

2013-2014

Thèse
Pour le
Diplôme d'État de Docteur en Pharmacie

**LA SNP ARRAY APPLIQUEE
A UNE SELECTION DE PATIENTS
AU CHU D'ANGERS**

NOUVELLE METHODE
D'ANALYSE DE L'ADN

Aurélie DESGUES COUSIN

Sous la direction du
Professeur Christophe VERNY

Membres du jury
Gérald Larcher | Professeur
Christophe Verny | Professeur
Dominique Bonneau | Professeur
Céline Bris | Assistante hospitalo-universitaire

Soutenu publiquement le :
16 décembre 2014

L'auteur du présent document vous autorise à le partager, reproduire, distribuer et communiquer selon les conditions suivantes :




- Vous devez le citer en l'attribuant de la manière indiquée par l'auteur (mais pas d'une manière qui suggérerait qu'il approuve votre utilisation de l'œuvre).
- Vous n'avez pas le droit d'utiliser ce document à des fins commerciales.
- Vous n'avez pas le droit de le modifier, de le transformer ou de l'adapter.

Consulter la licence creative commons complète en français : <http://creativecommons.org/licences/by-nc-nd/2.0/fr/>

ENGAGEMENT DE NON PLAGIAT

Je, soussigné(e) **Aurélié DESGUES COUSIN**
déclare être pleinement conscient(e) que le plagiat de documents ou d'une
partie d'un document publiée sur toutes formes de support, y compris l'internet,
constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée.
En conséquence, je m'engage à citer toutes les sources que j'ai utilisées
pour écrire ce rapport ou mémoire.

Le 5 novembre 2014,



**Cet engagement de non plagiat
doit être signé et joint
à tous les rapports, dossiers, mémoires.**

Présidence de l'université
40 rue de rennes – BP 73532
49035 Angers cedex
Tél. 02 41 96 23 23 | Fax 02 41 96 23 00

Année Universitaire 2014-2015

Liste des enseignants

Département Pharmacie

PROFESSEURS

Disciplines

BENOIT Jean-Pierre	Pharmacotechnie - Biopharmacie
DUVAL Olivier	Chimie Thérapeutique
ÉVEILLARD Matthieu	Bactériologie - Virologie
FAURE Sébastien	Pharmacologie
JARDEL Alain	Physiologie
LAGARCE Frédéric	Pharmacotechnie-Biopharmacie
LARCHER Gérald	Biochimie
MARCHAIS Véronique	Bactériologie - Virologie
PASSIRANI Catherine	Chimie générale – Chimie analytique
RICHOMME Pascal	Pharmacognosie
SAULNIER Patrick	Biophysique pharmaceutique et biostatistiques
SERAPHIN Denis	Chimie Organique
VENIER Marie-Claire	Pharmacotechnie - Biopharmacie

PAST

Disciplines

BRUNA Étienne	Industrie
---------------	-----------

MAITRES DE CONFERENCES

Disciplines

ANNAIX Véronique	Biochimie Générale et Clinique
BAGLIN Isabelle	Pharmaco - Chimie
BASTIAT Guillaume	Biophysique – biostatistiques -Rhéologie
BENOIT Jacqueline	Pharmacologie et Pharmacocinétique
CLERE Nicolas	Physiologie - Pharmacologie
DERBRÉ Séverine	Pharmacognosie
FLEURY Maxime	Immunologie
GUILLET David	Chimie Analytique

MAITRES DE CONFERENCES

HELESBEUX Jean-Jacques
LANDREAU Anne
MALLET Marie-Sabine
MAROT Agnès
PECH Brigitte
RIOU Jérémie
ROGER Émilie
SCHINKOVITZ Andréas
TRICAUD Anne

Disciplines

Chimie Organique
Botanique
Chimie Analytique et Bromatologie
Parasitologie et Mycologie médicale
Pharmacotechnie
Biostatistiques
Pharmacotechnie
Pharmacognosie
Biologie Cellulaire

A.H.U.

BRIS Céline
SPIESSER-ROBELET Laurence

Disciplines

Biochimie
Pharmacie clinique et Éducation Thérapeutique

PRC (Professeurs certifiés affectés dans l'enseignement supérieur)

BRUNOIS-DEBU Isabelle

Anglais

ATER (Assistants Enseignement Supérieur et Recherche).

BOISARD Séverine
DESHAYES Caroline
RODIER Marion
VERRIER Julie

Chimie analytique
Bactériologie
Pharmacologie
Parasitologie et mycologie médicale

Département ISSBA

PROFESSEURS

BOURY Franck
CALENDA Alphonse
MAHAZA Chetaou
MAURAS Geneviève

Disciplines

Biophysique
Biologie Moléculaire - Biotechnologie
Bactériologie - Virologie
Biologie Cellulaire

MAITRES DE CONFERENCES

BATAILLE Nelly
BILLAUD Sandrine
BONNIN Marie
CALVIGNAC Brice
DUBREUIL Véronique
FAISANT Nathalie
GIRAUD Sandrine
OGER Jean-Michel
RICHOMME Anne-Marie

Disciplines

Biologie Cellulaire et Moléculaire
Immunologie - Parasitologie
Management intégré / qualité logistique
Génie des procédés bioindustries
Chimie Analytique
Génie des produits industriels
Biologie moléculaire et cellulaire
Chimie
Valorisation des substances naturelles

PRAG (Professeurs Agrégés)

HANOTTE Caroline
ROUX Martine

Disciplines

Economie – Gestion
Espagnol

PRCE (Professeurs certifiés affectés dans l'enseignement supérieur)

LECOMTE Stéphane
MEENTS Ulrike

Disciplines

Anglais
Allemand

PAST

BERGER Virginie
BLOUIN Laurence
COLLE Stéphane

DELOUIS Anne-Laure
MATHIEU Éric
NORMAND Yves
POURIAS Marie-Annick
VERBORG Soisik

Disciplines

Sureté de fonctionnement des études cliniques
Management des structures des soins
Prévention des risques innovation et conception
HQS du bâti
Prévention des risques et sécurité
Ingénierie de projets dans les domaines de santé
Systèmes d'information santé
Projets professionnels – Formation continue
Management – Qualité

ATER (Assistants Enseignement Supérieur et Recherche).

MARTINEZ Émilie

Disciplines

Biologie et Physiologie de la nutrition

REMERCIEMENTS

La réalisation de cette thèse constitue une étape clé dans ma vie. Elle a commencé par mon stage dans le service de neurogénétique du CHU d'Angers au début de l'année 2012. Elle s'est poursuivie à Massy, à Marseille et prendra fin à Angers lors de la soutenance orale, à la fin de l'année 2014. Durant ces trois années, le déroulement des événements a été assez mouvementé mais je n'ai cessé d'avancer, jour après jour. Cela a été possible grâce à un grand nombre de personnes, de la vie professionnelle à la vie privée.

Si je suis aujourd'hui celle que je suis, c'est grâce en premier lieu aux personnes qui m'ont accompagnée depuis toujours, qui m'ont éduquée, qui m'ont appris ce qu'ils savaient, qui m'ont soutenue, qui m'ont donnée les moyens de réaliser des études. Je tiens donc à adresser mes premiers remerciements à mes parents ! Merci !

Au quotidien, une thèse n'est pas facile à vivre. Elle soulève de nombreux doutes. Elle entraîne parfois stress, mauvaise humeur, déceptions... Mais c'est aussi une source d'énergie, de joie, d'envie de réussir. Le passage d'une émotion à l'autre est quelque chose qui n'est pas facile à gérer, en particulier pour mon mari. Malgré cela, il n'a jamais cessé de me soutenir, de croire en moi et de m'aider à avancer. Il est à mes côtés depuis plus de onze années et de notre amour est né Erwen en décembre 2013. Ils sont ma joie de vivre et me donnent tous les jours l'envie d'aller plus loin. Aux hommes de ma vie, Merci !

Cette thèse a pris naissance dans le service de neurogénétique du CHU d'Angers. C'est le Professeur Verny qui m'a proposé ce sujet, original pour une thèse d'exercice en pharmacie, mais très intéressant et enrichissant d'un point de vue scientifique, médical. Il a su m'orienter et me guider dans ce travail de recherche. Je lui adresse donc un grand Merci !

Cette étude a impliqué de nombreux professionnels de santé. Dans un premier temps, ce sont les personnels du service de neurogénétique qui m'ont accueillie dans leurs locaux : Audrey, Marie, Marie-Anne. Elles m'ont fait participer à la vie du service et m'ont

REMERCIEMENTS

appris beaucoup sur les pathologies neurogénétiques et leurs prises en charge. Elles ont mis à ma disposition leurs bureaux lors de la consultation des nombreux dossiers patients. Merci ! Dans un second temps, ce sont des médecins qui ont contribué à mon travail de recherche. Le Professeur Bonneau a contribué au processus de sélection finale des dossiers. Le Dr Guichet a réalisé les analyses génétiques au sein du laboratoire du CHU d'Angers et m'a apporté de nombreux éléments sur la méthode utilisée, l'interprétation des résultats, la recherche des gènes candidats. Je leur dit Merci pour le temps qu'ils m'ont accordé !

Un grand Merci également au Professeur Larcher qui a accepté de m'encadrer et d'être mon référent au sein de la faculté de pharmacie. Il m'a consacré du temps, en particulier pour l'élaboration du plan du manuscrit et pour la relecture de ce dernier.

Je tiens à remercier aussi toutes les personnes qui m'entourent. Je pense tout d'abord à mes sœurs qui comptent énormément pour moi et avec qui on passe toujours de bons moments ! Je pense aussi au reste de ma famille, ma marraine, mes oncles, mes tantes, mes cousins et cousines ainsi qu'à ma belle famille. Je n'oublie pas mes amis, en particulier Charlotte, Sarah, Mary, Coline, Jessica qui sont là dans les bons moments comme dans les difficultés. A tous, Merci !

Table des matières

1	Introduction.....	1
2	Les bases de la génétique.....	3
2.1.	Le matériel génétique.....	4
2.1.1.	Structure primaire de l'ADN.....	4
2.1.2.	Structure secondaire de l'ADN.....	6
2.1.3.	Structure du chromosome	7
2.1.4.	L'ADN nucléaire.....	9
2.1.5.	L'ADN mitochondrial.....	9
2.2.	La réplication de l'ADN.....	10
2.2.1.	L'ADN polymérase.....	11
2.2.2.	Mécanisme de réplication.....	11
a)	L'étape d'initiation.....	11
b)	L'étape d'élongation.....	12
c)	L'étape de terminaison.....	13
2.2.3.	Risque d'erreur.....	13
2.2.4.	Lésions.....	13
2.2.5.	Exemple de l'anémie falciforme.....	15
2.2.6.	Quels sont les agents mutagènes ?	15
a)	Les agents chimiques.....	15
b)	Les agents intercalants.....	16
c)	Les agents analogues de bases.....	16
d)	Les agents physiques.....	16
2.2.7.	Réparation.....	16
2.3.	La transcription.....	17
2.3.1.	L'étape d'initiation.....	18
2.3.2.	L'étape d'élongation.....	18
2.3.3.	L'étape de terminaison.....	19
2.3.4.	La maturation des produits de transcription.....	19
a)	La formation de la coiffe ou capping.....	19
b)	L'épissage ou splicing.....	19
c)	La polyadénylation.....	20
2.3.5.	La régulation de la transcription.....	20
2.4.	La traduction.....	20
2.4.1.	Le code génétique.....	21
2.4.2.	Les molécules nécessaires à la traduction.....	22
a)	L'ARN messager, ARNm.....	22
b)	L'ARN de transfert, ARNt.....	22
c)	Les ARN ribosomiaux, ARNr.....	23
2.4.3.	L'étape d'initiation	23
2.4.4.	L'étape d'élongation.....	24

2.4.5.	L'étape de terminaison.....	24
2.4.6.	Altération de la traduction.....	25
2.4.7.	Action de certains médicaments sur la traduction.....	25
2.5.	Les principes de base de l'hérédité.....	25
a)	Maladie de Charcot Marie Tooth (CMT).....	27
b)	Chorée de Huntington.....	28
c)	Ataxie de Freidreich.....	28
d)	Neuropathie optique héréditaire de Leber.....	28
e)	Maladie de Steinert ou dystrophie myotonique de type I	28
3	Projet.....	29
3.1.	Matériel et méthodes.....	30
3.1.1.	Le principe d'hybridation de l'ADN.....	30
a)	Principe général de l'hybridation.....	30
b)	L'hybridation en pratique.....	30
c)	L'hybridation in situ fluorescente ou FISH.....	31
d)	Objectifs de l'hybridation.....	31
3.1.2.	La CGH array.....	31
a)	Origine et principe.....	31
b)	Les points forts de la CGH array.....	32
c)	Recueil des échantillons d'ADN.....	33
d)	Hybridation.....	33
e)	Applications.....	34
f)	Limites de la méthode.....	34
g)	Perspectives.....	34
3.1.3.	La SNP array.....	34
a)	Les points forts de la SNP array.....	35
b)	Principe.....	35
c)	Applications.....	36
d)	Limites de la méthode.....	36
3.1.4.	La sélection des patients.....	36
a)	Le suivi des patients au CHU d'Angers	37
b)	Consultation des dossiers.....	39
c)	Classification des patients	39
d)	Critères d'exclusion des dossiers.....	40
4	Résultats.....	44
4.1.	Les résultats quantitatifs.....	44
4.1.1.	La classification des patients.....	44
a)	Les pathologies les plus représentées.....	45
b)	Les pathologies les moins représentées.....	46
c)	Les pathologies très peu représentées.....	47
d)	La classification des autres patients.....	48

4.1.2.	L'exclusion des patients.....	48
4.2.	Les dossiers sélectionnés.....	51
4.2.1.	Patient A1.....	52
a)	Présentation du patient.....	52
b)	Résultats de l'analyse par SNP array.....	53
4.2.2.	Patient C3.....	53
a)	Présentation du patient.....	53
b)	Résultats de l'analyse par SNP array.....	53
4.2.3.	Patient D4.....	53
a)	Présentation du patient.....	53
b)	Résultats de l'analyse par SNP array.....	54
4.2.4.	Patient E6.....	54
a)	Présentation du patient.....	54
b)	Résultats de l'analyse par SNP array.....	54
4.2.5.	Patient F7.....	54
a)	Présentation du patient.....	54
b)	Résultats de l'analyse par SNP array.....	55
4.2.6.	Patient G8.....	55
a)	Présentation du patient.....	55
b)	Résultats de l'analyse par SNP array.....	55
4.2.7.	Patient G9.....	55
a)	Présentation du patient.....	55
b)	Résultats de l'analyse par SNP array.....	56
4.2.8.	Patient H10.....	56
a)	Présentation du patient.....	56
b)	Résultats de l'analyse par SNP array.....	56
4.2.9.	Patient I12.....	56
a)	Présentation du patient.....	56
b)	Résultats de l'analyse par SNP array.....	57
4.2.10.	Les régions LOH communes entre les patients G8 et G9.....	57
5	Discussion et conclusion.....	59

Index des figures

Figure 1: Structure simplifiée d'une cellule.....	3
Figure 2: Molécules de ribose et de désoxyribose.....	4
Figure 3: Les bases puriques.....	5
Figure 4: Les bases pyrimidiques.....	5
Figure 5: Molécule d'acide phosphorique.....	5
Figure 6: Liaison phosphodiester entre deux nucléotides.....	6
Figure 7: La double hélice de l'ADN.....	6
Figure 8: La double hélice enroulée autour des protéines histones.....	8
Figure 9: Les différents types de chromosomes.....	8
Figure 10: La mitochondrie.....	9
Figure 11: Schéma global de la réplication de l'ADN.....	10
Figure 12: Tableau de correspondance entre les codons et les acides aminés.....	21
Figure 13: Structure en trèfle de l'ARN de transfert.....	22
Figure 14: Structure du ribosome Eucaryote.....	23
Figure 15: Schéma d'une hybridation fluorescente (FISH).....	31
Figure 16: Exemple d'un arbre génétique.....	38
Figure 17: Exclusion des dossiers aboutissant à la sélection finale des patients.....	50

Index des tableaux

Tableau 1 : Extrait du tableau listant la totalité des patients.....	39
Tableau 2 : Classification 1 des patients.....	40
Tableau 3 : Classification 2 des patients.....	40
Tableau 4 : Gènes impliqués dans les CMT au CHU d'Angers.....	45
Tableau 5 : Récapitulatif des pathologies les plus représentées au CHU d'Angers....	46
Tableau 6 : Récapitulatif des pathologies les moins représentées au CHU d'Angers.	47
Tableau 7 : Listing des 14 patients sélectionnés.....	51

Index des annexes

Annexe I	Résultat de l'étude de l'équipe de M. Tayeh en 2009
Annexe II	Résultat de la SNP array du patient A1
Annexe III	Résultat de la SNP array du patient C3
Annexe IV	Résultat de la SNP array du patient D4
Annexe V	Résultat de la SNP array du patient E6
Annexe VI	Résultat de la SNP array du patient F7
Annexe VII 1/2	Résultat de la SNP array du patient G8
Annexe VII 2/2	Résultat de la SNP array du patient G8
Annexe VIII 1/2	Résultat de la SNP array du patient G9
Annexe VIII 2/2	Résultat de la SNP array du patient G9
Annexe IX	Résultat de la SNP array du patient H10
Annexe X	Résultat de la SNP array du patient I12

Avant-propos

Une page se tourne... Le document qui suit constitue l'aboutissement de six années d'études au sein de la faculté de pharmacie d'Angers. C'est au cours de ma cinquième année hospitalo-universitaire que j'ai choisi ce sujet avec l'aide du Professeur Verny, chef du service de neurogénétique du CHU d'Angers. En effet, au début de l'année 2012, j'ai passé trois mois au sein de ce service, j'ai eu l'opportunité d'accompagner l'équipe au quotidien et j'ai appris beaucoup au sujet des maladies génétiques et neurodégénératives, notamment leur diagnostics, leurs évolutions, leurs prises en charge. C'est un domaine peu connu par les pharmaciens et c'est pourquoi j'ai souhaité réaliser ma thèse dans ce domaine, qui, de plus, m'intéresse particulièrement.

Le Professeur Verny, après en avoir discuté avec le Professeur Bonneau, responsable du Département Biochimie et Génétique, m'a proposé de travailler sur la CGH array, l'*Hybridation Génomique Comparative* sur réseau d'ADN, et plus particulièrement, sur la SNP array qui est pratiquée au sein du laboratoire d'analyses du CHU d'Angers. J'ai accepté de relever le défi ! Il s'agit d'un sujet original pour une thèse d'exercice en pharmacie mais c'est une aventure avant tout scientifique et qui reste dans le domaine médical. J'espère vous transmettre mon envie de découverte de ce domaine très intrigant qu'est la génétique.

1 Introduction

Le service de neurogénétique du CHU d'Angers contient un grand nombre de dossiers médicaux parmi lesquels certains correspondent à des maladies qualifiées de génétiques mais dont l'origine n'est pas identifiée à ce jour. La nouvelle méthode d'hybridation de l'ADN¹ pratiquée au laboratoire de l'hôpital, la SNP² array, offre la possibilité à certains patients de voir leur pathologie identifiée sur le plan génétique après l'échec des analyses déjà employées. Cette méthode d'analyse apparaît donc aujourd'hui, pour eux, comme une piste sérieuse dans l'élaboration d'un diagnostic. Il va de soi que l'ensemble des dossiers ne peut être traité. Il faut sélectionner les patients pour qui cette méthode a le plus de chance de donner des résultats. Une démarche de sélection longue et rigoureuse est donc obligatoire. Cette dernière, après consultation de l'ensemble des dossiers médicaux du service d'Angers, a permis la mise en exergue d'un certain nombre de cas cliniques susceptibles d'être candidats à l'analyse par SNP array.

Durant les dix dernières années, le domaine de la génétique a considérablement évolué. Les procédés exceptionnels de séquençage de l'ADN, coûteux et réservés aux laboratoires de pointe il y a encore quelques temps, se sont démocratisés et sont donc à présent accessibles à un plus large comité. En parallèle, des techniques nouvelles et spécialisées ont vu le jour. C'est au sein de ce groupement de méthodes novatrices que l'on peut citer les dernières : l'*Hybridation Génomique Comparative* sur réseau d'ADN ou *CGH array* et la *Single Nucleotide Polymorphism array* ou *SNP array*.

Ces nouvelles approches par hybridation sont au centre des travaux présentés ici. Signe de l'intérêt et des perspectives prometteuses attendues, la CGH array fait l'objet de nombreuses études. Les premiers travaux sur le sujet sont publiés par Choi et al. [1]. Ces derniers démontrent l'utilité et la faisabilité de la technique dans le but d'établir des diagnostics de maladies génétiques. Suite à ces recherches, de nombreuses autres équipes vont faire de cette procédure l'une des méthodes phares dans l'élaboration de diagnostics complexes. On peut citer les travaux de John W. Cole en 2012 [2], M. H.

1 **ADN** : acide désoxyribonucléique

2 **SNP** : Single Nucleotide Polymorphism

Chahrour en 2012 [3] et A. Jimenez-Escrig et al. en 2011 [4]. Ce rôle d'outil de diagnostic est abordé par Ku CS et al. en 2012 [5].

La technique consiste en l'analyse d'une très faible partie (1.5 % environ) du programme génétique d'un individu. Ce cœur d'ADN, baptisé exome¹, constitue « l'unique » partie codante du génome. Le matériel étudié étant réduit, cela diminue d'autant le temps de l'analyse. L'utilisation de cette méthode se prête tout particulièrement aux études des maladies génétiques héréditaires. En effet, cette nouvelle approche a permis d'effectuer des diagnostics de maladies génétiques au sein de corps familiaux restreints, voire d'individus isolés – ce qui n'était pas réalisable via les anciens protocoles d'identification de mutations –. On lui reconnaît tout de même quelques points faibles. En effet, les mutations répétées ne sont pas accessibles. De plus, il est difficile d'identifier le nombre de copies réalisées au niveau d'un gène.

Ce manuscrit s'articule autour de trois sections. Dans un premier temps, en section 2, nous aborderons les bases de la génétique. Comment se structure le matériel génétique ? Quels mécanismes subit-il ? Comment se transmet-il ? Dans un second temps, en section 3, nous détaillerons la méthode d'analyse utilisée ainsi que le protocole de sélection de dossiers qui a été suivi. Enfin, dans un troisième temps, en section 4, nous discuterons les résultats des analyses et nous verrons les perspectives pour les patients.

1 **Exome** : ensemble des exons, parties codantes de l'ADN

2 Les bases de la génétique

La structure des acides nucléiques et leur implication dans le transfert d'informations chez les êtres vivants ont été découverts par des chercheurs de l'Université de Cambridge en 1953 : Francis Harry Compton Crick, James Dewey Watson et Maurice Hugh Frederik Wilkins. Cette immense découverte a été publiée dans la célèbre revue *Nature* et leur a donné l'opportunité de remporter le prix Nobel de médecine et de physiologie en 1962 [6].

Mais à quoi correspondent ces acides nucléiques ? Ce sont des macromolécules présentes dans toutes les cellules vivantes. On les retrouve soit sous leur forme libre soit sous leur forme combinée à des protéines et dans ce cas, elles forment des nucléoprotéines. Leur importance dans le monde vivant est capitale ! En effet, ces acides sont soit le support de l'information génétique, soit les agents permettant et régulant l'expression de cette information [7]. Il en existe deux types : l'ADN (ou DNA), Acide DésoxyriboNucléique et l'ARN (ou RNA), Acide Ribonucléique. L'ADN est séparé du cytoplasme, l'intérieur de la cellule, par une membrane que l'on appelle l'enveloppe nucléaire : il est donc cloisonné dans le noyau cellulaire, comme on peut le voir sur la figure 1 ci-après. L'ADN est l'élément constitutif des chromosomes, que l'on détaillera plus loin. On le retrouve également dans les mitochondries, qui sont des organites du cytoplasme des cellules. On l'appelle dans ce cas l'ADN mitochondrial. L'ARN, quant à lui, est essentiellement retrouvé au niveau du cytoplasme des cellules.

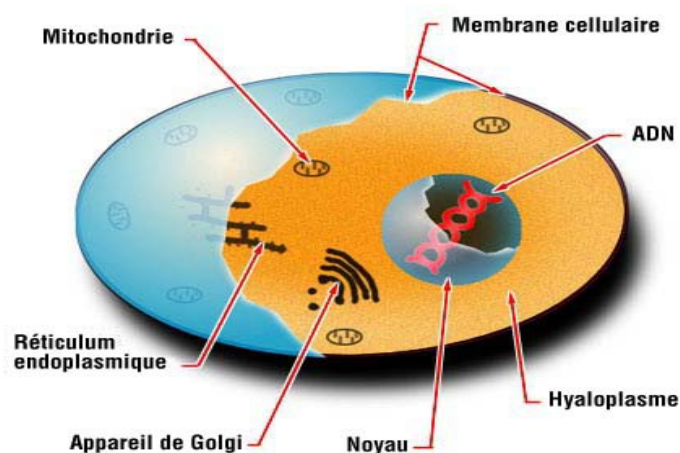


Figure 1: Structure simplifiée d'une cellule

Dans une première partie, nous aborderons la définition du matériel génétique, en détaillant bien évidemment l'ADN, de la constitution du nucléotide à la formation des chromosomes. Dans un second temps, nous verrons la réplication de l'ADN, mécanisme d'une importance capitale dans la transmission de l'information génétique. Puis nous verrons les mécanismes qui suivent logiquement la réplication, c'est-à-dire la transcription et la traduction. Ces étapes permettent l'expression de l'information génétique, afin de faire fonctionner la cellule. Enfin, seront traités les principes de bases de l'hérédité, soit comment l'information génétique est-elle transmise de génération en génération ?

2.1. Le matériel génétique

Le matériel génétique est constitué de molécules organiques assemblées de façon structurée donnant plusieurs niveaux d'organisation. La structure primaire correspond à l'enchaînement des nucléotides. Puis, les brins s'associent deux à deux en double hélice afin de constituer la structure secondaire. Puis, la double hélice d'ADN subit plusieurs compactages avant de former les chromosomes.

2.1.1. Structure primaire de l'ADN

La structure primaire désigne la séquence des sous-unités qui composent la molécule d'ADN. Dans le cas présent, les sous-unités sont des molécules présentant une structure particulière que l'on appelle les nucléotides. Un nucléotide est constitué de trois éléments principaux : le pentose, la base azotée et l'acide phosphorique.

Le pentose est un sucre constitué de cinq atomes de carbones et présentant une structure spécifique. Dans l'ADN, le pentose est un α -2'-désoxy-D-ribose alors que dans l'ARN, il s'agit d'un α -D-ribose. Leurs structures sont représentées dans la figure 2 ci-après.

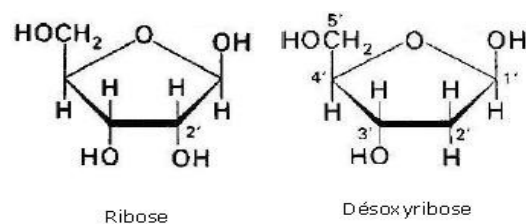


Figure 2: Molécules de ribose et de désoxyribose.

<http://www.cnrs.fr/insb/communication/glossaire/desoxyriboseillustr.htm>

Les bases ont, quant à elles, des structures hétérocycliques azotées. Elles dérivent pour certaines de la purine et pour d'autres de la pyrimidine. Il existe principalement deux bases puriques (dérivant de la purine, contenant un noyau imidazole) : l'adénine et la guanine, détaillées dans la figure 3.

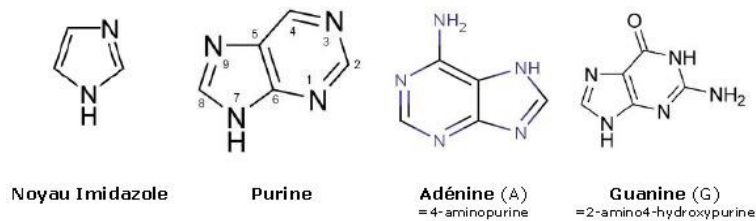


Figure 3: Les bases puriques.

http://fdanieau.free.fr/cours/bts/A1/biochimie/chapitre8/chap8_AcidesNucleiques1.php

Les bases pyrimidiques (dérivant de la pyrimidine) sont la cytosine et la thymine dans l'ADN et la cytosine et l'uracile dans l'ARN, détaillées dans la figure 4.

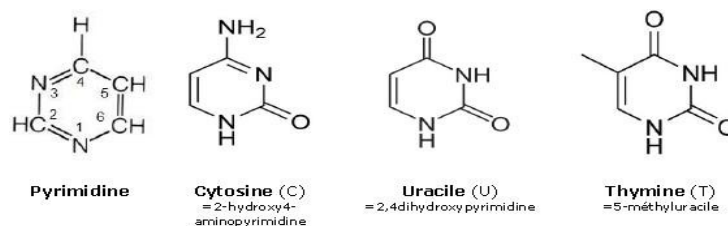


Figure 4: Les bases pyrimidiques.

http://fdanieau.free.fr/cours/bts/A1/biochimie/chapitre8/chap8_AcidesNucleiques1.php

Les bases sont symbolisées par leurs initiales : A, T(U), G et C.

L'acide phosphorique, représenté dans la figure 5, a pour rôle d'apporter l'énergie nécessaire à l'assemblage des nucléotides.

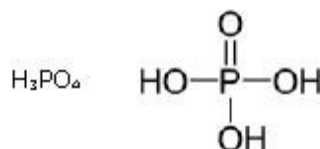


Figure 5: Molécule d'acide phosphorique.

http://fdanieau.free.fr/cours/bts/A1/biochimie/chapitre8/chap8_AcidesNucleiques1.php

Afin de former des séquences, les nucléotides sont reliés entre eux par des liaisons phosphodiester, qui s'établissent entre les atomes de carbone de deux nucléotides, plus précisément entre le carbone numéroté 3' de l'un et le carbone numéroté 5' de l'autre, comme le montre la figure 6. C'est à dire que l'enchaînement est constitué d'une

alternance de désoxyriboses et de phosphates, les bases se retrouvant ainsi « libres ». L'enchaînement est donc ordonné et orienté [8].

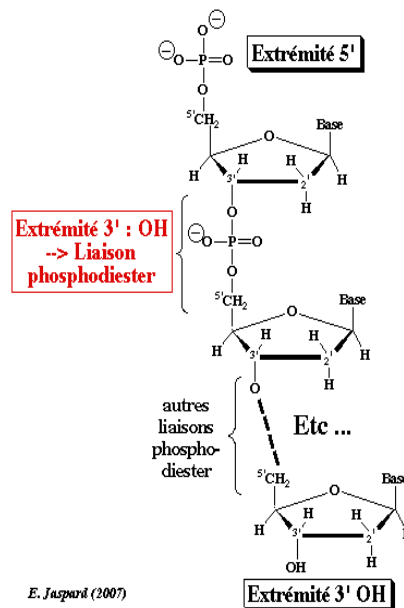


Figure 6: Liaison phosphodiester entre deux nucléotides.

<http://biochimej.univ-angers.fr/Page2/COURS/9ModulGenFoncVeg/5MethEtudGenFonc/1MethodeSEQUENGAGE/1SEQUENGAGE.htm>

2.1.2. Structure secondaire de l'ADN

Les nucléotides sont liés les uns aux autres pour former les polymères que l'on appelle des chaînes ou des brins d'ADN. Deux à deux, ces chaînes vont s'enrouler l'une autour de l'autre pour former ce que l'on appelle communément la double hélice de l'ADN, comme illustré en figure 7.

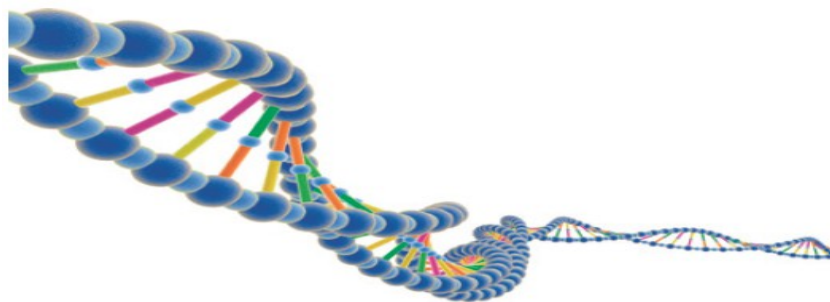


Figure 7: La double hélice de l'ADN.

<http://www.citedesgeometries.org/newcdg/de-lautre-cote-du-miroir-etre-soi-meme-ou-pas/>

Les deux brins d'ADN sont complémentaires. En effet, on retrouve toujours l'adénine en face de la thymine et la cytosine en face de la guanine. Cette

complémentarité permet d'associer face à face une base purique et une base pyrimidique, afin de les « emboîter » de façon optimale au niveau de la double hélice.

Aussi, les brins sont antiparallèles, c'est-à-dire que leur orientation, l'un en face de l'autre, est inversée. Cela est dû aux contraintes stériques – c'est à dire dans l'espace – imposées par la configuration spatiale des nucléotides. Le brin 5' → 3' est donc positionné face à un brin orienté dans le sens inverse, c'est-à-dire dans le sens 3' → 5'.

De plus, ils présentent une structure hélicoïdale. L'ADN s'enroule suivant une double hélice droite. Les deux brins sont associés grâce à des liaisons hydrogènes qui s'établissent entre les nucléotides qui se trouvent face à face, à l'intérieur de l'hélice. D'autres forces interviennent telles que des interactions hydrophobes et électroniques entre les bases.

2.1.3. Structure du chromosome

Le chromosome est une entité particulière. Sa taille est de l'ordre du μm alors que l'ADN déroulé a une taille de l'ordre du mètre. Ceci implique donc de compacter la molécule d'ADN de base afin de la faire tenir dans le noyau de la cellule.

La plupart du temps, les chromosomes se trouvent sous forme décondensée, ils sont dans ce cas très difficiles à observer. La condensation se produit au moment de la division cellulaire et c'est à cet instant qu'il est possible de les étudier car c'est là qu'ils sont le mieux visualisés.

Plusieurs mécanismes agissent sur les chromosomes dans le cadre du processus de compactage. En effet, l'ADN se lie à un certain nombre de protéines. On retrouve deux catégories de protéines qui se lient aux chromosomes pour former la chromatine : les histones et les protéines chromosomiques non histones. L'histone est un octamère, qui est donc composé de huit protéines histones, soit deux molécules de chacune des protéines suivantes : H2A, H2B, H3 et H4. Une partie de la double hélice s'enroule autour des histones, comme représenté dans la figure 8.

Les histones sont présents en très grande quantité et conduisent à la formation du premier niveau d'empaquetage de la chromatine : le nucléosome, qui a une taille d'environ 10 nm de diamètre. Les nucléosomes s'empilent pour aboutir à une structure encore plus dense : la fibre, qui a une taille de 30 nm de diamètre, qui elle-même pourra se condenser. Il y a donc plusieurs niveaux d'empaquetage de l'ADN [9].

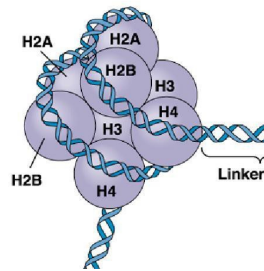


Figure 8: La double hélice enroulée autour des protéines histones.

<http://www.vdsclences.com/pages/sciences-biologiques/biologie-moleculaire/biologie-moleculaire.html>

Un chromosome fonctionnel a une forme de bâtonnet et est constitué de trois éléments essentiels : un centromère, une paire de télomères et des origines de réplication [10].

- Le centromère correspond au point d'attache des microtubules, filaments responsables du mouvement des chromosomes lors de la division cellulaire. Il apparaît comme une région resserrée. Il permet à chacune des copies de chaque chromosome dupliqué (forme de X) d'être dirigée de façon équilibrée vers chacune des deux cellules filles. On classe les chromosomes selon quatre types : métacentriques, submétacentriques, acrocentriques et télacentriques, voir la figure 9. Les bras courts des chromosomes sont désignés par la lettre p alors que les bras longs sont désignés par la lettre q.

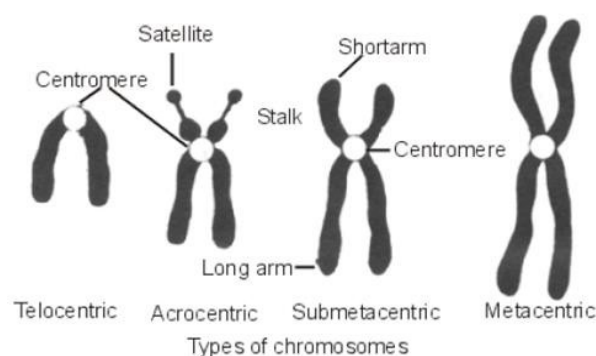


Figure 9: Les différents types de chromosomes.

<http://expertadn.fr/les-genes/>

- Les télomères sont les régions terminales des chromosomes, ils ont un rôle de protection ainsi que de stabilisation des extrémités du chromosome. Si un chromosome casse, découvrant une extrémité sans télomère, il sera détruit.

- Les origines de réplication sont des sites où la synthèse d'ADN pourra commencer. Elles correspondent à des séquences spécifiques qui seront reconnues comme origines de réplication.

2.1.4. L'ADN nucléaire

L'ADN nucléaire est linéaire et s'organise dans le noyau des cellules avec des protéines afin de former la chromatine. Comme nous l'avons vu précédemment, ces protéines jouent un rôle dans l'empaquetage et le contrôle des longues chaînes d'ADN. C'est lorsque cette chromatine se condense que l'on obtient la structure allongée et compacte des chromosomes. L'être humain n'est pas l'organisme qui possède l'ADN le plus long. En effet, certaines espèces végétales ont jusqu'à trente fois plus d'ADN que l'Homme.

L'ADN nucléaire est constitué de 46 chromosomes répartis en 23 paires, qui sont regroupés sous le nom de caryotype humain [9].

L'ADN nucléaire compte environ 3 milliards de paires de bases (pb) et 25 000 gènes. Les gènes représentent 30% du génome. 50% du génome correspond à des séquences intergéniques qui sont des séquences répétitives non codantes et/ou mobiles. Le reste se répartit notamment en zones de régulation. Sur l'ensemble des gènes, 90% codent pour des protéines et 10% pour des ARN [8].

2.1.5. L'ADN mitochondrial

L'ADN mitochondrial, ADN_m, désigne le matériel génétique contenu dans les mitochondries. La mitochondrie, représentée dans la figure 10, est l'« élément du cytoplasme de la cellule animale ou végétale dont le rôle essentiel est d'assurer l'oxydation, la respiration cellulaire, la mise en réserve de l'énergie par la cellule et le stockage de certaines substances » [11].

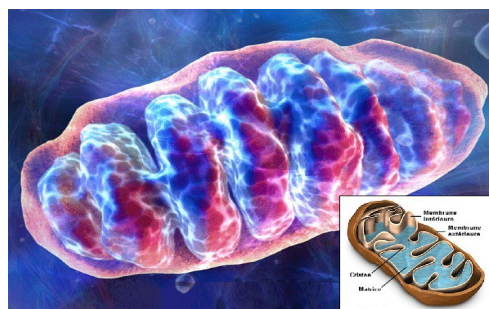


Figure 10: La mitochondrie.

La mitochondrie possède elle-même toute la machinerie ainsi que tous les éléments nécessaires à la synthèse des protéines. L'ADNm est particulier car il présente une structure circulaire. Il compte 16 569 pb et 37 gènes [8]. De plus, il provient uniquement de l'ADNm de la mère car au moment de la fécondation, le gamète masculin ne transmet pas de mitochondries. Ce sont celles de l'ovocyte fécondé qui seront utilisées par l'embryon. Ceci explique le fait que certaines maladies, dites mitochondriales, présentent une transmission exclusivement maternelle, comme par exemple le syndrome de MELAS.

2.2. La réplication de l'ADN

La biosynthèse des ADN est essentielle à la multiplication des cellules car elle permet de transmettre l'information génétique de la cellule mère aux cellules filles. C'est donc un mécanisme qui permet de transmettre et de conserver intact le matériel génétique d'une génération à l'autre. Il s'agit d'un processus complexe et primordial ! En effet, deux cellules filles doivent absolument être identiques entre elles et identiques à la cellule mère. La copie doit être fidèle et intégrale. Elle repose sur l'appariement de bases de façon complémentaire. On parle de réplication semi-conservative car chaque nouvelle molécule d'ADN va contenir un brin de l'ancien ADN et un brin nouvellement formé, synthétisé.

De nombreuses enzymes interviennent, dont l'ADN polymérase, qui permet de dupliquer les deux brins d'ADN en synthétisant les nouveaux brins. Les processus qui permettent de répliquer, de séparer et de répartir les deux copies d'ADN dans deux cellules filles, sont regroupés sous le terme « cycle cellulaire ». Le mécanisme de réplication est illustré en figure 11.

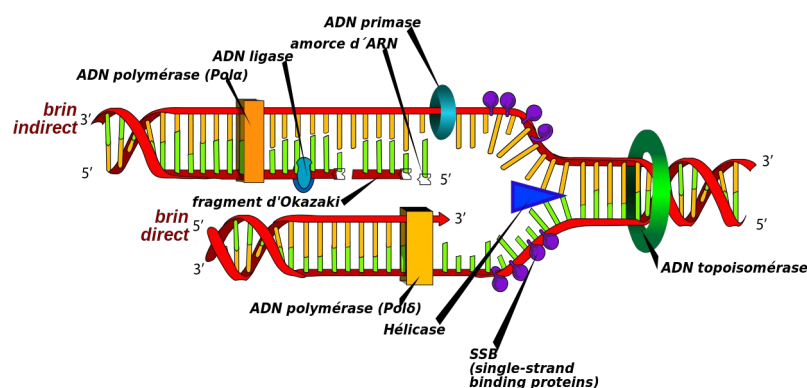


Figure 11: Schéma global de la réplication de l'ADN

2.2.1. L'ADN polymérase

Elle fut isolée en 1958 par Kornberg chez la bactérie *E.Coli* et par Bollum dans le foie de rat [12]. Par la suite, les ADN polymérases ont été trouvées dans toutes les cellules dans lesquelles on a pu observer une synthèse d'ADN.

Cette enzyme possède trois substrats :

- la chaîne d'ADN qui est copiée,
- la chaîne néosynthétisée ou chaîne en croissance,
- le 2'-désoxyribonucléotide-5'-triphosphate qui va être incorporé, obligatoirement complémentaire du nucléotide situé en face de lui sur la matrice.

Au niveau chimique, cette enzyme catalyse la formation de la liaison phosphodiester entre deux nucléotides.

2.2.2. Mécanisme de réplication

Comme expliqué précédemment, la réplication est un processus qui a pour objectif final la multiplication des cellules en transmettant un matériel génétique identique aux cellules produites, soit de générations en générations. Pour cela, elle doit réaliser une copie de l'ADN de la cellule mère, autrement dit, doubler la quantité d'ADN pour pouvoir répartir entre les deux cellules filles. La réplication de l'ADN peut être divisée schématiquement en trois étapes principales : l'étape d'initiation, l'étape d'élongation et l'étape de terminaison. Pour qu'elle puisse se réaliser, il est nécessaire que la double hélice de l'ADN, normalement très stable, puisse s'ouvrir localement et que cette ouverture soit maintenue pendant le processus. On parle de fourche de réplication.

a) L'étape d'initiation

Cette étape d'initiation est celle pendant laquelle les amorces sont synthétisées au niveau des origines de réplication. Une amorce est une séquence de nucléotides complémentaire produite à partir de la séquence de l'origine de réplication. Les amorces ainsi fabriquées seront par la suite allongées par les ADN polymérases afin de dupliquer l'ensemble de l'ADN. Cette initiation à la réplication se passe schématiquement en trois phases qui sont détaillées juste après : la reconnaissance de l'origine de réplication, l'ouverture de la double hélice au niveau de l'origine de réplication et la synthèse de l'amorce.

En premier lieu, il y a reconnaissance de l'origine de réplication par un complexe protéique, nommé ORC (Origin Recognition Complex). Les origines de réplication se distinguent du reste des brins d'ADN par des séquences nucléotidiques particulières. L'ADN humain contient environ 10 000 origines de réplication [9].

L'ouverture de la double hélice est ensuite réalisée au niveau de cette origine de réplication grâce à une enzyme ATP dépendante (Adénosine TriPhosphate dépendante) que l'on appelle une hélicase. Elle va se déplacer sur un brin d'ADN dans le sens 5' → 3' et supprimer les liaisons hydrogènes qui ont pour rôle de maintenir la double hélice dans sa configuration spatiale. Le complexe ORC devient le « complexe de pré-réplication ».

Enfin, il y a le recrutement de nouvelles protéines qui vont permettre de former le complexe de « pré-initiation ». Une fois les deux brins séparés, ils vont être maintenus à distance l'un de l'autre par des protéines appelées SSB (Single Strand DNA Binding) qui se fixent sur chacune des deux chaînes par des liaisons électrostatiques. La synthèse de l'amorce peut alors débuter. Elle nécessite une autre enzyme, une ARN polymérase appelée primase, qui aura pour mission de synthétiser une amorce, court fragment d'ARN (4 à 12 nucléotides), qui pourra s'apparier au brin parental.

La réplication débute au niveau, non pas d'une, mais de plusieurs origines de réplication (séquences d'initiation de réplication ou ARS (Autonom Replication System)) en même temps, localisées le long du chromosome, afin de le répliquer en totalité dans une durée moindre.

b) L'étape d'élongation

L'ADN polymérase peut alors se mettre en place au niveau de la fourche de réplication afin de commencer la « copie » des brins. Elle respectera bien sûr la complémentarité des bases. L'ADN polymérase peut entrer en action et réaliser la prolongation à partir d'une amorce seulement.

Il est important de souligner la différence de réplication entre les deux brins. En effet, la fourche de réplication avance dans un sens donné sur la double hélice. Or, les deux brins ne sont pas orientés dans le même sens étant donné qu'ils sont complémentaires. On parle alors de « brin avancé » et de « brin retardé ».

La synthèse s'effectuant toujours dans le sens 5' → 3', le brin retardé va être répliqué de manière discontinue, contrairement au brin avancé. C'est-à-dire que de

petits fragments d'ADN vont être synthétisés « à reculons ». Il s'agit de chaînes nucléotidiques qui comptent 1000 à 2000 résidus et qui ont été baptisées « fragments d'Okazaki ». Ces fragments seront ensuite associés, liés entre eux. C'est pourquoi l'on parle de réplication bi-directionnelle. La fourche de réplication est donc une entité asymétrique.

c) L'étape de terminaison

A la fin de la réplication, les amorces d'ARN sont hydrolysées par une enzyme spécifique : la RNase H et elles sont remplacées par des fragments d'ADN synthétisés par l'ADN polymérase, qui vient donc terminer son travail.

La réplication s'achève grâce à un autre type d'enzymes, les télomérases, qui ont pour rôle d'insérer des nucléotides à l'extrémité 3'.

C'est la polymérase qui se chargera de compléter la chaîne au niveau de l'extrémité 5'.

2.2.3. Risque d'erreur

Les polymérases sont des enzymes rapides puisqu'elles sont capables d'ajouter environ 1000 nucléotides à la seconde, avec un taux d'erreur très faible de $1/10^5$ nucléotides incorporés. Mais il faut tenter de limiter au maximum les erreurs de réplication, c'est pourquoi elles seront corrigées.

Ces enzymes possèdent en effet une activité de relecture du dernier nucléotide inséré. L'erreur est détectée car elle entraîne un retard dans l'avancée de la polymérase.

Lorsqu'une erreur est décelée, le nucléotide sera retiré et remplacé par le bon grâce à l'activité exonucléasique $3' \rightarrow 5'$ de ces enzymes. Cela permet de diminuer le taux d'erreur à $1/10^9$ [8].

2.2.4. Lésions

Malgré les systèmes de relecture et de correction lors de la réplication, de rares erreurs peuvent « passer entre les mailles du filet » et entraîner des lésions permanentes de l'ADN que l'on appelle des mutations.

Une mutation correspond à une erreur de copie : le brin fils n'est plus la copie exacte du brin parental. On parle de mutation ponctuelle lorsqu'elle ne touche qu'un seul nucléotide.

On distingue plusieurs types de mutations :

- les substitutions : le nucléotide inséré n'est pas celui attendu.
 - on appelle transition une mutation qui implique l'inversion de deux bases puriques ou de deux bases pyrimidiques.
 - on appelle transversion une mutation qui implique l'inversion entre une base purique et une base pyrimidique.
- les oublis : mutation par délétion.
- les ajouts : mutation par insertion.

Une modification d'un nucléotide peut entraîner par la suite un décalage de lecture lors de la traduction en protéine.

- Mutations qui n'entraînent pas de décalage lors de la lecture :
 - Mutations silencieuses : la modification du nucléotide ne change pas l'acide aminé qui sera lu. Ce type de mutation n'a pas de conséquence sur la protéine qui sera produite.
 - Mutations conservatives : l'acide aminé lu sera différent mais aura des propriétés proches de l'acide aminé qui aurait dû être lu. Les conséquences sont souvent insignifiantes dans ce cas.
 - Mutation non sens : la modification du nucléotide entraîne un arrêt de lecture lors de la production de la protéine. Elle ne sera donc pas produite en totalité.
 - Mutation faux sens : l'acide aminé inséré est très différent de celui qui aurait dû être lu.
- Mutations qui entraînent un décalage lors de la lecture :

Il s'agit dans ce cas de mutations de type insertion ou délétion. Cela entraîne un important décalage qui va modifier tous les acides aminés qui seront lus après la mutation. La protéine produite sera très différente de la protéine qui aurait dû être synthétisée.

Une mutation peut donc avoir de profondes conséquences, même lorsqu'elle n'affecte qu'un seul nucléotide, en particulier lorsqu'elle touche une séquence d'ADN qui porte des informations vitales.

2.2.5. Exemple de l'anémie falciforme

La modification permanente d'un nucléotide peut conduire la cellule à produire une hémoglobine avec une séquence en acides aminés incorrecte, ce qui est à l'origine de l'anémie falciforme. Les hématies formées sont beaucoup plus fragiles et se détruisent très facilement dans le flux sanguin. Cela aboutit à un nombre d'hématies circulantes plus faible et cela fait de cette pathologie une maladie potentiellement mortelle.

Cet exemple illustre l'importance de la protection des cellules, notamment des cellules reproductives, les cellules germinales, contre les mutations [9].

2.2.6. Quels sont les agents mutagènes ?

Un agent mutagène désigne une substance qui augmente le taux d'erreurs lors de la réplication de l'ADN.

Il est possible de lister quatre classes d'agents mutagènes :

- les agents chimiques.
- les agents intercalants.
- les agents analogues de bases.
- les agents physiques.

a) Les agents chimiques

Il s'agit de substances qui peuvent entraîner des modifications chimiques sur certaines bases. Nous pouvons citer par exemple :

- La désamination de la cytosine en uracile en milieu aqueux. Lors de la réplication, la cytosine devenue uracile sera appariée à une adénine : la paire initiale C-G est remplacée par la paire U-A puis T-A.
- La désamination de l'adénine par l'acide nitreux. Elle sera transformée en hypoxanthine capable de s'apparier à une cytosine : la paire initiale A-T est remplacée par la paire H-C puis G-C.
- Les radicaux libres (ions superoxydes, radicaux hydroxydes...) peuvent oxyder la guanine en 7,8-dihydro-8-oxo-guanine, appelée oxo-guanine, qui peut s'apparier aussi bien avec de l'adénine qu'avec de la cytosine. Lors de la réplication suivante, l'oxo-guanine s'apparie avec de l'adénine, puis, cette adénine s'apparie avec de la thymine : la paire initiale G-C est remplacée par une paire T-A.

b) Les agents intercalants

Ces agents sont des composés cycliques et plans qui ont la possibilité de s'insérer entre les plans des bases de l'ADN. Cela aboutit à des délétions ou insertions de bases.

Par exemple, ont été identifiés comme agents intercalants le bromure d'éthidium ou BET, la proflavine et le psoralène.

c) Les agents analogues de bases

On connaît principalement le 5-bromo-uracile qui se comporte comme la thymine, c'est-à-dire qu'il peut s'apparier avec l'adénine. Mais, l'équilibre de tautomérisation transforme cette molécule en forme éno. Lors de la réplication suivante, cette forme s'appariera avec la guanine qui elle-même s'appariera avec une cytosine. La paire initiale T-A est remplacée par une paire C-G.

d) Les agents physiques

Sont principalement en cause les radiations telles que les rayons X ou les UV, ultraviolets, qui sont absorbés par l'ADN, notamment au niveau des bases pyrimidiques.

2.2.7. Réparation

La machinerie répliquative empêche généralement l'apparition d'erreurs dans la copie de l'ADN. Néanmoins, des erreurs peuvent subsister. Chaque fois qu'une erreur de copie apparaît, il y a donc un nucléotide mal apparié dans la séquence nouvellement synthétisée. S'il était laissé sans correction, ce mésappariement aurait comme résultat une mutation permanente dans le cycle suivant de réplication.

C'est pourquoi l'erreur est repérée par un complexe protéique spécialisé dans la réparation des mésappariements. Ce complexe excise le brin d'ADN lésé et synthétise à nouveau la séquence manquante.

La réparation des mésappariements joue un rôle important dans la prévention des cancers. En effet, des mutations des gènes codant pour les protéines de réparation entraînent une susceptibilité héréditaire à certains cancers (notamment les cancers du côlon). Les individus qui héritent d'une copie endommagée de ce gène ne présente aucun symptôme jusqu'à ce que leur copie indemne du gène soit accidentellement endommagée dans une cellule somatique (cellule non germinale, non sexuelle). Une cellule dont le système de réparation est altéré présente un plus grand risque devenir cancéreuse. Finalement, l'héritage d'un gène muté prédispose un individu au cancer [9].

L'ADN peut être endommagé de bien d'autres manières, qui nécessitent des mécanismes de réparation différents. Ces lésions sont expliquées plus en détails dans la partie qui précède et qui cite les agents mutagènes.

Le système de base de réparation de l'ADN comprend trois étapes :

- Dans un premier temps, l'ADN endommagé est reconnu et excisé, notamment grâce à des nucléases. Cette étape est réalisée par des enzymes spécifiques du type de la lésion décelée.
- Dans un second temps, la lacune est comblée par une ADN polymérase de réparation qui insère le nucléotide complémentaire au brin parental.
- Dans un troisième temps, une fois la lacune comblée, il reste une cassure dans le squelette sucre-phosphate du brin réparé. C'est une ADN ligase qui se charge de réparer cette cassure.

Les deux dernières étapes sont presque les mêmes pour tous les types de réparation de l'ADN.

Les cassures double brins sont plus compliquées à réparer car il n'y a pas de matrice intacte pouvant servir de base à la synthèse. Des mécanismes spécifiques de réparation existent dans ce cas là. Le plus commun est le mécanisme appelé « jonction non homologue des extrémités », qui rapproche les deux extrémités de la brèche pour qu'une ligase puisse rétablir la continuité des deux brins. Mais cela entraîne souvent la perte de nucléotides au niveau du site de réparation. Une autre solution existe : la recombinaison homologue. Le principe de ce processus est l'échange d'informations génétiques entre deux paires de molécules d'ADN homologues, c'est à dire entre des ADN double brins qui sont similaires ou identiques de part leurs séquences nucléotidiques. Dans ce cas de figure, c'est l'ADN double brin intact qui sert de matrice pour réparer le double brin cassé [9].

2.3. La transcription

La transcription constitue la première étape dans la synthèse des protéines et sera suivie par la traduction qui sera abordée dans la partie suivante. Il s'agit du processus par lequel une cellule va copier un gène en ARN, soit en une séquence nucléotidique lui correspondant. Cette étape se déroule à l'intérieur du noyau (pour les cellules eucaryotes). De la même façon que dans le mécanisme de réplication, on observe 3

phases lors de la transcription, qui seront détaillées un peu plus loin: la phase d'initiation, la phase d'élongation et la phase de terminaison.

On distingue différents types de gènes, codant pour différents types d'ARN cellulaires. L'ARN messager est synthétisé à partir de gènes qui utilisent l'ARN polymérase II et qui codent pour des protéines. C'est ce mécanisme qui est décrit ci-dessous. Aussi, il est important de souligner que le gène sera transcrit dans un premier temps en ARN prémessager, qui subira lui-même une maturation avant d'être exporté vers le cytosol, soit à l'extérieur du noyau. On peut noter que l'ARN messager procaryote ne subit pas de maturation car il est traduit quasiment instantanément [13]. Les polymérases I et III transcrivent quant à elles des gènes qui codent pour des ARN de transfert, qui ont un rôle structural et catalytique dans la cellule.

2.3.1. L'étape d'initiation

L'ARN polymérase II est associée à des facteurs généraux de transcription appelés TF_{II} (Transcription Factor). Ces facteurs ont pour objectif de se fixer sur des régions spécifiques de l'ADN pour positionner l'ARN polymérase et séparer les deux brins d'ADN. On dénombre quatre régions spécifiques, la plus retrouvée étant la « TATA box », qui a pour séquence : TATAAT. Cette TATA box a un rôle essentiel car elle contribue à l'initiation du processus.

Les différents facteurs se fixent dans un ordre précis afin d'accéder au « complexe d'initiation de la transcription ».

2.3.2. L'étape d'élongation

Cette étape consiste à synthétiser l'ARN à partir du complexe de transcription en lisant l'ADN dans le sens $5' \rightarrow 3'$. Elle est réalisée par l'ARN polymérase qui ajoute un à un les ribonucléotides, en respectant encore une fois la complémentarité des bases. Elle n'a pas besoin d'amorce pour commencer à synthétiser, contrairement à l'ADN polymérase.

L'ARN nouvellement synthétisé ne reste pas fixé au brin d'ADN matrice par des liaisons hydrogènes, contrairement à ce qu'il se passe lors de la réplication. La libération immédiate du brin d'ARN nouvellement synthétisé permet la production de nombreuses copies d'ARN à partir d'un même gène, et ce, en très peu de temps [9].

2.3.3. L'étape de terminaison

L'ARN polymérase synthétise l'ARN jusqu'à ce qu'elle arrive au terminateur de transcription, dont la structure est à ce jour encore mal connue. L'ARN prémessager nouvellement synthétisé est clivé avant la fin de la synthèse et est polyadénylé par une enzyme, la polyA polymérase, au niveau du site de polyadénylation. C'est un processus de maturation.

2.3.4. La maturation des produits de transcription

L'ADN prémessager subit une maturation qui se fait en même temps que sa synthèse. La maturation se décline en trois processus : la formation de la coiffe, l'épissage des exons et la polyadénylation.

a) La formation de la coiffe ou *capping*

Le capping consiste en l'ajout d'un résidu guanylique, une guanine sur laquelle est fixé un groupement méthyle, à l'extrémité 5'P libre de l'ARN messenger. On appelle ce résidu la « coiffe ». Celle-ci a plusieurs rôles. En premier lieu, elle permet de protéger l'extrémité de cet ARN messenger des attaques des Rnases (enzymes dégradant les ARN). En second lieu, la coiffe permettra la reconnaissance de l'extrémité 5' par le ribosome pour l'initiation de la traduction, qui sera détaillée par la suite. L'extrémité 5' de l'ARN messenger peut subir ensuite plusieurs méthylations, notamment sur la position 7 du résidu guanylique rajouté, par une guanine-7-méthyltransférase.

b) L'épissage ou *splicing*

Chez le procaryote, la quasi totalité des séquences génomiques sont informatives. A contrario, chez l'eucaryote, moins de 10% de l'ADN semble être exprimé.

En effet, une grande partie de l'ADN consiste en répétitions de séquences qui ne sont pas traduites. On va distinguer deux types de séquences qui s'alternent sur l'ADN eucaryote : les séquences non codantes, intermédiaires, appelées introns (50 à 200 000 nucléotides) et les séquences codantes, exprimées, appelées exons (environ 50 à 200 nucléotides) [7].

Les sites d'épissage sont localisés aux jonctions introns/exons. Il est important de trouver les « bonnes » bornes et il faut que l'excision des introns se fasse de manière précise car une erreur d'un nucléotide pourrait avoir de graves conséquences sur la lecture future de l'ARN messenger.

Le repérage des bornes est effectué par des protéines spécifiques. L'épissage consiste en 2 étapes de trans-estérification afin d'aboutir à l'élimination des introns et donc à l'accolement des exons. Le spliceosome est le complexe qui permet ce mécanisme. L'épissage peut se faire de différentes façons et entraîner la production de protéines différentes à partir de la même séquence d'ADN de départ. On parle dans ce cas d'épissage alternatif.

c) La polyadénylation

Cette étape s'effectue avant même que l'ARN polymérase n'arrive au terminateur. Des endonucléases vont couper l'ARN puis plusieurs centaines d'adénosines vont être ajoutées à l'extrémité 3'-OH de l'ARN. Cette extrémité polyadénylée va permettre de protéger l'ARN des attaques de nucléases et interviendra par la suite dans le processus de traduction.

Les ARN messagers vont ensuite être exportés vers le cytoplasme de la cellule par le complexe de pores nucléaires, qui sélectionnent uniquement les ARN messagers matures. Pour être prêt à l'exportation, un ARN messager doit être lié à un ensemble de protéines [9].

2.3.5. La régulation de la transcription

Deux catégories de gènes peuvent être distinguées :

- Les gènes constitutifs, exprimés de la même manière tout au long de la vie de la cellule. Ces gènes sont indispensables au fonctionnement de la cellule. Ils ne subissent pas de processus de régulation de leur expression.
- Les gènes régulés qui représentent la majorité des gènes. Leur expression varie au cours de la vie de la cellule ou selon leur localisation dans un tissu.

2.4. La traduction

La traduction correspond au décodage des ARN messagers pour les transformer en protéines. Elle est la suite logique de la transcription et se déroule au niveau des ribosomes, donc dans le cytoplasme des cellules. Elle est possible grâce au code génétique qui regroupent les règles par lesquelles la séquence nucléotidique d'un gène, par l'intermédiaire d'un ARN messager, est traduite en une séquence d'acides aminés dans une protéine.

2.4.1. Le code génétique

Le code génétique est un système trouvé par la nature pour faire correspondre les 4 bases azotées aux 20 acides aminés constituant les protéines. En effet, l'ARN messager est lu par groupe de 3 nucléotides et chaque triplet ou codon correspond à un acide aminé. Dans la figure 12 sont représentées les correspondances entre les codons et les acides aminés.

		NUCLÉOTIDE 2 ^{ème} POSITION					
		U	C	A	G		
NUCLÉOTIDE 1 ^{ère} POSITION	U	UUU } phényl-alanine UUC } UUA } leucine UUG }	UCU } UCC } sérine UCA } UCG }	UAU } tyrosine UAC } UAA } non-sens UAG }	UGU } cystéine UGC } UGA } non-sens UGG } tryptophane	U C A G	NUCLÉOTIDE 3 ^{ème} POSITION
	C	CUU } CUC } leucine CUA } CUG }	CCU } CCC } proline CCA } CCG }	CAU } histidine CAC } CAA } glutamine CAG }	CGU } CGC } arginine CGA } CGG }	U C A G	
	A	AUU } AUC } isoleucine AUA } AUG } méthionine	ACU } ACC } thréonine ACA } ACG }	AAU } asparagine AAC } AAA } lysine AAG }	AGU } sérine AGC } AGA } arginine AGG }	U C A G	
	G	GUU } GUC } valine GUA } GUG }	GCU } GCC } alanine GCA } GCG }	GAU } acide GAC } aspartique GAA } acide GAG } glutamique	GGU } GGC } glycine GGA } GGG }	U C A G	

Figure 12: Tableau de correspondance entre les codons et les acides aminés.

<http://tpe-regime-hyperproteine.blogspot.fr/2012/02/partie-1.html>

On peut noter que ce code génétique est dit « universel » car il est le même chez une plante, un animal, un virus. Quelques exceptions ont tout de même été mises en évidence, notamment dans l'ADN mitochondrial dans lequel quelques codons varient. D'autres ont été décelées chez certaines levures et chez des protozoaires.

D'autre part, ce code génétique est dit « dégénéré » car un même acide aminé peut être codé par plusieurs combinaisons de trois bases azotées. C'est pourquoi il existe seulement 20 acides aminés pour 64 combinaisons possibles de nucléotides.

On dit aussi que ce code génétique est « non ambigu ». En effet, nous avons vu qu'un même acide aminé pouvait être codé par plusieurs codons. En revanche, un codon ne code qu'un seul acide aminé.

Il existe des codons particuliers. Le codon AUG est un codon d'initiation de la traduction. A l'inverse, il y a trois codons de terminaison de la traduction : UAA, UAG, UGA.

2.4.2. Les molécules nécessaires à la traduction

Plusieurs types de molécules interviennent dans le mécanisme de traduction : l'ARNm (ARN messenger), l'ARNt (ARN de transfert) et les ARN ribosomiaux.

a) L'ARN messenger, ARNm

L'ARNm est une séquence de nucléotides qui seront lu par codons dans le sens 5' → 3', permettant ainsi l'association dans un ordre précis des acides aminés. Cet ARNm contient un ORF¹ délimitant le début et la fin de la séquence de l'ARNm qui va être traduite en peptide. En effet, la séquence de l'ARNm est plus longue que la séquence qui sera effectivement traduite en peptide.

Au niveau de l'extrémité 5', on trouve le codon initiateur qui donne le signal de début de traduction. La plupart du temps chez les Eucaryotes, il s'agit du codon AUG, codant donc pour une méthionine.

b) L'ARN de transfert, ARNt

L'ARNt est la molécule qui va rendre possible la liaison entre l'ARNm et les acides aminés qui vont être intégrés.

Les ARNt présentent une structure particulière. En effet, on va retrouver dans cette molécule un grand nombre de nucléotides modifiés par méthylation, isomérisation, thiolation... Il semblerait que les bases inhabituelles ne soient pas essentielles au bon fonctionnement de l'ARNt.

Par ailleurs, on peut remarquer la structure spatiale caractéristique : l'aspect général de l'ARNt a une forme de trèfle, comme représenté en figure 13. Cela est la conséquence des appariements formés par des liaisons hydrogènes entre des bases complémentaires. Au niveau des boucles, les nucléotides ne sont pas appariés.

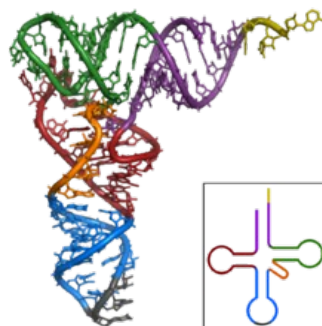


Figure 13: Structure en trèfle de l'ARN de transfert

¹ **ORF** : Open Reading Frame http://fr.wikipedia.org/wiki/Acide_ribonucl%C3%A9ique_de_transfert

Deux zones de l'ARNt possèdent des rôles bien identifiés :

- L'extrémité 3'OH est terminée par le triplet CCA où se produira une liaison ester entre l'ARNt et l'acide aminé.
- Une zone anticodon qui se lie par des liaisons faibles de type hydrogène au codon présent dans l'ARNm [8].

c) Les ARN ribosomaux, ARNr

Les ribosomes constituent les véritables machineries permettant la production des protéines. Ils se situent dans le cytoplasme des cellules. Ils sont caractérisés par leur constante de sédimentation exprimée en unité Svedberg, notée S.

Ils ont une structure différente chez les Procaryotes et chez les Eucaryotes. Chez les Eucaryotes, le ribosome est 80S, réparti en deux sous-unités : la grande sous-unité de 60S en jaune et la petite sous-unité de 40S en bleu comme illustré en figure 14. L'ARN est représenté en rouge.

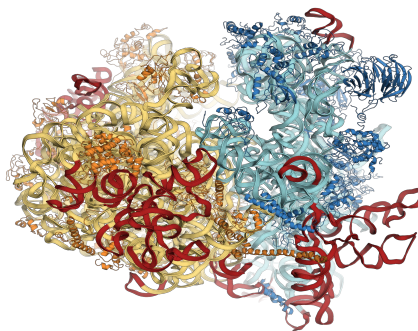


Figure 14: Structure du ribosome Eucaryote

<http://www.igbmc.fr/society/actualite/4/>

Nous pouvons préciser que la petite sous-unité contient de l'ARNr 18S et 33 protéines. La grande sous-unité, quant à elle, contient de l'ARNr 5S, 5,8S et 28S ainsi que 45 protéines.

2.4.3. L'étape d'initiation

Au départ de la traduction, les deux sous-unités du ribosome sont dissociées et libres dans le cytoplasme de la cellule. La petite sous-unité du ribosome va former un complexe avec le codon initiateur porté par l'ARNm et un ARNt. Plusieurs événements permettent à la petite sous-unité de se positionner correctement au niveau du codon initiateur porté par l'ARNm.

Il y a intervention à différents niveaux de nombreux facteurs appelés eIF :

- interaction entre un domaine de l'ARNm, le domaine RBS¹ et l'ARNr 18S de la petite sous-unité,
- liaison de facteurs eIF sur la petite sous-unité,
- prise en charge du premier ARNt porteur de la méthionine.

La coiffe 5' de l'ARNm est reconnue par le facteur eIF-4. Ce complexe va se positionner au niveau de la petite sous-unité du ribosome et va se déplacer sur l'ARNm afin de trouver le codon initiateur AUG. Une fois ce codon identifié, il y a séparation de l'ensemble des facteurs eIF. Puis, la grande sous-unité vient se mettre en place.

2.4.4. L'étape d'élongation

On peut décomposer l'élongation en trois étapes qui aboutissent à l'établissement de la liaison peptidique entre deux acides aminés.

Dans un premier temps, on observe l'arrivée d'un second ARNt lié à un acide aminé.

Dans un second temps, il y a formation d'une liaison peptidique. Pour cela, il faut qu'il y ait rupture entre le premier acide aminé, la méthionine et l'ARNt, qui est donc éjecté. La liaison peptidique peut donc être formée entre le groupement COOH de la méthionine et le groupement NH₂ du second acide aminé. Cela est réalisé par une enzyme, une peptidyl-transférase. On retrouve donc deux acides aminés liés entre eux et accrochés au deuxième ARNt.

Dans un troisième temps, la translocation se produit. C'est à dire que le ribosome va avancer sur l'ARNm pour lire le codon suivant.

Cette étape se reproduit tout au long de l'ARNm jusqu'au codon STOP qui annoncera la fin de la traduction.

2.4.5. L'étape de terminaison

Le ribosome se déplace le long de l'ARNm jusqu'à ce qu'il rencontre l'un des codons STOP. Il n'y a pas d'ARNt qui correspond à ces trois codons et donc pas d'acide aminé qui sera fixé. Cela induit la rupture entre le dernier acide aminé intégré et le dernier ARNt. Il y a intervention de facteurs de terminaison appelés eRF qui permettent

1 **RBS** : Ribosome Binding Site

la séparation du complexe ribosome-ARNm. Les deux sous-unités du ribosome se séparent et se retrouvent prêtes pour une nouvelle traduction.

2.4.6. Altération de la traduction

La synthèse des protéines se fait à partir de l'ADN, via les ARN messagers. Lorsqu'une mutation est présente au niveau de la séquence d'ADN, une modification de traduction pourra entraîner l'apparition, la modification ou l'absence d'une protéine. Cela peut avoir pour conséquence de développer une maladie par modification d'une activité métabolique. Par exemple, quand l'enzyme phosphogalactose-uridyl-transférase est absente, on parle de galactosémie congénitale. De même, dans la phénylcétonurie, c'est la phénylalanine hydrolase qui fait défaut. Il est important de noter qu'une mutation peut également apporter une activité métabolique supplémentaire et donc apporter un bénéfice à l'organisme. C'est sur cette possibilité que s'est basé Darwin pour établir sa théorie de l'évolution des espèces. En effet, il parle de sélection naturelle des mutations favorables. L'évolution est dans ce cas le résultat d'un enchaînement de mutations qui entraînent des modifications de protéines. L'évolution des espèces peut alors être retracée en suivant la séquence d'acides aminés d'une protéine donnée [7].

2.4.7. Action de certains médicaments sur la traduction

Certaines substances médicamenteuses peuvent avoir une action sur la traduction, notamment au niveau des ribosomes Procaryotes qui ont une structure différente des ribosomes Eucaryotes. Par exemple, la streptomycine inhibe la fixation de l'ARNt initiateur. La gentamycine et la néomycine se fixent sur l'ARNr 16S des ribosomes, empêchant ainsi l'appariement codon-anticodon. Le chloramphénicol inhibe l'activité peptidyl-transférase. Enfin, l'acide fusidique se lie à l'un des facteurs d'élongation et empêche de cette façon la fixation des ARNt liés aux acides aminés.

2.5. Les principes de base de l'hérédité

Les principes théoriques de base de l'hérédité ont été découverts par Gregor Johann Mendel. Suite à plusieurs années d'expérimentation sur des plants de pois, il publia ses résultats en 1866. Il décède en 1884. Ses travaux constitueront quelques dizaines d'années plus tard le fondement de l'analyse génétique par croisement. En effet, sa contribution à la génétique ne fut pas reconnue avant les années 1900 lorsque

ses expérimentations sont reprises par des chercheurs qui obtiendront les mêmes résultats que lui. Ils se rendent compte, alors, de la portée universelle du « mendélisme » car les principes énoncés semblent s'appliquer à l'ensemble des espèces animales ou végétales.

Ensuite, la théorie chromosomique de l'hérédité, fondée par des cytologistes tels que Sutton et Boveri, est étayée en abordant des mécanismes concrets, objectifs et cytologiques. Ces mécanismes correspondent à la ségrégation des chromosomes lors de la méiose ainsi qu'à la formation de gamètes haploïdes, c'est à dire avec un seul jeu de chromosomes, au lieu d'avoir des paires, suivis de l'union aléatoire des gamètes au moment de la fécondation. Les études qui en ont découlées ont été réalisées sur des levures qui ont l'avantage d'allier facilité et économie de culture. Elles confirment et affinent les concepts cités précédemment et introduisent plus facilement l'analyse fonctionnelle des gènes ainsi que leur régulation [14].

Vers le milieu du XXème siècle, les connaissances aboutissent à la compréhension du mode d'action des gènes et de l'effet de la mutation d'un gène sur le phénotype. Dans les mêmes temps, les biologistes moléculaires établissent que l'information génétique est chimiquement codée par l'ADN, puis expliquent la transcription, la traduction, les modalités de la régulation de l'expression des gènes. Dans les années 1970 est découverte l'organisation structurale des gènes eucaryotes en introns et en exons [14]. En parallèle, la cartographie des gènes s'est développée, s'affinant au fil des années grâce à l'évolution des connaissances et aux avancées technologiques.

Afin de mieux cerner ce domaine de la génétique, il est incontournable de définir les termes spécifiques utilisés. L'élément de base est l'ADN que l'on peut découper en gènes. Un gène peut être défini comme un segment d'ADN contenant toutes les instructions nécessaires à la biosynthèse d'une protéine donnée. C'est un facteur génétique qui détermine une caractéristique spécifique. Une corrélation existe entre la complexité d'un organisme et le nombre de gènes présents dans son génome. Chez l'homme on dénombre environ 25 000 gènes [9]. Il est fréquent de voir un gène se décliner en différentes versions (variation des séquences nucléotidiques) appelées allèles. On parle de façon générale de polymorphisme. C'est-à-dire qu'il existe des variations entre les individus au niveau de la séquence de gène. Ces variations existent

dans la population normale et ne sont pas pathogènes. Chaque gène occupe une place spécifique sur le chromosome : c'est le locus.

L'ensemble des gènes d'un individu forme le génotype. Lorsque sur deux chromosomes d'une même paire nous retrouvons deux allèles identiques, au même locus, on parle d'un individu homozygote pour ce gène. A contrario, lorsque sur deux chromosomes d'une même paire nous retrouvons deux allèles différents, au même locus, on parle d'un individu hétérozygote pour ce gène.

Le phénotype désigne les traits, l'apparence ou la manifestation d'un caractère chez un individu. Il résulte de l'expression du génotype [10].

C'est Mendel qui introduit le concept de dominance. En effet, il définit, dans ses conclusions qui font suite aux travaux qu'il a menés sur des plants de petits pois, le principe de séparation. Ce principe établit le fait que chaque individu diploïde possède deux allèles pour chacune de ses caractéristiques. Ces deux allèles se séparent au moment de la formation de deux gamètes : chaque gamète récupère l'un des deux allèles.

En découle ensuite le principe de dominance. Lorsque deux allèles différents sont présents dans un génotype, seul le trait codé par l'un d'entre eux, l'allèle dominant, est observé au niveau du phénotype. On dit alors du second allèle qu'il est récessif. Cela permet de comprendre les modes de transmission des maladies génétiques.

Les pathologies les plus souvent rencontrées aujourd'hui sont la maladie de Charcot Marie Tooth, la Chorée de Huntington, l'ataxie de Freidreich, la maladie de Leber et la maladie de Steinert. Ces maladies sont détaillées ci-après.

a) Maladie de Charcot Marie Tooth (CMT)

La maladie de Charcot-Marie-Tooth, ou CMT, est une neuropathie périphérique héréditaire qui atteint le système nerveux périphérique [15]. Sa prévalence est de 1 pour 2500. Elle se traduit par une atrophie musculaire et une neuropathie sensitive progressive qui concerne les extrémités des membres et qui aboutit au long terme à un handicap moteur [16]. Au niveau clinique, cette maladie se caractérise souvent par des « pieds creux ». Tous les modes de transmission peuvent être rencontrés, c'est pourquoi il existe de nombreux types de CMT. Des gènes sont connus pour les deux formes les plus courantes : CMT1 et CMTX1.

En Europe, la transmission autosomique dominante est prépondérante. A l'heure actuelle, il n'existe aucun traitement curatif. La rééducation permet de préserver aussi longtemps que possible la qualité de vie.

b) Chorée de Huntington

La chorée de Huntington est une atteinte du système nerveux central qui a une prévalence estimée entre 1 pour 10000 et 1 pour 20000 dans la population caucasienne [17]. Cette pathologie neurodégénérative, autosomique dominante, est diagnostiquée sur l'imagerie cérébrale et l'analyse génétique. Cette analyse consiste à évaluer l'expansion du triplet CAG dans l'exon 1 du gène HD. Les premiers symptômes apparaissent entre 30 et 45 ans : troubles moteurs, troubles du comportement (modification de la personnalité) et troubles cognitifs. L'évolution de cette maladie aboutit à une perte d'autonomie progressive. C'est pourquoi la prise en charge globale est importante afin d'accompagner au mieux le patient et son entourage [18].

c) Ataxie de Freidreich

Cette maladie correspond à des troubles de la coordination des mouvements, de l'articulation, associés à d'autres signes neurologiques, à une cardiomyopathie et parfois à un diabète [19]. La transmission se fait selon un mode autosomique récessif. Elle est l'ataxie héréditaire la plus fréquente. Le gène responsable, qui code pour la frataxine, est le gène FRDA et un diagnostic par analyse génétique est possible.

d) Neuropathie optique héréditaire de Leber

Cette neuropathie est un dysfonctionnement du nerf optique provoqué par une mutation sur l'ADN mitochondrial [20]. La transmission se fait selon un mode dominant.

e) Maladie de Steinert ou dystrophie myotonique de type I

Cette pathologie se caractérise par une myotonie et une atteinte multisystématique [21]. Elle est la plus fréquente des dystrophies musculaires chez l'adulte. La transmission se fait selon un mode autosomique dominant avec un phénomène d'anticipation. La mise en évidence de l'anomalie en 19q13-2 constitue le diagnostic de certitude.

3 Projet

L'avancée des connaissances en génétique médicale des dernières dizaines d'années a été considérable ! Les techniques d'analyses se sont développées en parallèles et sont de plus en plus précises et accessibles. Toutes ces méthodes d'étude de l'ADN peuvent être considérées comme une boîte à outils qui permet l'ouverture vers l'établissement de diagnostics dans les cas de pathologies rares, qui n'ont pas encore d'explications sur un plan génétique. La mise au point des dernières méthodes d'hybridation de l'ADN ouvre des perspectives dans la pratique quotidienne des neurogénéticiens. En effet, certains patients, connus et suivis depuis de nombreuses années, peuvent bénéficier de nouvelles analyses afin de tenter d'identifier la mutation de l'ADN, cause de leur pathologie. Pour cela, il est important de comprendre la méthode d'analyse afin de sélectionner au mieux les patients qui peuvent faire l'objet de cette recherche. C'est au sein du CHU d'Angers, dans le service de neurogénétique, que j'ai procédé à la consultation minutieuse des dossiers des patients, un à un, afin de trouver ceux dont les profils étaient les plus pertinents pour l'analyse. Les patients sélectionnés ont donc eu par la suite la possibilité de voir leur ADN analysé au laboratoire du CHU d'Angers. Cette partie aborde dans un premier temps et en détails les méthodes d'hybridation appelées CGH array et SNP array. Puis, nous nous focaliserons sur la méthode de sélection des dossiers sera ensuite expliquée.

3.1. Matériel et méthodes

Nous allons à présent aborder le principe d'hybridation de l'ADN qui nous permettra par la suite de mieux comprendre le fonctionnement des méthodes récentes utilisées en recherche génétique : la CGH array et la SNP array. Ensuite, nous détaillerons la classification des dossiers des patients puis nous aborderons les critères d'exclusion qui ont permis d'aboutir à la liste des patients sélectionnés pour cette étude.

3.1.1. Le principe d'hybridation de l'ADN

a) Principe général de l'hybridation

Pour étudier les acides nucléiques, l'une des méthodes universelles utilisée est l'hybridation. Cette analyse offre la possibilité de repérer, d'identifier ou de quantifier une molécule d'acide nucléique, ADN ou ARN. Elle permet aussi de comparer entre elles des séquences d'ADN ou d'ARN.

Son principe général se base sur les propriétés spécifiques des acides nucléiques. D'une part, lorsque l'on chauffe une molécule d'ADN, on observe une ouverture de la double hélice. Les brins se séparent. Puis, quand la température redescend, les hélice se reforment, les brins se ré-associent. On parle de dénaturation/renaturation de l'ADN. D'autre part, lorsque des mono-brins se rencontrent, ils ont tendance spontanément à se ré-associer grâce à la complémentarité qui existe entre les bases azotées.

b) L'hybridation en pratique

En pratique, on utilise une sonde, séquence d'ADN ou d'ARN, marquée par un isotope radioactif (souvent le ^{32}P). Chaque sonde est spécifique de la séquence à analyser. Un mélange de sondes et de molécules à étudier sont chauffées puis incubées.

Cela permet dans un premier temps de séparer les brins de la double hélice d'ADN puis, dans un second temps, de ré-associer les molécules complémentaires : c'est ce que l'on appelle l'étape d'hybridation. Les molécules non hybridées sont éliminées.

Les molécules à étudier sont fixées sur une membrane de nitrocellulose qui sera ensuite mise à incuber avec la sonde (l'excès de sonde sera éliminé). La radioactivité des sondes permettra, par la suite, de détecter la présence ou de mesurer la quantité des séquences d'acides nucléiques qui sont complémentaires à la sonde [7].

c) L'hybridation in situ fluorescente ou FISH

Dans le cas particulier de l'hybridation in situ fluorescente, les sondes marquées utilisées sont hybridées sur des brins d'ADN préalablement dénaturés. Les sondes sont par la suite révélées grâce à des anticorps couplés à un fluorochrome¹, comme représenté en figure 15. C'est ensuite grâce à un microscope que sont révélées les zones fluorescentes correspondant aux zones où les sondes sont fixées sur l'ADN étudié.

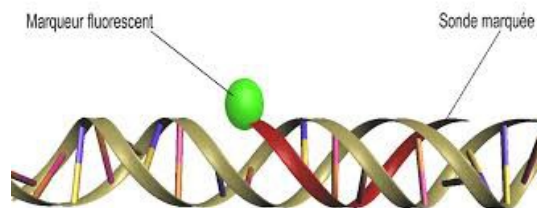


Figure 15: Schéma d'une hybridation fluorescente (FISH)

https://encrypted-tbn2.gstatic.com/images?q=tbn:ANd9GcRvKPWUAI-J7hsYyG_kBFAsmwQ2ebFVNIrexyCYOQqX3KkzUkuXWw

d) Objectifs de l'hybridation

L'hybridation est une méthode qui a pour objectif de cibler de courtes séquences d'ADN grâce aux différentes sondes utilisées. On l'utilise pour réaliser des caryotypes, détecter la présence de gènes, localiser les gènes ou encore analyser des anomalies chromosomiques. Deux nouvelles méthodes d'hybridation de l'ADN, détaillées ci-après, ont été mises en place ces dernières années: la CGH array et la SNP array. La SNP array peut-être considérée comme une variante de la CGH array.

3.1.2. La CGH array

La CGH, Hybridation Génomique Comparative, est une technique de cytogénétique mise au point en 1992 par l'équipe de Dan Pinkel [22]. Son principe général est la comparaison du nombre de molécules d'une même quantité d'ADN de deux individus.

a) Origine et principe

Vers la fin des années 1990, Solinas-Toldo et Pinkel et al. [23][24] proposent une variante de la technique de CGH classique, utilisant des lames sur lesquelles sont fixées des fragments d'ADN. C'est en grande partie grâce au séquençage du génome que l'on a pu fixer sur ces lames des séquences génomiques, appelées sondes. Ces lames ont

¹ **Fluorochrome** : substance chimique capable d'émettre de la lumière fluorescente après excitation.

donc été appelées « puces à ADN » ou « microarray ». On appelle donc cette technique la « CGH array » ou « CGH sur réseau d'ADN ».

Actuellement, cette technique est utilisée avec succès dans de nombreux laboratoires afin de détecter certaines altérations du génome humain. Dans le retard mental et les malformations congénitales (2 à 3% de la population), de nombreux travaux ont démontrés que cette technique permettait de détecter plus de 15% d'anomalies, en plus de celles détectées au moyen des techniques de cytogénétique classiques et moléculaires [25].

b) Les points forts de la CGH array

La CGH array prend le dessus aujourd'hui sur l'étude du caryotype afin de déceler des anomalies chromosomiques. En effet, cette méthode permet de mettre en lumière des anomalies avec une bien meilleure résolution car elle utilise uniquement des parties codantes de l'ADN : les exons. Le but de cette nouvelle méthode est d'établir une corrélation génotype/phénotype. Elle permet de faire le lien entre une pathologie et le(s) gène(s) impliqué(s). C'est pourquoi cette dernière constitue un outil de choix en recherche.

L'avantage de cette technique est qu'elle permet une approche à la fois globale et hautement résolutive.

Sa réalisation est rapide. En effet, il est possible de séquencer l'ensemble du génome d'un humain en l'espace de quelques semaines. En comparaison, il aura fallu des années pour séquencer l'ensemble du génome humain, avec le « Projet Genome Humain », débuté dans les années 1990 et achevé en 2003 [26].

C'est aussi une méthode sensible et automatisable. Elle offre la possibilité de tester par hybridation compétitive un nombre très élevé de régions du génome.

Le diagnostic est possible à prix raisonnable pour des centaines de maladies chromosomiques connues sur l'ensemble du génome à l'aide d'un seul et unique test. En effet, il faut compter 250€ pour une analyse. Ce coût correspond au prix des réactifs utilisés.

Cette technique peut permettre d'expliquer des pathologies telles que des handicaps mentaux ou des maladies du tissu conjonctif [27]. En conséquence, elle constitue aujourd'hui l'analyse incontournable en génétique clinique, en offrant une alternative puissante aux autres méthodes utilisées jusqu'à aujourd'hui [28].

c) Recueil des échantillons d'ADN

Suite au consentement signé par le patient ou son représentant légal lors d'une consultation, l'ADN est prélevé. Il est ensuite stocké au laboratoire en vue des futures analyses qui seront prescrites pour le patient. Conformément au protocole standard, l'ADN génomique est extrait à partir de salive et de leucocytes [29].

d) Hybridation

Cette technique d'hybridation génomique comparative utilise une même quantité d'ADN témoin (ou de contrôle) et d'ADN patient marqués par deux fluorochromes de couleurs différentes. Ces ADN sont déposés sur une lame. Sur une puce, on intègre des sondes en nombre variable. Plus il y a de sondes, plus la méthode sera résolutive. Puis, on met en compétition les deux ADN pour qu'ils s'hybrident avec la puce. Les quantités des deux ADN étant égales au départ, on interprétera la couleur des spots correspondant aux sondes : soit il n'y a pas de différence, soit on détecte une délétion, soit on détecte une duplication.

Les résultats sont interprétés par des logiciels spécifiques qui les représentent sous forme de graphiques.

Prenons comme exemple l'étude de l'équipe de Tayeh en 2009 [30]. Le protocole débute par l'extraction de l'ADN du patient à l'aide d'un kit d'extraction d'ADN. Des échantillons témoins homme et femme sont utilisés. Les échantillons du patients sont traités aux ultrasons afin d'obtenir des fragments d'ADN de 500 à 2000 bases, la taille des fragments étant vérifiée sur gel d'agarose à 1%. Ils ont ensuite utilisé une enzyme, Klenow enzyme, ainsi que des amorces, Cy3 ou Cy5 9mer, afin, respectivement, d'étiqueter les patients et de contrôler les échantillons. Les échantillons sont ensuite purifiés par précipitation à l'isopropanol. Treize microgrammes d'ADN de chaque patient et les contrôle de genre sont combinés et dessiqués, puis remis en suspension dans un tampon approprié pour permettre l'hybridation avec les séquences contrôles Cy3 et Cy5. Les puces à ADN sont ensuite hybridées pendant 16 à 20 heures à 42°C.

Puis, les résultats sont extraits à partir d'images fournies par un logiciel spécifique. Les données sont analysées par deux programmes différents qui donnent les résultats sous forme de graphiques, l'axe des ordonnées indiquant le gain ou la perte de matériel génétique (1 = gain, 0 = normal, -1 = perte) et l'axe des abscisses indiquant la position sur le chromosome (voir annexe I).

e) Applications

La CGH array est capable de détecter de nombreuses délétions et duplications, dans une large panoplie de gènes, en utilisant qu'un seul réseau [30]. Une étude d'Aradhya et al. [31] démontre que la CGH array est capable d'identifier, partiellement ou en totalité, des délétions et duplications dans approximativement 5% des cas dans une cohorte clinique envoyée pour analyse à un laboratoire. Lorsque les résultats sont détaillés en fonction du mode de transmission, on obtient une identification à hauteur de 5% des gènes autosomiques dominants, 10% des gènes autosomiques récessifs et 3,5% des gènes du chromosome X. Cette étude conclut que les mutations du nombre de copies intra-géniques sont plus fréquentes que précédemment suspectées dans les pathologies mendéliennes et que leur étude devrait faire systématiquement partie du bilan diagnostique de routine. Aussi, dans une étude de Wang et al. [32], un taux de détection des délétions et duplications de 3% apparaît grâce à la CGH array, lors de l'étude de pathologies mitochondriales et métaboliques, faisant apparaître cette méthode comme utile en tant qu'élément complémentaire de diagnostic.

f) Limites de la méthode

La CGH array est une méthode fournissant un nombre important de données. Leur interprétation reste encore délicate et peut contenir des erreurs. En effet, certaines régions du génome présentent un Nombre de Copies Variables, notées CNV. Ces CNV peuvent correspondre soit à des polymorphismes, soit à des anomalies moléculaires. C'est pourquoi il existe des bases de données de CNV qui permettent d'exclure les polymorphismes.

g) Perspectives

De nombreux cancers font l'objet d'études afin de faciliter leurs diagnostics. Cela ouvre des pistes pour la découverte de nombreux médicaments, qui constitue l'objectif principal de l'identification des altérations du matériel génétique et l'élaboration de thérapies. Les thérapies géniques font aujourd'hui l'objet de nombreux projets de recherches.

3.1.3. La SNP array

La Single Nucleotide Polymorphism array, ou SNP array, est la méthode utilisée au laboratoire du CHU d'Angers. C'est une « version » de la CGH array. En effet, cette

méthode utilise le même principe mais ne compare qu'une seule base à la fois et non une séquence de 10 à 15 paires de bases. Elle permet le génotypage d'un individu. Il existe peu de données expliquant cette méthode. Les éléments décrits ici sont ceux utilisés au laboratoire du CHU d'Angers.

Nous retrouvons tout de même de nombreuses publications récentes décrivant les applications de la SNP array dans de nombreuses pathologies humaines telles que les dystrophies musculaires, certains cancers, etc. Nous retrouvons, par exemple, une étude de L. Papic et al. qui étudie les LGMD, Limb Girdle Muscular Dystrophies, qui constitue un groupe hétérogène d'atteintes musculaires au niveau clinique et au niveau génétique [33]. Cette étude se fait par le biais d'une cartographie des zones homozygotes du génome, décelées par SNP array.

Aussi, cette technique permet aujourd'hui l'étude de certaines espèces animales. Par exemple, une étude de A. Drywa et al. décrit la différenciation de trois espèces de salmonidées grâce à la SNP array [34].

a) Les points forts de la SNP array

Le *Single Nucleotide Polymorphism* représente les variations génétiques dans le génome humain et correspond à environ une paire de bases sur mille dans le génome. C'est-à-dire qu'il représente plus de 90 % de toutes les différences entre les individus. Leur étude cible donc les régions codantes de l'ADN. Cela permet de réaliser l'étude du génome de façon plus optimale et d'augmenter les chances de trouver une mutation pertinente chez un patient.

La méthode est automatisée ce qui lui confère un gain de temps ainsi qu'une reproductibilité intéressante.

b) Principe

De la même façon que pour la CGH array, la SNP array utilise des puces à ADN. On place des billes de 3 μ m dans les puits d'une carte en PVC. Sur ces billes, on « branche » des oligonucléotides, c'est à dire des séquence de 10 à 15 bases. On hybride pour laisser apparaître uniquement le nucléotide ou SNP étudié. On répète 30 à 40 fois chaque bille de chaque séquence.

On recueille l'ADN du patient qui sera analysé à une concentration donnée. On le fragmente en oligonucléotides capable de se fixer sur les oligonucléotides fixés sur les billes, en laissant apparaître le SNP.

On procède ensuite à la révélation par scan : chaque pixel représente une bille. On étudie l'intensité pour chaque bille et pour chaque patient. C'est le logiciel qui réalise automatiquement la normalisation. L'analyse se fait par comparaison à un ADN de référence, en fonction de l'origine du patient : caucasien, africain, ...

Cette technique est automatisée et demande une durée de 3 jours sur la plateforme. Les ADN sont analysés par séries de 40 à 50.

c) Applications

La SNP array est une méthode qui peut être utilisée afin de réaliser des génotypages. Elle permet de comparer l'ADN des patients à des bases de données.

Elle détecte les délétions et les duplications. Par conséquent, elle peut mettre en évidence les translocations¹, dont la prévalence est de 0,2% [35]. Elle permet par ailleurs de déceler les cas de consanguinité en révélant les régions chromosomiques pour lesquelles on décèle une perte d'hétérozygotie. On parle de « régions avec LOH² ».

d) Limites de la méthode

La SNP array est une méthode plus compliquée à réaliser que la CGH array et les données obtenues sont différentes. Elle n'est pas réalisée en routine mais seulement dans des cas particuliers. Elle coûte chère, environ 1200€ par analyse. L'interprétation des résultats n'est pas informatisée, il est nécessaire d'avoir un œil avisé pour rendre un résultat au patient. Si rien n'est décelé, l'interprétation est rapide, environ 30 minutes. En revanche, si une anomalie est détectée, l'interprétation peut nécessiter plusieurs heures, d'autant plus s'il faut procéder à des comparaisons. Cette méthode permet de classer les résultats en 3 catégories : bénin, pathologique ou indéterminé. Dans le cas d'une découverte de variation pathologique, il faut utiliser une autre méthode d'analyse d'ADN pour confirmer l'anomalie retrouvée. Les résultats indéterminés représentent 5 à 10 % des analyses.

3.1.4. La sélection des patients

Il n'existe actuellement aucune base informatique permettant de sélectionner les dossiers selon leurs critères, notamment ceux qui nous intéressent dans le cadre de cette étude. En effet, l'archivage des dossiers médicaux ne s'effectue pas encore sous forme électronique. Les dossiers, sous forme papier, sont rangés par ordre

1 **Translocation** : échange de matériel génétique entre deux chromosomes non homologues.

2 **LOH** : Loss Of Heterozygosity

alphabétique, dans des locaux prévus à cet effet. L'étude se fait donc par l'ouverture successive de chaque dossier.

La consultation des dossiers du service de neurogénétique d'Angers a permis la sélection des patients qui pourront, par la suite, faire l'objet d'une analyse génétique par hybridation, au laboratoire du CHU d'Angers. En effet, comme décrit précédemment, la SNP array y est pratiquée depuis quelques années.

Le but de ce travail est de sélectionner les dossiers de patients présentant une maladie se transmettant selon un mode récessif. Dans les cas où un coefficient de consanguinité¹ est mis en évidence au sein d'une famille, la probabilité de retrouver des zones d'homozygoties est beaucoup plus importante. Ces dossiers seront traités en priorité.

a) Le suivi des patients au CHU d'Angers

Les patients qui sont amenés à consulter dans le service de neurogénétique sont des sujets qui présentent des symptômes jusque là inexpliqués ou des sujets qui possèdent des antécédents de maladie génétique dans leur famille.

Lors de la première consultation, un dossier médical est constitué : nom, prénom, date de naissance. Le neurogénéticien procède à un interrogatoire et à un examen clinique. L'arbre généalogique du sujet est complété. Cet arbre est construit systématiquement pour chaque patient lors de sa première consultation dans le service de neurogénétique. Il se présente sous la forme ci dessous, en figure 16.

Cette représentation des individus d'une même famille permet de retracer l'histoire de cette famille et de mettre en évidence les sujets atteints par la maladie. Lorsque l'on sait que les parents d'un individu sont originaires de la même région, cette information est précisée au niveau de l'arbre généalogique.

Lorsqu'il s'agit d'une pathologie génétique monofactorielle, la transmission suit les lois de Mendel. Dans ce cadre, on peut décrire quatre modes de transmission : autosomique ou lié à l'X, selon que le gène impliqué est localisé sur un autosome² ou sur le chromosome X ; dominant ou récessif.

Sur l'arbre génétique, on peut supposer le mode de transmission de la pathologie incriminée.

1 **Coefficient de consanguinité** : probabilité pour que les deux gènes allèles que possède un individu en un locus donné soient identiques *par descendance*.

2 **Autosome** : tout chromosome autre que les chromosomes sexuels X et Y.

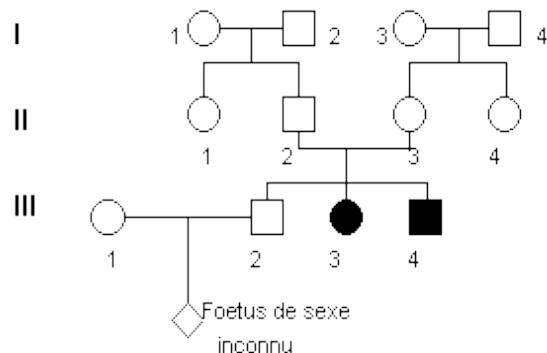


Figure 16: Exemple d'un arbre génétique.

<http://www2.ac-lyon.fr/enseigne/biologie/spip.php?article262>

Lorsque la maladie atteint un parent et que l'un de ses enfants est également atteint, fille ou garçon, on décèle un mode de transmission autosomique dominant.

Au contraire, si la pathologie atteint une fratrie, on suppose plutôt un mode de transmission autosomique récessif. Les parents sont alors tous les deux porteurs de la mutation à l'état hétérozygote et lorsqu'ils transmettent tous les deux la mutation à leur enfant, celui-ci développera la maladie. Lorsqu'un enfant hérite de la mutation d'un seul de ses parents, il est porteur mais ne développe pas la maladie.

D'autre part, on peut identifier une maladie qui atteint plusieurs générations mais uniquement des femmes. Dans ce cas, on peut identifier une transmission liée à l'X. Les hommes peuvent être porteurs du chromosome X pathogène et le transmettre à leurs filles.

Dans le cadre de la recherche génétique, une prise de sang peut être effectuée si nécessaire, après recueil du consentement du patient. Cet échantillon servira aux différentes analyses qui pourront être demandées par la suite.

Pour certaines maladies neurogénétiques, des tests de dépistages sont actuellement disponibles car les mutations leur correspondant sont connues. Certains patients ont pu bénéficier de ces tests, sur demande du neurogénétiicien qui les suit. Ces analyses sont effectuées soit au laboratoire de l'hôpital, soit après transfert du prélèvement dans des laboratoires ayant l'autorisation de pratiquer certains types d'analyses génétiques. Les résultats peuvent arriver plusieurs mois après la demande. Le résultat de chaque test est envoyé au médecin qui a établi la demande du test puis

il est classé dans le dossier du patient. Le patient sera informé du résultat lors de sa prochaine consultation.

Une grande majorité de patients vient en consultation de suivi une fois par an.

b) Consultation des dossiers

Au cours de l'année 2012, 1247 dossiers de patients ont été consultés, un à un, au sein du service de neurogénétique du CHU d'Angers.

Lors de la consultation des dossiers, un tableau informatique est construit en parallèle afin de répertorier les éléments clés de tous les dossiers : le nom du patient (3 premières lettres), son prénom (première lettre), sa date de naissance, le diagnostic et/ou les symptômes, le mode de transmission génétique, les antécédents familiaux, les mutations exclues, les mutations retrouvées, ...

On retrouve un extrait dans le tableau 1 ci-après.

NOM	PRENOM	NE(E) LE	PATHOLOGIE	TRANSMISSION	REMARQUES
ABR	M	19/10/35			Patient de neurologie
ADA	D	03/10/53	Suspicion CMT : pieds creux, orteils comme sa mère	AD	Gène MFN2 négatif
AFA	P	31/01/83	CMT	AD	PMP22
AFO	I	21/05/73	CMT type 1	AD	Mutation 17p11,2
AFO	M	27/11/54	Tableau clinique de CMT		Pas de mutation retrouvée
AGO	H	08/11/33	Amylose		Décédé, mut Ile 107 Val, gène TTR
...

Tableau 1 : Extrait du tableau listant la totalité des patients.

c) Classification des patients

A partir du tableau initialement construit, listant la totalité des dossiers consultés, un système de classification est établi.

Dans un premier temps, une classification selon la pathologie de chaque patient est réalisée. Pour chaque dossier, un numéro est attribué, correspondant à la maladie du patient ou, à défaut, lorsque la pathologie n'est pas identifiée, correspondant au symptômes du patient, comme illustré dans le tableau 2.

Classification 1	Pathologie
1	Patients non neurogénétique
2	Patients peut-être neurogénétiques
3	Patients sains
4	Patients qui vont bénéficier de tests de dépistages
5	Patients symptomatique ayant une pathologie non identifiée, « cas isolé »
6	Patients symptomatiques ayant une pathologie non identifiée, « avec antécédents familiaux »
7	Patients symptomatiques ayant une pathologie non identifiée, consanguinité
8	Patient dont la mutation a été identifiée
9	Patient porteur sain d'une mutation
10 à 72	Patients dont la pathologie a été identifiée : adrénoleucodystrophie, alpha-D-mannosidase, amyotrophie, ...

Tableau 2 : Classification 1 des patients

Dans un second temps, une classification selon les critères d'exclusion des dossiers est faite. Une lettre est associée à chaque dossier. Ces critères sont explicités dans le tableau 3.

Classification 2	Critère
a	Patients non neurogénétiques
b	Patients peut-être neurogénétiques
c	Patients sains
d	Patients porteurs de mutation
e	Patients dont la mutation a été identifiée
f	Patient ayant une pathologie autosomique dominante
g	Patients ayant une pathologie liée à l'X
h	Patients « cas isolée », peu d'information dans leurs dossiers
i	Patients « avec antécédents familiaux », peu d'information dans leurs dossiers
j	Patients pré-sélectionnés
k	Patients sélectionnés

Tableau 3 : Classification 2 des patients.

d) Critères d'exclusion des dossiers

La pré-sélection des dossiers a été réalisée par l'intermédiaire de critères dits d'exclusions. Ce sont des critères qui ont été déterminés et qui correspondent à des éléments des dossiers qui nous permettent d'éliminer les patients car ces éléments ne rentrent pas dans le cadre de notre analyse.

Cas des pathologies non neurogénétiques

Dans un premier temps, on a des patients dont les symptômes, l'histoire de la maladie ou les antécédents excluent une origine neurogénétique à leur pathologie. Ces patients ne peuvent pas intégrer notre étude.

Cas des patients peut-être neurogénétiques

Dans un second temps, certains patients présentent des symptômes pouvant entrer dans le cadre d'une pathologie neurogénétique mais sans aucune certitude. Nous ne pouvons donc pas les intégrer à notre sélection.

Cas des patients sains

Dans un troisième temps, on doit isoler les patients sains. Il s'agit des patients qui ont été reçus en consultation au CHU d'Angers suite à la découverte d'une pathologie chez un membre de leur famille. Le but de cette consultation était de vérifier la clinique du patient et, si nécessaire, de prescrire l'analyse permettant de confirmer ou d'infirmer la présence de la mutation familiale.

Cette démarche d'analyse est principalement réalisée dans le cadre des pathologies pour lesquelles on peut détecter facilement les mutations en cause, à savoir par exemple : la maladie de Huntington, la maladie de Steinert, la maladie de Charcot-Marie-Tooth. L'analyse peut aussi être demandée lorsqu'une mutation a été identifiée chez le membre de la famille atteint, la maladie n'étant pas forcément clairement identifiée. On peut donc dans ce cas rechercher précisément cette mutation.

Si ces analyses ne décèlent aucune mutation, les patients sont sains et ils ne seront donc plus reçus en consultation par la suite, dans le service de neurogénétique. Ils sont également exclus de notre étude.

Cas des porteurs sains de mutations

D'autre part, on retrouve des patients que l'on appelle des porteurs sains. Il s'agit de patients qui ne présentent pas de symptômes mais chez qui on a retrouvé une mutation à l'état hétérozygote. Soit il s'agit d'une maladie qui ne s'exprime qu'à l'état homozygote et dans ce cas, il s'agit de pathologies récessives. Soit il s'agit d'une mutation correspondant à une pathologie qui se transmet selon un mode dominant et

alors, il s'agit d'un patient présymptomatique. C'est-à-dire que le patient est dans une phase qui précède l'expression de la pathologie.

Cas des pathologies ou mutations déjà identifiées

Sur l'ensemble des dossiers consultés, certains patients présentent une pathologie déjà identifiée. Cela signifie que le diagnostic a déjà été posé, au vu du tableau clinique, ou que la mutation a déjà été retrouvée lors d'un test génétique. Dans ces cas là, l'ensemble de ces informations sont ajoutées dans le tableau et le dossier concerné est exclu de notre étude.

Cas des pathologies autosomiques dominantes

Pour un certain nombre d'autres patients, les résultats des tests génétiques demandés sont négatifs ou aucune analyse n'a, pour l'heure, été réalisée. Il n'y a pas encore d'anomalie identifiée ni de diagnostic posé. Les seuls éléments du dossier sont les symptômes du patients. Il faut alors envisager un autre protocole d'analyse.

La consultation de l'arbre génétique présent dans les dossiers permet de cibler rapidement le mode de transmission de la pathologie incriminée.

Cas des pathologies liées à l'X

Nous avons des patients qui présentent des pathologies liées à l'X. Ces patients ne rentrent pas dans notre étude car nous privilégions les maladies autosomiques récessives.

Cas isolés et trop peu d'informations

Enfin, le patient présente une pathologie « isolée » lorsqu'il n'y a pas d'autres cas connus dans la famille. Il est le seul atteint. De plus, en général, il manque d'informations dans les dossiers. Ces dossiers ne peuvent entrer dans notre sélection. Il est possible dans certains cas les parents soient consanguins mais lorsque cette information n'est pas signalée dans le dossier, il est impossible de l'intégrer à notre étude.

Cas avec antécédents familiaux et trop peu d'informations

Nous retrouvons des cas de patients ayant une pathologie non identifiée, ayant des antécédents familiaux. Lorsque le dossier est trop peu documenté, nous ne les incluons pas dans notre sélection.

Sujets pré-sélectionnés et sélectionnés

Suite à ces exclusions, les dossiers restants sont ceux des patients qui présentent une pathologie apparemment autosomique récessive (fratrie atteinte), avec ou sans consanguinité apparente. Il s'agit des dossiers les plus pertinents dans notre étude qui feront l'objet d'une réunion avec les neurogénétiiciens afin de réduire la liste et d'aboutir aux « patients sélectionnés » pour qui une demande de SNP array sera faite au laboratoire d'analyses du CHU d'Angers.

4 Résultats

Dans le service de neurogénétique d'Angers, des bases de données ont été établies mais aucune ne peut nous permettre de sélectionner les patients dans le cadre de notre étude, c'est pourquoi un tableau récapitulatif a été établi. C'est à partir de ce tableau qu'a ensuite été réalisé un travail de classification. Ce dernier a permis de quantifier le nombre de patients pour chaque catégorie de la classification, mais également de trier les patients à partir des critères d'exclusions. Ces critères ont permis d'aboutir à la sélection finale de patients qui bénéficieront de l'analyse par SNP array.

Dans une première partie, nous aborderons donc les résultats chiffrés qui sont ressortis de la consultation des dossiers. Nous verrons quelles sont les pathologies génétiques les plus représentées au sein de la patientèle angevine. Dans une seconde partie, nous nous plongerons au cœur de l'objet de ce travail, c'est-à-dire l'exclusion progressive des patients qui nous mènent à la liste finale de ceux qui ont été sélectionnés. Puis, nous verrons les résultats des analyses génétiques par SNP array des patients sélectionnés.

4.1. Les résultats quantitatifs

4.1.1. La classification des patients

Au sein du tableau récapitulatif, un numéro de code a été associé à chacun des 1247 patients étudiés. A chaque code correspond une catégorie. (voir section « c) Classification des patients » dans la section 3)

En premier lieu, ont été exclus les 28 patients n'ayant pas de pathologie neurogénétique ainsi que les 25 patients ayant probablement une maladie neurogénétique.

Puis, on peut isoler les 71 patients « sains ». En parallèle, 26 patients sont en attente des résultats des tests de dépistage pour la pathologie familiale ou pour une autre pathologie ou sont en attente de réaliser ces tests.

On peut noter que 14 patients sont porteurs sains de la mutation familiale.

Il reste 1083 patients parmi lesquels on peut dénombrer 892 patients qui ont une pathologie neurogénétique étiquetée, identifiée ou non sur le plan génétique. Nombreuses pathologies sont représentées.

a) Les pathologies les plus représentées

On définit les pathologies les plus représentées comme celles qui comptent plus de 10 patients au CHU d'Angers. La principale pathologie représentée est la maladie de Charcot-Marie-Tooth ou CMT. On observe 203 patients atteints de cette maladie. La mutation en cause n'est pas identifiée dans 128 cas. Pour les 75 autres patients, on peut lister les gènes porteurs des mutations impliquées dans les CMT que l'on a retrouvés, comme décrit dans le tableau 4 ci-après.

Gène	Nombre de patients
PMP22 (locus 17p11,2)	28
LITAF	11
GJB1 (lié à l'X, codant pour la connexine 32)	11
NEFL	7
MFN2 (codant pour la mitofusine 2)	6
MPZ	5
GDAP	4
SHT3C2	3

Tableau 4 : Gènes impliqués dans les CMT au CHU d'Angers.

La deuxième pathologie la plus représentée est la chorée de Huntington avec 130 patients suivis au CHU d'Angers. Parmi eux, de nombreux patients font partie de la même famille.

On retrouve ensuite les ataxies : 123 patients sont concernés. Au sein de ce groupe, 59 sujets présentent une ataxie spino-cérébelleuse ou SCA et 18 souffrent d'une ataxie de Friedreich. Parmi les 46 autres patients, 40 présentent un syndrome cérébelleux ou une ataxie cérébelleuse et 6 ont une ataxie autre.

Puis, on observe 71 patients qui présentent des neuropathies : 41 patients chez qui on a diagnostiqué une neuropathie héréditaire avec hypersensibilité à la pression (ou HNPP ou neuropathie tomaculaire) et 30 patients qui présentent une neuropathie optique dont 29 ont été identifiées comme étant des atrophies optiques de Leber.

On peut citer aussi les 52 patients présentant des paraparésies ou paraplégies et les 50 patients qui ont une neurofibromatose.

En ce qui concerne les cytopathies mitochondriales, elles sont représentées par 44 patients dont 7 ont un syndrome de MELAS.

Les atteintes musculaires, notamment les amyotrophies concernent 27 patients et les atrophies diverses concernent 23 patients.

On observe des retards de diverses natures, mental ou moteur, chez 21 patients.

On relève 14 patients qui souffrent de scléroses tubéreuses de Bourneville.

Enfin, on dénombre 16 patients affectés par la maladie de Steinert et 11 patients atteints de cavernomatose.

Tous ces chiffres sont résumés dans le tableau 5.

Pathologie	Nombre de patients
CMT	203
Chorée de Huntington	130
Ataxies (SCA, ataxie de Friedreich , syndromes cérébelleux ou autres)	123 (59,18,40,6)
Neuropathies (HNPP, neuropathie optique dont Leber)	71 (41, 30)
Paraparésies, paraplégies	52
Neurofibromatoses	50
Cytopathies mitochondriales (dont le syndrome de MELAS)	44 (7)
Atteintes musculaires, amyotrophies	27
Atrophies diverses	23
Retard mental ou moteur	21
Maladie de Steinert	16
Sclérose tubéreuse de Bourneville	14
Cavernomatoses	11

Tableau 5 : Récapitulatif des pathologies les plus représentées au CHU d'Angers.

D'autre part, 8 patients symptomatiques et pour lesquels la mutation génétique a été identifiée ont été regroupés. Les tableaux cliniques décrits chez ces patients ne correspondent actuellement à aucune pathologie connue. Par exemple, chez l'un de ces patient, une mutation du gène OPA1 a été retrouvée, qui entraîne une atrophie optique dominante et une surdité.

b) Les pathologies les moins représentées

On définit les pathologies les moins représentées comme étant celles qui comptent entre 2 et 10 patients au CHU d'Angers.

Elles sont listées dans le tableau 6.

Pathologies	Nombre de patients
Chorées (non Huntington)	9
Leucodystrophies	9
Maladie de Strümpell-Lorrain	9
Intolérances à l'effort	8
Maladie de Parkinson familiale	8
Maladie de Turner	6
Adrénoleucodystrophie	4
Syndrome Cadasil	4
Syndrome de Kearns Sayre	4
Syndrome de Leigh	4
Leucoencéphalopathie	3
Epilepsies	3
Hyposensibilité à la douleur	3
Neurinomes	3
Scléroses en plaques	3
Syndrome de Moya Moya	3
Syndrome de Wolfram	3
Alpha-D-mannosidase	2
Maladie de Refsum	2
Maladie de Von Hippel Lindau	2
Paramyotomie d'Eulenburg	2
Syndrome de Partington	2

Tableau 6 : Récapitulatif des pathologies les moins représentées au CHU d'Angers.

c) Les pathologies très peu représentées

Pour les 26 maladies listées ci-dessous, un seul patient est suivi à Angers :

- Céroïde lipofushinose infantile
- Déficit en OTC
- Dystrophie musculaire de Becker
- Hémochromatose
- Pachygyrie
- Liponeurocitome cérébelleux
- Maladie de Gilles de la Tourette
- Maladie de Kennedy
- Maladie de Riley-Day
- Maladie de Steel Richardson

- Maladie de Toriello Carrey
- Maladie de Waldenström
- Maladie de Wilson
- Mucopolysaccharidose de type III
- Neuropathie optique héréditaire de Kjer
- Syndrome de l'X fragile
- Pseudoxanthome élastique
- Syndrome de Joubert
- Syndrome de Bardet-Biedl
- Syndrome de Dravet
- Syndrome de Pfeiffer
- Syndrome de Rubinstein Taybi
- Syndrome de Sando
- Syndrome de Smith-Magenis
- Syndrome de Williams
- Tétralogie de Fallot

d) La classification des autres patients

Après avoir compté les patients non neurogénétique, les patients peut-être neurogénétique, les patients sains, les patients en attente de tests génétiques, les patients porteurs de mutation, les patients dont la pathologie a été étiquetée, on a pu construire les catégories détaillées ci-après.

Nous pouvons regrouper les 164 patients qui sont symptomatiques mais dont la pathologie n'est pas identifiée sur le plan génétique. Parmi ces patients, on peut noter qu'il y en a 107 qui sont des « cas isolés » et 57 qui possèdent des antécédents familiaux.

La dernière catégorie isole 4 patients symptomatiques pour lesquels on a retrouvé une consanguinité, qui peut éventuellement être à l'origine de leur pathologie.

4.1.2. L'exclusion des patients

Dans un premier temps, à l'aide du tableau listant la totalité des patients, on peut exclure les patients ne présentant pas de pathologie génétique, ceux qui ont éventuellement une pathologie neurogénétique et les patients sains. Cela représente

125 patients. Aussi, nous excluons 14 patients qui sont porteurs d'une anomalie génétique et 177 patients pour qui la mutation a été identifiée.

Ensuite, on recherche les maladies génétiques autosomiques récessives afin d'augmenter les chances de déceler une mutation. Il faut donc exclure les patients présentant une pathologie qui se transmet selon un mode autosomique dominant. Parmi les 507 sujets qui ont une pathologie autosomique dominante, 94 sont déjà exclus précédemment car la mutation est déjà identifiée. On exclut donc les 413 patients qui présentent une pathologie qui apparaît clairement autosomique dominante.

Aussi, nous pouvons exclure les 7 patients qui ont une pathologie qui est liée à l'X.

D'autres part, parmi les pathologie autosomiques récessives, on sélectionne les fratries atteintes, c'est à dire les sujets qui ont au moins un frère ou une sœur qui est atteint de la même maladie. Pour cela, on exclut les « cas isolé », c'est à dire les patients qui n'ont aucun antécédents familiaux, qui sont seuls dans leur famille à présenter les symptômes. On exclut donc 372 patients.

On peut aussi mettre de côté les 98 patients qui ont des antécédents familiaux mais pour lesquels le dossier médical ne contient pas assez d'information.

Il nous reste donc 44 dossiers présélectionnés. Ces dossiers ont été présentés en réunion au Pr Verny ainsi qu'au Pr Bonneau. L'objectif était de réaliser une dernière sélection, en fonction de leurs connaissances des dossiers pré-sélectionnés. Ce travail a aboutit à la conservation de 14 dossiers. Parmi ces 14 patients, certains ont été retenus en raison de leur lien familial avec l'un des autres patients dont le dossier était déjà retenu, malgré le fait que l'on ait peu d'informations à leur sujet.

L'ensemble des ces résultats sont représentés sous forme d'un diagramme en figure 17.

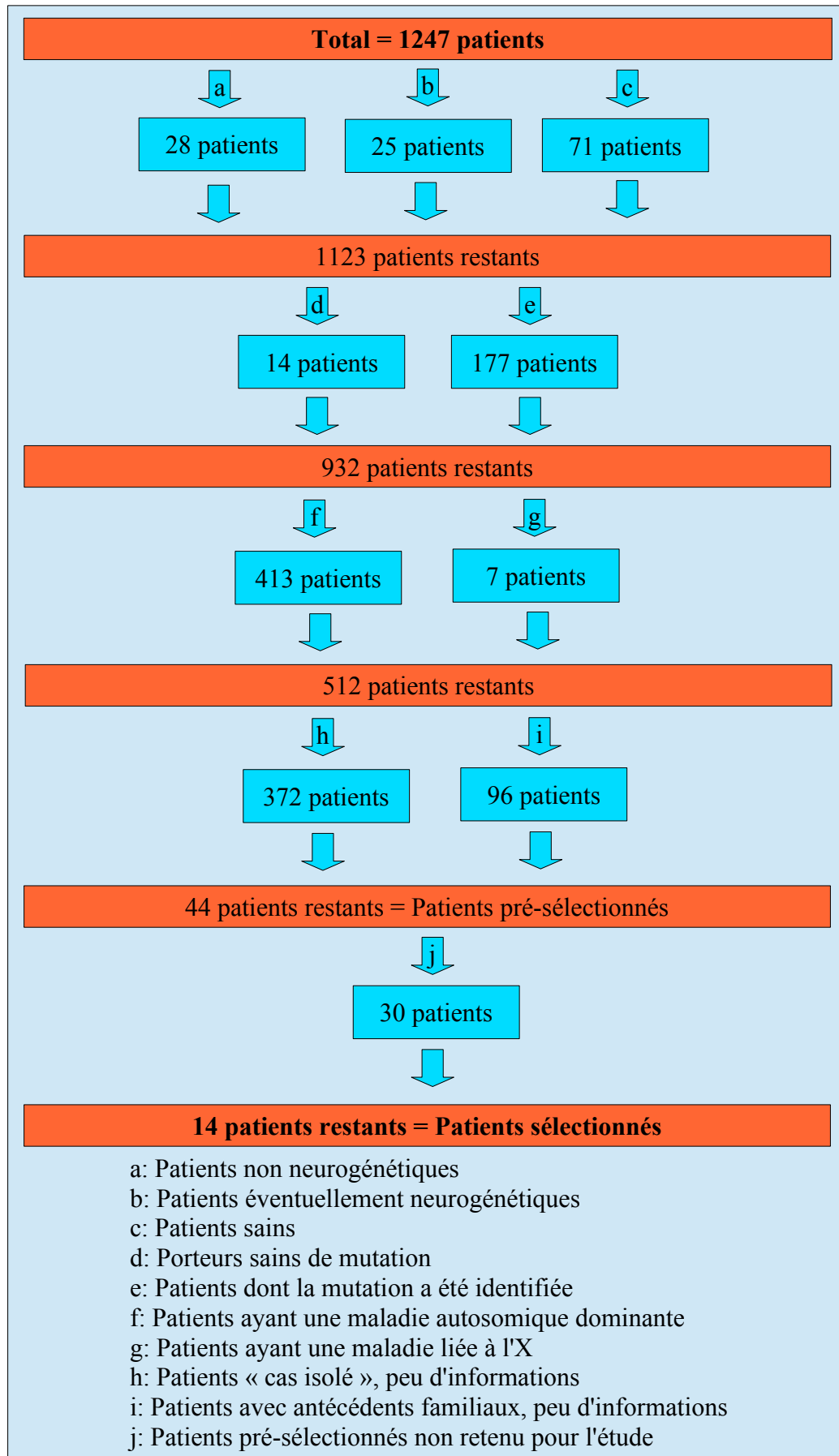


Figure 17: Exclusion des dossiers aboutissant à la sélection finale des patients

4.2. Les dossiers sélectionnés

Le récapitulatif de ces 14 dossiers, dans le tableau 7, a été réalisé afin de le transmettre au Dr Guichet, du laboratoire du CHU d'Angers, pour que les analyses soient réalisées. Pour chaque patient, un code de classification a été attribué afin de conserver leur anonymat, qui se présente sous la forme d'une lettre et d'un chiffre. La lettre correspond à la famille à laquelle appartient le patient car certains patients sélectionnés appartiennent à la même fratrie. Le chiffre correspond donc au patient, pour différencier les patients appartenant à la même famille.

Famille	Patient	Age (ans)	Symptômes	Commentaires
A	1	83	Ataxie spino-cérébelleuse, contexte clinique associant un syndrome cérébelleux et un syndrome tétrapyramidal évoluant depuis 20 ans	Parents consanguins, sœur atteinte ? Friedreich éliminé, pas de mutation dans les gènes SCA1, 2, 3, 6, 7, ni dans le gène POLG
B	2	72	Ataxie spinocérébelleuse autosomique récessive, sans neuropathie, avec un début tardif associé à une atrophie cérébelleuse très lentement progressive	1 frère atteint, gène AOA1 exclu en 2007
C	3	61	Ataxie spinocérébelleuse	1 sœur et 1 frère atteints, pas de consanguinité. Friedreich éliminé chez son frère en 1996
D	4	36	Tableau clinique d'HNPP	Pas de mutation retrouvée, 1 sœur atteinte, parents sains
E	5	52	Parkinson à début précoce	Parents consanguins, sœur atteinte
E	6	44		Sœur de E5
F	7	77	Ataxie spino-cérébelleuse, neuropathie axonale sensitivo-motrice périphérique associée	1 sœur atteinte, pas de consanguinité
G	8	21	Paraparésie spastique, retards des acquisitions, maculopathie	1 frère atteint, parents et grands-parents consanguins : gènes en cours à la Salpêtrière : les résultats ?

G	9	25	Paraparésie spastique, retard des acquisitions	1 frère atteint, parents et grands-parents consanguins
H	10	77	Insensibilité à la douleur, anosmie, pas de retard	1 sœur atteinte, parents non originaires de la même région
H	11		Insensibilité à la douleur, anosmie, pas de retard	sœur de H10
I	12	74	Ataxie cérébelleuse, léger retard mental, épilepsie	
C	13		Parents originaires de la même région	frère de C3
F	14			sœur de F7

Tableau 7 : Listing des 14 patients sélectionnés.

Pour effectuer les tests génétiques par SNP array, le Dr Guichet a utilisé les échantillons de sang veineux disponibles au CHU pour chaque patient. Les échantillons étaient à disposition pour 9 patients sur les 14 initialement prévus. Les 9 échantillons ont été passés sur la plate-forme au printemps 2014.

Il existe des bases de données qui répertorient les polymorphismes. Database of Genomic Variant [36] et la base Illumina [37] notamment. Illumina est une enseigne qui regroupe tous les services correspondant aux analyses de l'ADN : séquençage, génotypage, analyse des gènes de régulation, épigénétique, ... Ils fournissent le matériel d'analyse ainsi que les logiciels permettant de réaliser et de comparer les résultats des tests.

Les résultats de chaque analyse sont comparés à ces bases de données. Les résultats sont ensuite interprétés par le Dr Guichet. Ils sont détaillés ci après.

4.2.1. Patient A1

a) Présentation du patient

Le premier patient, numéroté A1, est un homme âgé de 83 ans. Il présente de nombreux symptômes. En effet, on peut décrire un tableau clinique associant un syndrome cérébelleux et un syndrome tétra-pyramidal qui évolue depuis plus de 20 ans.

On peut noter que ses parents sont consanguins. Il a une sœur atteinte des mêmes symptômes.

Au niveau génétique, la maladie de Friedreich a déjà été exclue. De plus, il n'a pas été retrouvé de mutation dans les gènes SCA 1, SCA2, SCA 3, SCA 6, SCA 7, ni dans le gène POLG.

b) Résultats de l'analyse par SNP array

Au niveau de l'analyse par SNP array, « aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence. » Aussi, la perte d'hétérozygotie au niveau de nombreuses régions chromosomique a été remarquée, à hauteur de 9,5% soit 284Mb cumulées. 13 chromosomes sont concernés. Cette particularité est en faveur d'une consanguinité. De plus, il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données ou chez Illumina. Ces résultats ont été mis en annexe II.

4.2.2. Patient C3

a) Présentation du patient

Le patient C3 est une femme âgée de 61 ans. Elle présente une ataxie spino-cérébelleuse.

On ne relève apparemment pas de consanguinité mais ses parents sont originaires de la même région. Cette patiente a un frère et une sœur atteints de la même pathologie. Chez son frère, le patient C13, les recherches génétiques ont éliminé la maladie de Friedreich. L'analyse par SNP array n'a pas pu être réalisée car son sang n'était pas disponible au laboratoire du CHU d'Angers.

b) Résultats de l'analyse par SNP array

Au niveau de l'analyse par SNP array, « aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence. » De plus, il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données ou chez Illumina. Ces résultats ont été mis en annexe III.

4.2.3. Patient D4

a) Présentation du patient

Le patient D4 est une femme âgée de 36 ans. Elle présente un tableau clinique caractéristique de la neuropathie héréditaire avec hypersensibilité à la pression, HNPP.

A l'heure actuelle, aucune mutation n'a été identifiée. Elle a une sœur atteinte des mêmes symptômes. Ses parents sont sains. Il n'y a à priori pas de consanguinité dans la famille.

b) Résultats de l'analyse par SNP array

Au niveau de l'analyse par SNP array, « aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence. » La délétion en 17p11,2 du gène PMP22, à l'origine des HNPP, n'a pas été observée. De plus, il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données ou chez Illumina. Ces résultats ont été mis en annexe IV.

4.2.4. Patient E6

a) Présentation du patient

Le patient E6 est une femme âgée de 44 ans. Les symptômes retrouvés chez elle indiquent un Parkinson à début précoce.

Ses parents sont consanguins et elle a un frère atteint de la même pathologie. Son frère est le patient numéroté E5. Il n'y avait pas pour lui de sang disponible pour réaliser la SNP array.

b) Résultats de l'analyse par SNP array

Au niveau de l'analyse par SNP array, « aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence. » Aussi, la perte d'hétérozygotie au niveau de nombreuses régions chromosomique a été remarquée, à hauteur de 7,5% soit 226Mb cumulées. 13 chromosomes sont concernés. Cette particularité est en faveur d'une consanguinité. De plus, il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données ou chez Illumina. Ces résultats ont été mis en annexe V.

4.2.5. Patient F7

a) Présentation du patient

Le patient F7 est un homme âgé de 77 ans. Il présente une ataxie spino-cérébelleuse associée à une neuropathie axonale sensitivo-motrice.

Il n'y a apparemment pas de consanguinité dans la famille. Il a une sœur atteinte de la même maladie. Sa sœur est la patiente F14, l'analyse n'a pas été réalisée pour elle car il n'y avait pas de sang disponible pour la SNP array au laboratoire du CHU d'Angers.

b) Résultats de l'analyse par SNP array

Au niveau de l'analyse par SNP array, « aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence. » De plus, il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données ou chez Illumina. Ces résultats ont été mis en annexe VI.

4.2.6. Patient G8

a) Présentation du patient

Le patient G8 est un homme âgé de 21 ans. Il présente une paraparésie spastique, des retards des acquisitions et une maculopathie.

Il a un frère atteint des mêmes symptômes, le patient G9, que nous détaillerons juste après. Dans sa famille, on peut noter la consanguinité entre ses parents ainsi qu'entre ses grands-parents. Il y aurait des analyses en cours à l'hôpital de la Salpêtrière mais il n'y a pas encore de résultats.

b) Résultats de l'analyse par SNP array

Au niveau de l'analyse par SNP array, « aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence. » Aussi, la perte d'hétérozygotie au niveau de nombreuses régions chromosomique a été remarquée, à hauteur de 14,5% soit 436Mb cumulées. 16 chromosomes sont concernés. Cette particularité est en faveur d'une consanguinité. Le patient partage avec son frère 118Mb. De plus, il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données ou chez Illumina. Ces résultats ont été mis en annexe VII.

4.2.7. Patient G9

a) Présentation du patient

Le patient G9 est un homme âgé de 25 ans. Il présente une paraparésie spastique, des retards des acquisitions.

Il a un frère atteint des mêmes symptômes, le patient G8, que nous avons vu précédemment. Dans sa famille, on peut noter la présence d'un coefficient de consanguinité important car ses parents ainsi que ses grands-parents sont consanguins.

b) Résultats de l'analyse par SNP array

Au niveau de l'analyse par SNP array, « aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence. » Aussi, la perte d'hétérozygotie au niveau de nombreuses régions chromosomique a été remarquée, à hauteur de 14% soit 430Mb cumulées. 14 chromosomes sont concernés. Cette particularité est en faveur d'une consanguinité. Le patient partage avec son frère 118Mb. De plus, il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données ou chez Illumina. Ces résultats ont été mis en annexe VIII.

4.2.8. Patient H10

a) Présentation du patient

Le patient H10 est un homme âgé de 77 ans. Il présente une insensibilité à la douleur, une anosmie. En revanche, il n'a pas été observé de retard.

Les parents ne sont pas originaires de la même région donc la consanguinité est très peu probable. Il a une sœur atteinte des mêmes symptômes, la patiente H11, qui n'a pas bénéficié de l'analyse par SNP array car son sang n'était pas disponible au laboratoire.

b) Résultats de l'analyse par SNP array

Au niveau de l'analyse par SNP array, « aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence. » De plus, il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données ou chez Illumina. Ces résultats ont été mis en annexe IX.

4.2.9. Patient I12

a) Présentation du patient

Le patient I12 est un homme âgé de 44 ans. Les symptômes décrits sont une ataxie cérébelleuse, un léger retard mental, une épilepsie.

Ce patient est originaire de Haïti, il y a une possible consanguinité. Il a une sœur et un frère qui sont atteints des mêmes symptômes mais ils ne sont pas suivis au CHU d'Angers.

b) Résultats de l'analyse par SNP array

Au niveau de l'analyse par SNP array, « aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence. » De plus, il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données ou chez Illumina. Ces résultats ont été mis en annexe X.

4.2.10. Les régions LOH communes entre les patients G8 et G9

Suite aux analyses par SNP array réalisées au laboratoire du CHU d'Angers, des régions LOH communes ont pu être mises en évidence entre les patients G8 et G9 qui sont frères. Ces frères présentent des symptômes similaires, notamment une paraparésie spastiques et un retard des acquisitions. Le patient G8 présentent aussi une maculopathie, ou lien ou pas avec la pathologie génétique. Nous pouvons alors supposer que leur pathologie a une origine génétique qui se trouve dans les régions LOH communes. La recherche génétique peut donc se poursuivre. Les régions LOH communes se situent sur 4 chromosomes le chromosome 1, le 12, le 13 et le 16. Les régions sont détaillées ci-après :

- chr1:28.607.616-67.642.223 soit 38 Mb
- chr12:9.737.858-60.028.954 soit 51 Mb
- chr13:76.786.952-92.374.033 soit 16 Mb
- chr16:53.429.205-70.177.571 soit 17 Mb

A partir de ces régions LOH, la recherche des gènes candidats peut être commencée. Pour cela, il existe un site internet qui recense les gènes et les mutations pathogènes déjà décrites dans la littérature :

<http://genome.ucsc.edu/cgi-bin/hgGateway>

A partir de ce site, chaque région est entrée dans le moteur de recherche. Pour chacune des régions, nous regardons alors les gènes impliqués dans des pathologies et nous recherchons en particulier des gènes pour lesquels les mêmes symptômes que les patients G8 et G9 ont été décrits. Nous avons un

résultat pour la région du chromosome 12. Nous trouvons 4 gènes impliqués dans des paraplégies spastiques autosomiques récessives :

- B4GALNT1
- GALGT
- GALNACT
- SPG26

Ces gènes deviennent des « gènes candidats ». C'est-à-dire qu'ils sont peut-être en cause dans la pathologie génétique étudiée. Ensuite, les publications décrivant ces gènes sont recherchées sur le moteur de recherche Pubmed. Nous retenons en particulier 3 publications [38] [39] [40]. Elles seront étudiées par les médecins afin de comparer les cas décrits dans les publications aux cas cliniques qui nous intéressent.

5 Discussion et conclusion

Le travail présenté dans ce manuscrit est centré sur une étude détaillée de dossiers médicaux du service de neurogénétique du CHU d'Angers. L'objectif de celle-ci s'articulait autour de la sélection de dossiers médicaux dont l'absence de diagnostic pourrait être palliée par une nouvelle technique d'analyse de l'ADN. Cette technique très récente, la SNP array, réalisée au laboratoire du CHU d'Angers, permet en effet d'identifier précisément certaines mutations génétiques pathogènes, localisées sur les parties codantes de l'ADN, les exons. Cette méthode étudie les SNP, Single Nucleotide Polymorphism, qui correspondent à des variations sur des sites uniques de l'ADN. Il en a été identifiés 50 millions dans le génome humain. L'objectif est d'identifier les variations pathogènes et de laisser de côté les polymorphismes non pathogènes.

Dans un premier temps, l'analyse de la patientèle du service de neurogénétique a pu être effectuée grâce à l'établissement d'un tableau recensant les patients ayant été reçu en consultation dans ce service. Nous pouvons noter que certains dossiers n'ont pas été consultés car ils n'étaient pas disponibles dans les locaux des archives. Ces dossiers se trouvaient auprès des secrétaires ou des médecins car la consultation de ces patients était récente. On estime à trente le nombre de dossiers dans ce cas. Ces dossiers non consultables réduisent d'autant l'échantillon analysé et limitent l'étendue de l'analyse réalisée ici.

Il a été possible, grâce au tableau récapitulatif qui a été construit, de quantifier le nombre de patients pour chaque pathologie déjà diagnostiquée, et de regrouper les patients symptomatiques qui n'ont pas encore été diagnostiqués. Ce dernier groupe compte 168 patients dont 4 présentaient un coefficient de consanguinité.

Dans un second temps, l'objectif de sélection des patients « pertinents » a été rempli. Il était nécessaire de réduire l'échantillon étudié car les analyses par SNP array ont un coût et on ne peut pas faire de recherche systématique pour tous les patients du service. Il a donc fallu effectuer un travail rigoureux de sélection. On recherchait en particulier les patients ayant une pathologie génétique symptomatique, non diagnostiquée jusqu'à présent, et présentant à priori un mode de transmission autosomique récessif. De plus, lorsque l'on observait un coefficient de consanguinité pour le patient, son dossier attirait tout particulièrement notre attention. En effet, la

présence d'un coefficient de consanguinité dans une famille permet d'augmenter les chances de découvertes de la mutation en cause. Du fait de leur consanguinité, deux membres atteints dans une fratrie présentent une fraction en commun de leur génome plus grande que la moyenne de la population générale. C'est pourquoi le critère de consanguinité est évalué en calculant le pourcentage de régions avec perte d'hétérozygotie, appelées régions avec LOH. C'est dans cette fraction commune d'ADN que l'on suspecte la présence de la mutation pathogène en cause dans cette fratrie. La mutation pathogène sera alors strictement identique chez les deux membres atteints si son expression est issue de la consanguinité familiale. En effet, la transmission du même gène pathogène à caractère récessif est favorisée dans le cas de parents consanguins.

Afin de procéder de façon adéquate, notre sélection a été réalisée par l'intermédiaire de critères dits « d'exclusion ». On a pu retenir un nombre restreint de patients pouvant entrer dans le cadre de notre étude. Ces critères d'exclusion ont permis de retirer de l'étude :

- les patients qui n'ont pas de pathologie neurogénétique
- les patients qui ont peut-être une pathologie neurogénétique
- les patients sains
- les patients porteurs sains de mutation
- les patients dont la mutation a déjà été identifiée
- les patients ayant une maladie autosomique dominante
- les patients ayant une pathologie liée à l'X.

Par ailleurs, les patients pour lesquels on a peu d'informations, considérés comme des « cas isolés » ou avec antécédents familiaux, ont également été mis de côté. Parmi eux, il y avait peut-être des patients pour qui l'analyse aurait donné un résultat mais le manque d'informations consignées dans leurs dossiers ne nous permettait pas de les inclure. Par exemple, certains patients « isolés » étaient peut-être issus d'une union consanguine mais l'information n'est pas toujours connue du neurogénéticien ou du patient lui-même. Le manque d'information limite les recherches génétiques dans les cas de certains patients.

Ce travail d'exclusion, basé sur les nombreux critères cités précédemment, a réduit de façon très importante l'échantillon initial de 1247 patient à un échantillon de

44 dossiers dits « pré-sélectionnés », représentant 3,5% de l'échantillon de départ. Au sein de ces 44 patients, 14 ont été retenus. Cela constitue au final 1,1% de l'échantillon de départ. Cette sélection finale des dossiers a été décidée en accord avec les neurogénétiiciens du service, grâce à leurs connaissances des dossiers des patients. Ces 14 dossiers sélectionnés ont pu faire l'objet d'une demande d'analyse génétique par SNP array auprès du laboratoire du CHU d'Angers et constituent le cœur de l'analyse présentée dans cette thèse.

Sur ces 14 dossiers, les analyses génétiques par SNP array ont pu être réalisées pour 9 d'entre eux. En effet, pour ce sous échantillon de 9 patients, du sang veineux était stocké au laboratoire et était donc directement disponible pour réaliser l'analyse. Pour clore l'étude de ces 14 dossiers, il faudrait réaliser des prises de sang chez les 5 patients restants, lors de leur prochaine consultation. Après recueil de leur consentement pour la réalisation de recherches génétiques, la comparaison aux dossiers des frères ou sœurs qui ont déjà bénéficié de l'analyse pourra être réalisée.

Les résultats d'analyses par SNP array des 9 patients n'ont pas permis de mettre directement en évidence des mutations à l'origine des troubles qu'ils présentent.

En revanche, il a été possible d'exclure la mutation 17p11,2 du gène PMP22 pour le patient D4. Cela constitue une avancée relative dans les recherches génétiques pour ce patient et un premier résultat périphérique de cette thèse. L'exclusion d'une mutation est toujours informative dans le cas d'une maladie génétique.

L'absence de résultats directs dans cette étude peut s'expliquer en partie de la façon suivante. Malgré la haute résolution de la méthode, certaines anomalies ne peuvent être détectées. La SNP array étant une méthode quantitative, elle ne peut mettre en évidence les anomalies chromosomiques équilibrées de type translocation ou inversion. Toute pathologie liée à ces types de mutations ne pourra être décelée par cette méthode. On pointe ici du doigt l'une des principales limites de la SNP array. Dans ce cadre, le caryotype conserve une place de choix dans les explorations génétiques. Il pourra être proposé aux patients qui n'ont pas encore bénéficié de cette analyse.

Suite aux analyses par SNP array des 9 échantillons sanguins, la consanguinité a pu être confirmée pour 4 patients : A1, E6, G8 et G9. Pour le patient A1, le pourcentage de régions avec LOH atteint 9,5%, 7,5% pour le patient E6, 14,5% pour le patient G8 et enfin 14% pour le patient G9. Ces valeurs élevées confirment l'existence d'une

consanguinité chez ces patients. Même si aucun lien explicite n'a pu être établi à ce jour entre cette consanguinité et les pathologies constatées chez les patients étudiés, ces preuves de consanguinité constituent un second résultat périphérique de cette thèse.

Pour les patients A1 et E6, il serait pertinent de comparer ces régions avec LOH à celle de leurs frères ou sœurs atteints pour lesquels il n'y avait pas de sang disponible au laboratoire.

Pour les patients G8 et G9, qui sont frères, les analyses ont révélé des régions LOH en commun dans leurs deux patrimoines génétiques. Après des recherches sur une base de données en ligne, Human (Homo sapiens) Genome Browser Gateway, des « gènes candidats » à l'origine de leur pathologie ont été identifiés. Ces gènes peuvent être à l'origine de la pathologie car des cas cliniques similaires (présentant une paraplégie spastique) ont été décrits en lien avec ces gènes et ont fait l'objet de publications [38] [39] [40]. La suite du processus d'analyse consiste en la comparaison des cas décrits dans ces publications aux cas cliniques qui nous intéressent, à savoir les patients G8 et G9. L'identification de ces gènes candidats, au sein de ce travail de thèse, constitue un résultat fort. La suite des explorations constitue une tâche qui demande de l'expérience et des compétences médicales. Les recherches vont donc être poursuivies par les neurogénétiiciens du service et le laboratoire du CHU d'Angers.

Malgré le faible nombre de dossiers pouvant aboutir à un résultat significatif, cette étude ouvre un certain nombre de pistes dans l'interprétation de cas qui sont à l'heure actuelle inexpliqués sur le plan génétique. Il est important de continuer les explorations car le fait d'identifier les mutations génétiques pathogènes offre de nombreuses perspectives.

En effet, lorsqu'une mutation est retrouvée et est mise en lien direct avec la pathologie que présente le patient, cela permet d'approfondir les connaissances du génome, de vérifier si plusieurs patients possèdent la même mutation. D'autre part, pour le patient, cela lui permet d'expliquer ses symptômes, de mieux comprendre sa maladie et donc de mieux l'accepter.

La perspective principale qu'ouvre l'identification des mutations pathogènes est à ce jour la thérapie génique. Ce concept a été introduit dans les années 1950 et les premiers essais chez l'homme ont été réalisés dans les années 1990. Le principe consiste à fournir à l'organisme une séquence nucléotidique correspondant à un gène

fonctionnel qui pourra substituer le gène pathologique. Cela est possible en utilisant un vecteur viral qui véhiculera le gène fonctionnel jusque dans les cellules cibles du patient.

Les champs d'applications de la thérapie génique apparaissent larges. En effet, son indication apparaît évidente et attendue en ce qui concerne les maladies héréditaires rares. En revanche, elle est testée aujourd'hui dans de nombreuses autres pathologies telles que la maladie de Parkinson, certains cancers, le SIDA, la mucoviscidose par exemple [41].

Malgré les nombreuses recherches et les nombreux essais menés aujourd'hui, la thérapie génique n'a à l'heure actuelle pas le succès que l'on pouvait en attendre. On pensait qu'elle allait révolutionner la médecine alors que nous en sommes encore au stade de la recherche expérimentale, après des années d'essais ! Nous avons appris beaucoup au niveau génétique avec l'avancée de la technologie mais nous voyons qu'il reste encore un long chemin à parcourir avant de maîtriser l'essence même de la vie : l'ADN.

Bibliographie

[1] Choi M., Sholl U. I., Ji W. et al. – Genetic diagnosis by whole exome capture and massively parallel DNA sequencing – 2009 – PNAS – Vol. 106.

[2] Cole J. W., Stine O. C., Liu X., et al. – Rare Variants in Ischemic Stroke: An Exome Pilot Study – 2012 – PloS ONE – Vol. 7.

[3] Chahrour M. H., Yu T. W., Lim E. T., et al. – Whole-Exome Sequencing and Homozygosity Analysis Implicate Depolarization-Regulated Neuronal Genes in Autism – 2012 – PloS Genetics – Vol. 8.

[4] Jimenez-Escrig A., Gobernado I., Garcia-Villanueva M., et al. – Autosomal recessive emery-dreifuss muscular dystrophy caused by novel mutation (R225Q) in the lamin A/C gene identified by exome sequencing – 2012 – Muscle & Nerve – 605.

[5] Ku C.-S., Cooper D. N., Polychronakos C., et al. – Exome Sequencing: Dual Role as a Discovery and Diagnostic Tool – 2012 – American Neurological Association – Vol. 71, N° 1.

[6] Nobelprize.org – The Nobel Prize in Physiology or Medicine 1962 – [en ligne] – http://www.nobelprize.org/nobel_prizes/medicine/laureates/1962/ – consulté le 06 octobre 2013.

[7] Weil J.-H. – Biochimie générale – ISBN 978-2-10-053010-6 – Paris – Dunod – 2009 – 725 p..

[8] Beaumont S. – Biologie moléculaire – ISBN 978-2-10-051041-2 – Paris – Dunod – 2007 – 306 p..

[9] Alberts B., Bray D., Hopkin K., et al. – L'essentiel de la biologie cellulaire, 3ème édition – ISBN 978-2-257-20402-8 – Paris – Lavoisier – 2012 – 731 p..

[10] Pierce B. A. – Genetics, A Conceptual Approach, 4th Edition – ISBN 978-1-4292-7606-1 – Basingstoke – W. H. Freeman and Company – 2012 – 721 p..

[11] Larousse – Mitochondrie – [en ligne] – <http://www.larousse.fr/encyclopedie/medical/mitochondrie/14578> – consulté le 1^{er} octobre 2013.

[12] Joyce C. M., Kelley W. S., Grindley N. D. – Nucleotide sequence of the Escherichia coli polA gene and primary structure of DNA polymerase I – 1982 – JBC – Vol. 257, N° 4.

[13] Lucetta P. – Biologie moléculaire en 30 fiches – ISBN 978-2-10-059237-1 – Paris – Dunod – 2013 – 155 p..

[14] Serre J. L. – Génétique, théorie, analyse et ingénierie – ISBN 978-2-10-057428-5 – Paris – Dunod – 2012 – 420 p..

[15] Orphanet, le portail des maladies rares et des médicaments orphelins – Maladie de Charcot-Marie-Tooth – [en ligne] – http://www.orpha.net/consor/cgi-bin/Disease_Search.php?Ing=FR&data_id=637&Disease_Disease_Search_diseaseGroup=charcot-marie-tooth&Disease_Disease_Search_diseaseType=Pat&Maladie%28s%29/groupe%20de%20maladies=Maladie-de-Charcot-Marie-Tooth&title=Maladie-de-Charcot-Marie-Tooth&search=Disease_Search_Simple – consulté le 20 août 2012.

[16] Azzedine H., Senderek J., Rivolta C., et al. – Molecular Genetics of Charcot-Marie-Tooth Disease : From Genes to Genomes – 2012 – Molecular Syndromology – 3:204-214.

[17] Roos R. – Huntington's disease : a clinical review – 2010 – Orphanet Journal of Rare Diseases – 5:40.

[18] Orphanet, le portail des maladies rares et des médicaments orphelins – Maladie de Huntington – [en ligne] – http://www.orpha.net/consor/cgi-bin/Disease_Search.php?Ing=FR&data_id=118&Disease_Disease_Search_diseaseGroup=huntington&Disease_Disease_Search_diseaseType=Pat&Maladie%28s%29/groupe%20de%20maladies=Maladie-de-Huntington&title=Maladie-de-Huntington&search=Disease_Search_Simple – consulté le 20 août 2012.

[19] Orphanet, le portail des maladies rares et des médicaments orphelins – Ataxie de Friedreich – [en ligne] – http://www.orpha.net/consor/cgi-bin/Disease_Search.php?Ing=FR&data_id=45&Disease_Disease_Search_diseaseGroup=friedreich&Disease_Disease_Search_diseaseType=Pat&Maladie%28s%29/groupe%20de%20maladies=Ataxie-de-Friedreich&title=Ataxie-de-Friedreich&search=Disease_Search_Simple – consulté le 20 août 2012.

[20] Orphanet, le portail des maladies rares et des médicaments orphelins – Neuropathie optique héréditaire de Leber – [en ligne] – http://www.orpha.net/consor/cgi-bin/Disease_Search.php?Ing=FR&data_id=167&Disease_Disease_Search_diseaseGroup=leber&Disease_Disease_Search_diseaseType=Pat&Maladie%28s%29/groupe%20de%20maladies=Neuropathie-optique-hereditaire-de-Leber&title=Neuropathie-optique-hereditaire-de-Leber&search=Disease_Search_Simple – consulté le 20 août 2012.

[21] Orphanet, le portail des maladies rares et des médicaments orphelins – Dystrophie myotonique de type 1 – [en ligne] – http://www.orpha.net/consor/cgi-bin/Disease_Search.php?Ing=FR&data_id=77&Disease_Disease_Search_diseaseGroup=steinert&Disease_Disease_Search_diseaseType=Pat&Maladie%28s%29/groupe%20de%20maladies=Dystrophie-myotonique-de-type-1--Maladie-de-Steinert-&title=Dystrophie-myotonique-de-type-1--Maladie-de-Steinert-&search=Disease_Search_Simple – consulté le 20 août 2012.

[22] Kallioniemi A., Kallioniemi O. P., Sudar D., et al. – Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors – 1992 – Science – 258(5083) : 818-21.

[23] Solinas-Toldo S., Lampel S., Stiglbauer S. et al. – Matrix-based comparative genomic hybridization : biochips to screen for genomic imbalances – 1997 – Genes Chromosome Cancer – 20(4) : 399-407.

[24] Pinkel D., Seagraves R., Sudar D. et al. – High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays – 1998 – Nat Genet – 20(2) : 207-11.

[25] Malan V., Lapiere J.-M., Vekemans M. et al. – La CGH array : un bouleversement de la pratique hospitalière en cytogénétique – 2007 – IRBM – Volume 28.

[26] GENOSCOPE, Centre National de Séquençage – Le projet Génome humain – [en ligne] – <http://www.genoscope.cns.fr/spip/Le-projet-Genome-humain.html> – consulté le 5 octobre 2014.

[27] Rauch A. – Le séquençage nouvelle génération, un saut quantique dans l'explication des maladies génétiques – 2012 – Génétique médicale – 12(3):44-46.

[28] Andrieux J. – Puces à ADN (CGH array) : application pour le diagnostic de déséquilibres cytogénétiques cryptiques – 2008 – Pathologie Biologie – Volume 56.

[29] Craigen W., Graham B., Wong L.-J., et al. – Exome sequencing of a patient with suspected mitochondrial disease reveals a likely multigenic etiology – 2013 – BMC Medical Genetics – 14:83.

[30] Tayeh M., Chin E., Miller V., et al. – Targeted comparative genomic hybridization array for the detection of single- and multiexon gene deletions and duplications – 2009 – Genetics in Medicine – Volume 11.

[31] Aradhya ., Lewis R., Bonaga T., et al. – Exon-level array CGH in a large clinical cohort demonstrates increased sensitivity of diagnostic testing for Mendelian disorders – 2012 – Genetics in Medicine – Volume 14.

[32] Wang J., Zhan H., Li F.-L., et al. – Targeted array CGH as a valuable molecular diagnostic approach : Experience in the diagnosis of mitochondrial and metabolic disorders – 2012 – Molecular Genetics and Metabolism – Volume 106.

[33] L. Papic, D. Fischer, S. Trajanoski, et al. – SNP-array based whole genome homozygosity mapping : A quick and powerful tool to achieve an accurate diagnosis in LGMD2 patients – 2011 – European Journal of Medical Genetics – Volume 54.

[34] A. Drywa, A. Pocwierz-Kotus, S. Dobosz, et al. – Identification of multiple diagnosis SNP loci for differentiation of three salmonid species using SNP-arrays – 2014 – Marine Genomics – Volume 15.

[35] B. Xiong, K. Tan, Y. Tan, et al. – Using SNP array to identify aneuploidy and segmental imbalance in translocation carriers – 2014 – Genomics Data – Volume 2.

[36] Database of Genomic Variant – A curated catalogue of human genomic structural variation – [en ligne] – <http://dgv.tcag.ca/dgv/app/home> – consulté le 6 octobre 2014.

[37] Illumina – [en ligne] – <http://www.illumina.com/> – consulté le 6 octobre 2014.

[38] P. A. Wilkinson, M. A. Simpson, L. Bastaki, et al. – A new locus for autosomal recessive complicated hereditary spastic paraplegia (SPG26) maps to chromosom 12p11.1-12q14 – 2005 – J Med Genet – volume 42.

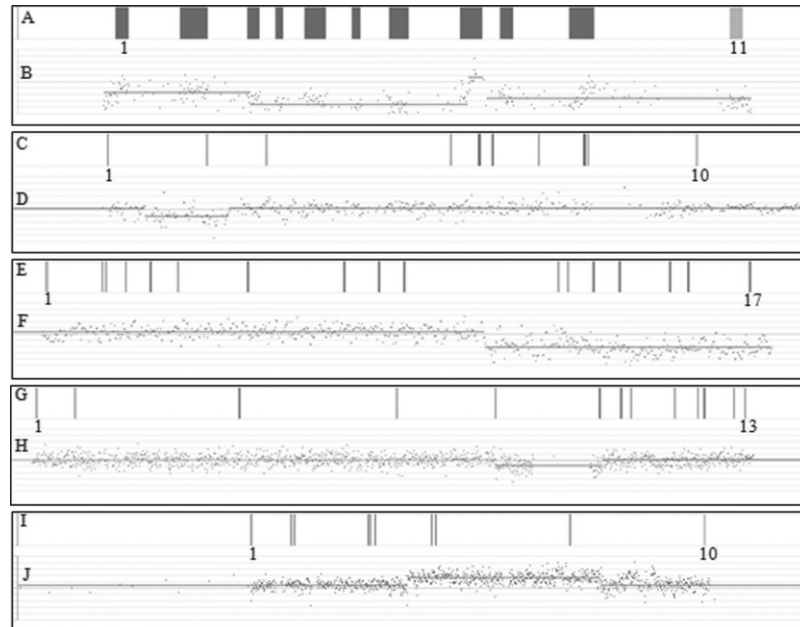
[39] A. Boukhris, R. Schule, J. L. Loureiro, et al. – Alteration of Ganglioside Biosynthesis Responsible for Complex Hereditary Spastic Paraplegia – 2013 – The American Journal of Human Genetics – volume 93.

[40] G. V. Harlalka, A. Lehman, B. Chioza, et al. – Mutations in BAGALNT1 (GM2 synthase) underlie a new disorder of ganglioside biosynthesis – 2013 – Brain, A Journal Of Neurology – Volume 136.

[41] O. Cohen-Haguenauer – Thérapie génique des maladies rares – 2011 – La Revue de médecine interne – Volume 32.

Annexe I

Résultat de l'étude de l'équipe de M. Tayeh en 2009





Targeted gene CGH array for detecting single- and multiexon gene deletion and duplication mutations. A, Horizontal lines represent all 11 exons of the GALT gene. B, Bar graph displays loss of genomic material that corresponds to all exons of the GALT gene, except for exon 8. C, Horizontal lines represent all 10 exons of the DBT gene. D, Bar graph displays loss of genomic material that corresponds to exon 2 of the DBT gene. E, Horizontal lines represent all 17 exons of the GALC gene. F, Bar graph displays loss of genomic material that corresponds to exon 11 through exon 17 of the GALC gene. G, Horizontal lines represent all 13 exons of the PAH gene. H, Bar graph displays loss of genomic material that corresponds to exon 5 through exon 6 of the PAH gene. I, Horizontal lines represent all 10 exons of the BCKDHB gene. J, Bar graph displays a gain of genomic material that corresponds to exon 11 through exon 14 of the BCKDHB gene. Loss and gain of genomic material are depicted as a shift down and shift up of the features on the x-axis, respectively.

Tayeh M., Chin E., Miller V., et al. – Targeted comparative genomic hybridization array for the detection of single- and multiexon gene deletions and duplications – April 2009 – Genetics in Medicine – Volume 11, Number 4.

Annexe II

Résultat de la SNP array du patient A1

 CENTRE HOSPITALIER UNIVERSITAIRE	SERVICE DE GENETIQUE Chef de service : Pr. Dominique BONNEAU Tél. secrétariat : 02 41 35 38 83
<u>ANALYSE MOLECULAIRE</u> <u>PAR SNP-ARRAY</u>	
Réf. : _____	A1
DOS n° : _____	Tissu : Sang veineux
NOM : _____	Prénom : _____
D.N. : _____	Indication : Postnatal 6- Autre ataxie spinocerebelleuse + syndrome tétrapyramidale
Médecin(s) : Pr Verny Christophe	
Date du Prélèvement : ?	Date de Réalisation : 21/05/2014
<hr/>	
Méthodologie : Puce Illumina à SNP (Single Nucleotide Polymorphism) Infinium assay: CytoSNP-850K, DNA Analysis BeadChip, composée de 850 000 marqueurs qui couvre 3300 gènes ainsi que les régions télomériques, péricentromériques et les chromosomes sexuels. La conception de cette lame a été réalisée conjointement avec le ICCG(auparavant appelé ISCA)(« International Collaboration for Clinical Genomics) et le CCMC (Cancer Cytogenetics Microarray Consortium). Ces SNP couvrent préférentiellement des régions du génome avec un espace médian de l'ordre de 1.8 kb soit une résolution effective de l'ordre de 18kb. Interprétation avec le logiciel Illumina Genome Studio v2011.1 et CNV partition v3.1.6 (Genotyping v.1.9.4 module et Illumina Genome Viewer 1.9.0)(Build 37)(hg19) Contrôle de qualité : Call rate 0.99	
<hr/>	
Résultats : arr(1-22)X2,(XY)X1 Aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence. On observe la présence de nombreuses régions chromosomiques avec perte d'hétérozygotie (LOH) qui concerne 13 chromosomes soit une taille cumulée de 284Mb (9.5%) ce qui est en faveur d'une consanguinité. Les régions avec LOH sont : chr4:104,241,267-144,282,005 (40Mb) chr5:42,526,025-110,401,872 (64Mb) chr6:80,614,816-108,321,035 (28Mb) chr7:112,497,725-130,673,153 (18Mb) chr10:82,708,485-114,726,596 (31Mb) chr16:25,224,437-52,016,720 (15Mb) chr17:25,329,949-57,408,572 (32Mb) chr22:26,814,608-41,526,364 (14Mb).	
<hr/>	
Dr Agnès GUICHET	Date : 19/08/2014
<hr/>	
CHU Angers 4, rue Larrey 49933 ANGERS cedex 9  www.chu-angers.fr	PÔLE DE BIOLOGIE Pr. Pascal REYNIER – Chef de pôle

Annexe III

Résultat de la SNP array du patient C3



SERVICE DE GENETIQUE

Chef de service : Pr. Dominique BONNEAU

Tél. secrétariat : 02 41 35 38 83

ANALYSE MOLECULAIRE PAR SNP-ARRAY

Réf. :

C3

DOS n° :

Tissu : **Sang veineux**

NOM :

Prénom :

D.N. :

Indication : **Postnatal 6- Autre Ataxie spinocerebelleuse**

Médecin(s) : **Dr Agnès GUICHET**

Date du Prélèvement : ?

Date de Réalisation : 22/04/2014

Méthodologie :

Puce Illumina à SNP (Single Nucleotide Polymorphism) Infinium assay: CytoSNP-850K, DNA Analysis BeadChip, composée de 850 000 marqueurs qui couvre 3300 gènes ainsi que les régions télomériques, péricentromériques et les chromosomes sexuels. La conception de cette lame a été réalisée conjointement avec le ICCG(auparavant appelé ISCA)(« International Collaboration for Clinical Genomics) et le CCMC (Cancer Cytogenetics Microarray Consortium). Ces SNP couvrent préférentiellement des régions du génome avec un espace médian de l'ordre de 1.8 kb soit une résolution effective de l'ordre de 18kb. Interprétation avec le logiciel Illumina Genome Studio v2011.1 et CNV partition v3.1.6 (Genotyping v.1.9.4 module et Illumina Genome Viewer 1.9.0)(Build 37)(hg19)

Contrôle de qualité : Call rate 0.99

Résultats :

arr(1-22,X)X2

Aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence.

Il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données(<http://projects.tcag.ca/variation>) ou chez Illumina.

Dr Agnès GUICHET

Date : 29/08/2014

CHU Angers

4, rue Larrey

49933 ANGERS cedex 9

www.chu-angers.fr



PÔLE DE BIOLOGIE

Pr. Pascal REYNIER – Chef de pôle

Annexe IV

Résultat de la SNP array du patient D4



SERVICE DE GENETIQUE

Chef de service : Pr. Dominique BONNEAU

Tél. secrétariat : 02 41 35 38 83

ANALYSE MOLECULAIRE PAR SNP-ARRAY

Réf. :

D4

DOS n° :

Tissu : **Sang veineux**

NOM :

Prénom :

D.N. :

Indication : **Postnatal 6- Autre HNPP**

Médecin(s) : **Pr Christophe VERNY**

Date du Prélèvement : ?

Date de Réalisation : 22/04/2014

Méthodologie :

Puce Illumina à SNP (Single Nucleotide Polymorphism) Infinium assay: CytoSNP-850K, DNA Analysis BeadChip, composée de 850 000 marqueurs qui couvre 3300 gènes ainsi que les régions télomériques, péricentromériques et les chromosomes sexuels. La conception de cette lame a été réalisée conjointement avec le ICCG (auparavant appelé ISCA) (« International Collaboration for Clinical Genomics ») et le CCMC (Cancer Cytogenetics Microarray Consortium). Ces SNP couvrent préférentiellement des régions du génome avec un espace médian de l'ordre de 1.8 kb soit une résolution effective de l'ordre de 18kb. Interprétation avec le logiciel Illumina Genome Studio v2011.1 et CNV partition v3.1.6 (Genotyping v.1.9.4 module et Illumina Genome Viewer 1.9.0)(Build 37)(hg19)

Contrôle de qualité : Call rate 0.99

Résultats :

arr(1-22,X)X2

Aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence.

On n'observe pas de délétion en 17p11.2 impliquant le gène PMP22.

Il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données (<http://projects.tcag.ca/variation>) ou chez Illumina.

Dr Agnès GUICHET

Date : 07/07/2014

CHU Angers

4, rue Larrey

49933 ANGERS cedex 9

www.chu-angers.fr



PÔLE DE BIOLOGIE

Pr. Pascal REYNIER – Chef de pôle

Annexe V

Résultat de la SNP array du patient E6



SERVICE DE GENETIQUE

Chef de service : Pr. Dominique BONNEAU

Tél. secrétariat : 02 41 35 38 83

ANALYSE MOLECULAIRE PAR SNP-ARRAY

Réf. :

DOS n° :

NOM :

D.N. :

E 6

Tissu : Sang veineux

Prénom :

Indication : ??

Médecin(s) : Pr Verny Christophe

Date du Prélèvement : ?

Date de Réalisation : 16/06/2014

Méthodologie :

Puce Illumina à SNP (Single Nucleotide Polymorphism) Infinium assay: CytoSNP-850K, DNA Analysis BeadChip, composée de 850 000 marqueurs qui couvre 3300 gènes ainsi que les régions télomériques, péricentromériques et les chromosomes sexuels. La conception de cette lame a été réalisée conjointement avec le ICCG(auparavant appelé ISCA)(« International Collaboration for Clinical Genomics ») et le CCMC (Cancer Cytogenetics Microarray Consortium). Ces SNP couvrent préférentiellement des régions du génome avec un espace médian de l'ordre de 1.8 kb soit une résolution effective de l'ordre de 18kb. Interprétation avec le logiciel Illumina Genome Studio v2011.1 et CNV partition v3.1.6 (Genotyping v.1.9.4 module et Illumina Genome Viewer 1.9.0)(Build 37)(hg19)

Contrôle de qualité : Call rate 0.99

Résultats :

arr(1-22)X2,(XY)X1

Aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence.

On observe la présence de nombreuses régions chromosomiques avec perte d'hétérozygotie (LOH) qui concerne 13 chromosomes soit une taille cumulée de 226 Mb (7.5%) et qui est en faveur d'une consanguinité :

chr2 : 44,044,680-67,224,057 soit 23 Mb
chr3 : 156,007,973-168,866,906 soit 13 Mb
chr5 : 74,208,396-91,893,792 soit 17 Mb
chr5 : 110,815,311-123,194,446 soit 12 Mb
chr7 : 113,963,334-142,044,942 soit 28 Mb
chr8 : 95,240,271-116,825,690 soit 21 Mb
chr9 : 26,044,500-39,179,289 soit 13 Mb
chr9 : 71,019,573-85,150,039 soit 14 Mb
chr12 : 102,204,244-119,530,488 soit 20 Mb.

Il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données(<http://projects.tcag.ca/variation>) ou chez Illumina.

Dr Agnès GUICHET

Date : 19/08/2014

CHU Angers

4, rue Larrey

49933 ANGERS cedex 9

www.chu-angers.fr



PÔLE DE BIOLOGIE

Pr. Pascal REYNIER – Chef de pôle

Annexe VI

Résultat de la SNP array du patient F7



SERVICE DE GENETIQUE

Chef de service : Pr. Dominique BONNEAU

Tél. secrétariat : 02 41 35 38 83

ANALYSE MOLECULAIRE PAR SNP-ARRAY

Réf. :

F7

DOS n° :

Tissu : Sang veineux

NOM :

Prénom :

D.N. :

Indication : Postnatal 6- Autre ataxie spinocerebelleuse,

neuropathie axonale sensitivo-motrice

Médecin(s) : Pr Verny Christophe

Date du Prélèvement : ?

Date de Réalisation : 16/06/2014

Méthodologie :

Puce Illumina à SNP (Single Nucleotide Polymorphism) Infinium assay: CytoSNP-850K, DNA Analysis BeadChip, composée de 850 000 marqueurs qui couvre 3300 gènes ainsi que les régions télomériques, péricentromériques et les chromosomes sexuels. La conception de cette lame a été réalisée conjointement avec le ICCG(auparavant appelé ISCA)(« International Collaboration for Clinical Genomics ») et le CCMC (Cancer Cytogenetics Microarray Consortium). Ces SNP couvrent préférentiellement des régions du génome avec un espace médian de l'ordre de 1.8 kb soit une résolution effective de l'ordre de 18kb. Interprétation avec le logiciel Illumina Genome Studio v2011.1 et CNV partition v3.1.6 (Genotyping v.1.9.4 module et Illumina Genome Viewer 1.9.0)(Build 37)(hg19)

Contrôle de qualité : Call rate 0.99

Résultats :

arr(1-22)X2,(XY)X1

Aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence.

Il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données(<http://projects.tcag.ca/variation>) ou chez Illumina.

Dr Agnès GUICHET

Date : 19/07/2014

CHU Angers

4, rue Larrey

49933 ANGERS cedex 9

www.chu-angers.fr



PÔLE DE BIOLOGIE

Pr. Pascal REYNIER – Chef de pôle

Annexe VII 1/2

Résultat de la SNP array du patient G8



SERVICE DE GENETIQUE

Chef de service : Pr. Dominique BONNEAU

Tél. secrétariat : 02 41 35 38 83

G8

ANALYSE MOLECULAIRE PAR SNP-ARRAY

Réf. : 32049

DOS n° :

Tissu : Sang veineux

NOM :

Prénom :

D.N. :

Indication : Postnatal 6- Autre paraparésie spastique, retard des

acquisitions et maculopathie

Médecin(s) : Dr Agnès GUICHET

Date du Prélèvement : ?

Date de Réalisation : 21/05/2014

Méthodologie :

Puce Illumina à SNP (Single Nucleotide Polymorphism) Infinium assay: CytoSNP-850K, DNA Analysis BeadChip, composée de 850 000 marqueurs qui couvre 3300 gènes ainsi que les régions télomériques, péricentromériques et les chromosomes sexuels. La conception de cette lame a été réalisée conjointement avec le ICCG(auparavant appelé ISCA)(« International Collaboration for Clinical Genomics ») et le CCMC (Cancer Cytogenetics Microarray Consortium). Ces SNP couvrent préférentiellement des régions du génome avec un espace médian de l'ordre de 1.8 kb soit une résolution effective de l'ordre de 18kb. Interprétation avec le logiciel Illumina Genome Studio v2011.1 et CNV partition v3.1.6 (Genotyping v.1.9.4 module et Illumina Genome Viewer 1.9.0)(Build 37)(hg19)

Contrôle de qualité : Call rate 0.99

Résultats :

arr(1-22)X2,(XY)X1

Aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence.

On observe la présence de nombreuses régions chromosomiques avec perte d'hétérozygotie (LOH) qui concerne 16 chromosomes soit une taille cumulée de 436 Mb (14.5%) et qui est en faveur d'une consanguinité. Voir liste ci dessous

Mohamed partage avec son frère 118Mb qui sont les régions soulignées.

Il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données(<http://projects.tcag.ca/variation>) ou chez Illumina.

Dr Agnès GUICHET

Date : 19/08/2014

CHU Angers

4, rue Larrey

49933 ANGERS cedex 9

www.chu-angers.fr



PÔLE DE BIOLOGIE

Pr. Pascal REYNIER – Chef de pôle

Annexe VII 2/2

Résultat de la SNP array du patient G8



SERVICE DE GENETIQUE

Chef de service : Pr. Dominique BONNEAU

Tél. secrétariat : 02 41 35 38 83

G8

Liste des régions avec perte d'hétérozygotie et celles partagées avec son frere _____ sont soulignées

chr1:1,472,201-13,783,354 soit 12 Mb
chr1:28,607,616-67,642,223 soit 38 Mb
chr1:235,771,151-249,218,992 soit 12 Mb
chr2:73,775,214-90,240,473 soit 16 Mb
chr5:117,714,620-134,456,005 soit 17 Mb
chr6:301,536-16,772,207 soit 16 Mb
chr6:100,288,317-112,012,707 soit 12 Mb
chr7:32,230,177-50,211,199 soit 18 Mb
chr8:29,743,361-42,450,441 soit 12 Mb
chr8:46,942,842-66,129,816 soit 19 Mb
chr10:54,335,375-73,264,879 soit 19 Mb
chr11:57,360,416-86,426,766 soit 28 Mb
chr12:9,737,858-60,103,232 soit 47 Mb
chr13:24,759,298-46,937,193 soit 22 Mb
chr13:76,786,952-92,374,033 soit 16 Mb
chr16:53,429,205-70,177,571 soit 17 Mb
chr18:3,747,619-21,693,892 soit 15 Mb
chr18:45,436,351-60,277,715 soit 15 Mb
chr20:17,670,934-40,912,801 soit 20 Mb.

CHU Angers

4, rue Larrey

49933 ANGERS cedex 9

www.chu-angers.fr



PÔLE DE BIOLOGIE

Pr. Pascal REYNIER – Chef de pôle

Annexe VIII 1/2

Résultat de la SNP array du patient G9



SERVICE DE GENETIQUE

Chef de service : Pr. Dominique BONNEAU

Tél. secrétariat : 02 41 35 38 83

ANALYSE MOLECULAIRE PAR SNP-ARRAY

Réf. :

G9

DOS n° :

Tissu : Sang veineux

NOM :

Prénom :

D.N. :

Indication : Postnatal 6- Autre paraparésie spastique, retard des

acquisitions,

Médecin(s) : Dr Agnès GUICHET

Date du Prélèvement : ?

Date de Réalisation : 21/05/2014

Méthodologie :

Puce Illumina à SNP (Single Nucleotide Polymorphism) Infinium assay: CytoSNP-850K, DNA Analysis BeadChip, composée de 850 000 marqueurs qui couvre 3300 gènes ainsi que les régions télomériques, péricentromériques et les chromosomes sexuels. La conception de cette lame a été réalisée conjointement avec le ICCG(auparavant appelé ISCA)(« International Collaboration for Clinical Genomics ») et le CCMC (Cancer Cytogenetics Microarray Consortium). Ces SNP couvrent préférentiellement des régions du génome avec un espace médian de l'ordre de 1.8 kb soit une résolution effective de l'ordre de 18kb. Interprétation avec le logiciel Illumina Genome Studio v2011.1 et CNV partition v3.1.6 (Genotyping v.1.9.4 module et Illumina Genome Viewer 1.9.0)(Build 37)(hg19)

Contrôle de qualité : Call rate 0.99

Résultats :

arr(1-22)X2,(XY)X1

Aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence.

On observe la présence de nombreuses régions chromosomiques avec perte d'hétérozygotie (LOH) qui concerne 14 chromosomes soit une taille cumulée de 430 Mb (14%) et qui est en faveur d'une consanguinité (voir liste ci-dessous).

Il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données(<http://projects.tcag.ca/variation>) ou chez Illumina.

Dr Agnès GUICHET

Date : 19/08/2014

CHU Angers

4, rue Larrey

49933 ANGERS cedex 9

www.chu-angers.fr



PÔLE DE BIOLOGIE

Pr. Pascal REYNIER – Chef de pôle

Annexe VIII 2/2

Résultat de la SNP array du patient G9



SERVICE DE GENETIQUE

Chef de service : Pr. Dominique BONNEAU

Tél. secrétariat : 02 41 35 38 83

G9

Liste des régions avec perte d'hétérozygotie et régions communes avec son frère
soulignées :

chr1:28,607,616-85,394,258 soit 56 Mb :

chr1:28,607,616-67,642,223 soit 38 Mb :

chr1:171,652,978-199,582,376 soit 28 Mb

chr5:36,902,291-66,276,513 soit 15 Mb

chr5:150,972,154-172,283,930 soit 21 Mb

chr7:116,801,277-136,124,848 soit 19 Mb

chr10:16,488,475-35,174,347 soit 19 Mb

chr11:83,159,958-115,183,128 soit 32 Mb

chr12:9,737,858-60,028,954 soit 51 Mb

chr13:58,248,789-92,374,033 soit 34 Mb

chr13:76,786,952-92,374,033 soit 16 Mb

chr15:61,699,175-74,651,231 soit 12 Mb

chr16:53,429,205-70,177,571 soit 17 Mb

chr18:45,436,351-55,786,359 soit 10 Mb

chr18:60,279,464-74,291,344 soit 14 Mb.

CHU Angers

4, rue Larrey

49933 ANGERS cedex 9

www.chu-angers.fr



PÔLE DE BIOLOGIE

Pr. Pascal REYNIER – Chef de pôle

Annexe IX

Résultat de la SNP array du patient H10



SERVICE DE GENETIQUE

Chef de service : Pr. Dominique BONNEAU

Tél. secrétariat : 02 41 35 38 83

ANALYSE MOLECULAIRE PAR SNP-ARRAY

Réf. :

H10

DOS n° :

Tissu : Sang veineux

NOM :

Prénom :

D.N. :

Indication : Postnatal 6- Autre insensibilité à la douleur, anosmie

Médecin(s) : Pr Verny Christophe

Date du Prélèvement : ?

Date de Réalisation : 16/06/2014

Méthodologie :

Puce Illumina à SNP (Single Nucleotide Polymorphism) Infinium assay: CytoSNP-850K, DNA Analysis BeadChip, composée de 850 000 marqueurs qui couvre 3300 gènes ainsi que les régions télomériques, péricentromériques et les chromosomes sexuels. La conception de cette lame a été réalisée conjointement avec le ICCG(auparavant appelé ISCA)(« International Collaboration for Clinical Genomics ») et le CCMC (Cancer Cytogenetics Microarray Consortium). Ces SNP couvrent préférentiellement des régions du génome avec un espace médian de l'ordre de 1.8 kb soit une résolution effective de l'ordre de 18kb. Interprétation avec le logiciel Illumina Genome Studio v2011.1 et CNV partition v3.1.6 (Genotyping v.1.9.4 module et Illumina Genome Viewer 1.9.0)(Build 37)(hg19)

Contrôle de qualité : Call rate 0.99

Résultats :

arr(1-22)X2,(XY)X1

Aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence.

Il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données(<http://projects.tcag.ca/variation>) ou chez Illumina.

Dr Agnès GUICHET

Date : 19/08/2014

CHU Angers

4, rue Larrey

49933 ANGERS cedex 9

www.chu-angers.fr



PÔLE DE BIOLOGIE

Pr. Pascal REYNIER – Chef de pôle

Annexe X

Résultat de la SNP array du patient I12



SERVICE DE GENETIQUE

Chef de service : Pr. Dominique BONNEAU

Tél. secrétariat : 02 41 35 38 83

ANALYSE MOLECULAIRE PAR SNP-ARRAY

Réf. :

DOS n° :

NOM :

D.N. :

mental, épilepsie

I 12

Tissu : Sang veineux

Prénom :

Indication : Postnatal 6- Autre ataxie cérébelleuse, retard

Médecin(s) : Pr Verny christophe

Date du Prélèvement : ?

Date de Réalisation : 21/05/2014

Méthodologie :

Puce Illumina à SNP (Single Nucleotide Polymorphism) Infinium assay: CytoSNP-850K, DNA Analysis BeadChip, composée de 850 000 marqueurs qui couvre 3300 gènes ainsi que les régions télomériques, péricentromériques et les chromosomes sexuels. La conception de cette lame a été réalisée conjointement avec le ICCG(auparavant appelé ISCA)(« International Collaboration for Clinical Genomics ») et le CCMC (Cancer Cytogenetics Microarray Consortium). Ces SNP couvrent préférentiellement des régions du génome avec un espace médian de l'ordre de 1.8 kb soit une résolution effective de l'ordre de 18kb. Interprétation avec le logiciel Illumina Genome Studio v2011.1 et CNV partition v3.1.6 (Genotyping v.1.9.4 module et Illumina Genome Viewer 1.9.0)(Build 37)(hg19)

Contrôle de qualité : Call rate 0.99

Résultats :

arr(1-22,X)X2

Aucune anomalie chromosomique considérée comme délétère à ce jour et potentiellement responsable de la pathologie actuelle n'a été mise en évidence.

Il existe plusieurs variants répertoriés comme polymorphisme de taille dans les bases de données(<http://projects.tcag.ca/variation>) ou chez Illumina.

Dr Agnès GUICHET

Date : 19/07/2014

CHU Angers

4, rue Larrey

49933 ANGERS cedex 9

www.chu-angers.fr



PÔLE DE BIOLOGIE

Pr. Pascal REYNIER – Chef de pôle

RÉSUMÉ

La SNP array appliquée à une sélection de patients au CHU d'Angers ; Nouvelle méthode d'analyse de l'ADN.

L'étude présentée ici s'articule autour de l'analyse par SNP array d'échantillons sanguins de 9 patients, suivis au sein du service de neurogénétique du CHU d'Angers. Ces patients ont été sélectionnés suite à une consultation rigoureuse de l'ensemble des dossiers présents dans les archives de ce service. Ce travail a permis de construire un tableau recensant la totalité des patients. A partir de ce tableau, il a été possible de classer et donc de sélectionner progressivement, par l'intermédiaire de critères d'exclusion, les cas pertinents de notre étude. L'objectif principal était de sélectionner des patients présentant une pathologie à priori autosomique récessive et ayant au moins un frère ou une sœur atteint. De plus, les patients présentant un coefficient de consanguinité ont attiré particulièrement notre attention. Tout ce processus a permis d'aboutir à la sélection de 14 patients dont 9 ont pu bénéficier de l'analyse par SNP array au sein du laboratoire du CHU d'Angers. Les résultats présentés ici ne mettent pas en évidence directe de mutations en lien avec les pathologies incriminées. En revanche, les informations que les analyses menées ont fourni permettent l'ajout de résultats périphériques aux dossiers des patients. Certains d'entre eux ouvrent même de sérieuses pistes de recherche.

Mots-clés : SNP array, CGH array, hybridation, consanguinité, neurogénétique, CHU d'Angers

ABSTRACT

SNP array applied to a sample of patients of the University Hospital of Angers ; New DNA analysis method.

The study presented here is based on SNP array analysis of blood samples of 9 patients followed in the Neurogenetics unit of the University Hospital of Angers. Using a rigorous reading of all medical reports of this unit, we have built a table. With this table, it is possible to classify and thus to select progressively, through a set of exclusion criteria, relevant cases for our study. The main objective was to select patients with autosomal recessive disease and having at least one brother or sister with the same pathology. In addition, patients having an inbreeding coefficient have attracted our attention.

The overall process has lead to a selection of 14 patients including 9 have benefited to a SNP array analysis performed in the genetic laboratory of the University Hospital of Angers.

The results presented here do not directly highlight any mutation linked to the different analysed pathologies. However, the results provided by these analysis have allowed to add some peripheral results in patient's reports. Some of them open serious perspectives of research.

Keywords : SNP array, CGH array, hybridization, inbreeding, neurogenetics, University Hospital of Angers