 755–760.
- Ng PC, Levy S, Huang J, Stockwell TB, Walenz BP, Li K, Axelrod N, Busam DA, Strausberg RL, Venter JC (2008). Genetic variation in an individual human exome. *PLoS Genet.* *15*;4(8).
- Ng, S.B., Turner, E.H., Robertson, P.D., Flygare, S.D., Bigham, A.W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E.E., et al. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* *461*, 272–276.
- Ng, S.B., Buckingham, K.J., Lee, C., Bigham, A.W., Tabor, H.K., Dent, K.M., Huff, C.D., Shannon, P.T., Jabs, E.W., Nickerson, D.A., et al. (2010a). Exome sequencing identifies the cause of a mendelian disorder. *Nat. Genet.* *42*, 30–35.
- Ng, S.B., Bigham, A.W., Buckingham, K.J., Hannibal, M.C., McMillin, M.J., Gildersleeve, H.I., Beck, A.E., Tabor, H.K., Cooper, G.M., Mefford, H.C., et al. (2010b). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* *42*, 790–793.
- Niessen, P., Rensen, S., van Deursen, J., De Man, J., De Laet, A., Vanderwinden, J.-M., Wedel, T., Baker, D., Doevendans, P., Hofker, M., et al. (2005). Smoothelin-a is essential for functional intestinal smooth muscle contractility in mice. *Gastroenterology* *129*, 1592–1601.
- Norman, P.E., Davis, T.M.E., Le, M.T.Q., and Golledge, J. (2007). Matrix Biology of Abdominal Aortic Aneurysms in Diabetes: Mechanisms Underlying the Negative Association. *Connect. Tissue Res.* *48*, 125–131.
- Ocana, E. (2003). Characterisation of T and B lymphocytes infiltrating abdominal aortic aneurysms. *Atherosclerosis* *170*, 39–48.

- Oderich, G.S., Panneton, J.M., Bower, T.C., Lindor, N.M., Cherry, K.J., Noel, A.A., Kalra, M., Sullivan, T., and Gloviczki, P. (2005). The spectrum, management and clinical outcome of Ehlers-Danlos syndrome type IV: A 30-year experience. *J. Vasc. Surg.* *42*, 98–106.
- Oderich, G.S., Sullivan, T.M., Bower, T.C., Gloviczki, P., Miller, D.V., Babovic-Vuksanovic, D., Macedo, T.A., and Stanson, A. (2007). Vascular abnormalities in patients with neurofibromatosis syndrome type I: clinical spectrum, management, and results. *J. Vasc. Surg.* *46*, 475–484.
- Ogata, T., Shibamura, H., Tromp, G., Sinha, M., Goddard, K.A.B., Sakalihan, N., Limet, R., MacKean, G.L., Arthur, C., Sueda, T., et al. (2005a). Genetic analysis of polymorphisms in biologically relevant candidate genes in patients with abdominal aortic aneurysms. *J. Vasc. Surg.* *41*, 1036–1042.
- Ogata, T., MacKean, G.L., Cole, C.W., Arthur, C., Andreou, P., Tromp, G., and Kuivaniemi, H. (2005b). The lifetime prevalence of abdominal aortic aneurysms among siblings of aneurysm patients is eightfold higher than among siblings of spouses: an analysis of 187 aneurysm families in Nova Scotia, Canada. *J. Vasc. Surg.* *42*, 891–897.
- Ohtsubo, H., Ichiki, T., Miyazaki, R., Inanaga, K., Imayama, I., Hashiguchi, Y., Sadoshima, J., and Sunagawa, K. (2007). Inducible cAMP early repressor inhibits growth of vascular smooth muscle cell. *Arterioscler. Thromb. Vasc. Biol.* *27*, 1549–1555.
- Olin, J.W., and Sealove, B.A. (2011). Diagnosis, management, and future developments of fibromuscular dysplasia. *J. Vasc. Surg.* *53*, 826–836.e1.
- Olin, J.W., Froehlich, J., Gu, X., Bacharach, J.M., Eagle, K., Gray, B.H., Jaff, M.R., Kim, E.S.H., Mace, P., Matsumoto, A.H., et al. (2012). The United States Registry for Fibromuscular Dysplasia: results in the first 447 patients. *Circulation* *125*, 3182–3190.
- Olin, J.W., Gornik, H.L., Bacharach, J.M., Biller, J., Fine, L.J., Gray, B.H., Gray, W.A., Gupta, R., Hamburg, N.M., Katzen, B.T., et al. (2014). Fibromuscular dysplasia: state of the science and critical unanswered questions: a scientific statement from the American Heart Association. *Circulation* *129*, 1048–1078.
- Ormezzano, O., Baguet, J.-P., Thony, F., Blin, D., Chavanon, O., Toussaint, B., Trocmé, C., Vanzetto, G., and Faure, P. (2010). Aminoterminal propeptide of type III procollagen (PIIINP) is associated with ascending aortic aneurysm growth rate. *Int. J. Cardiol.* *145*, 379–380.
- O’Roak, B.J., Stessman, H.A., Boyle, E.A., Witherspoon, K.T., Martin, B., Lee, C., Vives, L., Baker, C., Hiatt, J.B., Nickerson, D.A., et al. (2014). Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat. Commun.* *5*, 5595.
- Ouchi, Y., Tagawa, H., Yamakado, M., Takanashi, R., and Tanaka, S. (1989). Clinical significance of cerebral aneurysm in renovascular hypertension due to fibromuscular dysplasia: two cases in siblings. *Angiology* *40*, 581–588.
- Padmanabhan, S., Hastie, C., Prabhakaran, D., and Dominczak, A.F. (2010). Genomic approaches to coronary artery disease. *Indian J. Med. Res.* *132*, 567–578.
- Palubinskas, A.J., and Ripley, H.R. (1964). FIBROMUSCULAR HYPERPLASIA IN EXTRARENAL ARTERIES. *Radiology* *82*, 451–455.
- Pannier-Moreau, I., Grimbert, P., Fiquet-Kempf, B., Vuagnat, A., Jeunemaitre, X., Corvol, P., and Plouin, P.F. (1997). Possible familial origin of multifocal renal artery fibromuscular dysplasia. *J. Hypertens.* *15*, 1797–1801.

- Parthasarathy, S., Steinbrecher, U.P., Barnett, J., Witztum, J.L., and Steinberg, D. (1985). Essential role of phospholipase A2 activity in endothelial cell-induced modification of low density lipoprotein. *Proc. Natl. Acad. Sci. U. S. A.* *82*, 3000–3004.
- Parums, D.V. (1990). The spectrum of chronic periaortitis. *Histopathology* *16*, 423–431.
- Pate, G.E., Lowe, R., and Buller, C.E. (2005). Fibromuscular dysplasia of the coronary and renal arteries? *Catheter. Cardiovasc. Interv. Off. J. Soc. Card. Angiogr. Interv.* *64*, 138–145.
- Pays, E., Vanhollebeke, B., Uzureau, P., Lecordier, L., and Pérez-Morga, D. (2014). The molecular arms race between African trypanosomes and humans. *Nat. Rev. Microbiol.* *12*, 575–584.
- Pepin, M., Schwarze, U., Superti-Furga, A., and Byers, P.H. (2000). Clinical and genetic features of Ehlers-Danlos syndrome type IV, the vascular type. *N. Engl. J. Med.* *342*, 673–680.
- Pepin, M.G., Schwarze, U., Rice, K.M., Liu, M., Leistriz, D., and Byers, P.H. (2014). Survival is affected by mutation type and molecular mechanism in vascular Ehlers-Danlos syndrome (EDS type IV). *Genet. Med. Off. J. Am. Coll. Med. Genet.* *16*, 881–888.
- Perdu, J., Gimenez-Roqueplo, A.-P., Boutouyrie, P., Beaujour, S., Laloux, B., Nau, V., Fiquet-Kempf, B., Emmerich, J., Tichet, J., Plouin, P.-F., et al. (2006). Alpha1-antitrypsin gene polymorphisms are not associated with renal arterial fibromuscular dysplasia. *J. Hypertens.* *24*, 705–710.
- Perdu, J., Boutouyrie, P., Bourgain, C., Stern, N., Laloux, B., Bozec, E., Azizi, M., Bonaiti-Pellié, C., Plouin, P.-F., Laurent, S., et al. (2007). Inheritance of arterial lesions in renal fibromuscular dysplasia. *J. Hum. Hypertens.* *21*, 393–400.
- Perez-Iratxeta, C., Bork, P., and Andrade, M.A. (2002). Association of genes to genetically inherited diseases using data mining. *Nat. Genet.* *31*, 316–319.
- Perez-Iratxeta, C., Bork, P., and Andrade-Navarro, M.A. (2007). Update of the G2D tool for prioritization of gene candidates to inherited diseases. *Nucleic Acids Res.* *35*, W212–W216.
- Perko, M.J., Schroeder, T.V., Olsen, P.S., Jensen, L.P., and Lorentzen, J.E. (1993). Natural history of abdominal aortic aneurysm: a survey of 63 patients treated nonoperatively. *Ann. Vasc. Surg.* *7*, 113–116.
- Perry, N.A., Vitolo, M.I., Martin, S.S., and Kontogianni-Konstantopoulos, A. (2014). Loss of the obscurin-RhoGEF downregulates RhoA signaling and increases microtentacle formation and attachment of breast epithelial cells. *Oncotarget* *5*, 8558–8568.
- Persu, A., Touzé, E., Mousseaux, E., Barral, X., Joffre, F., and Plouin, P.-F. (2012). Diagnosis and management of fibromuscular dysplasia: an expert consensus: EXPERT CONSENSUS ON FIBROMUSCULAR DYSPLASIA. *Eur. J. Clin. Invest.* *42*, 338–347.
- Petersen, E., Boman, J., Wågberg, F., and Angquist, K.A. (2002). Presence of *Chlamydia pneumoniae* in abdominal aortic aneurysms is not associated with increased activity of matrix metalloproteinases. *Eur. J. Vasc. Endovasc. Surg. Off. J. Eur. Soc. Vasc. Surg.* *24*, 365–369.
- Phillips, J.K., McLean, A.J., and Hill, C.E. (1998). Receptors involved in nerve-mediated vasoconstriction in small arteries of the rat hepatic mesentery. *Br. J. Pharmacol.* *124*, 1403–1412.

- Pierce, S.B., Walsh, T., Chisholm, K.M., Lee, M.K., Thornton, A.M., Fiumara, A., Opitz, J.M., Levy-Lahad, E., Klevit, R.E., and King, M.-C. (2010a). Mutations in the DBP-deficiency protein HSD17B4 cause ovarian dysgenesis, hearing loss, and ataxia of Perrault Syndrome. *Am. J. Hum. Genet.* *87*, 282–288.
- Pilz, R.B., and Casteel, D.E. (2003). Regulation of gene expression by cyclic GMP. *Circ. Res.* *93*, 1034–1046.
- Platt, O.S., Lux, S.E., and Nathan, D.G. (1981). Exercise-induced hemolysis in xerocytosis. Erythrocyte dehydration and shear sensitivity. *J. Clin. Invest.* *68*, 631–638.
- Plouin, P.-F., Perdu, J., La Batide-Alanore, A., Boutouyrie, P., Gimenez-Roqueplo, A.-P., and Jeunemaitre, X. (2007). Fibromuscular dysplasia. *Orphanet J. Rare Dis.* *2*, 28.
- Poetsch, M., Wiegand, A., Harder, M., Blöhm, R., Rakotomavo, N., Freitag-Wolf, S., and von Wurmb-Schwark, N. (2013). Determination of population origin: a comparison of autosomal SNPs, Y-chromosomal and mtDNA haplogroups using a Malagasy population as example. *Eur. J. Hum. Genet.* *21*, 1423–1428.
- Pollak, M.R., Chou, Y.H., Cerda, J.J., Steinmann, B., La Du, B.N., Seidman, J.G., and Seidman, C.E. (1993). Homozygosity mapping of the gene for alkaptonuria to chromosome 3q2. *Nat. Genet.* *5*, 201–204.
- Poloskey, S.L., Kim, E.S., Sanghani, R., Al-Quthami, A.H., Arscott, P., Moran, R., Rigelsky, C.M., and Gornik, H.L. (2012). Low yield of genetic testing for known vascular connective tissue disorders in patients with fibromuscular dysplasia. *Vasc. Med.* *17*, 371–378.
- Powell, J.T., Vine, N., and Crossman, M. (1992). On the accumulation of D-aspartate in elastin and other proteins of the ageing aorta. *Atherosclerosis* *97*, 201–208.
- Price, A.L., Kryukov, G.V., de Bakker, P.I.W., Purcell, S.M., Staples, J., Wei, L.-J., and Sunyaev, S.R. (2010). Pooled association tests for rare variants in exon-resequencing studies. *Am. J. Hum. Genet.* *86*, 832–838.
- Probst, V., Le Scouarnec, S., Legendre, A., Jousseau, V., Jaafar, P., Nguyen, J.-M., Chaventré, A., Le Marec, H., and Schott, J.-J. (2006). Familial aggregation of calcific aortic valve stenosis in the western part of France. *Circulation* *113*, 856–860.
- Pyo, R., Lee, J.K., Shipley, J.M., Curci, J.A., Mao, D., Ziporin, S.J., Ennis, T.L., Shapiro, S.D., Senior, R.M., and Thompson, R.W. (2000). Targeted gene disruption of matrix metalloproteinase-9 (gelatinase B) suppresses development of experimental abdominal aortic aneurysms. *J. Clin. Invest.* *105*, 1641–1649.
- Quillon, A., Fromy, B., and Debret, R. (2015). Endothelium microenvironment sensing leading to nitric oxide mediated vasodilation: A review of nervous and biomechanical signals. *Nitric Oxide* *45*, 20–26.
- Ramshaw, A.L., and Parums, D.V. (1990). Immunohistochemical characterization of inflammatory cells associated with advanced atherosclerosis. *Histopathology* *17*, 543–552.
- Ranade, S.S., Qiu, Z., Woo, S.-H., Hur, S.S., Murthy, S.E., Cahalan, S.M., Xu, J., Mathur, J., Bandell, M., Coste, B., et al. (2014). Piezo1, a mechanically activated ion channel, is required for vascular development in mice. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 10347–10352.
- Randers, E., Jønler, M., Lund, A.M., and Danielsen, H. (1998). Atraumatic loss of a kidney in a patient with alpha1-antitrypsin deficiency. *Nephron* *80*, 71–75.
- Rauen, K.A. (2013). The RASopathies. *Annu. Rev. Genomics Hum. Genet.* *14*, 355–369.

- Razakandrainibe, R., Pelleau, S., Grau, G.E., and Jambou, R. (2012). Antigen presentation by endothelial cells: what role in the pathophysiology of malaria? *Trends Parasitol.* *28*, 151–160.
- Reed, D., Reed, C., Stemmermann, G., and Hayashi, T. (1992). Are aortic aneurysms caused by atherosclerosis? *Circulation* *85*, 205–211.
- Rees, D.C., Portmann, B., Ball, C., Mieli-Vergani, G., Nicolaou, A., Chetty, M.C., and Stewart, G.W. (2004). Dehydrated hereditary stomatocytosis is associated with neonatal hepatitis. *Br. J. Haematol.* *126*, 272–276.
- Regalado, E.S., Guo, D. -c., Villamizar, C., Avidan, N., Gilchrist, D., McGillivray, B., Clarke, L., Bernier, F., Santos-Cortez, R.L., Leal, S.M., et al. (2011). Exome Sequencing Identifies SMAD3 Mutations as a Cause of Familial Thoracic Aortic Aneurysm and Dissection With Intracranial and Other Arterial Aneurysms. *Circ. Res.* *109*, 680–686.
- Reilly, J.M., Sicard, G.A., and Lucore, C.L. (1994). Abnormal expression of plasminogen activators in aortic aneurysmal and occlusive disease. *J. Vasc. Surg.* *19*, 865–872.
- Ren, Y., Guo, L., Tang, X., Apparsundaram, S., Kitson, C., Deguzman, J., Fuentes, M.E., Coyle, L., Majmudar, R., Allard, J., et al. (2013). Comparing the differential effects of LPA on the barrier function of human pulmonary endothelial cells. *Microvasc. Res.* *85*, 59–67.
- Renard, M., Holm, T., Veith, R., Callewaert, B.L., Adès, L.C., Baspinar, O., Pickart, A., Dasouki, M., Hoyer, J., Rauch, A., et al. (2010). Altered TGF β signaling and cardiovascular manifestations in patients with autosomal recessive cutis laxa type I caused by fibulin-4 deficiency. *Eur. J. Hum. Genet.* *18*, 895–901.
- Resnick, N., Yahav, H., Shay-Salit, A., Shushy, M., Schubert, S., Zilberman, L.C.M., and Wofovitz, E. (2003). Fluid shear stress and the vascular endothelium: for better and for worse. *Prog. Biophys. Mol. Biol.* *81*, 177–199.
- Reusch, J.E.B., and Watson, P.A. (2004). Loss of CREB regulation of vascular smooth muscle cell quiescence in diabetes. *Rev. Endocr. Metab. Disord.* *5*, 209–219.
- Ritelli, M., Chiarelli, N., Dordoni, C., Quinzani, S., Venturini, M., Maroldi, R., Calzavara-Pinton, P., and Colombi, M. (2014). Further delineation of Loeys-Dietz syndrome type 4 in a family with mild vascular involvement and a TGFB2 splicing mutation. *BMC Med. Genet.* *15*, 91.
- Rivas, M.A., Pirinen, M., Conrad, D.F., Lek, M., Tsang, E.K., Karczewski, K.J., Maller, J.B., Kukurba, K.R., DeLuca, D.S., Fromer, M., et al. (2015). Human genomics. Effect of predicted protein-truncating genetic variants on the human transcriptome. *Science* *348*, 666–669.
- Roach, J.C., Glusman, G., Smit, A.F.A., Huff, C.D., Hubley, R., Shannon, P.T., Rowen, L., Pant, K.P., Goodman, N., Bamshad, M., et al. (2010). Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* *328*, 636–639.
- Robbins, S.H., Walzer, T., Dembélé, D., Thibault, C., Defays, A., Bessou, G., Xu, H., Vivier, E., Sellars, M., Pierre, P., et al. (2008). Novel insights into the relationships between dendritic cell subsets in human and mouse revealed by genome-wide expression profiling. *Genome Biol.* *9*, R17.
- Robin, N.H., Feldman, G.J., Mitchell, H.F., Lorenz, P., Wilroy, R.S., Zackai, E.H., Allanson, J.E., Reich, E.W., Pfeiffer, R.A., and Clarke, L.A. (1994). Linkage of Pfeiffer syndrome to chromosome 8 centromere and evidence for genetic heterogeneity. *Hum. Mol. Genet.* *3*, 2153–2158.

- Robinson, P.N., Köhler, S., Oellrich, A., Sanger Mouse Genetics Project, Wang, K., Mungall, C.J., Lewis, S.E., Washington, N., Bauer, S., Seelow, D., et al. (2014). Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome Res.* *24*, 340–348.
- Roder, C., Nayak, N.R., Khan, N., Tatagiba, M., Inoue, I., and Krischek, B. (2010). Genetics of Moyamoya disease. *J. Hum. Genet.* *55*, 711–716.
- Rose, A.G., and Dent, D.M. (1981). Inflammatory variant of abdominal atherosclerotic aneurysm. *Arch. Pathol. Lab. Med.* *105*, 409–413.
- Ross, R. (1999). Atherosclerosis--an inflammatory disease. *N. Engl. J. Med.* *340*, 115–126.
- Rudic, R.D. (2005). COX-2-Derived Prostacyclin Modulates Vascular Remodeling. *Circ. Res.* *96*, 1240–1247.
- Rushton, A.R. (1980). The genetics of fibromuscular dysplasia. *Arch. Intern. Med.* *140*, 233–236.
- Salem, J.-E., Bruguere, E., Iserin, L., Guiochon-Mantel, A., and Plouin, P.-F. (2012). Hypertension and aortorenal disease in Alagille syndrome: *J. Hypertens.* *30*, 1300–1306.
- Salem, M.K., Rayt, H.S., Hussey, G., Rafelt, S., Nelson, C.P., Sayers, R.D., Naylor, A.R., and Nasim, A. (2009). Should Asian men be included in abdominal aortic aneurysm screening programmes? *Eur. J. Vasc. Endovasc. Surg. Off. J. Eur. Soc. Vasc. Surg.* *38*, 748–749.
- Salyer, W.R., and Salyer, D.C. (1974). The vascular lesions of neurofibromatosis. *Angiology* *25*, 510–519.
- Sang, C.N., Whelton, P.K., Hamper, U.M., Connolly, M., Kadir, S., White, R.I., Sanders, R., Liang, K.Y., and Bias, W. (1989). Etiologic factors in renovascular fibromuscular dysplasia. A case-control study. *Hypertension* *14*, 472–479.
- Saracini, C., Bolli, P., Sticchi, E., Pratesi, G., Pulli, R., Sofi, F., Pratesi, C., Gensini, G.F., Abbate, R., and Giusti, B. (2012). Polymorphisms of genes involved in extracellular matrix remodeling and abdominal aortic aneurysm. *J. Vasc. Surg.* *55*, 171–179.e2.
- Saratzis, A., Abbas, A.A., Kiskinis, D., Melas, N., Saratzis, N., and Kitas, G.D. (2011). Abdominal aortic aneurysm: a review of the genetic basis. *Angiology* *62*, 18–32.
- Saratzis, A., Bown, M., Wild, B., Sayers, R.D., Nightingale, P., Smith, J., Johnson, C., and Kitas, G. (2014). C-reactive protein polymorphism rs3091244 is associated with abdominal aortic aneurysm. *J. Vasc. Surg.* *60*, 1332–1339.
- Sato, H., Kato, R., Isogai, Y., Saka, G., Ohtsuki, M., Taketomi, Y., Yamamoto, K., Tsutsumi, K., Yamada, J., Masuda, S., et al. (2008). Analyses of group III secreted phospholipase A2 transgenic mice reveal potential participation of this enzyme in plasma lipoprotein modification, macrophage foam cell formation, and atherosclerosis. *J. Biol. Chem.* *283*, 33483–33497.
- Savard, S., Steichen, O., Azarine, A., Azizi, M., Jeunemaitre, X., and Plouin, P.-F. (2012). Association Between 2 Angiographic Subtypes of Renal Artery Fibromuscular Dysplasia and Clinical Characteristics. *Circulation* *126*, 3062–3069.
- Saw, J., Aymong, E., Sedlak, T., Buller, C.E., Starovoytov, A., Ricci, D., Robinson, S., Vuurmans, T., Gao, M., Humphries, K., et al. (2014). Spontaneous coronary artery dissection: association with predisposing arteriopathies and precipitating stressors and cardiovascular outcomes. *Circ. Cardiovasc. Interv.* *7*, 645–655.

- Schauer, I.E., Knaub, L.A., Lloyd, M., Watson, P.A., Gliwa, C., Lewis, K.E., Chait, A., Klemm, D.J., Gunter, J.M., Bouchard, R., et al. (2010). CREB Downregulation in vascular disease: a common response to cardiovascular risk. *Arterioscler. Thromb. Vasc. Biol.* *30*, 733–741.
- Schievink, W.I. (2001). Spontaneous dissection of the carotid and vertebral arteries. *N. Engl. J. Med.* *344*, 898–906.
- Schievink, W.I., and Björnsson, J. (1996). Fibromuscular dysplasia of the internal carotid artery: a clinicopathological study. *Clin. Neuropathol.* *15*, 2–6.
- Schievink, W.I., Meyer, F.B., Parisi, J.E., and Wijdicks, E.F. (1998a). Fibromuscular dysplasia of the internal carotid artery associated with alpha1-antitrypsin deficiency. *Neurosurgery* *43*, 229–233; discussion 233–234.
- Schievink, W.I., Katzmann, J.A., and Piepgras, D.G. (1998b). Alpha-1-antitrypsin deficiency in spontaneous intracranial arterial dissections. *Cerebrovasc. Dis. Basel Switz.* *8*, 42–44.
- Schmolke, M., Viemann, D., Roth, J., and Ludwig, S. (2009). Essential impact of NF-kappaB signaling on the H5N1 influenza A virus-induced transcriptome. *J. Immunol. Baltim. Md 1950* *183*, 5180–5189.
- Schneiderman, J., Bordin, G.M., Engelberg, I., Adar, R., Seiffert, D., Thinnis, T., Bernstein, E.F., Dilley, R.B., and Loskutoff, D.J. (1995). Expression of fibrinolytic genes in atherosclerotic abdominal aortic aneurysm wall. A possible mechanism for aneurysm expansion. *J. Clin. Invest.* *96*, 639–645.
- Schuster, J.M., and Roberts, T.S. (1999). Symptomatic moyamoya disease and aortic coarctation in a patient with Noonan's syndrome: strategies for management. *Pediatr. Neurosurg.* *30*, 206–210.
- Seelow, D., Schuelke, M., Hildebrandt, F., and Nürnberg, P. (2009). HomozygosityMapper--an interactive approach to homozygosity mapping. *Nucleic Acids Res.* *37*, W593–W599.
- Seidl, M.D., Steingraber, A.K., Wolf, C.T., Sur, T.M.H., Hildebrandt, I., Witten, A., Stoll, M., Fischer, J.W., Schmitz, W., and Müller, F.U. (2015). Transcription factor cAMP response element modulator (Crem) restrains Pdgf-dependent proliferation of vascular smooth muscle cells in mice. *Pflug. Arch. - Eur. J. Physiol.* *467*, 2165–2177.
- Senior, R.M., Griffin, G.L., and Mecham, R.P. (1980). Chemotactic activity of elastin-derived peptides. *J. Clin. Invest.* *66*, 859–862.
- Senior, R.M., Griffin, G.L., Mecham, R.P., Wrenn, D.S., Prasad, K.U., and Urry, D.W. (1984). Val-Gly-Val-Ala-Pro-Gly, a repeating peptide in elastin, is chemotactic for fibroblasts and monocytes. *J. Cell Biol.* *99*, 870–874.
- Sharma, A., Yu, C., Leung, C., Trane, A., Lau, M., Utokaparch, S., Shaheen, F., Sheibani, N., and Bernatchez, P. (2010). A new role for the muscle repair protein dysferlin in endothelial cell adhesion and angiogenesis. *Arterioscler. Thromb. Vasc. Biol.* *30*, 2196–2204.
- Shi, G.P., Sukhova, G.K., Grubb, A., Ducharme, A., Rhode, L.H., Lee, R.T., Ridker, P.M., Libby, P., and Chapman, H.A. (1999). Cystatin C deficiency in human atherosclerosis and aortic aneurysms. *J. Clin. Invest.* *104*, 1191–1197.
- Shibamura, H., Olson, J.M., van Vlijmen-Van Keulen, C., Buxbaum, S.G., Dudek, D.M., Tromp, G., Ogata, T., Skunca, M., Sakalihan, N., Pals, G., et al. (2004). Genome scan for familial abdominal aortic aneurysm using sex and family history as covariates suggests genetic heterogeneity and identifies linkage to chromosome 19q13. *Circulation* *109*, 2103–2108.

- Shimizu, K., Libby, P., and Mitchell, R.N. (2005). Local cytokine environments drive aneurysm formation in allografted aortas. *Trends Cardiovasc. Med.* *15*, 142–148.
- Shimkets, R.A., Warnock, D.G., Bositis, C.M., Nelson-Williams, C., Hansson, J.H., Schambelan, M., Gill, J.R., Ulick, S., Milora, R.V., and Findling, J.W. (1994). Liddle's syndrome: heritable human hypertension caused by mutations in the beta subunit of the epithelial sodium channel. *Cell* *79*, 407–414.
- Silhol, F., Sarlon-Bartoli, G., Daniel, L., Bartoli, J.M., Cohen, S., Lepidi, H., Piquet, P., Bartoli, M.A., and Vaisse, B. (2015). Intranuclear expression of progesterone receptors in smooth muscle cells of renovascular fibromuscular dysplasia: a pilot study. *Ann. Vasc. Surg.* *29*, 830–835.
- Singh, K., Bønaa, K.H., Jacobsen, B.K., Bjørk, L., and Solberg, S. (2001). Prevalence of and risk factors for abdominal aortic aneurysms in a population-based study : The Tromsø Study. *Am. J. Epidemiol.* *154*, 236–244.
- Sirmaci, A., Walsh, T., Akay, H., Spiliopoulos, M., Sakalar, Y.B., Hasaneferendioglu-Bayrak, A., Duman, D., Farooq, A., King, M.-C., and Tekin, M. (2010). MASP1 mutations in patients with facial, umbilical, coccygeal, and auditory findings of Carnevale, Malpuech, OSA, and Michels syndromes. *Am. J. Hum. Genet.* *87*, 679–686.
- Slavin, R.E. (2009). Segmental arterial mediolysis: course, sequelae, prognosis, and pathologic-radiologic correlation. *Cardiovasc. Pathol. Off. J. Soc. Cardiovasc. Pathol.* *18*, 352–360.
- Smith, E.E., and Malik, H.S. (2009). The apolipoprotein L family of programmed cell death and immunity genes rapidly evolved in primates at discrete sites of host-pathogen interactions. *Genome Res.* *19*, 850–858.
- Sobreira, N.L.M., Cirulli, E.T., Avramopoulos, D., Wohler, E., Oswald, G.L., Stevens, E.L., Ge, D., Shianna, K.V., Smith, J.P., Maia, J.M., et al. (2010). Whole-genome sequencing of a single proband together with linkage analysis identifies a Mendelian disease gene. *PLoS Genet.* *6*, e1000991.
- Sölder, B., Streif, W., Ellemunter, H., Mayr, U., and Jaschke, W. (1997). Fibromuscular dysplasia of the internal carotid artery in a child with alpha-1-antitrypsin deficiency. *Dev. Med. Child Neurol.* *39*, 827–829.
- van Son, J.A.M., Edwards, W.D., and Danielson, G.K. (1994). Pathology of coronary arteries, myocardium, and great arteries in supravalvular aortic stenosis. *J. Thorac. Cardiovasc. Surg.* *108*, 21–28.
- Spinner, N.B., Colliton, R.P., Crosnier, C., Krantz, I.D., Hadchouel, M., and Meunier-Rotival, M. (2001). Jagged1 mutations in alagille syndrome. *Hum. Mutat.* *17*, 18–33.
- Stamm, C., Friehs, I., Ho, S.Y., Moran, A.M., Jonas, R.A., and Nido, P.J. del (2001). Congenital supravalvar aortic stenosis: a simple lesion? *Eur. J. Cardiothorac. Surg.* *19*, 195–202.
- Stanley, J.C., Gewertz, B.L., Bove, E.L., Sottiurai, V., and Fry, W.J. (1975). Arterial fibrodysplasia. Histopathologic character and current etiologic concepts. *Arch. Surg. Chic. Ill* *1960* *110*, 561–566.
- Stansfield, B.K., Bessler, W.K., Mali, R., Mund, J.A., Downing, B.D., Kapur, R., and Ingram, D.A. (2014). Ras-Mek-Erk Signaling Regulates Nf1 Heterozygous Neointima Formation. *Am. J. Pathol.* *184*, 79–85.
- Steinberger, D., Reinhartz, T., Unsöld, R., and Müller, U. (1996). FGFR2 mutation in clinically nonclassifiable autosomal dominant craniosynostosis with pronounced phenotypic variation. *Am. J. Med. Genet.* *66*, 81–86.
- Steinbicker, A.U., Liu, H., Jiramongkolchai, K., Malhotra, R., Choe, E.Y., Busch, C.J., Graveline, A.R., Kao, S.M., Nagasaka, Y., Ichinose, F., et al. (2011). Nitric oxide regulates pulmonary vascular smooth muscle cell

expression of the inducible cAMP early repressor gene. *Nitric Oxide Biol. Chem. Off. J. Nitric Oxide Soc.* 25, 294–302.

Stenson, P.D., Mort, M., Ball, E.V., Howells, K., Phillips, A.D., Thomas, N.S., and Cooper, D.N. (2009). The Human Gene Mutation Database: 2008 update. *Genome Med.* 1, 13.

Stonebridge, P.A., Draper, T., Kelman, J., Howlett, J., Allan, P.L., Prescott, R., and Ruckley, C.V. (1996). Growth rate of infrarenal aortic aneurysms. *Eur. J. Vasc. Endovasc. Surg. Off. J. Eur. Soc. Vasc. Surg.* 11, 70–73.

Stoyneva, Z. (2003). Current pathophysiological views on vibration-induced Raynaud's phenomenon. *Cardiovasc. Res.* 57, 615–624.

Strasser, A. (2005). The role of BH3-only proteins in the immune system. *Nat. Rev. Immunol.* 5, 189–200.

Sukhova, G.K., Shi, G.P., Simon, D.I., Chapman, H.A., and Libby, P. (1998). Expression of the elastolytic cathepsins S and K in human atheroma and regulation of their production in smooth muscle cells. *J. Clin. Invest.* 102, 576–583.

Sukhova, G.K., Wang, B., Libby, P., Pan, J.-H., Zhang, Y., Grubb, A., Fang, K., Chapman, H.A., and Shi, G.-P. (2005). Cystatin C deficiency increases elastic lamina degradation and aortic dilatation in apolipoprotein E-null mice. *Circ. Res.* 96, 368–375.

Sumida, H., Noguchi, K., Kihara, Y., Abe, M., Yanagida, K., Hamano, F., Sato, S., Tamaki, K., Morishita, Y., Kano, M.R., et al. (2010). LPA4 regulates blood and lymphatic vessel formation during mouse embryogenesis. *Blood* 116, 5060–5070.

Szabo, Z., Crepeau, M.W., Mitchell, A.L., Stephan, M.J., Puntel, R.A., Yin Loke, K., Kirk, R.C., and Urban, Z. (2006). Aortic aneurysmal disease and cutis laxa caused by defects in the elastin gene. *J. Med. Genet.* 43, 255–258.

Szekanecz, Z., Shah, M.R., Pearce, W.H., and Koch, A.E. (1994). Human atherosclerotic abdominal aortic aneurysms produce interleukin (IL)-6 and interferon-gamma but not IL-2 and IL-4: the possible role for IL-6 and interferon-gamma in vascular inflammation. *Agents Actions* 42, 159–162.

Takei, H., Ishikawa, S., Otaki, A., Sakata, K., Aizaki, M., Sato, Y., Suzuki, M., Ishikita, T., Iino, Y., and Yokoe, T. (1995). Screening for abdominal aortic aneurysm and occlusive peripheral vascular disease in Japanese residents. *Surg. Today* 25, 608–611.

Takeuchi, T., Futatsuka, M., Imanishi, H., and Yamada, S. (1986). Pathological changes observed in the finger biopsy of patients with vibration-induced white finger. *Scand. J. Work. Environ. Health* 12, 280–283.

Tang, Z.-Z., and Lin, D.-Y. (2014). Meta-analysis of sequencing studies with heterogeneous genetic associations. *Genet. Epidemiol.* 38, 389–401.

Taylor, H.E., Khatua, A.K., and Popik, W. (2014). The Innate Immune Factor Apolipoprotein L1 Restricts HIV-1 Infection. *J. Virol.* 88, 592–603.

Tennessen, J.A., Bigham, A.W., O'Connor, T.D., Fu, W., Kenny, E.E., Gravel, S., McGee, S., Do, R., Liu, X., Jun, G., et al. (2012). Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* 337, 64–69.

TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute, Crosby, J., Peloso, G.M., Auer, P.L., Crosslin, D.R., Stitzel, N.O., Lange, L.A., Lu, Y., Tang, Z., Zhang, H., et al.

- (2014). Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *N. Engl. J. Med.* *371*, 22–31.
- Thompson, R.W., and Parks, W.C. (1996). Role of matrix metalloproteinases in abdominal aortic aneurysms. *Ann. N. Y. Acad. Sci.* *800*, 157–174.
- Thompson, A.R., Golledge, J., Cooper, J.A., Hafez, H., Norman, P.E., and Humphries, S.E. (2009). Sequence variant on 9p21 is associated with the presence of abdominal aortic aneurysm disease but does not have an impact on aneurysmal expansion. *Eur. J. Hum. Genet. EJHG* *17*, 391–394.
- Thompson, R.W., Holmes, D.R., Mertens, R.A., Liao, S., Botney, M.D., Mecham, R.P., Welgus, H.G., and Parks, W.C. (1995). Production and localization of 92-kilodalton gelatinase in abdominal aortic aneurysms. An elastolytic metalloproteinase expressed by aneurysm-infiltrating macrophages. *J. Clin. Invest.* *96*, 318–326.
- Tilli, T.M., Ferreira, L.B., and Gimba, E.R.P. (2015). Osteopontin-c mediates the upregulation of androgen responsive genes in LNCaP cells through PI3K/Akt and androgen receptor signaling. *Oncol. Lett.* *9*, 1845–1850.
- Tofanelli, S., Bertocini, S., Castri, L., Luiselli, D., Calafell, F., Donati, G., and Paoli, G. (2009). On the Origins and Admixture of Malagasy: New Evidence from High-Resolution Analyses of Paternal and Maternal Lineages. *Mol. Biol. Evol.* *26*, 2109–2124.
- Tokunou, T., Shibata, R., Kai, H., Ichiki, T., Morisaki, T., Fukuyama, K., Ono, H., Iino, N., Masuda, S., Shimokawa, H., et al. (2003). Apoptosis induced by inhibition of cyclic AMP response element-binding protein in vascular smooth muscle cells. *Circulation* *108*, 1246–1252.
- Toma, C., Torricco, B., Hervás, A., Valdés-Mas, R., Tristán-Noguero, A., Padillo, V., Maristany, M., Salgado, M., Arenas, C., Puente, X.S., et al. (2014). Exome sequencing in multiplex autism families suggests a major role for heterozygous truncating mutations. *Mol. Psychiatry* *19*, 784–790.
- Touzé, E., Oppenheim, C., Trystram, D., Nokam, G., Pasquini, M., Alamowitch, S., Hervé, D., Garnier, P., Mousseaux, E., and Plouin, P.-F. (2010). Fibromuscular dysplasia of cervical and intracranial arteries. *Int. J. Stroke Off. J. Int. Stroke Soc.* *5*, 296–305.
- Tranchevent, L.-C., Barriot, R., Yu, S., Van Vooren, S., Van Loo, P., Coessens, B., De Moor, B., Aerts, S., and Moreau, Y. (2008). ENDEAVOUR update: a web resource for gene prioritization in multiple species. *Nucleic Acids Res.* *36*, W377–W384.
- Trerotola, M., Relli, V., Simeone, P., and Alberti, S. (2015). Epigenetic inheritance and the missing heritability. *Hum. Genomics* *9*, 17.
- Tsomi, K., Karagiorga-Lagana, M., Karabatsos, F., Fragodimitri, C., van Vliet-Konstantinidou, C., Premetis, E., and Stamoulakatou, A. (2001). Arterial elastorrhexis in beta-thalassaemia intermedia, sickle cell thalassaemia and hereditary spherocytosis. *Eur. J. Haematol.* *67*, 135–141.
- Tzemos, N., Therrien, J., Yip, J., Thanassoulis, G., Tremblay, S., Jamorski, M.T., Webb, G.D., and Siu, S.C. (2008). Outcomes in adults with bicuspid aortic valves. *JAMA* *300*, 1317–1325.
- Uddman, R., Tajti, J., Möller, S., Sundler, F., and Edvinsson, L. (1999). Neuronal messengers and peptide receptors in the human sphenopalatine and otic ganglia. *Brain Res.* *826*, 193–199.
- U.S. Preventive Services Task Force (2005). Screening for abdominal aortic aneurysm: recommendation statement. *Ann. Intern. Med.* *142*, 198–202.

- Vanhollebeke, B., and Pays, E. (2006). The function of apolipoproteins L. *Cell. Mol. Life Sci. CMLS* 63, 1937–1944.
- Vanhollebeke, B., and Pays, E. (2010). The trypanolytic factor of human serum: many ways to enter the parasite, a single way to kill. *Mol. Microbiol.* 76, 806–814.
- Van Laer, L., Dietz, H., and Loeys, B. (2014). Loeys-Dietz syndrome. *Adv. Exp. Med. Biol.* 802, 95–105.
- Vanwalleghem, G., Fontaine, F., Lecordier, L., Tebabi, P., Klewe, K., Nolan, D.P., Yamaryo-Botté, Y., Botté, C., Kremer, A., Burkard, G.S., et al. (2015). Coupling of lysosomal and mitochondrial membrane permeabilization in trypanolysis by APOL1. *Nat. Commun.* 6, 8078.
- Vendrov, A.E., Madamanchi, N.R., Hakim, Z.S., Rojas, M., and Runge, M.S. (2006). Thrombin and NAD(P)H oxidase-mediated regulation of CD44 and BMP4-Id pathway in VSMC, restenosis, and atherosclerosis. *Circ. Res.* 98, 1254–1263.
- Verloes, A., Sakalihasan, N., Koulischer, L., and Limet, R. (1995). Aneurysms of the abdominal aorta: familial and genetic aspects in three hundred thirteen pedigrees. *J. Vasc. Surg.* 21, 646–655.
- Visel, A., Zhu, Y., May, D., Afzal, V., Gong, E., Attanasio, C., Blow, M.J., Cohen, J.C., Rubin, E.M., and Pennacchio, L.A. (2010). Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* 464, 409–412.
- Visse, R., and Nagase, H. (2003). Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ. Res.* 92, 827–839.
- Vissers, L.E.L.M., de Ligt, J., Gilissen, C., Janssen, I., Stehouwer, M., de Vries, P., van Lier, B., Arts, P., Wieskamp, N., del Rosario, M., et al. (2010). A de novo paradigm for mental retardation. *Nat. Genet.* 42, 1109–1112.
- Vives Corrons, J.L., Besson, I., Aymerich, M., Ayala, S., Alloisio, N., Delaunay, J., Gonzalez, I., and Manrubia, E. (1995). Hereditary xerocytosis: a report of six unrelated Spanish families with leaky red cell syndrome and increased heat stability of the erythrocyte membrane. *Br. J. Haematol.* 90, 817–822.
- van Vlijmen-van Keulen, C.J., Pals, G., and Rauwerda, J.A. (2002). Familial abdominal aortic aneurysm: a systematic review of a genetic background. *Eur. J. Vasc. Endovasc. Surg. Off. J. Eur. Soc. Vasc. Surg.* 24, 105–116.
- Vonend, O., Stegbauer, J., Sojka, J., Habbel, S., Quack, I., Robaye, B., Boeynaems, J.-M., and Rump, L.C. (2005). Noradrenaline and extracellular nucleotide cotransmission involves activation of vasoconstrictive P2X1,3- and P2Y6-like receptors in mouse perfused kidney. *Br. J. Pharmacol.* 145, 66–74.
- Wahlgren, C.M., Larsson, E., Magnusson, P.K.E., Hultgren, R., and Swedenborg, J. (2010). Genetic and environmental contributions to abdominal aortic aneurysm development in a twin population. *J. Vasc. Surg.* 51, 3–7; discussion 7.
- Walsh, T., Shahin, H., Elkan-Miller, T., Lee, M.K., Thornton, A.M., Roeb, W., Abu Rayyan, A., Loulus, S., Avraham, K.B., King, M.-C., et al. (2010). Whole exome sequencing and homozygosity mapping identify mutation in the cell polarity protein GPSM2 as the cause of nonsyndromic hearing loss DFNB82. *Am. J. Hum. Genet.* 87, 90–94.

- Waltner-Romen, M., Falkensammer, G., Rabl, W., and Wick, G. (1998). A previously unrecognized site of local accumulation of mononuclear cells. The vascular-associated lymphoid tissue. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* *46*, 1347–1350.
- Walton, L.J., Powell, J.T., and Parums, D.V. (1997). Unrestricted usage of immunoglobulin heavy chain genes in B cells infiltrating the wall of atherosclerotic abdominal aortic aneurysms. *Atherosclerosis* *135*, 65–71.
- Wan, G., Zhaorigetu, S., Liu, Z., Kaini, R., Jiang, Z., and Hu, C. -a. A. (2008). Apolipoprotein L1, a Novel Bcl-2 Homology Domain 3-only Lipid-binding Protein, Induces Autophagic Cell Death. *J. Biol. Chem.* *283*, 21540–21549.
- Wandstrat, A., and Wakeland, E. (2001). The genetics of complex autoimmune diseases: non-MHC susceptibility genes. *Nat. Immunol.* *2*, 802–809.
- Wang, L., Guo, D., Cao, J., Gong, L., Kamm, K.E., Regalado, E., Li, L., Shete, S., He, W.-Q., Zhu, M.-S., et al. (2010). Mutations in Myosin Light Chain Kinase Cause Familial Aortic Dissections. *Am. J. Hum. Genet.* *87*, 701–707.
- Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T., et al. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* *38*, W214–W220.
- Weiner, R.B., Kiernan, T.J., Yan, B.P., Rosenfield, K., and Jaff, M.R. (2009). Images in cardiovascular medicine. Adventitial fibromuscular dysplasia of the renal artery: management challenges of nonatherosclerotic renal artery stenosis. *Circulation* *120*, e157–e158.
- Weng, Y.-C., Sonni, A., Labelle-Dumais, C., de Leau, M., Kauffman, W.B., Jeanne, M., Biffi, A., Greenberg, S.M., Rosand, J., and Gould, D.B. (2012). COL4A1 mutations in patients with sporadic late-onset intracerebral hemorrhage. *Ann. Neurol.* *71*, 470–477.
- Wentzel, J.J., Gijzen, F.J.H., Stergiopoulos, N., Serruys, P.W., Slager, C.J., and Krams, R. (2003). Shear stress, vascular remodeling and neointimal formation. *J. Biomech.* *36*, 681–688.
- White, C.R., Haidekker, M.A., Stevens, H.Y., and Frangos, J.A. (2004). Extracellular signal-regulated kinase activation and endothelin-1 production in human endothelial cells exposed to vibration. *J. Physiol.* *555*, 565–572.
- White, R., Woodward, S., Leppert, M., O’Connell, P., Hoff, M., Herbst, J., Lalouel, J.M., Dean, M., and Vande Woude, G. (1985). A closely linked genetic marker for cystic fibrosis. *Nature* *318*, 382–384.
- Whiteman, P., and Handford, P.A. (2003). Defective secretion of recombinant fragments of fibrillin-1: implications of protein misfolding for the pathogenesis of Marfan syndrome and related disorders. *Hum. Mol. Genet.* *12*, 727–737.
- Wick, G., Romen, M., Amberger, A., Metzler, B., Mayr, M., Falkensammer, G., and Xu, Q. (1997). Atherosclerosis, autoimmunity, and vascular-associated lymphoid tissue. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *11*, 1199–1207.
- Wienecke, T., Olesen, J., and Ashina, M. (2010). Prostaglandin I2 (epoprostenol) triggers migraine-like attacks in migraineurs. *Cephalalgia Int. J. Headache* *30*, 179–190.

- Wilkins, M.R., Aldashev, A.A., Wharton, J., Rhodes, C.J., Vandrovцова, J., Kasperaviciute, D., Bhosle, S.G., Mueller, M., Geschka, S., Rison, S., et al. (2014). α 1-A680T variant in GUCY1A3 as a candidate conferring protection from pulmonary hypertension among Kyrgyz highlanders. *Circ. Cardiovasc. Genet.* 7, 920–929.
- Wilmink, A.B., and Quick, C.R. (1998). Epidemiology and potential for prevention of abdominal aortic aneurysm. *Br. J. Surg.* 85, 155–162.
- Wolinsky, H., and Glagov, S. (1964). STRUCTURAL BASIS FOR THE STATIC MECHANICAL PROPERTIES OF THE AORTIC MEDIA. *Circ. Res.* 14, 400–413.
- Wu, M.C., Lee, S., Cai, T., Li, Y., Boehnke, M., and Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *Am. J. Hum. Genet.* 89, 82–93.
- Xia, S., Ozsvath, K., Hirose, H., and Tilson, M.D. (1996). Partial amino acid sequence of a novel 40-kDa human aortic protein, with vitronectin-like, fibrinogen-like, and calcium binding domains: aortic aneurysm-associated protein-40 (AAAP-40) [human MAGP-3, proposed]. *Biochem. Biophys. Res. Commun.* 219, 36–39.
- Xiong, W., MacTaggart, J., Knispel, R., Worth, J., Persidsky, Y., and Baxter, B.T. (2009). Blocking TNF-Attenuates Aneurysm Formation in a Murine Model. *J. Immunol.* 183, 2741–2746.
- Xu, B., Roos, J.L., Dexheimer, P., Boone, B., Plummer, B., Levy, S., Gogos, J.A., and Karayiorgou, M. (2011). Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nat. Genet.* 43, 864–868.
- Yamamoto, K., Sokabe, T., Matsumoto, T., Yoshimura, K., Shibata, M., Ohura, N., Fukuda, T., Sato, T., Sekine, K., Kato, S., et al. (2006). Impaired flow-dependent control of vascular tone and remodeling in P2X4-deficient mice. *Nat. Med.* 12, 133–137.
- Yamashita, M., Oka, K., and Tanaka, K. (1983). Histopathology of the brain vascular network in moyamoya disease. *Stroke J. Cereb. Circ.* 14, 50–58.
- Yanagida, K., Masago, K., Nakanishi, H., Kihara, Y., Hamano, F., Tajima, Y., Taguchi, R., Shimizu, T., and Ishii, S. (2009). Identification and Characterization of a Novel Lysophosphatidic Acid Receptor, p2y5/LPA6. *J. Biol. Chem.* 284, 17731–17741.
- Yeh, H.-L., Kuo, L.-T., Sung, F.-C., Chiang, C.-W., and Yeh, C.-C. (2013). GSTM1, GSTT1, GSTP1, and GSTA1 genetic variants are not associated with coronary artery disease in Taiwan. *Gene* 523, 64–69.
- Yen, H.C., Lee, F.Y., and Chau, L.Y. (1997). Analysis of the T cell receptor V beta repertoire in human aortic aneurysms. *Atherosclerosis* 135, 29–36.
- Yii, M.K. (2003). Epidemiology of abdominal aortic aneurysm in an Asian population. *ANZ J. Surg.* 73, 393–395.
- Yusuf, S., Hawken, S., Ounpuu, S., Dans, T., Avezum, A., Lanas, F., McQueen, M., Budaj, A., Pais, P., Varigos, J., et al. (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet Lond. Engl.* 364, 937–952.
- Zarychanski, R., Schulz, V.P., Houston, B.L., Maksimova, Y., Houston, D.S., Smith, B., Rinehart, J., and Gallagher, P.G. (2012). Mutations in the mechanotransduction protein PIEZO1 are associated with hereditary xerocytosis. *Blood* 120, 1908–1915.

Zhaorigetu, S., Yang, Z., Toma, I., McCaffrey, T.A., and Hu, C.-A.A. (2011). Apolipoprotein L6, induced in atherosclerotic lesions, promotes apoptosis and blocks Beclin 1-dependent autophagy in atherosclerotic cells. *J. Biol. Chem.* *286*, 27389–27398.

Zheng, Y., Guan, H., Li, Y., Liu, C., Liu, B., Sheng, Q., and Miao, S. (2002). [Microarray in screening for differentially expressed genes of cellular cycle and apoptosis in abdominal aortic aneurysms]. *Zhonghua Wai Ke Za Zhi* *40*, 817–819.

Zhou, C., Li, C., Li, D., Wang, Y., Shao, W., You, Y., Peng, J., Zhang, X., Lu, L., and Shen, X. (2013). BIG1, a brefeldin A-inhibited guanine nucleotide-exchange protein regulates neurite development via PI3K-AKT and ERK signaling pathways. *Neuroscience* *254*, 361–368.

Zhou, D., Hu, W., Wang, Q., and Jin, Y. (2014). Glutathione S-transferase M1 Polymorphism and Coronary Heart Disease Susceptibility: A Meta-analysis Involving 47,596 Subjects. *Heart Lung Circ.* *23*, 578–585.

Zhu, L., Vranckx, R., Khau Van Kien, P., Lalande, A., Boisset, N., Mathieu, F., Wegman, M., Glancy, L., Gasc, J.-M., Brunotte, F., et al. (2006). Mutations in myosin heavy chain 11 cause a syndrome associating thoracic aortic aneurysm/aortic dissection and patent ductus arteriosus. *Nat. Genet.* *38*, 343–349.

ARTICLE

The type of variants at the *COL3A1* gene associates with the phenotype and severity of vascular Ehlers–Danlos syndrome

Michael Frank^{1,2,8}, Juliette Albuissou^{1,2,3,8}, Brigitte Ranque^{2,3,4,8}, Lisa Golmard^{1,3}, Jean-Michael Mazzella¹, Laurence Bal-Theoleyre^{5,8}, Anne-Laure Fauret^{1,3,8}, Tristan Mirault^{1,3}, Nicolas Denarié¹, Elie Mousseaux^{3,6}, Pierre Boutouyrie^{2,3,7}, Jean-Noël Fiessinger^{1,3}, Joseph Emmerich^{1,3}, Emmanuel Messas^{1,2,3} and Xavier Jeunemaitre^{*,1,2,3}

Vascular Ehlers–Danlos syndrome (vEDS) is a rare and severe autosomal dominant disorder caused by variants at the *COL3A1* gene. Clinical characteristics and course of disease of 215 molecularly proven patients (146 index cases and 69 relatives) were analysed. We found 126 distinct variants that were divided into five groups: (1) Glycine substitutions ($n = 71$), (2) splice-site and in-frame insertions–deletions ($n = 36$), (3) variants leading to haplo-insufficiency ($n = 7$), (4) non-glycine missense variants within the triple helix ($n = 4$ variants), and (5) non-glycine missense variants or in-frame insertions–deletions, in the N- or C-terminal part of the protein ($n = 8$). Overall, our cohort confirmed the severity of the disease with a median age at first complication of 29 years (IQR 22–39), the most frequent being arterial (48%) and digestive (24%) ruptures. Groups 2 and 1 were significantly more severe than groups 3–5, with extreme median ages at first major complication of 23–47 years. Patients of groups 3–5 had a less typical phenotype and remarkably absence of digestive events. The distribution of glycine-replacing amino acids was strongly biased towards more destabilizing residues of the collagen assembly. Thus the natural course of vEDS and the clinical phenotype of patients are influenced by the type of *COL3A1* variant. This study also confirms that patients with variants located in the C- and N-termini or leading to haplo-insufficiency have milder course of the disease and less prevalent diagnostic criteria. These findings may help refine diagnostic strategy, genetic counselling and clinical care.

European Journal of Human Genetics advance online publication, 11 March 2015; doi:10.1038/ejhg.2015.32

INTRODUCTION

Vascular Ehlers–Danlos syndrome (vEDS, OMIM #130050) is a rare inherited autosomal dominant disorder with an estimated prevalence of 1/150 000. Patients are exposed to life-threatening complications, typically spontaneous arterial dissections and ruptures, bowel perforations and uterine ruptures.^{1,2} Clinical events usually occur from early adulthood onwards with unpredictable relapse intervals. Each complication is at high risk of death, and consequently, life expectancy is reduced, estimated in the early 2000s to a median of 48 years.²

Major and minor clinical diagnostic criteria of the Villefranche classification illustrate the variety of physical signs that may constitute the clinical phenotype and explain the difficulty of its diagnosis.³ Formal diagnosis of vEDS requires a positive genetic test, that is, the demonstration of a pathogenic variant within the *COL3A1* gene encoding the pro- $\alpha 1$ chain of type III procollagen. More than 570 *COL3A1* genetic variants (including the 126 described in the present article) are registered in the Ehlers–Danlos Syndrome Variant Database, a large majority of them being private.⁴ Most of the variants

are heterozygous missense substitutions affecting one of the glycine residues of the [Gly-X-Y]₃₄₃ repeat within the triple helical region of type III collagen, acting as dominant-negative variants.² vEDS might also be caused by heterozygous splice-site variants of exons encoding a triple helix sequence, the vast majority leading to an in-frame exon skipping and generation of a shorted translated product also altering the stable assembly of type III procollagen homotrimers.⁵ The rare frameshift, non-sense or large deletion variants result in a reduction of mature procollagen and haplo-insufficiency.⁶ Ultimately, both the alteration of type III collagen synthesis and assembly result in decreased resistance to mechanical stress of hollow organs, especially arteries and bowel.^{7,8}

Despite clinical and genetic heterogeneity, only variants leading to haplo-insufficiency have been discussed to result in a milder phenotype.^{5,6,9} In order to investigate the possibility of other associations between the nature and location of *COL3A1* variants and the phenotypic expression of the disease, the clinical characteristics and major complications in 215 patients with molecularly proven vEDS were analyzed.

¹AP-HP, Hôpital Européen Georges Pompidou, Département de Génétique, Service de Médecine Vasculaire et Centre de Référence des Maladies Vasculaires Rares, Paris, France; ²INSERM, U970, Paris centre de Recherche Cardiovasculaire—PARCC, Paris, France; ³Université Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine, Paris, France; ⁴AP-HP, Hôpital Européen Georges Pompidou, Service de Médecine Interne, Paris, France; ⁵AP-HM, Service de Médecine Vasculaire, Hôpital de la Timone, Av Jean Moulin, Marseille, France; ⁶AP-HP, Hôpital Européen Georges Pompidou, Service de Radiologie Cardiovasculaire, Paris, France; ⁷AP-HP, Hôpital Européen Georges Pompidou, Service de Pharmacologie, Paris, France

⁸These authors contributed equally to this work.

*Correspondence: Professor X Jeunemaitre, Département de Génétique—Centre de Référence des Maladies Vasculaires Rares, Hôpital Européen Georges Pompidou, APHP, 20-40, rue Leblanc, Paris 75015, France. Tel: +33 1 56 09 38 80; Fax: +33 1 56 09 38 84; E-mail: xavier.jeunemaitre@inserm.fr

Received 28 July 2014; revised 24 December 2014; accepted 28 January 2015

PATIENTS AND METHODS

Study population

Patients with vEDS referred between 2001 and 2014 to the French National Reference Centre for Rare Vascular Diseases and to the corresponding network of referral centres for either clinical care or genetic diagnosis were included in a retrospective cohort study. We have previously described the oral phenotype of 17 of them.¹⁰ All patients had a standardized assessment of personal medical history and a physical examination by a senior physician. During genetic counselling, first-degree relatives were proposed to be examined and genetically screened.

The patients' clinical characteristics were staged as major or minor diagnostic criteria according to the Villefranche nosology.³ Vascular, digestive and/or obstetrical events were assessed until last known follow-up visit in our department or in referring centres. Arterial events were defined by symptomatic dissections, aneurysms or spontaneous arterial ruptures. Digestive events were defined by spontaneous bowel perforations or ruptures requiring surgery. All postoperative or likely surgery-related digestive events were excluded from analysis. Obstetrical events were defined by uterine ruptures, arterial ruptures or dissections occurring during pregnancy, delivery or in the 1-month postpartum period.

Genetic testing was performed in compliance with French legislation on genetics diagnostics tests (French bioethics law no. 2004–800), and written informed consent was obtained for all patients. Decision of genetic testing was made after a multidisciplinary outpatient visit, associating a geneticist, a specialist in vascular medicine and a genetic counsellor. Psychological support was organized for each patient before and throughout the genetic testing process when necessary.

Genetic analysis

COL3A1 gene was screened for all patients either by direct sequencing from genomic DNA (gDNA) and Multiplex Ligation Probe Amplification or by sequencing from complementary DNA (cDNA) (for details of the genetic analysis methods, see Supplementary Method S1, Supplementary Tables S1 and S2). When a variant was detected, independent direct sequencing or arrayCGH was carried out as a control, when possible. Variants at the nucleotide and protein

levels are described according to the nomenclature recommended by the Human Genome Variation Society (www.hgvs.org/mutnomen), based on the *COL3A1* human cDNA and gene sequences (GenBank NM_000090.3 and NG_007404.1), which number the *COL3A1* gene exons from 1 to 51. Variants were registered in the Ehlers–Danlos Syndrome Variant Database (https://eds.gene.le.ac.uk/variants.php?action=search_unique&select_db=COL3A1).

Statistical analysis

Patients' characteristics are reported for the whole cohort and according to the ascertainment status (index case or relative) and the type of variant. Time-independent qualitative variables were compared using chi-square test or Fisher's exact test (when count was insufficient in the contingency table). The comparison of the incidence onset of the major complications was performed using Kaplan–Meier survival curves with Log-rank test. In respect of the comparison between the types of variant, Log-rank tests were stratified for the ascertainment status. Quantitative variables were reported using non-parametric descriptive statistics (median, interquartile range (IQR)) and compared using the Wilcoxon test. A sensitivity analysis was performed by replicating all comparisons in the index patients group only. A *P*-value <0.05 was considered significant. The statistical analysis was performed with the R software (version 3.1.1).¹¹

RESULTS

Clinical diagnostic criteria of index cases and first-degree relatives

A total of 215 patients (146 index cases, 69 first-degree relatives) were studied (Table 1). Almost all patients were of European origin and subjects were predominantly women (62%). The age of occurrence of the first major event did not differ significantly between men and women (Supplementary Figure S1). The median age at diagnosis was 33 (IQR 24–43) years for index cases and 38 (IQR 19–49) for relatives. As a consequence of recruitment bias, almost all index cases (90%) had experienced at least one complication at diagnosis compared with only 58% of relatives (*P*<0.001). They had also more frequent major diagnostic criteria (thin and translucent skin, extensive bruising and

Table 1 Main features and complications of genotyped individuals with vEDS (Villefranche major diagnostic criteria)

	All patients N = 215	Index cases N = 146	Relatives N = 69	Status	
				<i>P</i> -value	Test
Male gender	83 (38%)	58 (39%)	25 (36%)	NS	Chi-square
Median age at genetic diagnosis	34 (23–45)	33 (24–43)	38 (19–49)	NS	Wilcoxon
Median age at last follow-up visit	37 (24–46)	35 (26–45)	39 (21–50)	NS	Wilcoxon
Characteristic facial appearance	149 (77%)	111 (82%)	38 (66%)	*	Chi-square
Thin translucent skin	172 (90%)	121 (90%)	51 (87%)	NS	Chi-square
Extensive bruising	149 (76%)	109 (78%)	40 (68%)	NS	Chi-square
≥ 1 Major complication	168 (80%)	131 (90%)	37 (58%)	***	Log-rank
Median age of the first one	29 (22–39)	28 (22–36)	37 (26–48)	*	Wilcoxon
≥ 1 Vascular complication	150 (72%)	115 (80%)	35 (55%)	***	Log-rank
Median age of the first one	34 (25–42)	32 (25–40)	39 (30–50)	*	Wilcoxon
≥ 1 Digestive complication	63 (30%)	52 (36%)	11 (16%)	**	Log-rank
Median age of the first one	23 (19–34)	23 (19–32)	23 (17–37)	NS	Wilcoxon
≥ 1 Obstetrical complication ^a	27 (22%)	20 (25%)	7 (17%)	NS	Log-rank
Median age of the first one	29 (26–33)	29 (25–33)	26 (26–32)	NS	Wilcoxon
Median number of major criteria	4 (3–5)	4 (3–5)	3 (3–4)	***	Wilcoxon

Quantitative variables are given in medians (interquartile range); percentages were calculated on the basis of available data for each item; *P*-values stratified for index status: **P*<0.05; ***P*<0.01;

****P*<0.001.

^aFemales only.

characteristic facial appearance) (Table 1). The median prevalence of each minor diagnostic criteria was similar in both index cases and relatives, except for acrogeria and the occurrence of carotid cavernous fistula (Supplementary Table S3).

Survival analysis showed significant differences in the age at which the first major event occurred in index cases and relatives (Figure 1). At 28 years, half of index cases had experienced a first major complication, *versus* 25% of relatives only. At 43 years, half of relatives had experienced at least one major complication *versus* 85% of index cases.

First and second major events during the follow-up

In symptomatic patients, first major events were predominantly arterial (63%), followed by digestive complications (26%) (Table 2a). In women, the disease was revealed by a first obstetrical event in 12% of the cases. Arterial complications were predominantly local dissections and aneurysms at the iliac, renal and carotid arteries (80% with multiple locations), whereas digestive complications were spontaneous colonic perforation in 85% of the cases. We also observed that the second major event was influenced by the type of first event (Table 2b). It was more frequently digestive for patients with a first gastrointestinal event than for patients with a first arterial event (39 *versus* 7%, respectively; $P < 0.0001$).

COL3A1 variants distribution and classification

In the $n = 146$ index cases, 126 distinct variants were identified by Sanger sequencing on gDNA (Figure 2). Fifty of them were confirmed by cDNA sequencing (Supplementary Table S4). Variants were spread along the entire *COL3A1* gene, 90% ($n = 114$ variants) of them were located between exons 6 and 47 encoding the triple helix domain of type III procollagen. More than half of the 126 variants ($n = 68$) were not previously reported, including the complete gene deletion. The majority of them were glycine substitutions (Group 1; 71 distinct variants). The second most frequent variants were splice-site variants, which were predominantly (83%) located at the 5' donor site, or insertions/deletions resulting in a large in-frame anomaly located in the triple helix (Group 2; 36 distinct variants). We classified the other variants in three other groups: those resulting in haplo-insufficiency (group 3; 7 distinct variants), non-glycine missense variants located in the triple helix (group 4; $n = 4$), and non-glycine missense variants located in the C- or N-terminal regions of the protein (group 5; $n = 8$).

Most of the point substitutions were private and only 12 of them were common to several unrelated families (Figure 2). Among them,

the splice-site variant at intron 23, c.1662+1G>A, showed the highest level of recurrence in seven independent families (Supplementary Table S4a). Two splice-site variants, located at intron 48 and 49 5' donor sites, c.3283+1G>C (patient AN_002976) and c.4011+1G>A (patient AN_002972) led to an out of frame exon skipping, the occurrence of a premature termination codon and finally resulted in haplo-insufficiency. Details of the variants and their locations are presented in Figure 2, except for the complete deletion of the *COL3A1* gene found in one index case (see Supplementary Table S4 for detailed genotypic and phenotypic analysis of patients of each group).

Genotype-phenotype associations

According to the Villefranche nosology, the number of major criteria was consistent among groups 1–4 but significantly lower in group 5 (Table 3). However, the number of minor criteria differed according to the groups of patients (Supplementary Table S5): compared with group 1 (glycine substitutions), group 2 (splice-site variants) had more minor criteria whereas groups 3–5 had fewer ($P < 0.01$ for each comparison). However, when evaluated independently, the median prevalence of each minor diagnostic criteria was similar in all groups except for acrogeria, which was more prevalent in groups 1 and 2, when compared with groups 3–5 ($P < 0.01$).

Similarly, the age at diagnosis was lower in the groups 1–2 as compared with the three other groups (Table 3, global P -value < 0.001): group 2 had a low median age at diagnosis (25 years (18–33)), followed by group 1 with a delay of 10 years (34 years (27–42)). The groups 3–5 had a similar age at diagnosis, delayed by 12 years from group 1 (46 (38–52) years, 45 (44–48), and 45 (38–49), respectively). The same distribution was observed for the age at first major event. At the age of 40 years, proportion of event-free survival was 10% for patients of group 2, 24% in group 1; 70% in group 3; 80% in group 4 and 55% in group 5 ($P < 0.001$, Figure 3). Interestingly, none of the patients of groups 3–5 had digestive complications, *versus* 36 and 34% in groups 1 and 2, respectively (global $P < 0.001$, Table 3).

We also analysed the nature of the amino acids substituting the glycine residues, as a bias for several types of residues was previously

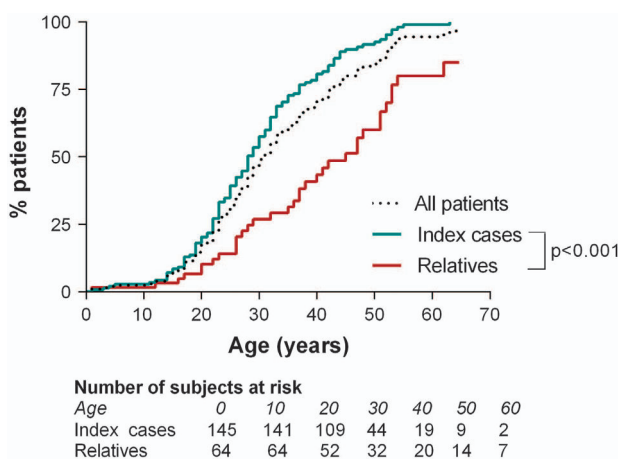


Figure 1 Age of onset of the first major complication in 215 vEDS patients.

Table 2a Type of the first complication according to index/relative status

	All patients N = 210	Index N = 146	Relatives N = 64
First complication			
Vascular	106 (50%)	82 (56%)	24 (37%)
Digestive	44 (21%)	37 (25%)	7 (11%)
Obstetrical	18 (9%)	12 (8%)	6 (10%)
None	42(20%)	15 (10%)	27 (42%)

Table 2b Type of second complication according to the type of first complication

Second complication	Vascular N = 106	Digestive N = 44	Obstetrical N = 18	Total N = 168
Vascular	64 (63%)	21 (49%)	8 (44%)	93 (57%)
Digestive	7 (7%)	17 (39%)	2 (11%)	26 (16%)
Obstetrical	1 (1%)	1 (2%)	5 (28%)	7 (4%)
None	29 (29%)	4 (9%)	3 (17%)	36 (21%)

Global comparison of distribution: chi-square P -value < 0.0001 . Comparison of the distribution of the second vascular and digestive events according to the first vascular and digestive complication: chi-square P -value < 0.0001 .

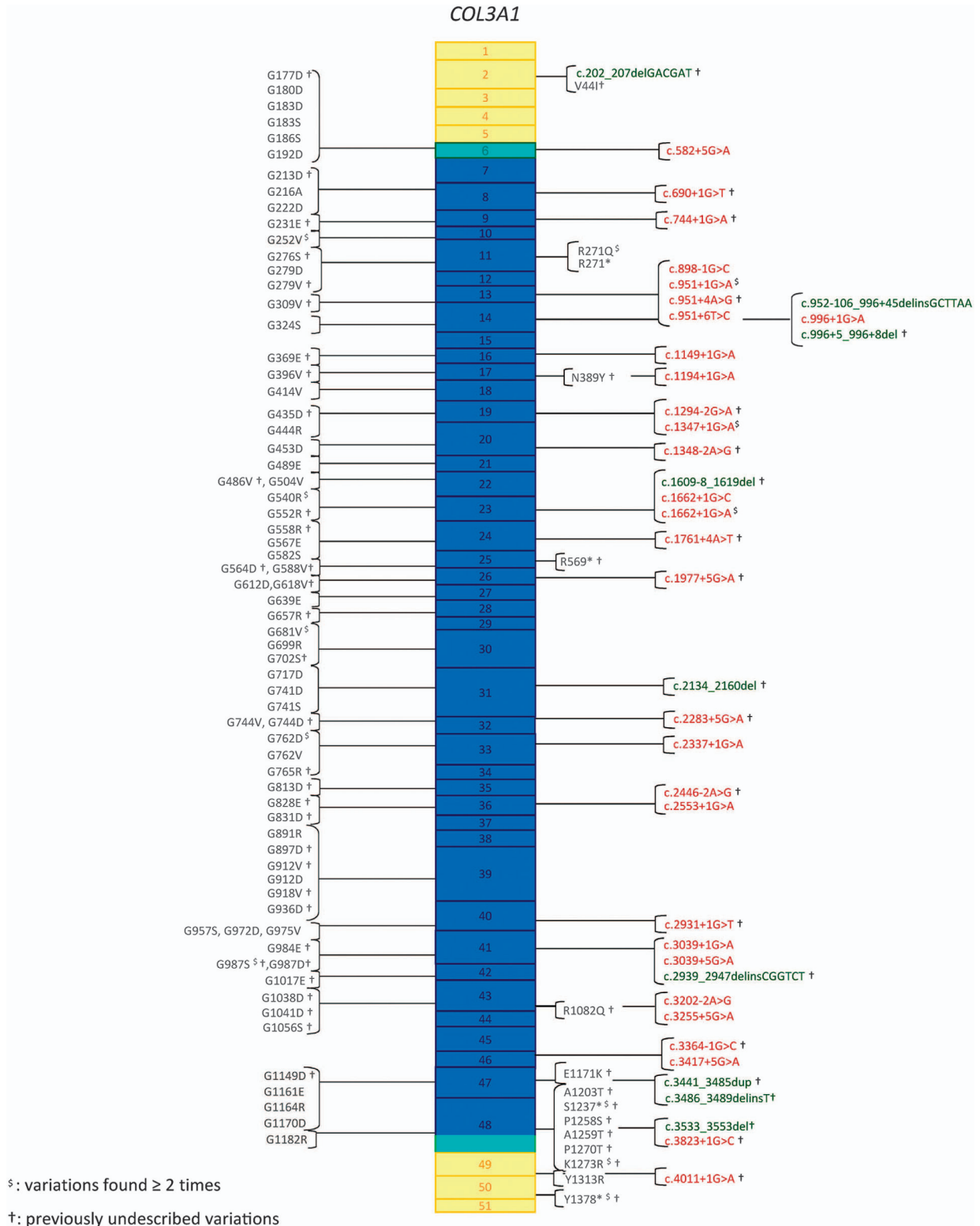


Figure 2 *COL3A1* gene variants detected in 146 unrelated index cases with vEDS. Boxes represent exons of the *COL3A1* gene (NG_007404.1) with colour code representing protein domains (yellow: N- and C-terminal domains, green: transition domains, blue: collagen helical domain). Glycine substitutions are presented on the left in a shortened, nonofficial nomenclature. The complete nomenclature of each variant can be found in the Supplementary Table S4, in which variants detected on the genomic but not on the cDNA sequence are indicated with parentheses, for example, p.(Gly324Ser). All of them have been reported in the Ehlers–Danlos Syndrome Variant Database (see Methods, Genetic analysis). Substitutions for other amino acids, splice-site variants and insertions/deletions are shown on the right. \$ indicates the variants found more than once in the index cases cohort; † indicates previously undescribed variants.

Table 3 Main features and complications of patients according to type of COL3A1 variation (Villefranche major diagnostic criteria)

	Gr.1 Gly missense N = 127	Gr.2 Splice/Ins/Del/Dup N = 55	Gr.3 Haplo-insufficiency N = 14	Gr.4 triple helix variants N = 7	Gr.5 C-/N-terminal variants N = 12	Gr.1 versus respectively: Gr.2, Gr.3, Gr.4, Gr.5 P-value	test
Index cases	79 (62%)	45 (81%)	8 (57%)	5 (71%)	9 (75%)	** , NS, NS, NS	Chi-square
Male gender	49 (39%)	25 (45%)	5 (36%)	2 (29%)	2 (17%)	NS, NS, NS, NS	Chi-square
Caucasian ethnicity	122 (98%)	50 (91%)	13 (93%)	7 (100%)	11 (92%)	NS, NS, NS, NS	Chi-square
Median age at genetic diagnosis ^a	34 (27–42)	25 (18–33)	46 (38–52)	45 (44–48)	45 (38–49)	***, *, NS, NS	Wilcoxon
Median age at last follow-up visit	37 (26–45)	27 (19–39)	47 (42–53)	48 (44–55)	46 (34–52)	** , *, *, NS	Wilcoxon
Characteristic facial appearance	91 (80%)	47 (94%)	7 (64%)	1 (14%)	3 (30%)	*, NS, ** , **	Chi-square
Thin translucent skin	102 (90%)	50 (98%)	11 (92%)	4 (67%)	5 (56%)	NS, NS, NS, *	Chi-square
Extensive bruising	88 (76%)	49 (94%)	7 (64%)	3 (43%)	2 (20%)	** , NS, NS, ***	Chi-square
≥ 1 Major complication	98 (80%)	44 (80%)	12 (86%)	6 (86%)	8 (73%)	***, *, *	Log rank
≥ 1 Vascular complication	89 (74%)	37 (68%)	13 (93%)	5 (71%)	6 (55%)	** , NS, NS, *	Log rank
≥ 1 Digestive complication	44 (36%)	19 (34%)	0 (0%)	0 (0%)	0 (0%)	NS, *, NS, *	Log rank
≥ 1 Obstetrical complication ^b	16 (22%)	6 (24%)	2 (22%)	1 (25%)	2 (25%)	NS, NS, NS, NS	Log rank
Median age of first major complication	29 (23–39)	23 (19–28)	45 (38–51)	47 (44–48)	41 (31–47)	** , ** , ** , *	Wilcoxon
Median number of major criteria	4 (2–5)	4 (3–5)	3 (2–3)	2 (2–3)	2 (1–2)	NS, NS, NS, ***	Wilcoxon

Quantitative variables are given in medians (interquartile range); percentages were calculated on the basis of available data for each item; P-values stratified for index status: * < 0.05; ** < 0.01; *** < 0.001. Between-groups comparison: Gly = missense variants at a Gly residue, Splice.
^aIndex cases only.
^bFemales only.

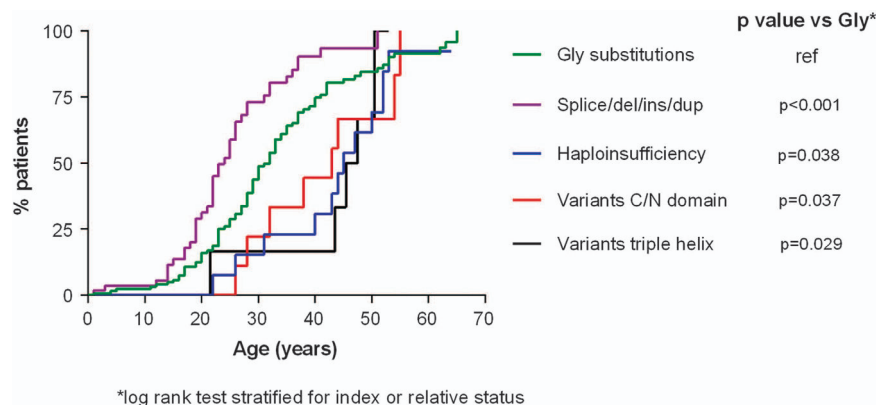
observed for α(I) and α(III) collagen chains in relation with the extent of destabilization of the triple helix.¹² The analysis of *n* = 79 index cases with Gly missense variants showed that the most destabilizing Asp, Val and Glu amino acids were over-represented (*n* = 29, 36%; *n* = 18, 23%; *n* = 9, 11%; respectively), whereas Ser and Ala residues were under-represented (*n* = 11, 15%; *n* = 1, 1%; respectively, global *P* < 0.0001) (Figure 4, Supplementary Table S6). This bias towards the more destabilizing residues was tested as to whether they were associated with a more severe phenotype in all patients. No marked significant difference was observed between the three subgroups (Asp/Val/Glu, Arg, Ser–Ala), except for the characteristic facial appearance (*P* = 0.003) and extensive bruising (*P* = 0.012) that could possibly be due to the relatively low number of subjects per subgroup (Supplementary Table S7).

DISCUSSION

We herein describe the largest series of European patients with molecularly proven vEDS systematically screened in the French referral Centre for this rare inherited disease. Our findings confirm the severity of the disease, positive individuals having experienced a first major vascular, digestive or obstetrical complication at the median age of 29 years. We show that gender does not influence the prognosis. More importantly, we identified subgroups of patients who have a later-onset of the disease and a better prognosis, and especially no digestive events. The five subgroups of variants have also differences in some major and minor criteria, acrogeria, characteristic facial appearance, thin translucent skin and extensive bruising being hallmarks of a severe disease. Finally, we show that for missense variants affecting a glycine residue, the distribution of mutated amino acids is strongly biased towards residues having the stronger destabilizing effect on collagen assembly.

The occurrence of clinical events, their frequency and distribution according to age were comparable to the available cohorts published in 2000 and 2014.^{2,9} Overall, 17% of our vEDS patients had a first complication by the age of 20 years, and 71% had at least one major complication by the age of 40 years. We did not observe a significant gender effect, contrary to the recent survival analysis made by Byers's group.¹³ However, our cohort differs from the US cohort^{2,13} in several aspects because of different recruitment and clinical monitoring modalities: index patients and their relatives described here were approximately 5 years older at the time of diagnosis, and number of symptom-free index cases at referral was only 11% in this cohort versus 39% in the US cohort. This important difference has multiple explanations: patients older in our cohort, possible referral bias for more severe subjects, increasing awareness of the disease, and improved imaging modalities and accessibility of genetic testing. We did not analyse intra-familial phenotypic heterogeneity or penetrance in this study, because we chose to limit the number of patients per family to three in order to avoid familial phenotypic bias. However, prospective systematic evaluation of individuals with the same variant currently performed in our centre may give further insight in the near future.

Interestingly, we found that some groups of patients are more prone to specific complications. First, all patients with digestive ruptures belonged either to group 1 (*n* = 44, 36% of group 1 patients) or to group 2 (*n* = 19, 34%). Patients belonging to the three other groups of variants had no digestive event at the median age of 45 years, whereas all colonic ruptures in groups 1 and 2 were observed in early adulthood (median 23 years, range 19–34). It is noteworthy that we observed no familial aggregation for digestive events. Second, as recently presented by Shalhub *et al*,⁹ aortic complications (aneurysms,



Number of subjects at risk

Age	0	10	20	30	40	50	60
Gly substitutions	123	118	101	56	26	12	7
Splice/del/ins/dup	54	53	31	13	6	3	2
Haploinsufficiency	14	14	14	12	10	5	2
Variants C/N domain	11	11	11	8	6	4	1
Variants triple helix	7	7	7	7	6	3	1

Figure 3 Age of onset of the first major complication as a function of type of *COL3A1* variant in 215 patients with vEDS. Abbreviations: Splice/del/ins/dup, insertion, deletion, duplication or splice-site variant in the triple helix; Gly, Glycine; C/N domain, C-terminal or N-terminal domain of the protein; TH, triple helix.

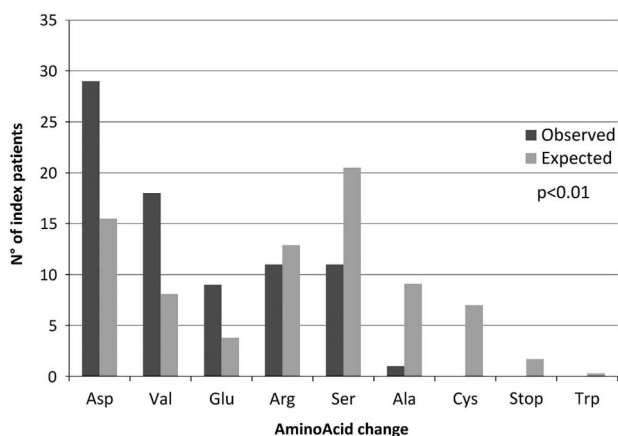


Figure 4 Distribution of observed and expected AA residues substituting a Glycine residue. Expected substitutions are calculated from the corrected expected numbers as previously reported.¹² Yates' chi-squared, $P<0.001$.

dissections and ruptures) appear to be more prevalent in patients with haplo-insufficiency (4/8, 50% of index cases), when compared with glycine substitution (19/79, 19%) or splice-site variant (6/45, 13%) patients (global $P=0.02$).

Genetic screening of the *COL3A1* gene showed a majority of missense variants involving a glycine residue (54% of index cases), followed by splice-site variants (31%), in agreement with previous reports and with the Ehlers–Danlos Syndrome Variant Database (https://eds.gene.le.ac.uk/home.php?select_db=COL3A1). The affected glycine residues were spread from exons 6 to 47, the region encoding chains of the triple helix domain of type III procollagen. As expected, almost all of these variants were private, only 5 out of the 71 glycine substitutions being recurrent. Whether the location of the mutated glycine residue could be associated to the phenotype has been the focus of previous studies. By the analysis of a limited set of 14 different

glycine substitutions, Pope *et al*¹ found an increasing abnormal phenotype when the variant shifted from exon 36 to exon 49. We could not confirm this tendency in this larger cohort of affected individuals (data not shown), consistently with the findings observed in the even greater US cohort.^{2,13}

We also found a strong distribution bias of the glycine replacements, consistently with previous reports suggesting that for the *COL3A1* gene the most destabilizing variants are over-represented in comparison to less destabilizing residues.¹² Indeed, in our series 73% (58/79) of the mutated residues in index cases were either Val, Glu or Asp, much more frequent than Ala and Ser (18/79, 23%), proportions that diverge significantly from that expected of the possible outcomes of a single base pair substitution in a glycine codon ($P<0.001$ for the two subgroups). This selection bias that was also recently observed by Pepin *et al*¹³ cannot be explained by the transition/transversion rate bias in the human coding sequences and is positively correlated to the triple helix destabilizing effect of the corresponding variants.¹⁴ The collagen triple helical structure is indeed disrupted more profoundly when Gly is replaced by a large charged amino acid such as Asp, than by a smaller residue such as Ala. A recent *in vitro* study tested the effect of two types of missense variants (Gly-to-Val and Gly-to-Ala) on the type III collagen folding using a bacterial system for production of homotrimeric model polypeptides.¹⁵ The effect of the Gly-to-Val variant was stronger compared with Gly-to-Ala, suggesting that some Ala variants might not be severe enough to cause vEDS.

The distribution of splice-site variants was characterized by an over-representation of variants at the 5' donor site (30 out of the 36 distinct variants) as already observed.^{13,16} Eighteen out of the 36 distinct variants were already shown to lead to exon skipping or to cryptic splice-site activation and to decreased secretion of procollagen trimers.^{16,17} Even though there was no variant hotspot in our cohort, a few recurrent substitutions were present, especially the

splice-site variant at intron 23, c.1662+1G>A, which is also the most frequent splice-site variant recorded in the Ehlers–Danlos Syndrome Variant Database. In our series, patients with splice-site variants had an earlier age at diagnosis than patients with glycine substitutions (25 *versus* 34 years, $P=0.0002$) and a comparable high prevalence of clinical events, of major and of minor diagnostic criteria. Almost all of these splice variants are in phase and lead to a product deleted of 54–108 residues that can act in a dominant-negative manner, strongly destabilizing the collagen assembly. To our best knowledge, whether this deleterious effect is more important than the one caused by glycine substitutions has not been investigated in biochemical studies.

One of the main findings of our study was the identification of yet unreported missense variants (groups 4 and 5) that were associated with mild clinical features of vEDS and arterial fragility. In addition to being evidenced in patients with suggestive features of the pathology, we also found arguments for their possible pathogenicity, which are given in the Supplementary Materials (Supplementary Results and Methods, Supplementary Tables S4c–e). Importantly, variants at the N- and C-terminal parts of the protein and the missense variants in the triple helix allowed identifying a subgroup of patients representing 10% of our cohort. These patients displayed a milder clinical phenotype with sometimes only one major and/or one minor clinical criterion. However, despite this mild clinical course, the absence of digestive complications and the presence of some of these variants in population databases at very low frequency, our findings suggest that these patients are at higher risk of arterial events and should be considered and followed as such. Heterogeneity in phenotypic expression in vEDS has been suggested previously, ranging from variable clinical phenotypes for non-glycine missense variants located in the C-terminal domain to overlapping with hypermobile EDS for non-glycine variants located in the triple helix domain of type III collagen.^{1,18} As for patients with Marfan syndrome that remained undiagnosed when staged by the Ghent nosology prior to its latest revision,¹⁹ it may be necessary to broaden indications for genetic testing in vEDS. For example, patients with only one major criterion and a familial history of sudden death, or patients with a major vascular event before the age of 50 years even in the absence of evocative clinical signs (as evidenced by the absence of clinical signs of the disorder in the relatives *versus* the index cases of our cohort (Table 1)),²⁰ may be discussed as candidates for *COL3A1* genetic screening, once common underlying causes of cardiovascular disease have been ruled out. As we specifically showed in group 3, the absence of digestive events associated with a high prevalence of aortic complications could mimic familial thoracic aortic aneurysm and dissection (FTAAD, OMIM #607086 for general informations). It is important to suspect vEDS also in these patients as therapeutic decisions are different from the ones for other inherited thoracic aortic aneurysms and dissections. Recent progress in sequencing technology should facilitate this wider genetic screening, even though some variants will require biochemical studies to establish their causal role on type III collagen alteration of assembly and/or production.

CONCLUSION

This study confirms in a large European cohort fully characterized in a dedicated diagnostic network that vEDS is a severe early-onset disease. In addition to the previously described subgroup of patients with variants leading to haplo-insufficiency, it also identifies subgroups of patients bearing non-glycine missense and/or genetic variations at the

C- and N-termini of type III procollagens, which are characterized by a later-onset and a milder phenotype, with no or very rare digestive complications. These findings may affect diagnostic strategy, genetic counselling and clinical follow-up of patients with vascular Ehlers–Danlos syndrome.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

In memory of Dr Jerome Perdu who initiated the Centre for Rare Vascular Diseases at the Hôpital Européen Georges Pompidou in Paris, France and was involved in the care of several patients and families reported here. We thank all patients and families and the Association Française des Syndromes d'Ehlers–Danlos (AFSED) for their confidence and continuing support. We acknowledge the contributions of clinicians of the 'Centres de Compétence' of the French network of vEDS patients. We also thank the technicians of the genetics laboratory (Magali Bousson, Valérie Boccio, Annabelle Venisse, Nelly Lepottier, Valérie Nau, Isabelle Roncelin) for their contribution to the *COL3A1* genetic analysis. This study was supported by the Assistance Publique des Hôpitaux de Paris, the INSERM, the French Ministry of Health through the program 'Maladies Rares 2006' and the Soutien aux Technologies Innovantes et Coûteuses (STIC 2006), the Agence Nationale pour la Recherche (10-BLAN-1129 GDPM2) and by the European Community's 7th Framework Program (HEALTH-F2-2007-Fighting Aneurysm Disease program).

- 1 Pope FM, Narcisi P, Nicholls AC, Germaine D, Pals G, Richards AJ: COL3A1 mutations cause variable clinical phenotypes including acrogeria and vascular rupture. *Br J Dermatol* 1996; **135**: 163–181.
- 2 Pepin M, Schwarze U, Superti-Furga A, Byers PH: Clinical and genetic features of Ehlers-Danlos syndrome type IV, the vascular type. *N Engl J Med* 2000; **342**: 673–680.
- 3 Beighton P, De Paepe A, Steinmann B, Tsipouras P, Wenstrup RJ: Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). *Am J Med Genet* 1998; **77**: 31–37.
- 4 Dagleish R: The Human Collagen Mutation Database 1998. *Nucleic Acids Res* 1998; **26**: 253–255.
- 5 Leistriz DF, Pepin MG, Schwarze U, Byers PH: COL3A1 haploinsufficiency results in a variety of Ehlers-Danlos syndrome type IV with delayed onset of complications and longer life expectancy. *Genet Med* 2011; **13**: 717–722.
- 6 Schwarze U, Schievink WI, Petty E *et al*: Haploinsufficiency for one COL3A1 allele of type III procollagen results in a phenotype similar to the vascular form of Ehlers-Danlos syndrome, Ehlers-Danlos syndrome type IV. *Am J Hum Genet* 2001; **69**: 989–1001.
- 7 Prockop DJ, Kivirikko KI: Heritable diseases of collagen. *N Engl J Med* 1984; **311**: 376–386.
- 8 Bornstein P, Sage H: Structurally distinct collagen types. *Annu Rev Biochem* 1980; **49**: 957–1003.
- 9 Shalhub S, Black 3rd JH, Cecchi AC *et al*: Molecular diagnosis in vascular Ehlers-Danlos syndrome predicts pattern of arterial involvement and outcomes. *J Vasc Surg* 2014; **60**: 160–169.
- 10 Ferre FC, Frank M, Gogly B *et al*: Oral phenotype and scoring of vascular Ehlers-Danlos syndrome: a case-control study. *BMJ Open* 2012; **2**: e000705.
- 11 CoreTeam: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria, 2013.
- 12 Persikov AV, Pillitteri RJ, Amin P, Schwarze U, Byers PH, Brodsky B: Stability related bias in residues replacing glycines within the collagen triple helix (Gly-Xaa-Yaa) in inherited connective tissue disorders. *Hum Mutat* 2004; **24**: 330–337.
- 13 Pepin MG, Schwarze U, Rice KM, Liu M, Leistriz D, Byers PH: Survival is affected by mutation type and molecular mechanism in vascular Ehlers-Danlos syndrome (EDS type IV). *Genet Med* 2014; **16**: 881–888.
- 14 Beck K, Chan VC, Shenoy N, Kirkpatrick A, Ramshaw JA, Brodsky B: Destabilization of osteogenesis imperfecta collagen-like model peptides correlates with the identity of the residue replacing glycine. *Proc Natl Acad Sci USA* 2000; **97**: 4273–4278.
- 15 Mizuno K, Boudko S, Engel J, Bachinger HP: Vascular Ehlers-Danlos syndrome mutations in type III collagen differently stall the triple helical folding. *J Biol Chem* 2013; **288**: 19166–19176.
- 16 Schwarze U, Goldstein JA, Byers PH: Splicing defects in the COL3A1 gene: marked preference for 5' (donor) splice-site mutations in patients with exon-skipping mutations and Ehlers-Danlos syndrome type IV. *Am J Hum Genet* 1997; **61**: 1276–1286.

- 17 Giunta C, Steinmann B: Characterization of 11 new mutations in COL3A1 of individuals with Ehlers-Danlos syndrome type IV: preliminary comparison of RNase cleavage, EMC and DHPLC assays. *Hum Mutat* 2000; **16**: 176–177.
- 18 Narcisi P, Richards AJ, Ferguson SD, Pope FM: A family with Ehlers-Danlos syndrome type III/articular hypermobility syndrome has a glycine 637 to serine substitution in type III collagen. *Hum Mol Genet* 1994; **3**: 1617–1620.
- 19 De Paepe A, Devereux RB, Dietz HC, Hennekam RC, Pyeritz RE: Revised diagnostic criteria for the Marfan syndrome. *Am J Med Genet* 1996; **62**: 417–426.
- 20 Faivre L, Collod-Beroud G, Callewaert B *et al*: Pathogenic FBN1 mutations in 146 adults not meeting clinical diagnostic criteria for Marfan syndrome: further delineation of type 1 fibrillinopathies and focus on patients with an isolated major criterion. *Am J Med Genet A* 2009; **149A**: 854–860.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)

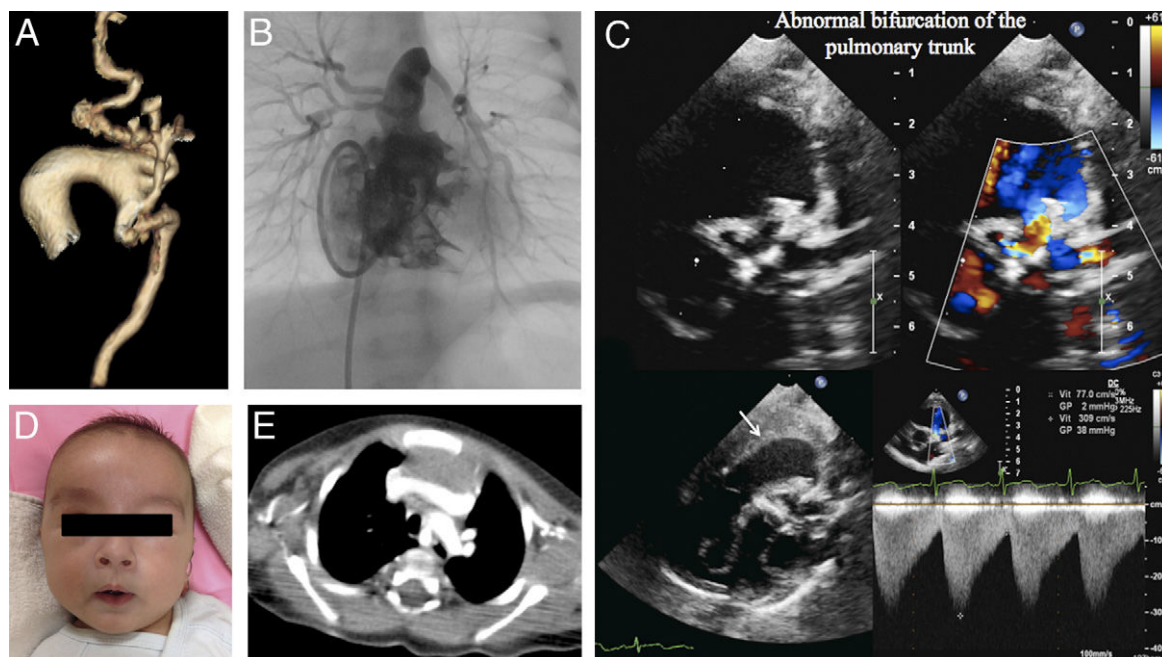
IMAGES IN CARDIOLOGY

Arterial Tortuosity Syndrome

Early Diagnosis and Association With Venous Tortuosity

Pamela Mocerì, MD,*† Juliette Albuissou, MD,‡ Marie Saint-Faust, MD,* Florence Casagrande, MD,* Fabienne Giuliano, MD,*§ Caroline Devos, MD,* Philippe Benoit, MD,* Nicolas Hugues, MD,|| Dorothee Ducreux, MD,¶ Pierre Cerboni, MD,† Christian Dageville, MD,* Xavier Jeunemaitre, MD, PhD‡

Nice and Paris, France; and Monaco



From the *Pediatric University Hospitals of Nice, GCS-CHU Lenval, Nice, France; †University Hospital of Nice, Cardiology Department, Pasteur Hospital, Nice, France; ‡Assistance Publique-Hôpitaux de Paris, Department of Genetics, National Referral Centre for Rare Vascular Diseases, European Hospital Georges Pompidou, Paris, France; §University Hospital of Nice, Department of Medical Genetics, Archet 2 Hospital, Nice, France; ||Cardio-thoracic Center of Monaco, Monaco; and the ¶University Hospital of Nice, Medical Imaging Department, Pasteur Hospital, Nice, France.

Manuscript received
June 18, 2012;
accepted June 26, 2012.

A newborn, first child of consanguineous parents was admitted to intensive care because of persistent pulmonary hypertension. The diagnosis of arterial tortuosity syndrome (ATS) was suspected due to severe tortuosity of the aorta (**A**) and pulmonary arteries (**B** and **C**, Online Video 1). Despite initial hospitalization, the first months of life were uneventful. He progressively developed peripheral pulmonary stenosis and moderate coarctation of the aorta. His facial features were characteristic (**D**), with elongated face, hypertelorism, and a long philtrum. He also presented with bilateral inguinal hernia, typical of connective tissue disease that is associated with ATS. The diagnosis was confirmed by genetic analysis: the patient was homozygous for a mutation in exon 2 (c.510G>A, p.Trp170X) of the *SLC2A10* gene. The particularity of this case is, for the first time to our knowledge, an early diagnosis of persistent pulmonary hypertension and an association of venous tortuosity (**E**; distorted innominate vein) with ATS.

CLINICAL UTILITY GENE CARD

Clinical utility gene card for: Arterial tortuosity syndrome

Juliette Albuissou^{*,1,2,3}, Pamela Mocerri^{4,5}, Elisabeth Flori⁶, Emre Belli⁷, Céline Gronier⁸ and Xavier Jeunemaitre^{1,2,3}

European Journal of Human Genetics (2015) **23**, doi:10.1038/ejhg.2014.294; published online 21 January 2015

1. DISEASE CHARACTERISTICS

1.1 Name of the disease (synonyms)

Arterial tortuosity syndrome (arterial tortuosity).

1.2 OMIM# of the disease

208050.

1.3 Name of the analysed genes or DNA/chromosome segments

SLC2A10, also named *GLUT10*, located in 20q13.12 chromosomal segment.

1.4 OMIM# of the gene(s)

606145.

1.5 Mutational spectrum

To date, <30 mutations have been described, approximately one-third of them are present in the public *SLC2A10* variants database (<http://databases.lovd.nl/shared/genes/SLC2A10>).

The mutation spectrum comprises transitions, transversions, large (exonic) deletions, small deletions and splice-site mutations.

Consequences:

frameshift, missense, nonsense, splice-site, haploinsufficiency by mRNA decay.

Comment:

The collection of variants in a shared database is of great importance, particularly in the case of very rare disorders like ATS, as it improves the diagnosis in laboratories by helping analysis and interpretation of these rare variants. All published and unpublished case with identified *SLC2A10* pathogenic variant(s) should be submitted in an open-access database.

1.6 Analytical methods

Sequencing of both strands is necessary. This can be completed if needed by a semi-quantitative technique seeking for partial or complete deletions of the gene, as described.¹ As this disorder is recessive and frequently caused by homozygous mutations, denaturation-based screening techniques can be performed on equimolar mixtures of patient and control DNA.

1.7 Analytical validation

Mutations are responsible for a loss-of-function of the corresponding allele. Analytical validation should systematically integrate conservation analysis of the mutant nucleotide(s) and aminoacid(s), as well as the surrounding sequence(s). The variation(s) should systematically be searched in databases of polymorphisms and pathogenic variations.

1.8 Estimated frequency of the disease (incidence at birth ('birth prevalence') or population prevalence. If known to be variable between ethnic groups, please report)

The incidence is extremely low although no accurate data exist. Estimates vary between 1/100 000 to 1/500 000 live births. The frequency does not seem to vary between different ethnical backgrounds, although in some populations a high level of consanguinity possibly favours the occurrence of new cases).²

1.9 Diagnostic setting

	Yes	No
A. (Differential) diagnostics	<input checked="" type="checkbox"/>	<input type="checkbox"/>
B. Predictive testing	<input type="checkbox"/>	<input checked="" type="checkbox"/>
C. Risk assessment in relatives	<input type="checkbox"/>	<input checked="" type="checkbox"/>
D. Prenatal	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Comment:

A first comment must be made concerning differential diagnostics: One main differential diagnosis has to be considered in the classical neonatal or infantile form of the disease. This syndromic condition, named Autosomal Recessive Cutis Laxa Type 1 (ARCL1b) is caused by mutations in *FBLN4* gene encoding fibuline 4. ARCL1b associates diffuse arterial tortuosity combining stenosis and aneurysms with severe to mild cutis laxa, inconstant diaphragmatic and inguinal hernias and facial dysmorphism.³⁻⁵ Differential diagnosis between those two conditions can be considered on the basis of slight dysmorphic differences of the face, a tendency to more generalised and severe cutis laxa in ARCL1b patients.⁴

In cases when the ATS is revealed by late childhood or adulthood cardiovascular complications, the differential diagnosis includes

¹AP-HP, Département de Génétique, Hôpital Européen Georges Pompidou, Paris, France; ²Université Paris Descartes, PRES Sorbonne Paris Cité, Paris, France; ³Centre de Référence des Maladies Vasculaires Rares, Paris, France; ⁴Unité de Cardiologie, Hôpital Pasteur, CHU de Nice, Nice, France; ⁵Unité de Cardiologie pédiatrique, Hôpitaux universitaires de Nice CHU-Lenval, Nice, France; ⁶Service de Cytogénétique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; ⁷Centre Chirurgical Marie Lannelongue, Université Paris-Sud, Centre de Référence Malformation Cardiaques Congénitales Complexes M3C, Paris, France; ⁸Pôle de Cardiologie, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

*Correspondence: Dr J Albuissou, AP-HP, Département de Génétique, Hôpital Européen Georges Pompidou, 20 rue Leblanc, 75015 Paris, France. Tel: +33 156093881; Fax: +33 156093884; E-mail: juliette.albuissou@egp.aphp.fr

Received 15 November 2013; revised 10 November 2014; accepted 9 December 2014; published online 21 January 2015

connective tissue disorders affecting the arterial system as Loey's–Dietz syndromes (LDS),⁶ and Ehlers–Danlos syndromes (EDS), particularly the vascular form. All differential diagnoses of ATS are available for diagnostic testing and should be explored in case of negativity of *SLC2A10* screening (see 2.3).

A second comment is necessary concerning prenatal diagnosis: ATS severity is variable and its outcome is currently unpredictable. As early severe morbidity and early demise have frequently been reported, prenatal diagnosis can be offered for this disease when the mutations have been identified.

2. TEST CHARACTERISTICS

		Genotype or disease		A: True positives	C: False negative
		Present	Absent	B: False positives	D: True negative
Test					
Positive	A	B	Sensitivity:	A/(A+C)	
			Specificity:	D/(D+B)	
Negative	C	D	Positive predictive value:	A/(A+B)	
			Negative predictive value:	D/(C+D)	

2.1 Analytical sensitivity

(proportion of positive tests if the genotype is present)

Almost 100%. When only one mutation has been identified by sequencing, the test should be completed by a semi-quantitative technique, allowing the detection of loss or gains of one or more exons of the gene. In these conditions the AS is close to 100%. When two heterozygous mutations are found, parental testing should always be performed to confirm compound heterozygosity.

2.2 Analytical specificity

(proportion of negative tests if the genotype is not present)

100%. As this is a recessive disorder, two loss-of-function mutations are expected, and false-positive results with two unknown variants in a patient's *SLC2A10* genomic sequence are not to be expected.

2.3 Clinical sensitivity

(proportion of positive tests if the disease is present)

The clinical sensibility can be dependent on variable factors such as age or family history. In such cases a general statement should be given, even if a quantification can only be made case by case.

When the clinical diagnosis has correctly been established, very few false negative patients are expected. The high level of global arterial tortuosity, associated with arterial stenosis (mostly of the pulmonary circulation) and specific dysmorphic features in infancy are generally sufficient to propose the diagnosis of ATS. However, as mentioned in 1.9, differential diagnoses can be proposed in infancy or in adulthood.⁹

Molecular screening of genomic sequences of *FBLN4* (for ARCL1b), *COL3A1*, *COL5A1* and *COL5A2* (for EDS) and *TGF β 2*, *TGF β RI/2* and *SMAD3* (for LDS) can be performed depending on the clinical presentation of the patient, in case of negativity of *SLC2A10* screening.

2.4 Clinical specificity

(proportion of negative tests if the disease is not present)

The clinical specificity can be dependent on variable factors such as age or family history. In such cases a general statement should be given, even if a quantification can only be made case by case.

Same comment as for 2.3.

2.5 Positive clinical predictive value

(life time risk to develop the disease if the test is positive).

100% if a prenatal diagnosis is performed. Not applicable in adulthood as this condition is a recessive childhood disorder.

2.6 Negative clinical predictive value

(probability of developing the disease if the test is negative).

Assume an increased risk based on family history for a non-affected person. Allelic and locus heterogeneity may need to be considered.

Index case in that family had been tested:

100%.

Index case in that family had not been tested:

Assuming the clinical diagnosis is right, 100%. In case of erroneous diagnosis, it depends on the differential diagnosis, which can either correspond to recessive or dominant conditions. All differential diagnoses corresponding to adult-onset, dominant disorders should be cautiously kept in mind as penetrance is highly variable depending on families and age of the patients.

3. CLINICAL UTILITY

3.1 (Differential) diagnostics: the tested person is clinically affected (To be answered if in 1.9 'A' was marked)

3.1.1 Can a diagnosis be made other than through a genetic test?

No (continue with

3.1.4)

Yes

Clinically

Imaging

Endoscopy

Biochemistry

Electrophysiology

Other (please describe) The diagnosis can be strongly suggested by classical

neonatal or infantile presentation. Though, all cases should be confirmed by a genetic test as explained above.

3.1.2 Describe the burden of alternative diagnostic methods to the patient

In classical cases, clinical diagnosis can be strongly suggested by combining family history, physical examination, echocardiography and arterial imaging like angioMRI or CT.¹

Patients are often born from consanguineous marriages. They generally present at birth with facial dysmorphism including beaked nose, blepharophimosis, hypertelorism, downslanting palpebral fissures and a small mouth with retrognathism. A hyperextensible skin or mild cutis laxa can exist. Fingers are long and thin. Inguinal hernias are frequent, but diaphragmatic hernias rarely occur.^{1,6}

Infants with ATS may develop dyspnoea revealing pulmonary hypertension, due to stenoses of the pulmonary arteries or branches, which can be confirmed by cardiovascular imaging (ultrasound, catheterisation, angioCT or angioMRI) showing truncal or branch stenosis of the pulmonary arteries followed by a dilated segment.

Generalised severe arterial tortuosity is present, and can be associated with venous tortuosity.⁷ The aortic arch often shows elongation and tortuosity, followed by a stenotic prediaphragmatic aortic segment.^{7,8}

All medium-sized arteries are highly tortuous,⁹ and may show localised aneurysms and stenoses. This high level of tortuosity may result in functional stenosis at multiple levels.

This typical presentation allows clinicians to make the diagnosis of ATS.

3.1.3 How is the cost effectiveness of alternative diagnostic methods to be judged?

Clinical and imaging methods are actually used to get an accurate clinical evaluation of the patients, which is necessary for their long-term follow-up, particularly concerning interventional or surgical treatment.

When a genetic test is available, it should always be performed to propose familial screening and genetic counselling of carriers and prenatal diagnosis, if necessary.

3.1.4 Will disease management be influenced by the result of a genetic test?

No	<input type="checkbox"/>
Yes	<input checked="" type="checkbox"/>
Therapy (please describe)	Not applicable.
Prognosis (please describe)	Not applicable.
Management (please describe)	Not applicable.

Comment:

Catheter interventions (balloon dilatation ± stenting), cardiovascular surgery of aneurysms and stenoses, associated with medical treatment of general and/or pulmonary hypertension, are the therapeutic cornerstones of the disorder. There is no modification of the treatment to expect after genotyping.

3.2 Predictive setting: the tested person is clinically unaffected but carries an increased risk based on family history (To be answered if in 1.9 'B' was marked)**3.2.1 Will the result of a genetic test influence lifestyle and prevention?**

If the test result is positive (please describe)

Not applicable.

If the test result is negative (please describe)

Not applicable.

3.2.2 Which options in view of lifestyle and prevention does a person at-risk have if no genetic test has been done (please describe)?

Not applicable.

3.3 Genetic risk assessment in family members of a diseased person (To be answered if in 1.9 'C' was marked)**3.3.1 Does the result of a genetic test resolve the genetic situation in that family?**

Not applicable.

3.3.2 Can a genetic test in the index patient save genetic or other tests in family members?

Not applicable.

3.3.3 Does a positive genetic test result in the index patient enable a predictive test in a family member?

Not applicable.

3.4 Prenatal diagnosis

(To be answered if in 1.9 'D' was marked)

3.4.1 Does a positive genetic test result in the index patient enable a prenatal diagnosis?

Yes, prenatal diagnosis can be performed in fetuses of parents having an affected child. It might also be proposed in the offspring of an affected adult and a carrier partner.

4. IF APPLICABLE, FURTHER CONSEQUENCES OF TESTING

Please assume that the result of a genetic test has no immediate medical consequences. Is there any evidence that a genetic test is nevertheless useful for the patient or his/her relatives? (Please describe)

In populations where consanguineous marriages occur frequently, heterozygous carrier identification might be useful in a patient's relative, if prenatal diagnosis for the offspring is considered.²

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by EuroGentest2 (Unit 2: 'Genetic testing as part of health care'), a Coordination Action under FP7 (Grant agreement Number 261469), the European Society of Human Genetics and the Centre de Référence des Maladies Vasculaires Rares.

- 1 Callewaert BL, Willaert A, Kerstjens-Frederikse WS *et al*: Arterial tortuosity syndrome: clinical and molecular findings in 12 newly identified families. *Hum Mutat* 2008; **29**: 150–158.
- 2 Faiyaz-Ul-Haque M, Zaidi SH, Wahab AA *et al*: Identification of a p.Ser81Arg encoding mutation in SLC2A10 gene of arterial tortuosity syndrome patients from 10 Qatari families. *Clin Genet* 2008; **74**: 189–193.
- 3 Sawyer SL, Dicke F, Kirton A *et al*: Longer term survival of a child with autosomal recessive cutis laxa due to a mutation in FBLN4. *Am J Med Genet A* 2013; **161A**: 1148–1153.
- 4 Kappanayil M, Nampoothiri S, Kannan R *et al*: Characterization of a distinct lethal arteriopathy syndrome in twenty-two infants associated with an identical, novel mutation in FBLN4 gene, confirms fibulin-4 as a critical determinant of human vascular elastogenesis. *Orphanet J Rare Dis* 2012; **7**: 61.
- 5 Renard M, Holm T, Veith R *et al*: Altered TGFβ signaling and cardiovascular manifestations in patients with autosomal recessive cutis laxa type I caused by fibulin-4 deficiency. *Eur J Hum Genet* 2010; **18**: 895–901.
- 6 Coucke PJ, Willaert A, Wessels MW *et al*: Mutations in the facilitative glucose transporter GLUT10 alter angiogenesis and cause arterial tortuosity syndrome. *Nat Genet* 2006; **38**: 452–457.
- 7 Mocerì P, Albuissou J, Saint-Faust M *et al*: Arterial tortuosity syndrome: early diagnosis and association with venous tortuosity. *J Am Coll Cardiol* 2013; **61**: 783.
- 8 Kalfa D, Gronier C, Ly M, Le Bret E, Roussin R, Belli E: Giant aortic aneurysm in an infant with arterial tortuosity syndrome. *Ann Thorac Surg* 2012; **94**: e51.
- 9 Castori M, Ritelli M, Zoppi N *et al*: Adult presentation of arterial tortuosity syndrome in a 51-year-old woman with a novel homozygous c.1411+1G>A mutation in the SLC2A10 gene. *Am J Med Genet A* 2012; **158A**: 1164–1169.