



Faculté des sciences et techniques

Master Biologie-Santé

Sciences chirurgicales et nouvelles technologies interventionnelles

2017/2018

GUIMBRETIERE Guillaume

Déterminants génétiques et non génétiques impliqués dans le développement de la sténose valvulaire aortique dans le cadre d'une forme familiale.

Directeur de recherche : Mr Thierry Le Tourneau

Co-directeur de recherche : Mme Solena Le Scouarnec



Institut du Thorax

INSERM UMR 1087 / CNRS UMR 6291

Equipe 1- Génétique des maladies cardiovasculaires

TABLE DES MATIÈRES

ABREVIATIONS	4
INTRODUCTION	6
1) CARDIOPATHIES VALVULAIRES DU CŒUR GAUCHE	6
2) ANATOMIE DE LA VALVE AORTIQUE	7
3) RETRECISSEMENT VALVULAIRE AORTIQUE	8
3.1 <i>Histoire naturelle</i>	8
3.2 <i>Traitement</i>	9
4) GENETIQUE ET RAC	10
4.1 <i>Evolution des approches et concepts génétiques</i>	10
4.2 <i>Variant fréquent et pathologies fréquentes</i>	12
➤ <i>Approche gène candidat</i>	12
➤ <i>Approche GWAS/TWAS</i>	13
4.3 <i>Variant rare et approche familiale</i>	14
4.4 <i>L'apport du séquençage de nouvelle génération</i>	15
5) OBJECTIF DU MASTER	16
MATERIELS ET METHODES	17
1) ENQUETE FAMILIALE	17
2) RECUEIL DES DONNEES	17
2.1 <i>Cliniques</i>	17
2.2 <i>Biologiques</i>	17
2.3 <i>Echocardiographiques</i>	17
3) SCORE VALVULAIRE AORTIQUE PAR CT SCAN	18
4) HISTOLOGIE	18
5) ANALYSES STATISTIQUES	19
6) INVESTIGATIONS GENETIQUES	19
6.1 <i>bio-collections</i>	19
6.2 <i>Séquençage génome entier (WGS)</i>	19
6.3 <i>Analyse bio-informatiques</i>	20
6.4 <i>Séquençage capillaire</i>	22
6.5 <i>Génotypage haut débit</i>	23
6.6 <i>Identité par descendance (IBD)</i>	23
7) SELECTION DES VARIANTS APRES SEQUENÇAGE GENOME ENTIER	24
7.1 <i>Filtre des variants identifiés à partir de l'outil knime4bio</i>	24
7.2 <i>Filtre des variants fonctionnels</i>	24
7.3 <i>Conservation des variants rares</i>	25
RESULTATS	26
1) ENQUETE FAMILIALE : ARBRE GENEALOGIQUE	26
2) CLINIQUE ET IMAGERIE :	26
3) SCORE CALCIQUE VALVULAIRE AORTIQUE	31
4) ANATOMOPATHOLOGIE	31
5) INVESTIGATIONS GENETIQUES	35
5.1 <i>Régions génomiques partagées par les individus atteints</i>	35
5.2 <i>Identification des variants rares</i>	36
5.3 <i>Analyses combinées des données séquençage génome entier et des régions IBD identifiées</i>	37
5.4 <i>Sélection des variants pour le séquençage capillaire après analyse combinée IBD/WGS</i>	37
5.5 <i>Séquençage capillaire</i>	37
DISCUSSION	40
1) CHOIX DU MATERIEL ET DES METHODES	40

2) RESULTATS	41
3) PROBLEMES RENCONTRES	42
CONCLUSION ET PERSPECTIVES	44
REFERENCES BIBLIOGRAPHIQUES	45
ANNEXES	49

REMERCIEMENTS

Un remerciement à l'association ADICARE qui m'a permis d'effectuer cette belle année de recherche qui je pense me sera utile tout au long de ma vie active.

ABREVIATIONS

ADN : Acide DésoxyriboNucléique

RAC : Rétrécissement Aortique Calcifié

SAVR : Surgical Aortic Valve Replacement

TAVI : Transcatheter Aortic Valve Implantation

PCR : Polymerase Chain Reaction

PGH : Projet Génome Humain

GWAS : Genome Wide Association

TWAS : Transcriptome Wide Association

SNP : Single Nucleotid Polymorphism

WGS : Whole Genome Sequencing

WES : Whole Exome Sequencing

PAPS : Pression Arterielle Pulmonaire Systolique

NGS : Next Generation Sequencing

VCF : Variant Call Format

VEP : Variant Effect Predictor

GERP : Genomic Evolutionnary Rate Profiling

MAF : Minor Allele Frequency

IBD : Identity By Descendance

ETT : Echographie Trans-Thoracique

SC : Surface Corporelle

CCVG : Chambre Chasse Ventricule Gauche

JST : Jonction Sino-Tubulaire

INTRODUCTION

Les valvulopathies cardiaques sont des pathologies relativement fréquentes et touchent près de 2,5% de la population générale dans les pays industrialisés (1,2) avec une forte prédominance pour les affections valvulaires dégénératives. Les seuls traitements curatifs des valvulopathies indiqués après échec des traitements médicaux sont les traitements interventionnels.

1) CARDIOPATHIES VALVULAIRES DU CŒUR GAUCHE

Les cardiopathies valvulaires cardiaques sont des affections fréquentes et hétérogènes. On retrouve des affections mono-valvulaires ou pluri-valvulaires non syndromiques (prolapsus valvulaire mitral, rétrécissement aortique etc.) et des formes syndromiques avec des désordres du tissu conjonctif généralisés comme le syndrome de Marfan.

La prévalence de ces pathologies valvulaires est difficile à évaluer au sein de la population générale. L'auscultation seule ne suffit pas à poser le diagnostic, il convient donc de réaliser des examens écho-cardiographiques sur un échantillon représentatif de la population générale. Une étude réalisée en ce sens aux États-Unis sur 11 911 individus a permis d'identifier que 1,7% de la population générale étaient porteurs d'une insuffisance mitrale, 0,1% d'un rétrécissement mitral, 0,5% d'une insuffisance aortique et 0,4% d'un rétrécissement aortique (1).

Des observations similaires ont été faites dans de plus petites cohortes en Europe (2,3). Les auteurs décrivent une augmentation de la prévalence des affections valvulaires avec l'âge (jusqu'à 13,3% au-delà de 75 ans). Les étiologies dégénératives apparaissent comme largement majoritaires du fait du vieillissement de la population et de l'amélioration des conditions sanitaires dans les pays occidentaux.

Ces cardiopathies valvulaires sont associées à une forte augmentation de la morbi-mortalité et à une augmentation importante des coûts de santé publique (2,4).

La majorité des patients présente des formes bénignes de valvulopathies avec de faibles répercussions sur les fonctions cardiaques. Cependant, certains patients présentent des formes sévères qui peuvent être à l'origine de complications allant de l'insuffisance ou la sténose valvulaire, au développement d'endocardites, d'une insuffisance cardiaque, d'accidents thromboemboliques, de troubles du rythme cardiaque, voire même de mort subite (5).

Pendant de nombreuses années, les valvulopathies ont été décrites comme résultant de processus dégénératifs passifs dus à l'âge et aux contraintes hémodynamiques répétées (« wear and tear »). Il est maintenant largement accepté et décrit que les valvulopathies résultent de processus actifs biologiquement régulés.

De plus, précédemment considérées comme des pathologies idiopathiques ou isolées, les pathologies valvulaires, comme le rétrécissement aortique calcifié ou le prolapsus valvulaire mitral, présentent un caractère familial suggérant que des facteurs génétiques sont impliqués dans leur apparition (6).

2) ANATOMIE DE LA VALVE AORTIQUE

La valve aortique est située entre la chambre de chasse du ventricule gauche et l'aorte ascendante, elle est constituée de trois cuspidés (ou valvules) semi-lunaires. Issues embryologiquement des bourgeons endocardiques les cuspidés ont un bord pariétal inséré sur un anneau fibreux aortique et un bord libre constituant par leur union la valve aortique et son rôle anti-reflux. Les cuspidés soutiennent les trois sinus de Valsalva (pour la première fois décrit par Leonardo da Vinci en 1508) (figure 1) qui sont des renflements de la partie initiale de l'aorte ayant un rôle crucial pour la fermeture des cusps et l'étanchéité de la valve aortique en diastole. Il y a deux valvules antérieures, une droite (qui est la plus antérieure) et une gauche, ainsi qu'une valvule postérieure (dans ses premières représentations Leonardo da Vinci considérait deux valvules postérieures et une antérieure). La partie postérieure de la valvule antéro-gauche a une insertion commune avec le grand feuillet de la valve mitrale (ou valve mitrale antérieure), on parle de « continuité mitro-aortique ». Un nodule fibreux (nodule d'Arantius) renfle la partie moyenne de chaque bord libre valvulaire. Au-dessus des valvules antérieures gauche et droite s'insèrent les orifices coronaires gauche et droit respectivement.

La valve aortique est en rapport avec le nœud atrio-ventriculaire, situé au sommet de la chambre de chasse ventriculaire gauche entre l'anneau atrio-ventriculaire gauche et la commissure séparant la cusp non coronaire de la cusp coronaire droite.

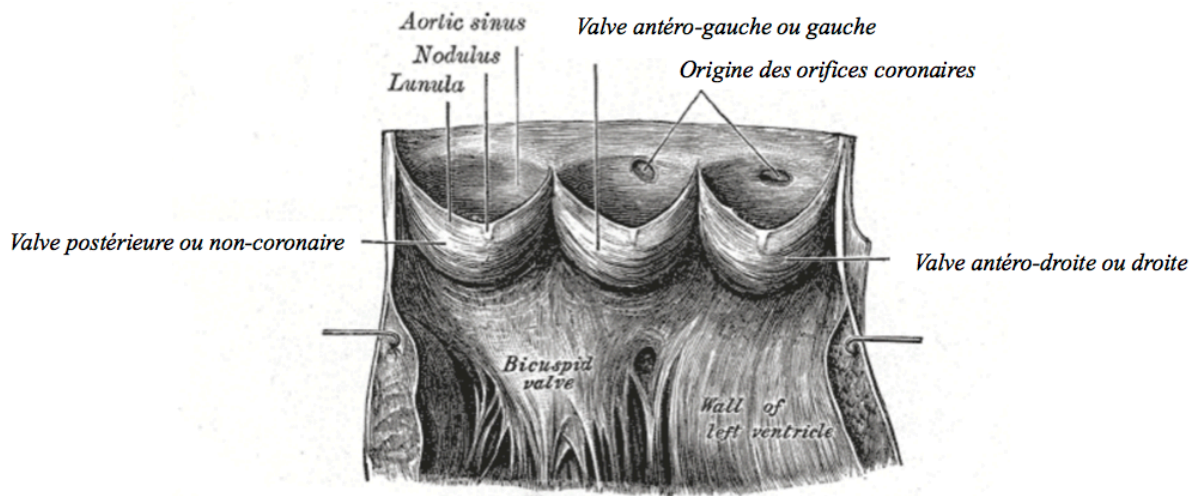


Figure 1 : Coupe anatomique d'une valve aortique normale adaptée des premiers croquis de Leonardo Da Vinci

3) RÉTRÉCISSEMENT VALVULAIRE AORTIQUE

3.1 HISTOIRE NATURELLE

La maladie de calcifiante valvulaire aortique ou rétrécissement aortique calcifié (RAC) est basée sur un processus de remodelage et de calcification pathologique conduisant à une modification morphologique des cusps aortiques. Ces modifications à type d'épaississement, de distorsion, de rigidification ainsi que de fibrose aboutissent à l'apparition de nodules cartilagineux et/ou osseux(7). La maladie a été divisée en deux catégories fonctionnelles traduisant une évolution naturelle de la maladie : la sclérose valvulaire aortique et la sténose valvulaire aortique. La distinction est basée sur l'obstruction de la voie d'éjection ventriculaire gauche, absente dans la sclérose valvulaire aortique (modification morphologique de la valve sans anomalie fonctionnelle significative) et présente dans la sténose valvulaire aortique (modification morphologique ayant un retentissement hémodynamique). Bien que cette classification soit utile pour la caractérisation du stade de la maladie dans la cadre de la recherche clinique, elle ne présage en rien de l'évolution de la pathologie d'une sclérose valvulaire aortique vers une sténose valvulaire aortique. En effet cette maladie dégénérative est basée sur un processus d'évolution lent, avec un continuum de la maladie débutant par des lésions de sclérose (proches des caractéristiques des plaques précoces d'athérosclérose) mais pouvant évoluer vers une sténose valvulaire aortique présentant des calcifications sévères. La prévalence du rétrécissement valvulaire aortique calcifiée (RAC), avec ou sans malformation valvulaire congénitale, augmente de façon exponentielle avec l'âge. Dans les sociétés industrialisées, la sclérose valvulaire aortique est présente chez plus de 25% des patients âgés de plus de 65 ans (8,9).

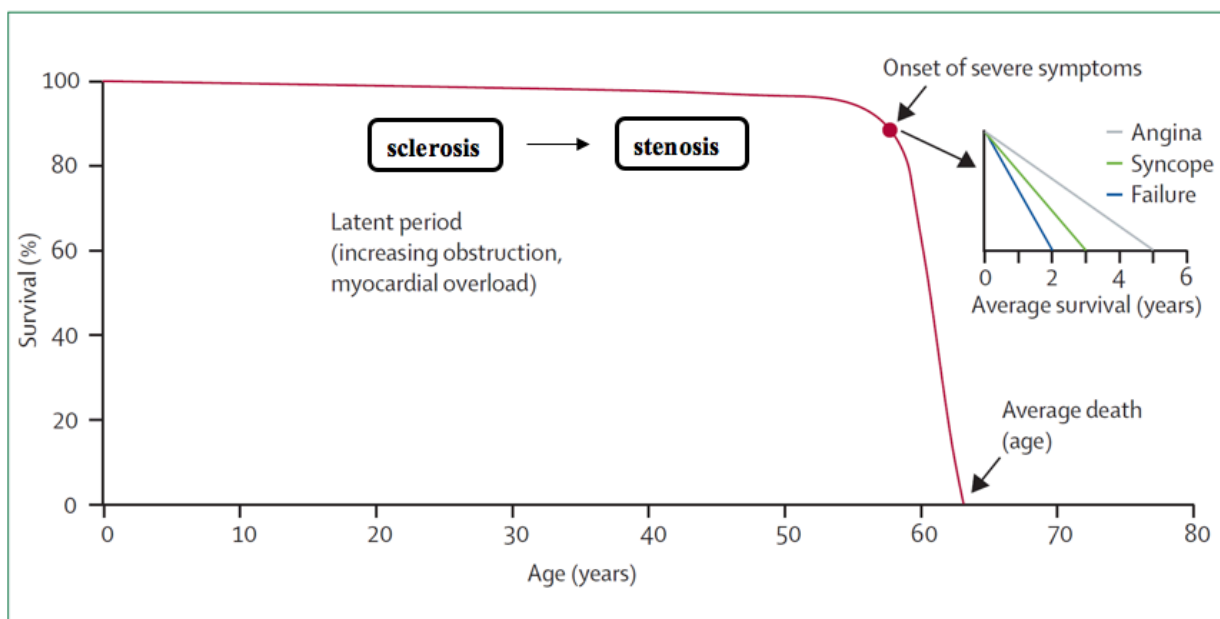


Figure 2 : Histoire naturelle du rétrécissement valvulaire aortique calcifié d'après Ross et Braunwald (Circulation 1968)

De nombreux facteurs ont été proposés pour expliquer la maladie calcifiante valvulaire aortique. Premièrement la bicuspidie valvulaire aortique (deux cuspidés au lieu de trois normalement), malformation congénitale présente chez 1 à 2 % de la population, et qui est un facteur majeur contribuant à la maladie calcifiante. Au cours de leur vie, la majorité des individus atteints d'un bicuspidie valvulaire aortique développent une pathologie de la valve

aortique le plus souvent à type de RAC (seulement 1% conserve une fonction valvulaire normale) (10). Un patient porteur d'une valve bicuspidé développent un RAC entre une à deux décennies avant un patient porteur d'une valve tricuspide. On soupçonne que les modifications hémodynamiques induites par cette conformation bicuspidé accélèrent le processus de calcification valvulaire. Cependant des anomalies biologiques intrinsèques doivent coexister et contribuer à un remodelage précoce chez ces individus. Au-delà des facteurs de risques irréfutables que sont les malformations congénitales de la valve aortique et l'âge d'autres facteurs de risques ont été identifiés tels que le tabagisme, le diabète, le sexe masculin, l'hypertension artérielle, l'hypercholestérolémie et le syndrome métabolique (7,11,12).

La figure 2 illustre le spectre de la maladie du patient « à risque » au patient présentant une sténose aortique grave en phase terminale. En effet une fois la sténose aortique établie, il existe un risque accru d'évènements cardiovasculaires. Dès le début des symptômes la survie chute et ce de façon spectaculaire lorsque la sténose aortique est sévère et très symptomatique (13). Le remplacement valvulaire aortique au début des symptômes prend ainsi tout son sens.

3.2 TRAITEMENT

Il n'y a pas de traitement préventif du rétrécissement valvulaire aortique, le seul traitement curatif est représenté par le remplacement valvulaire aortique. L'échocardiographie doppler est l'examen clé pour poser l'indication de l'intervention en dehors du caractère symptomatique ou non de cette sténose valvulaire aortique. Il existe aujourd'hui globalement deux types de prise en charge curatives de cette pathologie conditionnée par un score de risque type STS/Euroscore II. Le remplacement valvulaire aortique chirurgical classique avec clampage aortique et ouverture de l'aorte que nous nommerons SAVR (Surgical Aortic Valve Replacement) et la voie mini-invasive par les vaisseaux périphériques sans clampage aortique ni ouverture de l'aorte que nous nommerons TAVI (Transcatheter Aortic Valve Implantation). Les éléments à prendre en compte par la Heart Team pour choisir l'une ou l'autre des prises en charge sont citées dans le tableau 1 sachant qu'à ce jour et selon les recommandations de l'ESC/EACTS, le TAVI est seulement réservé aux patients à haut risque chirurgical (14).

Tableau 1 : Guidelines ESC/EACTS

Caractéristiques cliniques	TAVI	SAVR	Anatomie et aspects techniques	TAVI	SAVR
STS/Euroscore II<4% (logistic euroscore I <10%)		+	Accès favorable par les vaisseaux fémoraux	+	
STS/Euroscore II≥4% (logistic euroscore I ≥10%)	+		Pas d'accès (favorables) pour TAVI		+
Sévères comorbidités	+		Antécédent d'irradiation	+	
Age < 75 ans		+	Aorte porcelaine	+	
Age ≥ 75 ans	+		Pontages perméables	+	
Antécédent chirurgie cardiaque (redux)	+		Risque de mismatch (prothèse/patient)	+	
Fragilité	+		Déformation thoracique ou scoliose sévère	+	
Mobilité réduite	+		Distance courte entre ostia coronaire et anneau valvulaire aortique		+
Suspicion endocardite		+	Taille anneau valvulaire aortique en dehors taille disponible pour TAVI		+
Chirurgie concomitante	TAVI	SAVR	Anatomie valve aortique non favorable TAVI		+
Cardiopathie ischémique sévère nécessitant pontages aorto-coronaires		+	Morphologie valvulaire non favorable TAVI		+
Insuffisance mitrale primaire sévère		+	Thrombus aortique et / ou ventriculaire		+
Maladie tricuspidiennne sévère		+			
Anévrisme aorte ascendante		+			
Cardiomyopathie hypertrophique		+			

4) GÉNÉTIQUE ET RAC

4.1 EVOLUTION DES APPROCHES ET CONCEPTS GENETIQUES

Grace à la mise au point d'une technologie de séquençage par Sanger en 1977 puis au développement de la technique d'amplification en chaîne par polymérase (PCR : « polymerase chain reaction ») par Mullis en 1986, le « génie génétique » va permettre un développement important de la connaissance de la fonction des gènes.

Le concept « un gène responsable d'une pathologie » a rapidement été remis en cause. En effet les déterminants génétiques des pathologies apparaissent multiples et leur transmission peut être complexe. De plus, l'identification de variations génétiques fréquentes comme facteurs de prédisposition à certaines pathologies a modifié considérablement les approches génétiques des pathologies.

Ainsi trois catégories de variants responsables de pathologies génétiques ont été identifiées. On distingue (figure 3) :

- Les variants génétiques rares, qui possèdent des effets forts (variant causal à forte pénétrance vis à vis de la pathologie)
- Les variants fréquents à effets faibles (ou modérés)
- Les variants de fréquence intermédiaire présentant des effets modérés

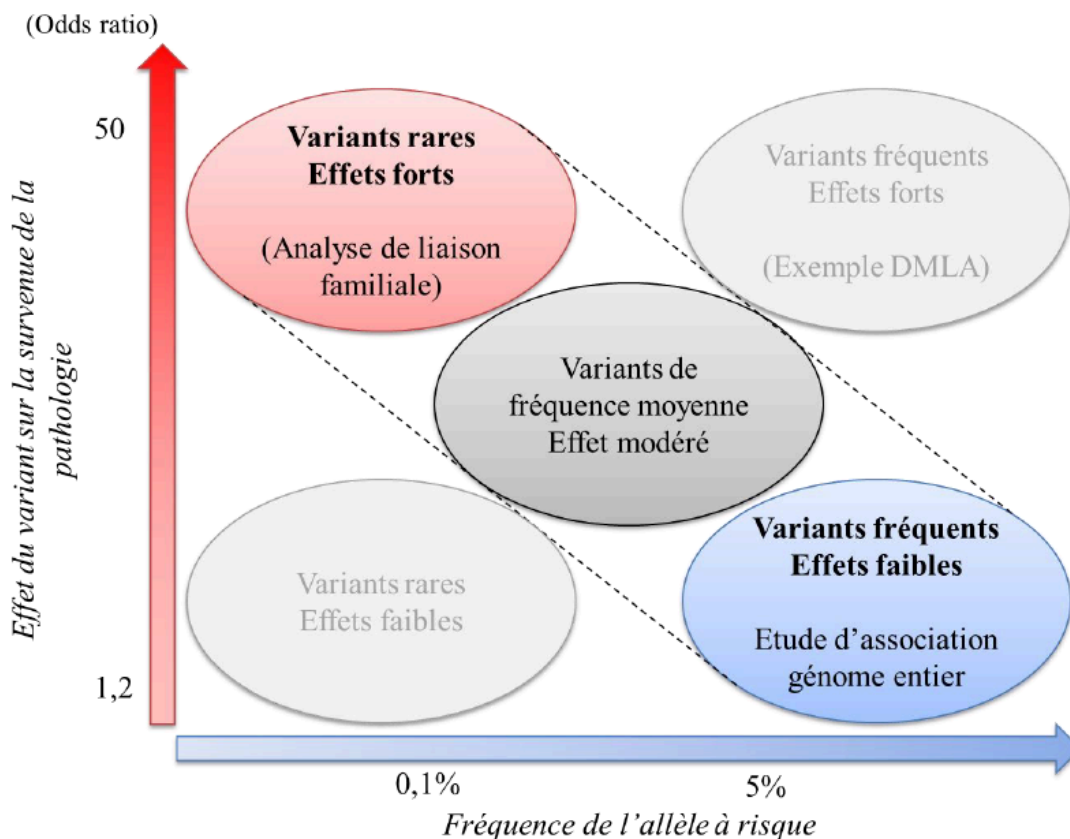


Figure 3 : Représentation schématique de l'effet pathogène de variants génétiques (odds ratio) en fonction de la fréquence de l'allèle à risque d'après Manolio et al 2009(15)

A partir de l'identification de variations génétiques rares et fréquentes, différentes hypothèses de causes génétiques des pathologies ont vu le jour. On considère ainsi que le spectre de fréquence des allèles à risque est à mettre en rapport avec le spectre de fréquence des pathologies. Ainsi les études génétiques se sont basées sur les théories du « common variant/common disease » (allèle 4 du gène APOE et maladie d'Alzheimer) et « rare variant rare disease » (CFTR et mucoviscidose) (16). Les approches génétiques utilisées pour identifier ces facteurs génétiques seront différentes en fonction du type de variant recherché et en fonction de la pathologie étudiée.

Pendant de nombreuses années le RAC était considéré comme une maladie dégénérative résultant seulement d'un processus d'usure mécanique (« wear and tear process »). Or, au cours des vingt dernières années des études moléculaires ont permis de définir des gènes et des voies moléculaires impliquées dans ce processus pathologique complexe.

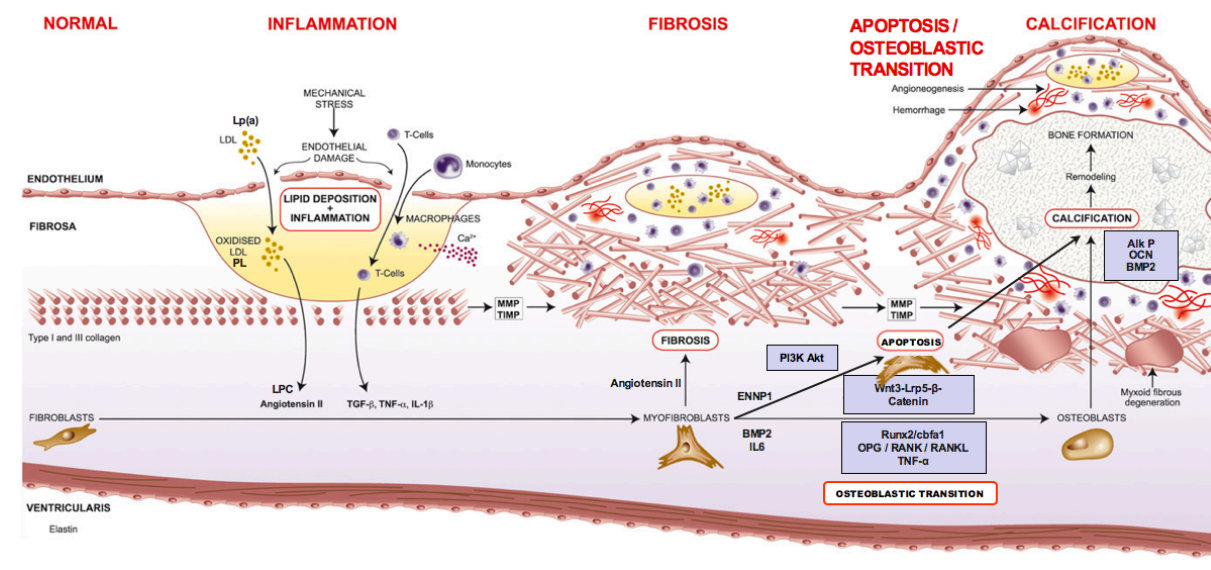


Figure 4 : Résumé des processus pathologiques survenant au cours de l'évolution de la sténose valvulaire aortique d'après Dweck et al.(17)

La potentielle interaction entre les cellules, les médiateurs moléculaires et les différentes voies de signalisation se déroule dans le compartiment sanguin et dans la région sous-endothéliale.

Tout d'abord (figure 4) une lésion ou une perturbation endothéliale peut permettre aux lipides de pénétrer dans l'endothélium valvulaire et s'accumuler dans les zones d'inflammation (18,19). Les lipoprotéines impliqués dans l'athérogenèse, notamment les lipoprotéines de basse densité et les lipoprotéines(a) sont présentes dans les lésions de la valve aortique précoce (18) et subissent ensuite une modification oxydative (19). Ces lipoprotéines oxydées sont hautement cytotoxiques et capable de stimuler une activité inflammatoire intense (20). La combinaison des lésions endothéliales et de dépôts lipidiques déclenche une inflammation de la valve. L'expression de molécules d'adhésion permet l'infiltration de la couche endothéliale par des monocytes qui se différencient en macrophages (21) et en cellules T. Les lymphocytes T activés libèrent des cytokines pro-inflammatoires, telles que le facteur de nécrose tumorale alpha (TNF alpha) (22), le facteur de croissance β1(TNF-β1) (23) et l'interleukine 1β (IL-1β).

L'interleukine IL-1 β stimule la production locale de métalloprotéines matricielles (MMP) ce qui contribue au remodelage de la matrice extracellulaire (MEC), les macrophages quant à eux expriment l'ostéopontine (OPN) une protéine associée à l'ossification. L'enzyme de conversion de l'angiotensine (ECA) qui est colocalisé avec l'apolipoprotéine B génère de l'angiotensine II (AgII) à partir de l'angiotensine I (AgI). Cette AgII stimule l'expression par les fibroblastes de protéoglycanes ayant une grande affinité pour les lipoprotéines ce qui piège les lipides dans le compartiment sous-endothélial.

Ce milieu athérogène régule aussi à la hausse l'expression de de la protéine morphogénique osseuse 2 (BMP2) et favorise les voies de signalisation ostéogéniques telles que les voies Runx2/Cbfa1 et les voies Wnt/Lrp5 et β caténine stimulant la différenciation des fibroblastes/myofibroblastes vers un phénotype ostéoblastique (figure 4). Ces cellules modifiées favorisent ainsi le développement de nodules de calcium et la formation osseuse (7).

4.2 VARIANT FREQUENT ET PATHOLOGIES FREQUENTES

➤ APPROCHE GENE CANDIDAT

Par cette approche un gène est considéré a priori comme candidat, c'est à dire potentiellement impliqué dans le phénotype d'intérêt. Il s'agit d'un gène connu codant pour une protéine suspecte d'être responsable du processus pathologique étudié. Plusieurs sources peuvent contribuer au choix des gènes candidats : les connaissances physiopathologiques, les modèles animaux, les analyses issues de l'analyse du transcriptome, la recherche in silico et les données issues du Projet Génome Humain (PGH).

L'hypothèse est ensuite testée en recherchant, par séquençage direct ou génotypage, des variants dans les séquences du gène.

Tableau 2 : Approche gène candidat dans la littérature

Article référence	Méthode	Protéine	Polymorphisme génique	Influence du polymorphisme génétique
Ortlepp JR. (2001) (24)	100 cas vs. 100 contrôles	Récepteur vitamine D	VDR BsmI allele B	Masse osseuse minérale
Avakian SD. (2001) (25)	62 cas vs. 62 contrôles	Apolipoprotéine B Apolipoprotéine E2	XbaI of ApoB	Taux total de cholestérol et LDL-c
Novaro GM. (2003) (26)	43 cas vs. 759 contrôles	Apolipoprotéine E	ApoE 4	Taux total de cholestérol et LDL-c
Nordstrom P. (2003) (27)	41 cas vs. 41 contrôles	→ Récepteur œstrogène α (ESR1) → Transforming Growth Factor (TGF β)	→918T/C(PvuII) →509C/T(AocI)	→Effet sur le serum lipidique →Formation et réparation tissu conjonctif
Ortlepp JR. (2004) (28)	187 valves aortiques sténosées prélevées	IL10 CCR5 CTCG	→1082 G/A →819 C/T →592 C/A →Déletion 32 pb →447 G/C	Inflammation

Ortlepp (24) qui a travaillé sur un variant génétique du récepteur de la vitamine D démontre dans son analyse d'association que la fréquence de l'allèle B est significativement plus importante dans le groupe patient atteint d'une maladie sténosante versus sans cette pathologie.

Avakian (25) quant à lui a trouvé une plus grande prévalence du polymorphisme de l'allèle de l'Apolipoprotéine B Xba1 et de l'Apolipoprotéine E2 chez des patients non diabétiques avec un RAC sévère isolé.

Dans une troisième étude (26) de la Cleveland Clinic, sur 759 patients ayant bénéficié d'une échographie trans-oesophagienne dans une démarche diagnostique classique 43 patients porteur d'un RAC ont été identifiés, les 716 autres patients constituant le groupe contrôle. Une plus grande prévalence de l'Apolipoprotéine E4 a été observée chez les patients porteurs d'un RAC versus groupe contrôle indépendamment du sexe, de l'âge et du taux de LDL cholestérol.

Dans une quatrième étude (27) Nordstrom s'est intéressé aux femmes ménopausées opérées d'un RAC, il a montré que le polymorphisme de PvuII du récepteur de l'œstrogène était indépendamment associé à un risque augmenté de sclérose valvulaire aortique. La combinaison des polymorphismes de PvuII et de AocI respectivement pour les gènes du récepteur de l'œstrogène et du TGF-1 étaient aussi associés à la sclérose aortique.

Ortlepp (28) trois ans après son travail sur le polymorphisme de gène du récepteur de la vitamine D, a travaillé sur 187 valves aortiques sténosées prélevées au bloc opératoire dans le cadre d'un remplacement valvulaire aortique classique. Il a finalement mis en évidence 3 polymorphismes (1082, 819 et 592) du promoteur de l'interleukine 10 significativement associés au degré de calcification mesuré par spectroscopie d'absorption atomique sur les valves prélevées. Il a aussi identifié dans ce travail une délétion de 32 paires de bases du gène récepteur de la chémokine 5 qui influence le degré de calcification. La limite de cette étude résidant dans la faible puissance (nombres de cas/contrôles).

➤ APPROCHE GWAS/TWAS

La combinaison de variants fréquents impliqués dans la survenue de pathologies fréquentes est étudiée depuis le milieu des années 2000 grâce au développement des puces à ADN permettant le génotypage d'un grand nombre de variant génétiques et au recrutement de cohortes importantes de patients.

Ces études d'associations sur génome entier (GWAS : Genome Wide Association) sont basées sur l'hypothèse que des variants géniques fréquents vont être transmis de génération en génération au sein de « blocs haplotypiques ». Ces hypothèses reposent sur les modèles de génétique de population, sur l'absence de résultats clairement significatifs des analyses de liaison pour les pathologies fréquentes et sur la difficulté d'identifier des variants à effet modeste par analyse de liaison (29). Initialement ces études se basaient en particulier sur les données du projet HapMap qui confèrent d'importantes informations sur les fréquences de variations génétiques, leur répartition et sur les groupes de SNPs en déséquilibre de liaison.

Le génotypage de variants décrits va permettre de comparer leurs fréquences alléliques dans une population de cas (patients non apparentés ayant un trait phénotypique étudié) à une population contrôle. Ainsi les différences de fréquences intergroupes sont comparées statistiquement à l'échelle du génome et permettent d'identifier des régions d'associations du génome avec le trait phénotypique étudié. L'hypothèse sous-jacente est qu'il existe, au sein de ces régions d'association, des gènes ou éléments régulateurs impliqués dans le développement de la pathologie.

L'approche TWAS (Transcriptome Wide Association Study) fonctionne de façon complémentaire et de façon dépendante à l'approche GWAS. En effet l'approche transcriptomique va permettre d'accéder à un niveau d'expression du gène en relation avec le SNP mis en évidence par l'étude GWAS pour un trait phénotypique donné.

Ces approches sont extrêmement puissantes de par leur étude sur le génome entier et sur des grandes cohortes de patients mais nécessitant un traitement bioinformatique complexe.

Tableau 3 : approche GWAS/TWAS dans la littérature

Article reference	Gene	Polymorphisme	Codant pour	Validation
Thanassoulis G. (2013) (19)	<i>LPA</i>	rs10455872	Lp(a)	GWAS 3 studies
Guauque-Olarte. (2015) (20)	<i>RUNX2</i>	rs114193529 rs144071310	Osteogenic transcription factor	GWAS
Helgadottir A. (2018) (21)	<i>PALMD</i> <i>TEX41</i>	rs7543130 rs1830321	Région régulatrice	GWAS
Thériault S. (2018) (22)	<i>PALMD</i>	rs6702619	Région régulatrice	TWAS

Thanassoulis et al (30) ont été les premiers à mettre en évidence un SNP dans un locus de la lipoprotéine A, Lp(a), qui atteint la significativité en GWAS (tableau 3)

Le deuxième travail (31) est une méta-analyse qui a été réalisée en combinant les résultats de deux approches GWAS avec 474 et 486 cas de l'équipe de Québec et Paris respectivement. Ce travail combiné a révélé deux SNPs localisés sur l'intron 1 de *RUNX2* qui étaient modérément associés à la sclérose aortique. *RUNX2* code pour un facteur de transcription ostéogénique, régulé par l'expression de *NOTCH1*.

Plus récemment, une grande étude (32) islandaise GWAS a identifié deux nouveaux loci sur le chromosome 1p21 proche de *PALMD* et sur le chromosome 2q22 in *TEX41*. Dans le même numéro de Mars 2018 de Nature Communication de l'équipe québécoise a confirmé l'implication de *PALMD* cette fois ci par une approche TWAS (33).

4.3 VARIANT RARE ET APPROCHE FAMILIALE

Une percée importante dans la génétique du RAC a été faite par la découverte de la voie de signalisation NOTCH par Garg et al (34). En effet, par l'analyse d'une famille atteinte de maladie valvulaire aortique, plus particulièrement d'une bicuspidie aortique, les auteurs ont identifié une liaison sur le chromosome 9q34-35. Après avoir étudié les gènes de la région, le travail s'est focalisé sur le gène *NOTCH1*, permettant l'identification de la mutation R1108X arrêtant la lecture du gène (codon stop). Tous les individus atteints de la famille étaient porteurs de l'allèle mutant et inversement absent chez les individus sains. Les mêmes auteurs ont identifié une seconde mutation *NOTCH1*(H1505del) qui ségrégeait avec la maladie valvulaire aortique dans une petite famille hispanique caractérisée par la présence d'une bicuspidie

aortique associée à une calcification valvulaire aortique. La voie NOTCH est connue pour être impliquée dans le développement embryonnaire et à partir de ces travaux associés au développement de la valve aortique. Ainsi la découverte de la voie NOTCH dans la maladie valvulaire aortique confirme le concept qu'une voie de développement embryonnaire peut être impliquée plus tard dans un processus pathologique. Ce travail fournit des preuves convaincantes de l'intérêt d'une approche familiale pour mettre en évidence des variants rares fonctionnels de la maladie valvulaire aortique ainsi que l'utilité de la recherche génétique pour identifier des nouvelles cibles moléculaires.

L'approche généalogique dans une population isolée (famille) a montré l'existence d'une part génétique dans les pathologies communes (35–37). Probst et al (6) sur ce même principe, ont évalué pour la première fois la contribution des facteurs familiaux dans le rétrécissement valvulaire aortique en réalisant une large investigation épidémiologique et généalogique pour identifier des grandes familles de RAC dans l'ouest de la France. A travers l'identification de cinq grandes familles de RAC, cette étude a montré pour la première fois que le RAC a une expression ou agrégation familiale fréquente. C'était le premier pas vers une possible identification d'un gène du RAC à travers une approche familiale associée au séquençage de nouvelle génération.

4.4 L'APPORT DU SEQUENÇAGE DE NOUVELLE GENERATION

La première génération de séquençage correspond au séquençage capillaire par la méthode Sanger, fiable mais de bas débit. L'apparition des technologies de séquençage de nouvelle génération (NGS : Next Generation Sequencing) a décuplé les capacités d'exploration du génome humain. Grâce à ces technologies, il est désormais possible de séquencer l'exome entier (WES : Whole Exome Sequencing), le génome entier (WGS : Whole Genome Sequencing), le transcriptome des patients. L'exome correspond à l'ensemble des régions codantes, soit 2% du génome. La quantité de données désormais accessibles a beaucoup augmenté. A l'inverse le prix ainsi que le temps de séquençage se sont considérablement réduits. Le WGS est particulièrement intéressant (en comparaison au WES) car il permet une couverture homogène de séquençage et ne nécessite pas de capture. En complément il permet d'explorer des régions non codantes, zones complexes à analyser mais qui représentent une partie de l'avenir de la génétique moderne.

Le NGS permet une approche pangénomique sans hypothèse physiopathologique et sans à priori sur les gènes candidats. Ces technologies aboutissent à la formation de « reads » qui sont ensuite alignés sur un génome de référence afin de détecter l'existence de variations génétiques. Chaque individu compte environ 4 millions de variants. Certains sont fréquents (dans la population générale ou dans une zone géographique donnée) d'autres sont rares.

L'association du WGS aux analyses génétiques familiales est un outil très puissant pour l'identification des variants causaux dans les familles. En effet lorsqu'une pathologie ségrège chez plusieurs individus d'une même famille, on pose l'hypothèse qu'une région du génome contenant une mutation est transmise de génération en génération. Un membre d'une famille partage un certain pourcentage de son génome en fonction de sa distance généalogique avec les autres membres de la famille : 50% avec un parent ou un frère/sœur, 25% avec un grand-parent ou oncle/tante, 12.5% avec un cousin germain. C'est pourquoi plus une famille est grande, plus elle sera informative pour identifier la région et le variant causal. Les analyses familiales sont ainsi réalisées sur des individus éloignés généalogiquement (cousins) afin de limiter le pourcentage de génome en commun et donc de variants partagés potentiellement causaux. A ce jour et malgré l'émergence du NGS, en dehors de Notch1 dans la bicuspidie valvulaire

aortique, aucun gène causal du RAC tricuspide n'a été identifié. La compréhension des facteurs favorisants pourra à terme permettre une prise en charge préventive de cette pathologie.

5) OBJECTIF DU MASTER

Le projet de recherche du master II s'inscrit dans la continuité des travaux réalisés au sein de l'équipe I de génétique cardiovasculaire de l'unité INSERM UMR 1087.

Grace à la constitution d'une bio-collection et au développement des outils de séquençage et de génotypage de nouvelle génération au sein du laboratoire, mon travail a consisté à rechercher un gène morbide de RAC au sein d'une grande famille de rétrécissement valvulaire aortique dite « précoce » de par la présence de sclérose valvulaire chez des sujets jeunes, famille précédemment identifiée et comportant plusieurs générations de sujets atteints.

L'objectif du master était d'étudier les caractéristiques phénotypiques et les déterminants génétiques et non-génétiques impliqués dans le développement d'une sclérose-sténose valvulaire aortique dans une grande famille avec peu de facteurs de risques cardiovasculaires.

MATÉRIELS ET MÉTHODES

1) ENQUÊTE FAMILIALE

Constitution d'un arbre généalogique permettant d'identifier 3 générations d'individus vivants et proposition de dépistage à tous ses membres pour identifier les individus porteurs ou non d'une sclérose valvulaire aortique et/ou sténose valvulaire aortique. Cette famille est issue de l'étude des grandes familles de RAC par l'Institut du Thorax à travers le recrutement chirurgical. Les malades sont classés avec un numéro d'anonymat issu du logiciel Integralis.

Ce projet a reçu l'accord du comité d'éthique de Nantes, et les patients ont signé un consentement éclairé pour leur inclusion et le prélèvement génétique.

2) RECUEIL DES DONNÉES

2.1 CLINIQUES

L'examen clinique des patients comprend un interrogatoire concernant les antécédents médicaux et un examen physique systématique, en recherchant en particulier des anomalies cardiovasculaires et des anomalies systémiques (cutanées, oculaires et squelettiques) qui pourraient être associées à une forme syndromique de pathologie valvulaire.

2.2 BIOLOGIQUES

Bilan lipidique complet (Cholestérol total, TG, HDL, LDL, Lp(a), ApoB), ainsi que d'autres paramètres biologiques (Hb, créatinine, PTH, Ph, Ca²⁺, marqueurs du métabolisme Ph-Ca²⁺)(38,39). Constitution d'une bibliothèque pour la réalisation de dosages subséquents.

2.3 ECHOCARDIOGRAPHIQUES

Le diagnostic phénotypique cardiaque des membres de la famille étudiée est basé sur un examen échocardiographique par un opérateur expérimenté avec analyse doppler/échographique (Vivid 9, GE MedicalSystems, Milwaukee, WI, USA) et le logiciel de mesures Echo PAC (GE Medical Systems, Milwaukee, WI, USA).

-L'évaluation de la valve aortique est faite selon les recommandations, par mesure des gradients moyen et maximum par voie apicale 5 cavités en doppler pulsé et continu(40) . La surface est estimée par l'équation de continuité à l'aide de la mesure de la chambre de chasse et de la mesure des ITV aortique et sous aortique. On précise la présence, le mécanisme et l'importance d'une insuffisance aortique associée.

-Le caractère bi ou tricuspide de la valve aortique est précisée par les coupes para sternale petit axe et sous costale lorsque cela est possible. L'aspect de la valve aortique est analysé de façon approfondi en évaluant la mobilité, l'épaississement et le degré de calcifications de chaque cusp.

-Les données sont reportées sur un diagramme avec un système de score de 0 à 27 points inspiré de celui utilisé dans la bicuspidie aortique par Michelena H et al (41).

-Mesure de la racine aortique et de l'aorte ascendante ainsi que celle de la crosse et de l'isthme aortique. Pour caractériser la forme de l'aorte nous avons calculé deux ratios avec la mesure de l'aorte ascendante (3 cm de la valve aortique) et la mesure de la jonction sino-tubulaire avec

comme dénominateur commun l'aorte initiale. Un rapport aorte ascendante/aorte initiale s'éloignant de 1 est en faveur d'un aspect fusiforme de l'aorte. De même un rapport jonction-sinotubulaire/aorte initiale s'éloignant de 1 est aussi en faveur d'une aorte fusiforme.

-La présence d'une valvulopathie mitrale associée (rétrécissement mitral ou insuffisance mitrale), ou de toute autre anomalie congénitale ou acquise cardiaque est également recherchée.

-L'existence de calcifications annulaires mitrale postérieure est recherchée .

-Le remodelage ventriculaire échographique est étudié par les mesures de diamètre, volume et masse ventriculaire gauche, effectuées selon les recommandations actuelles(42). La FEVG normale en Simpson Biplan est définie par une valeur $\geq 50\%$. La fonction diastolique est évaluée par le profil mitral E/A, et le rapport E/Ea, la taille de l'oreillette gauche et les PAPS (43).

- Les sujets apparentés sains > 35-40 ans servent de contrôle pour l'évaluation des anomalies valvulaires, en effet devant cette pathologie dégénérative le statut des individus < 35 ans reste indéterminé.

3) SCORE VALVULAIRE AORTIQUE PAR CT SCAN

Réalisation d'un scanner sans injection. Calcul du score calcique à la recherche de calcifications valvulaires. Analyse de la morphologie de l'aorte. Scanners réalisés au CHU de NANTES sur des appareils 64 barrettes (44,45). Deux acquisitions synchronisées à l'ECG se font à 40 et 70 % du cycle cardiaque. Les scores calciques sont calculés sur une console General Electric AW avec les séquences « Smart Score ».

La calcification est définie par deux pixels dont la densité est supérieure à 130 AU(arbitrary units). Le degré de calcification est mesuré semi automatiquement : les calcifications sont contourées manuellement par un opérateur entraîné puis le logiciel définit un score calcique. Le score calcique est également exprimé en volume (mm³) et en masse (mg) de calcification. Toutes ces valeurs sont calculées à 40% et 70 % du cycle cardiaque ; la valeur maximale est systématiquement retenue (46).

L'aspect de la valve aortique (tricuspide, bicuspide), de l'aorte, et le score calcique coronaire seront également analysés.

4) HISTOLOGIE

Chaque valve excisée au moment de la chirurgie a été fixée dans du formaldéhyde à 10% pendant 24 heures et analysée au service de pathologie. Après la fixation, des coupes de 3-5 nm d'épaisseur de la valve valvulaire aortique de chaque cas ont été coupées perpendiculairement à partir du bord annulaire jusqu'au bord du bord libre ; les sections ont été décalcifiées pendant 12 heures avant le traitement histologique. Des coupes de feuillet ont été obtenues en milieu de

bloc et colorées avec de l'hématoxyline-éosine. L'analyse histologique a porté (en binôme avec le Dr Toquet, anatomopathologiste) sur la calcification (présence et degré), la fibrose, les cellules inflammatoires, la néovascularisation et la métaplasie osseuse. L'échelle de sévérité était classée de 1 à 3 du moins sévère au plus sévère.

5) ANALYSES STATISTIQUES

Les variables continues sont présentées sous la forme de moyenne \pm DS ou médiane [intervalle interquartile] ; les variables catégorielles sous forme de fréquence et pourcentage. La comparaison des données entre les groupes (atteints vs sains) est faite à l'aide du test de t de Student ou du test de Mann-Whitney selon les besoins. Une valeur de $p < 0,05$ est considérée comme statistiquement significative.

6) INVESTIGATIONS GÉNÉTIQUES

6.1 BIO-COLLECTIONS

Les ADNs utilisés font partie de bio-collections gérées par l'équipe Bio-collection de l'Institut du thorax et par le Centre de Ressources Biologiques du CHU de Nantes. Les prélèvements sanguins, ont été réalisés chez les patients ayant signé un consentement éclairé et validé par un comité d'éthique.

6.2 SEQUENÇAGE GENOME ENTIER (WGS)

Les génomes ont été séquencés au Centre National de Recherche en Génomique Humaine. Le séquençage a été réalisé grâce à la technologie Illumina. Cette partie ayant été réalisée en amont de mon arrivée, le principe est brièvement décrit (figure 5).

La première étape consiste en la préparation de l'échantillon. L'ADN est extrait, purifié et fragmenté. Des adaptateurs sont apposés aux extrémités des fragments. L'ADN double brin est dénaturé en simple brin.

La deuxième étape est l'amplification en pont. L'extrémité des fragments simples brins se fixe aléatoirement à la surface de la flowcell (plaque en verre composée de deux types d'oligos complémentaires aux adaptateurs situés sur les fragments simples brins). Une polymérase crée le brin complémentaire du fragment hybridé. La molécule double brin obtenue est dénaturée, le fragment est dénaturé et le fragment initial est lavé. Le simple brin restant se pli et la région adaptatrice vient s'hybrider sur le deuxième type d'oligo de la flowcell. La polymérase génère le brin complémentaire, formant un pont double-brin. Le pont est dénaturé, résultant en deux fragments simples brins. Ce processus est répété simultanément pour des millions de clusters, résultant en l'amplification clonale de tous les fragments.

La troisième étape est le séquençage par synthèse. Le premier cycle de séquençage commence en ajoutant les amorces, l'ADN polymérase et quatre nucléotides marqués par des molécules fluorescentes (terminateurs réversibles marqués) ; ces nucléotides entrent en compétition pour se fixer sur la chaîne en extension. Un seul est incorporé en fonction de la séquence du fragment. Après chaque cycle, l'excitation des clusters par une source lumineuse permet l'émission d'un signal fluorescent caractéristique de la base assignée. La lecture est effectuée à chaque position sur toutes les séquences en parallèle (figure 5).

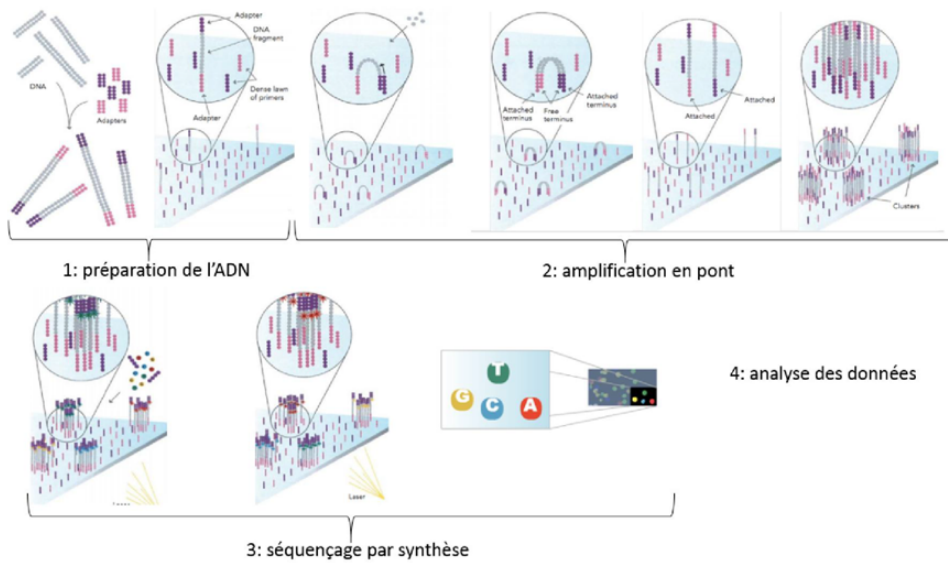


Figure 5 : Principe du séquençage Illumina

6.3 ANALYSE BIO-INFORMATIQUES

Le traitement des données est le principal défi qui accompagne le NGS.

Chaque fichier image généré lors du séquençage va être converti à chaque cycle en fichier BCL permettant d'identifier la base ajoutée au niveau de chaque cluster en y associant un score de qualité. Par la suite, ces fichiers vont être convertis en fichiers FASTQ grâce au logiciel CASAVA (logiciel du fournisseur Illumina). Ces fichiers FASTQ comportent des séquences de tous les « reads » séquencés et un score de qualité associé à chaque base (figure 6).

Chaque séquence obtenue « read » est ensuite alignée sur le génome de référence humain hg19/GRCh37. Cet alignement est facilité par le fait que les « reads » soient couplées deux à deux grâce à la méthode de séquençage en « paired-end ».

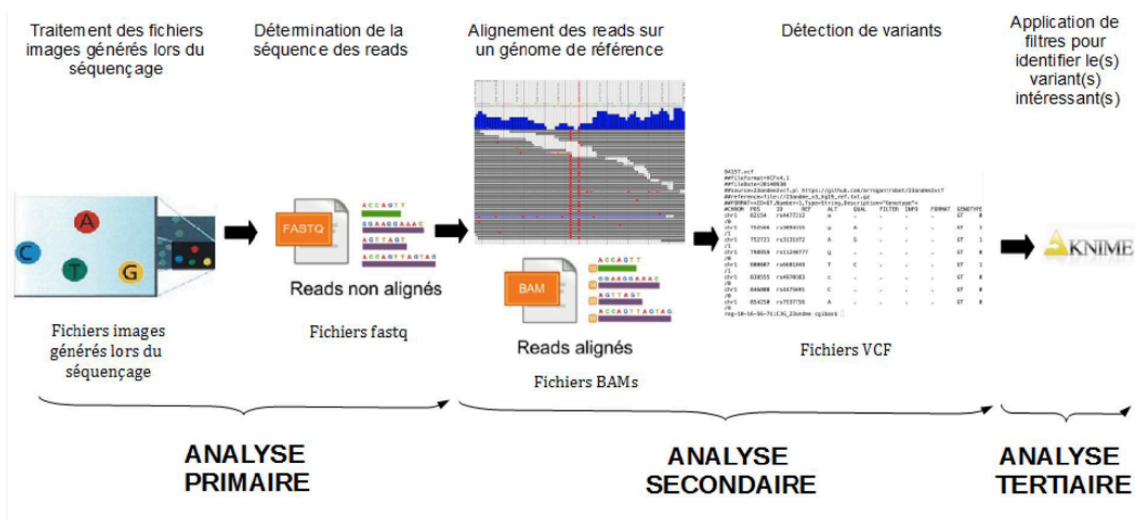


Figure 6 : Analyse bio-informatique des données de WGS

6.3.1 Alignement et annotation des séquences

➤ Alignement

L'outil BWA-MEM permet d'aligner chaque « read » sur le génome humain de référence (hg19/GRCh37) et va générer des fichiers SAM qui colligent de nombreuses informations relatives à tous les « reads » appariés (identifiant, séquence, scores de qualité, position sur le génome etc..). Ces fichiers SAM sont volumineux et vont donc être compressés en fichiers BAM (binaires). Les « reads » sont ensuite triés et hiérarchisés sur le génome. Les « reads » issus d'un même fragment d'ADN sont ensuite annotés avec Picard/MarkDuplicates.

➤ Détection de variants (« calling »)

La détection de variants a été réalisée grâce à un algorithme : GATK Unified Genotyper (version 2.8)(47). Le calling permet la genèse de fichiers VCF (« Variant Call Format ») (48). Les fichiers VCF sont des fichiers tabulaires listant l'ensemble des variations détectées par rapport à la séquence de référence du génome humain pour chaque patient séquencé. Chaque ligne correspond à un variant et chaque variant est annoté avec de nombreuses informations relatives à ce variant. On retrouve la position génomique du variant, son identifiant « rs » (si ce variant est référencé dans la base de données dbSNP (NCBI <http://www.ncbi.nlm.nih.gov/>), l'allèle de référence, l'allèle alternatif observé, l'état du variant (hétéro/homozygote) et une grande quantité d'informations relatives à celui-ci (score de conservation, fréquence allélique etc...) regroupées dans une même colonne.

Au terme du séquençage et de l'alignement des « reads », un contrôle permettant d'estimer la qualité des données et du séquençage est réalisé. Deux principaux paramètres sont évalués :

-la couverture de séquençage (pourcentage de bases couvertes)

-la profondeur de séquençage (nombre moyen de « reads » par base nucléotidique).

On peut ainsi estimer quelle partie du génome on peut sereinement analyser et identifier des régions mal couvertes.

➤ Annotation des variants

Les variants sont ensuite annotés en utilisant l'algorithme VEP (« Variant Effect Predictor »). Cette annotation, basée sur la base de données Ensembl (49), permet de connaître l'effet prédit du variant nucléotidique sur l'ensemble des transcrits référencés.

L'utilisation de bases de données publiques contenant des données de séquençage de nombreux individus contrôles et patients atteints de pathologies diverses de différentes populations, va permettre d'estimer la fréquence des variants recherchés.

Les bases de données utilisées sont :

- La base de données GnomAD (Genome Aggregation Database) regroupe des données de séquençage d'exome et de génome provenant d'un grand nombre de projets de séquençage à grande échelle (123 136 exomes et 15496 génomes).
- La base de données Genmed regroupe les génomes de 856 personnes de l'ouest de la France.
- La base de données Troucard regroupe les génomes de 43 patients avec une atteinte cardiovasculaire ou autre, étudiés au laboratoire.

De plus un score de conservation inter-espèces est attribué à chaque variant nucléotidique. Le score de conservation que nous avons utilisé est le score GERP (« Genomic Evolutionary Rate Profiling) compris entre -12.3 et 6.17 à partir de comparaisons basées sur les génomes de 29 espèces de mammifères (50).

6.4 SEQUENÇAGE CAPILLAIRE

6.4.1 PCR

La validation des variants et les analyses de coségrégation ont été réalisées par séquençage capillaire (méthode de Sanger). Les PCR sont réalisées avec la Taq DNA polymérase (Life Technologies®). La réaction de polymérisation en chaîne est une technique permettant d'amplifier *in vitro* une séquence cible d'ADN à l'aide d'amorces nucléotidiques (Primers) complémentaires et spécifiques des séquences flanquantes extérieures aux extrémités de la cible. Il s'agit d'une réaction en chaîne, les brins d'ADN néoformés étant utilisés comme matrice pour la synthèse d'ADN lors du cycle suivant.

Chaque couple de primers (forward et reverse) a été dessiné sur Primer 3+ à partir de séquence sélectionnée sur UCSC.com (hg18). Une vérification par « PCR *in silico* » sur le site UCSC, permet de sélectionner des couples de primers (T_m proches à moins de 1°C, couple spécifique de l'amplicon désiré, pas de mésappariements, pourcentage en GC de 50±5%)

Le mélange réactionnel est le suivant: (1) 50 ng d'ADN; (2) le tampon est ajusté à 1x (Invitrogen); (3) le MgCl₂ (à 1.5mM); (4) les dNTP à 200µM; (5) les amorces sens et antisens à 0.5µM; (6) 0.75 U de Taq DNA polymérase (Invitrogen); (7) dans un volume réactionnel de 20 µl.

La réaction a lieu dans un thermocycleur automatique GeneAmp PCR System 9700 (Applied Biosystems) selon le protocole suivant : (1) dénaturation initiale de l'ADN pendant 3 min à 94°C, (2) 30 cycles comprenant une dénaturation de l'ADN pendant 30 sec à 94°C, une hybridation des amorces pendant 30 sec à 55°C et une synthèse pendant 30 s à 72°C ; (3) extension finale de 10 min à 72°C ; et (4) température finale de 4°C.

La vérification des PCR est réalisée sur gel d'agarose 1% avec du BET. La migration est effectuée dans TpTAE pendant 30 min à 10 mV après dépôt de 5 µl de produit de PCR (ADN) et avec 2 µl de bleu de charge. Visualisation du gel sous UV.

6.4.2 Séquençage

Les produits de PCR afin d'éliminer les amorces et nucléotides restants par réaction enzymatique Exosap (Amersham Biosciences) selon protocole du fournisseur : 37°C pendant 15 minutes puis inactivation 80°C pendant 15 minutes.

Pour la réaction de séquence. La méthode utilisée est le séquençage par terminaison de chaîne (méthode de Sanger) avec le « Big Dye Terminator v3.1 Cycle Sequencing Kit » (Applied Biosystems) selon les instructions du fournisseur. La réaction de séquence a lieu en présence de 0.5 µl de BigDye Terminator (Applied Biosystems) et 0.25 µl d'amorce sens ou antisens (20 µM).

La réaction de séquence requiert une dénaturation de 1 minute à 96°C, puis d'une amplification de 25 cycles (comprenant 10 secondes de dénaturation à 96°C, 5 secondes d'appariement à 50°C et 2 minutes d'élongation à 60°C). Les produits d'élongation sont purifiés avec le « X Terminator Solution » (Applied Biosystems) selon instructions du fournisseur. Migration du produit de séquençage sur séquenceur capillaire AB 3730 DNA Analyzer (Applied Biosystems).

Enfin les séquences obtenues sont analysées avec le logiciel SeqScape v2.5 (Applied Biosystems). On peut ainsi déterminer si un individu est porteur de la variation génétique et si c'est le cas, à l'état homo ou hétérozygote.

6.5 GENOTYPAGE HAUT DEBIT

Le génotypage haut débit a été effectué sur les puces Axiom PMRA (Precision Medicine Research Array) commercialisées par Affymetrix (Santa Clara, CA, USA) par la plateforme de génomique de Nantes (<http://www.pf-genomique.univ-nantes.fr>).

Les puces Axiom PMRA permettent le génotypage de 865 720 variants fréquents (SNPs) sur l'ensemble du génome grâce à des puces composées de 1,2 Millions de sondes.

Les signaux fluorescents ont été mesurés grâce au « Gene Titan Multi-channel » (Affymetrix) et l'analyse primaire a été réalisée à partir de l'outil « Affymetrix Power Tool » en suivant les recommandations du fournisseur.

Les échantillons ayant un pourcentage de génotypage inférieur à 97 % ont été éliminés de l'analyse ainsi que ceux présentant un taux d'hétérozygotie anormalement différent de la valeur attendue. Les SNP présentant une fréquence de l'allèle mineur (MAF) inférieure à 10 % ont été exclus tout comme ceux ayant un taux de génotypage inférieur à 95 % ou étant en déséquilibre de Hardy-Weinberg ($P < 10^{-5}$).

6.6 IDENTITE PAR DESCENDANCE (IBD)

Les analyses d'identité par descendance ont pour but d'identifier des blocs haplotypiques partagés par des paires d'individus apparentés hérités d'un ancêtre commun sans événement de recombinaison. Ces analyses ont été réalisées avec l'algorithme « Identity By Descent Linkage Disequilibrium » (51) qui permet le calcul de probabilités d'identité par descendance (IBD) à partir de données de génotypage haut-débit décrites ci-dessus sans incidence de la taille ou de la complexité de la famille étudiée en tenant compte du déséquilibre de liaison. Les analyses ont été réalisées pour identifier des régions partagées par les individus atteints de la famille (paires concordantes) et en comparant les individus sains et atteints d'une famille (paires discordantes). Dans le cadre de nos analyses familiales, nous recherchons des zones haplotypiques partagées entre individus d'une même famille permettant de focaliser la recherche de variants sur le génome. Chaque marqueur génotypé (SNP), est comparé en prenant en considération le déséquilibre de liaison au sein du bloc haplotypique, et le nombre de paires d'individus partageant ces SNPs est ainsi attribué. Les données sont représentées sous forme de graphique représentant en abscisse la position chromosomique et en ordonnée le nombre de paires d'individus partageant ce SNP. Si une région haplotypique est partagée par toutes les paires concordantes, mais aussi par les paires discordantes, la région ne semblera pas intéressante vis-à-vis du trait phénotypique étudié (dans le cadre d'une pathologie à pénétrance complète).

L'ensemble de ces analyses a été réalisé au sein de l'équipe « Génétique et maladies cardiovasculaires » avec l'aide des biostatisticiens (Sidwell Rigade et Christian Dina).

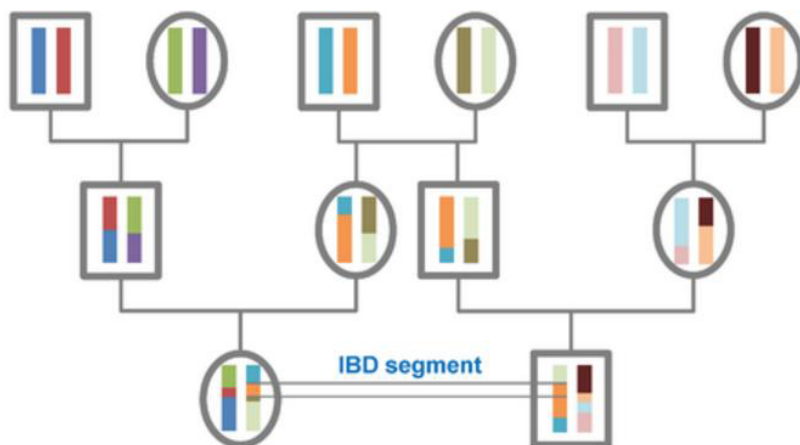


Figure 7 : Principe des IBD

La région orange sur cette figure correspond à une région IBD (partagée) entre deux cousins

7) SÉLECTION DES VARIANTS APRÈS SÉQUENÇAGE GÉNOME ENTIER

7.1 FILTRE DES VARIANTS IDENTIFIÉS À PARTIR DE L'OUTIL KNIME4BIO

Suite aux étapes successives de séquençage, l'analyse génétique des fichiers VCFs est réalisée grâce à l'outil Knime4bio développé au laboratoire (52). Cet outil permet de contourner la difficulté du code informatique pour analyser des fichiers contenant de nombreuses informations ne pouvant être gérées par des tableurs classiques. L'ajout de commandes est visualisé à l'aide d'une interface et permet d'appliquer successivement des actions de formatage et de filtrage présentées dans la figure 8. L'enchaînement des commandes est appelé « workflow ». Chaque filtre possède la particularité d'éliminer certains types de variants afin de ne sélectionner que le/les variants causaux parmi les 120 000 variants retrouvés en moyenne par patient.

7.2 FILTRE DES VARIANTS FONCTIONNELS

Dans un premier temps, le VCF généré par GATK est traité et introduit dans le workflow. L'étape suivante consiste à conserver les variants ayant un effet fonctionnel potentiel sur la fonction de la protéine sur les bases des transcrits ENSEMBL et des annotations VEP (53). Nous gardons ainsi les variants ayant des annotations non-synonymes et/ou impliqués dans les sites d'épissage des catégories suivantes grâce à la base de données sequence ontology (SO) (www.sequenceontology.org) : splice_donor_variant, splice_acceptor_variant, stop_gained, frameshift_variant, stop_lost, initiator_codon_variant, inframe_insertion, inframe_deletion, missense_variant, transcript_amplification, splice_region_variant, transcript_ablation, incomplete_terminal_codon_variant, coding_sequence_variant.

7.3 CONSERVATION DES VARIANTS RARES

Les filtres suivants utilisent différentes bases de données. Chacune de ces bases est constituée de variants identifiés dans des exomes ou génomes de populations contrôles afin de répertorier les fréquences alléliques de l'ensemble des variants des régions codantes du génome humain.

Pour chaque variant du fichier VCF de chaque individu, le chromosome, la position génétique, l'allèle de référence et l'allèle alternatif sont comparés à l'ensemble des variants référencés dans les bases de données. Le seuil maximal de la fréquence allélique est considéré comme rare (fréquence allélique <0.1%) et potentiellement impliqué dans la pathologie.

Un autre nœud filtre les variants partagés par tous les individus séquencés. Enfin le dernier filtre ne conserve que les variants identifiés seulement dans notre famille afin d'éliminer les faux positifs.

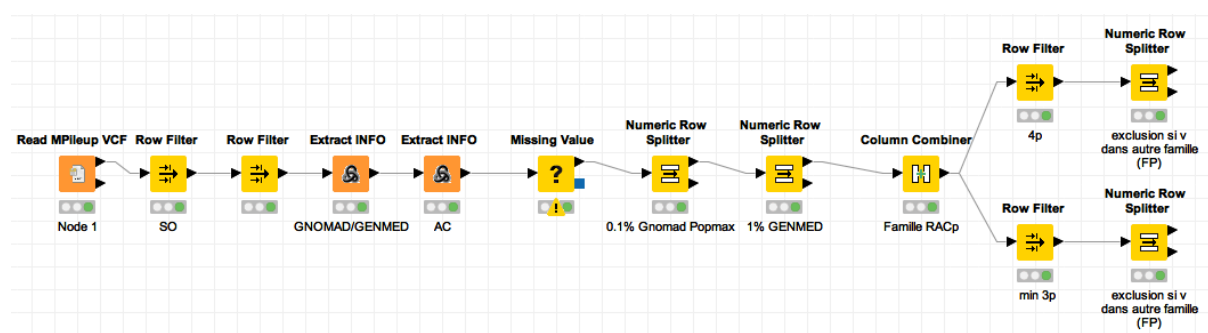


Figure 8 : Visualisation du « workflow » d'analyse réalisé avec l'outil Knime4bio afin de filtrer les fichiers VCF

Les bases de données mentionnées sont détaillées en 5.3.1.

RESULTATS

1) ENQUÊTE FAMILIALE : ARBRE GÉNÉALOGIQUE

Le recrutement de la famille a débuté suite à l'identification de trois frères ainsi qu'un cousin germain atteints d'un rétrécissement valvulaire aortique calcifié (RAC tricuspide confirmé chez 3 patients, 1 non précisé dans les compte-rendus) ou maladie de Monckeberg opérés d'un remplacement valvulaire aortique (âge d'intervention 60, 63, 70 et 79 ans).

Le recrutement familial a permis d'identifier plus de 100 individus (figure 9) sur 5 générations dont 38 individus examinés à ce jour, parmi ces 38 individus recueillis 15 ont été identifiés comme atteints ou probablement atteints et 23 sains suite à la consultation.

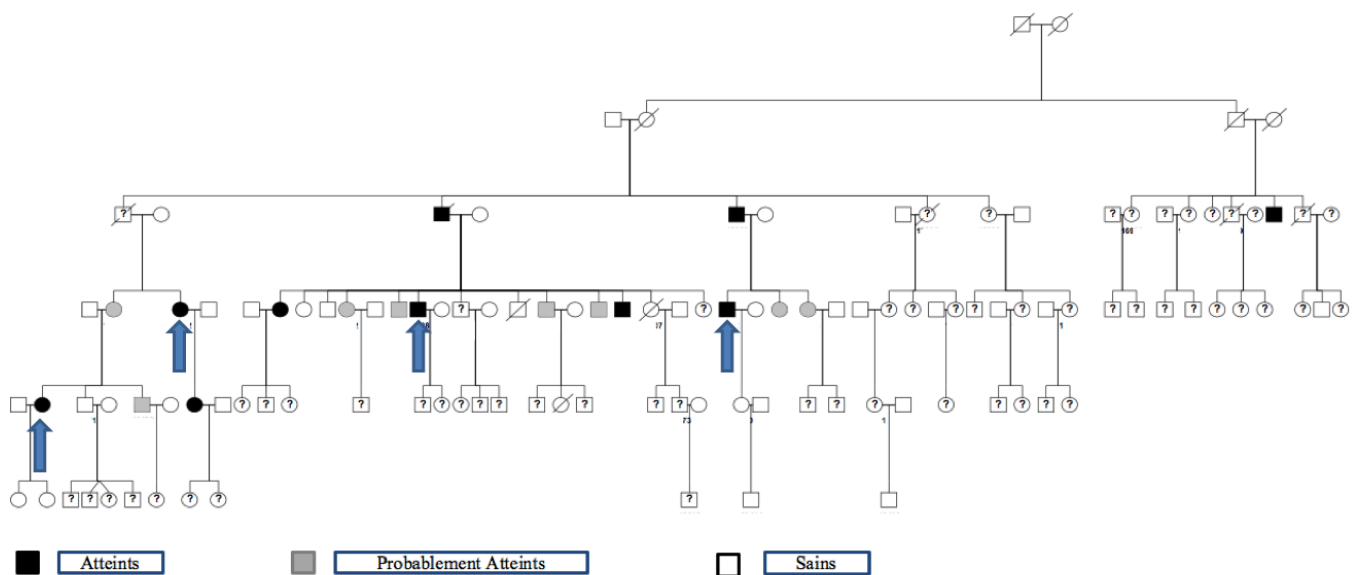


Figure 9 : arbre généalogique complet

Les individus fléchés sont les individus ayant bénéficiés du WGS.

2) CLINIQUE ET IMAGERIE :

Cette grande famille de RAC a la particularité unique de développer de façon précoce (36 ans pour le plus jeune sujet examiné à ce jour) (figure 10) une sclérose valvulaire aortique sur valves tricuspides, qui paraît évoluer vers une sténose calcifiée classique chez des sujets plus âgés. Nous avons ainsi nommé cette famille « RAC précoce ».

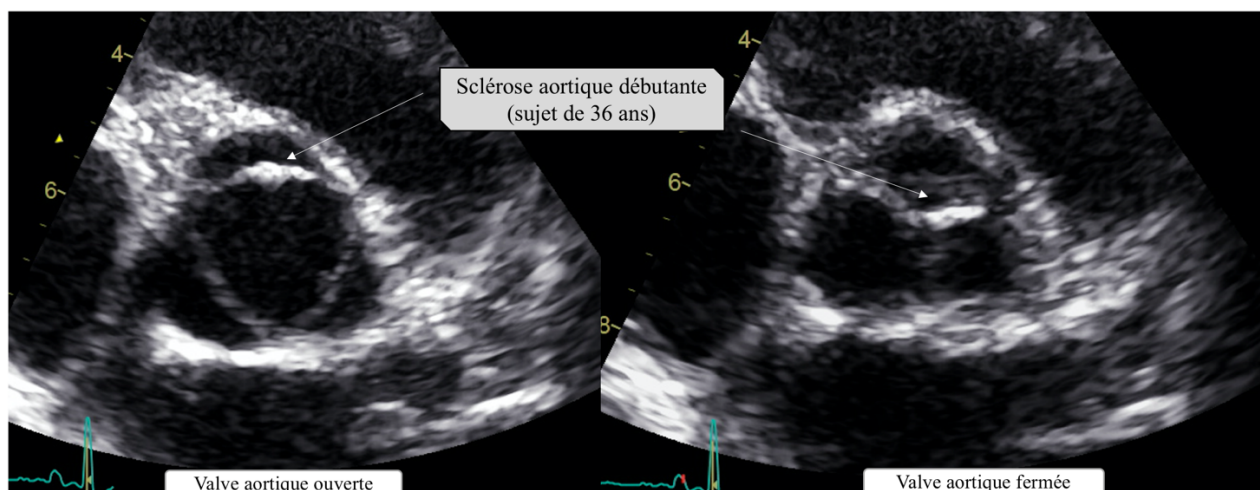


Figure 10 : Coupes échographiques parasternale petit axe d'analyse d'ouverture de la valve aortique chez une patiente de 36 ans présentant également une fuite aortique centrale de grade 1

Les données clinico-biologiques et les facteurs de risques cardiovasculaires des individus de la famille sont présentées dans le tableau 3 en comparant les individus atteints aux individus sains.

Les populations de malades atteints et sains sont relativement comparables sur le plan des facteurs de risques cardiovasculaires et des données clinico-biologiques.

A noter, les individus atteints sont en moyenne plus âgés ($62,4 \pm 12,7$ vs. $47,1 \pm 15,0$; $p=0,002$) confirmant le caractère dégénératif de cette pathologie. Le taux de cholestérol total et de LDL cholestérol est significativement plus élevé dans le groupe atteint tout en étant dans la norme internationale. Le taux de HDL protecteur est aussi plus élevé dans le groupe atteint. Le taux de Lp(a) est normal et comparable entre les groupes.

Tableau 3 : Caractéristiques cliniques et biologiques des individus de la famille

Variable	Tous (n=38)	Atteints (n=15)	Sains (n=23)	p
Age, ans	53,1±15,8	62,4±12,7	47,1±15,0	0,002
Homme, n (%)	19 (50%)	5 (33,3%)	14 (60,9%)	0,1
Surface Corporelle, m²	1,8±0,2	1,7±0,2	1,8±0,2	0,36
HTA, n (%)	8 (23,7%)	5 (33,3%)	3 (13,0%)	0,27
Diabète	0	0	0	-
Dyslipidémie, n (%)	7 (18,4%)	3 (20%)	4 (17,4%)	0,84
Hérédité	0	0	0	-
Tabac, n (%)	19 (50%)	5 (33,3%)	14 (60,9%)	0,1
Tabac sevré, n (%)	12 (31,6%)	5 (33,3%)	7 (30,4%)	0,86
Lp(a), g/L	0,53±0,36	0,52±0,51	0,54±0,49	0,92
PTH, pg/mL	38,1±4,2	40,4±1,1	36,7±5,13	0,53
Ca²⁺, mmol/L	2,43±0,15	2,37±0,13	2,59±0,07	0,55
Cholestérol Total, mmol/L	6,48±1,1	6,54±1,4	6,31±1,2	0,02
HDL, mmol/L	1,5±0,4	1,6±0,5	1,45±0,2	0,02
LDL, mmol/L	4,0±1,1	3,97±1,3	3,91±0,2	0,048
Apo A1, g/L	1,7±0,2	1,7±0,3	1,7±0,2	0,23
Apo B, g/L	1,3±0,3	1,3±0,3	1,3±0,1	0,57
NT-ProBNP, ng/L	80±78,2	108,8±105	63,3±54,1	0,13

HTA, Hypertension artérielle / Lp(a), Lipoprotéine(a) / PTH, parathyroid hormone/ Ca²⁺, calcémie / HDL, High Density Lipoprotein / LDL, Low Density Lipoprotein / Apo A1, Apolipoprotéine A1 / Apo B, Apolipoprotéine B / NT-ProBNP, N-terminal pro-brain natriuretic peptide

L'examen échocardiographique (ETT) (tableau 4) a révélé un phénotype héréditaire unique de sclérose/sténose valvulaire aortique calcifié associé à une aorte fusiforme (figure 1 annexe) chez 9 individus. Toutes les valves aortiques étaient tricuspides, avec une sclérose chez 14 individus (36,8%) et une sténose chez 2 (5,3%). Une patiente ultra-octogénaire a été classée saine malgré la présence d'une discrète sclérose à l'ETT, normale pour son âge avancé.

Sur les 15 individus atteints, 6 présentaient une insuffisance aortique ($p = 0,001$). L'aorte ascendante initiale était de plus grand diamètre chez les patients atteints ($34,2 \pm 4,3$ vs $27,1 \pm 4,6$ mm ; $p = 0,02$).

Un rapport du diamètre de la jonction sinotubulaire/diamètre initial de l'aorte ascendante a été calculé en échocardiographie et était plus faible chez les individus atteints ($0,88 \pm 0,08$ vs $0,94 \pm 0,05$, $p = 0,007$), ce qui confirme la présence d'un phénotype unique d'aorte fusiforme (annexe 1).

Tableau 4 : caractéristiques échocardiographiques des individus de la famille

Variable	Tous (n=38)	Atteints (n=15)	Sains (n=23)	p
SC, m ²	1,8±0,2	1,7±0,2	1,8±0,2	0,36
CCVG, mm	25,0±2,5	25,0±3,0	22,7±2,2	0,73
Sinus Valsalva, mm	30,7±6,9	32,6±4,8	29,3±7,8	0,17
Jonction Sino-tubulaire (JST), mm	27,9±4,5	28,9±5,5	27,2±3,6	0,20
Aorte initiale, mm	30,1±5,6	34,2±4,3	27,1±4,6	0,02
Aorte ascendante 3 cm, mm	37,0±6,2	38,8±4,4	34,3±8,1	0,07
Ratio Ao Asc/Ao initiale	1,12±0,12	1,12±0,12	1,12±0,12	0,97
Ratio JST/Ao initiale	0,91±0,07	0,88±0,08	0,94±0,05	0,007
Aorte fusiforme, n (%)	9 (23,7%)	7 (46,6%)	2 (8,7%)	0,02
Insuffisance aortique ≥ 1, n (%)	6(17,6%)	6(40,0%)	0	0,001
Tricuspidé	100%	100%	100%	—
Sclérose, n (%)	14 (36,8%)	13 (86,7%)	1 (4,3%)	<0,0001
Sténose, n (%)	2 (5,3%)	2 (13,3%)	0	0,08

Suite aux investigations cliniques et échocardiographiques nous avons sélectionné quatre individus atteints pour le WGS dont nous aborderons les résultats dans une partie propre. Devant la nécessité d'individus éloignés généalogiquement, afin de limiter le pourcentage de génome en commun et donc de variants partagés, potentiellement causaux, nous avons choisi

trois cousins et une jeune patiente de 36 ans qui est leur nièce (tableau 5) (flèches bleues figure 9).

Tableau 5 : individus sélectionnés pour le WGS

Individu	Phénotype	Type d'atteinte	Valve	Age inclusion
17008	Atteint	Sténose modérée	Tricuspide	67
17002	Atteint	Sténose modérée	Tricuspide	63
16952	Atteint	Sclérose avancée	Tricuspide	48
18470	Atteint	Sclérose débutante	Tricuspide	36

Tableau 6 : caractéristiques échocardiographiques et scannographiques des individus WGS

Variable	18470	17002	16952	17008
Age, ans	36	63	48	68
Sexe	Femme	Femme	Homme	Homme
SC, m ²	1,5	1,6	2,0	2,0
CCVG, mm	22,3	20,4	22,5	22,0
Sinus Valsalva, mm	31	28	35	35
JST, mm	29	22	30	34
Aorte initiale, mm	30,9	24,6	35,6	37,3
Aorte Ascendante, mm	32	30,7	40,7	45
Ratio Ao Asc/Ao initiale	1,04	1,25	1,14	1,21
Ratio JST/Ao initiale	0,94	0,89	0,84	0,91
Aorte fusiforme	0	1	1	1
Insuffisance aortique, grade	1	trace	trace	1
Tricuspide	1	1	1	1
Sclérose	1	0	1	1
Sténose	0	1	0	0
Score calcique valvulaire Ao (AU)	0	645	261	252

Nous avons ensuite présenté dans le tableau 6 les données échocardiographiques de ces individus sélectionnés pour le WGS.

Enfin nous avons comparé ces 4 individus WGS au reste de la population atteinte. Les individus atteints choisis pour le WGS sont représentatifs de la population d'individus atteints. En effet les données échocardiographiques étaient comparables, seul le score calcique valvulaire aortique était significativement plus élevé dans le groupe WGS, s'expliquant par le fait que nous avons sélectionné les individus les plus atteints.

3) SCORE CALCIQUE VALVULAIRE AORTIQUE

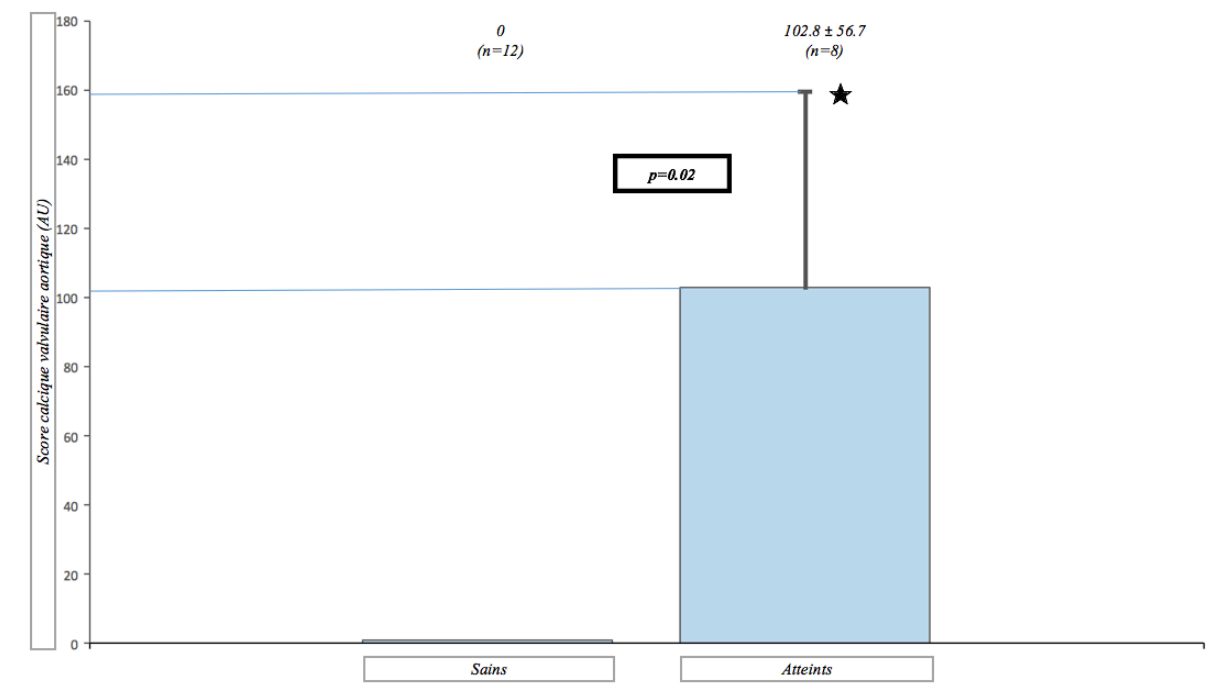


Figure 11 : Histogramme du score calcique valvulaire aortique des individus sains vs atteints

Les valeurs données sont la moyenne ± l'erreur standard à la moyenne (ESM).

Les individus de moins de 35 ans ne bénéficiaient pas de scanner car sans intérêt devant le statut indéterminé de leur atteinte dans le contexte de cette pathologie dégénérative. Au total 8 patients atteints (53.3%) et 12 sains (52.2%) ont eu un score calcique. Le score calcique valvulaire aortique était significativement supérieur dans le groupe atteint malgré le faible n (102.8 ± 196.4 vs 0 AU, $p= 0,02$). Le score calcique coronaire était comparable entre les groupes (84.3 ± 169.5 vs 122.8 ± 294.6 AU, $p= 0,75$).

4) ANATOMOPATHOLOGIE

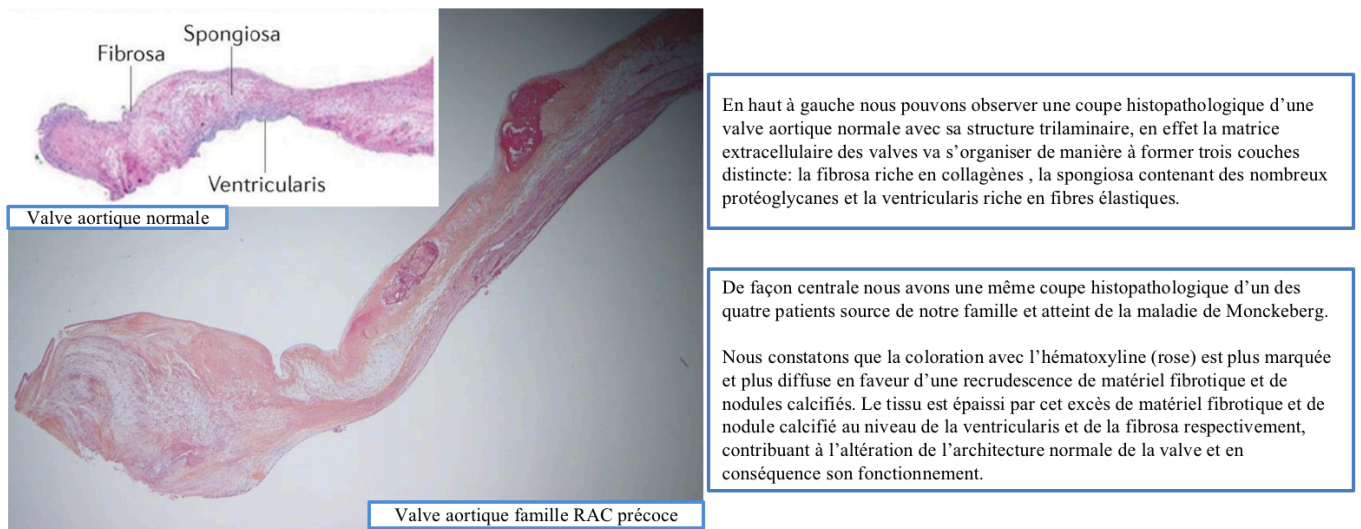


Figure 12 : coupes histologiques d'une valve aortique normale et d'une valve aortique d'un individu opéré de la famille RAC précoce

Nous avons dans un premier temps repris les coupes histologiques des valves des 4 patients sources opérés au CHU de Nantes pour les analyser selon les critères cités dans matériels et méthodes (figure 12).

Tout d'abord les données opératoires et l'analyse macroscopique. Pour rappel l'âge opératoire des 4 individus était de 60, 63, 70 et 79 ans. Macroscopiquement les valves étaient tricuspides pour 3 d'entre elles, le profil d'une valve reste indéterminé car pas de description de l'anatomopathologiste et du chirurgien, une seule valvule ayant été conservée il était difficile pour nous de conclure. Macroscopiquement ces valves étaient très calcifiées confirmant la maladie de Monckeberg.

Puis les données microscopiques (tableau 7), que nous avons analysées en double contrôle selon les critères précisés dans la partie matériels et méthodes (avec une échelle de sévérité de 1 à 3) (figure 13).

Tableau 7 : Analyse microscopique des 4 valves aortiques de la famille RAC précoce

Critères	17004	17001	16950	16988
Calcification	1 à 3	2	3	1 à 2
Localisation calcification	multifocale	multifocale	diffuse	multifocale
Métaplasie osseuse	0	0	1	0,5
Lipides	1	2	1	1
Spongiosa	normale	normale	normale	normale
Fibrosa	lipides +++	lipides +++	Peu visible	lipides +++
Ventricularis	épaissie, fibreuse	épaissie en focal, fibreuse	Peu visible	épaissie, fibreuse
Infiltrat cellulaire	1	0.5	2	1
Type	macrophages, rares PNE	lymphocytes	lymphocytes, plasmocytes, macrophages	lymphocytes, macrophages
Ulcération	0	0	0,5	0
Néo-vaisseaux	1	0	1	0

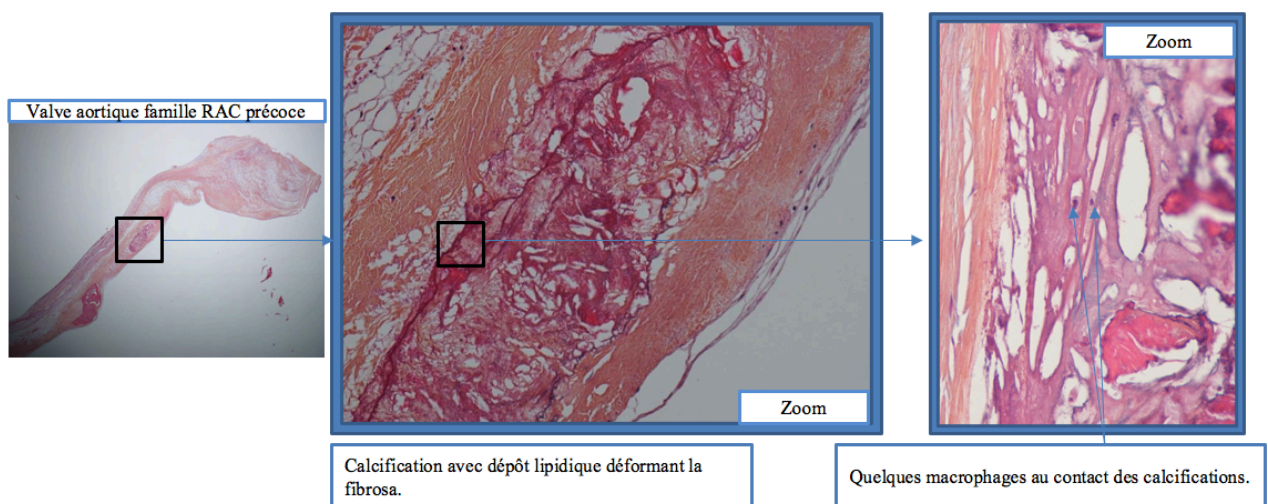
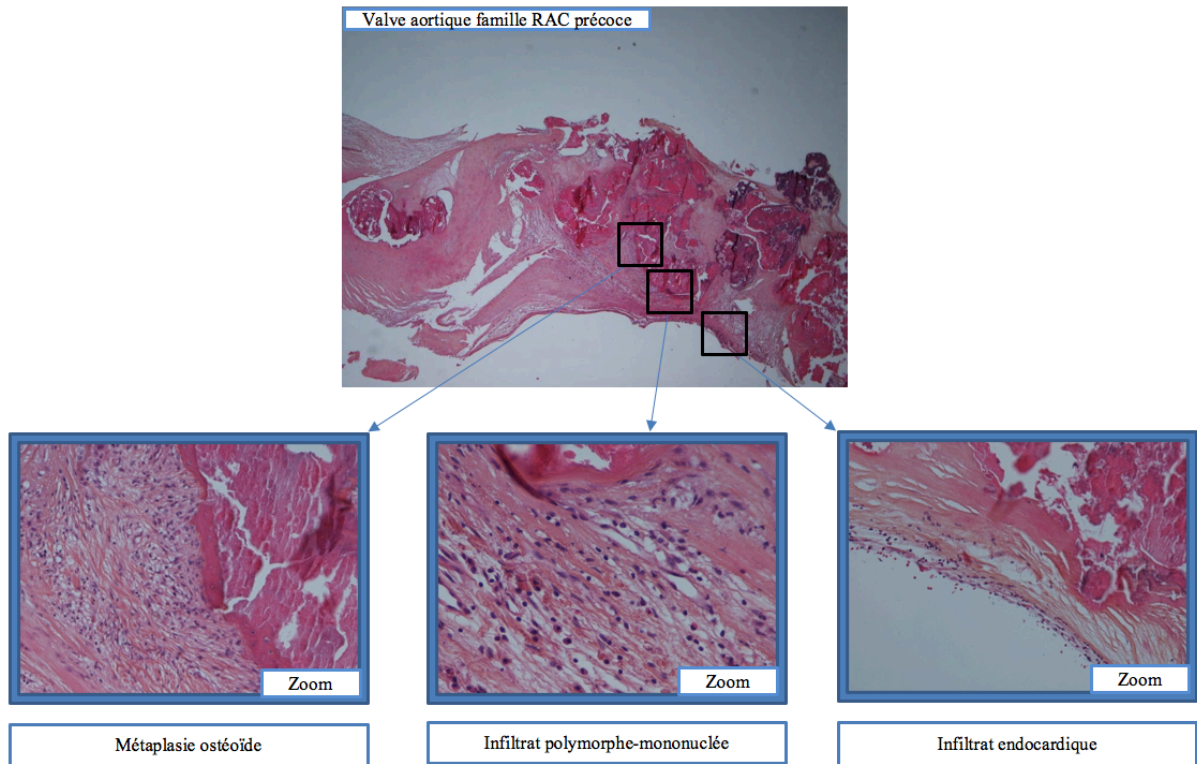


Figure 13 : Eléments clés des coupes histologiques des valves aortiques de la famille RAC précoce



Nous avons enfin en collaboration avec le bloc opératoire fait prélever 5 valves aortiques de patients atteints d'un rétrécissement valvulaire aortique au stade opératoire pour avoir un groupe contrôle de notre famille.

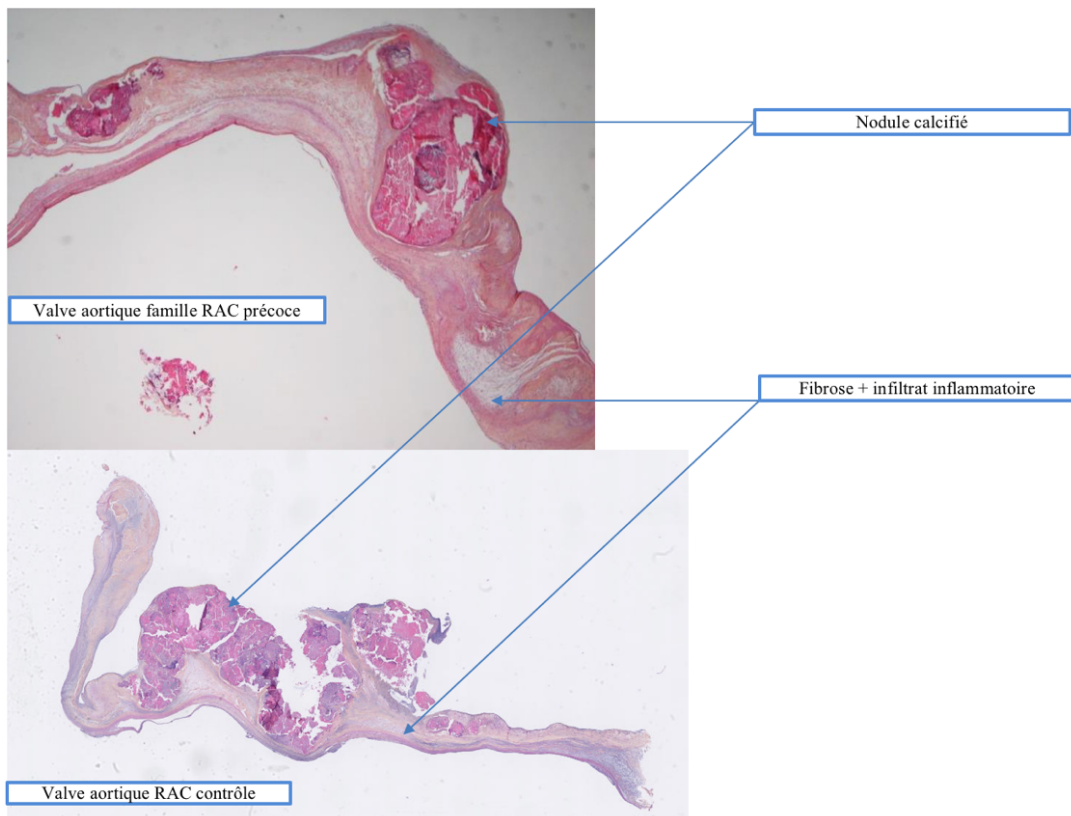


Figure 11 : Comparaison valve aortique de notre famille versus un RAC classique contrôle

Nous n'avons pas trouvé de différences significatives histologiques pour les différents critères étudiés entre notre famille de RAC « précoce » et les RAC contrôles (figure 14).

5) INVESTIGATIONS GÉNÉTIQUES

5.1 REGIONS GENOMIQUES PARTAGEES PAR LES INDIVIDUS ATTEINTS

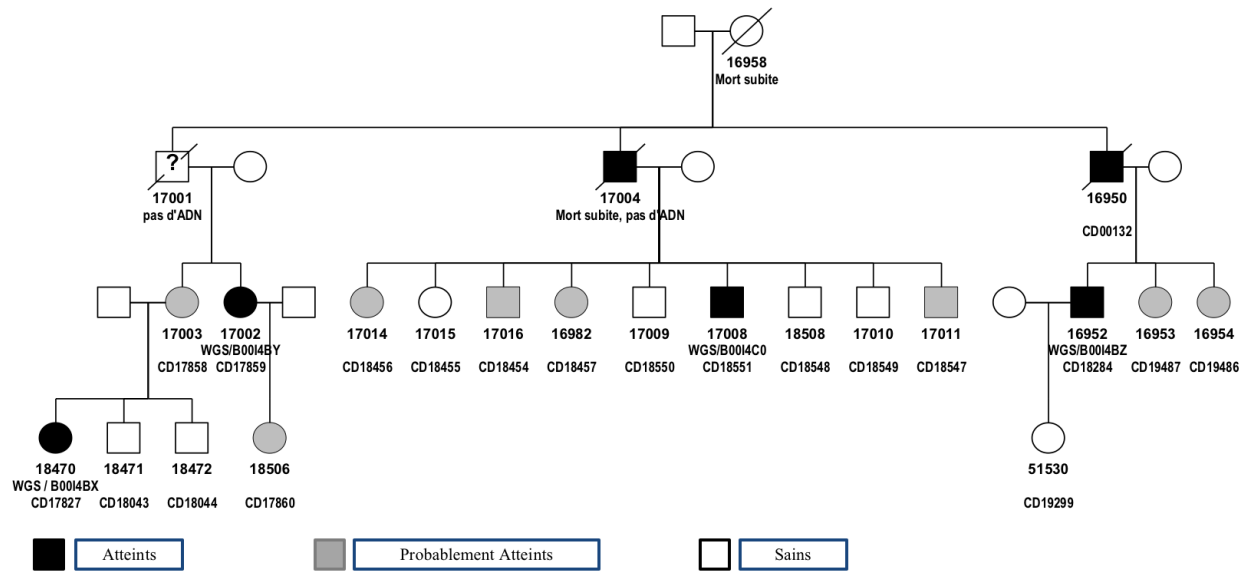


Figure 15 : Arbre généalogique IBD

A partir des données de génotypage haut débit de 20 individus de notre famille (figure 15) (12 atteints et 8 non atteints), nous avons recherché les régions identiques par descendance (IBD) partagées par les 12 individus atteints génotypés de la famille. Le nombre de paires d'atteints maximal partageant des régions génomiques IBD est de 62 ($\frac{n*(n-1)}{2}$)- $p^{P/E}$ avec n le nombre d'individus atteints génotypés ($n=12$) et $p^{P/E}$ le nombre de paires parent/enfant ($p^{P/E}=4$); les paires parent/enfant correspondent à la paire 17003/18470 et la paire 17002/18506. L'identification des régions IBD partagées par les individus de la famille sont présentés sur la figure 16.

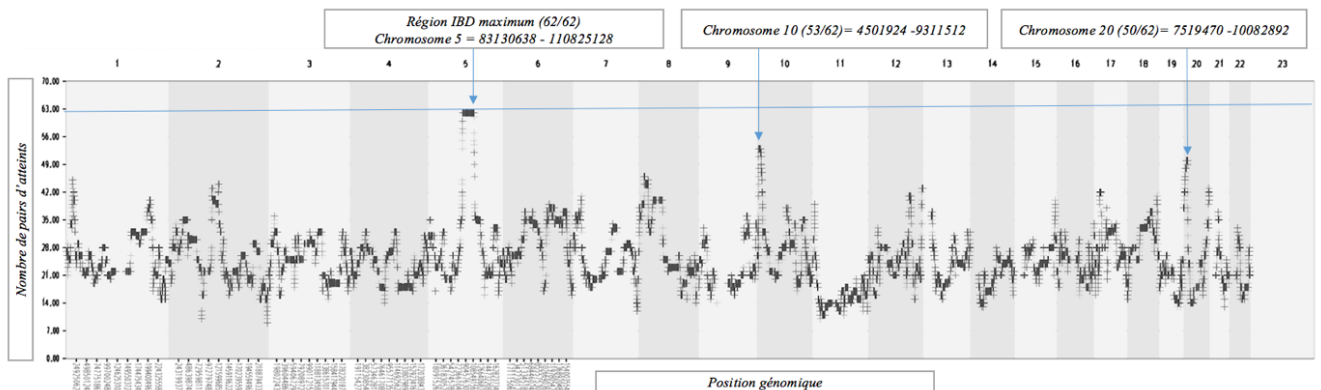


Figure 16 : représentation des régions IBD partagées par les individus génotypés de la famille

Chaque variant est représenté en fonction de sa position chromosomique (abscisse) et du nombre de paires d'individus atteints qui partage le même allèle du variant (ordonnées).

Le génotypage haut débit sur génome entier nous a permis d'identifier 1 région IBD maximale (62 paires sur 62) partagée par tous les individus atteints sur le chromosome 5 de 83130638 à 110825128 (rs 4129875 – rs 31619). Deux autres régions IBD ont été identifiées sur le chromosome 10 de 4501924 à 9311512 (53 paires sur 62) et sur le chromosome 20 de 7519470 à 10082892 (50 paires sur 62).

5.2 IDENTIFICATION DES VARIANTES RARES

Le séquençage des quatre individus (17002, 17008, 16952, 18470) génome entier (WGS) réalisé pour cette famille présente des données de couverture moyenne d'environ 40 fois nécessaires à l'analyse. (17002=42.57/17008=40.74/16952=43.74/18470=40.68)

L'analyse des données de séquençage génome entier n'a pas permis d'identifier un variant rare fonctionnel partagé par les 4 individus WGS, nous avons en revanche identifié 20 variants rares fonctionnels partagés par au moins 3 des 4 individus WGS successivement (tableau 8).

Tableau 8 : Variants fonctionnels rares retrouvés partagés par au moins 3 des 4 WGS successivement

Position génomique	Position ADNc	Position Protéique	MAF GM	MAF GN	Gene	Score GERP	IBD (paires d'atteints)
Chr1:g.39888514A>G	c.16202A>G	p.Asn5401Ser	-	2.0 E-04	MACF1	5.96	24
Chr3:g.150931612T>A	c.493A>T	p.Arg165Trp	-	-	P2RY14	0.89	19
Chr15:g.102197168CTTA T>C	c.2213_2216delATAA	p.Asn738fs	-	-	TARSL2	4.48	30
Chr12:g.10525783G>T	c.581C>A	p.Ser194*	-	-	KLRK1	4.7	26
Chr5:g.132585178A>G	c.818A>G	p.His273Arg	-	-	FSTL4	0	27
Chr7 :g.2978431C>A	c.899C>A	p.Ala300Asp	-	-	CARD11	5.72	23
Chr9 :g.15191192A>G	c.992A>G	p.Asn331Ser	-	-	TTC39B	4.9	28
Chr10 :g.5254979C>G	c.703C>G	p.Leu235Val	-	-	AKR1C4	3.3	53
Chr10 :g.17865094A>C	c.83A>C	p.Tyr28Ser	-	-	MRC1L1	4.17	
Chr10 :g.71883771T>G	c.70A>G	p.Ile24Val	-	-	AIFM2	5.46	38
Chr12 :g.12247479C>T	c.713C>T	p.Thr238Ile	5.83 E-04	-	BCL2L14	3.5	22
Chr1 :g.43888801G>A	c.2069G>A	p.Arg690Gln	-	-	SZT2	2.02	22
Chr3 :g.8668002C>T	c.680G>A	p.Ser227Asn	5.83 E-04	-	SSUH2	3.09	
Chr7 :g.148978652C>T	c.859C>T	p.Arg287Trp	-	-	ZNF783	2.78	31
Chr8 :g.52284546C>T	c.3788G>A	p.Ser1263Asn	-	-	PXDNL	0	40
Chr14 :g.105960185C>C A	c.600dupA	p.Arg201fs	-	-	C14orf80	3.37	22
Chr20 :g.1460623A>T	c.173T>A	p.Leu58Gln	-	-	SIRPB2	3.4	38
Chr1 :g.15701041G>A	c.3425G>A	p.Arg1142Gln	5.83 E-04	1.33 E-04	FHAD1	4.52	45
Chr1 :g.19018305G>A	c.644G>A	p.Arg215His	5.83 E-04	9.99 E-04	PAX7	4.98	42
Chr21 :g.14743897G>A	c.35C>T	p.Ala12Val	-	8.38 E-04	AL05030 2.1	0.1	21

La fréquence de l'allèle mineur (MAF) est donnée par les deux bases de données que sont GENMED (MAF GM) et GNOMAD (MAF GN). Les filtres successifs d'analyse de génome entier ont été réalisés par l'outil Knime4bio et permettent de conserver les variants fonctionnels décrits par Variant Effect Predictor (VEP), rares (bases de données publiques et internes) et retrouvés par l'algorithme de calling GATK. (Les étapes de l'analyse et les bases de données utilisées sont décrites dans matériels et méthodes).

5.3 ANALYSES COMBINÉES DES DONNÉES SEQUENÇAGE GÉNOME ENTIER ET DES RÉGIONS IBD IDENTIFIÉES

Ces analyses ont permis d'identifier qu'aucun des 20 variants sélectionnés n'était présent dans notre région IBD maximal sur le chromosome 5. En effet le variant *FSTL4* est en dehors de la région cible. Un variant rare fonctionnel est présent dans la région IBD du chromosome 10 partagés par 53 paires sur 62, il s'agit de *AKRIC4*. Le troisième pic sur le chromosome 20 ne coïncidait pas non plus avec l'un de nos 20 variants, en effet *SIRBP2* est en dehors de la région IBD cible.

5.4 SELECTION DES VARIANTS POUR LE SEQUENÇAGE CAPILLAIRE APRES ANALYSE COMBINÉE IBD/WGS

Devant le risque d'une atteinte différente de la jeune patiente (18470) et sa forme précoce de sclérose valvulaire aortique dans le cadre d'une maladie dégénérative, nous avons décidé de conserver les trois variants partagés par les 3 cousins (17002, 17008, 16952) que sont *MACF1/P2RY14* et *TARSL2*. Nous avons aussi choisi *KLRK1* car il est responsable d'un codon stop prématuré, *FSTL4*, *AKRIC4* et *SIRBP2* car ils sont sur le chromosome 5, le 10 et le 20 respectivement. Enfin nous avons aussi sélectionné *FHAD1* et *PAX7* car ils sont partagés par 45 et 42 paires/62 dans le cadre de l'analyse IBD (annexe 2).

5.5 SEQUENÇAGE CAPILLAIRE

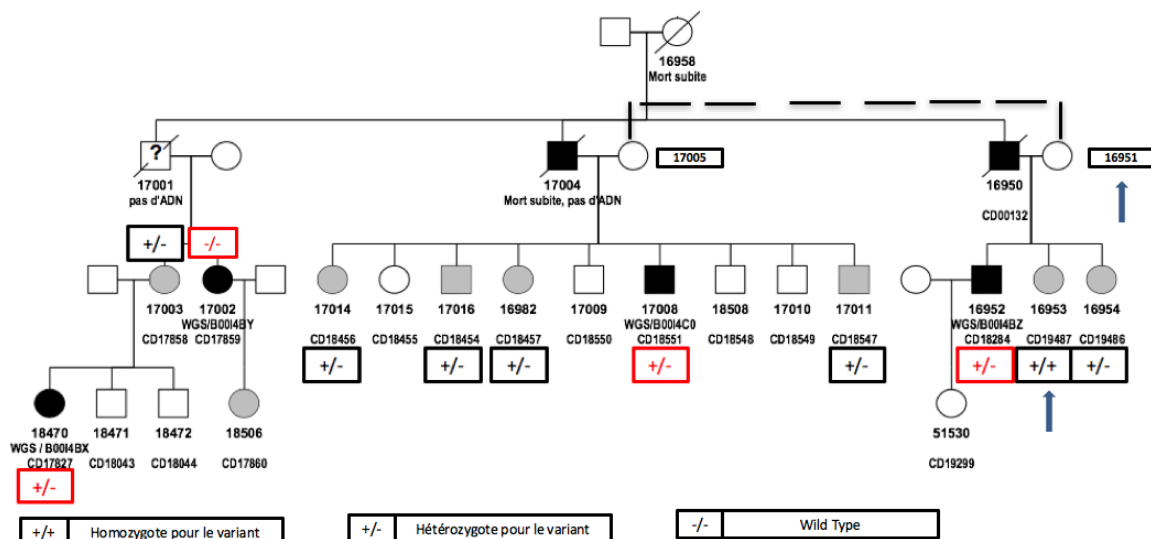


Figure 17 : Analyse de co-ségrégation du variant *PAX7* : p.Arg251His au sein de la famille RAC précoce

Les individus ayant bénéficié du WGS sont en rouge.

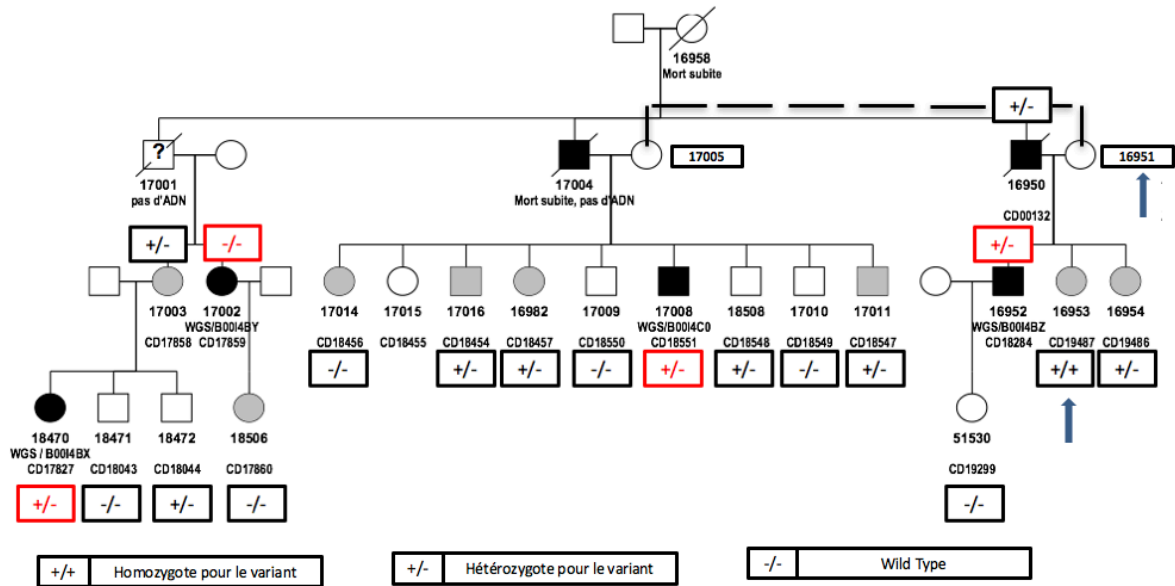


Figure 18 : Analyse de co-ségrégation du variant FHAD1 : p.Arg1142Gln au sein de la famille RAC précoce

Les patients WGS sont en rouge, pour chaque variant et comme cité plus haut nous avons 3 individus sur 4 qui sont porteurs du variant pour chaque variant testé. Le résultat du séquençage capillaire des variants de *PAX7* (figure 17) et *FHAD1* (figure 18) est assez similaire dans la problématique, au-delà d'être tous les deux sur le chromosome 1 nous pouvons constater que l'individu 16953 est homozygote pour ces deux variants, donc soit le variant est fréquent dans la région et n'a pas été filtré par la base de donnée interne soit les individus 16951 et 16950 sont apparentés à un certain degré dans l'ascendance.

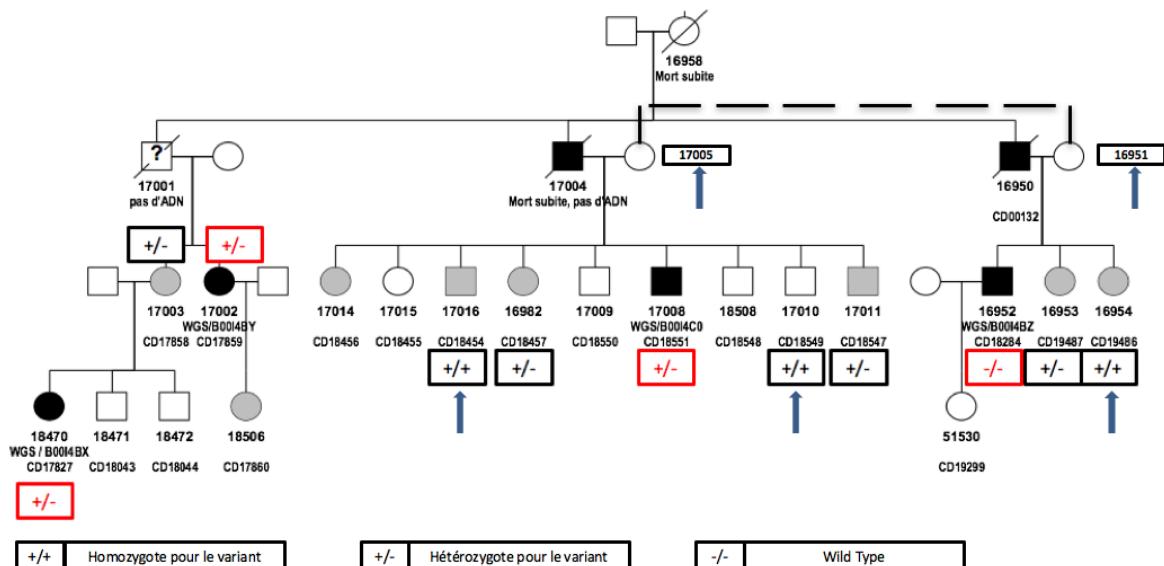


Figure 14 : Analyse de co-ségrégation du variant AKRIC4 : p.Leu235Val au sein de la famille RAC précoce

Le résultat pour le variant de *AKRIC4*, qui était un bon candidat car présent dans la région du chromosome 10 retrouvée en IBD, est encore plus surprenant car il soulève la même

interrogation que pour *PAX7* et *FHADI* avec l'individu 16954 qui est homozygote probablement car 16950 et 16951 sont apparentés à un certain degré, mais il met aussi en évidence la fraternité entre 17005 et 16951. Nous avons donc 3 descendants homozygotes pour ce variant.

Ces trois variants ne sont donc pas des candidats idéaux. Nous pouvons interrompre les investigations.

Le variant de *FSTL4* n'est pas présent chez tous les individus atteints ou probablement atteints. Il est présent sur le chromosome 5, en dehors de la région IBD max, mais n'est pas non plus le bon candidat.

Le séquençage capillaire des autres variants que sont *P2RY14*, *TARSL2*, *SIRPB2*, *KLRK1*, *MACF1* n'a pas permis de faire émerger de candidat idéal présent chez tous les patients atteints. Les arbres sont consultables en annexe.

DISCUSSION

L'identification de cette grande famille de RAC représente un enjeu majeur pour l'identification des causes génétiques du RAC. En effet cette famille présente peu de facteurs de risques cardiovasculaires, pouvant interférer avec le développement du RAC (7) et ainsi complexifier la recherche génétique. Cette famille a aussi pour particularité de développer de façon précoce une sclérose valvulaire aortique sur valves tricuspides, semblant évoluer vers une sténose calcifiée classique chez les sujets plus âgés.

1) CHOIX DU MATÉRIEL ET DES MÉTHODES

Les lésions de RAC présentent des caractéristiques identifiables en échocardiographie permettant d'avoir 3 générations de sujets vivants et considérés comme atteints, augmentant ainsi les capacités d'identification d'un facteur génétique impliqué dans le développement du RAC. Le phénotypage est la clé pour la réussite de l'analyse génétique. Il permet de limiter le nombre de phénotopies (augmentation de la fréquence du RAC dans la population générale) et de non pénétrant. Le risque de phénotopie doit être perpétuellement recherché et ce durant toute la durée de l'investigation. En effet la prévalence de la sténose calcifiée valvulaire aortique augmente de façon exponentielle avec l'âge. En guise d'exemple entre deux cas extrêmes, une patiente ultra-octogénaire présentant une sclérose valvulaire aortique normale pour l'âge doit-elle être considérée comme atteinte à la lumière du risque de phénotopie ? Au contraire un jeune patient d'une vingtaine d'années considéré comme sain/profil indéterminé devra être suivi et potentiellement sera étiqueté comme atteint des années plus tard. Ceci représente un challenge important pour l'identification d'un gène causal du RAC.

Dans notre travail de recherche et ce en corrélation avec ce qui est recommandé dans la littérature, nous avons classé l'atteinte valvulaire aortique en deux catégories fonctionnelles : la sclérose valvulaire aortique (anomalie morphologique de la valve aortique sans retentissement hémodynamique) et la sténose valvulaire aortique (anomalie morphologique avec retentissement hémodynamique) stade ultime de la maladie. Contrairement à la sténose valvulaire aortique dont nous pouvons quantifier de façon objective le retentissement hémodynamique, la sclérose reste quant à elle plus sujette à l'appréciation/expertise de l'opérateur qui doit être entraîné avec une connaissance parfaite de la famille étudiée. En effet l'ensemble des investigations génétiques qui vont suivre sera basé sur le phénotype attribué. De plus pour optimiser cette évaluation phénotypique, nous avons mis en place une approche multi-modalité couplant échocardiographie et scanner (score calcique) pour caractériser au mieux la pathologie dans cette famille.

Pour identifier un variant causal du RAC, nous avons utilisé une double approche d'analyse génétique tout à fait complémentaire, le génotypage haut débit (IBD) et le WGS. Cette double approche permet d'investiguer l'ensemble du génome sans aucune hypothèse a priori sur le mécanisme physiopathologique. Une des difficultés de l'analyse bio-informatique de ces données WGS est le choix des filtres/bases de données utilisées afin d'éliminer les variants fréquents. Il est indispensable d'appliquer des bases de données constituées d'individus démographiquement comparable avec le génome des patients afin d'éliminer les variants fréquents dans nos régions.

2) RÉSULTATS

L'étude de cette famille a permis de mettre en évidence un phénotype échocardiographique héréditaire unique de sclérose/sténose valvulaire aortique calcifiée associée à une aorte fusiforme sur valve tricuspide. Malgré l'appellation « RAC précoce », à travers cette atteinte précoce de type sclérose chez l'individu 18470, l'évolution de la maladie rentre dans un cadre dégénératif classique avec une atteinte partant d'une sclérose bénigne puis chez certains une évolution vers une sténose nécessitant un geste opératoire de remplacement valvulaire aortique. La moyenne d'âge opératoire dans notre famille est de 68 ans ce qui est comparable aux patients opérés de cette pathologie dans la population générale. L'aspect macroscopique et microscopique des valves explantées est comparable à ce qui est observé dans la population générale.

Cette famille présente peu de facteurs de risque cardiovasculaires ; le cholestérol total ainsi que le LDL sont certes plus élevés dans notre groupe d'individus atteints mais en restant dans la norme internationale. Thanassoulis et al (30) ont pour la première fois la significativité GWAS avec le SNP rs10455872 du gène *LPA* en association avec la présence d'une calcification valvulaire aortique. Leur travail suggère qu'une élévation chronique du niveau de Lp(a) augmente de façon marquée la prévalence d'une maladie valvulaire aortique calcifiée et implique ainsi la Lp(a) dans le développement de cette maladie. Une association entre le taux de Lp(a) et la maladie valvulaire aortique a ensuite été observée dans des études ultérieures(54–56). Ces études montraient que la Lp(a) s'accumule à la fois dans les lésions précoces et tardives de la valve aortique en se colocalisant avec le dépôt de calcium. Dans notre travail nous ne retrouvons pas ce lien causal, en effet le taux de Lp(a) est dans la norme et il n'y a pas de différence significative entre le groupe d'individus atteints et non atteints. Cela peut s'expliquer d'une part, par le manque de puissance (n=38) et d'autre part, par l'atteinte valvulaire dominante à type de sclérose et donc une faible atteinte calcique.

La sélection des individus pour le WGS était correcte avec quatre individus suffisamment éloignés généalogiquement, et représentatifs de la population d'individus atteints avec des données échocardiographiques comparables. Seul le score calcique valvulaire aortique était significativement différent s'expliquant par notre sélection des individus « les plus atteints ». L'apport du score calcique est intéressant pour quantifier la gravité de l'atteinte mais ne prend part que dans la deuxième partie de l'histoire naturelle et l'apparition de calcification valvulaire aortique. En effet l'aspect de fibrose valvulaire n'est pas pris en compte par le CT Scan, or son impact sur certaines formes de la maladie et notamment chez les femmes est à investiguer (57).

Bien que la maladie de Monckeberg soit regroupée au sein de familles (6), les facteurs génétiques expliquant cette hérédité restent difficilement identifiables d'où l'importance d'une sélection drastique de la famille «idéale» pour poursuivre les investigations génétiques modernes. L'identification par le génotypage haut débit d'une région IBD maximale, sur le chromosome 5, partagée par l'ensemble des individus génotypés nous confirme la bonne sélection familiale. En 2007, Martin et al. (58) par une approche GWAS de 38 familles de RAC (353 individus) a mis en évidence 3 loci sur le chromosome 18, le 13 et le chromosome 5, or le loci sur le 5 est le 5q21.2 soit en position génomique de 102 800 001 à 104 500 000 ce qui chevauche notre région IBD maximale allant de 83 130 638 à 110 825 128.

L'analyse bio-informatique via Knime4bio n'a pas permis de mettre en évidence de variants rares fonctionnels partagés par les 4 individus WGS. Nous avons donc sélectionné 20 variants rares fonctionnels partagés par au moins 3 des 4 individus. Parmi ces variants aucun ne résidait

dans la région IBD maximale du chromosome 5. Aucun des gènes sélectionnés n'était connu dans la littérature comme ayant un rapport avec le RAC et ces gènes n'ont pas été retrouvés dans l'étude des autres grandes familles de RAC investiguées dans notre unité. En étant moins stringent sur les filtres de fréquence des bases de données Gnomad et Genmed (de 0.1% à 5% et de 1% à 5% respectivement) nous mettons en évidence deux variants SERPINF1 et APC communs aux 4 individus WGS. SERPINF1 membre de la famille des serpinines est un gène qui code pour une protéine fortement anti-angiogénique. Des mutations dans ce gène ont été observées chez des individus présentant une ostéogénèse imparfaite de type IV. APC plus connu, code pour une protéine suppresseur de tumeur qui agit comme antagoniste de la voie de signalisation Wnt. Quand ce gène est défectueux, il peut être responsable de la polypose adénomateuse familiale (PAF). Ce gène candidat est intéressant pour deux raisons d'une part car il intervient dans la voie de signalisation Wnt, qui est une voie de signalisation ostéogénique impliquée dans le processus de calcification valvulaire aortique décrit dans la figure 4. Et d'autre part il est situé sur le chromosome 5 proche de notre région IBD maximale.

Le séquençage capillaire des 9 variants sélectionnés après analyse combinée WGS/IBD n'a pas permis de mettre en évidence de candidat idéal qui ségrége dans la famille. APC semble un bon candidat qu'il va falloir investiguer.

3) PROBLÈMES RENCONTRÉS

L'essence même de ce travail repose sur l'absolue nécessité d'étudier une famille d'individus avec une valve aortique tricuspide, or il persistait un doute au commencement des investigations sur la possible bicuspidie (contradiction entre compte-rendu opératoire et compte-rendu anatomopathologie) de deux de nos patients sources (17001/16950). Nous avons donc repris les différents compte-rendus et fait ressortir les coupes histologiques pour les analyser en détail. Nous avons pu conclure de façon formelle et collégiale pour l'individu 17001. En revanche persiste un doute pour l'individu 16950 car nous n'avons qu'une seule valvule en anatomopathologie donc difficile de conclure à une tricuspide. Le caractère très calcifié de cette valve et l'absence de la bicuspidie dans la descendance nous pousse à croire que cette valve était bien tricuspide.

Les analyses IBD ont pour but d'identifier des blocs haplotypiques partagés par des paires d'individus apparentés hérités d'un ancêtre commun sans événement de recombinaison. Ces analyses ont été réalisées avec un algorithme permettant le calcul de probabilités IBD à partir des données de génotypages haut-débit. Or les résultats préliminaires ont révélé qu'un ensemble d'individus étaient liés au reste de la famille de façon inattendue.

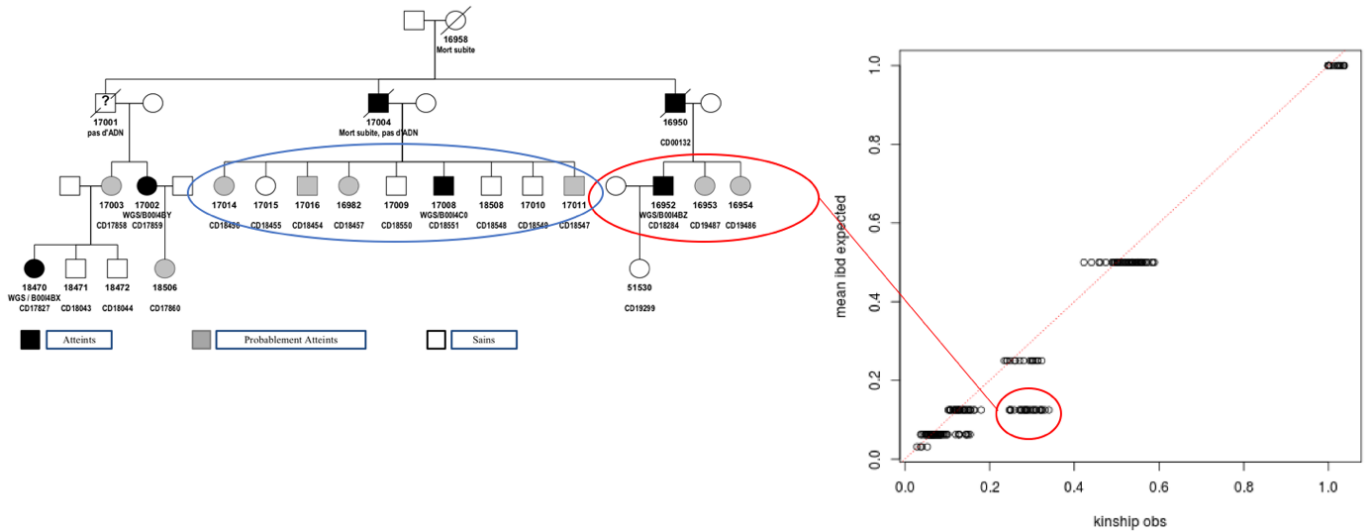


Figure 16 : résultats IBD préliminaires.

Contre toute attente les individus entourés en rouge ont un IBD plus haut avec les individus entourés en bleu (figure 16). Suite à cet épisode, découverte que les individus 16950 et 17005 étaient sœurs et donc mariées à deux frères. En rajoutant deux individus au pedigree comme étant les parents des individus 17005 et 16951 (figure 17), les apparentements observés sont bien ceux attendus pour tous les groupes d'individus. Nous avons ainsi pu ensuite accéder aux résultats présentés plus haut.

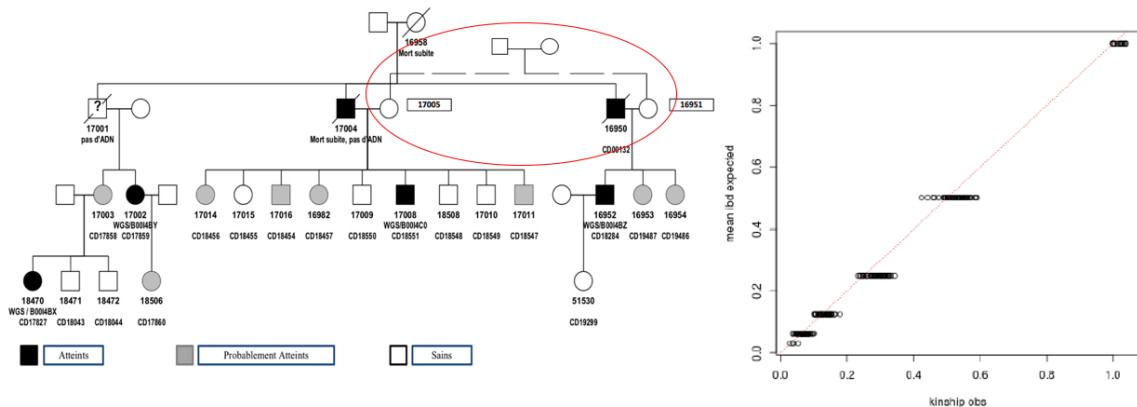


Figure 17 : apparentements attendus pour les groupes d'individus (IBD)

Nous avons ensuite fait face à des individus homozygotes pour plusieurs de nos variants sélectionnés et testés en séquençage capillaire. Soulevant deux hypothèses, soit les individus 16950 et 16951 ainsi que 17004 et 17005 sont apparentés entre eux à un certain degré dans l'ascendance soit les variants ne sont pas dans Genmed (peu probable qu'ils soient fréquents dans la région).

CONCLUSION ET PERSPECTIVES

L'approche phénotypique et génétique préliminaire d'une grande famille de RAC avec peu de facteurs de risques cardiovasculaires rapporte un portrait clinique homogène et unique de RAC tricuspide associé à une aorte fusiforme. Alors qu'un gène causal du RAC reste à identifier, le génotypage haut débit a permis de mettre en évidence une région IBD maximale sur le chromosome 5 associée au RAC dans cette famille.

L'importance d'une collaboration entre clinicien et généticien est cruciale dans la mesure où l'interprétation des résultats des investigations génétiques est étroitement liée au diagnostic phénotypique initial. Cette collaboration m'a permis de réaliser un travail transversal depuis le recrutement des familles, avec détermination des phénotypes, jusqu'à l'analyse génétique des familles.

Le variant rare fonctionnel du gène *APC*, gène impliqué dans la voie de signalisation Wnt, commun aux 4 individus WGS et proche région IBD maximale sur le chromosome 5 semble être notre candidat idéal qu'il faudra investiguer dans un avenir proche.

Le WGS permet l'étude des régions non codantes, régulatrices. Cependant l'étude de ces régions non codantes est encore au stade expérimental ne permettant pas de potentialiser totalement l'analyse du WGS. L'analyse de ces régions non codantes représente un des défis que va devoir relever la génétique moderne. A notre échelle locale, elle permettra d'investiguer cette région du chromosome 5 déjà mise en évidence par Martin et al, et semblant jouer un rôle dans le développement du RAC.

REFERENCES BIBLIOGRAPHIQUES

1. Nkomo VT, Gardin JM, Skelton TN, Gottdiener JS, Scott CG, Enriquez-Sarano M. Burden of valvular heart diseases: a population-based study. *Lancet* 2006 ;368:1005-11.
2. Iung B, Baron G, Butchart EG, Delahaye F, Gohlke-Bärwolf C, Levang OW, et al. A prospective survey of patients with valvular heart disease in Europe: The Euro Heart Survey on Valvular Heart Disease. *Eur Heart J*. 2003 ;24:1231-43.
3. Lindroos M, Kupari M, Heikkilä J, Tilvis R. Prevalence of aortic valve abnormalities in the elderly: An echocardiographic study of a random population sample. *J Am Coll Cardiol*. 1993 ;21:1220-5.
4. Avierinos J-F, Gersh BJ, Melton LJ, Bailey KR, Shub C, Nishimura RA, et al. Natural history of asymptomatic mitral valve prolapse in the community. *Circulation* 2002;106:1355-61.
5. Cosmi JE, Kort S, Tunick PA, Rosenzweig BP, Freedberg RS, Katz ES, et al. The Risk of the Development of Aortic Stenosis in Patients With Benign Aortic Valve Thickening. *Arch Intern Med*. 2002 ;162:2345-7.
6. Probst V. Familial Aggregation of Calcific Aortic Valve Stenosis in the Western Part of France. *Circulation* 2006 ;113:856-60.
7. Bossé Y, Mathieu P, Pibarot P. Genomics: The Next Step to Elucidate the Etiology of Calcific Aortic Valve Stenosis. *J Am Coll Cardiol*. 2008 ;51:1327-36.
8. Carabello BA, Paulus WJ. Aortic stenosis. *Lancet* 2009 ;373:956-66.
9. Lindman BR, Clavel M-A, Mathieu P, Iung B, Lancellotti P, Otto CM, et al. Calcific aortic stenosis. *Nat Rev Dis Primer*. 2016 ;2:16006.
10. Beppu S, Suzuki S, Matsuda H, Ohmori F, Nagata S, Miyatake K. Rapidity of progression of aortic stenosis in patients with congenital bicuspid aortic valves. *Am J Cardiol*. 1993 ;71:322-7.
11. Yan AT, Koh M, Chan KK, Guo H, Alter DA, Austin PC, et al. Association Between Cardiovascular Risk Factors and Aortic Stenosis: The CANHEART Aortic Stenosis Study. *J Am Coll Cardiol*. 2017 ;69:1523-32.
12. Lindman BR, Clavel M-A, Mathieu P, Iung B, Lancellotti P, Otto CM, et al. Calcific aortic stenosis. *Nat Rev Dis Primer*. 2016 ;2:16006.
13. Otto CM. Calcific aortic valve disease: outflow obstruction is the end stage of a systemic disease process. *Eur Heart J*. 2009 ;30:1940-2.
14. Baumgartner H, Falk V, Bax JJ, De Bonis M, Hamm C, Holm PJ, et al. ESC/EACTS Guidelines for the management of valvular heart disease. *Eur Heart J*. 2017 ;38:2739-91.
15. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009 ;461:747-53.
16. Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet TIG*. 2001 ;17:502-10.
17. Dweck MR, Boon NA, Newby DE. Calcific aortic stenosis: a disease of the valve and the myocardium. *J Am Coll Cardiol*. 2012 ;60:1854-63.
18. O'Brien KD, Reichenbach DD, Marcovina SM, Kuusisto J, Alpers CE, Otto CM. Apolipoproteins B, (a), and E accumulate in the morphologically early lesion of « degenerative » valvular aortic stenosis. *Arterioscler Thromb Vasc Biol*. 1996 ;16:523-32.
19. Olsson M, Thyberg J, Nilsson J. Presence of oxidized low density lipoprotein in nonrheumatic stenotic aortic valves. *Arterioscler Thromb Vasc Biol*. 1999 ;19:1218-22.
20. Parhami F, Morrow AD, Balucan J, Leitinger N, Watson AD, Tintut Y, et al. Lipid oxidation products have opposite effects on calcifying vascular cell and bone cell differentiation. A possible explanation for the paradox of arterial calcification in osteoporotic

patients. *Arterioscler Thromb Vasc Biol.* 1997 ;17:680-7.

21. Ghaisas NK, Foley JB, O'Briain DS, Crean P, Kelleher D, Walsh M. Adhesion molecules in nonrheumatic aortic valve disease: endothelial expression, serum levels and effects of valve replacement. *J Am Coll Cardiol.* 2000 ;36:2257-62.

22. Kaden JJ, Dempfle C-E, Grobholz R, Tran H-T, Kiliç R, Sarikoç A, et al. Interleukin-1 beta promotes matrix metalloproteinase expression and cell proliferation in calcific aortic valve stenosis. *Atherosclerosis.* 2003 ;170:205-11.

23. Jian B, Narula N, Li Q, Mohler ER, Levy RJ. Progression of aortic valve stenosis: TGF-beta1 is present in calcified aortic valve cusps and promotes aortic valve interstitial cell calcification via apoptosis. *Ann Thorac Surg.* 2003 ;75:457-65.

24. Ortlepp J, Hoffmann R, Ohme F, Lauscher J, Bleckmann F, Hanrath P. The vitamin D receptor genotype predisposes to the development of calcific aortic valve stenosis. *Heart* 2001 ;85:635-8.

25. Avakian SD, Annicchino-Bizzacchi JM, Grinberg M, Ramires J a. F, Mansur AP. Apolipoproteins AI, B, and E polymorphisms in severe aortic valve stenosis. *Clin Genet.* 2002 ;60:381-4.

26. Novaro GM, Sachar R, Pearce GL, Sprecher DL, Griffin BP. Association Between Apolipoprotein E Alleles and Calcific Valvular Heart Disease. *Circulation* 2003 ;108:1804-8.

27. Nordström P, Glader CA, Dahlén G, Birgander LS, Lorentzon R, Waldenström A, et al. Oestrogen receptor α gene polymorphism is related to aortic valve sclerosis in postmenopausal women. *J Intern Med.* 254:140-6.

28. Ortlepp JR, Schmitz F, Mevissen V, Weiss S, Huster J, Dronskowski R, et al. The amount of calcium-deficient hexagonal hydroxyapatite in aortic valves is influenced by gender and associated with genetic polymorphisms in patients with severe calcific aortic stenosis. *Eur Heart J.* 2004 ;25:514-22.

29. Chakravarti A. Population genetics--making sense out of sequence. *Nat Genet.* 1999 ;21:56-60.

30. Thanassoulis G, Campbell CY, Owens DS, Smith JG, Smith AV, Peloso GM, et al. Genetic Associations with Valvular Calcification and Aortic Stenosis. *N Engl J Med.* 2013 ;368:503-12.

31. Guauque-Olarte S, Messika-Zeitoun D, Droit A, Lamontagne M, Tremblay-Marchand J, Lavoie-Charland E, et al. Calcium Signaling Pathway Genes RUNX2 and CACNA1C Are Associated With Calcific Aortic Valve DiseaseCLINICAL PERSPECTIVE. *Circ Genomic Precis Med.* 2015 ;8:812-22.

32. Helgadóttir A, Thorleifsson G, Gretarsdóttir S, Stefansson OA, Tragante V, Thorolfssdóttir RB, et al. Genome-wide analysis yields new loci associating with aortic valve stenosis. *Nat Commun.* 2018 ;9:987.

33. Thériault S, Gaudreault N, Lamontagne M, Rosa M, Boulanger M-C, Messika-Zeitoun D, et al. A transcriptome-wide association study identifies PALMD as a susceptibility gene for calcific aortic valve stenosis. *Nat Commun.* 2018 ;9:988.

34. Garg V, Muth AN, Ransom JF, Schluterman MK, Barnes R, King IN, et al. Mutations in *NOTCH1* cause aortic valve disease. *Nature* 2005;437:270-4.

35. Rosenberg MJ, Agarwala R, Bouffard G, Davis J, Fiermonte G, Hilliard MS, et al. Mutant deoxynucleotide carrier is associated with congenital microcephaly. *Nat Genet.* 2002 ;32:175-9.

36. Sveinbjörnsdóttir S, Hicks AA, Jónsson T, Pétursson H, Guðmundsson G, Frigge ML, et al. Familial Aggregation of Parkinson's Disease in Iceland. *N Engl J Med.* 2009.

37. Aulchenko YS, Vaessen N, Heutink P, Pullen J, Snijders PJLM, Hofman A, et al. A Genome-Wide Search for Genes Involved in Type 2 Diabetes in a Recently Genetically Isolated Population From the Netherlands. *Diabetes.* 2003 ;52:3001-4.

38. Bellamy MF, Pellikka PA, Klarich KW, Tajik AJ, Enriquez-Sarano M. Association of cholesterol levels, hydroxymethylglutaryl coenzyme-A reductase inhibitor treatment, and progression of aortic stenosis in the community. *J Am Coll Cardiol.* 2002 ;40:1723-30.
39. Hekimian G, Boutten A, Flamant M, Duval X, Dehoux M, Benessiano J, et al. Progression of aortic valve stenosis is associated with bone remodelling and secondary hyperparathyroidism in elderly patients--the COFRASA study. *Eur Heart J.* 2013 ;34:1915-22.
40. Baumgartner H, Hung J, Bermejo J, Chambers JB, Evangelista A, Griffin BP, et al. Echocardiographic assessment of valve stenosis: EAE/ASE recommendations for clinical practice. *Eur J Echocardiogr.* 2009 ;10:1-25.
41. Michelena HI, Desjardins VA, Avierinos J-F, Russo A, Nkomo VT, Sundt TM, et al. Natural History of Asymptomatic Patients With Normally Functioning or Minimally Dysfunctional Bicuspid Aortic Valve in the Community. *Circulation* 2008 ;117:2776-84.
42. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, et al. Recommendations for Cardiac Chamber Quantification by Echocardiography in Adults: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr.* 2015 ;28:1-39.e14.
43. Nagueh SF, Smiseth OA, Appleton CP, Byrd BF, Dokainish H, Edvardsen T, et al. Recommendations for the Evaluation of Left Ventricular Diastolic Function by Echocardiography: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr.* 2016 ;29:277-314.
44. Cuff C, Serfaty J-M, Cimadevilla C, Laissy J-P, Himbert D, Tubach F, et al. Measurement of aortic valve calcification using multislice computed tomography: correlation with haemodynamic severity of aortic stenosis and clinical implication for patients with low ejection fraction. *Heart Br Card Soc.* 2011 ;97:721-6.
45. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol.* 1990 ;15:827-32.
46. Thaden JJ, Nkomo VT, Suri RM, Maleszewski JJ, Soderberg DJ, Clavel M-A, et al. Sex-related differences in calcific aortic stenosis: correlating clinical and echocardiographic characteristics and computed tomography aortic valve calcium score to excised aortic valve weight. *Eur Heart J.* 2016 ;37:693-9.
47. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010 ;20:1297-303.
48. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics.* 2011 ;27:2156-8.
49. Cunningham F, Amode MR, Barrell D, Beal K, Billis K, Brent S, et al. Ensembl 2015. *Nucleic Acids Res.* 2015 ;43:D662-9.
50. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol.* 2010 ;6:e1001025.
51. Han L, Abney M. Identity by descent estimation with dense genome-wide genotype data. *Genet Epidemiol.* 2011 ;35:557-67.
52. Lindenbaum P, Le Scouarnec S, Portero V, Redon R. Knime4Bio: a set of custom nodes for the interpretation of next-generation sequencing data with KNIME†. *Bioinformatics* 2011;27:3200-1.
53. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinforma Oxf Engl.* 2010 ;26:2069-70.
54. Owens DS, Katz R, Johnson E, Shavelle DM, Probstfield JL, Takasu J, et al. Interaction

of Age With Lipoproteins as Predictors of Aortic Valve Calcification in the Multi-Ethnic Study of Atherosclerosis. *Arch Intern Med.* 2008 ;168:1200-7.

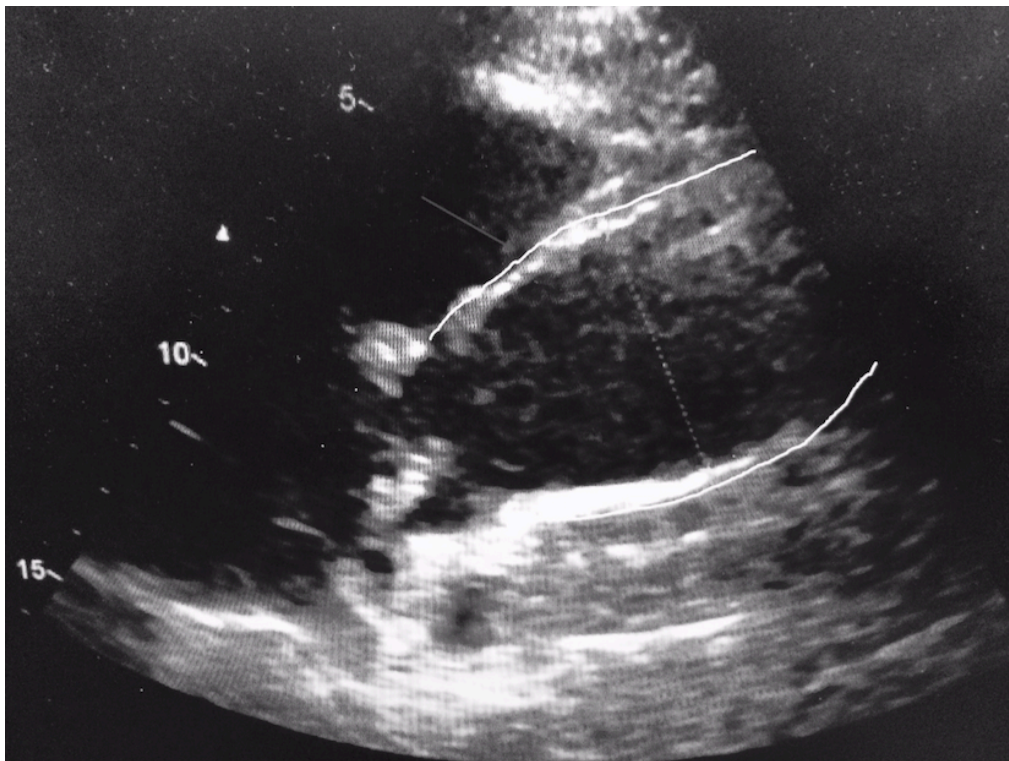
55. Capoulade R, Chan KL, Yeang C, Mathieu P, Bossé Y, Dumesnil JG, et al. Oxidized Phospholipids, Lipoprotein(a), and Progression of Calcific Aortic Valve Stenosis. *J Am Coll Cardiol.* 2015 ;66:1236-46.

56. Kamstrup PR, Tybjaerg-Hansen A, Nordestgaard BG. Elevated lipoprotein(a) and risk of aortic valve stenosis in the general population. *J Am Coll Cardiol.* 2014 ;63:470-7.

57. Simard L, Côté N, Dagenais F, Mathieu P, Couture C, Trahan S, et al. Sex-Related Discordance Between Aortic Valve Calcification and Hemodynamic Severity of Aortic Stenosis: Is Valvular Fibrosis the Explanation? *Circ Res.* 2017 ;120:681-91.

58. Martin LJ, Ramachandran V, Cripe LH, Hinton RB, Andelfinger G, Tabangin M, et al. Evidence in favor of linkage to human chromosomal regions 18q, 5q and 13q for bicuspid aortic valve and associated cardiovascular malformations. *Hum Genet.* 2007 ;121:275-84.

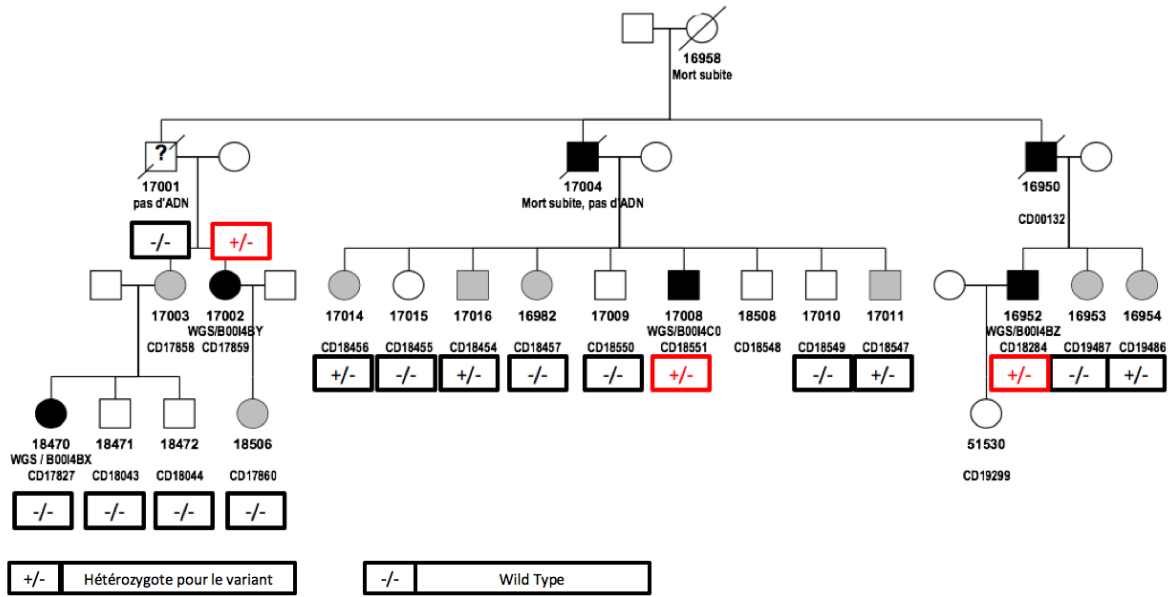
ANNEXES



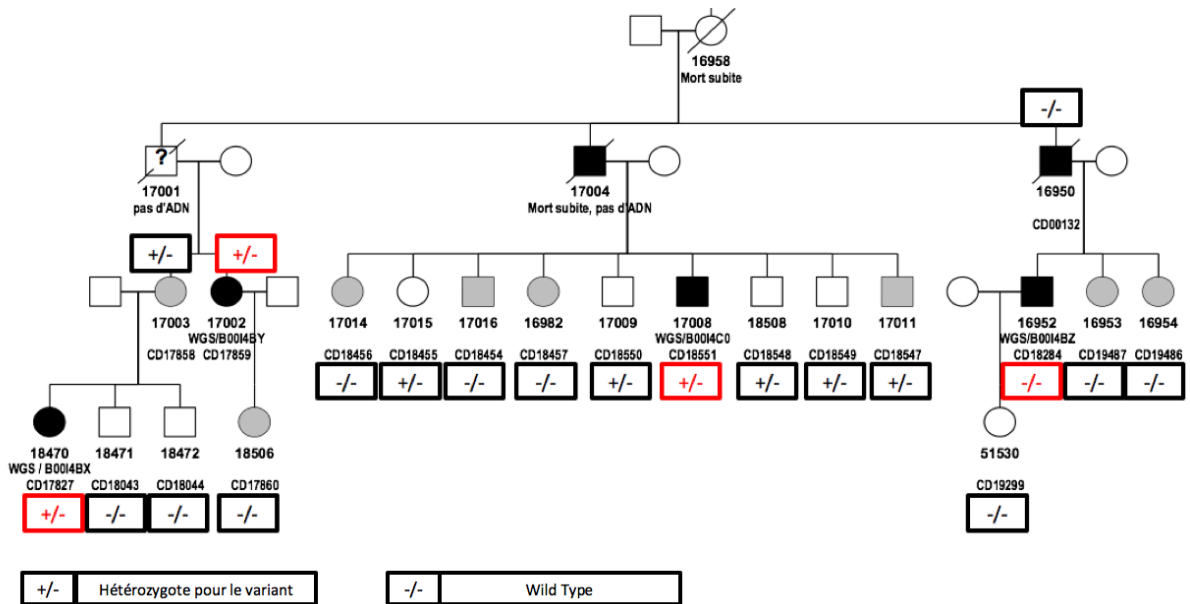
Annexe 1: Aorte fusiforme

Position génomique	Position ADNc	Position Protéique	MAF GM	MAF GN	Gene	Score GERP	IBD (paires d'atteints)
Chr1:g.39888514A>G	c.16202A>G	p.Asn5401Ser	-	-	MACF1	5.96	24
Chr3:g.150931612T>A	c.493A>T	p.Arg165Trp	-	-	P2RY14	0.89	19
Chr15:g.102197168CTTAT>C	c.2213_2216delATAA	p.Asn738fs	-	-	TARSL2	4.48	30
Chr12:g.10525783G>T	c.581C>A	p.Ser194*	-	-	KLRK1	4.7	26
Chr5:g.132585178A>G	c.818A>G	p.His273Arg	-	-	FSTL4	0	27
Chr10:g.5254979C>G	c.703C>G	p.Leu235Val	-	-	AKR1C4	3.3	53
Chr20:g.1460623A>T	c.173T>A	p.Leu58Gln	-	-	SIRPB2	3.4	38
Chr1:g.15701041G>A	c.3425G>A	p.Arg1142Gln	5.83 E-04	1.33 E-04	FHAD1	4.52	45
Chr1:g.19018305G>A	c.644G>A	p.Arg215His	5.83 E-04	9.99 E-04	PAX7	4.98	42

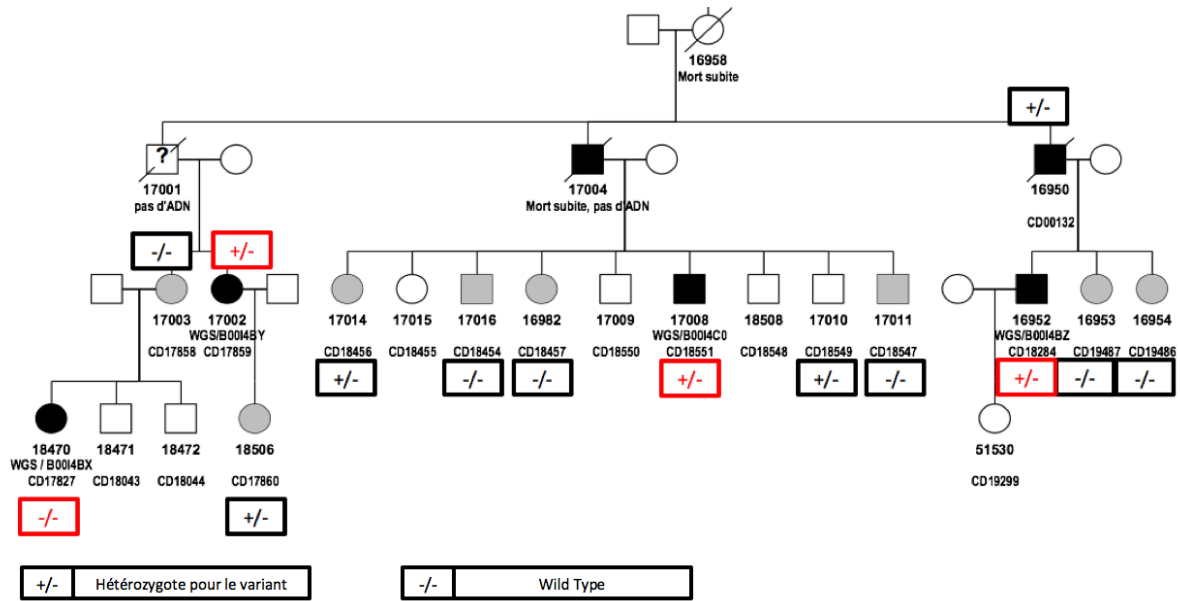
Annexe 2 : Variants fonctionnels rares sélectionnés pour le séquençage capillaire.



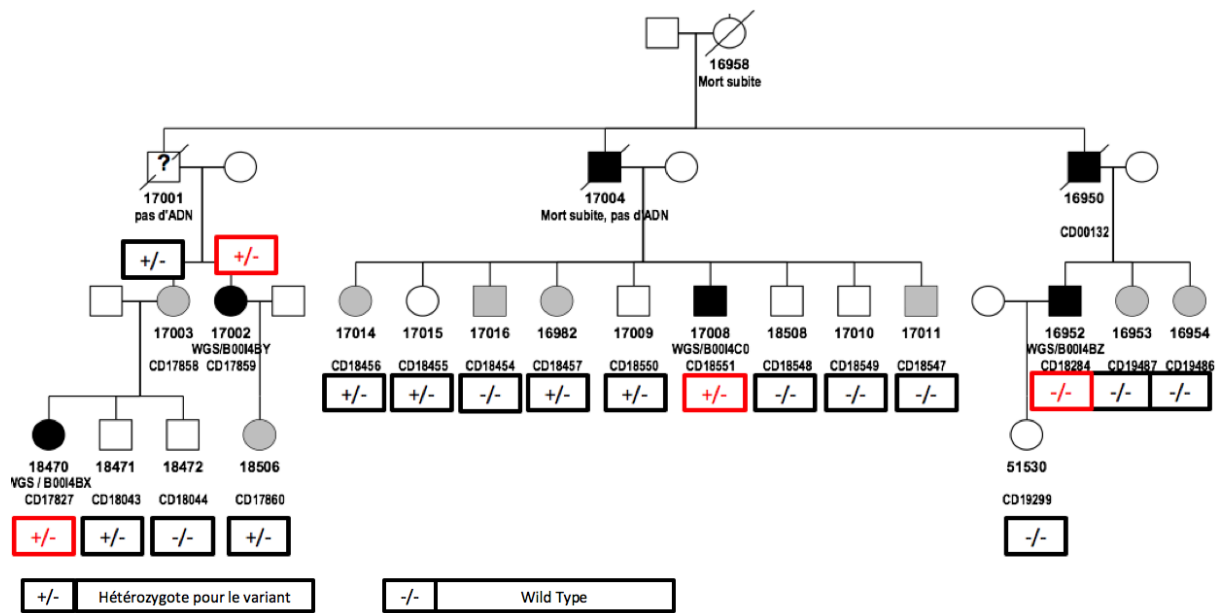
Annexe 3 : Analyse de co-ségrégation du variant TARSL2 : p.Asn738fs au sein de la famille RAC précoce



Annexe 4 : Analyse de co-ségrégation du variant KLRK1 : p.Ser194 au sein de la famille RAC précoce*



Annexe 5 : Analyse de co-ségrégation du variant P2RY14 : p.Arg165Trp au sein de la famille RAC précoce



Annexe 6 : Analyse de co-ségrégation du variant FSTL4 : p.His273Arg au sein de la famille RAC précoce

Titre du mémoire :

DÉTERMINANTS GÉNÉTIQUES ET NON GÉNÉTIQUES IMPLIQUÉS DANS LE DÉVELOPPEMENT DE LA STÉNOSE VALVULAIRE AORTIQUE DANS LE CADRE D'UNE FORME FAMILIALE.

RESUMÉ :

Introduction : La maladie calcifiante valvulaire aortique est considérée comme une conséquence du risque cardiovasculaire et de facteurs génétiques. Le rétrécissement valvulaire aortique calcifié (RAC) est une cause majeure de morbi-mortalité chez environ 5% des patients de plus de 65 ans. La physiopathologie du RAC est complexe, multifactorielle et malgré les investigations précédentes, la connaissance des déterminants génétiques du RAC reste pauvre.

Objectif : Etudier les caractéristiques phénotypiques et les déterminants génétiques et non-génétiques impliqués dans le développement d'une sclérose-sténose valvulaire aortique dans une grande famille avec peu de facteurs de risques cardiovasculaires.

Méthodes : Nous avons identifié une grande famille de RAC avec 3 générations d'individus vivants. Nous avons proposé un dépistage complet à tous les individus avec une évaluation clinique, biologique et échocardiographique. Le score calcique valvulaire aortique a été évalué par CT scan. Double approche génétique avec combinaison du génotypage des individus atteints permettant d'accéder aux IBD (Identity by Descendance) et du séquençage génome entier (WGS) chez 4 individus atteints génétiquement distants (3 cousins et leur nièce) pour identifier un variant rare causal du RAC.

Résultats : La famille comprend plus de 100 individus. A ce jour 38 membres ($53,1 \pm 15,8$ ans ; 19 hommes) ont été dépistés, 15 sont atteints ($62,4 \pm 12,7$ ans) et 23 sains ($47,1 \pm 15,0$ ans ; $p < 0,002$). La plus jeune patiente est une femme de 36 ans atteinte à type de sclérose valvulaire aortique. L'examen échocardiographique a révélé un phénotype héréditaire unique de sclérose/sténose valvulaire aortique calcifié associé à une aorte fusiforme chez 9 individus. Toutes les valves étaient tricuspides avec une sclérose dans 14 cas (36,8%) et une sténose dans 2 cas (5,3%). Sur les 15 individus atteints, 6 présentaient une insuffisance aortique ($p = 0,001$). L'aorte ascendante initiale était de plus grand diamètre chez les patients atteints ($34,2 \pm 4,3$ vs $27,1 \pm 4,6$ mm ; $p = 0,02$). Un rapport du diamètre de la jonction sino-tubulaire/diamètre initial de l'aorte ascendante a été calculé en échocardiographie et était plus faible chez les individus atteints ($0,88 \pm 0,08$ vs $0,94 \pm 0,05$, $p = 0,007$), ce qui confirme la présence d'un phénotype unique d'aorte fusiforme. Le score calcique valvulaire aortique était significativement supérieur dans le groupe atteint ($102,8 \pm 196,4$ vs 0 AU, $p = 0,02$). Le génotypage haut débit de 20 individus (12 atteints et 8 non atteints) nous a permis d'identifier 1 région IBD maximale (62 paires sur 62) partagée par tous les individus atteints sur le chromosome 5 de 83130638 à 110825128 (rs 4129875 – rs 31619). L'analyse des données de séquençage génome entier n'a pas permis d'identifier un variant rare fonctionnel partagé par les 4 individus WGS ; nous avons en revanche identifié 20 variants rares fonctionnels partagés par au moins 3 des 4 individus WGS successivement. L'analyse combinée des données de séquençage génome entier et du génotypage (IBD) ainsi que le séquençage capillaire des 20 variants sélectionnés n'ont pas fait émerger de candidat idéal. En étant moins stringent sur les filtres de fréquence des bases de données, un variant rare fonctionnel du gène APC a été mis en évidence, commun aux 4 individus WGs et sur le chromosome 5 proche de la région IBD maximale.

Conclusion : L'approche phénotypique et génétique préliminaire d'une grande famille de RAC avec peu de facteurs de risques cardiovasculaires rapporte un portrait clinique homogène et unique de RAC tricuspide associé à une aorte fusiforme. Alors qu'un gène causal du RAC reste à identifier, le génotypage haut débit a permis de mettre en évidence une région IBD maximale sur le chromosome 5 associée au RAC dans cette famille.

Mots-clés : Rétrécissement Aortique Calcifié (RAC), Pathologie dégénérative, Echocardiographie, Score calcique valvulaire, CT scan, Génétique, Séquençage génome entier (WGS), Séquençage haut débit de régions cibles.