



RECHERCHE DE NOUVEAUX FACTEURS GENETIQUES ASSOCIES AUX CARDIOMYOPATHIES DILATEES

Carole Proust

► **To cite this version:**

Carole Proust. RECHERCHE DE NOUVEAUX FACTEURS GENETIQUES ASSOCIES AUX CARDIOMYOPATHIES DILATEES. Génétique. 2012. <hal-01470753>

HAL Id: hal-01470753

<https://hal-ephe.archives-ouvertes.fr/hal-01470753>

Submitted on 17 Feb 2017

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

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

**MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE L'EDUCATION
NATIONALE**

ECOLE PRATIQUE DES HAUTES ETUDES

Sciences de la Vie et de la Terre

MEMOIRE

Présenté par

Carole PROUST

Pour l'obtention du diplôme de l'Ecole Pratique des Hautes Etudes

RECHERCHE DE NOUVEAUX FACTEURS GENETIQUES

ASSOCIES AUX CARDIOMYOPATHIES DILATEES.

Soutenu le 15 novembre 2012 devant le jury suivant :

Sophie Gad-Lapiteau	Président
Daniel Stockholm	Rapporteur
Alain Carrié	Examineur
Susan Saint-Just	Examineur
Laurence Tiret	Examineur

Laboratoire de Pharmacologie Cellulaire
et Moléculaire de l'EPHE

Centre de Recherche des Cordeliers
15, rue de l'Ecole de Médecine, 75006 Paris

Directeur : Dr. Jean Chambaz

Tutrice pédagogique : Susan Saint Just

susan.saint-just@crc.jussieu.fr

Laboratoire de Génomique Cardiovasculaire
UMRS 937 INSERM

Faculté de Médecine Pierre et Marie Curie

Directeur : Dr. François Cambien

Tutrice scientifique : Laurence Tiret

laurence.tiret@upmc.fr

91, Bd de l'Hôpital, 75013 Paris

ECOLE PRATIQUE DES HAUTES ETUDES

Sciences de la Vie et de la Terre

RECHERCHE DE NOUVEAUX FACTEURS GENETIQUES

ASSOCIES AUX CARDIOMYOPATHIES DILATEES.

Carole Proust

RESUME

Les cardiomyopathies dilatées (CMD) sont une cause majeure d'insuffisance cardiaque. Jusqu'à présent, la génétique de ces maladies est mal comprise, en effet les CMD sont, pour la majeure partie, des formes sporadiques alors que pour un tiers des patients seulement, une transmission familiale a été mise en évidence. A partir de cette composante génétique, de l'identification de gènes mutés et de protéines déficientes, il est possible d'espérer à la fois comprendre les mécanismes des formes familiales, et ouvrir quelques pistes dans les mécanismes des formes sporadiques.

A travers les recherches menées dans notre laboratoire, les connaissances récentes du génome humain et les nouvelles technologies mises à notre disposition, nous avons tenté d'identifier de nouveaux loci d'intérêt contribuant aux CMD sporadiques en réalisant la première étude pangénomique, ou « GWAS », de ces pathologies dans des pools d'ADN de patients et de sujets témoins.

La première phase de l'étude génétique a porté sur l'analyse de plus de mille patients et autant de sujets témoins, dont l'ADN a été mis en commun (pools) pour tester l'association des CMD avec près de 600 000 polymorphismes génétiques (SNPs). Les deux SNPs trouvés significativement associés à la maladie ont été validés par une seconde étape de génotypage individuel au sein de nouvelles cohortes de patients et de sujets témoins.

Ainsi, deux locus ont été mis en évidence par cette étude, le locus 1p36.13 proche du gène *HSPB7*, déjà mis en cause lors de précédentes analyses génétiques, et le locus 10q26 autour du gène *BAG3*. Les résultats du séquençage du gène *BAG3* dans des familles de CMD suggèrent pour la première fois qu'un gène pourrait être impliqué à la fois dans les formes sporadiques mais aussi familiales des CMD.

MOTS-CLES : Cardiomyopathie dilatées – GWAS – Pools d'ADN – *HSPB7* – *BAG3*

TABLE DES MATIERES

<u>INTRODUCTION</u>	5
<u>LES CARDIOMYOPATHIES</u>	5
<u>1. HISTORIQUE ET DEFINITION</u>	5
<u>2. APPROCHE MOLECULAIRE DES CARDIOMYPATHIES</u>	6
<u>3. LES CARDIOMYOPATHIES DILATEES (CMD)</u>	6
<u>3.1 Généralités</u>	6
<u>3.2 Etiologie</u>	7
<u>3.3 Hétérogénéité génétique et phénotypique</u>	7
<u>3.3.1 Génétique des formes monogéniques</u>	7
<u>3.3.2 Génétique des formes multifactorielles</u>	9
<u>STRATEGIES D'ETUDES GENETIQUES</u>	10
<u>1. STRATÉGIE D'ETUDE GÉNÉTIQUE DES FORMES FAMILIALES DE CMD</u>	10
<u>2. STRATÉGIE D'ETUDE GÉNÉTIQUE DES FORMES MULTIFACTORIELLES DE CMD</u>	11
<u>3. PRINCIPE DES ETUDES D'ASSOCIATION</u>	11
<u>3.1 Stratégie d'association « gène candidat »</u>	12
<u>3.2 Stratégie « association génome entier » (GWAS)</u>	13
<u>4. LES BIOPUCES OU « MICRO-ARRAYS »</u>	13
<u>4.1 Historique et principe</u>	13
<u>4.2 Stratégie de sélection des polymorphismes</u>	14
<u>4.3 Applications</u>	15
<u>5. OJECTIF DE L'ETUDE</u>	16

ABREVIATIONS

MCV : Maladies Cardiovasculaires

OMS : Organisation Mondiale de la Santé

CMD: Cardiomyopathie Dilatée

CMH : Cardiomyopathie Hypertrophique

CMR : Cardiomyopathie Restrictive

CVDA : Cardiomyopathie Ventriculaire Droite Arythmogène

AD : Autosomique Dominant

SNP : Single Nucleotide Polymorphism

SSCP : Single Strand Conformation polymorphism

DGGE : Denaturing Gradient Gel Electrophoresis

DHPLC : Denaturing High Performance Liquid Chromatography

DL : Déséquilibre de Liaison

GWAS : Genome Wide Association Study ou études pangénomiques

HapMap : Haplotype Map ou carte des haplotypes

EDTA : Ethylène Diamine Tétracétique

OR : Odds Ratio

CI : Corrélation Intraclasse

MAF : Fréquence de l'allèle Mineur

INTRODUCTION

LES CARDIOMYOPATHIES

1. HISTORIQUE ET DEFINITION

C'est en 1957 que le terme de « cardiomyopathie » apparaît pour la première fois, et en 1980 que l'Organisation Mondiale de la Santé (OMS) définit la pathologie comme étant « une maladie du muscle cardiaque de cause inconnue » {OMS, 1980}. En 1996, les cardiomyopathies sont définies par Richardson {Richardson, 1996} comme des « maladies du myocarde associées à une dysfonction ventriculaire » au sein desquelles on distingue, d'une part, les cardiomyopathies primitives idiopathiques (sans cause connue) ou familiales, qu'elles soient dilatées, hypertrophiques ou restrictives, et d'autre part, les cardiomyopathies secondaires à une cause infectieuse, endocrinienne, métabolique, toxique ou infiltrante (amylose, hémochromatose).

La classification à l'intérieur de ce groupe hétérogène de maladies est rendue difficile par le chevauchement de certaines catégories de cardiomyopathies, que se soit d'un point de vue anatomique ou physiologique, mais aussi génétique. Ainsi une maladie peut être classée dans deux catégories à la fois. Prenant en considération toutes les difficultés de classification et de compréhension de ces maladies, ainsi que l'avancée de la recherche génétique dans le domaine, Maron et al., en 2006, {Maron, 2006} définissent les cardiomyopathies comme : « un groupe hétérogène de maladies du myocarde associées à un dysfonctionnement mécanique et / ou électrique qui généralement (mais pas toujours) se caractérisent par une hypertrophie ventriculaire inappropriée ou dilatation et sont dues à une variété de causes qui sont fréquemment génétiques ».

En 2008, la Société Européenne de Cardiologie quant à elle, définit les cardiomyopathies comme étant une « pathologie du myocarde dans laquelle le muscle cardiaque est structurellement et fonctionnellement anormal et cela en l'absence de maladie coronaire, d'hypertension, de valvulopathie et de maladie cardiaque congénitale » {Elliott, 2008}. L'étiologie de ces maladies se dévoile progressivement au fur et à mesure des études effectuées, cependant leur physiopathologie reste peu connue. Les cardiomyopathies sont classées en sous-groupes phénotypiques en fonction de critères morphologiques et fonctionnels, on distingue ainsi :

- Les cardiomyopathies dilatées (CMD), caractérisées par une dilatation des cavités cardiaques associée à une altération de la fonction systolique ventriculaire.
- Les cardiomyopathies hypertrophiques (CMH), caractérisées par une hypertrophie du ventricule gauche principalement au niveau du septum interventriculaire.

- Les cardiomyopathies restrictives (CMR), caractérisées par une rigidité des parois cardiaques altérant la fonction diastolique.
- Les cardiomyopathies ventriculaires droites arythmogènes (CVDA), caractérisées par des arythmies et le remplacement du tissu myocardique du ventricule droit par un tissu fibro-adipeux.

Toutes ces atteintes cardiaques sont de gravité et d'évolution variable, certaines sont à haut risque de mort subite, en particulier chez le jeune adulte ou le sportif de haut niveau (CMH et CVDA). L'évolution peut se faire aussi vers une altération de la fonction systolique avec le développement d'une insuffisance cardiaque.

2. APPROCHE MOLECULAIRE DES CARDIOMYPATHIES

Les causes génétiques de ces pathologies se dévoilent progressivement grâce aux avancées considérables et récentes de la génétique moléculaire. Dès 1990, des études moléculaires sont menées pour les cardiomyopathies hypertrophiques avant d'être étendues aux phénotypes dilatés et restrictifs. Les cardiomyopathies arythmogènes du ventricule droit quant à elles, n'ont été étudiées d'un point de vu génétique qu'à partir de 2002.

L'ensemble de ces études génétiques a révélé une grande hétérogénéité génétique, c'est-à-dire que de multiples gènes peuvent être impliqués dans le développement des cardiomyopathies. Des mutations ont ainsi été identifiées dans divers gènes du sarcomère (CMH, CMD) ; du cytosquelette (CMD) ; des jonctions intercellulaires ; la desmine, la transthyréline, le fibrinogène A ou le gène HFE (pour les CMR). Cette hétérogénéité génétique peut impliquer que diverses mutations résultent en un même phénotype mais également que plusieurs gènes soient impliqués dans le développement de la maladie. Les études génétiques sont de plus rendue compliquées par l'intervention de facteurs environnementaux dans la mise en place du phénotype ce qui classe la plupart des cardiomyopathies dans le domaine des maladies multifactorielles. Bien souvent les gènes de susceptibilités ne sont pas encore identifiés, et en outre, lorsqu'ils le sont, leurs mécanismes d'action dans la maladie restent souvent hypothétiques.

Le travail présenté au sein de ce manuscrit, traitera spécifiquement de l'étude des cardiomyopathies dilatées.

3. LES CARDIOMYOPATHIES DILATEES (CMD)

3.1 Généralités

Les cardiomyopathies dilatées (CMD) sont des affections sévères caractérisées par une dilatation des ventricules avec une altération de la contractilité du myocarde qui peut conduire au développement d'une insuffisance cardiaque et à la mort subite. C'est d'ailleurs la principale indication de transplantation cardiaque {Dec, 1994 ; Everly, 2008}. Cette maladie touche les deux sexes et toutes les tranches d'âge, {Maron, 2006 ; Roger, 2011}, sa

prévalence est de 1 adulte sur 2500 avec une fréquence plus élevée chez les hommes que chez les femmes {Codd, 1989}. Chez l'enfant, l'incidence annuelle est de 0,57 cas pour 100 000, et là encore elle semble être plus élevée chez les garçons que chez les filles. Bien que la physiopathologie des CMD soit encore mal comprise {Taylor, 2006}, elle est considérée comme multifactorielle impliquant des facteurs environnementaux et la présence d'une forte prédisposition génétique {Jefferies, 2010}.

3.2 Etiologie

Les causes de CMD sont multiples. Elles peuvent être d'une part, d'origine toxique (médicamenteuse, alcoolique), virale et/ou auto-immune, ou, d'autre part, avoir une origine génétique. Dans le cas où la maladie est d'origine génétique on va distinguer la forme dite « familiale », touchant différents membres d'une même famille, de la forme dite « sporadique » ou après enquête génétique, aucun autre membre de la famille d'un malade que ce soit de sa génération ou de générations antérieures n'est atteint.

Les études génétiques portant sur l'identification des gènes impliqués dans les pathologies humaines ont débuté depuis des décennies mais se sont surtout orientées vers l'identification des gènes impliqués dans le développement de forme monogénique et familiale des maladies. En effet lorsque l'on s'intéresse aux maladies multifactorielles, beaucoup plus fréquemment sporadiques, les études sont rendues beaucoup plus difficiles du fait de l'hétérogénéité génétique et environnementale.

3.3 Hétérogénéité génétique et phénotypique

3.3.1 Génétique des formes monogéniques

Les formes familiales, et donc génétiques, de la maladie, sont le plus souvent monogéniques et représentent 20 à 35% des cas de CMD {Grunig, 1998 ; Goerss, 1995}. Elles sont caractérisées par une grande hétérogénéité, que ce soit au niveau des modes de transmission, des différentes formes cliniques, ou bien du point de vue de la multiplicité des gènes et *loci* impliqués.

- Mode de transmission

Les formes familiales se transmettent dans 80 à 90% des cas selon un mode autosomique dominant (AD) {Mestroni, 1999}. Cependant diverses études ont rapporté une transmission récessive liée à l'X {Bione, 1996 ; Vesel, 2003 ; Towbin, 1993} touchant essentiellement de jeunes garçons, et plus rarement un mode de transmission mitochondrial {Zhang, 2003} ou autosomique récessif {Arbustini, 1998 ; Mestroni, 1999} .

- Phénotype

Une analyse plus précise du phénotype montre qu'au sein même des formes à transmission AD, une stratification existe : les CMD « isolées » (forme la plus fréquente) et les CMD associées à des troubles de la conduction et/ou à des myopathies. Il existe également

une hétérogénéité allélique au sein des CMD puisqu'en fonction de la mutation présente au sein du gène impliqué le phénotype clinique peut varier. C'est par exemple le cas pour les mutations affectant les Desmines {Li, 1999}. Ceci ajoute un élément de complexité supplémentaire dans la compréhension de ces maladies.

- Pénétrance et expressivité variable

La pénétrance de la maladie est incomplète, c'est-à-dire qu'un individu porteur d'un génotype à risque ne va pas forcément développer le phénotype, et de plus l'expressivité est variable. En effet, des individus de même génotype n'expriment pas tous le phénotype à un même degré. Mangin et al {Mangin, 1999}, évaluent la pénétrance des CMD au sein de familles françaises atteintes de CMD à transmission AD, à 72% chez l'adulte. Selon les mêmes auteurs, l'âge de survenue de la maladie est significativement influencé par l'âge et le sexe.

- Génétique moléculaire

Un grand nombre de gènes ou *loci* ont été identifiés comme étant impliqués dans le développement de la maladie. En 2000, Towbin et al. répertorient des mutations associées à la maladie dans plus de 20 gènes {Towbin, 2000}. A l'heure actuelle plus de 30 gènes sont recensés comme étant associés à la maladie {Hershberger, 2011}.

Il existe donc une grande diversité dans les causes génétiques incriminées {Shaw, 2002}. Cependant peu d'études valident de véritables voies physiopathologiques, seules des hypothèses sont émises quant aux mécanismes potentiels d'action des gènes incriminés. Ainsi dans le cadre des CMD à transmission AD, les mutations identifiées touchent :

- des gènes codant des protéines du sarcomère (unité fondamentale de la contraction de toutes les cellules du muscle strié) identiques à celles impliquées dans les CMH (*MYBPC3*, *MYH7*, *TNNT2*, *TNNI3*, *TPM1*, *ACTC*). Les mutations dans ces gènes pourraient altérer la production de la force contractile mais également altérer sa transmission d'un sarcomère à un autre {Seidman, 2001}.

- des gènes codant des protéines du cytosquelette (dystrophine, α -sarcoglycanes...) et des protéines des filaments intermédiaires (desmine), et pouvant altérer la transmission de la force contractile {Olson, 1998 ; Schonberger, 2001}, cette dernière serait alors moins bien transmise par les protéines filamenteuses porteuses de mutation.

- des gènes codant des protéines nucléaires (émerine, lamines A/C) qui pourraient altérer notamment la stabilité nucléaire ainsi que d'autres protéines qui pourraient affecter différents mécanismes intervenant dans la fonction cardiaque comme la signalisation calcique {Schmitt, 2003} ou l'apoptose {Wencker, 2003}.

Parmi les gènes décrits (tableau 1), et contrairement aux CMH, aucun d'entre eux ne prédomine, la fréquence des mutations au sein de ces gènes est souvent <1% et leur analyse ne permet de trouver l'anomalie causale que dans une faible proportion de patients {Hershberger, 2011}. Récemment, une étude portant sur le séquençage à haut débit du gène de la Titine chez des sujets atteints de cardiomyopathie dilatée, des sujets atteints de cardiomyopathies hypertrophiques et des sujets contrôles, a identifié des mutations dans les CMD. Cependant, les mutations tronquantes (mutations conduisant à la formation d'un codon stop prématuré aboutissant à la formation d'une protéine tronquée) surviennent dans environ 25% des cas familiaux de la maladie, dans 18% des cas sporadiques mais sont aussi présentes chez les sujets contrôles {Herman, 2012}. A ce jour il n'existe donc aucune preuve d'un défaut moléculaire généralisé pour ces maladies. L'ensemble de ces résultats suggère l'existence d'autres gènes qu'il reste à identifier.

3.3.2 Génétique des formes multifactorielles

Les formes sporadiques de CMD constituent la majorité des cas (environ 70 %), et la maladie est considérée ici comme multifactorielle avec une composante génétique probable. Il n'existe pas de données précises, cependant, en 2001, le Docteur Charron {Charron, 2001} distingue deux types de gènes parmi les facteurs génétiques potentiellement impliqués dans ces formes multifactorielles, d'une part des gènes de prédisposition qui représenteraient des facteurs de risque de développer la maladie (tableau 2), d'autre part des gènes modificateurs qui modèleraient l'évolution et la sévérité de la maladie lorsque celle-ci est apparue.

- Gènes de prédisposition associés aux CMD

De nombreux gènes sont connus pour être impliqués dans les maladies cardiovasculaires. Afin d'établir une relation entre ces gènes et les cardiomyopathies dilatées, la méthodologie consiste habituellement à analyser des polymorphismes génétiques selon une stratégie dite « gène-candidat », souvent dans le cadre d'études cas-témoins. Le principe de cette stratégie sera développé dans la partie suivante.

Plusieurs gènes de prédisposition ont ainsi été étudiés dans le cadre des CMD sporadiques afin d'identifier une éventuelle association d'un ou plusieurs de leurs polymorphismes avec la maladie. Les premières études de ce type ont analysé le gène de l'Enzyme de Conversion de l'Angiotensine I (*ECA*) à travers le polymorphisme I/D (insertion/délétion) situé dans l'intron 16 du gène, mais des résultats contradictoires ont été rapportés quant au rôle délétère du génotype homozygote DD. Dans les CMD aucune association n'a été trouvée. De même plusieurs gènes du système rénine-angiotensine, des récepteurs beta-adrénérgiques, TNF alfa, TGF beta, ... connus pour jouer un rôle dans les pathologies cardiovasculaires ont été étudiés mais ne se sont pas révélés associés aux CMD.

En revanche d'autres études ont permis de suggérer l'association de gènes de prédisposition aux CMD.

- Gènes modificateurs associés aux CMD

Le principe des analyses est ici de comparer l'évolution et/ou la sévérité de la maladie au sein d'une population de malades, en fonction d'un polymorphisme génétique donné (génotype ou allèle). Andersson et al {Andersson, 1996} ont analysé le rôle du polymorphisme I/D de l'ACE sur l'évolution de la maladie, par la construction de courbes de survie. Le génotype D/D était significativement associé à un plus mauvais pronostic dans une population de 193 patients atteints d'insuffisance cardiaque idiopathique. Ce polymorphisme apparaît donc comme un facteur pronostique dans l'évolution de l'insuffisance cardiaque idiopathique multifactorielle. D'autres gènes ont été associés au pronostic de l'insuffisance cardiaque.

Dans les formes multifactorielles, plusieurs polymorphismes de gènes ont été associés à la prédisposition ou à la gravité du phénotype. Cependant, ces associations ont principalement été observées sur un petit nombre de cas et devront être confirmées. La part génétique pour ces formes de la maladie est donc loin d'être élucidée, pourtant l'identification de gènes de susceptibilité pourrait permettre d'identifier des voies physiopathologiques d'intérêt et l'identification de gènes modificateurs pourrait s'avérer utile pour identifier les patients à haut risque de complication.

Pour étudier la composante génétique des maladies complexes telles que les cardiomyopathies dilatées, qu'elles soient familiales ou sporadiques, plusieurs stratégies d'étude existent et ne cessent d'évoluer parallèlement à l'évolution des connaissances du génome humain.

STRATEGIES D'ETUDES GENETIQUES

1. STRATEGIE D'ETUDE GENETIQUE DES FORMES FAMILIALES DE CMD.

Dans le cadre des cardiomyopathies dilatées familiales, du fait du taux élevé de mortalité liée à la maladie et de la pénétrance incomplète, les études de liaison génétique (analyse de la coségrégation des allèles d'un marqueur avec la maladie au sein d'une famille) ne constituent pas une stratégie appropriée pour découvrir les gènes impliqués. C'est pourquoi

la recherche de gènes de prédisposition est réalisée grâce à des analyses d'associations, le plus souvent par une approche gène candidat. Cette stratégie consiste à étudier des gènes connus dont la fonction suggère qu'ils pourraient être impliqués dans l'étiologie de la maladie. C'est par cette approche que la plupart des mutations associées aux CMD de forme autosomique dominante ont été décelées.

2. STRATEGIE D'ETUDE GENETIQUE DES FORMES MULTIFACTORIELLES DE CMD

Les études d'association dites « pangé-nomique » (sans hypothèse biologique *a priori*) constituent une approche privilégiée pour étudier la composante génétique des maladies complexes telles que les formes sporadiques de CMD. Elles prennent en compte de multiples facteurs génétiques et environnementaux, elles ont la capacité de détecter des effets génétiques relativement modestes et ont bénéficié d'avancées récentes spectaculaires tant dans le domaine de la connaissance du génome humain et de sa variabilité que du développement d'outils technologiques de plus en plus performants alliant le haut débit à la miniaturisation. Centrées au départ sur un petit nombre de polymorphismes de gènes candidats, ces approches sont maintenant étendues au génome entier avec la perspective de pouvoir découvrir de nouveaux mécanismes physiopathologiques dans les maladies complexes.

Afin de mieux appréhender les différentes stratégies d'étude génétique permettant d'étudier la maladie, il est nécessaire de rappeler en quoi consistent les études d'association génétique.

3. PRINCIPE DES ETUDES D'ASSOCIATION

La séquence des 3×10^9 paires de bases (bp) qui composent le génome humain n'est pas identique d'un individu à l'autre {Lander, 2001 ; Venter, 2001}. Il existe un grand nombre de variations, ou polymorphismes (environ 1 toutes les 1000 bp), dont la fréquence varie d'une population à une autre. Les variations les plus courantes sont des substitutions d'un nucléotide par un autre (SNP pour "Single Nucleotide Polymorphism"). Les autres types de variations sont les polymorphismes de répétition, d'insertion/délétion et les polymorphismes de structure qui affectent de grandes régions chromosomiques. La plupart des polymorphismes sont situés dans des régions non fonctionnelles du génome et n'ont pas d'impact phénotypique. Ils sont dits "neutres" et sont généralement désignés sous le terme de "marqueurs". Cependant, quand ces polymorphismes affectent des régions codantes ou régulatrices de gènes, ils peuvent modifier la séquence de la protéine ou le niveau d'expression du gène et se traduire par des effets phénotypiques observables.

Le principe des études d'association est de mettre en relation ces variations génétiques avec un phénotype (qui peut être la maladie ou un de ses facteurs de risque) dans des études

épidémiologiques ou cliniques réalisées à partir d'échantillons d'individus non apparentés (même si les études d'association peuvent également s'appliquer à des familles).

L'approche repose sur le phénomène fondamental de déséquilibre de liaison (DL) entre polymorphismes. En effet, les polymorphismes, quand ils sont physiquement proches sur le génome, ne sont souvent pas indépendants les uns des autres. Quand une nouvelle mutation survient, elle va être associée à une combinaison particulière des allèles des autres polymorphismes déjà existants sur le même chromosome. Les recombinaisons survenant au cours de la méiose ont tendance par la suite à éroder cette association, mais pour des polymorphismes proches (par exemple au sein d'un même gène), la corrélation, dénommée DL, persiste au cours du temps. Pour cette raison, et d'autres liées à l'histoire des populations et à la sélection, des associations statistiques fortes entre polymorphismes voisins sont souvent observées au niveau populationnel. Cela implique que même si un polymorphisme impliqué dans le processus pathologique étudié n'est pas directement mesuré, son association avec la maladie peut être "capturée" par un polymorphisme voisin en DL avec lui. Ceci constitue la base des études d'association dont le principe est de tester l'association statistique entre des marqueurs génétiques mesurés et la maladie ou ses phénotypes intermédiaires. Quand une association est mise en évidence, cela n'implique pas forcément que le polymorphisme a un rôle causal dans la maladie, mais qu'il est situé à proximité d'un ou plusieurs polymorphismes fonctionnels qu'il reste à identifier. La localisation dans des régions potentiellement fonctionnelles du gène, ainsi que la nature du polymorphisme (par exemple, induisant un changement d'acide aminé) constituent des arguments en faveur d'un rôle causal, mais seules des expériences analysant la fonctionnalité du variant génétique dans des modèles cellulaires ou animaux permettront d'établir ce rôle.

3.1 Stratégie d'association « gène candidat »

Initialement, les études d'association consistaient à tester la relation entre la maladie et des polymorphismes de gènes candidats choisis pour leur implication potentielle dans le processus physiopathologique étudié. Les toutes premières études d'association ont été conduites avec des polymorphismes identifiés au niveau protéique, comme les polymorphismes du système HLA ou le polymorphisme de l'apolipoprotéine E. Dans les années 80, l'apparition de nouvelles techniques de biologie moléculaire ont permis la caractérisation de nouveaux polymorphismes de type RFLP (Restriction Fragment Length Polymorphism). Les années 90 ont vu apparaître des techniques plus performantes (SSCP, DGGE, DHPLC, spectrométrie de masse, séquençage direct) permettant de détecter directement les variations de séquence de l'ADN, tels que les SNPs, et ont rendu possible le criblage moléculaire systématique des gènes pour identifier l'ensemble de leurs polymorphismes fréquents.

La stratégie gène candidat se limite à un nombre restreint de gènes connus supposés contribuer à la maladie, ceux-ci ne représentant qu'une fraction de ceux potentiellement impliqués.

3.2 Stratégie « association génome entier » (GWAS)

Une approche plus récente est apparue avec l'achèvement du Projet Génome Humain en 2003 et les premiers résultats du projet international HapMap en 2005 (projet visant à créer une ressource publique de SNPs caractérisant la variabilité du génome humain dans quatre populations d'origines ethniques différentes) qui ont complètement révolutionné la recherche des facteurs génétiques impliqués dans les maladies communes. Il existe maintenant des bases de données publiques qui contiennent la séquence du génome humain, et un catalogue des variations génétiques humaines fréquentes qui s'enrichit au cours du temps (The International HapMap Consortium, 2005, <http://hapmap.ncbi.nlm.nih.gov/>).

Ces connaissances, couplées aux nouvelles technologies de génotypage à haut débit (puces à ADN) permettent d'analyser rapidement et avec précision le génome entier à la recherche de variants pouvant contribuer à l'apparition d'une maladie sans faire d'hypothèse physiopathologique *a priori*. L'idée est que, s'il existe quelque part dans le génome des variants génétiques liés à la maladie, il devrait être possible de les détecter grâce à leur DL avec des SNPs génotypés (tagSNPs), à condition que la carte des SNPs génotypés soit suffisamment dense. Cette approche consistant à cribler le génome entier sans aucune hypothèse *a priori* est dénommée GWAS (genome-wide association study). Contrairement à l'approche gène candidat, l'approche GWAS est complètement indépendante de la biologie traditionnelle et offre donc des perspectives importantes pour découvrir de nouveaux mécanismes étiologiques des maladies {Hirschhorn, 2005}.

Cette approche, qui reste avant tout exploratoire, est source de nouvelles hypothèses et de nouveaux gènes candidats à tester ensuite de manière plus approfondie.

4. LES BIOPUCES OU « MICRO-ARRAYS »

L'avènement des connaissances nouvelles du génome ainsi que l'émergence des approches visant à explorer la variabilité de l'ensemble du génome ont nécessité la mise au point de technologies de biologie moléculaire nouvelles. Ainsi, de nombreux laboratoires académiques et surtout plus récemment des sociétés privées ont mis au point de nouveaux outils performants qui ne cessent d'évoluer : « Les biopuces ».

4.1 Historique et principe

Le principe d'une biopuce repose sur l'analyse en quelques heures de milliers de séquences d'ADN ou d'ARN. Grâce à cette technologie, il est possible de génotyper très rapidement l'ADN de nombreux individus pour des variants connus *a priori*. Du fait du

potentiel qu'on leur attribue en génomique, les biopuces suscitent un intérêt croissant de la part de la communauté scientifique et des industriels.

Une puce comporte des milliers d'unités d'hybridation appelées « spots », chacune étant constituée d'un dépôt de fragments d'ADN ou oligonucléotides correspondant à des sondes de séquence connue. L'hybridation de la puce avec un échantillon biologique d'ADN ou d'ARN, marqué par une molécule fluorescente, permet de détecter et de quantifier l'ensemble des cibles qu'il contient en une seule expérience.

Les puces de génotypage, puces utilisées au cours de mon travail, intègrent des sondes ciblant les séquences flanquantes des SNPs (séquences entourant le variant). Pour chaque SNP, deux sondes correspondant à la séquence de ses deux allèles sont intégrées. L'ADN de l'échantillon à analyser est hybridé sur la puce de manière à fournir une mesure pour chaque allèle, marqué par un fluorophore différent. Le niveau relatif de ces deux mesures permet alors de déterminer le génotype. Chaque type de puce nécessite des traitements informatiques et statistiques spécifiques.

Le génotypage par biopuce a pour objectif de mesurer simultanément le génotype de plusieurs centaines de milliers de SNPs pour un individu. Depuis leur apparition, la capacité des puces de génotypage n'a cessé de croître, en effet si en 2004 les premières puces permettaient de génotyper 10 000 SNPs à la fois, aujourd'hui la société Illumina propose des puces permettant de génotyper 4,3 millions de SNPs (Illumina Omni5). Une telle densité a été permise par l'accroissement des connaissances sur la variabilité du génome humain résultant des étapes successives du projet HapMap (www.hapmap.org) et plus récemment, du projet 1000 Génomes (www.1000genomes.org).

4.2 Stratégie de sélection des polymorphismes

Outre les différences de construction des puces, les puces de génotypage diffèrent par les stratégies de sélection des SNPs. Le choix des SNPs à inclure sur une puce de génotypage visant à capturer un maximum d'information sur les SNPs non mesurés, ceci impose une connaissance *a priori* du DL à partir des informations fournies par le projet HapMap. Ce choix est un point déterminant pour le design des biopuces.

Les SNPs présents sur les puces sont pour la plupart des polymorphismes fréquents (fréquence allélique > 0.01) sélectionnés pour établir une couverture fine du génome permettant d'utiliser les SNPs mesurés pour « marquer » les SNPs non génotypés, soit par une forte densité et un espacement régulier (choix de la société Affymetrix), soit par sélection de « proxys » caractérisant les haplotypes de la région et permettant d'inférer au mieux les génotypes des SNPs manquant par imputation (choix d'Illumina).

4.3 Applications

La première publication faisant date dans le domaine des GWAS est celle du Wellcome Trust Case Control Consortium {WTCCC, 2007}, dans laquelle l'approche pan-génomique à l'aide des puces a été appliquée à 14,000 malades pour 7 maladies communes et 3 000 témoins, et qui a permis l'identification de plusieurs nouveaux loci de susceptibilité pour les différentes maladies étudiées. Depuis cette publication, de nombreuses études GWAS ont conduit à la découverte de nouveaux loci de susceptibilité aux maladies complexes, répliqués pour la plupart dans des études indépendantes (<http://www.genome.gov/gwastudies/>).

Dans le domaine cardiovasculaire, l'exemple le plus marquant est celui du locus 9p21 qui a été initialement découvert dans une étude GWAS sur la maladie coronaire et s'est révélé ensuite impliqué dans d'autres pathologies cardiovasculaires telles les accidents vasculaires cérébraux et l'anévrisme aortique {Helgadottir, 2008 ; Chen, 2009} ainsi que le diabète de type 2 {Shea, 2011}.

Cependant, les technologies permettant ces études restent coûteuses pour des laboratoires académiques, c'est pourquoi dans le cadre de notre projet une alternative au génotypage individuel a été proposée, consistant à effectuer les analyses par biopuces sur un mélange d'ADN (ou « pool d'ADN ») de plusieurs individus de même phénotype. Chaque pool contient l'ADN de dizaines, voire de centaines d'individus, cela permet de réduire considérablement le coût des études tout en préservant une puissance adéquate {Sham, 2002 ; Liu, 2005}. L'intérêt majeur des études sur ADN poolés est non seulement de réduire considérablement les coûts, mais aussi la quantité d'ADN requise, un aspect non négligeable pour des études anciennes dans lesquelles la quantité d'ADN disponible est limitée.

Cependant, en contrepartie, il existe une perte d'information dans cette approche comparativement à des études utilisant le génotypage individuel. Plusieurs raisons contribuent à cette perte d'information :

- La première raison provient des erreurs expérimentales propres à la formation des pools d'ADN.
- La seconde raison réside dans l'imprécision relative de l'estimation de la fréquence allélique des SNPs, qui est mesurée à partir d'un signal quantitatif proportionnel au nombre d'allèles présents dans le pool, et non pas dérivée du comptage des génotypes individuels comme dans une GWAS classique.
- La troisième raison est que la conception des pools ne permet pas d'évaluer l'effet de variables qui n'entrent pas dans la définition *a priori* des groupes.

Toutefois, l'impact de ces inconvénients peut être la plupart du temps limité par un protocole expérimental rigoureux et une méthodologie appropriée pour la conception des pools.

Pour notre projet, le challenge résidait d'une part dans la mise au point d'un protocole fiable et reproductible pour la constitution des pools d'ADN, et d'autre part dans le développement d'outils statistiques et informatiques spécifiquement adaptés à l'analyse des pools.

5. OJECTIF DE L'ETUDE

En dépit des succès des GWAS pour d'autres pathologies cardiovasculaires {Yasuno, 2010 ; Smith, 2010 ; Schunkert, 2011}, aucune GWAS sur les CMD non familiales n'avait encore été réalisée au moment du lancement de notre projet. L'une des raisons tenait probablement à la grande difficulté d'obtenir des cohortes suffisamment nombreuses et homogènes de patients pour cette maladie relativement rare.

A l'aide de puces de génotypage à haute densité couvrant l'ensemble du génome, nous avons cherché à identifier des marqueurs génétiques associés aux CMD idiopathiques, dans une étude rassemblant plus de 2000 malades et 2000 témoins.

Afin de diminuer les coûts d'une telle étude et d'économiser les précieux ADN de nos patients nous avons mis au point une approche visant à analyser sur des biopuces, des mélanges d'ADN ou « pools » d'ADN, stratifiés selon l'âge, le sexe et l'origine géographique d'une population de sujets atteints de CMD sporadique d'une part, et de sujets témoins d'autre part.

Le projet final comporte quatre étapes distinctes :

- Mise au point d'un protocole visant à pooler les échantillons entre eux.
- Etude d'association génome entier à partir des pools d'ADN constitués.
- Vérification des résultats significatifs issus de la GWAS par génotypage individuel des mêmes individus.
- Réplication des résultats les plus significatifs issus des deux dernières étapes par génotypage individuel dans des cohortes différentes.

BIBLIOGRAPHIE

1. Abelmann WH. Classification and natural history of primary myocardial disease. *Prog Cardiovasc Dis* 1984;27(2):73-94.
2. Abraham R, Moskvina V, Sims R, et al. A genome-wide association study for late-onset Alzheimer's disease using DNA pooling. *BMC Med Genomics* 2008;1:44.
3. Andersson B, Sylven C. The DD genotype of the angiotensin-converting enzyme gene is associated with increased mortality in idiopathic heart failure. *J Am Coll Cardiol* 1996;28(1):162-7.
4. Arbustini E, Diegoli M, Fasani R, et al. Mitochondrial DNA mutations and mitochondrial abnormalities in dilated cardiomyopathy. *Am J Pathol* 1998;153(5):1501-10.
5. Arimura T, Hayashi T, Terada H, et al. A Cypher/ZASP mutation associated with dilated cardiomyopathy alters the binding affinity to protein kinase C. *J Biol Chem* 2004;279(8):6746-52.
6. Arimura T, Hayashi YK, Murakami T, et al. Mutational analysis of fukutin gene in dilated cardiomyopathy and hypertrophic cardiomyopathy, 2009.
7. Arimura T, Ishikawa T, Nunoda S, Kawai S, Kimura A. Dilated cardiomyopathy-associated BAG3 mutations impair Z-disc assembly and enhance sensitivity to apoptosis in cardiomyocytes. *Hum Mutat* 2011;32(12):1481-91.
8. Arimura T, Matsumoto Y, Okazaki O, et al. Structural analysis of obscurin gene in hypertrophic cardiomyopathy. *Biochem Biophys Res Commun* 2007;362(2):281-7.
9. Arimura T, Nakamura T, Hiroi S, et al. Characterization of the human nebulin gene: a polymorphism in an actin-binding motif is associated with nonfamilial idiopathic dilated cardiomyopathy. *Hum Genet* 2000;107(5):440-51.
10. Arola AM, Sanchez X, Murphy RT, et al. Mutations in PDLIM3 and MYOZ1 encoding myocyte Z line proteins are infrequently found in idiopathic dilated cardiomyopathy. *Mol Genet Metab* 2007;90(4):435-40.
11. Barbaux S, Tregouet DA, Nicaud V, et al. Polymorphisms in 33 inflammatory genes and risk of myocardial infarction--a system genetics approach. *J Mol Med* 2007;85(11):1271-80.
12. Barbaux SC, Blankenberg S, Rupprecht HJ, et al. Association between P-selectin gene polymorphisms and soluble P-selectin levels and their relation to coronary artery disease. *Arterioscler Thromb Vasc Biol* 2001;21(10):1668-73.
13. Bienengraeber M, Olson TM, Selivanov VA, et al. ABCC9 mutations identified in human dilated cardiomyopathy disrupt catalytic KATP channel gating. *Nat Genet* 2004;36(4):382-7.
14. Bione S, D'Adamo P, Maestrini E, Gedeon AK, Bolhuis PA, Toniolo D. A novel X-linked gene, G4.5, is responsible for Barth syndrome. *Nat Genet* 1996;12(4):385-9.
15. Borjesson M, Magnusson Y, Hjalmarson A, Andersson B. A novel polymorphism in the gene coding for the beta(1)-adrenergic receptor associated with survival in patients with heart failure. *Eur Heart J* 2000;21(22):1853-8.

16. Bosse Y, Bacot F, Montpetit A, et al. Identification of susceptibility genes for complex diseases using pooling-based genome-wide association scans. *Hum Genet* 2009;125(3):305-18.
17. Brauch KM, Karst ML, Herron KJ, et al. Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. *J Am Coll Cardiol* 2009;54(10):930-41.
18. Cambien F, Poirier O, Nicaud V, et al. Sequence diversity in 36 candidate genes for cardiovascular disorders. *Am J Hum Genet* 1999;65(1):183-91.
19. Cappola TP, Matkovich SJ, Wang W, et al. Loss-of-function DNA sequence variant in the CLCNKA chloride channel implicates the cardio-renal axis in interindividual heart failure risk variation. *Proc Natl Acad Sci U S A* 2011;108(6):2456-61.
20. Carballo S, Robinson P, Otway R, et al. Identification and functional characterization of cardiac troponin I as a novel disease gene in autosomal dominant dilated cardiomyopathy. *Circ Res* 2009;105(4):375-82.
21. Carniel E, Taylor MR, Sinagra G, et al. Alpha-myosin heavy chain: a sarcomeric gene associated with dilated and hypertrophic phenotypes of cardiomyopathy. *Circulation* 2005;112(1):54-9.
22. Charron P, Komajda M. Are we ready for pharmacogenomics in heart failure? *Eur J Pharmacol* 2001;417(1-2):1-9.
23. Charron P, Komajda M. Genes and their polymorphisms in mono- and multifactorial cardiomyopathies: towards pharmacogenomics in heart failure. *Pharmacogenomics* 2002;3(3):367-78.
24. Charron P, Tesson F, Poirier O, et al. Identification of a genetic risk factor for idiopathic dilated cardiomyopathy. Involvement of a polymorphism in the endothelin receptor type A gene. *CARDIGENE* group. *Eur Heart J* 1999;20(21):1587-91.
25. Chen Z, Qian Q, Ma G, et al. A common variant on chromosome 9p21 affects the risk of early-onset coronary artery disease. *Mol Biol Rep* 2009;36(5):889-93.
26. Codd MB, Sugrue DD, Gersh BJ, Melton LJ, 3rd. Epidemiology of idiopathic dilated and hypertrophic cardiomyopathy. A population-based study in Olmsted County, Minnesota, 1975-1984. *Circulation* 1989;80(3):564-72.
27. Comabella M, Craig DW, Camina-Tato M, et al. Identification of a novel risk locus for multiple sclerosis at 13q31.3 by a pooled genome-wide scan of 500,000 single nucleotide polymorphisms. *PLoS One* 2008;3(10):e3490.
28. Combadiere C, Godin O, Vidal C, Cangialosi A, Proust C, Tzourio C. Common CX3CR1 alleles are associated with a reduced risk of headaches. *Headache* 2008;48(7):1061-6.
29. Conrad DF, Jakobsson M, Coop G, et al. A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nat Genet* 2006;38(11):1251-60.
30. Consortium. IH. A haplotype map of the human genome. *Nature* 2005;437(7063):1299-320.
31. Daehmlow S, Erdmann J, Knueppel T, et al. Novel mutations in sarcomeric protein genes in dilated cardiomyopathy. *Biochem Biophys Res Commun* 2002;298(1):116-20.
32. Dec GW, Fuster V. Idiopathic dilated cardiomyopathy. *N Engl J Med* 1994;331(23):1564-75.

33. Doran P, Gannon J, O'Connell K, Ohlendieck K. Aging skeletal muscle shows a drastic increase in the small heat shock proteins alphaB-crystallin/HspB5 and cvHsp/HspB7. *Eur J Cell Biol* 2007;86(10):629-40.
34. Duboscq-Bidot L, Charron P, Ruppert V, et al. Mutations in the ANKRD1 gene encoding CARP are responsible for human dilated cardiomyopathy. *Eur Heart J* 2009;30(17):2128-36.
35. Duboscq-Bidot L, Xu P, Charron P, et al. Mutations in the Z-band protein myopalladin gene and idiopathic dilated cardiomyopathy. *Cardiovasc Res* 2008;77(1):118-25.
36. Elliott P, Andersson B, Arbustini E, et al. Classification of the cardiomyopathies: a position statement from the European Society Of Cardiology Working Group on Myocardial and Pericardial Diseases. *Eur Heart J* 2008;29(2):270-6.
37. Evans A, Salomaa V, Kulathinal S, et al. MORGAM (an international pooling of cardiovascular cohorts). *Int J Epidemiol* 2005;34(1):21-7.
38. Everly MJ. Cardiac transplantation in the United States: an analysis of the UNOS registry. *Clin Transpl* 2008:35-43.
39. Fatkin D, MacRae C, Sasaki T, et al. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med* 1999;341(23):1715-24.
40. Ferreira-Maldent N, Maillot F, Quilliet L, Guilmot JL, Charbonnier B. [Dilated cardiomyopathy, diabetes and deafness related to a mutation of mitochondrial DNA]. *Arch Mal Coeur Vaiss* 2007;100(2):149-52.
41. Franks PW. Identifying genes for primary hypertension: methodological limitations and gene-environment interactions. *J Hum Hypertens* 2009;23(4):227-37.
42. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. *Science* 2002;296(5576):2225-9.
43. Gamberdinger M, Hajieva P, Kaya AM, Wolfrum U, Hartl FU, Behl C. Protein quality control during aging involves recruitment of the macroautophagy pathway by BAG3. *Embo J* 2009;28(7):889-901.
44. Geier C, Perrot A, Ozcelik C, et al. Mutations in the human muscle LIM protein gene in families with hypertrophic cardiomyopathy. *Circulation* 2003;107(10):1390-5.
45. Gerull B, Gramlich M, Atherton J, et al. Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nat Genet* 2002;30(2):201-4.
46. Goerss JB, Michels VV, Burnett J, et al. Frequency of familial dilated cardiomyopathy. *Eur Heart J* 1995;16 Suppl O:2-4.
47. Goldbourt U, Neufeld HN. Genetic aspects of arteriosclerosis. *Arteriosclerosis* 1986;6(4):357-77.
48. Grisoni ML, Proust C, Alanne M, et al. Lack of association between polymorphisms of the IL18R1 and IL18RAP genes and cardiovascular risk: the MORGAM Project. *BMC Med Genet* 2009;10:44.
49. Grunig E, Tasman JA, Kucherer H, Franz W, Kubler W, Katus HA. Frequency and phenotypes of familial dilated cardiomyopathy. *J Am Coll Cardiol* 1998;31(1):186-94.
50. Gunderson KL, Steemers FJ, Ren H, et al. Whole-genome genotyping. *Methods Enzymol* 2006;410:359-76.

51. Hassel D, Dahme T, Erdmann J, et al. Nexilin mutations destabilize cardiac Z-disks and lead to dilated cardiomyopathy. *Nat Med* 2009;15(11):1281-8.
52. Hayashi T, Arimura T, Ueda K, et al. Identification and functional analysis of a caveolin-3 mutation associated with familial hypertrophic cardiomyopathy. *Biochem Biophys Res Commun* 2004;313(1):178-84.
53. Helgadóttir A, Thorleifsson G, Magnusson KP, et al. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet* 2008;40(2):217-24.
54. Herman DS, Lam L, Taylor MR, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* 2012;366(7):619-28.
55. Herrmann S, Schmidt-Petersen K, Pfeifer J, et al. A polymorphism in the endothelin-A receptor gene predicts survival in patients with idiopathic dilated cardiomyopathy. *Eur Heart J* 2001;22(20):1948-53.
56. Hershberger RE, Kushner JD, Parks SB. Dilated Cardiomyopathy Overview. 1993.
57. Hershberger RE, Kushner JD, Parks SB. Dilated Cardiomyopathy Overview. 2007.
58. Hershberger RE, Norton N, Morales A, Li D, Siegfried JD, Gonzalez-Quintana J. Coding sequence rare variants identified in MYBPC3, MYH6, TPM1, TNNC1, and TNNI3 from 312 patients with familial or idiopathic dilated cardiomyopathy. *Circ Cardiovasc Genet* 2010;3(2):155-61.
59. Hershberger RE, Siegfried JD. Update 2011: clinical and genetic issues in familial dilated cardiomyopathy. *J Am Coll Cardiol* 2011;57(16):1641-9.
60. Hiroi S, Harada H, Nishi H, Satoh M, Nagai R, Kimura A. Polymorphisms in the SOD2 and HLA-DRB1 genes are associated with nonfamilial idiopathic dilated cardiomyopathy in Japanese. *Biochem Biophys Res Commun* 1999;261(2):332-9.
61. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005;6(2):95-108.
62. Hoffmann B, Schmidt-Traub H, Perrot A, Osterziel KJ, Gessner R. First mutation in cardiac troponin C, L29Q, in a patient with hypertrophic cardiomyopathy. *Hum Mutat* 2001;17(6):524.
63. Holden MJ, Haynes RJ, Rabb SA, Satija N, Yang K, Blasic JR, Jr. Factors affecting quantification of total DNA by UV spectroscopy and PicoGreen fluorescence. *J Agric Food Chem* 2009;57(16):7221-6.
64. Holweg CT, Baan CC, Niesters HG, et al. TGF-beta1 gene polymorphisms in patients with end-stage heart failure. *J Heart Lung Transplant* 2001;20(9):979-84.
65. Homma S, Iwasaki M, Shelton GD, Engvall E, Reed JC, Takayama S. BAG3 deficiency results in fulminant myopathy and early lethality. *Am J Pathol* 2006;169(3):761-73.
66. Ichihara S, Yamada Y, Yokota M. Association of a G994-->T missense mutation in the plasma platelet-activating factor acetylhydrolase gene with genetic susceptibility to nonfamilial dilated cardiomyopathy in Japanese. *Circulation* 1998;98(18):1881-5.
67. Inagaki N, Hayashi T, Arimura T, et al. Alpha B-crystallin mutation in dilated cardiomyopathy. *Biochem Biophys Res Commun* 2006;342(2):379-86.
68. Jee SH, Suh I, Won SY, Kim MY. Familial correlation and heritability for cardiovascular risk factors. *Yonsei Med J* 2002;43(2):160-4.

69. Jefferies JL, Towbin JA. Dilated cardiomyopathy. *Lancet* 2010;375(9716):752-62.
70. Jung M, Poepping I, Perrot A, et al. Investigation of a family with autosomal dominant dilated cardiomyopathy defines a novel locus on chromosome 2q14-q22. *Am J Hum Genet* 1999;65(4):1068-77.
71. Kamisago M, Sharma SD, DePalma SR, et al. Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy. *N Engl J Med* 2000;343(23):1688-96.
72. Kimura A. Contribution of genetic factors to the pathogenesis of dilated cardiomyopathy: the cause of dilated cardiomyopathy: genetic or acquired? (genetic-side). *Circ J* 2011;75(7):1756-65; discussion 1765.
73. Kimura A, Harada H, Park JE, et al. Mutations in the cardiac troponin I gene associated with hypertrophic cardiomyopathy. *Nat Genet* 1997;16(4):379-82.
74. Kirov G, Nikolov I, Georgieva L, Moskvina V, Owen MJ, O'Donovan MC. Pooled DNA genotyping on Affymetrix SNP genotyping arrays. *BMC Genomics* 2006;7:27.
75. Knoll R, Hoshijima M, Hoffman HM, et al. The cardiac mechanical stretch sensor machinery involves a Z disc complex that is defective in a subset of human dilated cardiomyopathy. *Cell* 2002;111(7):943-55.
76. Knoll R, Postel R, Wang J, et al. Laminin-alpha4 and integrin-linked kinase mutations cause human cardiomyopathy via simultaneous defects in cardiomyocytes and endothelial cells. *Circulation* 2007;116(5):515-25.
77. Knuiman MW, Divitini ML, Welborn TA, Bartholomew HC. Familial correlations, cohabitation effects, and heritability for cardiovascular risk factors. *Ann Epidemiol* 1996;6(3):188-94.
78. Krajinovic M, Pinamonti B, Sinagra G, et al. Linkage of familial dilated cardiomyopathy to chromosome 9. Heart Muscle Disease Study Group. *Am J Hum Genet* 1995;57(4):846-52.
79. Krief S, Faivre JF, Robert P, et al. Identification and characterization of cvHsp. A novel human small stress protein selectively expressed in cardiovascular and insulin-sensitive tissues. *J Biol Chem* 1999;274(51):36592-600.
80. Kruglyak L. Power tools for human genetics. *Nat Genet* 2005;37(12):1299-300.
81. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409(6822):860-921.
82. Landstrom AP, Weisleder N, Batalden KB, et al. Mutations in JPH2-encoded junctophilin-2 associated with hypertrophic cardiomyopathy in humans. *J Mol Cell Cardiol* 2007;42(6):1026-35.
83. Lee JS, Choi M, Yan X, Lifton RP, Zhao H. On optimal pooling designs to identify rare variants through massive resequencing. *Genet Epidemiol* 2011;35(3):139-47.
84. Levitas A, Muhammad E, Harel G, et al. Familial neonatal isolated cardiomyopathy caused by a mutation in the flavoprotein subunit of succinate dehydrogenase. *Eur J Hum Genet* 2010;18(10):1160-5.
85. Li D, Parks SB, Kushner JD, et al. Mutations of presenilin genes in dilated cardiomyopathy and heart failure. *Am J Hum Genet* 2006;79(6):1030-9.
86. Li D, Tapscoft T, Gonzalez O, et al. Desmin mutation responsible for idiopathic dilated cardiomyopathy. *Circulation* 1999;100(5):461-4.

87. Liggett SB, Wagoner LE, Craft LL, et al. The Ile164 beta2-adrenergic receptor polymorphism adversely affects the outcome of congestive heart failure. *J Clin Invest* 1998;102(8):1534-9.
88. Lindquist S, Craig EA. The heat-shock proteins. *Annu Rev Genet* 1988;22:631-77.
89. Liu QR, Drgon T, Walther D, et al. Pooled association genome scanning: validation and use to identify addiction vulnerability loci in two samples. *Proc Natl Acad Sci U S A* 2005;102(33):11864-9.
90. Loh E, Rebbeck TR, Mahoney PD, DeNofrio D, Swain JL, Holmes EW. Common variant in AMPD1 gene predicts improved clinical outcome in patients with heart failure. *Circulation* 1999;99(11):1422-5.
91. Macgregor S, Zhao ZZ, Henders A, Nicholas MG, Montgomery GW, Visscher PM. Highly cost-efficient genome-wide association studies using DNA pools and dense SNP arrays. *Nucleic Acids Res* 2008;36(6):e35.
92. Mangin L, Charron P, Tesson F, et al. Familial dilated cardiomyopathy: clinical features in French families. *Eur J Heart Fail* 1999;1(4):353-61.
93. Maron BJ, Towbin JA, Thiene G, et al. Contemporary definitions and classification of the cardiomyopathies: an American Heart Association Scientific Statement from the Council on Clinical Cardiology, Heart Failure and Transplantation Committee; Quality of Care and Outcomes Research and Functional Genomics and Translational Biology Interdisciplinary Working Groups; and Council on Epidemiology and Prevention. *Circulation* 2006;113(14):1807-16.
94. Matkovich SJ, Van Booven DJ, Hindes A, et al. Cardiac signaling genes exhibit unexpected sequence diversity in sporadic cardiomyopathy, revealing HSPB7 polymorphisms associated with disease. *J Clin Invest* 2010;120(1):280-9.
95. Matsuzaki H, Loi H, Dong S, et al. Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. *Genome Res* 2004;14(3):414-25.
96. McNair WP, Ku L, Taylor MR, et al. SCN5A mutation associated with dilated cardiomyopathy, conduction disorder, and arrhythmia. *Circulation* 2004;110(15):2163-7.
97. McNamara DM, Holubkov R, Postava L, et al. Effect of the Asp298 variant of endothelial nitric oxide synthase on survival for patients with congestive heart failure. *Circulation* 2003;107(12):1598-602.
98. Meaburn E, Butcher LM, Liu L, et al. Genotyping DNA pools on microarrays: tackling the QTL problem of large samples and large numbers of SNPs. *BMC Genomics* 2005;6:52.
99. Messina DN, Speer MC, Pericak-Vance MA, McNally EM. Linkage of familial dilated cardiomyopathy with conduction defect and muscular dystrophy to chromosome 6q23. *Am J Hum Genet* 1997;61(4):909-17.
100. Mestroni L, Maisch B, McKenna WJ, et al. Guidelines for the study of familial dilated cardiomyopathies. Collaborative Research Group of the European Human and Capital Mobility Project on Familial Dilated Cardiomyopathy. *Eur Heart J* 1999;20(2):93-102.
101. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;16(3):1215.
102. Mizon-Gerard F, de Groote P, Lamblin N, et al. Prognostic impact of matrix metalloproteinase gene polymorphisms in patients with heart failure according to the aetiology of left ventricular systolic dysfunction. *Eur Heart J* 2004;25(8):688-93.

103. Mogensen J, Murphy RT, Shaw T, et al. Severe disease expression of cardiac troponin C and T mutations in patients with idiopathic dilated cardiomyopathy. *J Am Coll Cardiol* 2004;44(10):2033-40.
104. Mohapatra B, Jimenez S, Lin JH, et al. Mutations in the muscle LIM protein and alpha-actinin-2 genes in dilated cardiomyopathy and endocardial fibroelastosis. *Mol Genet Metab* 2003;80(1-2):207-15.
105. Moulik M, Vatta M, Witt SH, et al. ANKRD1, the gene encoding cardiac ankyrin repeat protein, is a novel dilated cardiomyopathy gene. *J Am Coll Cardiol* 2009;54(4):325-33.
106. Mullen MP, Creevey CJ, Berry DP, et al. Polymorphism discovery and allele frequency estimation using high-throughput DNA sequencing of target-enriched pooled DNA samples. *BMC Genomics* 2012;13:16.
107. Murphy RT, Mogensen J, Shaw A, Kubo T, Hughes S, McKenna WJ. Novel mutation in cardiac troponin I in recessive idiopathic dilated cardiomyopathy. *Lancet* 2004;363(9406):371-2.
108. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 1997;349(9064):1498-504.
109. Neaton JD, Blackburn H, Jacobs D, et al. Serum cholesterol level and mortality findings for men screened in the Multiple Risk Factor Intervention Trial. Multiple Risk Factor Intervention Trial Research Group. *Arch Intern Med* 1992;152(7):1490-500.
110. Nicaud V, Francomme C, Ruidavets JB, et al. Lack of association between complement factor H polymorphisms and coronary artery disease or myocardial infarction. *J Mol Med (Berl)* 2007;85(7):771-5.
111. Niimura H, Patton KK, McKenna WJ, et al. Sarcomere protein gene mutations in hypertrophic cardiomyopathy of the elderly. *Circulation* 2002;105(4):446-51.
112. Olson TM, Illenberger S, Kishimoto NY, Huttelmaier S, Keating MT, Jockusch BM. Metavinculin mutations alter actin interaction in dilated cardiomyopathy. *Circulation* 2002;105(4):431-7.
113. Olson TM, Kishimoto NY, Whitby FG, Michels VV. Mutations that alter the surface charge of alpha-tropomyosin are associated with dilated cardiomyopathy. *J Mol Cell Cardiol* 2001;33(4):723-32.
114. Olson TM, Michels VV, Thibodeau SN, Tai YS, Keating MT. Actin mutations in dilated cardiomyopathy, a heritable form of heart failure. *Science* 1998;280(5364):750-2.
115. OMS. Report of the WHO/ISFC task force on the definition and classification of cardiomyopathies. *Br Heart J* 1980;44(6):672-3.
116. Osio A, Tan L, Chen SN, et al. Myozenin 2 is a novel gene for human hypertrophic cardiomyopathy. *Circ Res* 2007;100(6):766-8.
117. Out AA, van Minderhout IJ, Goeman JJ, et al. Deep sequencing to reveal new variants in pooled DNA samples. *Hum Mutat* 2009;30(12):1703-12.
118. Pearson JV, Huentelman MJ, Halperin RF, et al. Identification of the genetic basis for complex disorders by use of pooling-based genomewide single-nucleotide-polymorphism association studies. *Am J Hum Genet* 2007;80(1):126-39.
119. Poetter K, Jiang H, Hassanzadeh S, et al. Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. *Nat Genet* 1996;13(1):63-9.

120. Richardson P, McKenna W, Bristow M, et al. Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of cardiomyopathies. *Circulation* 1996;93(5):841-2.
121. Roger VL, Go AS, Lloyd-Jones DM, et al. Heart disease and stroke statistics--2011 update: a report from the American Heart Association. *Circulation* 2011;123(4):e18-e209.
122. Schmitt JP, Kamisago M, Asahi M, et al. Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban. *Science* 2003;299(5611):1410-3.
123. Schonberger J, Kuhler L, Martins E, Lindner TH, Silva-Cardoso J, Zimmer M. A novel locus for autosomal-dominant dilated cardiomyopathy maps to chromosome 7q22.3-31.1. *Hum Genet* 2005;118(3-4):451-7.
124. Schonberger J, Levy H, Grunig E, et al. Dilated cardiomyopathy and sensorineural hearing loss: a heritable syndrome that maps to 6q23-24. *Circulation* 2000;101(15):1812-8.
125. Schonberger J, Seidman CE. Many roads lead to a broken heart: the genetics of dilated cardiomyopathy. *Am J Hum Genet* 2001;69(2):249-60.
126. Schonberger J, Wang L, Shin JT, et al. Mutation in the transcriptional coactivator EYA4 causes dilated cardiomyopathy and sensorineural hearing loss. *Nat Genet* 2005;37(4):418-22.
127. Schrauwen I, Ealy M, Huentelman MJ, et al. A genome-wide analysis identifies genetic variants in the RELN gene associated with otosclerosis. *Am J Hum Genet* 2009;84(3):328-38.
128. Schunkert H, Konig IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* 2011;43(4):333-8.
129. Seidman JG, Seidman C. The genetic basis for cardiomyopathy: from mutation identification to mechanistic paradigms. *Cell* 2001;104(4):557-67.
130. Sham P, Bader JS, Craig I, O'Donovan M, Owen M. DNA Pooling: a tool for large-scale association studies. *Nat Rev Genet* 2002;3(11):862-71.
131. Shaw T, Elliott P, McKenna WJ. Dilated cardiomyopathy: a genetically heterogeneous disease. *Lancet* 2002;360(9334):654-5.
132. Shea J, Agarwala V, Philippakis AA, et al. Comparing strategies to fine-map the association of common SNPs at chromosome 9p21 with type 2 diabetes and myocardial infarction. *Nat Genet* 2011;43(8):801-805.
133. Sile S, Velez DR, Gillani NB, Alexander CA, George AL, Jr., Williams SM. Haplotype diversity in four genes (CLCNKA, CLCNKB, BSND, NEDD4L) involved in renal salt reabsorption. *Hum Hered* 2008;65(1):33-46.
134. Singer VL, Jones LJ, Yue ST, Haugland RP. Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation. *Anal Biochem* 1997;249(2):228-38.
135. Skibola CF, Bracci PM, Halperin E, et al. Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. *Nat Genet* 2009;41(8):873-5.
136. Smith NL, Felix JF, Morrison AC, et al. Association of genome-wide variation with the risk of incident heart failure in adults of European and African ancestry: a prospective meta-analysis from the cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. *Circ Cardiovasc Genet* 2010;3(3):256-66.
137. Stark K, Esslinger UB, Reinhard W, et al. Genetic association study identifies HSPB7 as a risk gene for idiopathic dilated cardiomyopathy. *PLoS Genet* 2010;6(10):e1001167.

138. Stary HC, Chandler AB, Dinsmore RE, et al. A definition of advanced types of atherosclerotic lesions and a histological classification of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association. *Arterioscler Thromb Vasc Biol* 1995;15(9):1512-31.
139. Steemers FJ, Gunderson KL. Whole genome genotyping technologies on the BeadArray platform. *Biotechnol J* 2007;2(1):41-9.
140. Sylvius N, Tesson F, Gayet C, et al. A new locus for autosomal dominant dilated cardiomyopathy identified on chromosome 6q12-q16. *Am J Hum Genet* 2001;68(1):241-6.
141. Taylor MR, Carniel E, Mestroni L. Cardiomyopathy, familial dilated. *Orphanet J Rare Dis* 2006;1:27.
142. Taylor MR, Slavov D, Gajewski A, et al. Thymopoietin (lamina-associated polypeptide 2) gene mutation associated with dilated cardiomyopathy. *Hum Mutat* 2005;26(6):566-74.
143. Thiene G, Corrado D, Basso C. Cardiomyopathies: is it time for a molecular classification? *Eur Heart J* 2004;25(20):1772-5.
144. Thomas DC, Haile RW, Duggan D. Recent developments in genomewide association scans: a workshop summary and review. *Am J Hum Genet* 2005;77(3):337-45.
145. Tiret L, Mallet C, Poirier O, et al. Lack of association between polymorphisms of eight candidate genes and idiopathic dilated cardiomyopathy: the CARDIGENE study. *J Am Coll Cardiol* 2000;35(1):29-35.
146. Tiret L, Poirier O, Nicaud V, et al. Heterogeneity of linkage disequilibrium in human genes has implications for association studies of common diseases. *Hum Mol Genet* 2002;11(4):419-29.
147. Towbin JA, Bowles NE. Genetic abnormalities responsible for dilated cardiomyopathy. *Curr Cardiol Rep* 2000;2(5):475-80.
148. Towbin JA, Bowles NE. Sarcoglycan, the heart, and skeletal muscles: new treatment, old drug? *J Clin Invest* 2001;107(2):153-4.
149. Towbin JA, Hejtmancik JF, Brink P, et al. X-linked dilated cardiomyopathy. Molecular genetic evidence of linkage to the Duchenne muscular dystrophy (dystrophin) gene at the Xp21 locus. *Circulation* 1993;87(6):1854-65.
150. Tregouet DA, Escolano S, Tiret L, Mallet A, Golmard JL. A new algorithm for haplotype-based association analysis: the Stochastic-EM algorithm. *Ann Hum Genet* 2004;68(Pt 2):165-77.
151. Tregouet DA, Garelle V. A new JAVA interface implementation of THESIAS: testing haplotype effects in association studies. *Bioinformatics* 2007;23(8):1038-9.
152. Tsubata S, Bowles KR, Vatta M, et al. Mutations in the human delta-sarcoglycan gene in familial and sporadic dilated cardiomyopathy. *J Clin Invest* 2000;106(5):655-62.
153. Vasile VC, Edwards WD, Ommen SR, Ackerman MJ. Obstructive hypertrophic cardiomyopathy is associated with reduced expression of vinculin in the intercalated disc. *Biochem Biophys Res Commun* 2006;349(2):709-15.
154. Vatta M, Mohapatra B, Jimenez S, et al. Mutations in Cypher/ZASP in patients with dilated cardiomyopathy and left ventricular non-compaction. *J Am Coll Cardiol* 2003;42(11):2014-27.
155. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science* 2001;291(5507):1304-51.

156. Vesel S, Stopar-Obreza M, Trebusak-Podkrajsek K, Jazbec J, Podnar T, Battelino T. A novel mutation in the G4.5 (TAZ) gene in a kindred with Barth syndrome. *Eur J Hum Genet* 2003;11(1):97-101.
157. Villard E, Perret C, Gary F, et al. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J*;32(9):1065-76.
158. Wang T, Pradhan K, Ye K, Wong LJ, Rohan TE. Estimating allele frequency from next-generation sequencing of pooled mitochondrial DNA samples. *Front Genet* 2011;2:51.
159. Watkins H, Conner D, Thierfelder L, et al. Mutations in the cardiac myosin binding protein-C gene on chromosome 11 cause familial hypertrophic cardiomyopathy. *Nat Genet* 1995;11(4):434-7.
160. Wencker D, Chandra M, Nguyen K, et al. A mechanistic role for cardiac myocyte apoptosis in heart failure. *J Clin Invest* 2003;111(10):1497-504.
161. WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447(7145):661-78.
162. Yasuno K, Bilguvar K, Bijlenga P, et al. Genome-wide association study of intracranial aneurysm identifies three new risk loci. *Nat Genet* 2010;42(5):420-5.
163. Zak I, Niemiec P, Balcerzyk A, Krauze J. Combined "pro-atherosclerotic" variants of the ACE and APOE genes increase the risk of the coronary artery disease associated with the presence of cigarette smoking. *Acta Cardiol* 2008;63(6):741-7.
164. Zeggini E, Rayner W, Morris AP, et al. An evaluation of HapMap sample size and tagging SNP performance in large-scale empirical and simulated data sets. *Nat Genet* 2005;37(12):1320-2.
165. Zhang D, Mott JL, Farrar P, et al. Mitochondrial DNA mutations activate the mitochondrial apoptotic pathway and cause dilated cardiomyopathy. *Cardiovasc Res* 2003;57(1):147-57.
166. Zhao ZZ, Nyholt DR, James MR, Mayne R, Treloar SA, Montgomery GW. A comparison of DNA pools constructed following whole genome amplification for two-stage SNP genotyping designs. *Twin Res Hum Genet* 2005;8(4):353-61.