



HAL
open science

Le séquençage nouvelle génération du génome mitochondrial dans le diagnostic des cytopathies mitochondriales

Nada Houcinat

► **To cite this version:**

Nada Houcinat. Le séquençage nouvelle génération du génome mitochondrial dans le diagnostic des cytopathies mitochondriales. Médecine humaine et pathologie. 2015. dumas-01218641

HAL Id: dumas-01218641

<https://dumas.ccsd.cnrs.fr/dumas-01218641>

Submitted on 21 Oct 2015

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

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

Université de Bordeaux
U.F.R. DES SCIENCES MEDICALES

Année 2015

Thèse n°3097

**Thèse pour l'obtention du
DIPLOME d'ETAT de DOCTEUR EN MEDECINE**

Présentée et soutenue publiquement

Le 21/09/2015

Par Nada HOUCINAT

Née le 22/01/1985

**Le séquençage nouvelle génération du génome
mitochondrial dans le diagnostic des cytopathies
mitochondriales**

Directeur de thèse : Madame le Pr Marie-Laure LEIBREICH-NEGRIER

Rapporteur de thèse : Madame le Dr Sylvie BANNWARTH

Membres du jury :

Monsieur le Pr Didier LACOMBE (président)
Madame le Dr Patrizia AMATI-BONNEAU (juge)
Monsieur le Pr Benoit ARVEILER (juge)
Madame le Dr Patricia FERGELOT (juge)
Monsieur le Pr Cyril GOIZET (juge)

Remerciements :

Je remercie tous les membres du jury du temps consacré à la lecture de ce travail.

Je remercie monsieur le Pr Didier LACOMBE d'avoir accepté la présidence de ce jury, je le remercie également de m'avoir accueillie au sein du service de génétique médicale de Bordeaux et pour tous ses enseignements.

Je remercie monsieur le Pr Benoit ARVEILER pour ses enseignements et pour m'avoir permis de réaliser ce travail au sein du plateau technique de biologie moléculaire.

Je remercie le Pr Cyril GOIZET de ses encouragements, de tous ses enseignements et de ses conseils tout le long de mon parcours d'interne.

Je remercie le Dr Patricia FERGELOT pour tous les moments d'échanges, ses conseils avisés et pour m'avoir fait bénéficier de ses vastes connaissances.

Je remercie le Dr Patrizia AMATI-BONNEAU du temps consacré à la lecture de ce mémoire de thèse et de s'être déplacée pour cette soutenance.

Je remercie madame le Pr Marie-Laure LEIBREICH-NEGRIER, ma directrice de thèse, pour la confiance qu'elle m'a accordée pour la réalisation de ce travail, de son soutien et de ses corrections.

Je remercie particulièrement le Dr Sylvie BANNWARTH, mon rapporteur de thèse du temps consacré à examiner ce travail malgré un emploi du temps chargé.

Je la remercie également ainsi que tous les membres de l'équipe de génétique médicale du CHU de Nice pour leur accueil chaleureux et pour toute l'aide qu'ils m'ont apportée pour initier ce travail, avec une pensée particulière pour Sonia et Gaëlle que je remercie profondément du temps qu'elles m'ont accordé.

Je remercie toute l'équipe technique du plateau de biologie moléculaire de Bordeaux sans lesquels ce travail n'aurait pas été aisé, je les remercie pour leur bonne humeur, leur bienveillance, leur disponibilité, leurs soutiens et encouragements de tous les jours. Un grand merci à Claudio et à Armelle pour leur investissement dans ce projet, un grand merci à Sylvie, à Gwenaëlle, à Brigitte, aux 2 Julie, à Marie-Hélène, à Catherine et j'en oublie sûrement...

Un grand merci à toute l'équipe de génétique médicale bordelaise qui est une équipe chaleureuse et humaine, c'est un véritable privilège de travailler auprès de vous.

Un grand merci à Danielle pour son amitié et son oreille attentive, un merci aussi à Sonia, Sophie, Fanny, Laetitia, Monique, Cécile, Virginie, Stéphanie, Eulalie, Armelle et là encore j'en oublie sûrement.

Un merci également à toutes les équipes et tous les services qui m'ont accueillie ces dernières années avec une pensée particulière pour l'équipe de biochimie métabolique auprès de laquelle j'ai beaucoup appris, un grand merci aux docteurs Samir MESLI et Isabelle REDONNET-VERNHET.

A Marie-Laure merci de m'avoir accompagnée toutes ces années et pour ton amitié, c'est un réel plaisir pour moi de te connaître sur les deux plans personnel et professionnel. Rares de nos jours sont les personnes comme toi avec une telle bonté et une telle douceur... merci pour tous les moments de partage, j'espère qu'il y en aura beaucoup d'autres.

A ma famille et tous mes amis et en particulier Aurélie, ma compagne de toujours ;-), sans lesquels je ne serai pas là où je suis aujourd'hui. Je les remercie du fond du cœur pour avoir fait de moi la personne que je suis, j'ai beaucoup de chance de vous avoir à mes côtés.

A mon mari Sébastien, il n'existe pas de mots assez forts pour décrire tout l'amour, la reconnaissance et le respect que j'éprouve à ton égard... pour tout merci.

Table des matières

Avant-Propos.....	7
Abréviations.....	8
Première section : Introduction.....	10
I. Les mitochondries :	10
1. Origine et historique.....	10
2. Structure et organisation cellulaire des mitochondries :	11
2.1. Structure et morphologie mitochondriales.....	11
2.2. Le réseau mitochondrial :	13
2.3. Le « turnover » mitochondrial.....	17
3. Le génome mitochondrial :	18
3.1. Structure :	18
3.2. Etudes phylogénétiques	19
3.3. La réplication	20
3.4. La transcription.....	22
3.5. La traduction et le code génétique mitochondrial.....	23
3.6. L'ADNmt au cours du développement embryonnaire : devenir des mitochondries d'origine paternelle	23
3.7. Les séquences mitochondriales nucléaires	24
4. Principaux rôles de la mitochondrie.....	26
4.1. Fonctions métaboliques de la mitochondrie.....	26
4.2. Autres fonctions mitochondriales	35
II-Cytopathies mitochondriales	36
1. Epidémiologie	36
2. Expressions cliniques.....	36
3. Explorations paracliniques	38
3.1. Bilan métabolique	38
3.2. Imagerie du système nerveux central	40
3.3. Etudes histopathologiques	43
3.4. Investigations enzymologiques	45
3.5. Détection des complexes I à V par électrophorèse Blue Native (BN-PAGE)	46
3.6. Complémentation fonctionnelle	47

4. Les déficits génétiques dans les cytopathies mitochondriales :.....	47
4.1. Le déficit des gènes nucléaires.....	48
4.2. Déficit des gènes mitochondriaux.....	53
4.2.1. Généralités :.....	53
4.2.2. L'Hérédité mitochondriale.....	54
4.2.3. La notion d'hétéro/homoplasmie.....	54
4.2.4. Physiopathologie et théorie du goulot d'étranglement mitochondrial.....	55
4.2.5. Les syndromes associés à des délétions du génome mitochondrial.....	55
4.2.5.1. Les syndrome Kearns-Sayre (KSS).....	56
4.2.5.2. Le syndrome de Pearson.....	56
4.2.5.3. Ophtalmoplégie externe progressive (PEO : progressive external ophtalmoplegia).....	57
4.2.6. Les principaux syndromes liés à des mutations ponctuelles du génome mitochondrial....	57
4.2.6.1. Le syndrome de MELAS.....	57
4.2.6.2. Le syndrome de MERRF.....	58
4.2.6.3. Les syndromes de Leigh et NARP.....	58
4.2.6.4. La neuropathie optique héréditaire de Leber (LHON).....	59
4.2.6.5. Mutations de l'ADNmt et surdité.....	60
4.2.6.6. Mitochondries et diabète.....	61
5. Stratégie diagnostiques des cytopathies mitochondriales.....	63
6. Prise en charge et traitement des maladies mitochondriales.....	62
7. Techniques d'études moléculaires du génome mitochondrial.....	64
III- Les techniques de séquençage de nouvelle génération.....	65
1. Les étapes de la techniques de NGS.....	66
1.1. La préparation de la librairie :.....	66
1.1.1. Les techniques d'enrichissement.....	66
1.1.2. Le marquage des fragments par ligation des adaptateurs.....	67
1.1.3. Les techniques d'amplification clonale de la librairie :.....	68
1.2. Les techniques de séquençage.....	70
2. Le traitement des données.....	74
2.1. Le « base calling ».....	75
2.2. Les scores qualités et fichier FASTQ.....	76
2.3. L'alignement des « reads » sur le génome de référence :.....	77
2.4. L'identification des variants ou « variant calling ».....	78
3. Autres applications du NGS :.....	78
Deuxième section : matériels et méthodes.....	83
1. Première partie : la mise au point de la technique.....	81

1.2. Sélection des échantillons à tester.....	81
1.3. Extraction d'ADN	82
1.4. Les étapes de la technique de préparation de la librairie et séquençage.....	82
1.5. Analyses des données : pipeline bio-informatique	85
2. Deuxième partie : recherche de nouvelles mutations :	87
2.1. Sélection des patients à étudier :	87
2.2. Stratégie d'interprétation des variants retrouvés :	88
Troisième section : Résultats et discussion.....	96
1. Première partie : mise au point de la technique.....	92
1.1. Evaluation des coûts:.....	92
1.2. Rendement de la micro-puce de séquençage	93
1.3. Longueurs moyennes des « reads » et optimisation de la digestion	94
1.4. L'alignement	96
1.5. La couverture.....	100
1.6. La détection des réarrangements et leurs points de cassure	106
1.7. Détection des variants et hétéroplasmie	114
1.7.1. Efficacité de détection des variations connues	114
1.7.2. Les effets de l'alignement sur la détection des variants	118
2. Deuxième partie : recherche de nouvelles mutations	120
2.1. Caractéristiques cliniques des patients de la série	120
2.2. Résultats des PCR longues.....	124
2.3. Résultats du séquençage de nouvelle génération	126
2.3.1. Les variants générés et fichiers de sortie.....	126
2.3.2. Tri des variants	128
2.3.3. Outil de tri « maison ».....	131
2.3.4. Les mutations et nouveaux variants détectés dans notre série.....	137
Quatrième section : Conclusions et perspectives.....	141
Annexes	144
Références bibliographiques	167
Summary.....	176
Résumé.....	177

Abréviations :

ADNmt : acide désoxyribonucléique mitochondrial

ARNt : acide ribonucléique de transfert

ARNr : acide ribonucléique ribosomal

OXPHOS : phosphorylation oxydative

ROS : dérivés réactifs à l'oxygène (Reactive oxygen species)

RRF : fibre rouge déchiquetée (Ragged Red Fiber)

COX : cytochrome c oxydase

KSS : syndrome Kearns-Sayre

MELAS: Mitochondrial Encephalomyopathy, Lactic Acidosis, and Stroke-like episodes

MERRF: Myoclonic Epilepsy with Ragged Red Fibers

NARP: neuropathie, ataxie, rétinite pigmentaire

PEO: ophtalmoplégie externe progressive (Progressive External Ophtalmoplegia)

PCR : Polymerase chain Reaction

RFLP : polymorphisme de longueur des fragments de restriction (Restriction Fragment Length Polymorphism)

NGS : séquençage de nouvelle génération (Next Generation Sequencing)

Avant-Propos

Les cytopathies mitochondriales sont un groupe de maladies génétiques dont l'origine commune est un déficit de la chaîne respiratoire mitochondriale. Cette dernière a pour rôle essentiel de fournir l'énergie nécessaire au fonctionnement cellulaire normal par synthèse d'ATP, grâce aux réactions de phosphorylation oxydative. La chaîne respiratoire est constituée de 5 complexes multi-protéiques dont l'origine génétique est double : nucléaire et mitochondriale. En effet, la mitochondrie est un organe semi-autonome qui possède son propre génome sous la forme d'une molécule d'ADN circulaire de 16569 paires de bases. Ce génome code pour 13 des protéines de la chaîne respiratoire, 22 ARN de transfert et deux ARN ribosomiaux, dont le déficit peut être responsable de cytopathies mitochondriales.

Les déficits de la chaîne respiratoire ont une expression clinique hétérogène, le plus souvent multi systémiques, s'associant à des altérations métaboliques telle que l'hyper-lactacidémie. De plus, l'activité enzymatique des différents complexes de la chaîne respiratoire est possible sur des tissus ou des cellules en culture, pouvant révéler un déficit isolé ou combiné. Ces anomalies métaboliques sont cependant loin d'être constantes et des bilans métaboliques normaux n'éliminent pas une cytopathie mitochondriale. Tout ceci rend compte de la complexité diagnostique de ce groupe de maladies et de l'importance des analyses moléculaires pour poser formellement le diagnostic.

Depuis la mise en évidence des premières associations syndromiques liées à des anomalies de l'ADNmt (Syndrome de Kearns et Sayre, Syndrome MELAS..) dans les années 1980, et jusqu'à récemment l'étude de l'ADNmt consistait en la détection des grands réarrangements (délétions simples, multiples..) et des principales mutations ponctuelles décrites. En effet, le séquençage complet de l'ADNmt par méthode Sanger classique n'est pas réalisable de façon systématique au regard de la lourdeur et du coût de la technique.

L'objectif de ce travail était de mettre en place la technique de séquençage de nouvelle génération (Next Generation Sequencing ou NGS) pour l'étude du génome mitochondrial au CHU de Bordeaux. Cette nouvelle technique permet une étude rapide de la totalité du génome mitochondrial pour chaque patient suspect de cytopathie mitochondriale. Nous avons évalué dans un second temps, le rendement diagnostique de cette nouvelle technique dans une série de patients ayant une présentation clinique très évocatrice de cytopathie mitochondriale et pour lesquels les techniques conventionnelles n'avaient pas retrouvé d'anomalies de l'ADN mitochondrial.

Ce mémoire de thèse est divisé en quatre parties :

Dans l'introduction générale, nous avons cherché à replacer cette étude dans le contexte des connaissances actuelles, concernant la structure et le fonctionnement des mitochondries, les aspects cliniques, diagnostiques, la prise en charge thérapeutique et le conseil génétique des cytopathies mitochondriales et enfin les méthodes de séquençage de nouvelle génération.

Dans la section matériel et méthodes, nous avons exposé dans une première partie la technique de mise au point du NGS et le traitement bioinformatique des données effectué. Dans une deuxième partie, nous avons présenté les critères d'inclusion des patients de notre série pour la recherche de nouvelles mutations et la stratégie d'interprétation et de validation des variants trouvés.

Dans la section résultat et discussion, nous avons présenté et discuté les résultats concernant la mise au point du séquençage sur le plateau technique de biologie moléculaire et la détection de nouvelles mutations dans la série de patients étudiés.

Dans la section conclusions et perspectives, nous avons cherché à tirer les enseignements de nos investigations en particulier dans la perspective du transfert en diagnostic de routine de cette technique.

Première section : Introduction

I. Les mitochondries :

1. Origine et historique

Selon la théorie endosymbiotique énoncée par Margulis (1981), l'endosymbiose d'une α -protéobactérie dans les cellules eucaryotes primitives serait à l'origine des mitochondries. L'hypothèse de Margulis, qui a aujourd'hui rallié la majorité des biologistes, a recueilli de nombreuses évidences en sa faveur:

- Les analyses phylogénétiques ont clairement démontré que plastides et mitochondries dérivait de lignées bactériennes apparentées respectivement aux cyanobactéries et aux protéobactéries actuelles.
- La présence d'une double membrane serait un vestige de la phagocytose initiale : la membrane interne maintenant repliée sur elle-même serait l'ancienne membrane bactérienne, tandis que la membrane externe proviendrait de la cellule hôte elle-même.
- L'existence d'un génome propre sous forme d'ADN circulaire. Toutefois, les protéines codées par cet ADN ne couvrent pas l'ensemble des protéines mitochondriales. On pense que le procaryote aurait perdu certains gènes au profit du noyau de la cellule, un processus connu sous le nom de « transfert de gènes endosymbiotique ». Pour cette raison, les mitochondries sont aujourd'hui dépendantes de leur hôte pour la synthèse de la plupart de leurs composants.
- L'existence d'un mécanisme de chimiosmose comparable à celui des bactéries pour la production d'énergie sous forme d'ATP

Deux points néanmoins restent non expliqués par cette théorie : le code génétique du génome mitochondrial diffère du code universel des procaryotes et l'existence d'introns dans le génome mitochondrial de la levure alors que les génomes bactériens n'en possèdent pas.

Les mitochondries ont été découvertes il y a près de 150 ans lorsque Kölliker, en 1857, décrivit pour la première fois leur morphologie dans le tissu musculaire. Elles furent appelées mitochondries du grec mitos (filament) et chondros (grain) en référence à leur forme. La présence de ces organites dans toutes les cellules fut démontrée en 1890 par Altmann, qui mit au point une technique de coloration des mitochondries et suggéra leur autonomie métabolique et génétique. En 1937, Krebs construisit un

modèle métabolique siégeant dans la mitochondrie chez les eucaryotes, il l'appela cycle de l'acide citrique. L'isolement de mitochondries fut réalisé pour la première fois en 1940 par Claude qui effectua un fractionnement cellulaire à partir de cellules hépatiques. Cette technique fut par la suite améliorée par Hogeboom et Schneider qui utilisèrent du sucrose isotonique ayant la propriété de préserver l'intégrité et la morphologie mitochondriale. Ceci permit d'observer la structure mitochondriale au microscope électronique et ainsi de les décrire comme un organe délimité par deux membranes, la membrane interne présentant des invaginations appelées crêtes. En 1948, Kennedy et Lehninger démontrèrent que le cycle de l'acide citrique ou cycle de Krebs, la β -oxydation des acides gras et la phosphorylation oxydative (OXPHOS) avaient lieu dans la mitochondrie. En 1978, Peter Mitchell obtint le prix Nobel pour sa théorie chimio-osmotique postulant que le couplage entre la chaîne d'oxydoréduction et le mécanisme de phosphorylation repose sur la formation d'un gradient de protons à travers la membrane interne, servant de réservoir d'énergie pour la synthèse d'ATP. L'ADN mitochondrial fut découvert en 1963 (Nass et Nass.1963) et la structure génétique de l'ADN mitochondrial humain fut découverte en 1981 par Anderson (ANDERSON *et al.* 1981). La première mutation pathogène fut identifiée par Wallace en 1988 (WALLACE *et al.* 1988). Et enfin, Boyer et Walker obtinrent le prix Nobel en 1997 pour leurs travaux sur la structure et le fonctionnement de l'ATP synthase.

2. Structure et organisation cellulaire des mitochondries :

2.1. Structure et morphologie mitochondriales

Les mitochondries sont des organites retrouvés dans la plupart des cellules eucaryotes à l'exception des érythrocytes. Elles mesurent entre 0.5 et 1 μm de diamètre et peuvent atteindre plus de 10 μm de long. Les mitochondries sont constituées d'une membrane externe et d'une membrane interne, délimitant deux compartiments, l'espace inter-membranaire et la matrice mitochondriale (MILANE *et al.* 2015) (figure 1).

La membrane externe constitue une barrière semi-perméable aux ions et aux petites molécules, ce qui fait que l'espace inter-membranaire a une composition proche de celle du cytoplasme. Cette perméabilité est assurée en grande partie par la présence d'une protéine formant un canal transmembranaire, le VDAC (voltage-dependent anion channel). Cette protéine permet le passage des molécules d'une taille inférieure à 10 kDa (anions, cations, acides gras, pyruvate et nucléotides). Les molécules supérieures à 10 kDa, lorsqu'elles possèdent un signal d'adressage mitochondrial, sont prises en charge par les complexes d'importation TOM/TIM (translocase of the outer membrane / translocase of the inner membrane) localisés au niveau d'une zone d'accolement entre les membranes externe et interne. La membrane externe contient par ailleurs des complexes d'importation spécifiques du cholestérol, pour la synthèse des stéroïdes.

La membrane interne a une composition qui diffère des autres membranes biologiques avec environ 80% de protéines pour seulement 20% de phospholipides. Elle est caractérisée par son imperméabilité liée notamment à sa richesse en un lipide particulier la cardiolipine. Cette barrière sélective nécessite donc la présence de transporteurs facilitateurs tels que le transporteur des nucléotides adényliques (ANT) qui réalise l'échange ATP/ADP entre le cytosol et la matrice mitochondriale, le transporteur du phosphate inorganique (PiC), le transporteur du pyruvate, la navette glycérol 3-phosphate qui transfère les électrons du NADH à l'ubiquinone (CoQ) située dans la membrane interne et la navette malate/aspartate formée par les perméases malate/alpha-cétoglutarate et aspartate/glutamate. Cette dernière transfère le malate en direction de la matrice et régénère les NADH cytosoliques en NADH mitochondriaux. La membrane interne s'invagine et forme des crêtes mitochondriales qui permettent d'augmenter la surface membranaire d'échange. Par ailleurs, la cardiolipine joue un rôle dans la fixation des protons à leurs groupements phosphates et contribuent ainsi pour beaucoup au maintien du gradient électrochimique. Les cardiolipines interagissent également avec divers transporteurs dont l'ANT (antiport d'échange [ADP/ATP]) et les complexes III, IV et l'ATP synthase de la chaîne respiratoire. Ils contribuent ainsi à l'optimisation de leur structure et à leur orientation pour le transfert des électrons au sein de la chaîne respiratoire.

Le compartiment inter-membranaire, du fait des perméabilités différentes entre les membranes mitochondriales internes et externes constitue un lieu de transit pour les molécules de taille inférieure à 10 kDa. Il contient principalement des ions H^+ et des protéines telles que le cytochrome C, les procaspases impliquées dans l'apoptose et la créatine Kinase.

La matrice mitochondriale contient quant à elle les molécules d'ADN mitochondrial (2 à 10 copies par mitochondrie), des ARN ribosomiaux (les mitoribosomes) et des ARN de transfert, ainsi que les éléments nécessaires à la réplication et à l'expression des gènes mitochondriaux. C'est également le siège de nombreuses voies métaboliques telles que le cycle de Krebs et la β -oxydation des acides gras.

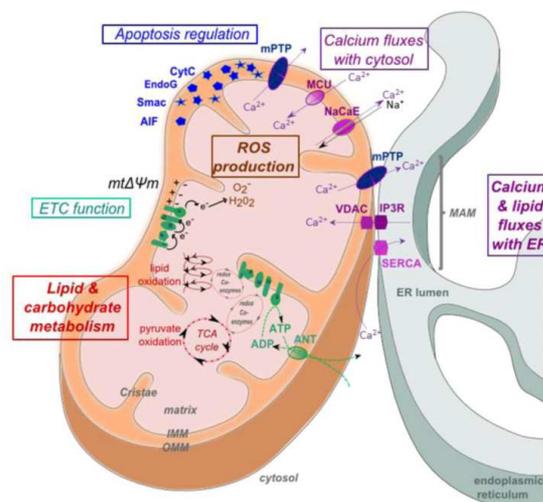


Figure 1 : structure de la mitochondrie

2.2. Le réseau mitochondrial :

2.2.1. Généralités

Durant de nombreuses années, la mitochondrie a été considérée comme un organe statique, une structure bien individualisée, ovoïde, telle que caractérisée sur les coupes 2D des premières observations en microscopie électronique (figure 2). Cependant, il a été récemment découvert notamment grâce à la microscopie confocale (BAKEEVA *et al.* 1978) que les mitochondries constituaient un ensemble de structures cytosoliques interconnectées qui fusionnent et fissionnent constamment formant ainsi le réseau mitochondrial (figure 3). Ce réseau est animé de nombreux mouvements dans les cellules vivantes qui lui confèrent une organisation réticulée ou fragmentée formée de tubules plus ou moins longs, selon le statut énergétique de la cellule. Ainsi par exemple, les fibroblastes possèdent habituellement des mitochondries de type longs filaments de 1–10 μm de long et un diamètre constant de 700 nm alors que les mitochondries hépatocytaires sont plus sphériques et ovoïdes. (DAS *et al.* 2012)



Figure 2 : Mitochondries en microscopie électronique à transmission

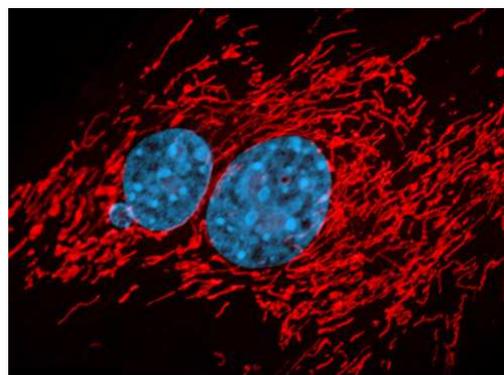


Figure 3 : le réseau mitochondrial dans les fibroblastes. Immunofluorescence au MitoTracker Red

Cet assemblage tubulaire de mitochondries s'étend dans le cytosol et paraît être en contact avec le noyau, le réticulum endoplasmique, l'appareil de Golgi et le cytosquelette (AMCHENKOVA *et al.* 1988) (SZABADKAI *et al.* 2003). De nombreuses études ont également montré que le réseau mitochondrial se déplaçait dans le cytosol le long des microtubules grâce à une GTPase : la protéine MIRO présente à la surface de la membrane externe mitochondriale et son effecteur MILTON qui interagit avec la chaîne lourde de la kinesin-1 (Kinesin heavy chain [KHC] ou KIF5B) qui permet le déplacement le long des microtubules en présence de GTP (figure 4) (GLATER *et al.* 2006)

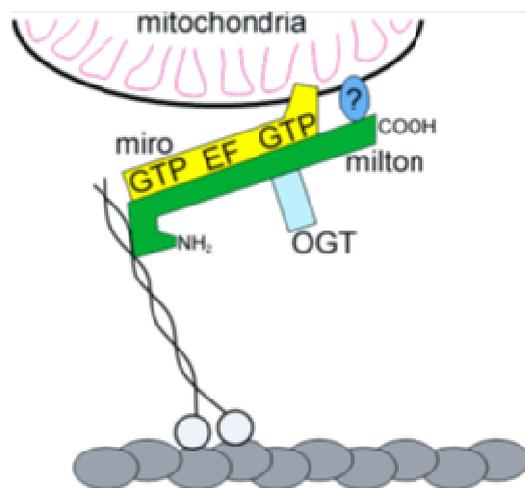


Figure 4 : Représentation schématique du complexe protéique assurant le transport axonal mitochondrial (GLATER *et al.* 2006)

Il semble que ce déplacement ne soit pas aléatoire mais répondrait aux besoins locaux des cellules. Ainsi par exemple, dans les cellules des acini pancréatiques, il existe trois groupes individualisés de réseaux mitochondriaux : un localisé dans la région basolatérale, sous la membrane plasmique, un deuxième réseau autour du noyau et enfin un troisième semblant se positionner préférentiellement autour des granules sécrétoires. Chacun de ces trois réseaux, semble répondre de manière indépendante à des stimuli cytotogiques distincts. (PARK *et al.* 2001)

2.2.2. Processus de fission et fusion mitochondriales

Les mécanismes de fission et de fusion mitochondriales font intervenir des protéines clés appartenant à la superfamille des GTPases. La fission des mitochondries passe par la protéine « dynamin related protein » Drp1, une protéine cytosolique qui permet d’initier la fission mitochondriale quand celle-ci est recrutée à la surface de la membrane externe des mitochondries. La fusion mitochondriale est médiée par les mitofusines MFN1 et MFN2 pour la membrane externe et par la protéine OPA1 pour la membrane interne (figure 5).

Les fission et fusion mitochondriales semblent répondre à de multiples évènements cellulaires tels que la division cellulaire, la modification du métabolisme, la différenciation cellulaire ou encore l’apoptose. La plupart des signaux connus permettant de réguler ce processus agissent par modification post-traductionnelle des protéines impliquées (phosphorylation, ubiquitination ou sumoylation...).(ZORZANO and CLARET 2015)

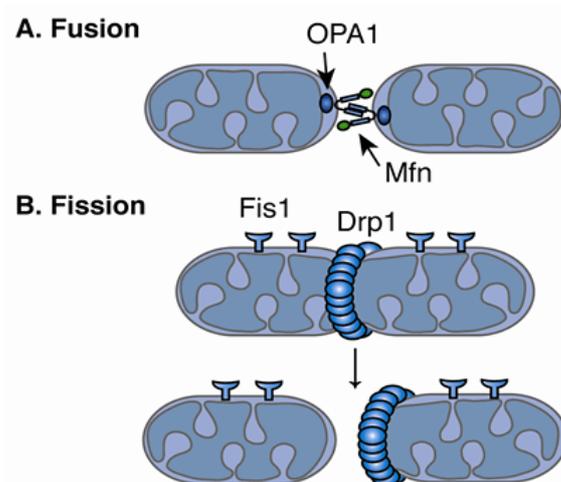


Figure 5 : schématisation des fission et fusion mitochondriales et les protéines impliquées (CHEN and CHAN 2006)

2.2.2.1. La fission mitochondriale

Le rôle de DRP1 dans la fission mitochondriale a été mis en évidence grâce à l'inhibition de son expression par siRNA ou par expression d'un mutant dominant négatif montrant une augmentation de la longueur des tubules mitochondriaux et une réticulation du réseau ((SMIRNOVA *et al.* 2001) (LEE *et al.* 2004). Cette protéine appartient à la superfamille des dynamines GTPases, possédant un domaine GTPase en N-terminal. De ce fait, il a été proposé que cette dernière puisse agir comme une mécanoenzyme utilisant l'hydrolyse du GTP pour permettre la constriction puis la scission membranaire mitochondriale au niveau de sites de fission, et ce par analogie avec les dynamines classiques qui interviennent dans la constriction et l'invagination des vésicules d'endocytose.

Des observations ont montré que DRP1 est une protéine cytoplasmique qui est recrutée pendant le phénomène de fission au niveau des futurs sites de fission de la membrane mitochondriale. Elle s'assemble ensuite formant des oligomères complexes et donne un aspect de foci. On ne sait cependant toujours pas le mode de recrutement de la protéine à la membrane, surtout en l'absence de domaine d'adressage mitochondrial identifié. Ce recrutement doit probablement se faire via d'autres protéines membranaires adaptatrices telle que Fis (LEE *et al.* 2004).

2.2.2.2. *La fusion mitochondriale*

La fusion mitochondriale nécessite la fusion concomitante des deux membranes interne et externe. Au niveau de la membrane externe, celle-ci est médiée par deux protéines homologues MFN1 et MFN2 qui semblent avoir des fonctions redondantes, mais un pattern d'expression différent. La fusion de membrane interne semble quant à elle dépendre principalement de OPA1 (CIPOLAT *et al.* 2004) (GRIPARIC *et al.* 2004), une protéine localisée dans l'espace inter-membranaire mitochondrial qui serait retrouvée en quantité plus importante au niveau des crêtes mitochondriales (OLICHON *et al.* 2002). Elle est ubiquitaire, mais prédominante au niveau du cerveau, des testicules, des muscles squelettiques et de la rétine (DELETTRE *et al.* 2001). L'étude d'OPA1 par Western Blot a permis d'identifier plusieurs isoformes: deux isoformes longues ancrées à la membrane interne et 3 isoformes courtes solubles. Toutes semblent être indispensables à la fusion mitochondriale même si leur mécanisme d'action reste obscur.

La fusion des deux types de membranes, bien que mettant en jeu des protéines différentes, paraît étroitement coordonnée, le mécanisme sous-jacent à ce processus étant encore mal compris. Il semblerait néanmoins, que ce sont les protéines MFN1 et MFN2, localisées au niveau de la membrane externe qui vont initier en premier lieu l'interaction entre mitochondries adjacentes, puis coordonner la fusion de leurs membranes internes via une courte boucle protéique localisée dans l'espace inter-membranaire permettant une interaction avec OPA1.

2.2.2.3. *Implication des protéines de fusion et fission mitochondriales en pathologies humaines :*

DRP1 a été impliquée dans un cas d'encéphalopathie néonatale sévère lié à une mutation dans le gène *DRP1* avec vraisemblablement un effet dominant négatif qui a été rapporté par (WATERHAM *et al.* 2007). En effet, l'examen des fibroblastes de la patiente en immunofluorescence avait révélé un marquage fusiforme allongé des mitochondries mais également des peroxyosomes comme ce qui est observé dans les modèles cellulaires de dominant négatif ou de KO^{-/-} DRP1. Il est à noter également que DRP1 a été impliquée dans les mécanismes de dégénérescence neuronale observée dans la maladie de Huntington. Des études récentes ont montré que la protéine huntingtin mutée (polyQ) aurait une action hyper-activatrice de DRP1, directement en stimulant son domaine GTPasique et indirectement en inhibant la phosphorylation de la DRP1 active. Ceci conduirait à une fragmentation excessive du réseau mitochondrial et à une inhibition du transport axonal des mitochondries ce qui participerait au processus de mort neuronale. (SONG *et al.* 2011)

Les mutations dans les gènes de la fusion mitochondriale *MFN2* et *OPA1* sont responsables de neuropathies : optique pour *OPA1* (ALEXANDER *et al.* 2000) (DELETTRE *et al.* 2000) et sensitivomotrice

des fibres longues pour *MFN2* (Charcot Marie Tooth 2A) (*ZUCHNER et al. 2004*). Quelques cas d'atrophie optique ont également été rapportés en lien avec un déficit de *MFN2*. Les mécanismes exacts liant le défaut de fusion et l'atteinte neuropathique ne sont pas encore élucidés. Néanmoins, il semblerait que la perturbation de la dynamique mitochondriale altérerait la survie neuronale à cause de l'atteinte du transport axonal et du déficit énergétique. Les mutations *OPA1* sont à l'origine d'une affection dégénérative spécifique des cellules ganglionnaires de la rétine, où l'équilibre entre le processus de fusion et fission semble être critique. En effet, des observations ont montré que la portion intra-orbitaire des axones des cellules ganglionnaires était non myélinisée et riche en mitochondrie sphérique, à cause d'une demande en énergie accrue (nécessaire à la transmission de l'influx nerveux). Après le passage de la lame criblée (couche fibreuse au pôle postérieur de l'orbite), les fibres acquièrent une gaine de myéline, permettant une conduction nerveuse rapide et des besoins énergétiques moindres, dans cette portion, les mitochondries sont moins nombreuses et plutôt allongées. (*CARELLI et al. 2004*)

Par ailleurs, il existe un lien étroit entre la pathologie neurodégénérative et le contrôle de qualité des mitochondries. Par exemple, la protéine Parkin (MARCH-V ou MITOL) mutée dans les formes familiales et précoces de la maladie de Parkinson, est une des protéines régulatrices des mécanismes de fusion/fission mitochondriales. Elle régule notamment l'ubiquitylation et la dégradation de MFN1 et MFN2 par le protéasome (*GEGG et al. 2010*) (*PARK et al. 2010*). Plus généralement Parkin semble avoir un rôle important dans le contrôle de qualité des mitochondries et régule les phénomènes de mitophagie (*YOSHII et al. 2011*) (*DURCAN and FON 2015*)

2.3. Le « turnover » mitochondrial

Le contenu mitochondrial cellulaire peut être ajusté pour adapter la capacité de production énergie à la demande physiologique. La quantité de mitochondries à l'état stationnaire est le résultat de la balance entre deux processus : la biogénèse et la dégradation. Dans la cellule hépatique par exemple les mitochondries sont renouvelées avec une demi-vie d'environ 10 jours.

2.3.1. La biogénèse

La biogénèse mitochondriale passe par la duplication de l'ADN mitochondrial et la mitochondriogénèse. Elle fait intervenir l'expression de gènes appartenant aux génomes nucléaire et mitochondrial (*HOOD et al. 2006*). L'activation de la biogénèse fait intervenir différents facteurs de transcription. Les facteurs NRF (*Nuclear Respiratory Factor*) 1 et 2 sont les principaux facteurs de transcriptions nucléaires qui induisent l'activation de la transcription des gènes codant pour des protéines mitochondriales. NRF1 et NRF2 activent la transcription de mtTFA (*Mitochondrial*

Transcription Factor A) qui se fixe à l'ADNmt pour induire la transcription des gènes mitochondriaux (HOOD *et al.* 2006) (LAGOUGE *et al.* 2006) (COLLU-MARCHESE *et al.* 2015).

2.3.2. La mitophagie

Pour maintenir l'efficacité de cette machinerie complexe, un système de contrôle-qualité existe pour dégrader les protéines des complexes de la chaîne respiratoire qui seraient mal-assemblées ou endommagées. Des fragments entiers de mitochondries peuvent également être éliminés par une forme sélective d'autophagie, la mitophagie. Les mitochondries sont séquestrées dans une membrane autophagosomale qui fusionne avec les lysosomes pour dégrader son contenu. Ce processus est régulé par différentes voies de signalisation cellulaires (la voie des protéines mitochondriales BH3, le complexe Beclin-1/Bcl-2-Bcl-xL, les voies mTOR et AMPK) (BEAU *et al.* 2011). Ces différentes voies peuvent activer ou inhiber la mitophagie indépendamment ou de concert en réponse à divers stimuli spécifiques. La mitophagie peut être activée pour débarrasser la cellule des mitochondries endommagées, par exemple, lorsque la production énergétique est défectueuse (RODRIGUEZ-HERNANDEZ *et al.* 2009) (PRIAULT *et al.* 2005), lorsque la production de ROS est élevée (SCHERZ-SHOVAL and ELAZAR 2007), ou encore, pour adapter la quantité de mitochondrie à la disponibilité ou à la consommation d'oxygène. (SCHERZ-SHOVAL *et al.* 2007)

3. Le génome mitochondrial :

3.1. Structure :

Le génome mitochondrial est une molécule d'ADN circulaire double brin de 16569 paires de bases, localisée dans la matrice mitochondriale. Il a la particularité par rapport au génome nucléaire d'être de transmission uniquement maternelle (cf Chapitre Devenir des mitochondries d'origine paternelle)

Chaque mitochondrie comporte plusieurs molécules d'ADNmt (2 à 10 copies). Les molécules d'ADN mitochondrial ne sont pas libres dans la matrice mais sont rattachées à la membrane mitochondriale interne et forment des complexes appelés « nucléoïdes » avec plusieurs protéines matricielles, notamment celles impliquées dans les machineries de la réplication de la réparation et de la transcription. Des protéines chaperonnes ont été également identifiées : HSP90-β et HSP70, elles semblent être importantes pour la stabilité de l'ADNmt. (BOGENHAGEN 2012)

Chaque molécule comporte un brin dit lourd (ou H pour Heavy), car riche en résidus guanines ce qui lui confère un plus haut poids moléculaire, et un brin léger (ou L pour Light). Le brin léger est le brin sens de l'orientation génomique (de 1 à 16569) (ANDERSON *et al.* 1981). Chaque molécule possède 37 gènes codant pour 2 ARN ribosomiques (ARNr 12S et 16S), 22 ARN de transfert (ARNt) et 13 sous-unités protéiques : 7 appartiennent au complexe I (ND1-6, et ND4L), 1 au complexe III (cytb), 3 au

complexe IV (COXI-COXIII) et 2 au complexe V (ATPase 6 et 8). Le génome mitochondrial ne possède pas de séquences introniques et les régions inter-géniques sont quasi-inexistantes. En plus, certaines région géniques se chevauchent, c'est le cas entre les gènes ATP6 et ATP8 et les gènes ND4 et ND4L. A peu près 94% de l'ADNmt est codant, la seule région non codante régulatrice est une séquence de 1118 nucléotides (16028-577) appelée « D-loop » qui contient les promoteurs des brins lourd et léger ainsi que l'origine de réplication du brin lourd (figure 6).

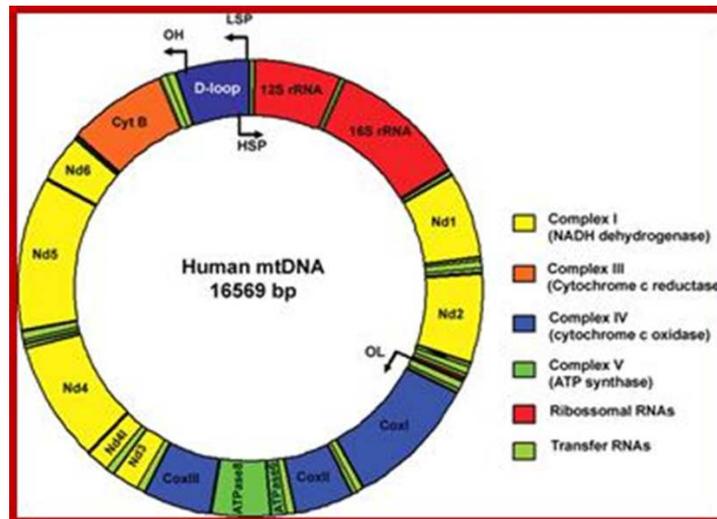


Figure 6 : le génome mitochondrial (*source mitomap*)

3.2. Etudes phylogénétiques

L'ADN mitochondrial possède un degré de polymorphisme important lié à un taux de mutation qui est jusqu'à 10 fois plus élevé que celui de l'ADN nucléaire (environ 1.7^{10} substitutions par site et pas an). Cette variabilité est surtout présente au niveau de trois régions de la D-Loop, les régions hypervariables HV1 (16024 à 16365), HV2 (73 à 340) et HV3 (438 à 576) (ANDERSON *et al.* 1999) (STONEKING 2000). Ce fort taux mutationnel a été attribué d'une part à une moins bonne efficacité de la polymérase gamma lors de la réplication par rapport aux polymérases nucléaires, à la proximité de l'ADN mitochondrial de la chaîne respiratoire et donc de la toxicité des ROS, et d'autre part à un système de réparation moins efficace. En effet, même si plusieurs mécanismes de réparation ont été décrits récemment dans la mitochondrie (BOESCH *et al.* 2011), ces derniers ne suffisent pas pour assurer une réparation aussi efficace que celle de l'ADN nucléaire. Parmi ces mécanismes, le système de réparation par excision de base est celui qui a été le mieux caractérisé à ce jour (BOESCH *et al.* 2010).

La variabilité de l'ADNmt associée à son origine monoparentale et donc à l'absence de recombinaison font de lui un très bon marqueur phylogénétique

A ce jour, 65 recombinaisons de mutations ont été identifiées et permettent de distinguer de nombreux haplogroupes mitochondriaux (PhyloTree build.16). Ceci a permis de remonter génétiquement, par matrilinearité, à une unique hypothétique ancêtre commune (Ève mitochondriale) qui aurait vécu il y a 140 000 ans sur le continent africain.

3.3. La réplication

Selon le modèle de Clayton (CLAYTON 1982) (CLAYTON 2003), la réplication débute avec la synthèse du brin lourd H au niveau de son origine de réplication OH, située dans de la D-loop et progresse dans le sens des aiguilles d'une montre ($5' \rightarrow 3'$). Lorsque la fourche de réplication a parcouru 2/3 de la distance génomique, elle atteint l'origine de réplication du brin léger OL. Celui-ci se trouve alors exposé et sera répliqué en sens inverse (figure 7). Ce modèle de réplication est dit « bidirectionnel et asynchrone ». Un autre modèle a également été proposé chez les mammifères à partir de multiples origines de réplication.

L'origine de la réplication du brin lourd (OH) est située à proximité du site d'initiation de la transcription du brin léger ITL (en position 561). L'ARN polymérase va initier au niveau de ce site, la transcription d'une courte séquence d'ARN complémentaire qui servira d'amorce à l'ADN polymérase gamma (FUSTE *et al.* 2010). L'ADN polymérase gamma va ensuite poursuivre l'élongation par la synthèse d'un brin lourd H fille, et une RNase-H va dégrader l'amorce d'ARN. Les mécanismes qui vont orienter le processus vers la transcription du brin léger ou la synthèse du brin lourd à partir de l'amorce d'ARN ne sont pas élucidés (figure 8), la transition ARN/ADN semble se produire de manière variable au niveau de trois blocs de séquences conservées (CSB), qui constitue vraisemblablement l'origine réelle de réplication (OH). La fourche de réplication comme au niveau nucléaire va être déroulée grâce à une DNA-Hélicase appelée Twinkle. Des protéines SSB (single strand binding) vont stabiliser le fragment simple brin libre et une topoisomérase va lutter contre la tension de torsion engendrée par le déroulement de la molécule double brin d'ADN (figure 9). L'origine de réplication du brin léger OL est activée grâce au déroulement de la fourche de réplication du brin lourd. Une fois exposée, cette séquence nucléotidique semble adopter une structure en forme de tige-boucle par auto-complémentarité permettant ainsi sa reconnaissance par une RNA primase.

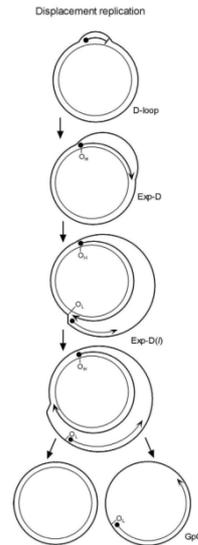
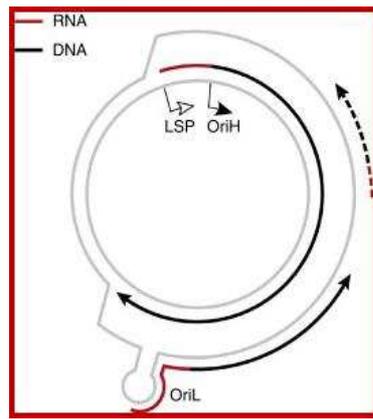


Figure 7: Modèle de réplication asymétrique de l'AND mitochondrial

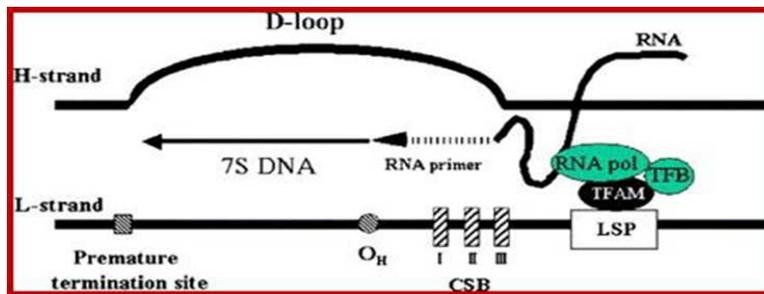


Figure 8 : l'origine e la réplication du brin lourd débute au niveau de la D-loop par une amorce d'ARN synthétisée par l'ARN polymérase mitochondriale ((FUSTE *et al.* 2010)

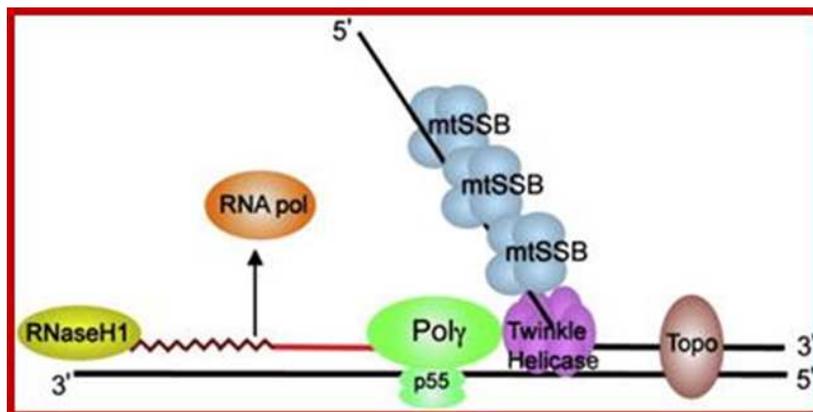


Figure 9 : représentation schématique de la fourche de réplication mitochondriale

3.4. La transcription

Les 2 brins d'ADNmt sont transcrits en ARN polycistroniques dont la maturation génère des ARN ribosomaux, des ARN de transfert et des ARN messagers (ARNm). Deux facteurs de transcription mitochondriale ont été identifiés à ce jour mTFA et mTFB. Les deux séquences promotrices pour chacun des brins léger L et lourd H sont situées en cis dans la D-loop. Il s'agit de la même séquence de 15pb, motif consensus de 5'CANACC(G)CC(A)AAAGAYA3', entourant les sites d'initiation de transcription (le A). Les deux sites d'initiation de transcription sont situés en position 561 pour le brin lourd (IT H) et en 407 pour le brin léger (ITL). En dépit de la proximité des deux promoteurs, les transcriptions des deux brins semblent être indépendantes tout du moins in vitro. Ceci suggère l'existence probable d'autres séquences régulatrices enhanceurs. Le brin lourd est le plus transcrit, il sert de brin matrice à la synthèse des deux ARN ribosomaux, de 14 ARN de transfert et à des ARN messagers codant pour 12 protéines. Le brin léger est moins transcrit, il contient les 9 gènes restant, servant à la synthèse de 8 ARN de transfert et à la protéine ND6 du complexe I (figure 10) (TAANMAN 1999).

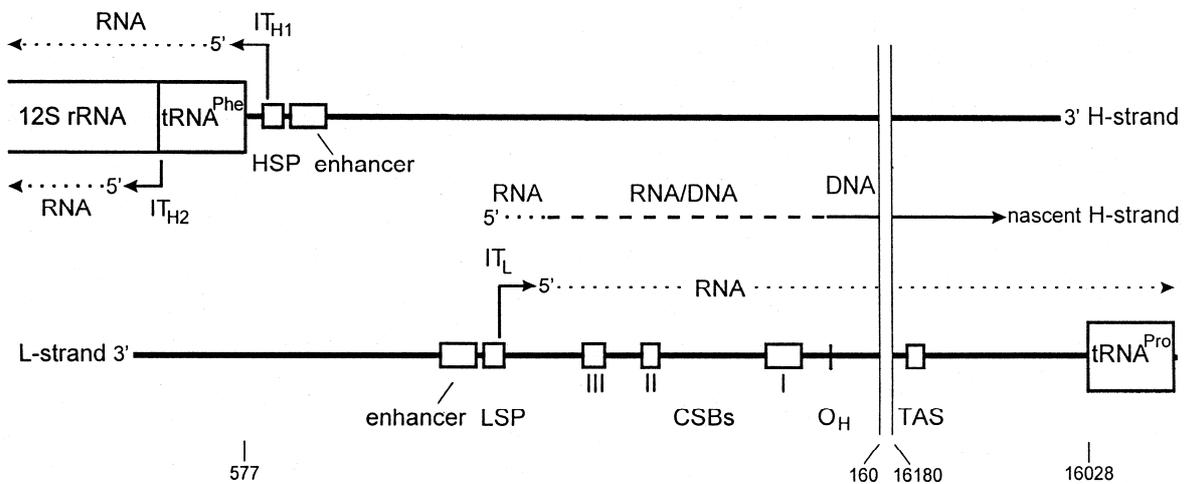


Figure 10 : initiation de la transcription des deux brins au niveau de la D-loop, les deux sites d'initiation ITH1 (et ITH2 accessoire) pour le brin lourd et IT L pour le brin léger (TAANMAN 1999).

3.5. La traduction et le code génétique mitochondrial

La mitochondrie possède sa propre machinerie de traduction avec 22 ARN de transfert et 2 ARN ribosomiaux, une grande sous-unité (40s) contenant un ARNr 16s et une petite sous-unité (30s) contenant un ARNr 12s. Il est à noter également que le code génétique mitochondrial diffère du code universel des eucaryotes (code de la traduction nucléaire (tableau 1))

Phe F	UUU UUC	Thr T	ACU ACC ACA ACG	Asp D	GAU GAC
Leu (1) L (UUA/G)	UUA UUG	Ala A	GCU GCC GCA GCG	Glu E	GAA GAG
Leu (2) L (CUN)	CUU CUC CUA CUG	Tyr Y	UAU UAC	Cys C	UGU UGC
Ile I	AUU AUC	Ter	UAA UAG	Trp W	UGA UGG
Met M	AUA AUG	His H	CAU CAC	Arg R	CGU CGC CGA CGG
Val V	GUU GUC GUA GUG	Gln Q	CAA CAG	Ser (2) S (AGU/C)	AGU AGC
Ser (1) S (UCN)	UCU UCC UCA UCG	Asn N	AAU AAC	See note below*	AGA AGG
Pro P	CCU CCC CCA CCG	Lys K	AAA AAG	Gly G	GGU GGC GGA GGG

Tableau 1 : Le code génétique mitochondrial

3.6. L'ADNmt au cours du développement embryonnaire : devenir des mitochondries d'origine paternelle

L'ADN mitochondrial est de transmission uniquement maternelle, ceci implique une destruction totale de l'ADNmt d'origine spermatique dans l'œuf fécondé. La reconnaissance et la destruction des mitochondries paternelles reposent sur le processus d'ubiquitinylation d'une protéine membranaire mitochondriale (la prohibitine), commencé au cours de la spermatogenèse. Lors de la décondensation du spermatozoïde dans le cytoplasme ovocytaire, les sites ubiquitinylés des membranes mitochondriales paternelles (sites initialement masqués par des ponts disulfures durant le transit épидidymaire) seraient alors démasqués et pris pour cible par les enzymes protéolytiques du

cytoplasme ovocytaire. Cette destruction se produirait au plus tard lors de la troisième division de segmentation de l'embryon. Ce mécanisme semble "spécifique" au sens propre du terme, car dans des embryons "chimères", issus de croisements murins interspécifiques (entre espèces de souris différentes), on observe une persistance des mitochondries paternelles. Ce mécanisme est également spécifique des mitochondries de la lignée germinale mâle, puisque des mitochondries d'origine tissulaire différente (de foie par exemple) injectées dans un ovocyte de souris ne sont pas détruites. (SATO and SATO 2013).

Ainsi, après la fécondation, seules les mitochondries apportées par l'ovocyte sont conservées et serviront de "réserve" initiale des mitochondries du nouvel individu. Il est à noter que l'inhibition de la transcription et de la traduction mitochondriale ne semble pas perturber pas le développement du zygote jusqu'au stade blastocyste (chez la souris), on peut supposer donc que le contingent mitochondrial maternel suffit à assurer les besoins énergétiques de l'embryon jusqu'à ce stade. Cependant, il ne semble pas y avoir de réplication de l'ADNmt avant l'implantation de l'embryon. La plupart des études menées chez la souris montrent un taux d'ADNmt constant jusqu'à la gastrulation. Comme le montre la baisse de la capacité de développement des embryons dans le cas d'une réduction expérimentale du taux d'ATP ovocytaire (chez la souris), le contenu en ATP constitue un bon indicateur de la vitalité de l'embryon. Il est possible en effet, de "sauver" (chez l'animal) des ovocytes de mauvaises qualités par transfert de cytoplasme provenant d'un ovocyte normal, ou de façon plus spécifique, par transfert de mitochondries purifiées, capable à lui seul d'augmenter la production d'ATP et de prévenir l'apoptose de l'ovocyte receveur (FRAGOULI *et al.* 2015).

3.7. Les séquences mitochondriales nucléaires

Dernièrement, certaines équipes se sont rendues compte de l'existence de séquences similaires à des régions de l'ADNmt, tout le long des différents chromosomes du génome nucléaire (tableau 2) et (figure 11). La première observation de ces séquences appelées « séquences mitochondriales nucléaires » ou « pseudogènes mitochondriaux » remonte à l'année 1967, quand des travaux ont montré l'hybridation possible entre les molécules de l'ADNmt et le génome nucléaire.

A ce jour, on ne connaît ni les mécanismes impliqués dans l'intégration de séquences mitochondriales dans le génome nucléaire ni la période de l'évolution à partir de laquelle ces insertions ont commencé à se produire. Plus récemment, une hypothèse a été formulée, évoquant la possibilité d'une fragmentation de l'ADNmt sous l'effet du stress cellulaire ou d'agents mutagènes, ces fragments s'inséreraient dans les chromosomes pendant la réparation des cassures doubles brins grâce à la machinerie de réparation « non-homologous end joining ou NHEJ » (SIMONE *et al.* 2011).

Table 2 NumtS number and percentage for each chromosome

Chr name	NumtS per Chr	Chr length (bp)	Total NumtS span (bp)	NumtS bp % per Chr
1	46	247249719	57670	0.023
2	92	242951149	108317	0.045
3	39	199501827	24203	0.012
4	38	191273063	36179	0.019
5	24	180857866	40163	0.022
6	31	170899992	11708	0.007
7	35	158821424	51356	0.032
8	37	146274826	46805	0.032
9	28	140273252	36854	0.026
10	33	135374737	21445	0.016
11	28	134452384	26640	0.020
12	40	132349534	8551	0.006
13	19	114142980	11159	0.010
14	8	106368585	8991	0.008
15	12	100338915	9002	0.009
16	11	88827254	17205	0.019
17	17	78774742	24285	0.031
18	7	76117153	1340	0.002
19	15	63811651	19373	0.030
20	10	62435964	5905	0.009
21	8	46944323	7210	0.015
22	8	49691432	9025	0.018
X	13	154913754	30845	0.020
Y	12	57772954	11510	0.020

Chromosome 2 displays the highest bp percentage of NumtS, while chromosome 18 the lowest. Chr. chromosome.

Tableau 2 : représentation du pourcentage de pseudogènes mitochondriaux par chromosome. Le chromosome 2 est le plus riche en séquences mitochondriales (SIMONE *et al.* 2011).

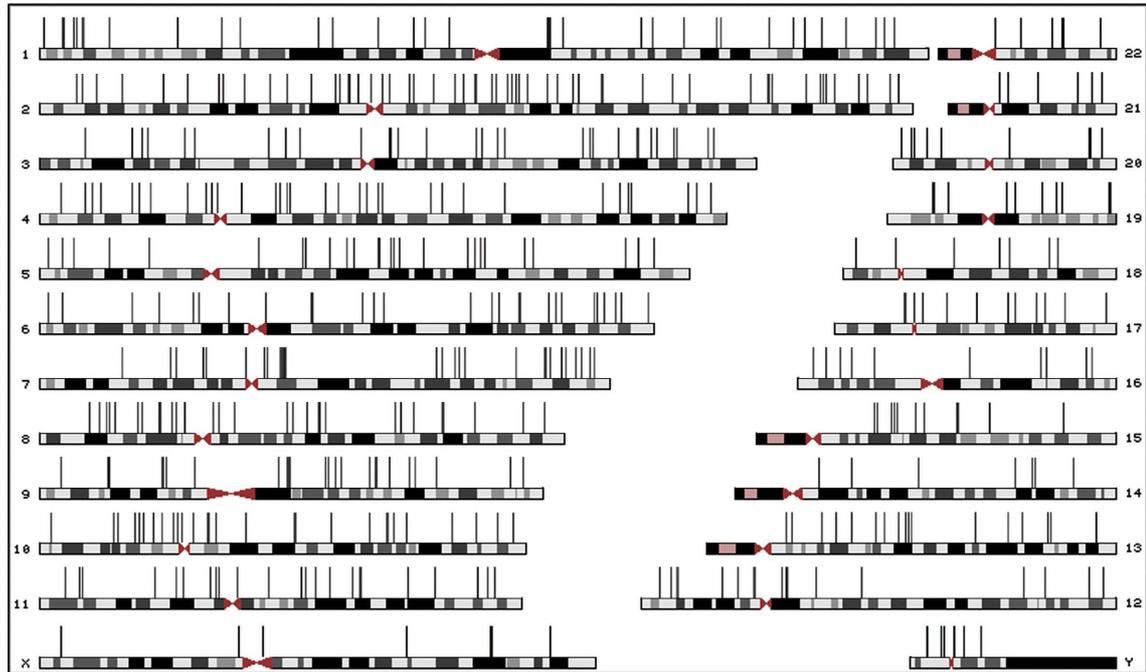


Figure 11 : localisation des pseudogènes mitochondriaux sur le caryotype humain (SIMONE *et al.* 2011).

4. Principaux rôles de la mitochondrie

4.1. Fonctions métaboliques de la mitochondrie

La mitochondrie est impliquée dans de nombreuses voies métaboliques telle que le cycle de Krebs, la β oxydation des acides gras et la cétogenèse. Elle est également le siège, en interaction avec le cytosol de certaines étapes du métabolisme des acides aminés, de la néoglucogenèse, du cycle de l'urée et de la biosynthèse des stérols à partir du cholestérol (grâce au Cytochromes P450). Cependant, sa principale fonction consiste en la production d'énergie utilisable par la cellule sous la forme de molécules d'ATP, et ce grâce au système des phosphorylations oxydatives (OXPHOS) qui se déroulent au niveau de la membrane interne mitochondriale

4.1.1. Quelques définitions

4.1.1.1. L'ATP ou adénosine triphosphate

L'adénosine triphosphate est un nucléotide constitué d'un ribonucléoside, c'est-à-dire d'une base : l'adénine et d'un ribose, associé à trois groupements phosphate. Elle est synthétisée par phosphorylation de l'ADP (Adénosine di-phosphate) (figure 12) Les liaisons entre ces groupements phosphate sont riche en énergie. De ce fait cette molécule possède un haut potentiel de transfert d'énergie et est utilisée dans de nombreuses réactions et processus biologiques notamment pour les contractions musculaires, le transport actif, l'amplification de signaux et les biosynthèses... Cette molécule rend favorable des réactions thermodynamiquement défavorables.

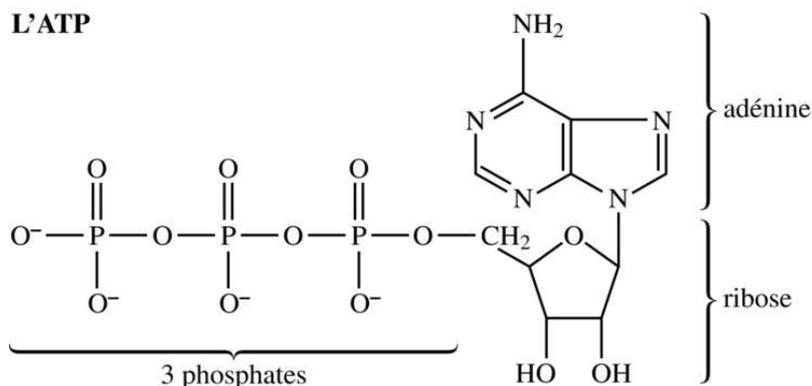


Figure 12: la molécule d'ATP est constituée de trois groupements phosphates, un ribose et une base azotée l'adénine.

4.1.1.2. L'Oxydoréduction et couples Redox

Une réaction d'oxydoréduction ou redox est une réaction chimique au cours de laquelle se produit un échange d'électrons. L'espèce chimique qui donne les électrons est appelée « réducteur » et sera sous sa forme oxydée une fois les électrons cédés ; celle qui les accepte est appelée « oxydant » qui sera sous forme réduite, une fois les électrons captés. Les deux éléments forment ainsi un « couple Redox ». La première réaction d'oxydoréduction qui a été décrite impliquait une molécule d'oxygène O₂ qui est une molécule très électronégative et possède la propriété de capter facilement les électrons. En effet, Lavoisier fut le premier à mettre en évidence cette réaction en 1772 entre le mercure et l'oxygène. Le nom d'oxydation a été donné à la réaction de perte des électrons, car celle-ci s'accompagnait d'une combinaison avec l'oxygène, premier élément « oxydant » décrit. Cependant, le caractère « oxydant » ou « réducteur » est relatif dans le cadre d'une réaction chimique. En effet, un élément réducteur dans une réaction peut être oxydant dans une autre, ceci va dépendre de la force oxydante (ou, dans l'autre sens, de force réductrice) : c'est le potentiel d'oxydoréduction. En outre, ce potentiel peut dépendre du contexte chimique et notamment du pH, des effets de la lumière ou de la température. Dans les conditions physiologiques de la cellule, de nombreuses réactions d'oxydoréduction surviennent dans différentes voies métaboliques et impliquent trois principaux couples redox : NAD⁺ / NADH,H⁺, NADP⁺ / NADPH, FAD / FADH₂.

En effet, le NAD (nicotinamide adénine dinucléotide), le nicotinamide adénine dinucléotide phosphate (NADP) et la Flavine adénine dinucléotide (FAD), sont des coenzymes d'oxydoréduction des cellules vivantes. Ils sont synthétisés à partir de molécules de vitamines B₃ (PP) pour les deux premiers (NAD et NADP) et à partir de la vitamine B₂ (riboflavine) pour le troisième (FAD). Ils vont pouvoir de par leur capacité à transporter les électrons de manière réversible, permettre les réactions métaboliques qui nécessitent une oxydoréduction et assurer le transport des électrons dans la cellule, ce transport d'électrons s'accompagne également d'un transfert d'ions H⁺ (figure 13).

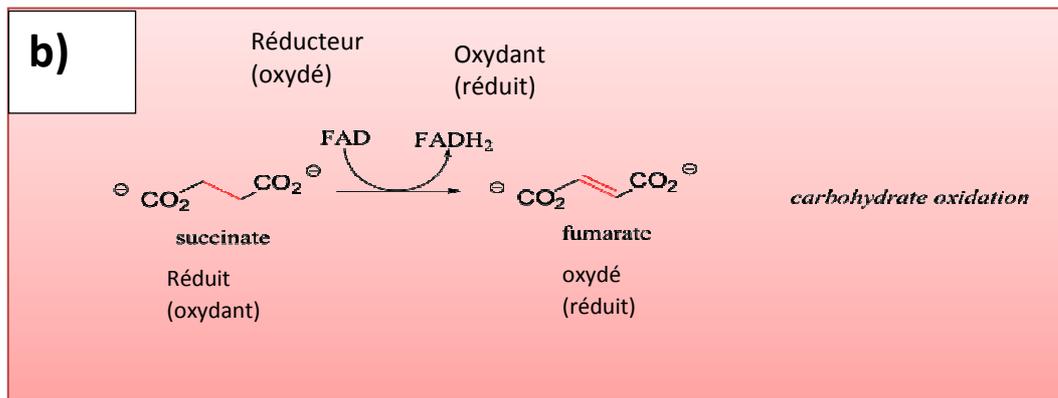
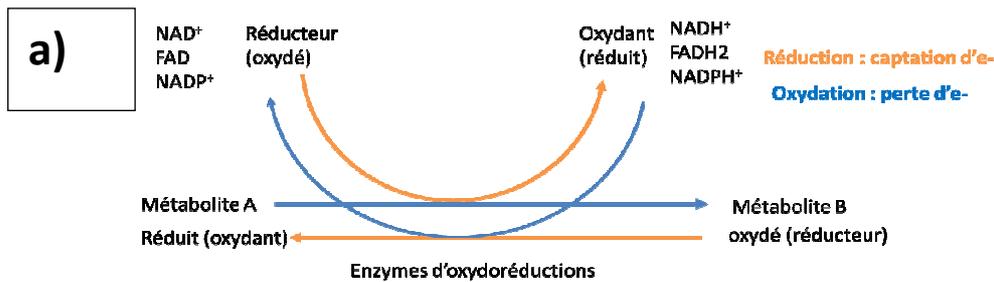


Figure 13 : a) principe d'une réaction d'oxydoréduction et couples Redox b) exemple d'une réaction d'oxydoréduction mitochondriales couple succinate/fumarate et FAD/FADH₂

Dans la mitochondrie, deux couples Redox parmi les trois cités : le NAD⁺/NADH et le FAD/FADH₂, vont rentrer en jeu lors de divers réactions enzymatiques mitochondriales et notamment à certaines étapes du cycle de Krebs et de la bêta-oxydation des acides gras. En effet, les équivalents réduits générés par ces réactions vont transporter les électrons jusqu'aux complexes de la chaîne respiratoire, au niveau de la membrane interne pour une série d'oxydoréduction qui permettra la production finale d'une molécule d'ATP. Le NADPH⁺ est produit principalement par les réactions de la voie des pentoses phosphate.

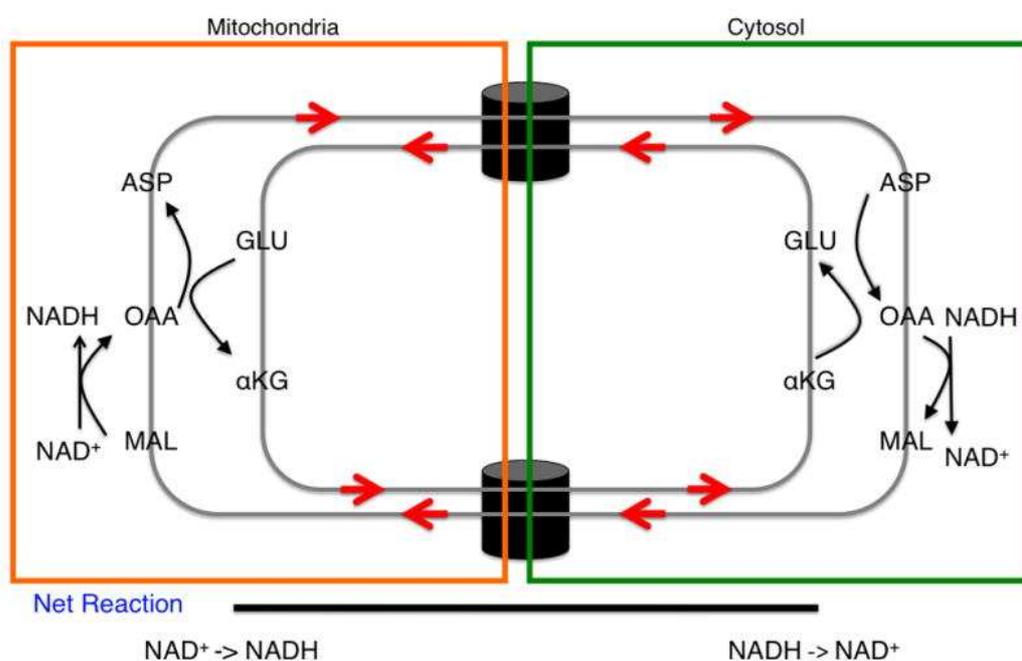
4.1.2. Origines des équivalents réduits FADH₂ et NADH

4.1.2.1. Production mitochondriale

Les équivalents réduits sont fournis essentiellement par le Cycle de Krebs vers lequel converge toutes les voies du métabolisme énergétique grâce à la production de l'Acétyl-CoA. En effet, l'Acétyl-CoA peut être produit par déshydrogénation du pyruvate issu du catabolisme des carbohydrates mais aussi par la bêta- oxydation des acides gras et par la cétolyse à l'état de jeûne. Chaque cycle de Krebs produit un total de 3 NADH et un FADH₂. Il est à noter que chaque cycle de la bêta-oxydation des acides gras (hélice de Lynen) en plus de fournir un Acétyl-CoA produit une molécule de FADH₂ et un NADH.

4.1.2.2. Import des NADH cytosoliques

Leur source principale est la glycolyse et ils sont importés vers les mitochondries via la navette malate/aspartate (BARRON *et al.* 1998). Le NADH transfère ses deux électrons à une molécule d'oxaloacétate cytosolique, générant ainsi du malate qui transporté dans la mitochondrie est de nouveau oxydé en oxaloacétate, libérant une molécule de NADH par réduction de NAD⁺ (figure 14).



(BARRON *et al.* 1998)

Figure 14 : Navette malate/aspartate : le NADH transfère ses deux électrons à une molécule d'oxaloacétate cytosolique sous l'action d'une malate déshydrogénase du cytosol qui synthétise une molécule de malate cytosolique. Celle-ci va emprunter un antiport malate- α -cétoglutarate pour franchir la membrane mitochondriale interne. Au niveau de la matrice mitochondriale le malate est oxydé à nouveau en oxaloacétate par une malate déshydrogénase mitochondriale convertissant une molécule de NAD⁺ en NADH qui est régénéré, l'oxaloacétate est ensuite converti en aspartate sous l'action d'une aspartate aminotransférase mitochondriale par transfert d'un groupe amine du glutamate l'aspartate emprunte un antiport glutamate-aspartate pour regagner le cytosol. L'oxaloacétate cytosolique est régénéré à partir de l'aspartate grâce à une aspartate aminotransférase cytosolique qui produit également du glutamate à partir de l' α -cétoglutarate. L'oxaloacétate cytosolique pourra de nouveau servir comme substrat à la malate déshydrogénase cytosolique pour produire du malate et oxyder le NADH cytosolique. Ce mécanisme présente l'avantage de préserver le potentiel de transfert des électrons produits par les oxydations cytosoliques, car ces électrons sont transférés indirectement d'un NADH cytosolique à un NADH mitochondrial.

4.1.3. La chaîne respiratoire et la série d'oxydations phosphorylantes

4.1.3.1. Généralités

Les équivalents réduits NADH et FADH₂ sont réoxydés par transfert de leurs électrons sur un accepteur final : l'oxygène. Ce transfert se fait par l'intermédiaire d'une chaîne de transport, située dans la membrane interne et sur laquelle les électrons vont transiter en fonction de leur potentiel d'oxydo-réduction. Ainsi, le NAD⁺ ayant une faible affinité pour les électrons, sa forme réduite, le NADH, peut facilement perdre deux électrons. L'oxygène moléculaire possède une très forte affinité pour les électrons et capture facilement ceux des autres molécules. Il va donc s'établir le long de la chaîne respiratoire un gradient de potentiels d'oxydoréduction et les électrons vont transiter de molécules à faible potentiel (tels que le NAD⁺ qui a un potentiel de - 320 mV) vers des molécules à fort potentiel (tels que l'oxygène qui a un potentiel de + 816 mV) (figure 15). Au niveau des complexes I, III et IV de la chaîne respiratoire le transfert des électrons s'accompagne d'un pompage de protons de la matrice vers l'espace inter-membranaire, générant ainsi une force proton-motrice. Ce gradient électrochimique est ensuite utilisé par l'ATP synthase ou complexe V pour la phosphorylation de l'ADP en ATP. L'ATP ainsi formé est alors exporté dans le cytosol grâce à la translocase des nucléotides adényliques (ANT). Ce phénomène est la base de la théorie chimio-osmotique émise par Peter Mitchell en 1961 (prix Nobel en 1978). Dans ce processus, l'oxydation des substrats est donc couplée à la synthèse d'ATP par la force proton-motrice. En effet, la translocation des protons dans l'espace inter-membranaire induit une différence de concentration en protons et de charges électriques entre la matrice et l'espace inter-membranaire. Cette différence génère la force proton-motrice qui résulte du gradient électrochimique créé. Somme des gradients de pH et de potentiel membranaire.

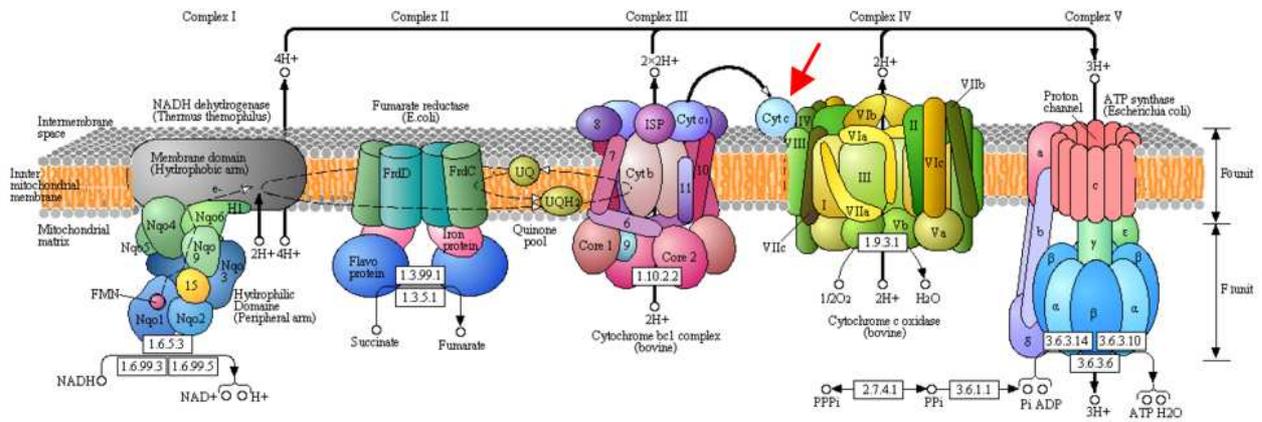


Figure 15 : schématisation de la chaîne des oxydations phosphorylantes

4.1.3.2. Les complexes de la chaîne respiratoire

Le complexe I : ou NADH ubiquinone oxydoreductase est la première enzyme de la chaîne respiratoire. Il catalyse le transfert de deux électrons du NADH à l'ubiquinone, couplé à la translocation de quatre protons au travers de la membrane, ce qui participe à la force protonmotrice (SAZANOV 2007). Cette enzyme est constituée de plus de 40 sous-unités (CARROLL *et al.* 2006) (CHOMOVA and RACAY 2010) (SAZANOV 2015), la plupart sont codées par le génome nucléaire, tandis que 7 autres sont codées par le génome mitochondrial. Avec son poids moléculaire de 980 kDa environ, elle constitue l'un des plus gros complexes protéiques membranaires. Le domaine hydrophile de ce complexe contient le site de liaison du NADH, la flavine mononucléotide qui est le premier accepteur d'électrons et huit ou neuf centres Fer/Soufre (FeS) constituant le centre catalytique de l'enzyme (SAZANOV 2007). Le système de pompage des protons est quant à lui localisé dans la partie membranaire du complexe.

Le complexe II : ou succinate deshydrogénase est composé de quatre sous-unités (A à D). Les sous-unités A et B constituent le domaine hydrophile de l'enzyme dépassant dans la matrice tandis que les sous-unités C et D constituent le domaine d'ancrage du complexe dans la membrane interne mitochondriale. Les quatre gènes codant pour ces quatre sous-unités sont uniquement nucléaires. Ce complexe joue un rôle important dans la mitochondrie puisqu'il fait à la fois partie de la chaîne respiratoire et du cycle de Krebs (BRIERE *et al.* 2005). Il participe à la chaîne respiratoire en transférant les 2 électrons du succinate au Coenzyme Q générant une molécule de fumarate. Le complexe II possède un groupement FAD qui accepte initialement les électrons du succinate et les transmet au site de liaison du CoQ via une série de centre Fe-S. Le transfert d'électrons par le complexe II ne s'accompagne pas de passage de protons dans l'espace inter-membranaire.

Le Coenzyme Q : CoQ (forme quinone) est un composé liposoluble qui peut se mouvoir à l'intérieur de la membrane interne. Il est donc réduit par le complexe I ou le complexe II en CoQH₂ (forme quinol) (HORSEFIELD *et al.* 2006). L'oxydation des équivalents réduits des nutriments conduit donc à la création d'un pool de CoQH₂, porteurs des électrons, qui se déplacent dans la membrane mitochondriale jusqu'au complexe III.

Le complexe III : ou ubiquinol-cytochrome c réductase est composé au total de 11 sous-unités protéiques et contient une centre Fer/Souffre. Une seule sous-unité est codée par le génome mitochondrial (cytochrome b) (BORISOV 2002). Cette enzyme catalyse le transfert de deux électrons

de l'ubiquinol au cytochrome c. Cette étape est associée au transfert de 4 protons de la matrice vers l'espace inter membranaire.

Le cytochrome C : est une hémoprotéine hydrosoluble de 104 acides aminés, Elle est codée par l'ADNmt et localisée dans l'espace inter-membranaire mitochondrial dans les conditions normales (HUTTEMANN *et al.* 2011).

Le complexe IV : ou cytochrome c oxydase appartient à la superfamille des oxydases à hème-cuivre. Son isoforme humaine est composée de 13 sous-unités dont trois (COX I, COX II et COX III) sont codées par le génome mitochondrial (BORISOV 2002). Les électrons cédés par le cytochrome c entrent dans ce complexe protéique, se fixent au niveau du site actif (hème-cuivre) pour être transférés à l'oxygène qui sera réduit en molécule d'eau (ABRAMSON *et al.* 2001). Durant cette étape, 4 protons sont consommés et 4 autres sont transloqués de la matrice vers l'espace inter membranaire participant à la force protonmotrice.

L'ATP synthase : elle est constituée de 2 segments fonctionnels : le segment Fo et le segment F1. Le segment Fo est un segment transmembranaire composé de multiples sous-unités formant le canal à protons (HABERSETZER *et al.* 2013). Le segment F1 est constitué par des sous-unités α , β , γ , ϵ et δ , est le composant catalytique de l'enzyme. En effet, la sous-unité α possède les sites de liaison pour l'ADP et le Pi. L'association des sous-unités α et β forme une structure en forme d'hélice. Elle est associée au segment Fo par la sous-unité γ . Le passage des protons dans le canal du segment Fo va induire la rotation de la structure en hélice via la sous-unité γ . Cette rotation permet la liaison entre l'ADP et le Pi (phosphate inorganique) et la formation de l'ATP.

4.3.1.3. L'existence de super-complexes

Des expériences de co-précipitation de complexes et des observations en microscopie électronique ont permis de mettre en évidence l'existence, dans la membrane interne mitochondriale, d'unités fonctionnelles composées de « supercomplexes » regroupant plusieurs complexes de la chaîne respiratoire (GENOVA and LENAZ 2014) (BOUMANS *et al.* 1998) (DUDKINA *et al.* 2005). La formation de ces structures serait possible grâce à la présence des cardiolipines (ZHANG *et al.* 2005).

Quatre associations principales ont été retrouvées : dimères d'ATP synthase, supercomplexe III₂+IV₁-2, supercomplexe I+III₂ et le supercomplexe I+III₂+IV₁. La formation de ces supercomplexes semble avoir une importance physiologique, puisque ces structures pourraient augmenter la vitesse de transfert des électrons, jouer un rôle régulateur sur le fonctionnement de la chaîne respiratoire, influencer la structure des crêtes de la membrane interne, augmenter la stabilité des complexes et enfin

augmenter la capacité d'insertion des protéines dans la membrane interne mitochondriale (BOEKEMA and BRAUN 2007).

Ces interactions ont cependant été, pour la majorité, identifiées par des expériences de co-immunoprécipitation et la fonctionnalité de ces structures *in vivo* doit encore être démontrée.

4.3.1.4. Production mitochondriale des espèces réactives de l'oxygène

Les espèces réactives de l'oxygène (reactive oxygen species ou ROS) sont des espèces chimiques oxygénées pouvant induire l'oxydation d'autres molécules. Elles comprennent, notamment, l'anion superoxyde ($O_2^{\cdot-}$), le peroxyde d'hydrogène (H_2O_2) et les radicaux peroxydes ($RCOO^{\cdot}$ et OH^{\cdot}). Lors du fonctionnement de la phosphorylation oxydative, il arrive que des électrons quittent la chaîne respiratoire, ils ont alors la possibilité d'interagir avec l' O_2 (qui possède un fort potentiel d'oxydoréduction). Le passage d'un électron sur la molécule d' O_2 forme un $O_2^{\cdot-}$ (SARASTE 1999). C'est une réduction incomplète de l' O_2 , contrairement à la réduction réalisée pendant la chaîne de phosphorylation oxydative qui transfère 4 électrons (figure 16). Cet anion est presque instantanément pris en charge par une enzyme anti-oxydante mitochondriale, la superoxyde dismutase qui catalyse la formation d' H_2O_2 et d' O_2 à partir de 2 molécules d' $O_2^{\cdot-}$ et de 2 protons. L' H_2O_2 possède une durée de vie beaucoup plus longue que l' $O_2^{\cdot-}$, et il peut induire l'oxydation des molécules environnantes telles que les protéines, l'ADN et les membranes cellulaires et peut diffuser à travers les membranes biologiques. Son élimination est prise en charge par des enzymes anti-oxydantes, principalement par la glutathion peroxydase .

La formation d' $O_2^{\cdot-}$ nécessite une grande proximité entre l' O_2 et les sites de transport des électrons, limitant le nombre de sites de formation de ROS possibles. Les études réalisées sur mitochondries isolées ont permis de mettre en évidence 2 sites de formation de ROS au sein de la chaîne respiratoire : le complexe I et le complexe III (KUSSMAUL and HIRST 2006). Les premières estimations avaient suggéré que la production de ROS représentait approximativement 1 à 2 % de la consommation d'oxygène d'un individu (ROLFE and BROWN 1997). Ce chiffre, basé sur les expériences menées *in vitro*, semble être surévalué. Les travaux réalisés avec des concentrations de substrats et une pression partielle en oxygène plus physiologiques montrent que la production de ROS, *in vivo*, représenterait, en fait, moins de 1% de notre consommation d'oxygène (MURPHY 2009).

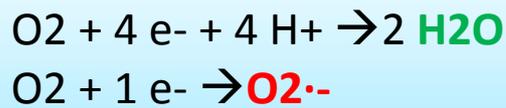


Figure 16: la réduction incomplète de molécule d'O₂ est à l'origine de la production d'espèces réactives.

Dans certaines situations, il apparaît un déséquilibre provoqué par une production exagérée de radicaux libres ou par une diminution des systèmes de défense (enzymatiques et non enzymatiques), ou encore par une association de ces deux phénomènes. Un tel déséquilibre entre systèmes producteurs de ROS et systèmes de défense, caractérise l'état de stress oxydant, sans qu'il soit aisé de déterminer si ce dernier est causal ou s'il constitue seulement une réponse de l'organisme à des stimuli, notamment inflammatoire ou à des agressions. Des expérimentations, réalisées sur des souris invalidées pour le gène de la superoxyde dismutase mitochondriale, montrent que la perte de l'une des enzymes majeures des défenses anti-oxydantes induit des phénotypes pathologiques sévères, impliquant notamment une dégénérescence neuronale (LEBOVITZ 1996). Les ROS peuvent aussi avoir un impact direct sur le fonctionnement de la phosphorylation oxydative, en altérant les complexes de la chaîne respiratoire. Ainsi, un régime hypercalorique est associé à une augmentation de la production de ROS et à une diminution de l'activité des complexes I et III. Ce processus serait lié à la peroxydation des cardiolipines, étroitement associées aux complexes. En effet, l'ajout de cardiolipines par fusion de liposomes permet la restauration de l'activité du complexe et ce processus n'est pas observé lors d'ajout d'autres phospholipides (phosphatidylcholine ou phosphatidyléthanolamine) ou par l'ajout de cardiolipines peroxydées (PARADIES *et al.* 2001). Par ailleurs, à des taux physiologiques, les ROS sont des acteurs essentiels du fonctionnement cellulaire normal, ils sont notamment impliqués dans la prolifération, la croissance et la survie cellulaire et participent à de nombreuses voies de signalisation.

4.3.1.5. La régulation de la phosphorylation oxydative

Il n'existe pas de stockage cellulaire de l'ATP, la mitochondrie doit donc s'adapter en permanence à la demande énergétique de la cellule ou du tissu dont elle dépend. Cette demande d'énergie peut varier d'un facteur 1 à 10 et nécessite donc une régulation rapide de la phosphorylation oxydative (KADENBACH *et al.* 2010), tout en évitant d'augmenter la production de ROS. Celle-ci passe par la régulation de l'intensité et de la vitesse de production d'ATP à plusieurs niveaux.

Il existe par exemple une régulation des complexes au niveau transcriptionnel, certaines de leurs nombreuses sous-unités possèdent plusieurs isoformes dont l'expression dépend du tissu ou des conditions environnementales. L'expression de ces différentes isoformes pourrait avoir des répercussions sur les propriétés des complexes et leur activité (ANTHONY *et al.* 1993). Ces mêmes complexes peuvent aussi faire l'objet d'une inhibition post-traductionnelle par phosphorylations réversibles.

Le taux d'ATP peut aussi réguler directement le fonctionnement de la chaîne respiratoire. En effet, l'ATP aurait un effet allostérique sur le complexe IV, en se fixant au niveau de sa face matricielle. Cette fixation induirait un changement de conformation de l'enzyme et un ralentissement de la vitesse de transfert des électrons (ARNOLD and KADENBACH 1997). A l'inverse, si la consommation d'ATP est augmentée, il y aura une augmentation du transport de l'ATP vers la cytosol et inversement de l'ADP vers la matrice mitochondriale, l'ADP va ensuite stimuler directement l'activité de l'ATP synthase. Par ailleurs, l'efficacité de la phosphorylation oxydative est déterminée par le degré du couplage entre le fonctionnement de la chaîne respiratoire (reflété par la quantité d'oxygène utilisée) et l'activité de l'ATP synthase (reflétée par la quantité d'ATP produite). Lors de la baisse de la demande énergétique, il y a un découplage entre le transfert des électrons et le pompage des protons. Dans ce cas, le nombre de protons pompés par électron transféré sera diminué (découplage intrinsèque). Il semble également exister une certaine dissipation de cette force par diffusion passive des protons à travers la membrane mitochondriale (découplage extrinsèque). Cette diffusion pourrait être autorisée par une modification des propriétés membranaires ou par l'activation de transporteurs (rôle des protéines découplantes chez les mammifères ?) (CAPITANIO *et al.* 1996).

4.2. Autres fonctions mitochondriales

En plus d'être un carrefour énergétique important, la mitochondrie intervient dans beaucoup d'autres voies physiologiques, notamment dans la thermogénèse, l'homéostasie du calcium, la synthèse des complexes fer-soufres etc... La mitochondrie joue surtout un rôle fondamental dans la mort cellulaire programmée. En effet, la plupart des signaux de mort convergent vers la mitochondrie qui tient le rôle d'un centre intégrateur et exécuteur pour l'apoptose. La mitochondrie est d'ailleurs une cible thérapeutique intéressante dans les pathologies tumorales (cf. Annexe 1 pour plus de détails)

II-Cytopathies mitochondriales

Les cytopathies mitochondriales regroupent une grande variété de pathologies dont le l'origine commune est le déficit génétique de la chaîne respiratoire mitochondriale, qu'il soit nucléaire ou mitochondriale.

1. Epidémiologie

Il est difficile d'estimer réellement la survenue des cytopathies mitochondriales dans la population générale. Néanmoins, une équipe a estimé la prévalence des cytopathies mitochondriales sur une période de 10 ans (tous déficits confondus), dans une population d'individus âgés de plus de 14 ans à 11,5/100 000 personnes soit 1/8500 personnes (ARPA *et al.* 2003). Les mutations ponctuelles du génome mitochondrial ont été évaluées à 5.24/100 000 soit environ 1/20 000 personnes dans les régions du Nord de l'Angleterre (CHINNERY *et al.* 2000).

2. Expressions cliniques

Ces pathologies ont une expression très hétérogène aussi bien dans la présentation clinique que dans l'âge d'apparition des symptômes. En effet, les cytopathies mitochondriales peuvent apparaître à n'importe quel âge, de la période néonatale voire dans certains cas anténatale à l'âge adulte, parfois avancé et peuvent être d'une sévérité variable. Néanmoins, la plupart du temps, une atteinte de la chaîne respiratoire va s'exprimer préférentiellement de manière multisytémique et ce de par la fonction ubiquitaire des mitochondries. Ainsi, une pathologie mitochondriale doit être évoquée devant une association de symptômes qui vont s'additionner sur un mode évolutif, touchant plusieurs organes ou tissus et pouvant s'aggraver dans les situations d'hypercatabolisme et de besoins énergétiques accrus. Les symptômes fréquemment rencontrés sont résumés dans le tableau 1 (**en Annexe 2**). Il est important de noter que dans certains cas, le déficit de la chaîne respiratoire peut donner une atteinte d'organe isolée comme la surdité, une atrophie optique ou encore une myopathie. Lors des 40 dernières années, plusieurs associations syndromiques ont été identifiées qu'il est bon de connaître (tableau 2 en **Annexe 2**). Il ne faut cependant pas oublier que ces syndromes, même s'ils peuvent être très évocateurs voir spécifiques, ne constituent pas tous les modes de présentations des cytopathies mitochondriales. De plus, les corrélations génotype phénotype sont très faibles pour ce groupe de pathologies, la caractérisation d'une association syndromique ne présage pas de la mutation causale. En effet, pour une même présentation clinique, plusieurs mutations peuvent être impliquées aussi bien dans le génome mitochondrial que nucléaire, et de même, une mutation peut donner lieu à des expressions cliniques très différentes. On peut citer comme exemple : les syndromes de Kearns-Sayre et

de Pearson, causés par une même délétion de l'ADNmt, le syndrome de Leigh et les associations diabète surdité qui peuvent être dus aussi bien à des mutations de l'ADNmt qu'à des mutations du génome nucléaire.

A ce jour, on ne connaît pas les facteurs qui déterminent la sévérité et le type d'atteinte causés par une mutation donnée. Une ébauche d'explication est l'existence d'un effet seuil, au-delà duquel le déficit de la chaîne respiratoire provoque des dommages et devient symptomatique. Ce seuil semble varier entre les différents organes et tissus, ce qui expliquerait peut-être certaines atteintes préférentielles d'organe liées à certaines mutations. Néanmoins, on ne sait toujours pas pourquoi une même mutation va donner d'un côté une myopathie modérée chez un adulte et de l'autre provoquer une encéphalomyopathie néonatale sévère, comme ce qui est le cas dans les mutations du gène *TK2* (BEHIN *et al.* 2012). Il faut souligner l'importance de l'identification des facteurs qui influencent la sévérité du phénotype des cytopathies mitochondriales car ils pourraient être des cibles thérapeutiques intéressantes. Et ce, d'autant plus qu'il n'y a pas de toujours de corrélation entre le taux de molécules d'ADNmt mutées (hétéroplasmie) et la sévérité clinique. Cette même variabilité s'applique aussi aux déficits liés aux mutations nucléaires qui sont présentes dans tous les tissus.

Les manifestations anténatales :

Les manifestations anténatales sont peu décrites, quelques cas ont néanmoins été rapportés dans la littérature. Il semblerait que le retard de croissance intra-utérin (RCIU) isolé soit le symptôme le plus fréquent (dans 20 % des cas) avec des grossesses menées à terme (VON KLEIST-RETZOW *et al.* 1998). D'autres manifestations ont été plus rarement décrites, à type d'atteintes d'organes isolées, s'exprimant par exemple par une anasarque fœtale dans les atteintes hépatiques, cardiaques ou hématologiques. Les anomalies cérébrales peuvent également se voir : une agénésie du corps calleux, une leucomalacie periventriculaire (BRUN *et al.* 1999), des calcifications intracérébrales et des ventricules élargis (SAMSON *et al.* 1994), et enfin des associations malformatives de type « VATER » ont également été décrites (DAMIAN *et al.* 1996). Néanmoins, ni le RCIU ni les autres anomalies ne sont spécifiques d'une maladie mitochondriale.

Cette grande hétérogénéité clinique et l'absence de signes spécifiques d'orientation, en dehors de quelques rares syndromes, rendent le diagnostic extrêmement difficile, et ce d'autant plus que les investigations biologiques et les explorations fonctionnelles ne sont pas toujours concluantes. Plusieurs équipes ont tenté d'établir des scores cliniques prédictifs et des arbres décisionnels, qui sont plus au moins efficaces. Dans la figure 17, est illustré un exemple de score clinique, établi pour des populations pédiatriques en se basant sur des critères cliniques et paracliniques (WOLF and SMEITINK 2002).

Table Mitochondrial disease criteria (simplified version for bedside use)*

I. Clinical signs and symptoms, 1 point/symptom (max. 4 points)				
A. Muscular presentation (max. 2 points)	B. CNS presentation (max. 2 points)	C. Multisystem disease (max. 3 points)	II. Metabolic/imaging studies (max. 4 points)	III. Morphology (max. 4 points)
Ophthalmoplegia†	Developmental delay	Hematology	Elevated lactate†	Ragged red/blue fibers‡
Facies myopathica	Loss of skills	GI tract	Elevated L/P ratio	COX-negative fibers‡
Exercise intolerance	Stroke-like episode	Endocrine/growth	Elevated alanine†	Reduced COX staining‡
Muscle weakness	Migraine	Heart	Elevated CSF lactate†	Reduced SDH staining
Rhabdomyolysis	Seizures	Kidney	Elevated CSF protein	SDH positive blood vessels†
Abnormal EMG	Myoclonus	Vision	Elevated CSF alanine†	Abnormal mitochondria/EM†
	Cortical blindness	Hearing	Urinary TA excretion†	
	Pyramidal signs	Neuropathy	Ethylmalonic aciduria	
	Extrapyramidal signs	Recurrent/familial	Stroke-like picture/MRI	
	Brainstem involvement		Leigh syndrome/MRI†	
			Elevated lactate/MRS	

* Score 1: mitochondrial disorder unlikely; score 2 to 4: possible mitochondrial disorder; score 5 to 7: probable mitochondrial disorder; score 8 to 12: definite mitochondrial disorder.

† This specific symptom scores 2 points.

‡ This symptom in a higher percentage scores 4 points.

GI = gastrointestinal; L/P = lactate/pyruvate; COX = cytochrome c oxidase; SDH = succinate dehydrogenase; EM = electron microscopy; EMG = electromyography; TA = tricarboxylic acid.

Figure 17 : exemple de score clinique (MDC) (WOLF and SMEITINK 2002).

3. Explorations paracliniques

Du fait de l'atteinte multi-systémique, une évaluation paraclinique globale de tous les organes doit être réalisée à titre systématique, devant toute suspicion d'une cytopathie mitochondriale, hépatique, rénale, hémérologique, cardiaque... Certains examens sont cependant plus spécifiques et aident à l'orientation diagnostique et notamment le bilan métabolique, l'imagerie cérébrale ainsi que les explorations enzymologiques et histologiques.

3.1. Bilan métabolique

3.1.1. Les points redox

Le défaut de transfert des électrons par déficit de la chaîne respiratoire provoque une modification profonde des équilibres des réactions d'oxydoréduction dans la cellule par accumulation d'équivalents réduits (NADH, FADH₂). Ceci peut se traduire par une hyperlactatémie persistante (> 2,5 mM), une élévation du rapport lactate/pyruvate (L/P > 20) et du rapport des corps cétoniques 3-hydroxybutyrate/acétoacétate (3OHB/AcAc > 2). En effet, dans la mitochondrie, les corps cétoniques : acétoacétate et 3-hydroxybutyrate sont en équilibre, de même que le lactate et le pyruvate dans le

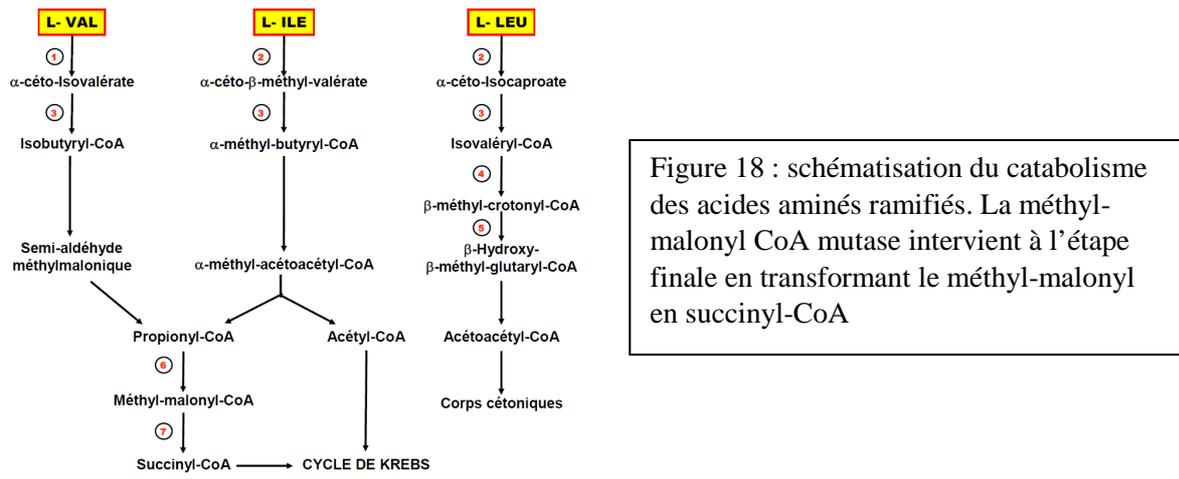
cytoplasme. L'accumulation de NADH pousse les deux réactions dans le sens de la réduction des métabolites, ainsi, l'acétoacétate est préférentiellement transformé en 3-hydroxybutyrate et le pyruvate en lactate, entraînant l'élévation des rapports 3-hydroxybutyrate/acétoacétate et Lactate/pyruvate (DEBRAY *et al.* 2007). Ces anomalies ne sont pas constantes, elles manquent de spécificité et de sensibilité, globalement moins de 10% des hyperlactatémies sont confirmées par un 2^{ème} dosage. L'hyperlactémie dépend probablement de l'intensité de l'atteinte mitochondriale, des organes préférentiellement touchés et de la balance sécrétion/élimination du lactate plasmatique. Certaines formes de cytopathies mitochondriales s'accompagnent rarement d'une élévation du lactate sanguin comme dans les mutations de la polymérase gamma (*POLG1*), l'atrophie optique de Leber, le syndrome de Kearns-Sayre et le syndrome de Leigh par déficit du complexe I (TRIEPELS *et al.* 1999). Les dosages sanguins des points redox doivent, de manière générale être réalisés au repos, répétés au cours de la journée, à jeun et en postprandial, afin de démasquer une hyperlactacidémie latente. Une élévation des lactates doit également être recherchée dans le liquide céphalo-rachidien, surtout dans les formes neurologiques et dans les urines, notamment en cas de tubulopathie. Par ailleurs, les anomalies des lactates sanguins ne sont pas propres aux cytopathies mitochondriales. Les lactates peuvent être augmentés dans le sang, dans diverses autres conditions physiologiques et pathologiques, par exemple, dans les conditions d'hypoxie (insuffisance respiratoire et cardiaques), en période post-critique dans les épilepsies, dans les accidents vasculaires cérébraux et aussi pendant les processus inflammatoires et tumoraux.... Cette élévation peut être secondaire à d'autres déficits métaboliques, tels le déficit en thiamine, de la bêta-oxydation des acides gras ou le déficit de l'enzyme pyruvate déshydrogénase, enzyme d'entrée du pyruvate dans le cycle de Krebs. L'élévation peut aussi être artéfactuelle, liée à des mauvaises conditions de prélèvement et d'acheminant ; en effet, le prélèvement doit se faire en dehors de toute période d'activité physique, sans garrot et doit être transposé à + 4°. (HAAS 1988) (CHOW *et al.* 2005).

3.1.2. Les chromatographies des acides organiques et amines

Les chromatographies des acides aminés et organiques peuvent révéler des anomalies inconstantes et non spécifiques.

La chromatographie des acides organiques urinaires peut retrouver une excrétion anormale des dérivés intermédiaires du cycle de Krebs (succinate, fumarate...), de l'acide 3-méthylglutaconique qui est un intermédiaire du catabolisme de la leucine. Cette accumulation est en partie due à une accumulation de succinyl-CoA, résultat de la saturation du cycle de Krebs par ralentissement du transfert des électrons (figure 18). Une acidurie dicarboxylique peut également se voir, c'est le témoin du retentissement du

déficit de la chaîne respiratoire sur la bêta-oxydation des acides gras, de même que la diminution de la carnitine libre sur le profil des acyl-carnitine avec un rapport acylcarnitine /carnitine libre augmenté. Les chromatographies des acides aminés sanguins peuvent parfois révéler de manière aspécifique, une augmentation de l'alanine (cycle de Cori), une augmentation de la proline avec une hypocitrullinémie (inhibition de la proline oxidase), témoins d'une hyperlactatémie latente. Quelques observations ont également rapporté une élévation de la sarcosine (intermédiaires du métabolisme de la choline) dans le sang et les urines (HAAS *et al.* 2008)



3.2. Imagerie du système nerveux central

3.2.1. IRM cérébrale

L'IRM cérébrale est un examen très utile dans le diagnostic des cytopathies mitochondriales, même si les anomalies retrouvées ne sont pas spécifiques (RUSSELL and TURNBULL 2014). En effet, tous les types d'anomalies de morphologie et de signal peuvent s'observer, notamment dans les formes neurologiques. L'IRM peut parfois aussi être normale sans éliminer pour autant le diagnostic. Il peut exister une atrophie du tissu cérébral, d'évolution progressive ou parfois rapide, lors d'accès paroxystique rapporté notamment dans les déficits du complexe V (CASTRO-GAGO *et al.* 1999). L'atrophie peut toucher le cervelet dans -entre autre- les syndromes de MELAS, de Leigh, de Kearns-Sayre...etc (LINCKE *et al.* 1996) (VALANNE *et al.* 1998). La substance grise profonde et notamment les ganglions de la base est sélectivement vulnérable à une défaillance du métabolisme énergétique, elle est donc souvent anormale dans les cytopathies mitochondriales. Cette atteinte concerne aussi bien les ganglions de la base que le thalamus ou la substance grise périaqueducule, qui se présentent en hyposignal T1 et hypersignal T2spin-écho et FLAIR, préférentiellement bilatérale et symétrique (**Annexe 3A**). Les descriptions histologiques de ces lésions ont retrouvé un aspect nécrotique associé à une démyélinisation, une prolifération vasculaire et une gliose (CAVANAGH and HARDING 1994). Le

syndrome de Leigh est le modèle historique de ce type d'atteinte, mais il faut noter que ces lésions ne sont pas spécifiques des déficits primaires de la chaîne respiratoire, et peuvent également être observées lors d'une intoxication au monoxyde de carbone, d'une toxicité de la bilirubine, d'un déficit en biotinidase, des troubles du métabolisme des acides gras ou dans les aciduries organiques. La souffrance métabolique de la substance blanche se traduit par des anomalies de la myélinisation avec un aspect de leucoencephalopathie à l'IRM (**Annexe 3B**). Elle peut aller du simple hypersignal T2 FLAIR péri-ventriculaire à la leucodystrophie diffuse, respectant les fibres en U (syndrome MNGIE) ou pas (déficit en cytochrome c oxidase) (MORATO *et al.* 2014). Enfin, d'autres atteintes comme les dysgénésies du corps calleux ou des calcifications des noyaux gris centraux ont été également rapportées.

3.2.2. Imagerie des pseudo-AVC

Les lésions de pseudo-AVC sont retrouvées typiquement dans le syndrome de MELAS mais pas seulement (MALFATTI *et al.* 2007) (JANSSEN *et al.* 2006). Sur le plan histologique, ces lésions ont un aspect de plages nécrotiques, des micro-vacuolisations, une prolifération capillaire, une activation macrophagique et une astrogliose (IZUKA and SAKAI 2005). Elles se caractérisent sur le plan radiologique des AVC vrais, par un aspect d'ischémie sans systématisation artérielle, avec une angio-IRM classiquement normale. Elles siègent préférentiellement en régions corticales pariéto-occipitales. Par ailleurs, les pseudo-AVC s'accompagnent en phase aigüe, d'un œdème vasogénique objectivé à l'IRM en diffusion et ADC, contrairement aux AVC d'origine artérielle qui sont plutôt associés à un œdème cytotoxique.

L'objectif des séquences IRM pondérées en diffusion est d'obtenir des images, dont le contraste est influencé par les différences de mobilité de molécule d'eau. Cette séquence permet de faire la différence entre l'œdème vasogénique (molécules d'eau extracellulaire) et l'œdème cytotoxique (molécules d'eau intracellulaires). C'est ainsi que le premier apparaît en hyposignal, dit en restriction de diffusion (mime le signal de LCR : molécules d'eau à très grande mobilité) et le deuxième apparaît en hypersignal (molécule d'eau intracellulaire à moyenne mobilité).

A partir des séquences de diffusion et FLAIR on peut calculer un coefficient dit « coefficient de diffusion apparente » ou ADC qui est la différence de signal entre ces deux séquences. Ce coefficient est élevé dans les œdèmes vasogéniques et diminué dans l'œdème cytotoxique, permettant de différencier à la phase aigüe, les AVC ischémiques des AVC métaboliques (**Annexe 3C**). Il n'est cependant pas toujours si aisé de faire la différence entre les deux, car de plus en plus de cas de MELAS avérés, présentant des AVC avec un ADC diminué sont rapportés.

Tout ceci soulève la question des mécanismes physiopathologiques des pseudo-AVC dans les cytopathies mitochondriales. Cette question n'est pas encore élucidée, néanmoins, trois hypothèses majeures ont été proposées pour expliquer ce phénomène :

-l'origine vasculaire ischémique ou angiopathie mitochondriale : elle se base sur des observations d'atteintes dégénératives des petites artères et artérioles cérébrales, dans les autopsies de patients atteints de syndrome de MELAS (OHAMA *et al.* 1987) (KISHI *et al.* 1988). Cette théorie suggère l'existence d'une dysfonction vasculaire avec un trouble segmentaire de la relaxation du muscle lisse par inhibition de la NO synthétase, par l'accumulation d'équivalents réduits.

- le mécanisme cytopathique : fait référence à une souffrance neuronale et/ou gliale généralisée, par effet direct du déficit de la phosphorylation oxydative avec des vaisseaux préservés.

- le mécanisme neurovasculaire non ischémique : suggère que les pseudo-AVC seraient le résultat d'un état d'hyperexcitabilité (activité épileptique) neuronale locale, ce qui augmenterait les besoins énergétiques cellulaires, créant un déséquilibre entre la demande et la disponibilité d'ATP.

Il semblerait plus probable que tous ces processus concourent de manière conjointe aux lésions observées dans les pseudo-AVC à des étapes et degré variables.

Tout ceci, impose la recherche de cytopathies mitochondriales dans les AVC, même cytotoxiques notamment, chez le sujet jeune surtout en l'absence de systématisation artérielle et de facteur de risques cardiovasculaires.

Certaines équipes ont essayé d'établir des corrélations génotype-phénotype neuroradiologiques plus précises, en étudiant des séries de patients atteints de cytopathies mitochondriales. Ainsi par exemple, dans le syndrome de Leigh, les noyaux préférentiellement atteints sont les putamen++, le pallidum et/ou le noyau caudé et les pseudo-AVC occipitaux sont classiques dans le syndrome de MELAS. Par ailleurs, une étude récente de 30 cas, atteints d'un déficit du complexe I a montré que les déficits d'origine mitochondriale sont plus souvent associés à des épisodes de pseudo-AVC, alors que les déficits d'origine nucléaire s'associent plus volontiers à une leucoencephalopathie. Cette équipe (DEVAUX-BRICOUT *et al.* 2014) a proposé des algorithmes décisionnels en fonction des images de l'IRM cérébrale (**Annexe 3D**). Ces diagrammes sont donnés à titre indicatif et sont à prendre avec précaution mais pourraient parfois, avec une bonne interprétation radiologique aider dans les cas complexes à l'orientation diagnostique.

3.2.3. La spectroscopie par résonance magnétique

La spectroscopie par résonance magnétique permet une exploration non invasive et in vivo de la composition moléculaire des tissus (MANSFIELD and MAUDSLEY 1977) (FILIPPI *et al.* 2014) . Elle permet d'identifier certains constituants moléculaires, les métabolites, impliqués dans des processus physiologiques ou pathologiques. Le signal étudié dans ce cas de figure est celui du noyau hydrogène.

Deux types d'acquisition existent :

-Mono voxel : centré sur un point, elle a l'avantage d'être rapide (<1min) mais elle est limitée, elle trouve son intérêt dans l'exploration d'une anomalie de signal dans un site donné.

-Imagerie métabolique : enregistre l'information spectroscopique sur un ensemble de voxel et donne une information sur un volume cérébral plus important mais l'acquisition est plus longue (13 minutes).

L'acquisition de séquences spectroscopiques (Temps d'Echo long + Temps Echo court) permet au total d'explorer un certain nombre de métabolites :

- Le myo-inositol qui est un marqueur des cellules gliales,
- La choline, qui est un métabolite capital lors de la synthèse et de la dégradation de la myéline ; le pic de choline comprend également la phosphocholine et la glycéro-phosphocholine ; c'est bon marqueur de la myélinisation
- La créatine et la phosphocréatine
- Le N-acétyl-aspartate (NAA), c'est un marqueur neuronal il est diminué dans les la souffrance et la mort des neurones.
- Le glutamate et la glutamine
- Le lactate

Les mesures qu'on peut faire, sont reproductibles et il est donc possible d'avoir une bonne idée des valeurs des rapports entre les différents métabolites dans un cerveau sain. Il convient cependant, de noter que ces valeurs ou ces rapports peuvent varier en fonction de la localisation des mesures (voxel), dans le plan de coupe, et de l'âge du sujet. Ceci rend compte de l'interprétation difficile des spectres, notamment dans la population pédiatrique. Dans les cytopathies mitochondriales, les anomalies retrouvées sont principalement une baisse du pic du NAA, témoin d'une souffrance neuronal et la présence d'un doublet de lactates (**Annexe 3E**). La spectroscopie par résonance magnétique est toutefois, plus contributive quand il y a des anomalies de signal, permettant d'orienter la localisation des acquisitions.

3.3. Etudes histopathologiques

L'examen histopathologique de la biopsie musculaire peut orienter le diagnostic en présence d'une accumulation segmentaire anormale des mitochondries (figure 19A) et/ou d'une surcharge lipidique et par l'existence de fibres présentant un déficit total ou partiel de l'activité de la cytochrome oxydase aux études histoenzymologiques (figure 19B). L'étude en microscopie électronique peut venir confirmer les anomalies en mettant en évidence la présence d'anomalies mitochondriales avec notamment la présence d'inclusions paracrystallines dans la matrice mitochondriale (figure 19 C).

L'accumulation anormale de mitochondries peut se traduire par différents aspects révélés par la coloration du trichrome de Gomori :

- Le plus caractéristique est celui des ragged red fibers (RRF) (fibres rouges déchiquetées) correspondant à une accumulation de mitochondries sous le sarcolemme et au niveau du réseau

intermyofibrillaire produisant un aspect déchiqueté de la fibre.

- l'accumulation peut se limiter à une augmentation des mitochondries uniquement sous le sarcolemme donnant un aspect « en coup d'ongle ».

La nature mitochondriale de cette accumulation peut être confirmée par la mise en évidence d'activités enzymatiques localisées au niveau de la mitochondrie (Succinate deshydrogénase (SDH), cytochrome oxydase (COX)).

La présence de RRF est caractéristique des pathologies mitochondriales mais peut s'observer dans d'autres pathologies notamment inflammatoires et au cours du vieillissement et inversement les RRF sont rares dans certaines pathologies mitochondriales comme les syndromes de Leigh ou de NARP.

Les cytopathies mitochondriales étant secondaires à des déficits enzymatiques de la chaîne respiratoire, il est possible par technique histoenzymologique de révéler ce déficit. L'étude de l'activité de la SDH et de la COX est très informative, la présence de fibres déficientes en cytochrome oxydase avec une activité normale ou augmentée de la SDH étant très suggestive d'une pathologie secondaire à une anomalie de l'ADNmt. Les fibres COX négatives ne sont pas spécifiques d'un déficit en complexe IV de la chaîne respiratoire et sont généralement associées aux anomalies de l'ADNmt touchant les gènes des ARN de transfert.

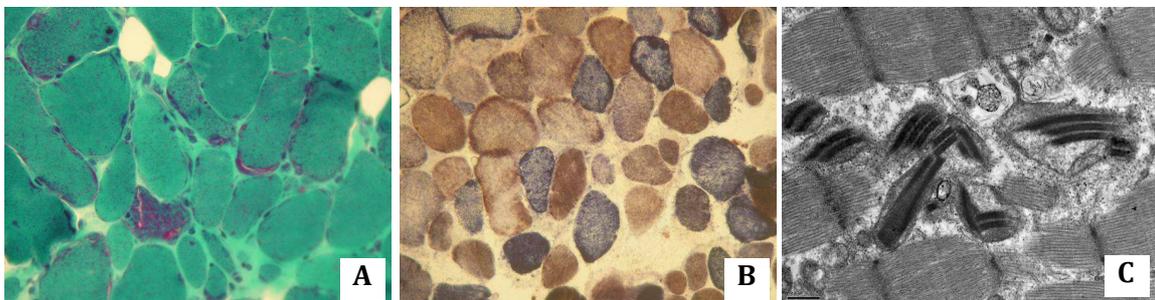


Figure 19 : Myopathie mitochondriale : A-Coloration du trichrome de Gomori : Présence de RRF et d'accumulations sous sarcolemmale de mitochondries au niveau de certaines fibres. B-Détection simultanée des activités SDH et COX. L'activité de la SDH se traduit par un précipité bleu et celui de la COX par un précipité brun. Le déficit en COX au niveau de certaines fibres se traduit par une coloration bleue des fibres contrairement aux autres qui apparaissent brun-bleue. C-Microscopie électronique à transmission : mise en évidence d'inclusions paracrystallines au niveau de mitochondries située dans les espaces intermyofibrillaires

3.4. Investigations enzymologiques

L'étude de l'activité des différents complexes enzymatiques de la chaîne respiratoire est possible à partir de fragments de tissu de ou de cellules en culture, dans des milieux empêchant la contre-sélection des cellules déficitaires (ex : présence d'uridine et de pyruvate). Les explorations peuvent ainsi définir l'existence de déficits isolés ou combinés des complexes.

Du fait de l'hétérogénéité des tissus et organes atteints, il est toujours préférable d'étudier le ou les tissus atteints quand cela est possible. Ainsi, on préférera la biopsie musculaire dans les présentations myopathiques et les lymphocytes dans les tableaux hématopoiétiques (syndrome de Pearson). Des atteintes s'exprimant de façon prédominante dans le foie ou le cœur, pourront être étudiées sur des biopsies hépatiques à l'aiguille ou des biopsies endomyocardiques, quand celles-ci sont réalisables. Quel que soit l'organe atteint, il est toujours utile de prélever une biopsie de peau des patients (même en post-mortem immédiat) pour de futures investigations enzymologiques ou moléculaires sur fibroblastes en culture.

Deux techniques peuvent être utilisées : la polarographie et la spectrophotométrie (RUSTIN *et al.* 1994).

3.4.1. Polarographie

Les études polarographiques permettent de mesurer la consommation d'oxygène par des fractions enrichies en mitochondries à l'aide d'une électrode de Clark (électrode à oxygène), en présence de différents substrats (malate + pyruvate, malate + glutamate, succinate, palmitate...). Un déficit du complexe I se traduira par une oxydation diminuée du NADH, tandis que l'oxydation de substrats produisant du FADH₂ (succinate) sera normale. La situation opposée s'observera dans le cas d'un déficit du complexe II, alors qu'un déficit du complexe III ou du complexe IV affectera l'oxydation de tous les substrats. De façon similaire, dans le cas d'un déficit du complexe V, l'oxydation de tous les substrats, en présence d'ADP, sera déficitaire. Cependant dans ce dernier cas, l'addition d'un agent découplant tel que le cyanure de carbonyle m-chloro-phénylhydrazone stimulera fortement les oxydations indiquant que la phosphorylation de l'ADP en ATP est l'étape limitative.

Ces études requéraient, il y a encore quelques années, des quantités de muscle de l'ordre du gramme (environ 2 g), mais des miniaturisations permettent maintenant d'obtenir des préparations enrichies en mitochondries à partir de petites quantités de tissu (100-200 mg), ce qui intéresse chez les très jeunes enfants. Par contre, cette technique n'est réalisable que sur du tissu frais non congelé, l'autre

alternative étant l'étude de cellules: lymphocytes circulants (isolés à partir de 10 ml de sang) ou cellules en culture perméabilisées (lignées lymphoblastoïdes, fibroblastes)

3.4.2. Spectrophotométrie

Les études spectrophotométriques permettent de mesurer les activités des complexes de la chaîne respiratoire seuls ou par groupe, en utilisant des donneurs ou des accepteurs d'électrons spécifiques. Dans ce cas, il n'est pas nécessaire d'isoler les mitochondries. En conséquence, la quantité de matériel nécessaire est beaucoup moins importante (10-20 mg) et peut être obtenue par biopsie à l'aiguille (foie, rein) ou même par biopsie endomyocardique. La spectrophotométrie peut être réalisée sur du tissu frais ou bien sur du tissu immédiatement congelé, conservé à -80°C ou dans l'azote liquide.

L'interprétation des résultats se fait sur les rapports des valeurs d'activité des différents complexes entre elles, qui sont constantes en toutes circonstances et constituent un bon indicateur de l'activité enzymatique à chaque étape du transfert des électrons. Les valeurs absolues sont par contre moins fiables car peuvent varier entre les tissus. Cependant, comme tout ce qui concerne la mitochondrie, l'interprétation des résultats n'est pas si aisée. Une activité normale de la chaîne respiratoire ne permet pas totalement d'exclure un déficit, même si le tissu étudié est celui qui exprime la maladie. Il est possible d'être en présence d'une anomalie touchant les propriétés cinétiques d'une enzyme (non détectables par les techniques utilisées) ou d'une hétérogénéité tissulaire.

Il d'autant plus difficile d'interpréter ces résultats qu'il existe des déficits secondaires de la mitochondrie. En effet, d'autres déficits génétiques tel que la maladie de Wilson ou le déficit de la bêta-oxydation des acides gras, le vieillissement physiologique et les traitements pharmacologiques (comme par exemple les antirétroviraux) peuvent être à l'origine d'une diminution des activités des complexes de la chaîne respiratoire en plus des anomalies histologiques mitochondriales.

3.5. Détection des complexes I à V par électrophorèse Blue Native (BN-PAGE)

Il s'agit d'une technique d'électrophorèse non dénaturante développée par SCHAGGER et VON JAGOW en 1991 (SCHAGGER and VON JAGOW 1991), pour l'analyse de complexes protéiques membranaires, tels que le système des complexes de la chaîne respiratoire. Cette technique permet de maintenir les complexes dans leur état natif, biochimiquement actif, sans dénaturer les protéines ni dissocier les complexes, permettant ainsi d'étudier leur organisation structurale. Afin de conserver

l'état natif des complexes protéiques, le Serva Blue G est utilisé (aussi appelé Bleu de Coomassie G250) dans les tampons d'extraction et de migration électrophorétique. Le Serva Blue G permet, tout comme le SDS (sodium dodecyl sulfate) utilisé dans l'électrophorèse conventionnelle, d'obtenir des complexes protéiques chargés, ce qui est nécessaire à leur mobilité dans le champ électrique. À la suite de la migration, les bandes de protéines obtenues sont transférées du gel de polyacrylamide sur une membrane de nitrocellulose et révélées à l'aide d'anticorps spécifiques des complexes (technique Western Blot). L'intensité des bandes est mesurée par densitométrie et les résultats sont exprimés sous forme de ratio en les rapportant au complexe II. Le complexe II, entièrement codé par l'ADN nucléaire (non affecté par des mutations à l'ADNmt), est un bon marqueur de l'abondance des mitochondries (DiMAURO 2013). Cette technique permet d'analyser individuellement la taille, l'assemblage et l'abondance relative des complexes.

3.6. Complémentation fonctionnelle

Faire la part entre l'origine nucléaire et mitochondriale du déficit n'est pas aisé. L'utilisation de lignées cellulaires commerciales « cybrides » peut être utile pour distinguer les deux types de déficits. Les cybrides ou hybrides cytoplasmiques sont des cellules chimères, construites à partir de la fusion de cellules énuclées avec des cellules déplétées en ADNmt (cellule hôte appelée rho-zero)). Ainsi, il est possible d'introduire l'ADNmt à étudier (du patient) dans la cellule, si ce dernier affecte le fonctionnement mitochondrial, on s'orientera plutôt vers une origine mitochondriale et inversement pour les déficits nucléaires (VITHAYATHIL *et al.* 2012).

Concernant les déficits d'origine nucléaire, il est possible de préciser le chromosome responsable du déficit par des études de complémentations fonctionnelles, grâce au transfert d'un chromosome. Le principe de cette technique est d'introduire, dans des cellules de patients mises en culture, en l'occurrence des fibroblastes, les différents chromosomes humains et d'observer les conséquences de ce transfert sur le phénotype des cellules (DE LONLAY *et al.* 2002). Les chromosomes humains sont obtenus par fragmentation de cellules hybrides homme-rongeur stables qui ne contiennent qu'un seul chromosome humain : les *microcells*. Après le transfert des microcells dans les cellules de patients, si le déficit biochimique est restauré, cela signifie que le chromosome introduit est celui qui porte le gène déficitaire.

4. Les déficits génétiques dans les cytopathies mitochondriales :

Comme déjà évoqué précédemment, les protéines mitochondriales ont une double origine génétique nucléaire et mitochondriale, avec une participation nucléaire majoritaire (plus de 99% des protéines sont d'origine nucléaire). Chez les mammifères, environ 1500 protéines sont requises afin d'assurer le

fonctionnement et/ou la structure des mitochondries. Chez l'Homme, environ 1000 gènes ont été identifiés comme étant potentiellement impliqués dans les différentes voies, fonctions et structures mitochondriales (LOPEZ *et al.* 2000). Différentes stratégies ont été utilisées pour l'identification de ces protéines, telles que les études de liaisons génétiques dans les familles informatives, les prédictions *in silico* des signaux protéiques d'adressage mitochondrial, les études de co-immunoprécipitation et plus récemment les techniques d'étude protéomique à haut débit, grâce à la spectrométrie de masse...etc

Tous les modes d'hérédité ont été observés dans les cytopathies mitochondriales, en fonction du gène impliqué : autosomique dominant, autosomique récessive et lié à l'X pour les gènes nucléaires, hérédité maternelle pour les gènes mitochondriaux.

A ce jour, plus de 200 gènes nucléaires ont été clairement impliqués dans la pathologie humaine mitochondriale et une cinquantaine sont étudiés en quasi-routine diagnostique, mais ce nombre est en augmentation constante (ZHU *et al.* 2009).

4.1. Le déficit des gènes nucléaires :

Les principaux déficits génétiques nucléaires impliqués dans les cytopathies mitochondriales sont résumés dans le tableau 1 de **l'Annexe 4**. Cette liste n'est pas exhaustive, elle représente les déficits affectant directement les complexes de la chaîne respiratoire (CHINNERY and HUDSON 2013). Il est difficile aujourd'hui d'obtenir une liste complète de tous les gènes rapportés dans les cytopathies mitochondriales.

4.1.1. Les gènes de structure et d'assemblage des complexes de la chaîne respiratoire (figure 20)

Des mutations pathogènes ont été décrites dans 13 parmi la quarantaine de gènes nucléaires de structure du complexe I et cette liste est en constante augmentation. Elles entraînent dans la majorité des cas un syndrome de Leigh ou un tableau neurologique, associé ou non à une cardiomyopathie (DISTELMAIER *et al.* 2009).

Des mutations dans le gène *SDHA*, sous-unité flavoprotéine du complexe II a été impliqués dans un tableau de syndrome de Leigh ou d'atrophie optique avec ataxie. Les sous unité *SDHB*, *SDHD* et *SDHC*, ont surtout été impliqués dans des formes dominantes de syndromes tumoraux notamment des phéochromocytomes, des paragangliomes et des Cowden-like syndromes. Parmi les gènes des 10 sous-unités nucléaires du complexe III, 2 (*UQCRCB* et *UQCRCQ*) ont été impliqués respectivement dans un tableau d'hypoglycémie avec acidose lactique et dans un phénotype neurologique avec retard psychomoteur sévère et signes extrapyramidaux. Seul un gène de structure du complexe IV (*COX6B1*), responsable d'une encéphalomyopathie sévère, a été identifié et, à notre connaissance, aucune mutation dans un gène nucléaire de structure du complexe V n'a pu être à ce jour, mise en évidence.

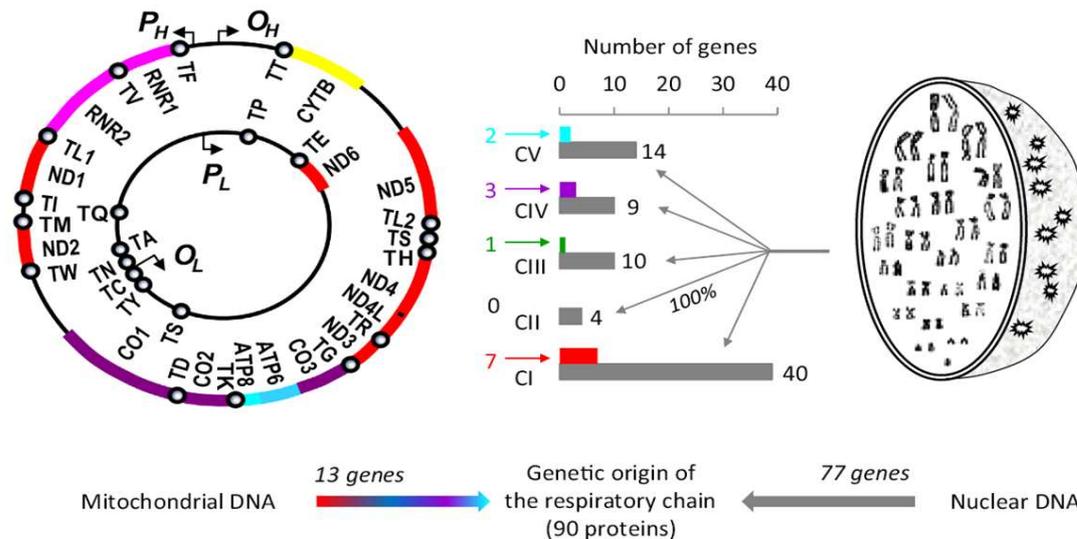


Figure 20 : participation nucléaire et mitochondriale à la formation des complexes de la chaîne respiratoire (TURNBULL and RUSTIN 2015)

Le nombre de gènes d'assemblage impliqués dans les cytopathies mitochondriales ne cesse de s'accroître. Ils ont été décrits en premier lieu pour les complexes I, III, IV puis V et plus récemment pour le complexe II avec le déficit du gène *SDHAF*, impliqué dans des cas de leucodystrophie (GHEZZI *et al.* 2009). Ils ont été initialement rapportés dans des présentations neurologiques isolées puis dans des atteintes multisystémiques (neurologiques, cardiaques, hépatiques et rénales).

4.1.2. Gènes impliqués dans la traduction des protéines mitochondriales :

Ceci concerne les gènes codant pour les protéines des mitoribosomes, les protéines impliquées dans la maturation des ARNt et ceux codant pour des facteurs d'élongation. Ils sont le plus souvent responsables de tableaux sévères d'encéphalopathie associée à une atteinte multi-systémique

4.1.3. Gènes impliqués dans la répllication et la maintenance de l'ADN mitochondrial

Le déficit de ces gènes est à l'origine des syndromes de déplétion de l'ADN mitochondrial qui correspond à une diminution de la quantité (nombre de copies) des molécules de l'ADN dans les mitochondries (généralement inférieur à 10%). Plusieurs tableaux cliniques ont été rapportés, s'exprimant en général dans la population pédiatrique : myopathies, encéphalo-myopathies ou atteintes hépato-cérébrales. Ils ont comme particularité de respecter l'activité du complexe II, seul complexe codé totalement par le génome nucléaire sans participation mitochondriale. Certains déficits

peuvent être à l'origine de délétions multiples par l'accumulation d'intermédiaires de réplication anormaux (POLG, Twinkle et TP) (COPELAND 2012). Deux catégories de protéines ont été impliquées : les protéines de la réplication et des protéines du maintien du pool de dNTP (figures 21 et 22)

Les protéines de la réplication :

- La polymérase gamma est constituée de deux sous-unités : l'unité catalytique POLG1 et l'unité accessoire POLG2. Les mutations dans *POLG1* ont d'abord été identifiées dans des ophtalmoplégies externes (PEO), souvent progressives familiales autosomiques ou récessives, classiquement associées à des délétions multiples dans le muscle. Son spectre s'est étendu ces dernières années au syndrome d'Alpers et à différentes atteintes neurologiques dont l'ataxie. Les mutations dans *POLG2* sont beaucoup plus rares et sont responsables d'ophtalmologie externe progressive familiale autosomique dominante.
- Twinkle, hélicase de la réplication est à l'origine de tableaux de PEO familiale dominante avec parfois des myopathies et des neuropathies. Les mutations récessives sont responsables de tableaux sévères d'encéphalo-hépathopathie précoces.

Les voies de régulation du pool de dNTP mitochondrial:

- La thymidine kinase 2 (TK2) est une kinase mitochondriale des nucléosides de type pyrimidine (thymidine T, cytosine C, uridine U), elle est donc indispensable à la synthèse des desoxyribonucléotide monophosphates (dCMP, dTMP et dUMP). Son déficit entraîne un tableau de myopathies de sévérité variable chez l'enfant et l'adulte.
- La Deoxyguanosine kinase (DGUOK) est la kinase mitochondriale des nucléotides de type purine (guanidine G et adenine A), elle est importante pour la synthèse des dAMP, et dGMP. Ses mutations sont à l'origine d'un tableau hépato-cérébral chez l'enfant
- La protéine p53-R2 est une petite sous -unité de la ribonucléase réductase, enzyme de réduction du ribose au niveau des ribonucléotide, permettant l'obtention de dNTP (désoxyribonucléotide triphosphate) directement à partir des NTP (nucléotide-triphosphate). Le déficit de cette enzyme a été rapporté dans des encéphalopathies néonatales associées des tubulopathies.
- La thymidine phosphorylase (TP), enzyme qui catalyse la déphosphorylation de la thymidine ou deoxyuridine en thymine ou uracile. Son déficit entraîne un déséquilibre dans le pool des dNTP vers l'accumulation de dTTP proportionnellement aux autres dNTP. Ses mutations sont à l'origine du syndrome MNGIE (Encéphalopathie myo-neuro-gastrointestinale).
- *SUCLA2* et *SUCLG1* sont des sous-unités de la Succinyl-CoA ligase (SUCL), enzyme du cycle de Krebs qui transforme le succinyl-CoA en présence d'ADP ou de GDP en succinate, synthétisant ainsi une molécule d'ATP ou de GTP. Les mutations dans *SUCLA2* et *SUCLG1*

sont associées à des encéphalomyopathies précoces. *SUCLG1* est responsable de tableaux plus sévères avec détresse métabolique néonatale et décès dans les premiers mois de vies. Les déficits en SUCL se caractérisent par une excrétion urinaire anormale d'acide méthylmalonique, dépistée sur la chromatographie organique urinaire. Celle-ci s'explique par l'accumulation de succinyl-CoA qui inhibe l'enzyme Méthylmalonyl-CoA mutase responsable de la conversion de méthylmalonyl-CoA en succinyl-CoA dans la voie du catabolisme des acides aminés (enzyme déficitaire dans les aciduries méthylmaloniques)

- La MPV17 est une protéine de la membrane mitochondriale interne dont la fonction est inconnue. Elle a été impliquée dans des formes hépatocérébrales de déplétion mitochondriale.
- L'ANT ou translocase mitochondriale ADP/ATP a été retrouvée mutée dans des formes dominantes d'ophtalmoplogie externe progressive
- Des mutations dans le gène *RRM2B* ont été associées à des délétions multiples de l'ADNmt avec un phénotype de PEO familiale dominante ou récessive.

Il est à noter que les déficit des protéines de la fusion mitochondriale (OPA1 et MFN2) sont également responsables d'instabilité du génome mitochondrial et peuvent être à l'origine de déplétion ou de délétions de l'ADNmt

Table 1. Nuclear genes associated with instability of mitochondrial DNA in patients.

Gene	Clinical disorder	Consequence of mtDNA	Chromosome locus	Function
POLG(POLG1)	Ar Alpers or ataxia Ar PEO / adPEO	Depletion Deletions	15q25	Mitochondrial DNA polymerase
POLG2	AdPEO		17q23-24	Pol γ accessory subunit
DNA helicase(C10orf2)	AdPEO Ar ataxia / encephalopathy	Deletions Depletion	10q24	Mitochondrial DNA helicase, Twinkle
TP(TYMP)	Ar MNGIE	Depletion Deletions	22q13.32	Thymidine phosphorylase
DGUOK	Ar encephalohepatopathy	Depletion	2p13	Mitochondrial deoxyguanosine kinase
TK2	Ar congenital myopathy	Depletion	16q22	Mitochondrial thymidine kinase
RRM2B	Ar encephalomyopathy Ad PEO	Depletion	8q23.1	p53-inducible small subunit of ribonucleotide reductase
ANT1(SLC25A4)	AdPEO	Deletions	4q34-35	Adenine nucleotide translocator
MPV17	Ar neurohepatopathy	Depletion	2p21-23	Mitochondrial inner membrane protein
SUCLA2	Ar encephalomyopathy	Depletion	13q12.2-q13.3	β -subunit, Succinate-CoA ligase
SUCLG1	Ar encephalomyopathy	Depletion	2p11.1	α -subunit, Succinate-CoA ligase
OPA1	Ad optic atrophy	Deletions	3q28-q29	Dynamin related GTPase

Ad, autosomal dominant; Ar, autosomal recessive.

Figure 21 : les déficits génétiques du maintien de l'ADN mitochondrial (COPELAND 2012)

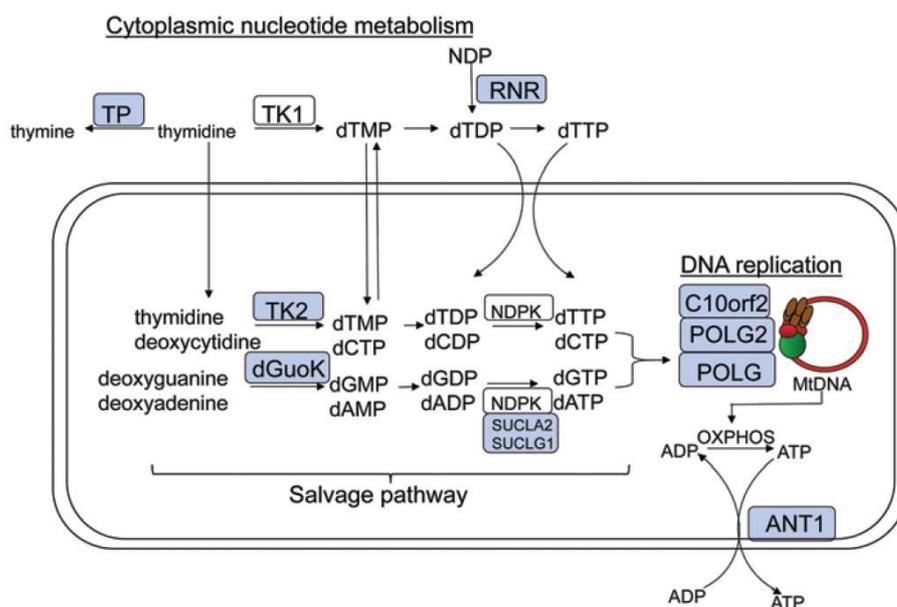


Figure 22 : représentation schématique des rôles des protéines impliquées dans les voies de maintenance et de stabilité du génome mitochondrial (COPELAND 2012)

4.1.4. Autres déficits d'origine nucléaire :

D'autres protéines de localisations mitochondriales ont été identifiées et impliquées dans des atteintes mitochondriales avec des tableaux variables :

- Le syndrome de Barth : est dû à des mutations dans le gène *TAZ* codant pour la Taffazin,, enzyme de la biosynthèse des cardiolipines, il se caractérise par principalement par une cardiomyopathie (non compaction ventriculaire), une atteinte musculaire, une neutropénie, un retard de croissance et une acidurie 3-méthylglutaconique (SCHLAME and REN 2006).
- L'ataxie de Friedrich : ataxie d'origine génétique la plus fréquente avec cardiomyopathie hypertrophique, diabète et neuropathie. Elle est causée par une expansion homozygote de triplets dans le premier intron du gène codant pour la frataxine. Cette protéine possède une localisation mitochondriale et a été impliquée dans la voie de biosynthèse des complexes fer-souffre (DELATYCKI *et al.* 2000).
- Plusieurs formes de paraplégies spastiques héréditaires (SPG) ont pour origine un déficit génétique de protéines mitochondriales, parmi lesquelles, la paraplegin dans SPG 7 dont les mutations ont également été associées à des formes d'ophtalmoplégie externe et d'atrophie optique (CASARI *et al.* 1998).
- Déficit primaire de la biosynthèse du coenzyme Q10 (QUINZII and HIRANO 2011) : déficit décrit plus récemment il a été associé à :

- ✓ Un tableau des encephalomyopathies, associant intolérance à l'exercice, une épilepsie ainsi qu'une ataxie
- ✓ Une encéphalopathie avec tubulopathie
- ✓ Une rhabdomyolyse avec une atteinte musculaire isolée
- ✓ Une encéphalopathie avec anomalies des ganglions de la base

4.2. Déficit des gènes mitochondriaux

4.2.1. Généralités :

Les premières mutations de l'ADNmt ont été décrites en 1988 quand l'équipe de Wallace (WALLACE *et al.* 1988) rapportèrent une mutation ponctuelle (m.11778G>A) dans le gène *ND4*, dans une famille atteinte d'atrophie optique de Leber. L'équipe de Holt (HOLT *et al.* 1988) a décrit la même année, une grande délétion en lien avec un cas de myopathie mitochondriale. Depuis, la liste des mutations pathogènes du génome mitochondrial ne cesse de s'allonger (DIMAURO and SCHON 2001), avec plus de 100 mutations ponctuelles plus au moins fréquentes décrites à ce jour (figure 23). La pathogénicité d'au moins une soixantaine de mutations n'est plus à démontrer, pour les autres elles ont été rapportées dans plusieurs cas mais leur rôle exclusif ou non reste plus débattu. Le tableau 2 en **Annexe 4** résume les principales mutations décrites à ce jour.

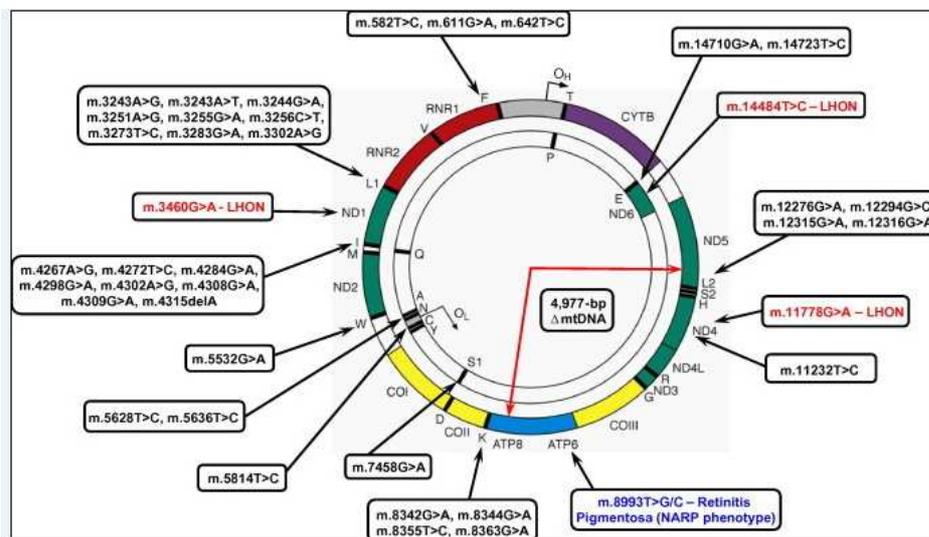


Figure 23 : carte génétique des principales mutations mitochondriales décrites (source mitomap)

Les mutations du génome mitochondrial (grands réarrangements ou mutations ponctuelles) expliquent à elles seules environ 10 à 20 % des cas de cytopathies mitochondriales (DARIN *et al.* 2001). Elles sont

impliquées, aussi bien dans des associations syndromiques bien identifiées (Atrophie optique de Leber, MELAS, MERRF, Leigh, NARP, Kearns-Sayre, Pearson...) que dans des atteintes isolées. En réalité, parmi cette longue liste de mutations, certaines sont beaucoup plus fréquentes dites récurrentes dans des gènes qui codent pour des protéines ou des ARNt. Elles constituent de véritables « hot spots » mutationnels et sont souvent recherchées en première intention : la mutation m.3243A>G dans l'ARNtleu (UUR) dans les tableaux de MELAS ou de diabète-surdité à hérédité maternelle, la mutation m.8344A>G dans l'ARNt leu pour le syndrome de MERRF, les mutations dans les gènes codants pour des protéines du complexe I m.11778G>A (*MT-ND4*), m.3460A>G (*MT-ND1*), m.14484TT>C (*MT-ND6*) dans l'atrophie optique de Leber, et la mutation m.8393T>G (*ATPase6*) dans le syndrome NARP.

Les cytopathies mitochondriales liées à des mutations dans des ARNt sont souvent associées à une prolifération mitochondriale compensatrice, se manifestant sous forme d'une accumulation importante des fibres rouges déshiquetées (RRF), avec un marquage en histoenzymologie positif pour la succinate déshydrogénase (complexe II) mais faible voir négatif pour la cytochrome c oxydase (COX). Ceci est probablement expliqué par l'origine uniquement nucléaire du complexe II, contrairement à la COX. La seule exception est le syndrome de MELAS par mutation m.3243G>A où le marquage COX est positif (c'est le paradoxe du MELAS) qui pourrait être dû à un effet seuil élevé de la mutation dans le muscle, le taux de la mutation ne suffirait donc pas pour diminuer l'activité COX musculaire (PETRUZZELLA *et al.* 1994).

4.2.2. L'Hérédité mitochondriale

Comme expliqué précédemment, seul l'ADN mitochondrial de l'ovocyte (maternel) est transmis à la descendance. Les mitochondries d'origine paternelle sont détruites dès les premiers stades du développement embryonnaire. Ceci implique que seules les mères peuvent transmettre les mutations du génome mitochondrial à leurs enfants, c'est l'hérédité mitochondriale ou cytoplasmique.

4.2.3. La notion d'hétéro/homoplasmie

La notion d'hétéroplasmie/homoplasmie est une particularité propre au génome mitochondrial lié au fait que l'ADNmt est présent en de nombreuses copies dans une même cellule. On parlera d'homoplasmie si un seul type de molécule d'ADN mt (muté ou non) est présent, dans ce cas-là une seule population d'ADN mt est représentée, et inversement si les deux populations moléculaires (sauvage et muté) coexistent, on parlera d'hétéroplasmie. Au cours des divisions mitotiques, les mitochondries ont tendance à se répartir au hasard dans les cellules filles (ségrégation mitotique), en fonction de la répartition on aura pour un même individu des taux d'hétéroplasmie variables dans les différents tissus (DOWLING 2014).

4.2.4. Physiopathologie et théorie du goulot d'étranglement mitochondrial

La physiopathologie des cytopathies mitochondriales et notamment celles liées à des mutations de l'ADNmt est loin d'être claire. Leur extrême hétérogénéité phénotypique n'est pas élucidée non plus. Néanmoins, un début d'explication a émergé ces dernières années grâce à des travaux d'observation de lignées maternelles mitochondriales, sur plusieurs générations chez des mammifères. Ces travaux ont montré que la fréquence de certains variants mitochondriaux avait tendance à augmenter au fil du temps aux dépens d'autres variants. Ainsi, et par analogie avec la dérive génétique de populations d'individus, ceci aboutirait à la fixation d'une des deux populations moléculaires par un effet fondateur. Cette théorie, appelée « la théorie du goulot d'étranglement mitochondrial », pourrait expliquer en moins en partie l'hétérogénéité des cytopathies mitochondriales. En effet, à l'échelle de l'organisme il pourrait y avoir une dérive locale dans une région tissulaire ou un organe, qui fixerait au hasard l'espèce moléculaire mutée ou sauvage. Chaque tissu exprimerait ou pas le déficit en fonction du taux de mutations présentes. A cela se rajouterait un effet seuil à partir duquel le déficit de la chaîne respiratoire parlerait, un seuil qui serait variable selon les tissus (DOWLING 2014). Il semble de surcroît, que ce phénomène de « goulot d'étranglement » se produise aussi pendant l'ovogenèse où seule une petite quantité d'ADNmt maternel passe dans les ovocytes primaires. Ainsi, au moment de la transmission, seulement quelques molécules maternelles d'ADNmt sont représentées chez le fœtus. Ceci pourrait être à l'origine de la variabilité des taux d'hétéroplasmie transmis qui dépend de l'échantillonnage mitochondrial de départ et la fixation d'une espèce moléculaire aux dépens d'une autre.

4.2.5. Les syndromes associés à des délétions du génome mitochondrial

Le génome mitochondrial peut être le siège fréquent de grands réarrangements et notamment de grandes délétions médiées par des régions répétées. De nombreuses délétions uniques récurrentes ont été rapportées au sein du génome mitochondrial, leurs tailles varient de 1 à 10Kb voire plus. Contrairement aux délétions multiples celle-ci sont accidentelles et se produisent le plus souvent *de novo*. Les récurrences dans les fratries sont exceptionnelles, et les mères atteintes transmettent rarement les délétions (1/24), en cela le conseil génétique peut être rassurant.

Les délétions semblent se produire de manière accidentelle dans les ovocytes ou dans les cellules embryonnaires puis se répartissent aléatoirement dans les différents tissus et organes. Certaines équipes ont montré que des délétions sont présentes naturellement dans 0,1% des molécules d'ADNmt des cellules germinales (CHEN *et al.* 1995), ainsi comme une minorité seulement est transmises, les risques pour qu'une molécule délétée passe dans l'embryon sont faibles.

L'expression clinique va dépendre de la répartition des molécules délétées pendant le développement embryonnaire. Si les molécules délétées sont présentes dans les trois feuillets embryonnaires (ecto, méso et endoderme), le déficit concernera tous les types de tissus : neurologique, musculaire et viscéral et sera plutôt à l'origine d'un syndrome de Kearns-Sayre. Si en revanche, la délétion prédomine dans les tissus hématopoïétiques, elle sera à l'origine d'un syndrome de Pearson et de la même façon si le déficit concerne le tissu musculaire de manière plus isolée, celui-ci sera plutôt responsable d'un phénotype d'ophtalmoplégie externe progressive. Ces trois syndromes font en réalité partie d'un continuum phénotypique avec parfois des atteintes chevauchantes. Il ne semble pas y avoir de corrélations entre la taille de la délétion et l'expression clinique, seule la répartition tissulaire de la délétion influence le phénotype.

Il est important de noter également que ces mêmes atteintes peuvent être causées par des insertions/duplications du génome mitochondrial qui sont beaucoup plus rares mais sont transmissibles d'une génération à une autre (POULTON *et al.* 1989).

4.2.5.1. Les syndrome Kearns-Sayre (KSS)

Décrit pour la première fois en 1965 (KEARNS 1965), il se définit par une triade clinique : rétinite pigmentaire, d'ophtalmoplégie externe progressive et un début avant 20 ans à laquelle se rajoute un de ces critères : 1) troubles de la conduction cardiaque, 2) ataxie cérébelleuse ou 3) hyperprotéïnorachie (>100 mg/dL). Il peut s'y associer également un retard de croissance staturopondéral avec ou sans déficit en hormone de croissance, un diabète, une hypoparathyroïdie, une surdité neurosensorielle, un déficit cognitif de sévérité variable ou parfois, une évolution chez les adultes vers une démence précoce. L'atteinte du système nerveux central dans le syndrome de KSS se caractérise de manière inexplicable par la rareté des manifestations épileptiques et des pseudo-AVC, contrairement aux autres encéphalopathies mitochondriales. Sur le plan moléculaire, plus de 150 délétions différentes ont été rapportées dans le syndrome de KSS. Une délétion est particulièrement fréquente de 4977 pb (m.8470_13446del4977) (PHADKE *et al.* 2012).

4.2.5.2. Le syndrome de Pearson

Décrit pour la première fois par Pearson en 1979 (PEARSON *et al.* 1979), il se caractérise par une anémie sidéroblastique réfractaire, une vacuolisation des précurseurs de la moelle et une insuffisance pancréatique externe. Les manifestations hématologiques débutent dans la première enfance, plus exceptionnellement à la période néonatale. Elles se caractérisent par une anémie macrocytaire sidéroblastique associée parfois à une neutropénie ou à une thrombopénie. L'aspect du myélogramme est évocateur avec la présence de vacuoles dans les progéniteurs granuleux et érythroblastiques. La coloration de Perls révèle la présence de sidéroblastes en couronne. S'y associent une insuffisance

pancréatique externe par fibrose avec malabsorption et diarrhée, et une hyperlactatémie intermittente. D'autres atteintes d'organe sont également possibles, notamment rénale avec une tubulopathie et une amino-acidurie et hépatique avec une hépatomégalie, une cytolyse et une cholestase. Plus rarement, il peut exister une atteinte cardiaque ou neuromusculaire. En général, les enfants décèdent dans les trois premières années de vies et ceux qui survivent évoluent avec un phénotype semblable au syndrome de Kearns-Sayre, les manifestations hématologiques s'amendent d'elles-mêmes, alors qu'apparaissent ou se majorent les symptômes neurologiques et myopathiques.

4.2.5.3. Ophtalmoplégie externe progressive (PEO : progressive external ophtalmoplegia)

Il s'agit d'une myopathie mitochondriale, touchant de manière sélective les muscles oculaires. La paralysie est habituellement bilatérale, graduellement progressive d'un ou plusieurs muscles extra-oculaires. Elle commence par une diplopie et un ptosis puis évolue au stade terminal vers une paralysie complète des yeux, voire une mydriase et parfois une atrophie du nerf optique. Dans certains cas l'atteinte musculaire est plus généralisée avec une faiblesse proximale des membres appelée « PEO plus ». Il peut également exister tout un spectre phénotypique entre les deux atteintes KSS et PEO plus au moins variables. Il faut rappeler que des déficits nucléaires sont également à l'origine de PEO avec notamment des mutations dans le gène *POLG1*, autosomiques dominantes ou récessives.

4.2.6. Les principaux syndromes liés à des mutations ponctuelles du génome mitochondrial

Les mutations ponctuelles de l'ADNmt (syndromes MELAS, MERRF, NARP et l'atrophie optique de Leber) sont presque toujours à l'état hétéroplasmique à l'exception de l'atrophie optique de Leber, sont transmises selon un mode maternel et dans la grande majorité des cas sont retrouvées chez les mères et certains apparentés maternels des patients.

4.2.6.1. Le syndrome MELAS

Le syndrome MELAS (Mitochondrial Encephalomyopathy, Lactic Acidosis, and Stroke-like episodes) est une maladie multisystémique débutant habituellement dans l'enfance (PAVLAKIS *et al.* 1984). Le développement psychomoteur est initialement normal mais il existe souvent un retard de croissance. Les symptômes débutent souvent entre 2 et 10 ans par des crises généralisées tonico-cloniques, des céphalées récurrentes, une anorexie et des vomissements réguliers. L'intolérance à l'exercice ou la faiblesse musculaire des ceintures peuvent également être le symptôme initial. La succession de pseudo-AVC débutant souvent dès l'adolescence ou chez l'adulte jeune et récupérant partiellement est

responsable d'une altération progressive des capacités motrices, visuelles et cognitives. L'IRM cérébrale montre un hypersignal T2 non systématisé à un territoire vasculaire (touchant typiquement les régions occipitales). Des calcifications des NGC à la TDM peuvent être présentes. Un déficit auditif de perception est habituel. Les autres manifestations incluent un diabète sucré, une cardiomyopathie, des anomalies de conduction cardiaque, une rétinopathie pigmentaire, une ataxie cérébelleuse (EL-HATTAB *et al.* 2015).

Sur le plan biologique, le lactate est souvent élevé dans le sang et le LCR notamment après un exercice physique. Les protéines dans le LCR sont souvent supérieures à 1 g/L. L'analyse de la chaîne respiratoire montre habituellement des déficits partiels de plusieurs complexes, en particulier I et/ou IV mais elle peut également être normale.

Sur le plan moléculaire, la mutation récurrente m.3243A>G (*MT-TL1*) est identifiée dans 80% des cas, *MT-ND5* est le 2^{ème} gène mitochondrial impliqué dans le syndrome MELAS.

4.2.6.2. Le syndrome MERRF

Le syndrome MERRF (myoclonic epilepsy with ragged red fibers) (DIMAURO and HIRANO 1993) est dû dans plus de 80% des cas à la mutation m.8344A>G dans le gène *MT-TK*. C'est une atteinte multisystémique débutant dans l'enfance par des myoclonies et se manifestant par la suite par une épilepsie généralisée, une ataxie, une faiblesse musculaire et une démence. Les autres symptômes fréquemment retrouvés comprennent un déficit auditif, un retard statural, une atrophie optique, et une intolérance à l'exercice. Une cardiomyopathie avec des troubles de la conduction (Wolff-Parkinson-White), une rétinopathie pigmentaire, une paralysie des oculomoteurs, et une lipomatose touchent la région cervicale sont également présents dans moins de 50% des cas.

4.2.6.3. Les syndromes de Leigh et NARP

Ces deux syndromes sont probablement l'expression clinique d'un même spectre phénotypique avec des degrés de sévérité variable. Le syndrome de Leigh étant vraisemblablement la présentation la plus sévère du syndrome NARP (CARELLI *et al.* 2002). En effet, ces deux syndromes sont causés le plus souvent par une mutation récurrente du gène *MT-ATP6*, la m.8993T>G qui doit être recherchée en priorité, la substitution m.8993T>C a également été rapportée (CHAKRAPANI *et al.* 1998). Il semble y avoir une corrélation entre le taux d'hétéroplasmie et la sévérité clinique. Ainsi la mutation retrouvée à l'état homoplasmique est à l'origine du syndrome de Leigh, alors que des faibles taux d'hétéroplasmie sont plutôt responsables d'un syndrome NARP.

Sur le plan clinique, le syndrome NARP (neurogenic weakness with ataxia and retinitis pigmentosa) est caractérisé par une ataxie cérébelleuse et des troubles des apprentissages débutant dans l'enfance, une neuropathie périphérique sensitivo-motrice qui peut débuter plus tardivement à l'âge adulte, une ataxie et une atteinte rétinienne variable, allant d'un aspect poivre et sel à la maculopathie. Des anomalies des noyaux gris centraux, de l'électrorétinogramme et une atrophie optique peuvent aussi être mises en évidence. Une atrophie cérébrale et cérébelleuse peut être observée à l'IRM cérébrale. Les autres symptômes pouvant être retrouvés sont un retard statural, un déficit auditif de perception, une ophtalmoplégie externe progressive et des troubles de la conduction cardiaque.

Le syndrome de Leigh est caractérisé par un déclin neurologique plus précoce et sévère sur les plans moteur et cognitif de manière progressive au cours d'épisodes d'encéphalomyélopathie subaiguë nécrosante, débutant classiquement dans les 1^{ers} mois de vie au cours d'une infection virale. Les épisodes de décompensation s'accompagnent d'une élévation du lactate sanguin et/ou du LCR et des anomalies des NGC à type de lésions nécrotiques multifocales. Les manifestations associent retard psychomoteur voire régression, hypotonie, spasticité, mouvements anormaux dont la chorée, ataxie cérébelleuse et neuropathie périphérique. Une cardiomyopathie hypertrophique est souvent retrouvée. Le décès est souvent précoce dans les 3 premières années de vie. Une atteinte moins sévère est néanmoins constatée dans 25% des cas. La plupart des patients ont un déficit du complexe I ou IV de la chaîne respiratoire. La CAA plasmatique et la CAO urinaires peuvent montrer une élévation d'alanine et une lactacidurie secondaires. Une hypocitrullinémie peut être identifiée selon le type de mutation. Enfin, la spectro-IRM peut être utile pour confirmer une élévation du lactate.

4.2.6.4. La neuropathie optique héréditaire de Leber (LHON)

La LHON est caractérisée par une perte rapide et sans douleur de la vision centrale qui débute en général chez l'adulte jeune, l'âge moyen de survenue se situant entre 18 et 35 ans. C'est une maladie à expressivité variable et pénétrance incomplète affectant préférentiellement les hommes. Les manifestations cliniques comportent classiquement une phase aiguë de baisse de la vision généralement sévère (moins de 1/10), brutale et souvent unilatérale se bilatéralisant dans un second temps. La vision centrale est généralement profondément altérée avec un scotome centrocaecal et une préservation plus ou moins importante de la périphérie du champ visuel. La vision des couleurs est également très perturbée. Après six mois d'évolution, la phase chronique se caractérise par la dégénérescence des fibres ganglionnaires et à une atrophie optique globale non spécifique. Dans de plus rares cas, l'atrophie optique s'accompagne d'une atteinte du SNC avec une encéphalopathie, des troubles moteurs tels qu'une dystonie ou une ataxie... Ces formes complexes sont appelées « Leber plus ». L'atrophie optique de Leber est due à des mutations de l'ADNmt au niveau de gènes codant

pour les sous-unités du complexe I. Les trois mutations les plus fréquentes affectent les sites m.11778, m.3460 et m.14484 dans les gènes *MT-ND4*, *MT-ND1* et *MT-ND6* respectivement. Dans la plupart des cas, les mutations sont à l'état homoplasmique mais quelques cas d'hétéroplasmie ont également été retrouvés (MAN *et al.* 2003). Le taux d'hétéroplasmie semble influencer l'expression de la maladie puisque certaines études rapportent que le risque d'atteinte visuelle est réduit chez les patients ayant un taux correspondant à moins de 60% d'ADNmt mutés dans le sang (CHINNERY *et al.* 2001). Il existe également une corrélation génotype-phénotype, la mutation m.11778G>A donnant généralement le phénotype le plus sévère alors que la mutation m.14484 T>C est associée au meilleur pronostic, la mutation m.3460G>A donne un phénotype intermédiaire. Comme cité précédemment, l'atrophie optique de Leber peut aussi être causée par des mutations dans le gène nucléaire *OPA1*.

4.2.6.5. Mutations de l'ADNmt et surdité

Les mutations de l'ADNmt peuvent être à l'origine de surdités génétiques isolées ou syndromiques et représentent 1% des surdités génétiques isolées. Les mutations qui en sont responsables sont essentiellement situées sur le gène de l'ARNr 12S, dont deux fréquentes : m.1555A>G (PREZANT *et al.* 1993) et m.1494C>T. Ces deux mutations sont à l'origine de cas familiaux, ou plus rarement sporadiques de surdité isolée cochléaire ou retro-cochléaire sans atteinte vestibulaire. La sévérité phénotypique peut être très variable même au sein des membres d'une même famille, pouvant aller d'une surdité profonde et précoce à une simple prédisposition à l'ototoxicité des aminosides. Les aminosides sont une classe d'antibiotiques qui ont pour cible thérapeutique l'ARNr 30S des bactéries, ainsi les mutations de l'ARNr des mitochondries, descendant des bactéries par endosymbiose, créent des sites de sensibilité à l'action des aminosides et désorganisent la transcription mitochondriale. La toxicité prédominante au niveau de l'oreille interne peut être expliquée par le fait que cet antibiotique s'accumule préférentiellement dans la péri et l'endolymphe (HENLEY and SCHACHT 1988), aboutissant à la destruction irréversible des cellules ciliées de la cochlée par apoptose, la toxicité semble par contre être réversible au niveau vestibulaire (GUAN 2011) (figure 24). Les surdités syndromiques mitochondriales sont majoritairement représentées par l'association diabète-surdité d'hérédité maternelle. Le plus souvent due à la mutation m.3243G>A, variant allélique du syndrome MELAS. Une autre mutation particulière, la m.7445A>G (RNAtSer) est à l'origine de l'association surdité et hyperkératose palmoplantaire, diagnostic différentiel des mutations hétérozygotes du gène de la connexine 26 ou *GJB2*.

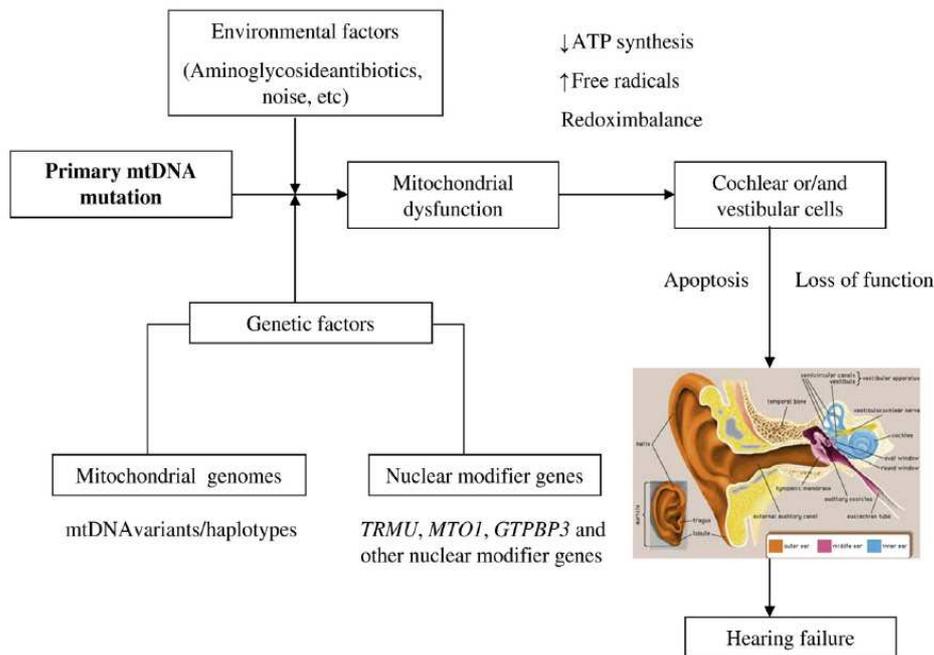


Figure 24 : schématisation des voies impliquées dans la toxicité mitochondriale des aminosides (GUAN 2011)

4.2.6.6. Mitochondries et diabète

Dans la plus part des cas le diabète mitochondrial est syndromique avec notamment l'association diabète-surdité liée à la mutation m.3243G>A. Néanmoins, de plus en plus de mutations causales de diabète plus au moins tardif et plus ou moins isolé sont rapportées (LI *et al.* 2014).

Les mécanismes physiopathologiques du diabète d'origine mitochondriale ne sont pas complètement expliqués. Plusieurs études ont cependant objectivé une insuffisance d'excrétion de l'insuline par les cellules β du pancréas, associée aux mutations mitochondriales. Les canaux K^+ responsable de la sécrétion pancréatique d'insuline sont stimulés par l'augmentation des rapports ATP/ADP cytoplasmiques, or le déficit mitochondrial pourrait conduire à une diminution de production d'ATP à l'origine d'une baisse des rapports ATP/ADP et une inhibition de ces canaux (figure 25).

Possible Metabolic Basis of Diabetes Mellitus

<http://www.mitomap.org/MITOMAP/DiabetesMitoMetab.pdf>
Copyright 2002 @ Mitomap.org

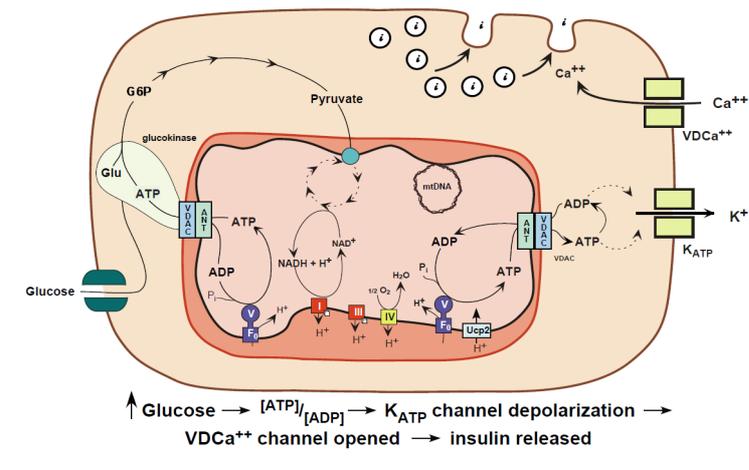


Figure 25 : schématisation des voies impliquées dans le diabète d'origine mitochondrial (source : [Mitomap](http://www.mitomap.org))

5. Stratégies diagnostiques des cytopathies mitochondriales

Devant cette grande hétérogénéité, la démarche diagnostique des cytopathies mitochondriales est complexe (cf. [Annexe 5](#)), elle se base sur la présence de faisceaux d'arguments cliniques et paracliniques : radiologiques, métaboliques, histologiques et enzymologiques. En dehors de quelques présentations spécifiques (ex : Leber, Pearson), l'orientation du diagnostic moléculaire n'est pas simple. Ce dernier repose classiquement sur l'étude de l'ADNmt à la recherche de grands réarrangements, de déplétion et de quelques mutations ponctuelles fréquentes. L'étude des gènes nucléaires arrive le plus souvent en deuxième intention, elle peut être orientée en fonction du ou des complexes déficitaires et de l'existence ou non d'instabilité de l'ADNmt. En effet, à cause de nombre important de gènes nucléaires impliqués et l'absence de corrélation génotype/phénotype, les méthodes classiques de criblage mutationnel ne permettent pas d'explorer la totalité des locus responsables de cytopathies mitochondriales. Le recours à des méthodes complémentaires complexes comme la complémentation fonctionnelle, le BN-Page et l'étude de liaison génétique demeurent nécessaire dans beaucoup de cas.

6. Prise en charge et traitement des maladies mitochondriales

Il n'y a actuellement pas de thérapie des maladies mitochondriales, et tous les essais thérapeutiques se sont révélés plutôt décevants. A l'exception des déficits primaires en coenzyme Q10 (CoQ) qui répondent bien à la substitution en CoQ, quand celle-ci est débutée précocement et à fortes doses, les

traitements spécifiques curatifs sont inexistants (KERR 2013). Il faut néanmoins souligner que le traitement par L-arginine dans les pseudo-AVC liés au MELAS semble apporter une amélioration clinique. Une administration intra-veineuse à la phase aigüe puis orale au long court semble normaliser les taux plasmatiques de L-arginine et diminuer de manière significative la fréquence et la gravité des « pseudo-strokes » par amélioration des fonctions endothéliales (KOGA *et al.* 2010). Par ailleurs, le régime cétogène a été recommandé dans les déficits en complexe I associés à une hyperlactacidémie, et ce dans l'hypothèse qu'un apport important en glucose ne peut être utilisé du fait du bloc enzymatique immédiatement en aval. Mais ceci n'a pas du tout fait la preuve de son efficacité et doit être discuté. Par ailleurs, ni les stratégies de substitutions enzymatiques ou les greffes allogéniques de tissus souches n'ont montré d'efficacité

Le reste de la conduite thérapeutique consiste en un traitement symptomatique avec une supplémentation vitaminique et en co-facteurs tels que l'acide folique dans les atteintes hématologiques. Certaines équipes ont testé des suppléments en molécules telles que la riboflavine ou le succinate sans réel effet. Aussi, la correction des déséquilibres métaboliques, en acides aminés ou des lactates est importante et l'utilisation de certains médicaments doit être prudente. Certaines molécules doivent être contre-indiquées comme le Valproate de sodium, l'aspirine, les anti-inflammatoires non stéroïdiens et les tétracyclines qui agissent sur la β -oxydation des acides gras ou par la séquestration du coenzyme A. L'utilisation des anesthésiants doit également être prudente. Enfin, les risques liés aux vaccinations ont également été discutés, notamment dans les formes neurologiques. Les vaccins pourraient, en effet, favoriser la survenue de poussées dans le syndrome de Leigh.

Le conseil génétique et le diagnostic anté-natal

La place du conseil génétique et du diagnostic pré-natal reste prépondérante dans la prise en charge des cytopathies mitochondriales. La transmission de la maladie peut être maternelle dans les cas de mutations de l'ADNmt, ou autosomique récessive le plus souvent, mais aussi dominante et liée à l'X dans les cas où le génome nucléaire est impliqué. La possibilité d'un diagnostic prénatal se présente uniquement pour des familles dans lesquelles le diagnostic de cytopathie mitochondriale a été formellement établi chez le cas index, soit par mise en évidence d'un déficit enzymatique soit par identification de la mutation en cause. La mesure des activités de la chaîne respiratoire est possible sur les villosités choriales entre 9 à 11 SA et sur les amniocytes en culture à 16 SA, celles-ci ne sont cependant pas toujours d'interprétation facile. De plus, ces analyses ne sont faisables que dans la mesure où les fibroblastes du cas index expriment le déficit, ce qui est loin d'être toujours le cas. Un résultat positif révèle une récurrence de la maladie. En revanche, un résultat normal n'exclut en aucun cas la possibilité d'une expression ultérieure du déficit au cours de la grossesse ou même en période postnatale. En effet, le mode d'expression des gènes de la chaîne respiratoire et la demande énergétique au cours du développement foetal sont encore inconnus. Un résultat normal ne peut être

complètement rassurant et demande impérativement un contrôle ultérieur sur amniocytes. Concernant les anomalies de l'ADNmt, les remaniements de grande taille sont le plus souvent sporadiques. Le conseil génétique peut donc être *a priori* rassurant, surtout si la mère n'a pas de remaniement de l'ADNmt. Il est cependant possible, dans le but essentiellement de rassurer la famille, de rechercher ce remaniement dans les villosités chorales du prochain foetus. En effet, comme exposé précédemment, les cas de transmissions maternelles de délétions de l'ADNmt sont extrêmement rares (1/24). Les mutations ponctuelles de l'ADNmt, transmises selon un mode maternel et dans la grande majorité des cas retrouvées chez les mères, sont presque toujours à l'état hétéroplasmique. Il est ainsi toujours très difficile de se prononcer dans ce type de DPN. En effet, la proportion de molécules d'ADNmt mutées dans les villosités chorales ne permet pas d'estimer sa proportion dans d'autres tissus fœtaux, ni son évolution au cours du développement embryonnaire. La présence de moins de 20% ou plus de 80% de molécules mutées chez le foetus est respectivement de bon et de mauvais pronostic. En revanche, des résultats intermédiaires sont extrêmement difficiles à interpréter.

Ce n'est que dans les cas où une mutation d'un gène nucléaire a été identifiée qu'il est possible de proposer un DPN moléculaire fiable, qui ne repose pas uniquement sur l'investigation enzymologique mal connue chez le foetus, et qui n'est pas soumis à l'incertitude de l'hétéroplasmie comme c'est le cas pour les mutations de l'ADNmt.

7. Techniques d'études moléculaires du génome mitochondrial

L'étude de l'ADN mitochondrial se faisait jusqu'à présent par les techniques de biologies moléculaires classiques, à la recherche de mutations ponctuelles et de grands réarrangements. La recherche de grandes délétions et de duplications est réalisée grâce à une amplification par PCR longues du génome mitochondrial, suivies par une migration lente sur gel d'Agarose et révélées aux UV. Ces anomalies sont également recherchées par techniques de Southern Blot, qui est notamment plus efficace pour objectiver les duplications. La quantification de l'ADNmt à la recherche de déplétion se fait quant à elle par PCR quantitative. Enfin, la recherche de mutations ponctuelles est plus difficile et habituellement seul le screening des mutations récurrentes, fréquentes est réalisé. Une des méthodes les plus utilisées reste la PCR avec études des polymorphismes de longueur des fragments de restriction (Restriction Fragment Length Polymorphism) au niveau de sites des mutations.

Lorsque les principales mutations ont été éliminées, une analyse plus complète de l'ADNmt peut être réalisée mais le séquençage systématique de tout le génome mitochondrial (16.5Kb) par méthode Sanger reste fastidieux. Cette méthode est aussi limitée dans la détection des mutations

hétéroplasmiques faibles, d'où l'émergence, ces dix dernières années de techniques spécifiques plus adaptées à l'étude du génome mitochondrial l'ADNmt :

- La méthode Surveyor ®(Transgenomic) : consiste en une première étape d'amplification par plusieurs PCRs chevauchantes de tout le génome mitochondrial. Les fragments obtenus sont ensuite dénaturés puis renaturés, permettant la formation d'hétéroduplexes qui sont ensuite détectés grâce à une digestion enzymatique par une endonucléase spécifique. Cette enzyme a la particularité de cliver spécifiquement au niveau des sites de mésappariement et donc d'objectiver l'existence de variants hétéroplasmiques (BANNWARTH *et al.* 2005).
- L'analyse par puce de re-séquençage Affymetrix® (GeneChip Mitochondrial Resequencing 2.0 Array) (MAITRA *et al.* 2004): il s'agit d'une puce à ADN contenant des sondes oligonucléotidiques (25pb) spécifiques du génome mitochondrial et plus exactement des régions codantes. Ainsi, pour chaque position nucléotidique 4 sondes avec les quatre versions de base possibles existent (A, T, G ou C), permettant après hybridation, de renseigner sur la nature de la base pour toutes les positions de la séquence et d'identifier ainsi les variants homoplasmiques. La détection des variants hétéroplasmiques étaient théoriquement possible mais moins efficace.

Au total, l'étude complète de l'ADNmt jusqu'à récemment était fastidieuse, longue et coûteuse, et n'était pas toujours réalisée en routine diagnostique. Les nouvelles techniques de séquençage dites à haut débit, générant une grande quantité de séquences en un temps court, peuvent apporter une grande aide dans le diagnostic des cytopathies mitochondriales, qu'elles soient liées à des mutations de l'ADN mitochondrial ou nucléaire. Cette technologie rend possible pour chaque patient, le séquençage rapide d'un grand nombre de locus impliqués dans les maladies mitochondriales (panel de gènes ciblés) mais aussi de séquencer la totalité du génome mitochondrial. De plus la génération de plusieurs lectures pour chaque position nucléotidique (ou profondeur) est particulièrement adaptée pour la détection des variations hétéroplasmiques.

III- Les techniques de séquençage de nouvelle génération

Pendant longtemps la méthode Sanger fut la technique de référence de séquençage d'ADN, elle a d'ailleurs permis le séquençage du génome humain en 2003. Cette technique décrite dans la seconde moitié des années 70 repose sur la synthèse enzymatique par complémentarité de bases, à partir d'un brin d'ADN grâce à une polymérase (SANGER *et al.* 1977). Cette élongation est interrompue de manière aléatoire par l'incorporation de di-désoxyribonucleotide (ddNTP) fluorescents (à la place de

dNTP). Ainsi, à partir d'un produit d'amplification, on obtient des fragments dont la taille varie d'une base pouvant être séparés par électrophorèse capillaire et la fluorescence détectée à chaque position permet d'identifier la nature du nucléotide. Cette technique reste, néanmoins longue et coûteuse, elle nécessite l'exploration des fragments d'ADN un par un, pour chaque patient et avec une taille de séquençage limitée par la longueur des capillaires qui ne dépasse en général pas les 800 nucléotides par fragment.

La mise au point d'un nouveau procédé de séquençage, a amorcé un retournement de tendance permettant un gain considérable en terme de coût et de temps. C'est le séquençage de nouvelle ou deuxième génération dont le principe est un séquençage massif en parallèle avec la lecture de millions de petits fragments en même temps. Il s'agit d'une méthode qui se base sur le suivi en temps réel de l'intégration des nucléotides grâce à un procédé de miniaturisation des réactions qui se fait sur puce à ADN ou (micro-array). Les premières techniques de séquençage massif en parallèle sont apparues vers la fin des années 1990, elles reposaient sur la méthode de pyroséquençage (RONAGHI *et al.* 1996). Depuis, plusieurs technologies ont émergé et ne cessent de progresser, elles restent cependant toutes assez proches dans leur fonctionnement et leurs résultats (Quail *et al.* 2012). Elles se basent globalement sur trois étapes : la préparation de la librairie, le séquençage en lui-même et l'analyse des données générées.

1. Les étapes de la techniques de NGS

1.1. La préparation de la librairie :

Le but de cette étape est d'obtenir des fragments identifiés représentant la région cible à séquencer. La taille médiane des fragments est variable et définie selon la chimie employée: 100, 200, 300 ou 400 bases.

1.1.1. Les techniques d'enrichissement :

Concerne le séquençage de régions ciblées du génome (une région donnée, un exome ou un panel de gènes...), le « whole genome sequencing » s'affranchit donc de cette étape. Plusieurs stratégies sont possibles (figure 26):

-L'amplification par PCR :

-PCR longues : elles sont suivies par une fragmentation par digestion enzymatique ou par sonication.

-Techniques de PCR multiplexe : elles permettent d'amplifier simultanément un grand nombre de petits fragments d'intérêt (*Multiplicom*, *AmpliSeq* (Life Technologies), *TruSeq Amplicons* -Illumina).

-La PCR par technique microfluidique (PCR en micro-gouttes) : elle permet la réalisation concomitante d'un très grand nombre de PCR simplex (dans des micro-gouttes) en parallèle dans un appareil dédié (*RainDance technologies®*).

- la capture par hybridation :

Cette méthode consiste en une première étape de fragmentation de l'ADN séquencer par sonication ou digestion, puis la capture des fragments d'intérêts par des sondes d'oligonucléotides ADN ou ARN simple brin spécifiques.

- l'enrichissement par circularisation ou hybridation à haute spécificité (*HaloPlex®* de Agilent): Il s'agit d'une méthode particulière consistant à digérer l'ADN génomique par plusieurs enzymes de restriction à extrémités cohésives et hybrider des sondes biotinylées double brin dont le centre d'environ 30pb est commun à toutes les sondes et les extrémités d'environ 20pb sont spécifiques des extrémités des régions ciblées entraînant la circularisation du fragment. Les ADN non circularisés sont éliminés à l'aide de streptavidine fixée sur des billes magnétiques puis de multiples PCR sont effectuées.

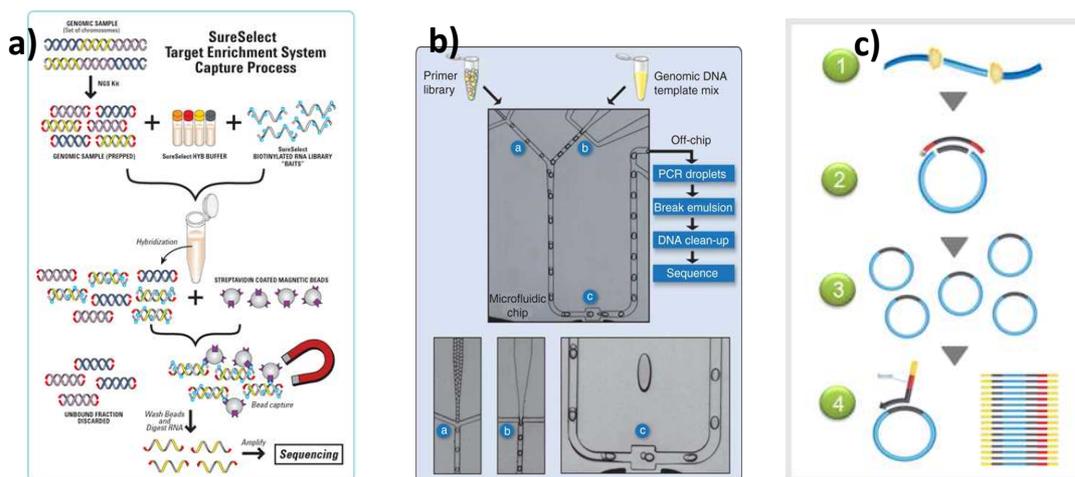
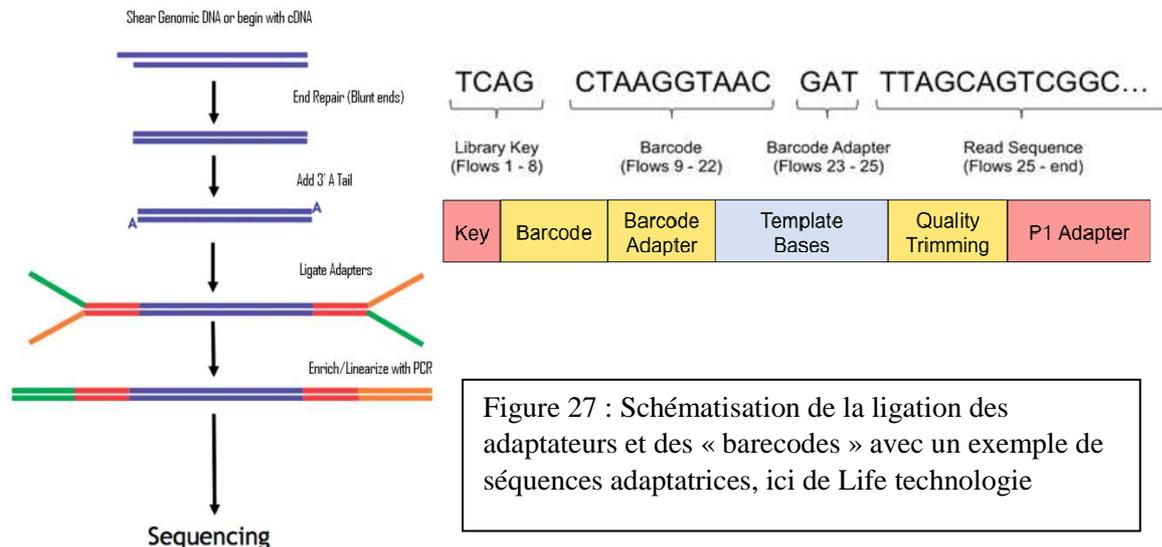


Figure 26 : schématisation des étapes des principales techniques d'enrichissement : dans le séquençage ciblé a) capture (SureSelect®) b) RainDance® c) HaloPlex® ou capture par circularisations

1.1.2. Le marquage des fragments par ligation des adaptateurs

Afin d'être reconnus et fixés sur les supports de séquençage, les fragments doivent être identifiés par addition de séquences nucléotidiques spécifiques de quelques paires de bases (figure 27), appelés adaptateurs. Les adaptateurs vont d'une part aider à la fixation des molécules sur les supports de séquençage et d'autre part servir d'amorces universelles pour le séquençage.

Il est également possible de rajouter à cette étape une autre séquence nucléotidique appelée index ou « barcodes » qui identifie les fragments appartenant à un même échantillon pour permettre d'en mélanger plusieurs dans un même « run » de séquençage.



1.1.3. Les techniques d'amplification clonale de la librairie :

Une fois les régions d'intérêt sélectionnées, une amplification clonale est nécessaire afin d'augmenter la quantité d'ADN et améliorer le recueil du signal pendant le séquençage. Deux techniques peuvent être utilisées en fonction des technologies :

- PCR sur support plan (bridge PCR sur flow cell) :

Les amorces spécifiques sont fixées sur une lame. Chaque molécule d'ADN va se fixer aux 2 amorces situées sur la lame en formant un pont et va être amplifiée de façon successive à un endroit donné de la lame formant un cluster de molécules d'ADN toutes identiques (Figure 28)

- PCR en émulsion sur billes :

Cette technique consiste en l'utilisation de microréacteurs de PCR situés dans des gouttelettes en émulsion dans lesquelles les molécules d'ADN se fixent sur les billes (une bille par gouttelette) grâce à des amorces. Les molécules sélectionnées sont incorporées dans les gouttelettes par loi d'action de masse avec des concentrations en billes et en ADN optimisées pour qu'une seule molécule d'ADN soit incorporée dans chaque microréacteur. Ceci permettra l'amplification clonale de cette molécule. (Figure 29)

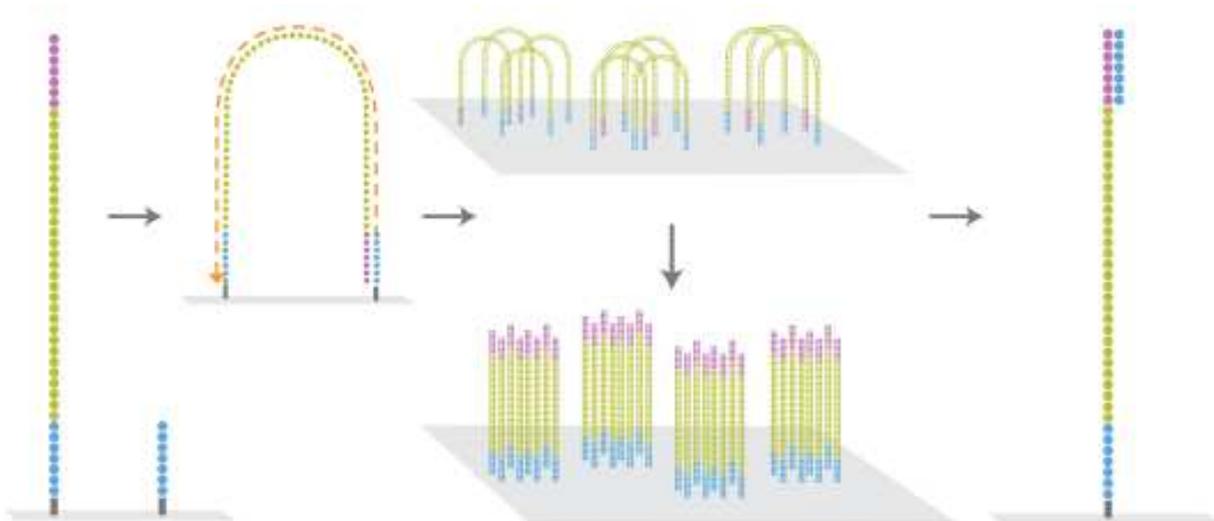


Figure 28 : Amplification clonale sur support solide (Bridge PCR)

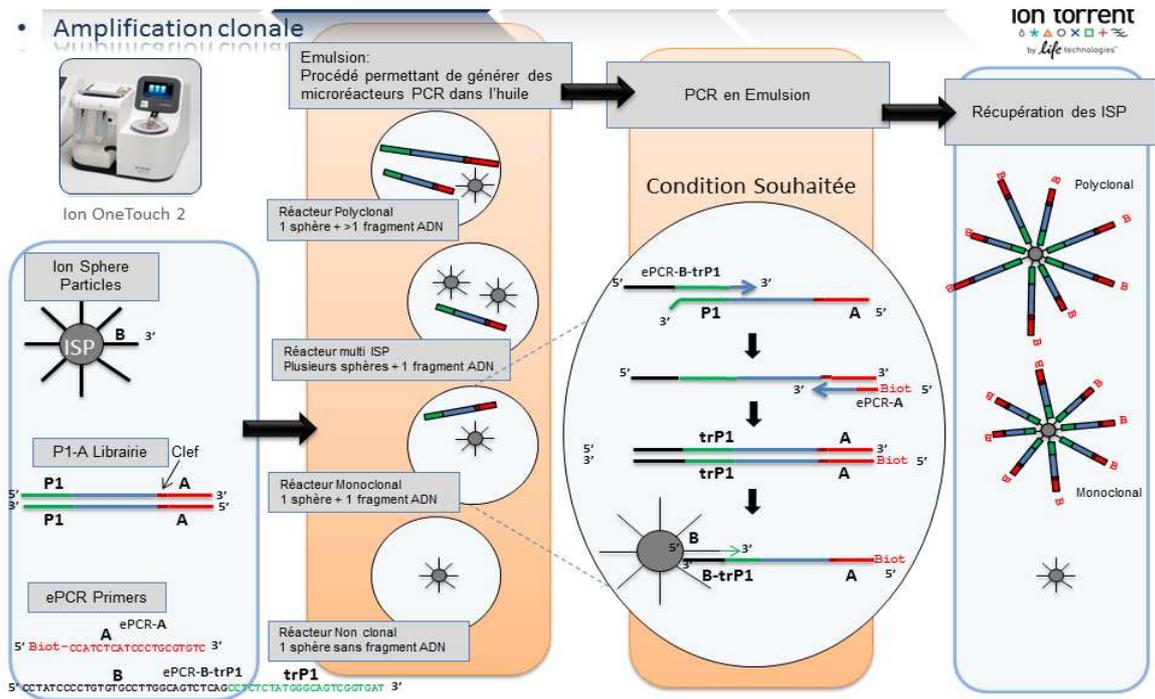


Figure 29 : Amplification clonale : PCR en émulsion sur billes

1.2. Les techniques de séquençage

1.2.1. Les technologies principales

Les principaux appareils et technologies de séquençage des quinze dernières années sont résumés dans les figures 30 et 31)

- le séquençage par synthèse :

La technique de pyroséquençage (de 454 Life Sciences/Roche®) a été la première à être commercialisée dès la fin des années 1990 (RONAGHI *et al.* 1996). La réaction de séquençage fait suite à une PCR en émulsion, le produit d'amplification clonal obtenu est déposé sur une plaque où chaque emplacement (puit) ne peut contenir qu'une seule molécule d'ADN amplifiée. La réaction de séquençage par synthèse est alors initiée base par base. La lecture de chaque base incorporée est révélée à l'aide d'une réaction chemo-luminescente et détectée par une caméra. Des cycles successifs d'ajout d'un dNTP puis lavage permettent de connaître la séquence d'incorporation dans chaque puit. L'autre méthode reposant sur le séquençage par synthèse est la technologie Ion Torrent®, basée sur l'utilisation de puces semi-conductrices remplies de puits (Ion Torrent semiconductor sequencing) de Life Technologies (RUSK 2011). Comme la technique précédente, un seul type de dNTP est ajouté dans le milieu à chaque cycle dans un ordre bien établi puis lavé (par cycle de 32 nucléotides TACGTACGTCTGAGCATCGATCGATGTACAGC), mais cette fois-ci le signal détecté n'est pas une fluorescence mais une variation de pH. En effet, quand l'incorporation du dNTP est réalisée par la polymérase, un proton est libéré par la réaction et le changement de pH local est mesuré par des capteurs. La variation de pH est proportionnelle au nombre de dNTP incorporés.

-*Séquençage par ligation réversible* (SOLID = *Sequencing by Oligonucleotide Ligation and Detection* de Applied Biosystems/ Life Technologies ®) :

Par cette technique le séquençage se fait également sur billes qui sont fixées sur un support solide. Le processus consiste à ajouter par ligase et de manière répétitive, des sondes de 8 bases contenant un fluorochrome et où les deux premières sont bien identifiées (toutes les combinaisons AA, AT, AG.. sont utilisées soit $2^4=16$) suivies de 6 bases dégénérées ou universelles (pouvant se fixer sur n'importe quelle séquence nucléotidique, comme la deoxyinosine, une base G sans son amine). Quatre fluorochromes sont utilisables pour les 16 types de sondes, ce qui veut dire qu'à chaque fluorescence va correspondre 4 sondes possibles. Ainsi, la fluorescence détectée pendant le séquençage ne fournit qu'une information partielle sur la séquence des deux bases incorporées. L'astuce consiste à répéter les cycles successifs d'addition des sondes en décalant à chaque fois d'une base, grâce à des amorces différentes en taille (d'une seule base). Cela permet par recoupement d'identifier les bases de l'ADN et qui plus est, de les mesurer chacune deux fois. Grâce à cette méthode, il est possible de lire jusqu'à

un milliard et demi de séquences en parallèle de 75 bases de long. Le système incorporé de correction d'erreurs associé à l'utilisation de la ligase rend cette technologie très fiable.

-Séquençage à l'aide de terminateurs réversibles (ou Dye reversible d'Illumina®/Solexa)

Cette technique a été développée par S.Balasubramanian and D.Klenerman co-fondateurs de la compagnie Solexa (appartenance aujourd'hui à Illumina) au début des années 2000 et a été publiée en 2008 (BENTLEY *et al.* 2008).

Pour cette méthode, l'amplification de l'échantillon à analyser ne s'effectue pas en solution mais sur un support solide. La réaction de séquençage est alors réalisée directement sur le support où l'ADN a été amplifié. Elle se déroule position après position, en ajoutant un mélange contenant les quatre (A, T, C, G) bases, associée chacune à un fluorophore différent. L'extrémité de ces bases est protégée pour empêcher l'addition de bases supplémentaires à chaque cycle d'incorporation. Une lecture laser permet alors de détecter simultanément toutes les positions incorporées. Le clivage des fluorophores permet ensuite l'incorporation de la base suivante, la lecture est effectuée ainsi cycle après cycle. Cette méthode permet l'acquisition en parallèle de plus de 3 milliards de séquences de 100 bases de long. Chaque position étant lue l'une après l'autre, les erreurs principales de cette technologie sont des erreurs de substitution d'une base par une autre.

	454 FLX/Roche	Solexa/Illumina	Solid/Life Tech
Préparation de banques de NGS	<p>ADN génomique → Fragmentation → Ligation d'adaptateurs</p> <p>But ajout d'adaptateurs : Utiliser toujours la même amorce pour PCR et Séquençage quelque soit l'ADN à séquençer</p>		
Amplification clonale	PCR en émulsion sur une bille Amplification clonale dans des microréacteurs	« Bridge » PCR sur un support plan (=Flow Cell) Colonies d'ADN = polonies	PCR en émulsion sur une bille Amplification clonale dans des microréacteurs
Système miniature d'amplification			
Support pour le séquençage	Billes insérées dans des micropuits 	Flow Cell ayant servi à l'amplification 	Billes fixées à une lame de verre
Séquençage en temps réel : 1 bille ou 1 cluster → 1 Read	*Séquençage par synthèse *Pyroséquençage *Longs fragments	*Séquençage par synthèse *dye-terminator réversible *petits fragments	*Séquençage réversible par ligation *encodage de 2 bases *petits fragments

Figure 31 : Récapitulatif des étapes de la technique de séquençage : les plus gros fournisseurs du marché : les sociétés **Illumina** et **Life Technologies** et le plus ancien : **Roche**

1.2.2. Le séquençage single versus paired-end

La plupart des technologies offrent un séquençage dit unidirectionnel, c.à.d. que chaque molécule ou brin est séquençé une seule fois dans un seul sens, c'est le séquençage « single-end ». Mais il est également possible de réaliser un séquençage bidirectionnel, où chaque molécule est lue deux fois. Pour un même point ou « cluster », la molécule est séquençée dans un sens puis l'autre (forward et reverse), un peu à l'image de ce qui réalisé en méthode Sanger, c'est le séquençage « paired-end ». Dans ce cas-là, les deux « reads » générés pour une même molécule sont identifiées par paires, ce qui a pour avantage de faciliter et d'améliorer l'alignement (figure 32)

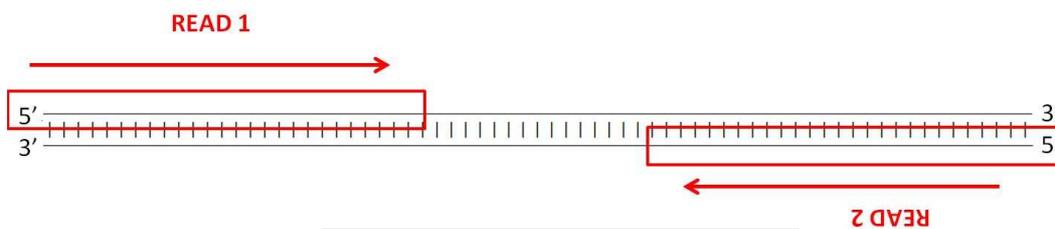


Figure 32: séquençage paired-end

1.2.3. Les différents appareils de séquençage

Les plus gros fournisseurs du marché actuellement, sont les sociétés, **Illumina** et **Life Technologies** qui n'ont cessé de faire évoluer leur gamme, tant sur le plan des équipements que sur le plan des capacités de séquençage. Plusieurs appareils existent avec des capacités variables (figure 33), le choix

de l'appareil et des technologies dépend de l'application souhaitée. La technologie de **Roche** était la première sur le marché mais elle n'est plus commercialisée.

Séquenceurs 2 ^{ème} génération											
Société	Roche			Illumina				Life Technologies			
Plateforme											
Technologie	Titanium	FLX Titanium	FLX +					Chip 314	Chip 316	Chip 318	
Acides nucléiques (matrice)											
Ligation adaptateurs											
Méthode d'amplification	PCR en émulsion			« Bridge PCR »				PCR en émulsion			
Durée de séquençage/run	10h	10h	20h	26h	8jrs	8jrs	14jrs	2h	12jrs	8jrs	8jrs
Capacité (Mb) séquençage/run	50	500	900	1500	100000	200000	95000	>10 >100 >1000	70000	80000	150000
Taille moyenne des reads	400	400	700	150+150	100+100	100+100	150+150	100 >100 >100	50+35	75+35	75+35
Coût (\$) /run	1100	6200		750	10000	20000	11500	500 750 950	8150	6100	10500
Coût machine + annexes ((k\$))	110+25	500+30		125	560	690	250	50+20	480+55	350+55	600+55
Exactitude de séquençage (%)	99	99		99,9	99,9	99,9	99,9	99	99,95	99,95	99,99

Figure 33 : récapitulatif des principaux appareils et technologies de séquençage

2. Le traitement des données

Le problème de ces nouvelles techniques est la génération de quantités importantes de données en peu de temps qu'il va falloir traiter. Ceci nécessite des outils et logiciels plus performants, adaptés aux traitements de données à haut débit. Les méthodes de traitement bio-informatique des données peuvent varier d'une technologie à l'autre mais se basent globalement sur les mêmes principes

2.1. Le « base calling »

Tout d'abord, le signal brut original qu'il soit luminescent ou électrique (variation de pH) est converti en séquences nucléotidiques c'est l'étape du « base calling ». La succession des nucléotides en chaque point de la plaque correspond à une lecture ou « read ».

A cette étape une évaluation de la qualité est également réalisée et les séquences vont subir un prétraitement qui consiste en :

- un filtre : c'est l'élimination des « reads » en leur totalité quand ils sont inutilisables car de trop petites tailles, de mauvaise qualité ou issus de signaux de billes polyclonales.
- un « trimming » : c'est élimination des séquences des adaptateurs et des « barecodes » ainsi que la suppression des bases qui ont un score qualité insuffisant .Le « trimming » se fait en général à partir de la partie 3'terminal de chaque « read » car l'efficacité du séquençage diminue au fur et à mesure des cycles, à cause d'une polymérase moins opérationnelle et de l'accumulation des bruits de fonds (figure 34)

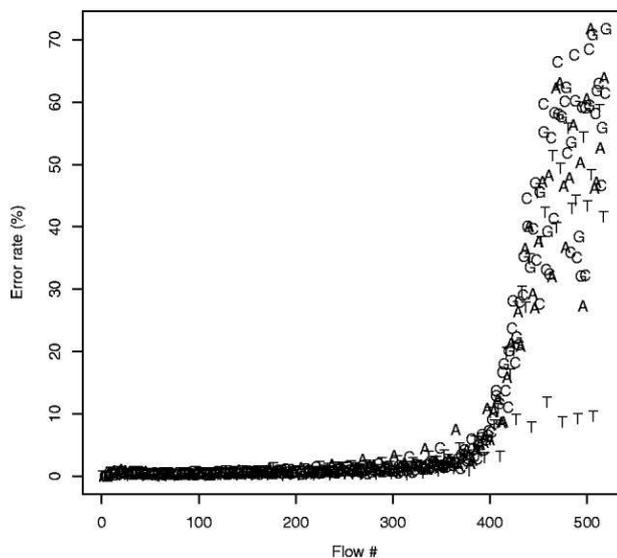


Figure 34 : Le taux d'erreur de séquençage au cours des cycles de séquençage sur avec la technologie Ion Torrent (PGM). Augmentation importante des erreurs après 380 cycles. (GOLAN and MEDVEDEV 2013)

Après le « base calling » et l'évaluation qualité, les séquences obtenues sont présentés sous la forme de fichier .FASTQ.

2.2. Les scores qualités et fichier FASTQ

2.2.1. Le fichier .FASTQ

Il utilise en principe 4 lignes par « read » la ligne 1 commence par un caractère "@" suivi de l'identifiant du « read » et éventuellement d'une description (l'identifiant du « run » ou de l'appareil par exemple). La ligne 2 contient la séquence nucléotidique brute en lettres (A, T, C, G). La ligne 3 est optionnelle, c'est un vestige historique, elle commence par un caractère "+", parfois suivi par la répétition de l'identifiant de la séquence. Enfin, la ligne 4 contient les scores de qualité associés à chacune des bases de la séquence de la ligne 2, ainsi les lignes 2 et 4 et doivent avoir exactement le même nombre de symboles (figure 35).

```
1 @SEQ_ID
2 GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTTT
3 +
4 !'!*((( (***) ) %%%++) (%%% ) .1***-+*'! ) **55CCF>>>>>CCCCCCC65
```

Figure 35 : exemple d'une séquence au format FASTQ

2.2.2. Les scores de qualité du séquençage :

A chaque base lue, va être assignée une valeur de score qualité. En fonction des technologies, il existe plusieurs possibilités pour évaluer la qualité du signal, mais toutes se basent sur l'évaluation des différents critères métriques du signal obtenu, en les comparants à ceux attendus.

Le plus classique est le score Qphred (Q) (standard Sanger). Ce dernier, adapté aujourd'hui aux techniques de NGS a été développé à l'origine pour aider à l'automatisation du séquençage Sanger de l'ADN, lors du Projet Génome Humain. La qualité est évaluée sur les différents critères métriques du signal pour chaque base en comparaison à ceux attendus.

Le score Phred est relié de façon logarithmique à la probabilité d'erreur d'identification d'une base P: $Q = -10 \times \log_{10} P \rightarrow P = 10^{-Q/10}$. Par exemple, si Phred assigne un score de qualité de 30 à une base, la probabilité que cette base ait été identifiée incorrectement est de 1 pour 1000.

À l'origine, les scores étaient codés sous un format numérique, mais ce système n'est pas applicable au format FASTQ car pour chaque base codée par un seul caractère (A, C, G ou T) doit correspondre un score devant être lui aussi codé par un caractère unique. L'utilisation de la table d'encodage ASCII permet de surmonter cette contrainte. Cette table contient 128 caractères (le code ASCII), 93 (de 33 à 126) sont utilisés c'est la notation Phred+33 en NGS (Table ASCII en [Annexe 6](#)).

L'autre variant de score de qualité classiquement utilisé est le score Solexa/Illumina 1.0, utilisant les codes ASCII de 59 à 126 (Phred+64). Ce dernier se base sur une équation différente pour calculer la probabilité d'erreur.

2.3. L'alignement des « reads » sur le génome de référence :

L'alignement des « reads » contre le génome de référence est indispensable avant l'analyse des données. Pour cela, plusieurs algorithmes d'alignement existent, parmi les plus utilisés le Bowtie qui s'appuie sur la transformée de Burrows-Wheeler (LANGMEAD *et al.* 2009). Le but de cette opération est de retrouver la correspondance de chaque « read » au niveau de génome de référence avec le moins d'erreurs possibles, tout en autorisant les mésappariements et les indels. Ceci, nécessite des données de bonne qualité avec un nombre limité d'erreurs. Ainsi, si la séquence s'aligne sur plus d'une position, elle sera éliminée de l'analyse, ceci concerne notamment les régions de faible complexité dont les homopolymères.

Le fichier obtenu est alors au format SAM (Sequence Alignment/Map). A ce stade, il est possible de calculer des statistiques d'alignement (profondeur moyenne, couverture...). A partir du fichier SAM, un fichier BAM est généré (fichier compressé). Il existe également une version indexée dans laquelle les lectures sont classées par localisation génomique (*.bam.bai).

Selon le type de méthode d'enrichissement, certaines étapes doivent être effectuées pour améliorer la qualité des séquences. Par exemple, si une capture a été utilisée, un logiciel peut être utilisé pour éliminer les duplicats de PCR afin d'éviter d'accroître l'impact d'erreurs de séquençage ou d'artéfacts ou l'enrichissement de régions homologues qui ne correspondent pas aux cibles initiales (pseudogènes). Certains outils permettent aussi de réaligner les bases autour des indels afin d'améliorer leur détection au cours des étapes suivantes. Une re-calibration des qualités des bases permet de corriger les biais d'assignement des scores de qualité des séquenceurs. En fait, il est possible, après suppression des bases situées en région 3, de recalculer pour chaque « read » le score d'alignement, si ce celui-ci est trop bas. Ainsi, c'est la séquence du « read » qui s'aligne avec le score le plus haut est qui est gardée, ce procédé est appelé « clipping ». On parlera de « hard clipping » si le nombre de bases supprimées est constant et imposé et de « soft clipping » si ce nombre dépend du score d'alignement obtenu.

2.4. L'identification des variants ou « variant calling »

Sur les « reads » alignés, la recherche de divergence par rapport à la séquence de référence va pouvoir se faire selon des filtres qui sont déterminés par l'expérimentateur en fonction des objectifs du travail.

Au départ, toutes les variations détectées sont annotées mais seules les variants qui remplissent les critères des paramètres choisis sont appelés c'est le « variant calling ». Les variants sont générés sous la forme d'un fichier dit VCF pour « Variant Caller Format »

Les variants peuvent ensuite être sélectionnés en fonction des hypothèses de travail : localisation dans des gènes OMIM, mutations déjà référencées dans des bases de données (dbSNP, ClinVar, HGMD), sélection selon le mode de transmission supposé lorsque les données de plusieurs apparentés sont disponibles.

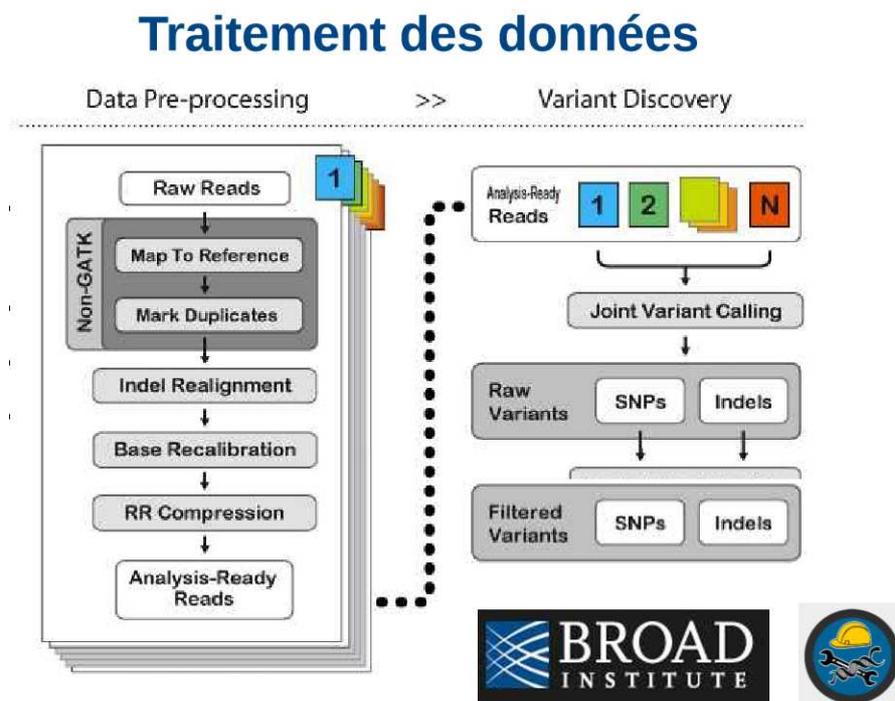


Figure 36 : Récapitulatif des traitements de données bioinformatiques du

3. Autres applications du NGS :

Tout le procédé décrit plus haut est applicable pour les technique dite re-séquençage de génomes connues ce qui est utilisé classiquement pour détecter des mutations en routine diagnostique ou dans la recherche de nouveaux gènes. Néanmoins d'autres applications du NGS sont possibles (BUERMANS and DEN DUNNEN 2014):

- détection des variations chromosomiques de structure (translocations réciproques, inversions, délétions, duplications) grâce à des algorithmes d'alignement différents et des normalisations des profondeurs de lectures
- RNA-seq : séquençage de transcriptome, de miRNA...
- CHIP-seq : séquençage de régions chromosomiques sélectionnées sur la base d'interactions protéiques (étude des sites d'interaction de facteurs de transcription, des modifications post-traductionnelles des histones)
- chromosome conformation capture (Hi-C) : exploration des régions chromosomiques interagissant physiquement dans le noyau des cellules
- séquençage de génomes inconnus : assemblage de novo grâce à la constitution de contigs

L'application désirée va conditionner le choix de la plateforme et de la technologie à utiliser, le pipeline bio-informatique sera également adapté pour obtenir les résultats plus fiables (figure 37). Par exemple, la détection de variations de structure sera facilitée par l'utilisation d'un séquençage bidirectionnel (paired-end) ; l'alignement des ARN messagers, des fragments obtenus par Hi-C ou la détection de transcrits de fusion devront prendre en compte le fait que la séquence correspond à plusieurs régions à cause de l'épissage (notamment alignement par k-mer, fragment par fragment) ; le séquençage de nouveaux génomes passe quant à lui par des algorithmes d'assemblage de novo.

Séquenceurs 2^{ème} génération (2013)

Société	Roche		Illumina				Life technologies						
Plateforme													
Technologie	Titanium	GS FLX+					Chip 314 v2	Chip 316 v2	Chip 318 v2	Chip PI	Chip PII		
Génome humain	✗	✗	✗	✓	✓	✓	✗	✗	✗	✗	✗	✓	✓
Exome	✗	✗	✓	✓	✓	✓	✗	✗	✗	✓	✓	✓	✓
Petit génome (Bactéries, levures)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Régions ciblées	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Transcriptome	✗	✗	✗	✓	✓	✓	✗	✗	✗	✓	✓	✓	✓
Chip-Seq	✗	✗	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓
Métagénomique	✗	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓

Figure 37 : récapitulatif des principales applications NGS possibles sur les différents appareils de séquençage

Deuxième section : Matériels et méthodes

Les deux objectifs de ce travail sont : la mise au point des méthodes de séquençage de nouvelle génération pour l'étude du génome mitochondrial sur le plateau technique de biologie moléculaire de Bordeaux, et la recherche de mutations peu communes chez les patients pour lesquelles une cytopathie mitochondriale a été évoquée mais avec une recherche des principales mutations négative.

1. Première partie : la mise au point de la technique

Le plateau technique de biologie moléculaire au sein duquel se trouve le centre de diagnostic moléculaire des cytopathies mitochondriales disposait déjà d'un appareil de séquençage de nouvelle génération. Il s'agit du système Ion Torrent™ Personal Genome Machine® (PGM), utilisé pour le séquençage de gènes nucléaires dans la cadre d'autres pathologies d'origine génétique. Cette technologie comme décrit précédemment, se base sur l'utilisation de puces à semi-conduction avec détection de variation de pH..Nous avons donc opté pour l'adaptation de cette technologie déjà existante au séquençage du génome mitochondrial dans un but de mutualisation des appareils et réactifs inhérente au fonctionnement d'un plateau technique. De plus, cette technologie avait déjà été testée et publiée par d'autres équipes et semblait convenir au séquençage du génome mitochondrial (SENECA *et al.* 2015) (SEO *et al.* 2015) (PARSON *et al.* 2013).

1.1. Sélection des échantillons à tester

Les ADN séquencés ont été collectés à partir d'échantillons de patients envoyés dans le centre dans le cadre d'une suspicion de cytopathies mitochondriales. Pour chacun d'entre eux, une étude du génome mitochondrial avait été demandée à la recherche de grands réarrangements et de mutations communes

Les échantillons disponibles étaient soit du sang prélevé sur tubes EDTA soit du tissu musculaire congelé prélevé par biopsie. Nous avons sélectionné des patients « contrôles positifs » chez lesquelles une mutation ou un réarrangement avaient été préalablement identifiés dans le laboratoire par les techniques conventionnelles. Des PCR longues et Southern Blot pour la recherche de réarrangements, PCR-RFLP, séquençage Sanger pour les mutations ponctuelles. Le but était d'obtenir une bonne représentation des différentes anomalies rencontrées classiquement dans le laboratoire : délétions

simples, duplications, délétions multiples et mutations ponctuelles avec des taux d'hétéroplasmie variables. Dans le laboratoire, la quantification de l'hétéroplasmie est réalisée pour une seule mutation m. 3243A>G qui a été utilisée pour l'évaluation de la détection de l'hétéroplasmie par NGS.

La quantification se fait par une technique de PCR quantitative utilisant une amorce marquée, couplée à une analyse de migration des fragments après digestion enzymatique (HaeIII) au site de la mutation. Un ADN témoin du laboratoire a également été utilisé comme contrôle négatif.

1.2. Extraction d'ADN

Nous avons utilisé des techniques d'extraction d'ADN génomique total nucléaire et mitochondrial. L'ADN a été purifié partir du sang total prélevé sur tubes EDTA de manière automatisée sur l'appareil *Tecan Freedom EVO*® (Promega®), grâce à la technologie de résine magnétique. Pour le tissu musculaire congelé, une extraction manuelle a été réalisée en utilisant le kit de purification d'ADN *Wizard*® (Promega®). Après une digestion par l'enzyme protéinase K, une lyse des membranes cytoplasmiques et nucléaires est réalisée suivie d'une précipitation des protéines. L'ADN génomique contenu dans le surnageant a ensuite été précipité grâce à de l'isopropanol puis réhydraté.

Le dosage et le contrôle de qualité de l'ADN extrait ont été évalués par spectrophotométrie au NanoDrop® (mesure d'absorbance à 260 nm et à 280 nm pour les contaminants protéiques).

1.3. Les étapes de la technique de préparation de la librairie et séquençage

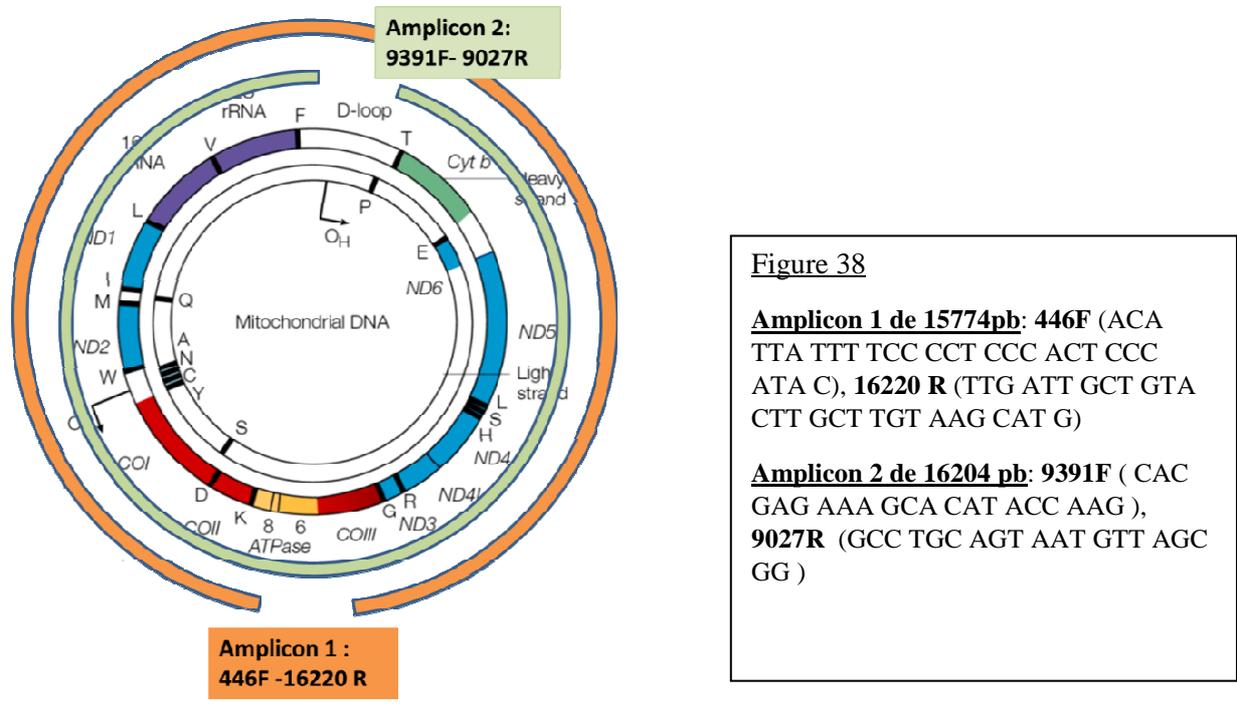
Le séquençage de nouvelle génération se base sur le principe du séquençage massif en parallèle. Le but est donc d'obtenir des fragments identifiés de petites tailles représentant la séquence de la région cible à séquencer. Pour cela plusieurs étapes sont nécessaires (cf [Annexe 7](#) pour plus de détails):

- L'enrichissement :

Travaillant à partir d'extraction d'ADN génomique total c.à.d. contenant aussi bien de l'ADN nucléaire que mitochondrial, une étape d'enrichissement ciblant le génome mitochondrial était nécessaire.

Nous avons utilisé deux PCRs longues pour amplifier deux fois la quasi-totalité du génome mitochondrial en un seul amplicon, et ce grâce à deux couples d'amorces spécifiques de l'ADNmt (figure 38).

Le kit de la polymérase *TaKaRa* (Ozyme®) a été utilisé pour la réalisation des PCRs longues. La vérification des produits de PCR a été réalisée sur gel d'Agarose 1% révélés aux Ultraviolets (UV) grâce à l'incorporation de *Midori Green*, un agent intercalant fluorescent aux UV.



- Fragmentation des produits de PCR :

Les produits de PCRs ont été purifiés sur billes Agencourt® AMPure® XP (Beckman Coulter®), puis fragmentés par digestion enzymatique grâce au kit Ion Shear™ Plus (Life Technologies®), en suivant les instructions du fournisseur. Une étape d'optimisation du temps de digestion à 37° a été nécessaire afin d'obtenir des fragments de la taille souhaitée. Au total, 100ng de produits de PCR (50ng pour chacune des deux PCRs) ont été digérés par échantillons, après dosage au *Qubit*®. Cette méthode de dosage des acide nucléiques ou fluoromètre est plus précise que la spectrophotométrie classique (*NanoDrop*®) car elle utilise des agents fluorophores qui ne fluorescent que quand elles sont liées aux molécules d'ADNs permettant ainsi une quantification précise des acide nucléiques sans les contaminants. Le système du *Qubit*® possède également l'avantage de pouvoir doser spécifiquement les molécules d'ADN doubles brins.

- Ligation des adaptateurs et addition des barcodes :

Sur les fragments digérés et purifiés, l'addition des adaptateurs et des « barcodes » (Ion Xpress Barcode Adapters 1-32) a été réalisée grâce au kit Ion Plus Fragment Library Kit™ (Life Technologies®). Une amplification fidèle de 5 cycles a été ensuite réalisée en suivant les recommandations du fournisseur.

- Sélection des fragments d'intérêts :

La sélection des fragments dans les tailles d'intérêt, à savoir entre 150 et 380 pb (pour un séquençage de 200 pb) s'est faite grâce à une purification par billes Agencourt® AMPure® XP (Beckman Coulter®). En jouant sur le volume de billes utilisé par rapport au volume d'acides désoxyribonucléiques à purifier, l'affinité des billes pour les molécules d'ADN n'est pas la même et permet ainsi de purifier les fragments en fonction de leurs tailles.

- Contrôle de la taille des fragments et leur quantification :

L'analyse des fragments obtenus se fait ensuite par électrophorèse capillaire grâce au système de la Tape Station (Agilent® 2200 TapeStation System). Ce système se base sur l'électrophorèse capillaire des échantillons sur cassettes microfluidiques, il permet de vérifier la distribution des tailles de fragments de la librairie et leurs quantifications après acquisition de l'image et analyse automatisée des électrophorogrammes (figure 39)

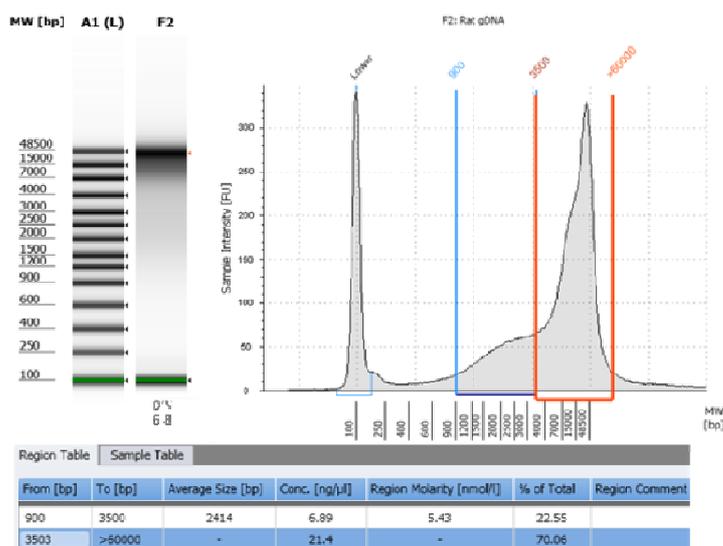


Figure 39 : Quantification de la librairie à la *TapeStation*

Après le dosage, tous les échantillons identifiés grâce à un « barcode » sont poolés aux mêmes concentrations pour les étapes suivantes (PCR en émulsion sur billes et dépôt sur puce). En effet, le multiplexage de plusieurs échantillons sur une même puce de séquençage exige une préparation de librairies équimolaires avec la même quantité de molécule d'ADN à séquencer pour chaque patient, afin d'éviter tout déséquilibre dans la répartition des lectures.

- La PCR en émulsion sur billes (Ion Spheres™ Particles) et le dépôt sur puce 316v2 ont été faits de manière robotisée grâce au système Ion Chef™ (Life Technologies®) toujours en suivant les instructions du fournisseur.

- Le séquençage en lui-même a été réalisé sur l'appareil Ion Torrent™ Personal Genome Machine®, en utilisant le kit Ion PGM™ IC 200, il s'agissait d'un séquençage unidirectionnel de taille de 200pb. Une étape d'initialisation de l'appareil est d'abord réalisé en suivant les instructions du fournisseur, elle permet notamment une homogénéisation des valeurs de pH ~ 7,8 au sein des différents réactifs de l'appareil (« Auto Ph »)

Par ailleurs, les mutations retrouvées en NGS ont été vérifiées par RFLP ou séquençage Sanger sur un appareil ABI 3500 xL Dx (Applied Biosystems®)

1.4. Analyses des données : pipeline bio-informatique

L'analyse des données a été réalisée directement via le logiciel fourni *Torrent Suite*® (Life Technologies).

Les données brutes générées par le séquenceur sont produites sous la forme d'un signal électrique ou différence de potentiel (fichier.DAT), qui est transformé dans un second temps en séquence nucléotidique (Base calling). (Figure 40)

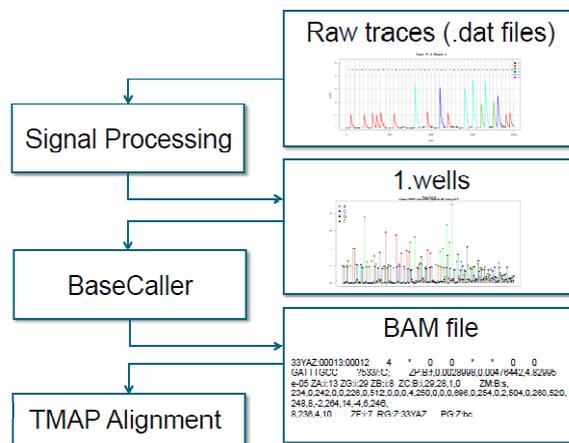


Figure 40 : Traitement du signal brut et conversion en séquences nucléotidiques « signal processing » et « bsase calling »

Une deuxième étape d'évaluation de qualité des séquences a été ensuite réalisée car en effet, seules les « reads » de qualité suffisante ont été gardées pour l'alignement contre le génome de référence.

Pour cela, nous avons maintenu les paramètres de qualité présents par défaut et recommandés par le fournisseur, ont été donc exclus de l'analyse :

- Les bases dont le score de qualité *Phred* était supérieur à 15 ce qui correspond à un risque d'erreur supérieur à 3%
- Les « reads » dont les tailles étaient inférieures à 8 bases après « trimming »

Les séquences contaminantes, les « reads » issus de signaux polyclonaux ou qui ne possédaient pas de séquences de marquage (Key signal : ATCG.) ont été bien entendu supprimées.

Les « reads » d'une qualité suffisante ont ensuite été alignés contre le génome de référence mitochondrial grâce à l'algorithme « *TMAP* » fourni par le logiciel. Cet algorithme est basé sur l'algorithme BWA « *Burrows-Wheeler Aligner* » largement utilisé (LANGMEAD *et al.* 2009) qui a été en plus optimisé pour la technologie du séquençage Ion Torrent®. C'est en réalité une combinaison de plusieurs algorithmes qui s'adaptent aux différentes tailles des « reads » et tendent à améliorer l'alignement des régions de faibles complexités (contenant des homopolymères).

La détection des variants, substitutions nucléotidiques (SNP) ou petites insertions/délétions (indels) s'est faite sur tous les « reads » alignés sur la région cible, à savoir tout le génome mitochondrial.

Le filtre de détection des variants choisi était le « variant caller v4.4.2.1 » fourni par le logiciel Torrent Suite® avec les paramètres de détection de mosaïques et à faible stringence. Ces paramètres ont pour principale caractéristique un seuil de détection des substitutions à 2% ce qui permet de détecter des mutations avec un faible taux d'hétéroplasmie. Ce paramétrage a aussi une très grande sensibilité avec un seuil des scores de qualité d'alignement bas et une grande tolérance pour les biais de brins allant jusqu'à 95%. (**cf annexe 7** pour plus de détails)

Les données concernant l'alignement des « reads » et l'identification des variants ont été respectivement générés sous format de fichiers .BAM et VCF. Pour chaque patient, ces fichiers ont été lus grâce au logiciel Alamut® qui permet la visualisation aussi bien des variants que des « reads » tout le long la séquence de référence. Un autre logiciel gratuit de visualisation avait été testé initialement, le « *Integrative Genomics Viewer* » (IGV®) mais ce dernier présentait plusieurs inconvénients dont la lenteur d'exécution incompatible avec une utilisation en routine diagnostique et l'impossibilité d'annoter au fur et à mesure les variants retrouvés chez les patients.

Il faut noter qu'il était également possible de visualiser directement les variations nucléotidiques dans un fichier EXCEL généré par le logiciel *Torrent Suite*® sous la forme d'un tableau comportant plusieurs colonnes qui contiennent notamment les données de qualité et de couverture pour chaque variant identifié.

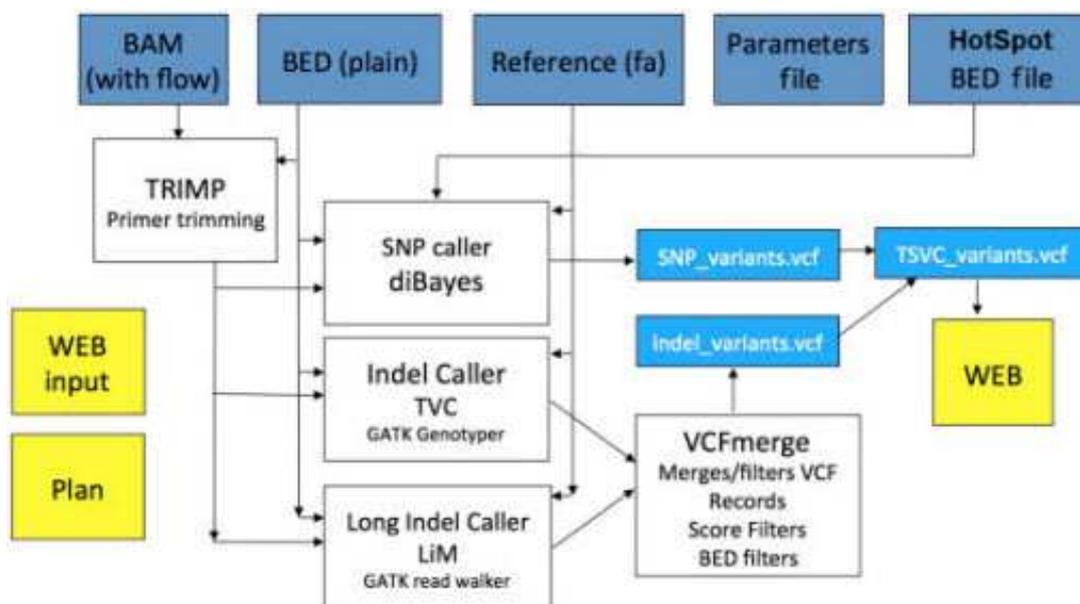


Figure 41 : Schématisation d'un Workflow en NGS pour la détection de variants

2. Deuxième partie : recherche de nouvelles mutations :

2.1. Sélection des patients à étudier :

Pour l'évaluation des techniques de « NGS » dans la détection des mutations peu communes, nous avons sélectionné des patients avec une présentation clinique très en faveur d'une cytopathie mitochondriale mais pour lesquels l'étude moléculaire du génome mitochondrial par techniques conventionnelles s'était révélée négative.

Pour la population adulte les critères d'inclusion, en dehors des caractéristique cliniques étaient la présence d'anomalies mitochondriales à la biopsie musculaire de type : 1) des mitochondries anormales à la microscopie électronique, 2) des « ragged red fibres » à la coloration au trichome de Gomori, 3) des fibres COX négatives en histoenzymologie. Les anomalies histologiques étant en général moins présentes dans la population pédiatrique, nous avons inclus les enfants sur la base d'une présentation clinique compatible et la présence d'un déficit d'un ou plusieurs complexes de la

chaîne respiratoire. Ont été exclus, les patients issus d'une union consanguine pour lesquelles une origine nucléaire semblait plus plausible.

2.2. Stratégie d'interprétation des variants retrouvés :

Pour chaque patient, les variants identifiés ont été annotés, triés puis classés en trois catégories : les variants polymorphiques, les mutations pathogènes et les variants de signification inconnue. Cette classification s'est faite grâce à l'utilisation des bases de données existantes, répertoriant toutes les variations du génome mitochondrial. Il est important de noter que même si les bases de données classiques, regroupant les variations du génome nucléaire tel que « dbSNP » ou « 1000genomes » contiennent un certain nombre de données concernant le génome mitochondrial, celles-ci ne sont pas assez fournies pour permettre l'interprétation des variations génomiques mitochondriales. En effet, des bases de données publiques plus spécifiques à l'ADNmt existent depuis plusieurs années dont deux : « Mitomap » et « mtDB » sont largement utilisées par les communautés scientifique et médicale.

La base de données « Mitomap » (<http://www.mitomap.org/MITOMAP>) a été créée en 1996 (KOGELNIK *et al.* 1998) et elle est mise à jour régulièrement. C'est l'une des plus exhaustives puisqu'elle répertorie actuellement plus de 30000 séquences mitochondriales publiées ou non, soumises par différentes sources. Les variants dans « Mitomap » sont classés en plusieurs groupes :

- Les polymorphismes qui sont les variations fréquemment observées
- Les mutations confirmées « confirmed » sont celles dont la preuve de pathogénicité n'est plus à démontrer, car elles ont été retrouvées dans plusieurs familles avec des phénotypes similaires ou ont été validées sur le plan fonctionnel.
- Les mutations dites rapportées ou « reported » ont été associées à des phénotypes pathologiques mais leur implication en n'a pas encore été prouvée.

La base mtDB pour « *Human Mitochondrial Genome Database* » (<http://www.genpat.uu.se/mtDB>) (INGMAN and GYLLENSTEN 2006) est une base de données plus récente, elle est plus réduite que la précédente avec environ 2700 patients mais plus informative car contient des échantillons provenant d'origines ethniques variables et a pour objectif d'être la plus représentative possible des différentes populations humaines et leurs haplogroupes. Ainsi, la base mtDB offre la possibilité de détecter les polymorphismes même rares car spécifiques à certains groupes ethniques. Par ailleurs, les sujets donateurs ont été évalués et étaient tous sains au moment de l'étude moléculaire, conférant à cette base une grande fiabilité quant à la détection des polymorphismes. Il est maintenant établi que les variants

retrouvés avec une fréquence supérieure à 0.2% dans mtDB sont des variants bénins ou polymorphismes (WANG *et al.* 2012)

Dans cette étude nous avons considéré comme bénins, les variants qui étaient présent dans au moins 0.2% des sujets de mtDB. Les autres ont été recherchés dans « Mitomap » et classés en mutations pathogènes si répertoriés en tant que telles. Les variations absentes ou très rares de « Mitomap » et les mutations non confirmées « reported mutations » ont été classées en variants de signification inconnue.

Dans un second temps, l'effet pathogène des variants inconnus a été exploré grâce à des outils de prédiction *in silico* tel que « SIFT » et « Polyphen2 » (consultation via l'outil mitImpact (CASTELLANA *et al.* 2015)). Ces outils se basent sur des algorithmes qui établissent des scores de pathogénicité en fonction de plusieurs paramètres dont la conservation nucléotidiques entre espèces, l'effet sur la conformation protéique 3D ou la modification des sites catalytiques. Ces outils sont donc surtout utiles pour les substitutions en acides aminés non synonymes dans des gènes codant pour des protéines. L'impact des mutations dans les ARN de transfert et ribosomiques a été évalué à l'aide de la base de donnée « Mamit-tRNA » (<http://mamit-trna.u-strasg.fr/human.asp>) qui répertorie tous les ARNt mitochondriaux, leurs structures, les variants et mutations déjà identifiés ainsi que des références bibliographiques étendues sur le sujet.

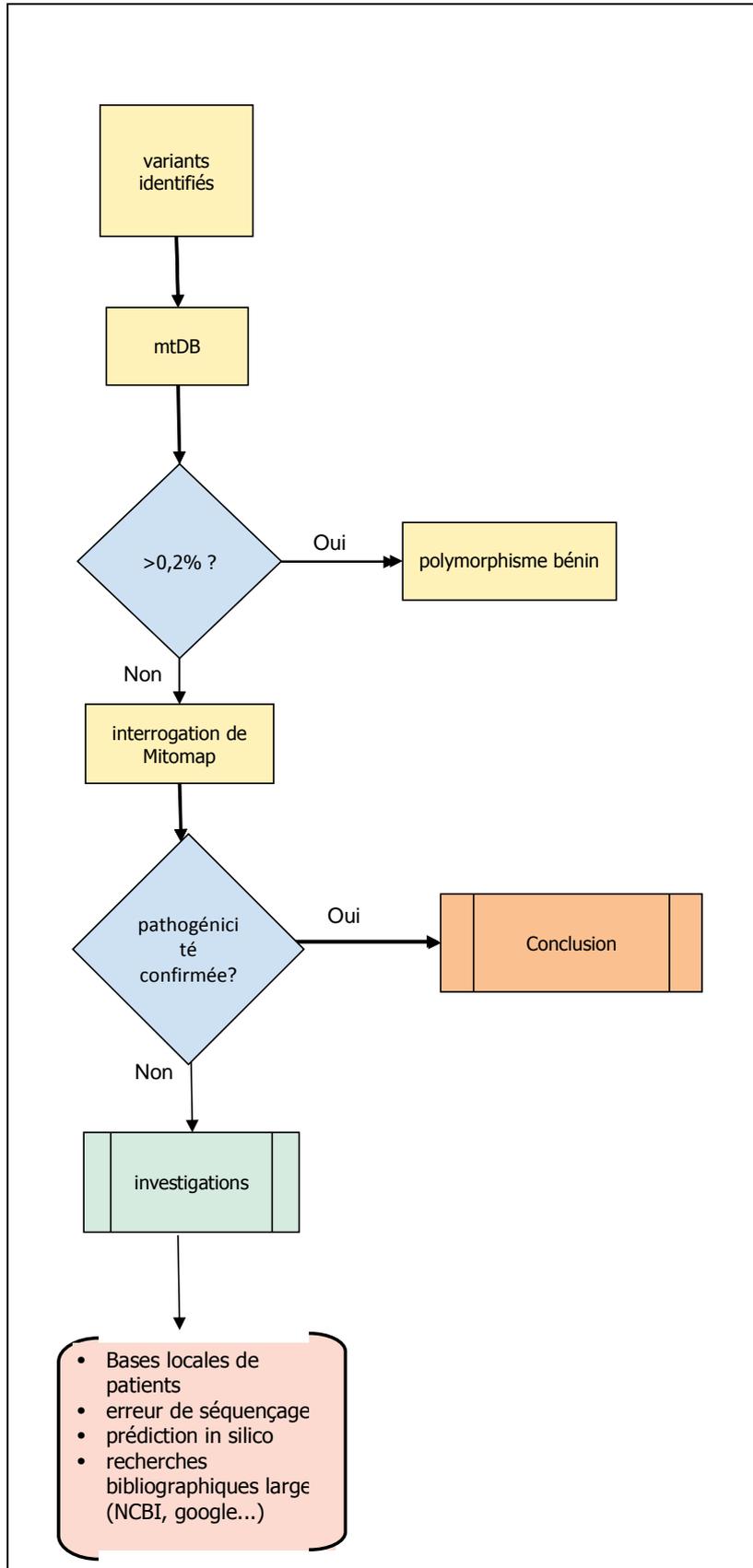


Figure 42 : Algorithme décisionnel d'interprétation des variants

Troisième section : Résultats et discussion

Au total quatre « runs » de séquençage 200 pb sur puce 316v2 ont été réalisés pendant ce travail, les trois premiers ont fait l'objet d'une optimisation pour améliorer le rendement des puces. Les échantillons tests pour la détection de mutations connues ont été séquencés sur les deux premières puces : 8 sur la première et 5 sur la deuxième dont 2 étaient des duplicats de la première (tableau 3). La recherche de mutations peu communes a été réalisée sur les deuxième, troisième et quatrième puces de séquençages : 3 échantillons sur la première, 9 sur chacune des troisième et quatrième puces.

n° run	P	Tissu	mutations	taux
run 1	P1	S	m.3243G>A	Hétéroplasmie : 60%
	P2	S	Délétion simple	
	P3	M	Délétions multiples	
	P4	M	m.3243G>A	Hétéroplasmie : 3%
	P5	S	m.3243G>A	Hétéroplasmie : 12%
	P6	S	m.11778G>A	homoplasmique
	P7	S	Témoin Normal	
	P8	S	Témoin Normal	
run 2	P9	M	Non étiqueté	
	P10	M	Non étiqueté	
	P11	M	Non étiqueté	
	P12	M	délétion simple	
	P13	M	Non étiqueté	
	P14	S	Duplication	
	P2bis	S	délétion simple duplicat	
	P4bis	M	m.3243G>A duplicat	hétéroplasmie: 3%
run 3	P15	M	Non étiqueté	
	P16	S	Non étiqueté	
	P17	M	Non étiqueté	
	P18	M	Non étiqueté	
	P19	M	Non étiqueté	
	P20	M	Non étiqueté	
	P21	M	Non étiqueté	
	P22	M	Non étiqueté	
	P23	M	Non étiqueté	
run 4	P24	M	Non étiqueté	
	P25	M	Non étiqueté	
	P26	M	Non étiqueté	

P27	M	Non étiqueté	
P28	M	Non étiqueté	
P29	M	Non étiqueté	
P30	S	Non étiqueté	
P31	S	Non étiqueté	
P32	M	Non étiqueté	

Tableau 3 : Récapitulatif de tous les patients étudiés durant le travail. P = patient, S=sang EDTA, M = muscle congelé

1. Première partie : mise au point de la technique

1.1. Evaluation des coûts:

L'étude de l'ADN mitochondrial dans le centre de diagnostic moléculaire de Bordeaux, consistait principalement en la recherche de 11 mutations ponctuelles par PCR-RFLP, et des grands réarrangements par plusieurs PCRs longues +/- Southern Blot. En 2014, le centre a reçu 117 demandes d'étude d'ADN mitochondrial pour un coût total de 10192.29 euros, ce qui fait une moyenne de 87 euros/patient. Nous avons souhaité connaître les surcoûts qu'engendrerait cette nouvelle technique par rapport aux techniques conventionnelles.

Le tableau 4 résume le coût de chaque réactif utilisé pour le NGS. Ainsi, le séquençage du génome mitochondrial coûterait environ 115.19 euros/patient ce qui paraît relativement peu coûteux étant donné la somme d'informations et de données obtenues grâce à cette technique (0.007 euros/base étudiée)

Produit	Fournisseur	référence	prix CHU euros	Nb d'échantillons analysés par kit	prix par patient
QUBIT ASSAY TUBES (500)	Life Technologies	Q32856	57,35	500	0,11 €
QUBIT DsDNA HS ASSAY Kit	Life Technologies	Q32851	71	100	0,71 €
Ion 316™ Chip Kit v2 BC (4-pack)	Life Technologies	4488145	1047,1	36	29,09 €
Ion PGM™ IC 200 Kit	Life Technologies	4484080	2192,25	72	30,45 €
Ion Xpress™ Plus Fragment Library Kit	Life Technologies	4471269	748,6	20	37,43 €
Ion Xpress Barcode Adaptors 1-16 Kit	Life Technologies	4471250	635	160	3,97 €
8 X PCR Strip 960 échantillons	Agilent	401428	73,92	960	0,08 €
CAPS for 8 X PCR Strip	Agilent	401425	33,6	960	0,04 €
Loading Trips (1pack)	Agilent	5067-5153	90,72	384	0,24 €
High sensitivity D1000 screentape	Agilent	5067-5584	230,16	112	2,06 €
High sensitivity D1000 screentape regen	Agilent	5067-5585	78,12	112	0,70 €
Agencourt AMPure XP 60ml	Beckman Coulter	A29152	1131,6	270	4,19 €
Enzyme Takara 1x125U	Ozyme	RR002A	156	25	6,24 €
					115,29 €

Tableau 4 : des coûts du séquençage du génome mitochondrial par le système PGM®

1.2. Rendement de la micro-puce de séquençage

Une étape d'optimisation du chargement de la micro-puce de séquençage 316v2 a été réalisée. Cette puce a la capacité théorique de générer 2 à 3 millions de « reads » par « run » ce qui correspond à un rendement d'environ 300 Méga bases utilisables pour les kits de séquençage de 200 pb. Ce rendement dépend entre autres de l'étape de la PCR en émulsion sur billes magnétiques et de la qualité du chargement de la puce.

En effet, le but est :

- d'enrichir un maximum de billes avec une seule molécule d'ADN par bille.
- de déposer une seule bille par puits lors de l'étape de chargement.

L'objectif final est d'éviter les puits vides ou engendrant des signaux polyclonaux (figure 43). Ce taux d'enrichissement dépendra par conséquent de l'utilisation de la bonne proportion de molécules d'ADN -molarité des échantillons - par rapport à la quantité de billes pendant la PCR en émulsion.

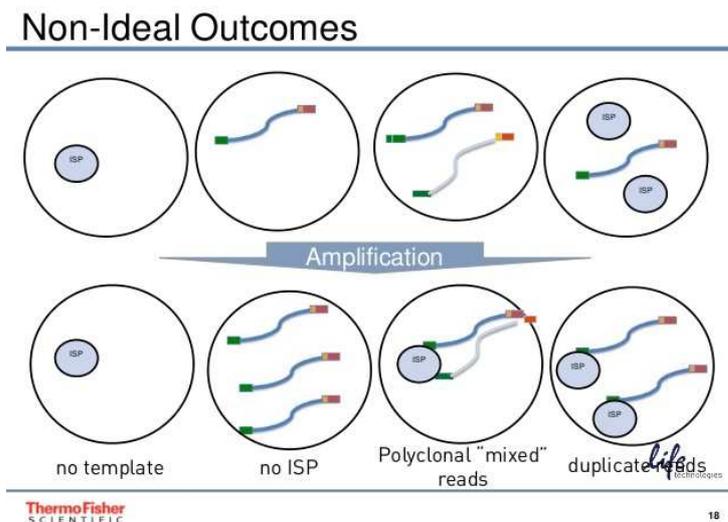


Figure 43: dans les conditions idéales de la PCR en émulsion : un puits=une bille=une molécule d'ADN. *IPS = ion sphere particles*

La figure 44 résume les résultats de rendement des 3 premières puces réalisées. On remarquera que ce dernier s'est nettement amélioré entre le premier et le dernier « run », se rapprochant des rendements théoriques annoncés : 2,6 million de « reads » utilisables avec 34% de polyclonalité. Pour cela, nous avons augmenté la molarité des bibliothèques utilisées de 26 pM (ce qui était utilisé localement sur le plateau pour le séquençage des gènes nucléaires) à 30 pM. L'amélioration du rendement nous a permis d'augmenter le nombre d'échantillons séquencés de 8 à 9 par micro-puce avec de bons taux de couverture.

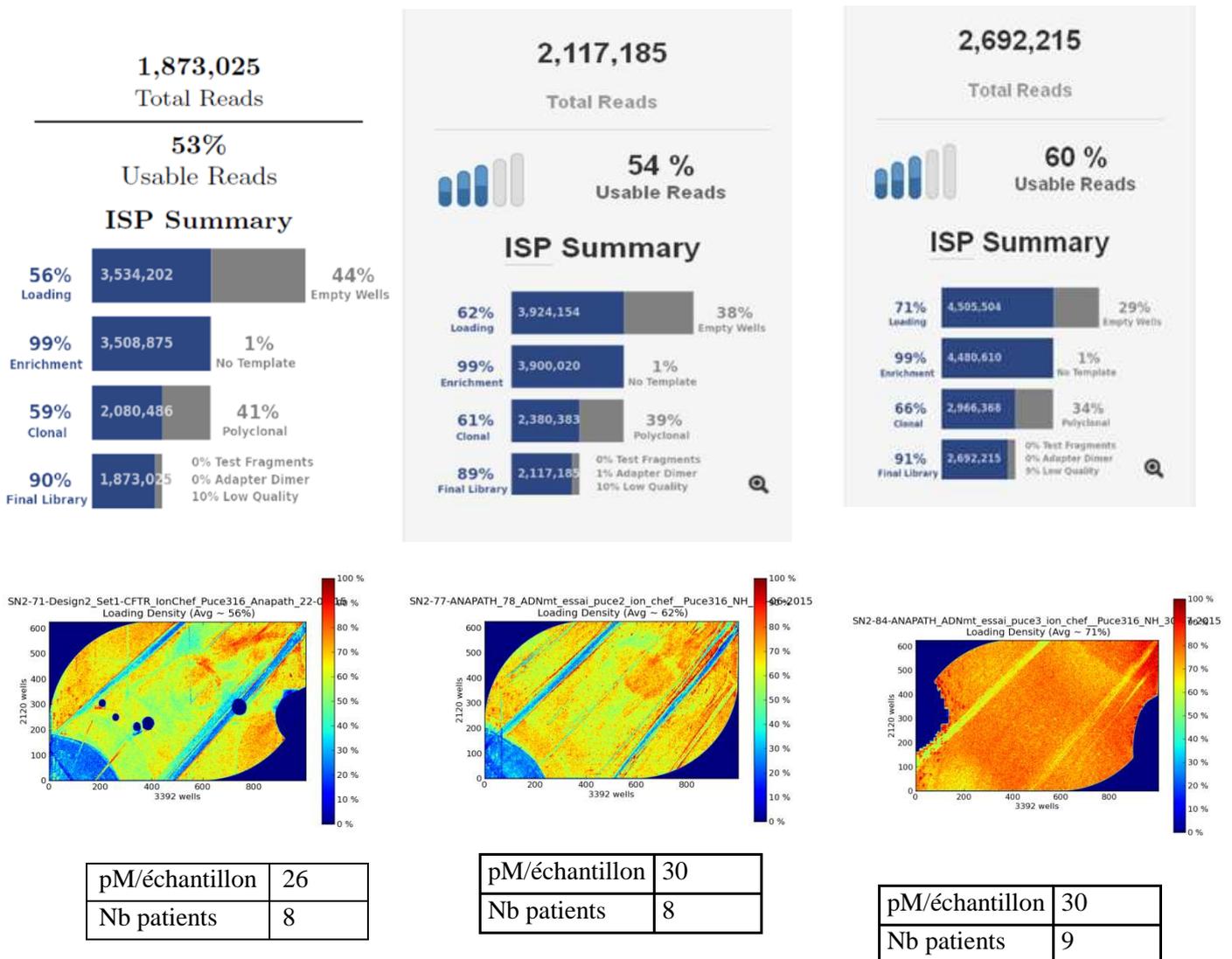


Figure 44 : résultats de rendement des 3 premières puces réalisées. Indication de la quantité totale de reads générés et des étapes de sélection conduisant à la librairie finale.

1.3. Longueurs moyennes des « reads » et optimisation de la digestion

La longueur moyenne des « reads » est passée de 149 à 183 pb entre le premier et le troisième « run » grâce à l'optimisation du temps de fragmentation par digestion enzymatique.

Le temps de digestion enzymatique à 37°C nécessaire pour une fragmentation optimale pouvait varier en fonction des conditions et appareils utilisés dans chaque laboratoire (bain-marie, sec, agitation...). Le fournisseur recommandait donc de tester plusieurs temps de digestion pour obtenir le profil de fragmentation désiré. Pour un séquençage cible de 200 pb (Ion PGM™ IC 200 kit), le fournisseur préconisait d'obtenir une distribution des tailles de fragments entre 100 et 700 pb pour une médiane située entre 200 et 300 pb. Ce profil a été obtenu dans notre laboratoire grâce à un temps de digestion

de 13 minutes, mais la longueur finale des « reads » obtenue avec ce profil n'était pas satisfaisante (149 pb). Nous avons donc décidé de réduire la durée de digestion à 11 min ce qui a permis d'obtenir des longueurs de « reads » dans la fourchette souhaitée (183 pb) (figure 45).

Au final le profil de digestion permettant d'obtenir la longueur de « reads » cible n'était pas celui recommandé par le fournisseur (décalage de la médiane d'environ 100 pb). Ceci pourrait s'expliquer par les étapes de purification qui succèdent à la digestion, avec une perte plus importante des grands fragments par notre technique de préparation de la librairie. De plus, ces recommandations ne sont pas forcément actualisées pour s'adapter à l'évolution constante des kits de séquençage. Ce qui était au début une taille maximale de séquençage est aujourd'hui une taille moyenne avec, pour les kits de 200 pb, des tailles de lecture pouvant aller jusqu'à 350 pb.

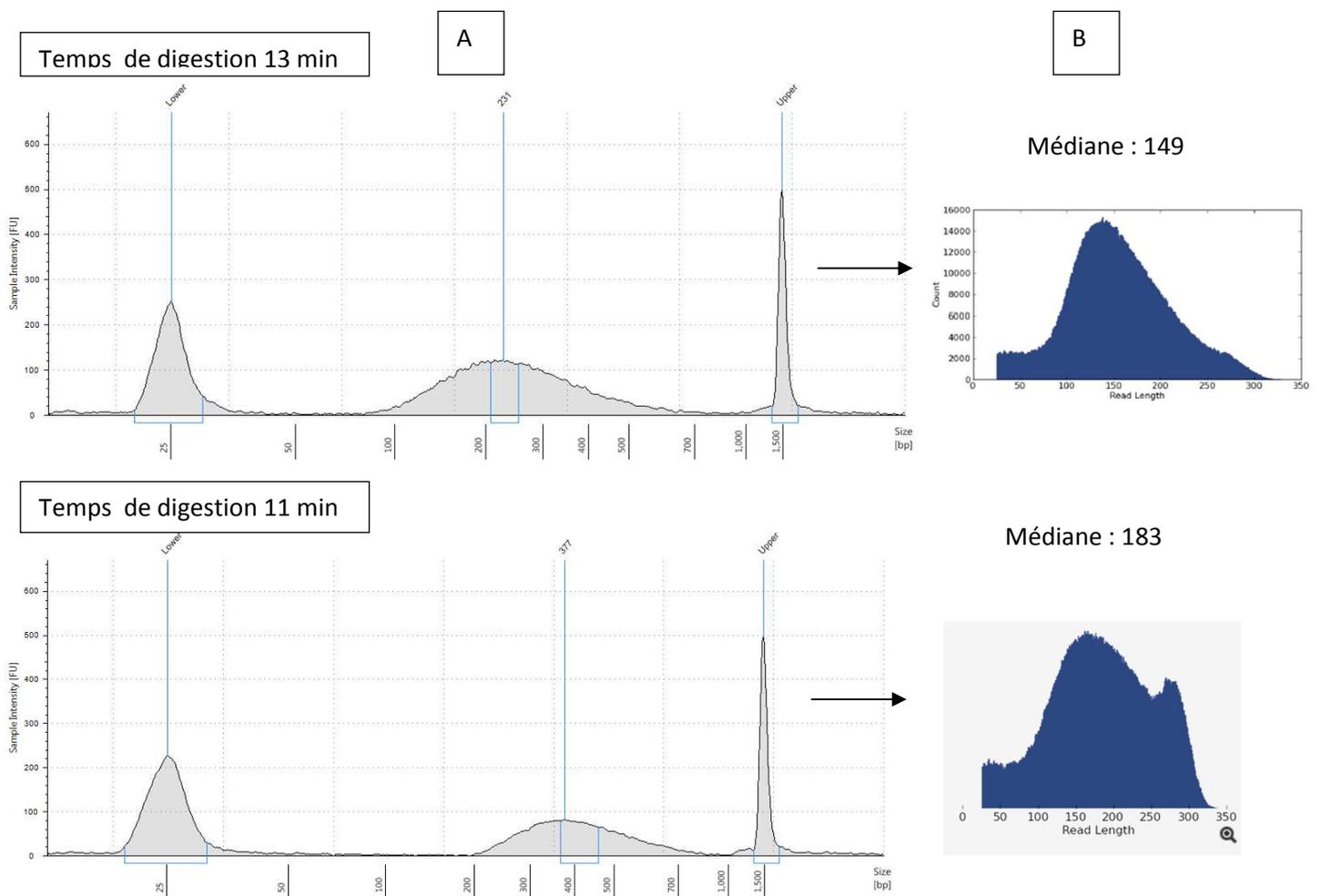


Figure 45 : Contrôle de la qualité de la digestion enzymatique (*TapeStation*®) (a) et distribution des longueurs des « reads » (b) à 13 et 11 minutes.

1.4. L'alignement

Le génome mitochondrial est référencé depuis 1981 (ANDERSON *et al.* 1981) , c'est l' « *original Cambridge Reference Sequence* » (oCRS). Cette version a ensuite été révisée en 1999, en apportant 7 corrections c'est la référence utilisée aujourd'hui : « *the revised Cambridge Reference Sequence* » (The rCRS : NC_012920 gi:251831106) (ANDREWS *et al.* 1999). Mais la séquence du génome mitochondrial a aussi été incluse dans la référence du génome humain et notamment dans les dernières versions (Hg 19/*Genome Reference Consortium GRCh37*) en tant que chromosome, appelé le chromosome M.

Classiquement l'alignement des « reads » de séquençage se fait sur la « rCRS », afin d'éviter les alignements non appropriés sur les séquences mitochondriales nucléaires. Le problème que pose l'alignement contre la séquence de référence mitochondriale est l'impossibilité de lire les fichiers .BAM (les « reads » alignés) générés dans les logiciels de visualisations dédiés habituels, comme *Alamut*® qui ne sont pas paramétrés pour reconnaître la « rCRS ». Ces logiciels sont pourtant très utiles pour la validation et l'annotation des variants, l'évaluation de la qualité réelle de l'alignement et éventuellement la visualisation des grandes délétions et de leurs points de cassure. Aussi, nous avons tenté d'évaluer l'alignement sur l'« Hg19 » par rapport à celui réalisé directement sur la « rCRS », afin d'identifier d'éventuelles objections à son utilisation en routine diagnostique pour la recherche de mutations de l'ADNmt. Pour cela, nous avons comparé deux alignements différents à partir de données issues d'un même « run » de séquençage (le run n°1) : alignement sur l'« Hg19 » versus la « rCRS ».

1.4.1. Les résultats d'alignement (run n°1)

➔ Alignement contre la « rCRS »

Sur les 1839103 « reads » générés utilisables du run n°1, 98.9% ont été alignés sur la « rCRS ». Les scores d'alignement étaient bons, puisque plus de 260 millions de bases avaient un score d'alignement supérieur à Q17 soit un risque d'erreur inférieur à 2% (tableau 7).

Total Reads	1,839,103	100%
Aligned Reads	1,818,877	98.9%
Unaligned Reads	20,226	1.1%

	AQ17	AQ20	Perfect
Total Number of Bases [bp]	260 M	237 M	181 M
Mean Length [bp]	150	142	114
Longest Alignment [bp]	350	339	331

Tableau 5 : Résultats de l'alignement sur la « rCRS » issus du logiciel *Torrent Suite*®

➔ Alignement contre la « Hg19 »

Nous avons ensuite réaligné les mêmes données sur la référence « Hg19 » : sur un total de 1838847 « reads » générés, 1827423 à savoir 99.4% ont été alignés avec des scores de qualités équivalents à ceux de l'alignement sur la « rCRS » (tableau 6).

Total Reads	1,838,847	—
Aligned Reads	1,827,423	99.4%
Unaligned Reads	11,424	0.6%

	AQ17	AQ20	Perfect
Total Number of Bases [bp]	263 M	239 M	183 M
Mean Length [bp]	150	142	114
Longest Alignment [bp]	350	339	331

Tableau 6 : Résultats de l'alignement sur l'« Hg19 » issus du logiciel *Torrent Suite*®

Nous avons ensuite cherché à connaître les « reads » réellement alignés sur le génome mitochondrial (le chromosome M). Le tableau 7 ci-dessous, résume pour les 8 patients du premier « run » le nombre de « reads » effectivement alignés sur le chromosome M (génome mitochondrial réel) parmi tous les « reads » alignés (sur l'« Hg19 »). Ainsi, en moyenne 215766 « reads » ont été correctement alignés par patient, pour un total moyen de 228428 « reads » soit 94.5%.

	P1	P2	P3	P4	P5	P6	P7	P8	Moyenne
nb de reads alignés sur chrM	232131	264530	217148	228427	158547	198548	260039	166758	215766
total reads aligné (Hg19)	244710	278914	226703	238572	168277	212668	279951	177628	228427,875

Tableau 7 : Reads alignés sur le chrM lors de l'alignement sur l'« Hg19 ».

→ Evaluation de l'efficacité de l'alignement sur la référence « Hg19 »

En résumé, l'alignement global sur la référence « Hg19 » a généré un total de 1827423 « reads » alignables sur tout le génome, soit 8546 « reads » de plus que l'alignement sur la « rCRS ». Cette différence est probablement due aux « reads » dont les scores d'alignement étaient insuffisants pour le génome mitochondrial mais étaient alignables sur le reste du génome. Néanmoins, parmi tous ces « reads », 1726128 seulement ont été effectivement alignés sur le chromosome M. Ainsi, 92749 « reads » ont été perdus par mauvais alignement sur les autres chromosomes (séquences mitochondriales nucléaires). Les pertes réelles d'alignement en « reads » et en nombre de bases totales pour chaque patient sont représentées respectivement dans le tableau 8. En moyenne la perte des « reads » alignés était de 5.10% (92749 « reads ») ce qui correspondait à 1155824 bases soit 3.34% des bases totales. Cette différence s'explique probablement par le fait que les mauvais alignements concernent le plus souvent les petits « reads » ce qui est inhérent au processus et aux algorithmes d'alignement.

Nb reads	P1	P2	P3	P4	P5	P6	P7	P8	total	Moyenne
chrM	232131	264530	217148	228427	158547	198548	260039	166758	1726128	
RCS	245453	279681	226971	239221	166961	209786	274471	176333	1818877	
différence	13322	15151	9823	10794	8414	11238	14432	9575	92749	11593,625
pourcentage	5,43%	5,42%	4,33%	4,51%	5,04%	5,36%	5,26%	5,43%		5,10%
Nb de bases	P1	P2	P3	P4	P5	P6	P7	P8	total	Moyenne
ChrM	35088527	40954639	34231518	35733392	25601429	30358103	39626372	25085291	266679271	
RCS	36435847	42554038	35232116	36812960	26432700	31453399	41005933	25998875	275925868	
différence	1347320	1599399	1000598	1079568	831271	1095296	1379561	913584	9246597	1155824,63
pourcentage	3,70%	3,76%	2,84%	2,93%	3,14%	3,48%	3,36%	3,51%		3,34%

Tableau 8 : le séquençage de l'ADNmt, alignement sur l'hg19 (chrM) versus rCRS comparatif en terme de nombre de « reads » (haut) et nombre de bases (bas)

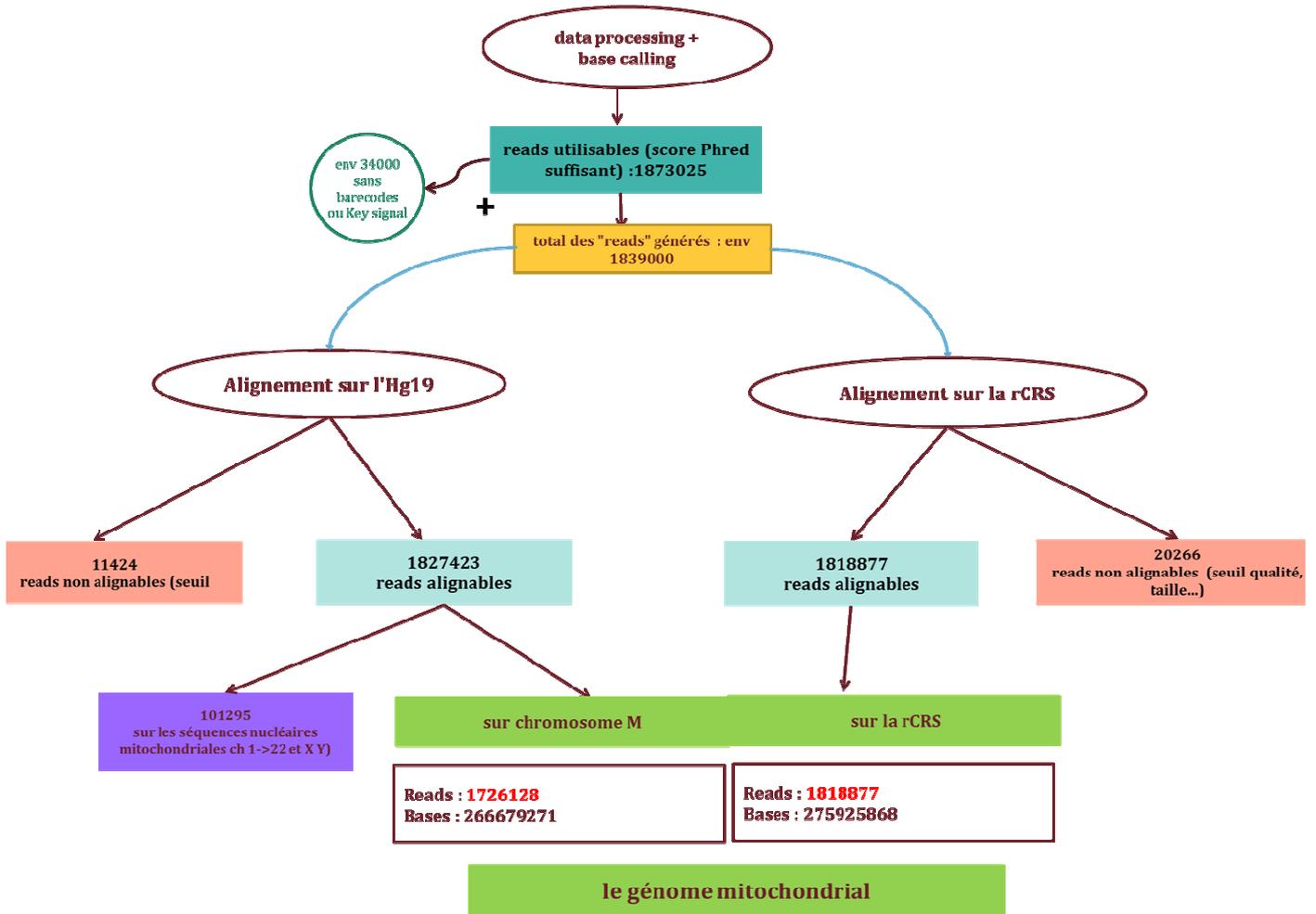


Figure 46 : schématisation des résultats de l'alignement sur la « rCRS » versus l' « Hg19 »

➔ L'alignement sur les autres chromosomes :

Nous avons ensuite essayé d'identifier les mauvais alignements sur les différents chromosomes, tel que classés dans le tableau 9, ci-dessous. On notera que le chromosome 1 est celui qui contient le plus de « reads » alignés avec 3.3%. Les autres ne sont que très peu représentés et ont une participation minimale dans la distribution des « reads ». L'alignement sur le chromosome 2 ne contient que 0.2% des « reads », malgré le fait que ce dernier est le plus riche en séquences mitochondriales. Ceci pourrait être attribué à une moins bonne similitude des séquences.

	P1	P2	P3	P4	P5	P6	P7	P8	moyenne	pourcentage
chrom	total_reads									
chrM	232131	264530	217148	228427	158547	198548	260039	166758	215766	94,5%
chr1	8616	10125	5583	6973	5484	7379	9652	6194	7500,75	3,3%
chr17	864	519	644	752	601	949	1158	754	780,125	0,3%
chr5	771	847	616	672	564	868	1168	675	772,625	0,3%
chr13	646	1234	948	695	425	656	912	533	756,125	0,3%
chr2	428	169	327	158	328	473	805	247	366,875	0,2%
chr7	142	260	227	167	268	472	792	428	344,5	0,2%
chr11	272	335	364	203	171	340	487	256	303,5	0,1%
chr3	88	148	130	71	210	316	535	186	210,5	0,1%
chr4	127	109	93	61	215	343	470	188	200,75	0,1%
chr6	65	68	83	45	141	273	441	142	157,25	0,1%
chr8	63	56	71	35	160	247	429	149	151,25	0,1%
chrX	82	79	50	51	140	157	414	143	139,5	0,1%
chr12	73	60	57	49	142	206	379	117	135,375	0,1%
chr10	51	69	49	27	141	216	364	144	132,625	0,1%
chr9	56	49	58	31	129	191	317	124	119,375	0,1%
chr14	36	42	46	24	99	165	232	85	91,125	0,0%
chr16	28	29	32	20	97	144	244	102	87	0,0%
chr19	24	22	20	16	78	164	232	73	78,625	0,0%
chr15	30	27	31	19	80	131	200	80	74,75	0,0%
chr18	30	38	35	23	92	111	184	73	73,25	0,0%
chr20	25	26	33	12	67	133	221	64	72,625	0,0%
chr21	23	40	22	12	48	81	126	47	49,875	0,0%
chr22	19	14	15	10	38	51	107	40	36,75	0,0%
chrY	20	19	21	19	12	54	43	26	26,75	0,0%
total	244710	278914	226703	238572	168277	212668	279951	177628	228427,875	100,0%

Tableau 9: répartition des lectures alignées sur les différents chromosomes **Rouge>jaune>vert**

L'alignement sur l' « *Hg19* » n'a pas engendré de perte quantitativement importante du nombre de lectures et n'impactait pas ou très peu sur les profondeurs de couverture. Néanmoins, ceci ne présage pas de la qualité de l'alignement et nous avons également évalué l'impact sur la détection des variants en cherchant un schéma d'erreur inhérent à l'alignement sur l' « *Hg19* » versus « *rCRS* », celui-ci sera évoqué dans le chapitre « détection des variants ».

1.5. La couverture

Les objectifs de couverture admis dans le cadre du diagnostic des mutations mitochondriales sont : 200 X pour les tissus notamment musculaires et 500 X pour les leucocytes circulants afin de pouvoir détecter les mutations hétéroplasmiques même à des taux faibles.

Au total, tous les échantillons séquencés dans cette étude avaient un profil de couverture très similaire, à l'exclusion des patients porteurs de grands réarrangements non équilibrés, observés à la migration de la PCR longue, qui altéreraient inévitablement le profil de couverture en fonction des régions délétées.

Ce profil bien qu'identique entre tous les patients, était très hétérogène. Les taux de couverture variaient selon les régions du génome mitochondrial avec des profondeurs allant en moyenne de 854 X à 3935 X (pour les 3 runs confondus). Le tableau 10 résume les moyennes de profondeurs, de lectures maximales et minimales pour chacun des trois runs (à l'exclusion des patients porteurs de délétions). La figure 47 représente les profondeurs de lecture en fonction des régions du génome mitochondrial (1-16569). Celle-ci illustre bien la grande hétérogénéité de couverture et marque la sur-représentativité de certaines régions par rapport à d'autres.

	run3	run2	run1	3 runs confondus
Min	1133,06	796,87	546,97	854,89
Max	5117,40	4027,34	2965,73	3935,64

Tableau 10 : moyennes des maximales et minimales des profondeurs de lectures pour les trois « runs »

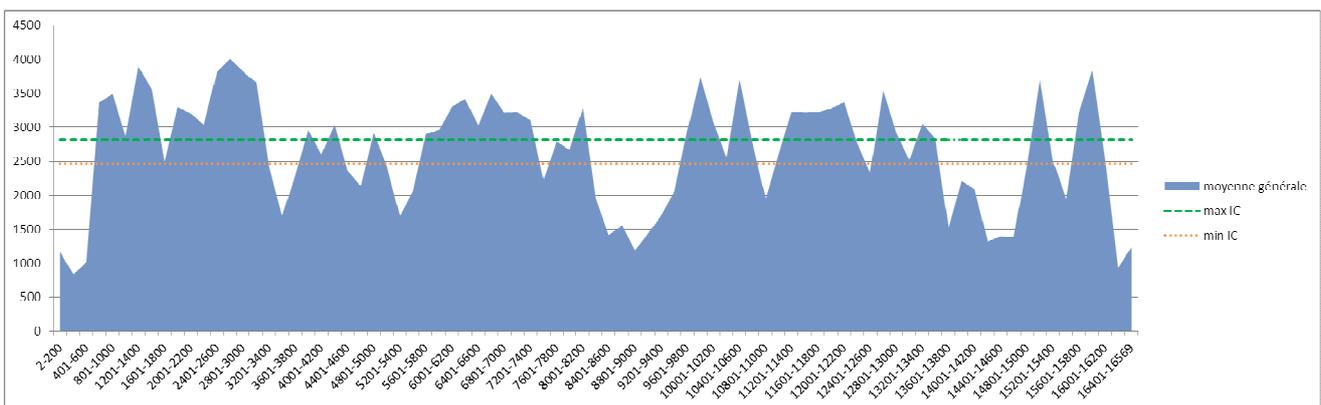


Figure 47 : Taux de couverture le long du génome mitochondrial (1-16569pb) : moyenne des trois runs de mises au point

Max IC-min IC = intervalle de confiance à 95% de la moyenne des profondeurs

Malgré cette hétérogénéité, la couverture obtenue était très bonne puisque nous avons atteint sur le dernier run d'optimisation (puce n°3) des profondeurs ≥ 200 X et ≥ 500 X pour respectivement 100% et plus de 99% des bases (hors aspects de grandes délétions à la PCR, en caractère **gras** dans le tableau) (tableau 11)

	P15	P16	P17	P18	P19	P20	P21	P22	P23
moyenne de profondeur de lecture	2062	5032	3359	3233	2068	3440	3260	3440	3285
couverture à 1x	100%	100%	100%	100%	100%	95%	100%	100%	100%
couverture à 20x	100%	100%	100%	100%	100%	95%	95%	100%	100%
couverture à 100x	100%	100%	100%	100%	100%	90%	95%	100%	100%
couverture à 200x	100%	100%	100%	100%	100%	80%	83%	100%	100%
couverture à 500x	99,42%	99,99%	99,90%	99,98%	99,75%	55,93%	42,78%	99,99%	99,99%

Tableau 11 : récapitulatif des couvertures pour les patients séquencés du « run » n°3, en caractère gras les patients porteurs de délétions

Au final, nous avons gardé comme seuil de diagnostic une profondeur de 500 X car il n'est pas possible en routine diagnostique de faire des « runs » séparés pour les prélèvements de sang et de tissu, ceci engendrerait des délais d'attente trop longs. En effet, le centre de Bordeaux a un recrutement moyen d'une centaine d'échantillons par an et donc moins d'une dizaine par mois (prélèvements de muscle et de sang confondus), ce qui n'est pas suffisant pour faire un « run » par type de prélèvements.

Lors de ce travail, nous avons séquencé de manière aléatoire aussi bien des échantillons de sang que de muscle issus de différents patients, en fonction de leur disponibilité dans le laboratoire. Il n'y avait pas de différences évidentes observées entre les résultats de séquençage des deux types de prélèvements, notamment en ce qui concerne la couverture et la qualité. Nous n'avons cependant pas approfondi cet aspect de l'étude ni comparé directement les résultats issus du séquençage de muscle et de sang d'un même patient. Ceci aurait été un meilleur reflet de la réalité en éliminant la variabilité interindividuelle. Par ailleurs, il serait intéressant d'évaluer dans un second temps le séquençage par cette technique d'autres tissus ou cellules que le muscle et le sang sur EDTA.

→Les causes de l'hétérogénéité de couverture du génome mitochondrial par NGS

Nous nous sommes intéressés aux régions moins bien couvertes pour essayer d'identifier les facteurs en cause. Il semblerait qu'un grand nombre de ces régions se caractérise par un biais de brin plus important, le plus souvent à cause d'une sous-représentation du brin sens (Forward), ceci est illustré dans la figure 48.

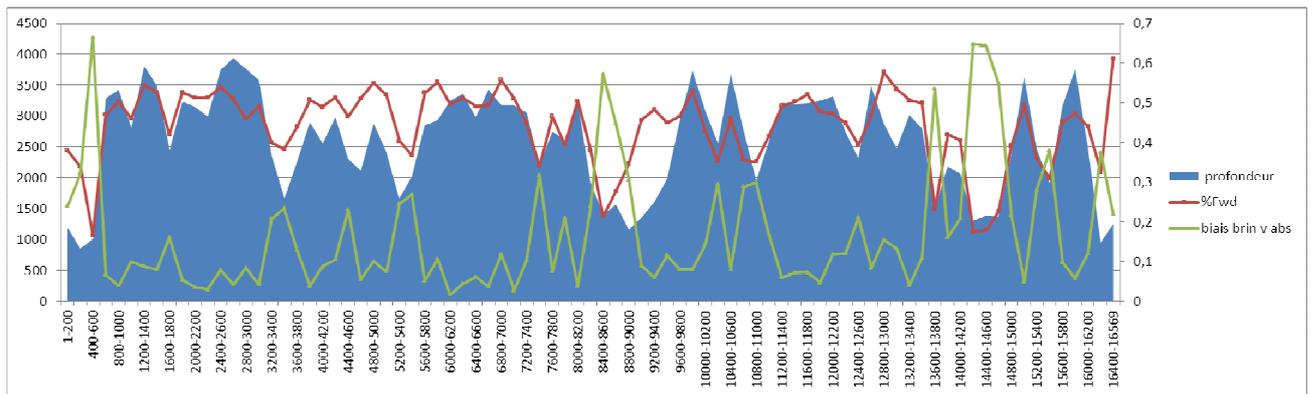


Figure 48 : Visualisation graphique par région génomique des profils de couverture (aire en bleu), de pourcentage des reads en forward (rouge) et du biais de brin (vert). Les régions de couvertures très basses sont celles qui possèdent le plus grand biais de brin, majoritairement par perte de reads forwards aux dépens des reverses

Plusieurs équipes ont en réalité rapporté cette hétérogénéité de séquençage avec le même profil que celui décrit ici, et ce quel que soit le type d'appareil, de technologie (Illumina, Life technologies), de technique de fragmentation des PCRs (sonication , digestion enzymatique) ou d'algorithme d'alignement. A chaque fois, les patterns de couverture étaient similaires et ressemblaient aux résultats que nous avons obtenus (figure 49)

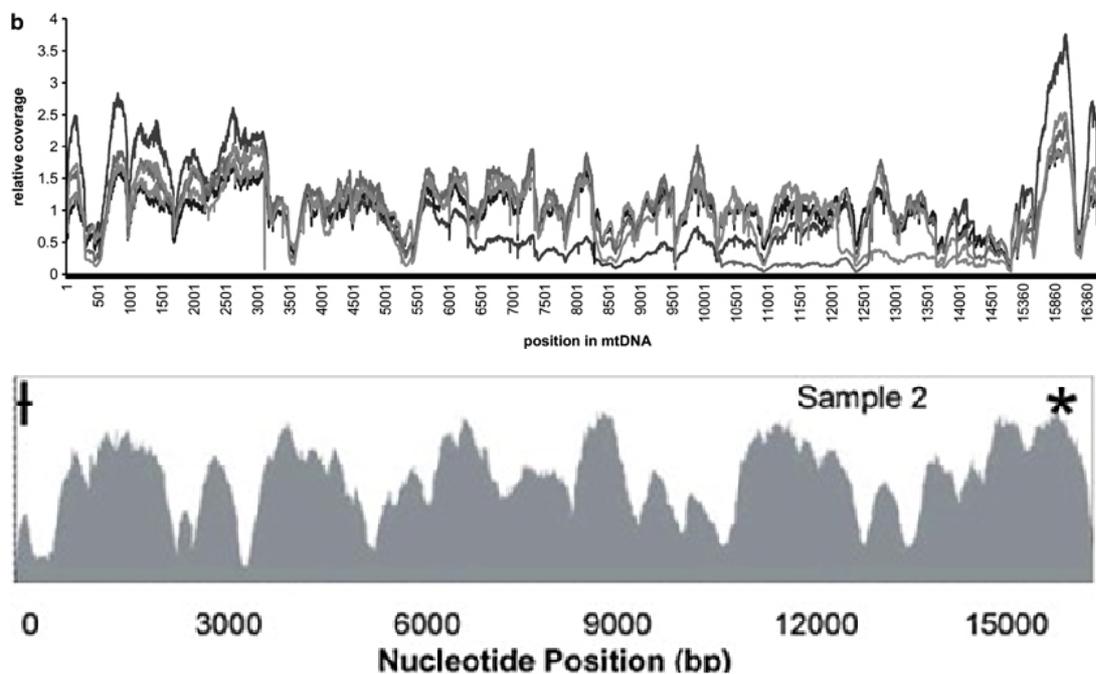


Figure 49 : distribution des lectures le long de l'ADNmt avec le séquençage par PGM (SENECA *et al.* 2015) (haut) et Miseq (bas) (MCELHOE *et al.* 2014)

Les régions où le taux de couverture chutaient ne semblaient pas être plus riches en GC, et nous n'avons pas trouvé de corrélation entre le contenu en GC et la profondeur de lecture (figure 50) . Ceci avait déjà été démontré par d'autres équipes et n'est guère surprenant puisque l'ADNmt a la particularité d'avoir deux brins de composition différente, avec pour la même séquence un brin léger (ou forward) riche en C et en regard un brin lourd (ou reverse) riche en G.

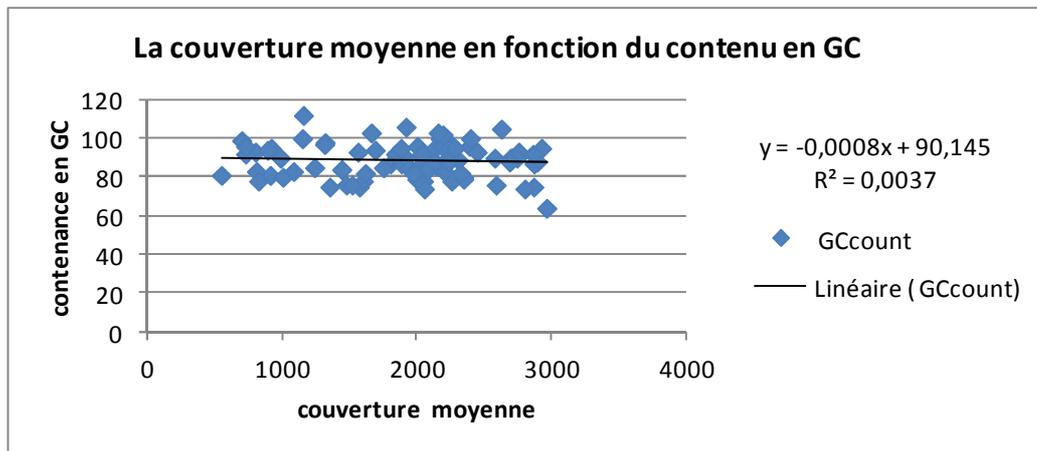


Figure 50 : Aucune corrélation entre les moyennes de couverture et le contenu en GC pour chaque région

L'autre caractéristique du génome mitochondrial est sa richesse en homopolymères et particulièrement en cytosine polyC (polyG sur le brin reverse) (SEO *et al.* 2015) ont comparé le contenu des régions moins bien couvertes aux autres, celles-ci semblent être particulièrement riches en séquences polyC (figure 51). Néanmoins, l'existence d'un réel lien causal entre les deux observations reste à prouver, et le mécanisme par lequel la présence de ces homopolymères influencerait le séquençage n'est pas clair.

Table 1 Comparison of homopolymers between high coverage and low coverage areas across 24 samples

Homopolymer type	High coverage areas ¹	Low coverage areas ²
	Number of Homopolymers	Number of Homopolymers
AA	276	320
CC³	275	347
GG	167	81
TT	238	225
AAA	104	95
CCC	76	124
GGG	31	12
TTT	82	42
AAAA	22	33
CCCC	31	54
GGGG	4	2
TTTT	18	12
AAAAA	10	11
CCCCC	8	28
GGGGG	2	0
TTTTT	4	3
AAAAAA	4	6
CCCCCC	1	8
GGGGGG	1	0
TTTTTT	2	1
AAAAAAA	2	2
CCCCCCC	0	1
TTTTTTT	1	0
AAAAAAA	0	1

Figure 51 : Contenu en homopolymères de C des régions peu couvertes versus région bien couvertes (SEO *et al.* 2015) (**High coverage vs low coverage**)

Par ailleurs, l'utilisation d'autres algorithmes d'alignement ne change pas non plus le profil de couverture, les « reads » non alignés à cause d'une qualité insuffisante semblent se distribuer sur le même pattern que les « reads » de bonne qualité (MCELHOE *et al.* 2014). Il ne semble donc pas y avoir de perte d'alignement prépondérante au niveau de ces régions, le problème se situerait plutôt au niveau de la réaction de séquençage elle-même et ce quelle que soit la technologie utilisée.

De plus, les « reads » manquants sont en majorité des « forward » riches en séquence polyC, on peut donc supposer que le séquençage des molécules composées de polyG est difficile et que ces régions sont peut-être incompatibles avec le séquençage. Il est clair que la technique d'enrichissement par PCR longue s'affranchit des limites de la faisabilité de la PCR pour chaque région prise isolément, contrairement à l'amplification par PCRs multiplexes. Néanmoins, ceci se répercute probablement à l'étape du séquençage en lui-même avec la perte des couvertures au niveau de ces régions. L'hétérogénéité n'est cependant pas gênante pour le séquençage de l'ADNmt en routine diagnostique, car n'empêche pas d'avoir une profondeur suffisante comme décrit au chapitre précédent.

1.6. La détection des réarrangements et leurs points de cassure

L'amplification de tout le génome mitochondrial par une PCR longue permet d'orienter vers l'existence de grands réarrangements non équilibrés par la visualisation sur gel des différentes tailles de fragments générés. Le séquençage en « whole genome » des produits de PCR permet ensuite d'obtenir des contigs et de délimiter ainsi les points de cassures (figure 52).

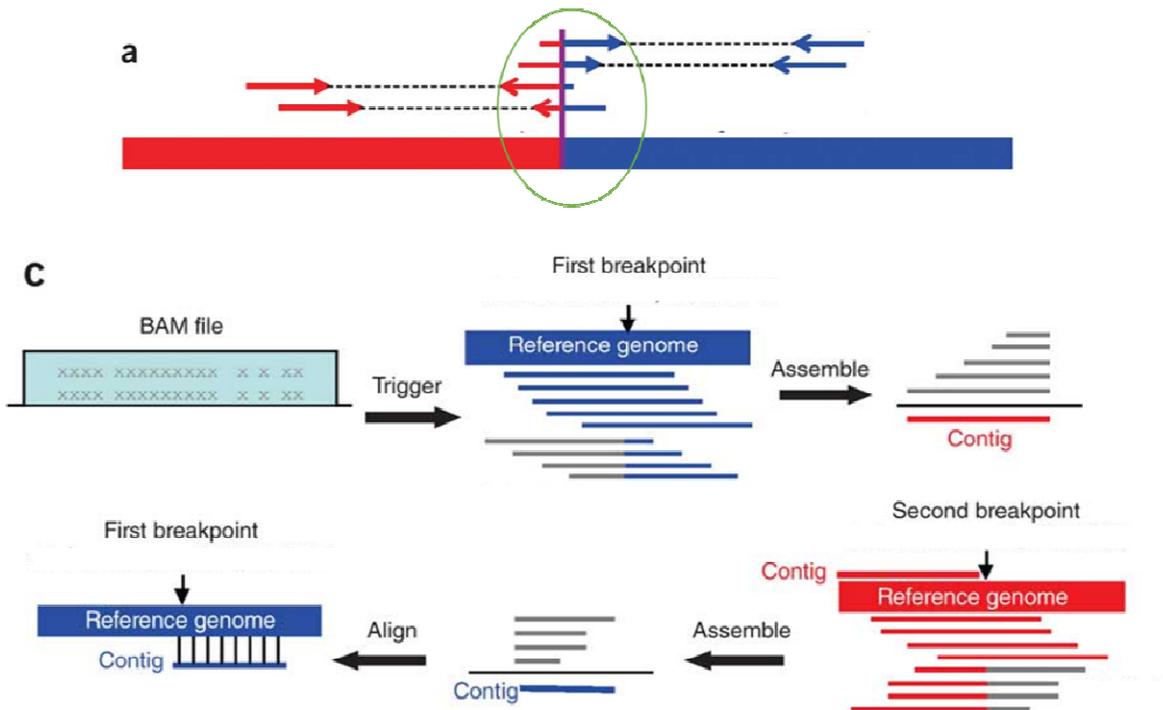


Figure 52: Réassemblage des reads situés au niveau des points de cassure ayant une correspondance partielle avec la séquence de référence pour reconstituer la région du remaniement (BARTENHAGEN and DUGAS 2015)

→Les délétions uniques :

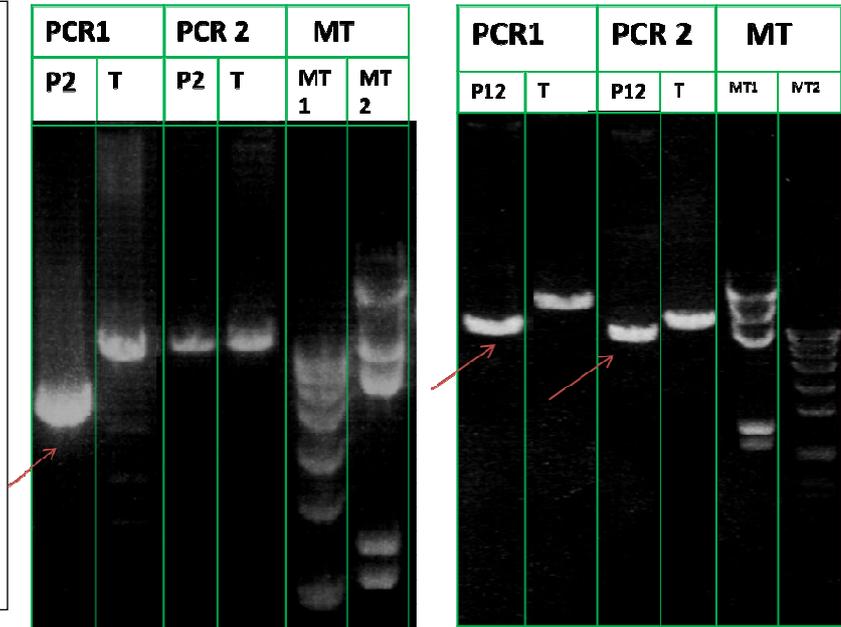
Nous avons testé, dans notre série deux patients (P2 et P12) porteurs chacun d'une grande délétion simple connue, identifiée par les techniques conventionnelles (PCRs + séquençage Sanger) (figure 53).

Figure 53 :

Migration sur gel d'agarose 1% des PCRs longues des deux patients P2 et P12 : délétion de $\approx 8\text{kb}$ et 5kb respectivement. On note un seul fragment d'une plus petite taille

Pour le P2 : amplification d'un fragment normal par la PCR 2 car le couple d'amorces est situé dans la région de la délétion \rightarrow sélection des molécules non délétées par la PCR

PCR 1=446F-16620R, PCR 2=9391F-9027R, T= témoin normal, MT=marqueur de taille



Les profils de lectures pour ces deux patients se distinguaient de ceux des autres par une chute nette de la profondeur de lecture au niveau de la région délétée (figures 54 et 55). La visualisation sur *Alamut*[®] des « reads » alignés au niveau de ces régions a montré une correspondance partielle avec la séquence de référence (existence de mésappariements sur une grande partie du read); leur réassemblage a permis ensuite une identification précise des points de cassure (figure 56).

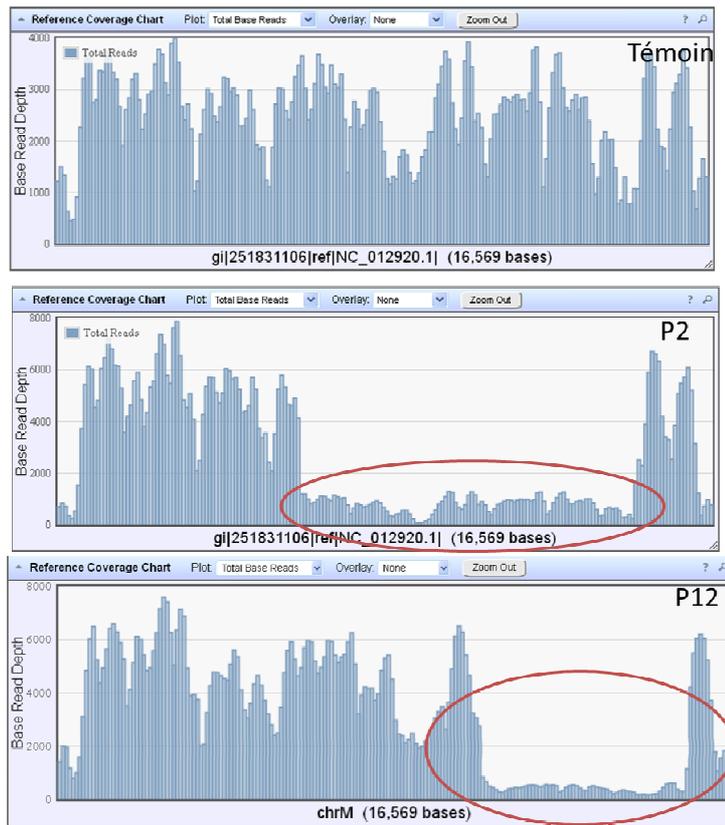


Figure 54 : Les couvertures de séquençage tels que visualisés sur le logiciel Torrent Suite pour les deux patients porteurs de délétions par rapport à un témoin normal

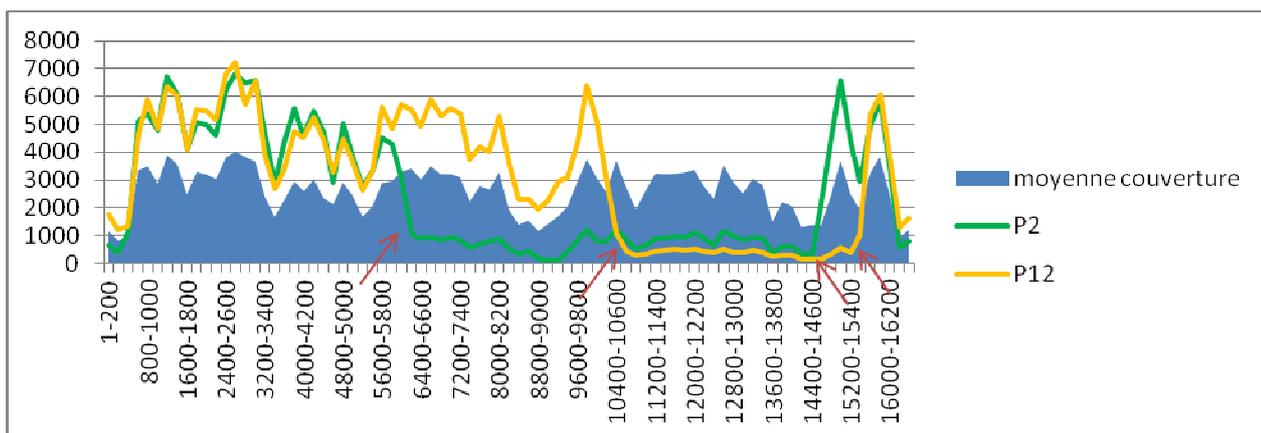


Figure 55 : Graphique montrant le pattern de couverture des deux patients porteurs d'une délétion par rapport à la moyenne globale de couverture tout le long du génome mitochondrial : on note la chute brutale de couverture au niveau des points de cassures (flèches)

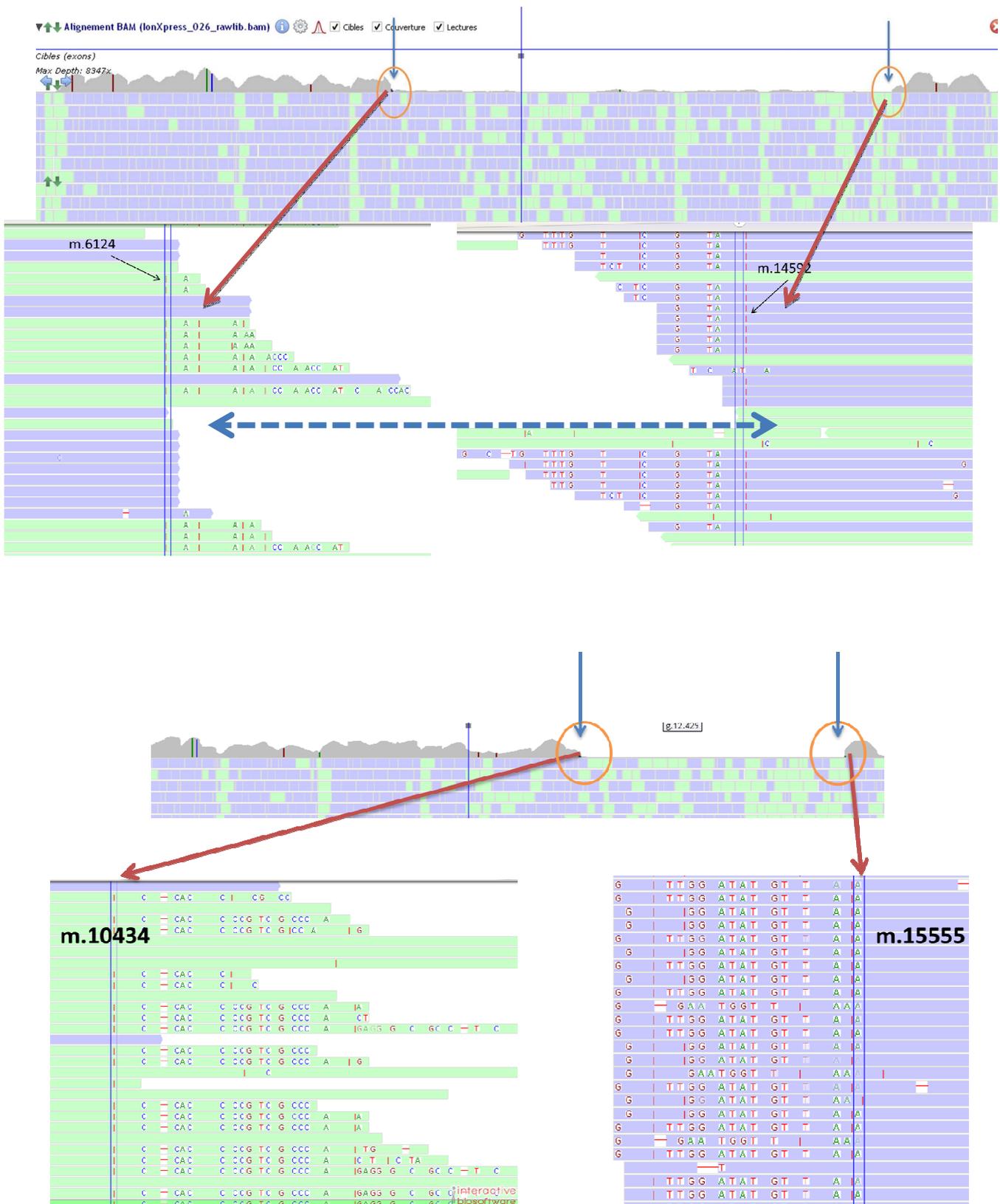


Figure 56 : Visualisation Alamut® : le réassemblage des reads mal alignés a permis l'identification des points de cassures (m.6124-14592) et (m.10435-m.15555) pour les patients P2 (haut) et P12 (bas)

→Les délétions multiples :

Dans notre série, un seul patient (P3) était porteur d'un syndrome de délétions multiples par mutation dans le gène *POLG1*. Dans ce cas, la migration sur gel d'agarose des produits de PCR a mis en évidence la présence d'au moins trois fragments de tailles différentes. Ceci s'est reflété sur les résultats de séquençage par un profil de couverture altéré par rapport à celui du témoin normal (figure 57)

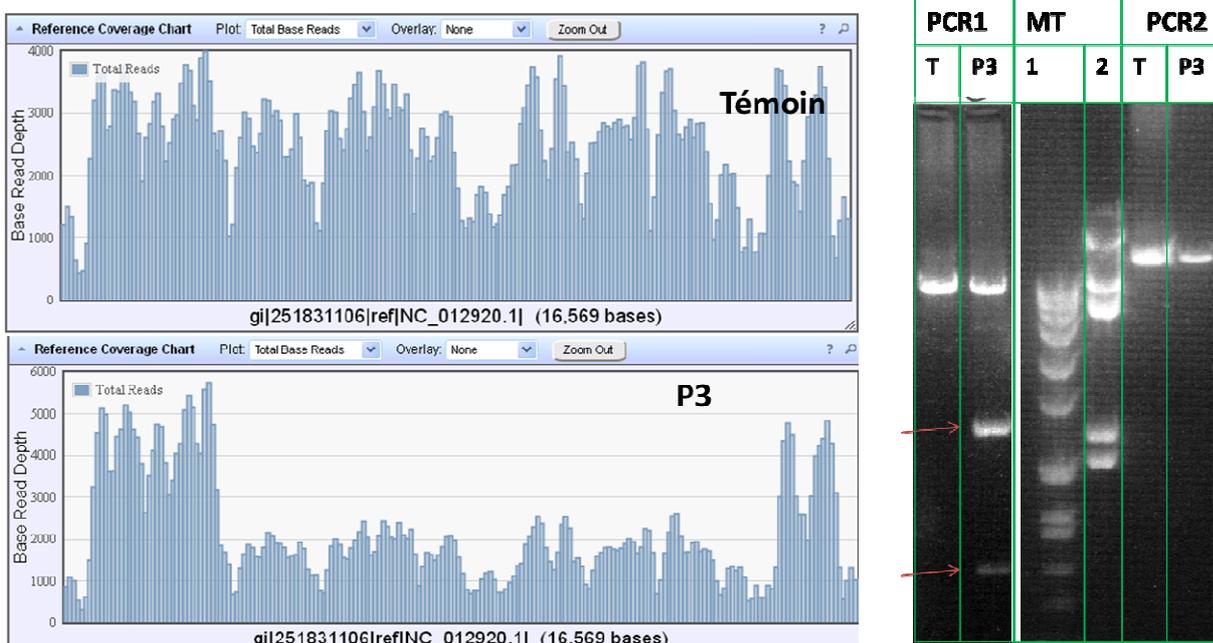


Figure57 : Profil altéré de couverture correspond à la présence de plusieurs fragments à la migration de la PCR.A) Visualisation de la distribution des lectures sur le logiciel *Ion Torrent* ®P3 (délétions multiples) versus témoin normal. B) la migration sur gel des deux PCR1 et 2 pour le P3 et un témoin sans délétions montre la présence de plusieurs fragments intermédiaires amplifiés en plus de l'amplicon normal à 16Kb

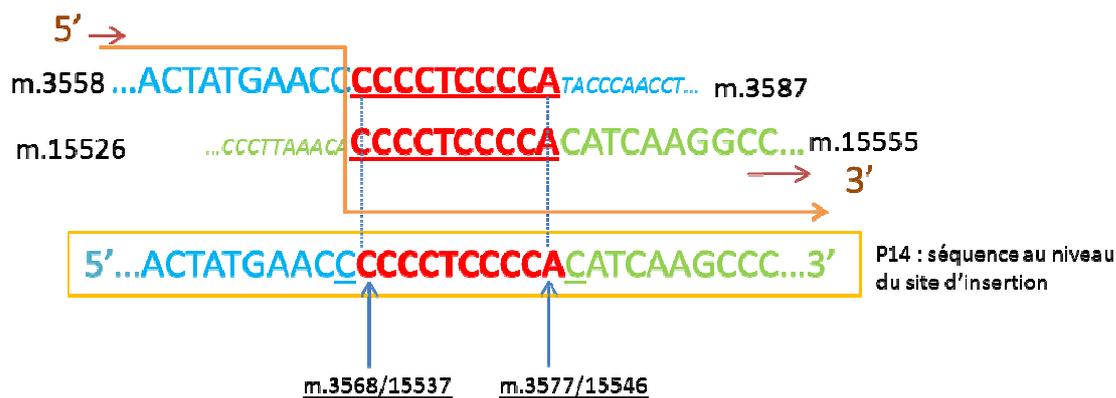
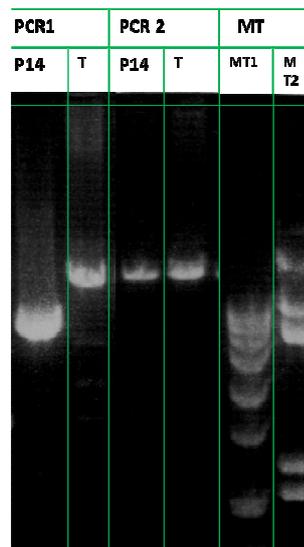


Figure 59 : Visualisation de la séquence au niveau du site de l'insertion médiée par la région répétée 5'ccctcccca3'

L'amplification par PCRs longues de l'ADNmt du patient P14 est illustrée dans la figure ci-dessous (figure). L'amplification grâce au couple d'amorce 9391F-9027R n'a pas permis d'obtenir la molécule de 21kb, seules les molécules normales ont été amplifiées (fragment de ≈16kb visualisé sur le gel). L'amplification par le couple d'amorce 446F-16220R, a quant à elle généré un seul fragment de petite taille (≈4 kb) (figure 60).

Figure 60 : Migration des PCRs longues du P14 versus un témoin T
 PCR 1=446F-16620R, PCR 2=9391F-9027R,
 T= témoin normal, MT=marqueur de taille



Après séquençage, il est apparu que ce fragment correspondait en réalité aux séquences m.446-3567 et m.15537-16220 (≈3+1 kb), comme le montre la visualisation des fichiers .BAM dans *Alamut*® (figure 61). Au niveau des régions de chute de profondeurs, les séquences correspondaient à la fusion des deux segments et contenaient le point d'insertion m.3868-15537.

La technique de NGS est limitée dans l'identification des réarrangements complexes dont les duplications/insertions par les limites d'amplification de la PCR longue, elle-même conditionnée par la taille et la position de l'insert. Par ailleurs, le fait que l'ADNmt soit une molécule circulaire rajoute un degré de complexité et ne permet pas toujours l'identification des positions relatives des différentes séquences.

Le Southern Blot reste l'examen clé pour le diagnostic des duplications du génome mitochondrial mais le NGS pourrait intervenir en complément, dans un deuxième temps pour l'identification rapide des points de cassure, grâce à l'assemblage des contigs.

1.7. Détection des variants et hétéroplasmie

1.7.1. Efficacité de détection des variations connues

En plus des patients porteurs de délétions, quatre patients dont les mutations ponctuelles étaient connues ont été testés afin de s'assurer de leurs détections par NGS. De plus, afin d'évaluer les seuils de détection de l'hétéroplasmie nous avons sectionné trois patients avec des taux variables de la mutation m.3243G>A : 60%, 3% et 12% pour respectivement les patients P1, P4 et P5 (Tableau 12)

Au total, les mutations homoplasmiques et hétéroplasmiques avec des taux élevés ou moyens ont été retrouvées à chaque fois par NGS.

La détection de la mutation m.3243G>A à un taux de 3% demeurait par contre aléatoire. Le patient P4 porteur de cette mutation a été étudié deux fois sur deux « runs » différents, le premier a permis de détecter la mutation alors qu'elle n'apparaît pas dans le deuxième.

Patient	Tissu	Run	mutations	taux	VCF (NGS)	
					fréquence	Quality
P1	S	run1	m.3243G>A	Hétéroplasmie : 60%	87.3%	22584.1
P4	M	run1	m.3243G>A	Hétéroplasmie : 3%	3.0%	12.22
P5	S	run1	m.3243G>A	Hétéroplasmie : 12%	11.6%	548.9
P6	S	run1	m.11778G>A	homoplasmique	100%	31664.7
P4bis	M	run2	m.3243G>A en duplicat	hétéroplasmie: 3%	<i>Non détectée</i>	

Tableau 12 : détection des mutations connues par NGS

Nous avons essayé de mettre en évidence les facteurs qui ont empêché sa détection dans le deuxième run, malgré un seuil de détection identique du « variant calling » à 2%. En effet, la visualisation du fichier .BAM du patient P4 sur *Almut*® a bien objectivé l'existence de la substitution A>G dans 3% des « reads » (figure 63).



Figure 63 :*ALAMUT*® on note bien la présence de 3% de G au niveau du site m.3243 dans les deux runs

Nous avons essayé plusieurs combinaisons de filtre possible du « variant calling » en modifiant les paramètres de sélection. Nous avons par exemple autorisé un biais de brin total et des seuils de qualité très bas sans aucun résultat, ce variant restait indétectable.

En réalité la sélection des variants par le système Torrent Suite® se fait en deux étapes, la première consiste en la recherche de toutes les variations par rapport à la séquence de référence. Dans un second temps, cette liste globale est filtrée et seuls les variants qui correspondent aux critères de sélection appliqués par le « variant calling » sont appelés dans le VCF généré. Ainsi, il est possible de visualiser la liste complète des variants non sélectionnés, dits « NO CALL » ainsi que leur paramètre de qualité et les raisons de leur exclusion.

Le fichier des variants non appelés « NO CALL » du patient P4 a été étudié, ce dernier ne contient pas la substitution m.3243G>A ce qui signifie qu'elle n'est pas identifiée par le système comme étant une variation par rapport à la séquence de référence (figure 64).

68	chrM	398	.	C	CC	40.4666	NOCALL	AF=0;AO=93;DP=
69	chrM	420	.	C	CT	93.3554	NOCALL	AF=0;AO=65;DP=571
70	chrM	432	.	A	AC	18.1513	NOCALL	AF=0;AO=55;DP=527
71	chrM	1689	.	C	A	152.29	NOCALL	AF=0;AO=68;DP=
72	chrM	2150	.	T	TA	166.248	NOCALL	AF=0;AO=589;DP=
73	chrM	3105	.	AC	A	16.9897	NOCALL	AF=0;AO=4174;DP=
74	chrM	3166	.	T	C	147.184	NOCALL	AF=0;AO=138;DP=
75	chrM	3167	.	TC	T,TT	84.3892	NOCALL	AF=0;AO=
76	chrM	3174	.	TA	T	345.146	NOCALL	AF=0;AO=624;DP=
77	chrM	3179	.	G	T	173.404	NOCALL	AF=0;AO=154;DP=
78	chrM	3343	.	AT	A	402.934	NOCALL	AF=0;AO=509;DP=
79	chrM	3390	.	AG	A	425.196	NOCALL	AF=0;AO=474;DP=
80	chrM	3715	.	GCCC	GCC,GCCA	27945.4	NOCALL	A
81	chrM	3959	.	GC	G	422.754	NOCALL	AF=0;AO=425;DP=
82	chrM	4136	.	A	AC	320.54	NOCALL	AF=0;AO=242;DP=
83	chrM	4175	.	G	GA	162.143	NOCALL	AF=0;AO=239;DP=
84	chrM	5281	.	C	CA	53.301	NOCALL	AF=0;AO=226;DP=
85	chrM	5302	.	TC	T	210.817	NOCALL	AF=0;AO=357;DP=
86	chrM	5407	.	T	A	147.876	NOCALL	AF=0;AO=105;DP=
87	chrM	5746	.	GA	G	270.293	NOCALL	AF=0;AO=231;DP=
88	chrM	5884	.	AT	A	301.313	NOCALL	AF=0;AO=397;DP=
89	chrM	6004	.	GCC	G	236.213	NOCALL	AF=0;AO=154;DP=
90	chrM	6293	.	TCC	TC,T	286.143	NOCALL	AF=0;AO=
91	chrM	6419	.	AC	A	275.244	NOCALL	AF=0;AO=417;DP=
92	chrM	6956	.	TG	T	428.625	NOCALL	AF=0;AO=1261;DP=
93	chrM	7422	.	GAA	AGA,GA,GAC	66.0395	NOCALL	AF=0;AO=
94	chrM	7535	.	AT	A	419.517	NOCALL	AF=0;AO=272;DP=
95	chrM	7819	.	C	T	164.302	NOCALL	AF=0;AO=116;DP=
96	chrM	8027	.	GC	G	344.818	NOCALL	AF=0;AO=342;DP=
97	chrM	8118	.	GT	G	171.491	NOCALL	AF=0;AO=454;DP=
98	chrM	8232	.	TA	T	305.005	NOCALL	AF=0;AO=555;DP=
99	chrM	8271	.	AC	A	227.146	NOCALL	AF=0;AO=413;DP=
100	chrM	8280	.	AC	A	233.552	NOCALL	AF=0;AO=293;DP=
101	chrM	8405	.	AC	A,ACC	140.919	NOCALL	AF=0,0;AO=
102	chrM	8495	.	A	C	16.6063	NOCALL	AF=0.0119048;DP=
103	chrM	8896	.	GC	G	87.6743	NOCALL	AF=0;AO=65;DP=
104	chrM	8902	.	GC	G	91.4379	NOCALL	AF=0;AO=73;DP=

Figure 64 : Fichier « No Call » du patient P4 : on remarque l'absence du variant sur le site m.3243

L'erreur de détection intervient probablement en amont de l'étape de filtre ou « variant calling ». Les raisons de sa non visibilité ne sont pas encore résolues, nous sommes à priori aux limites de détection du NGS ou du moins du système Ion Torrent®. Tout ceci peut poser un réel problème diagnostique dans le cadre de la recherche de mutations mitochondriales du fait de l'hétéroplasmie. En effet, il est important de détecter les mutations hétéroplasmiques même à taux faibles dans un tissu donné car celles-ci peuvent être plus élevées dans d'autres tissus et peuvent donc être pathogènes. Il semble donc important de toujours vérifier la composition des « reads » au niveau des positions des mutations les plus fréquentes de l'ADNmt afin de s'assurer de leur absence.

Nous ne sommes pas allés plus loin dans le cadre de la validation de détection des variants car le système Ion Torrent a déjà été évalué par plusieurs équipes pour la recherche de SNP et d'indels (BRAGG *et al.* 2013) (QUAIL *et al.* 2012). Il est aujourd'hui largement utilisé dans le cadre du diagnostic moléculaire en pathologie humaine à travers diverses applications dont l'étude de l'ADNmt (SENECA *et al.* 2015) (SEO *et al.* 2015) (PARSON *et al.* 2013). Les taux d'erreurs admis sont assez faibles pour les SNP inférieurs à 1 % mais plus de l'ordre de 1 à 2 % pour les indels (BRAGG *et al.* 2013).

De manière générale, l'utilisation du PGM® appliquée au re-séquençage d'ADN « DNA resequencing » a permis de mettre en évidence quelques erreurs systématiques inhérentes au système Ion Torrent®. Ces erreurs de séquençage étaient surtout en rapport avec la difficulté à bien identifier les séquences de répétitions et d'homopolymères (surestimation des petites séquences d'homopolymères et sous-estimation des plus longues). Le PGM se caractérise aussi par un excès de faux positifs d'indels qui peuvent être présentes avec des fréquences de lectures assez hautes pouvant être trompeuses. Certaines semblaient être redondantes sur des sites dits de prédisposition (BRAGG *et al.* 2013) en fonction des génomes étudiés mais d'autres pouvaient être aléatoires et non reproductibles entre les « runs ».

Dans notre série, quelques erreurs systématiques de séquençage n'ont pas été filtrées correctement par le « variant calling ». Il s'agissait majoritairement d'erreurs liées aux régions d'homopolymères et de répétitions dont les deux principales sont la région polymorphique d'homopolymère de C en m.311 et la séquence polyA en m.4604 (tableau 13). Une substitution nucléotidique récurrente a également été observée (variation m.8860A>C). Au niveau de ce site, le nucléotide de référence est un A alors qu'il s'agit en réalité d'un polymorphisme plus rare que la base C, mais qui a tout de même été gardée dans la « rCRS ». De ce fait, le variant m.8860A>G a été retrouvé chez tous les patients. Certains (13) présentaient également le variant m.8860A>C à une fréquence très faible (<5%) et un biais de brin majeur était observé systématiquement (prédominance des « reads » forwards).

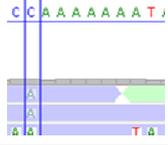
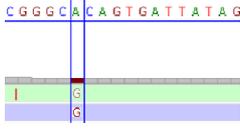
site	séquence	couverture	Score qualité	N°P
Site polymorphique polyC autour de m.310 m.310T>C		Fréquence des lectures <10% Biais de brin +++ aux dépend du Reverse (Fw=0)	Bons >1000	27
m.4604C>A Régions d'homopolymères A		Fréquence >10% des lectures Biais de brin++ (Rev=0)	Mauvais <20	19
m.8860A>C		Fréquences <5% des lectures Biais de brin (Rev=0)	Bons >1000	13

Tableau 13 : erreurs récurrentes de séquençage

1.7.2. Les effets de l'alignement sur la détection des variants :

Afin d'évaluer l'influence de l'alignement sur la détection des variants, nous avons systématiquement comparé les fichiers VCF issus des deux types d'alignement (« rCRS versus « Hg19 »), pour un même échantillon, et ce à partir des données de deux « runs » différents (run 1 et 3) (tableau 14).

Les différences de détection concernaient à chaque fois les variations à taux très faible, toutes ayant une fréquence $\leq 4\%$ et/ou des scores de qualité bas. Dans la plupart des cas, le « variant calling » a détecté plus de variations à partir de l'alignement sur l'« Hg19 » que celui réalisé que la « rCRS », mais celles-ci étaient majoritairement des erreurs récurrentes de séquençage (m.4604C>A et m.8860A>C). Il est à noter, cependant que la mutation m.3243G>A du patient P4 n'a été détectée dans le « run » n°1 qu'à partir des « reads » alignés sur l'« Hg19 », celle-ci était absente du fichier VCF généré à partir de l'alignement sur la « rCRS ».

Globalement, la détection de ce type de variations, avec des scores de qualité et des fréquences dans les seuils limites, semble être aléatoire et n'est pas toujours reproductible entre les « runs » et les processus du « signal processing ». Ceci est illustré par la détection de la mutation m.3243G>A dans un « run » et pas l'autre.

	Hg19	rCRS
P1	m.4604 C>A	
P2	m.8860A>C	
P3	m.4604 C>A m.8860A>C	
P4	m.3243G>A m.8860A>C	
P5	Pas de différence	
P6	m.8860A>C	m.308C>T
P7	m.4604 C>A m.8860A>C	
P8	m.4604 C>A m.8860A>C	
P15	m.4604 C>A	
P16	m.4604 C>A m.8860A>C m.12418insA	
P17	m.8860A>C	m.565G>A
P18	m.4604 C>A	
P19		m.6692A>G
P20	m.4604 C>A	
P21	Pas de différence	
P22	m.8860A>C	

Tableau 14 : la différence entre les variants détectés avec l'alignement hg19 versus rCRS

Le mécanisme par lequel la recherche de variants à partir de « reads » alignés sur l'Hg19 génère plus de variations (erreurs ou non) n'est pas clair. Ces erreurs systématiques plus importantes pourraient être dues à un alignement moins précis au niveau de ces régions, notamment dans la séquence d'homopolymères A en m.4604. Néanmoins, il n'y avait pas plus d'indels avec l'alignement sur l'« Hg19 », et la détection n'était pas moins efficace que ce qui est obtenu après alignement sur la « rCRS », autorisant ainsi son utilisation en routine

2. Deuxième partie : recherche de nouvelles mutations

2.1. Caractéristiques cliniques des patients de la série

Pour la deuxième partie de cette étude nous avons sélectionné 22 patients avec une forte suspicion de cytopathie mitochondriales : 4 enfants et 18 adultes, dont les caractéristiques clinico-biologiques sont résumées dans le tableau ci-dessous (tableau 15)

Patient	Année de naissance	Sexe	année de prélèvement	présentation clinique	histologie	biologie	chaîne respiratoire	imagerie cérébrale et SRM
P9	1049	M	2011	plusieurs épisodes neurologiques : AVC multiples ?	5COX-	NR	NR	NR
P10	1948	M	2012	PEO, fatigabilité à la 30aine, troubles de déglutition à 60 ans, Sd extra-pyramidal à 63 ans	qq RRF, nombreuses COX-	NR	NR	DATscan dénévation dopaminergique
P11	1959	M	2013	PEO	10 RRF	L : N, CK : x2N	NR	NR
P13	1963	M	2013	surdit�, n�phropathie, diab�te, contexte familial (m�re diab�tique)	8 RRF et 6 COX-	L, NH3, CK : N	NR	NR
P15	1956	M	2000	Surdit�, myopathie, atteinte du SNC : Sd c�r�belleux progressif, sd tetrapyramidal, d�mence	Nb RRF et COX-	NR	NR	NR
P16	2002	M	2002	hypotrophie � la naissance, d�tresse respiratoire et acidose lactique insuffisance h�patocellulaire et r�nale	dizaine RRF et nb COX-	acidose lactique	d�ficit combin� de tous les complexes	NR
P17	1963	F	2004	enc�phalopathie progressive depuis l'enfance avec �pilepsie partielle	7RRF et nb COX-	hyperlactacid�mie persistance	NR	atrophie c�r�brale diffuses
P18	1996	F	2005	retard du d�veloppement moteur, ataxie c�r�belleuse progressive, cardiomyopathie hypertrophique	NR	hyperlactacid�mie 5mmol/l	d�ficit des complexes I et IV	hypoplasie c�r�belleuse gauche

P19	1993	M	2006	myopathie : fatigabilité	présence de mitochondries anormales au ME	NR	NR	NR
P20	1922	F	2007	SNC : chorée; ataxie statique et discrètement cinétique.	3 RRF et 10 COX-	CK : x2-3N	NR	TDMc N pour trauma
P21	1946	M	2007	Epilepsie généralisée, ptosis, surdit�, diab�te.	Mitochondries anormales, d�p�ts fuschinophiles sous le sarcolemme d'une fibre musculaire 8 COX-	L : N	NR	IRM: leuco-enc�phalopathie post�rieure.
P22	2001	F	2008	retard psychomoteur �voluant avec DI, surdit� mixte, tubulopathie, atteinte musculaire	nb COX- et fibres avec de discr�ts d�p�ts fuscino-philos sous sarcolemmal et un d�but d'aspect d�chiquet�	L : N	NR	L�sions �vocatrices de cytopathies mitochondriales (NR?)
P23	1951	M	2009	Ataxie c�r�belleuse Par�sie des nerfs oculomoteurs, faiblesse MI	4RRF nb COX-	L pr�P : N L postP : 2,38	NR	
P24	2008	F	2009	surdit�, perception compl�te gauche et retard psychomoteur	Normale	hyperlactaci�mie persistante rapport L/P et BHB/AA �lev�	d�ficit des complexes I et IV	quelques hypersignaux T2 et Flair de la substance blanche p�ri-ventriculaire en regard des carrefours ventriculaires
P25	1957	F	2010	Epilepsie r�fractaire depuis l'enfance, AVC multiples � l'�ge	1RRF 8COX-	NR	NR	NR

				adulte. EMG: NP sensitive axonale + tracés myogènes				
P26	1932	M	2010	paralysie supranucléaire progressive et une atteinte neurovégétative (malaises, troubles vésico-sphinctériens).	3RRF et 3COX-	L, CK : N	NR	TDMc N pour vertiges
P27	1943	M	2010	ataxie neuropathie périphérique surdité	3COX-	L : N, 1 épisode rhabdomyolyse	NR	DATscan : dénervation dopaminergique
P28	1941	F	2013	Camptocormie , myopathie	3RRF 9COX-	NR	NR	NR
P29	1973	F	2011	Ataxie cérébrale avec atrophie optique et neuropathie périphérique	10 COX- RRF très remaniées	NR	NR	hypotrophie cérébelleuse et moëlle cervicale
P30	1948	F	2013	Démence, dépression, multiples AVC, diabète IR de début < 35 ans	NR	Lsg=2,2 ; L/P=11,7 L lcr=2,4	NR	multiples AVC / microangiopathie
P31	1971	F	2014	Paresthésies hémicorporelles G, troubles de concentration	discrète atteinte myogène, 4 COX-	L, CK : N synthèse intratéchale d'Ig	NR	Sd de Fahr : calcification des NGC : striato-palido-dentelées
P32	1992	M	2014	AVC à 12 et 16 ans, troubles de l'attention, difficultés scolaires	NR	NR	NR	2 AVC, ACSC

Tableau 15 : résumé des renseignements cliniques et paracliniques chez les patients inclus dans cette étude. AA : acétoacétate ; ACSC : atrophie cortico-sous-corticale ; BHB : beta-hydroxybutyrate ; CK : creatines kinases ; COX : cytochrome c oxydase ; L : lactate ; N : normal ; NH3 : ammoniac ; P : pyruvate ; préP /postP : pré et post-prandial ; RRF : ragged-red fibers.

2.2. Résultats des PCR longues

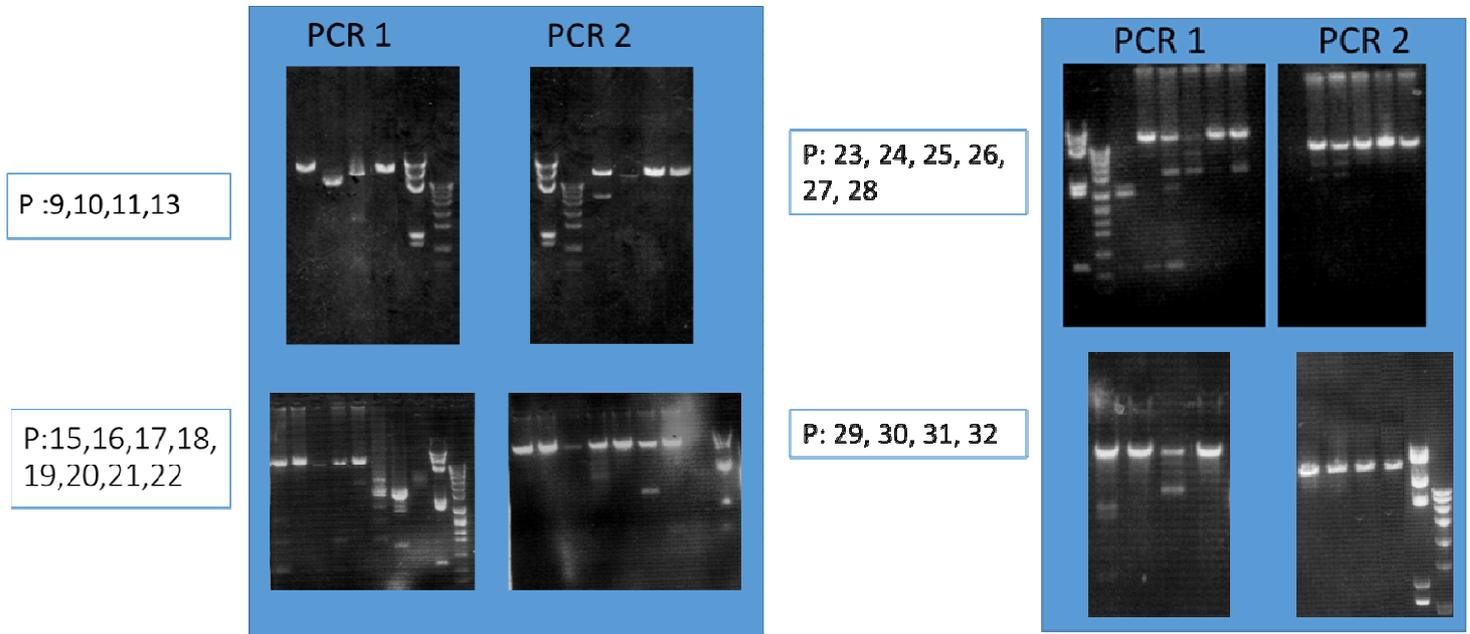


Figure 65 : Migration des PCR longues des patients à étudier, PCR1 =446F-16220R, PCR2=9391F-9022R

De manière assez surprenante, certains patients avaient des profils de migration de PCR anormaux alors qu'ils avaient déjà été testés dans le laboratoire à la recherche de grands réarrangements avec d'autres couples d'amorces qui n'avaient pas montré d'anomalies.

Ainsi, nous avons classé ces patients selon les profils de migration en trois catégories :

- ✓ Normales avec un seul fragment ~16kb dans les deux PCRs, incluant également l'existence d'un fragment d'intensité faible (notés normales + dans le tableau)
- ✓ Aspect de délétions multiples avec plusieurs fragments intermédiaires classiquement à la PCR1 associée à une PCR2 faibles (parfois normales ou avec un petit fragment additionnel en plus de celui à 16Kb)
- ✓ Des aspects plus atypiques avec : une suspicion de déplétion devant l'intensité plus faible des bandes dans les deux PCRs (P17), l'existence d'un smear à la PCR1 et l'absence d'amplification de la PCR2 (P22 et 23)
- ✓

Patient P	Profils PCRs
9	<i>Normales +</i>
10	Délétion simple ou multiple?
11	Délétions multiples?
13	<i>Normales</i>
15	<i>Normales +</i>
16	<i>Normales</i>
17	Déplétion?
18	<i>Normales+</i>
19	<i>Normales +</i>
20	Délétions multiples?
21	Délétions multiples?
22	Atypique : smears PCR1, rien à la PCR2
23	Atypique : smears PCR1, rien à la PCR2
24	<i>Normales</i>
25	Délétions multiples?
26	Délétions multiples?
27	<i>Normales</i>
28	<i>Normales +</i>
29	<i>Normales+</i>
30	<i>Normales</i>
31	<i>Normales+</i>
32	<i>Normales</i>

Tableau 16 : résumés des profils de PCRs chez les 22 patients testés

Toutes les PCRs anormales ont été re-testées à plusieurs reprises, notamment en présence de deux témoins, un normal et un deuxième positif pour les délétions multiples, afin de nous assurer de l'efficacité de la PCR. La qualité de l'ADN a également été évaluée par contrôle direct de l'ADN total par migration puis par dosage au Qubit®, ce dernier avait pour intérêt de doser spécifiquement les molécules doubles brins et donc d'exclure une dénaturation importante des molécules. Il n'y avait pas de différence de qualité ou de quantité entre les ADNs avec des PCRs anormales et les autres. A priori la qualité de l'ADN n'était pas à l'origine de ces profils de PCR altérés.

L'interprétation des profils de PCRs longues peut être difficile. En effet, les fragments intermédiaires de PCR s'observent dans les syndromes de délétions multiples par déficit de la stabilité du génome mitochondrial (mutations dans les gènes nucléaires tels que *POLG*, *PEO1*...), mais ils peuvent aussi être le résultat d'une accumulation physiologique de molécules d'ADNmt délétées notamment en raison de l'âge ou d'un artefact de PCR (KAJANDER *et al.* 1999). Ces profils posent donc un réel problème diagnostique, d'autant que les délétions multiples ne sont pas toujours détectées par techniques de Southern Blot à cause de leurs faibles taux, et ce qu'elles soient réelles (pathologiques ou non) ou artéfactuelles.

Ne pouvant exclure des causes secondaires à l'origine des anomalies de PCR dans notre série de patients, nous avons décidé de séquencer tous les échantillons, y compris ceux avec des profils de PCRs altérés, pour exclure l'existence de mutations de l'ADNmt à responsables du phénotype.

2.3. Résultats du séquençage de nouvelle génération

2.3.1. Les variants générés et fichiers de sortie

La recherche de variants sur tout le génome mitochondrial a généré de 15 à 59 variants par patient.

Les variants étaient produits par le logiciel Torrent Suite® sous deux formats de fichiers de sortie possibles :

Format VCF : fichier texte tabulé, comprenant le chromosome, la position, le nucléotide de référence et celui détecté, intégrant ensuite pour chaque variation des paramètres de qualité de couvertures, les informations du filtre...

Fichier Excel (figure 66) : il fournit les mêmes informations que le VCF mais sous la forme d'un tableau. Il a l'avantage de pouvoir se lire directement contrairement aux fichiers VCF qui utilisent des caractères spécifiques pour encoder l'information.

Les fichiers .VCF ont surtout l'avantage dans notre cas de pouvoir être ouverts dans le logiciel Alamut®, permettant une visualisation rapide des variants annotés.

The image shows a screenshot of an Excel spreadsheet titled "alleles_IonXpress_025 - Excel (Échec de l'activation du produit)". The spreadsheet contains a large table of genetic variant data. A red magnifying glass icon is positioned over the first few columns of the table, with a red arrow pointing downwards from the handle of the magnifying glass. The table has the following columns: A (Chrom), B (Position), C (Ref), D (Variant), E (Allele Call), F (Filter), G (Frequency), H (Quality), I (Filter), J (Type), K (Allele Sour), L (Allele Nam), M (Gene ID), N (Region), O (Nan), P (VCF Positio), Q (VCF Ref), R (VCF Variant), S (Original Co), T (Coverage), U (Filter), and V (Coverage). The data rows start with "chrM" and include various variant details such as position, reference, variant, and quality scores.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
Chrom	Position	Ref	Variant	Allele Call	Filter	Frequency	Quality	Filter	Type	Allele Sour	Allele Nam	Gene ID	Region	Nan	VCF Positio	VCF Ref	VCF Variant	Original Co	Coverage	Filter	Coverage
chrM	73	A	G	Homozygou	-	100.0	15023.5	-	SNP	Novel	---	unknown	chromosom	73	A	G	947	949	-	-	-
chrM	263	A	G	Homozygou	-	100.0	11902.9	-	SNP	Novel	---	unknown	chromosom	263	A	G	746	745	-	-	-
chrM	292	T	A	Heterozygo	-	5.0	26.63	-	SNP	Novel	---	unknown	chromosom	292	T	A	508	502	-	-	-
chrM	310	T	C	Heterozygo	-	27.2	2461.63	-	SNP	Novel	---	unknown	chromosom	309	CT	CC	123	228	-	-	-
chrM	311	-	C	Heterozygo	-	72.8	2461.63	-	INS	Novel	---	unknown	chromosom	309	CT	CTC	123	228	-	-	-
chrM	489	T	C	Homozygou	-	100.0	9685.76	-	SNP	Novel	---	unknown	chromosom	489	T	C	614	614	-	-	-
chrM	514	-	CA	Heterozygo	-	85.9	8473.49	-	INS	Novel	---	unknown	chromosom	513	G	GCA	726	711	-	-	-
chrM	709	G	A	Homozygou	-	100.0	31938.2	-	SNP	Novel	---	unknown	chromosom	709	G	A	4575	1993	-	-	-
chrM	750	A	G	Homozygou	-	100.0	31772	-	SNP	Novel	---	unknown	chromosom	750	A	G	5022	1991	-	-	-
chrM	1303	G	A	Homozygou	-	100.0	31604.8	-	SNP	Novel	---	unknown	chromosom	1303	G	A	3338	1987	-	-	-
chrM	1438	A	G	Homozygou	-	100.0	31943.9	-	SNP	Novel	---	unknown	chromosom	1438	A	G	2847	1998	-	-	-
chrM	1888	G	A	Homozygou	-	100.0	29750.2	-	SNP	Novel	---	unknown	chromosom	1888	G	A	1875	1871	-	-	-
chrM	2706	A	G	Homozygou	-	100.0	31838.6	-	SNP	Novel	---	unknown	chromosom	2706	A	G	2883	1994	-	-	-
chrM	3243	A	G	Heterozygo	-	87.3	22584.1	-	SNP	Novel	---	unknown	chromosom	3243	A	G	1816	1830	-	-	-
chrM	3921	C	T	Homozygou	-	98.7	31078.6	-	SNP	Novel	---	unknown	chromosom	3921	C	T	2456	1999	-	-	-
chrM	3954	C	T	Homozygou	-	100.0	31604.6	-	SNP	Novel	---	unknown	chromosom	3954	C	T	2211	1987	-	-	-
chrM	4604	C	A	Heterozygo	-	6.5	18.09	-	SNP	Novel	---	unknown	chromosom	4604	C	A	65	77	-	-	-
chrM	4769	A	G	Homozygou	-	100.0	7785.92	-	SNP	Novel	---	unknown	chromosom	4769	A	G	486	489	-	-	-
chrM	4916	A	G	Homozygou	-	99.1	31267.8	-	SNP	Novel	---	unknown	chromosom	4916	A	G	2619	1998	-	-	-
chrM	6461	A	G	Homozygou	-	100.0	31848.2	-	SNP	Novel	---	unknown	chromosom	6461	A	G	2204	1994	-	-	-
chrM	7028	C	T	Homozygou	-	100.0	19152.7	-	SNP	Novel	---	unknown	chromosom	7028	C	T	1197	1196	-	-	-
chrM	8701	A	G	Homozygou	-	100.0	14173.9	-	SNP	Novel	---	unknown	chromosom	8701	A	G	894	896	-	-	-
chrM	8860	A	G	Heterozygo	-	81.2	129.47	-	SNP	Novel	---	unknown	chromosom	8860	A	G	16	16	-	-	-
chrM	9540	T	C	Homozygou	-	99.9	25410.6	-	SNP	Novel	---	unknown	chromosom	9540	T	C	1607	1606	-	-	-
chrM	9833	T	C	Homozygou	-	100.0	31663.4	-	SNP	Novel	---	unknown	chromosom	9833	T	C	2929	1989	-	-	-
chrM	10398	A	G	Homozygou	-	100.0	31774.2	-	SNP	Novel	---	unknown	chromosom	10398	A	G	3152	1993	-	-	-
chrM	10400	C	T	Homozygou	-	100.0	31570.1	-	SNP	Novel	---	unknown	chromosom	10400	C	T	3173	1986	-	-	-
chrM	10873	T	C	Heterozygo	-	97.2	16499.6	-	SNP	Novel	---	unknown	chromosom	10873	T	C	999	1117	-	-	-
chrM	11719	G	A	Homozygou	-	100.0	31579.8	-	SNP	Novel	---	unknown	chromosom	11719	G	A	2575	1985	-	-	-
chrM	12477	T	C	Homozygou	-	100.0	23568.2	-	SNP	Novel	---	unknown	chromosom	12477	T	C	1471	1476	-	-	-
chrM	12705	C	T	Homozygou	-	100.0	31458.7	-	SNP	Novel	---	unknown	chromosom	12705	C	T	3003	1987	-	-	-
chrM	14323	G	A	Homozygou	-	100.0	9494.31	-	SNP	Novel	---	unknown	chromosom	14323	G	A	591	611	-	-	-
chrM	14766	C	T	Homozygou	-	100.0	11752	-	SNP	Novel	---	unknown	chromosom	14766	C	T	717	739	-	-	-
chrM	14783	T	C	Homozygou	-	98.9	12338.3	-	SNP	Novel	---	unknown	chromosom	14783	T	C	786	791	-	-	-

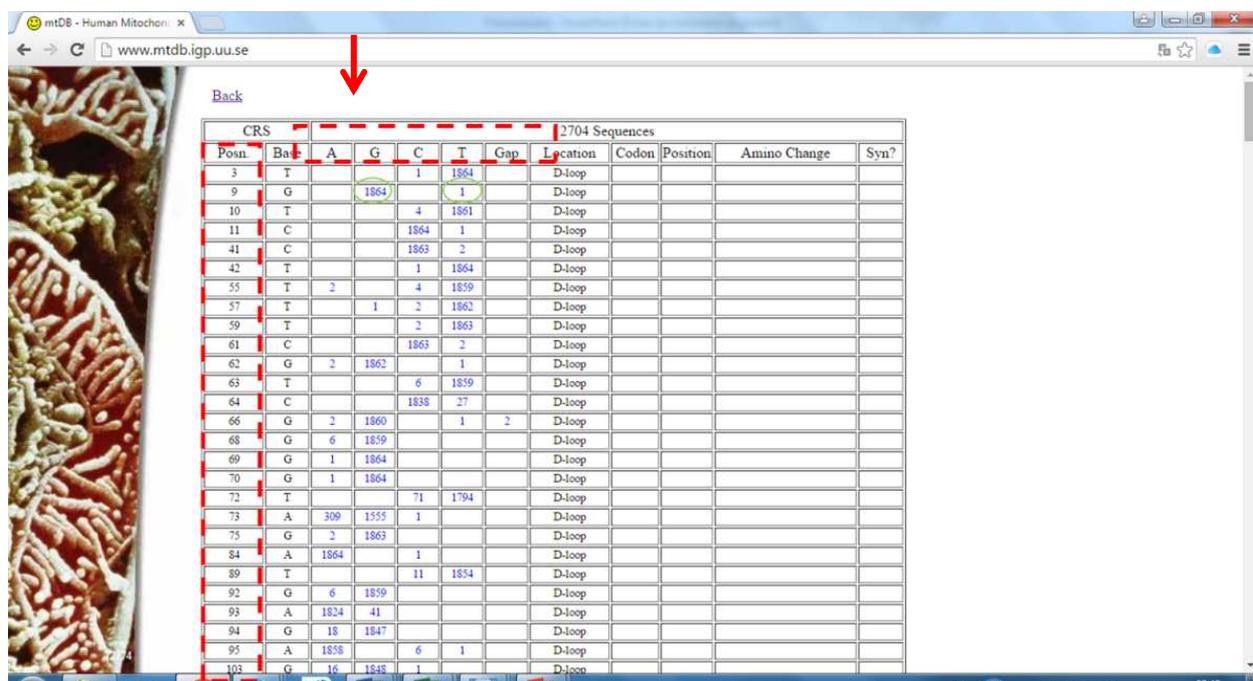
Figure 66 : présentation du fichier Excel de sortie des variants, nombre très important de colonnes et beaucoup d'information non triées

2.3.2. Tri des variants

Comme décrit dans la section « matériel et méthodes », le tri des variants s’est fait à partir des bases de données dédiées « mtDB » et « mitomap », consultables via internet. En effet, il n’existe pas à notre connaissance d’outil facile compatible avec l’explorateur Windows qui permette d’établir un lien direct entre ces bases et les fichiers de sortie, pour un tri rapide des variants (polymorphismes, mutations, variants inconnus), comme ce qui est le cas par exemple pour *Alamut*® et les bases de données de l’ADN nucléaire (dbSNP...).

Ce tri se fait donc au prix de manipulations manuelles pour interroger directement ces bases de données variant par variant.

Concernant « mtDB », celle-ci se présente sous la forme d’un tableau à deux entrées : les positions génomiques et les nucléotides (A,T,G,C), donnant pour chaque type de bases le nombre de fois où celle-ci a été détectée au niveau de chaque position. (figure 67)



Posn	Base	A	G	C	T	Gap	Location	Codon	Position	Amino Change	Syn?
3	T		1	1864			D-loop				
9	G		1864		1		D-loop				
10	T			4	1861		D-loop				
11	C			1864	1		D-loop				
41	C			1863	2		D-loop				
42	T				1	1864	D-loop				
55	T	2		4	1859		D-loop				
57	T		1	2	1862		D-loop				
59	T			2	1863		D-loop				
61	C			1863	2		D-loop				
62	G	2	1862				D-loop				
63	T			6	1859		D-loop				
64	C			1838	27		D-loop				
66	G	2	1860			1	2				
68	G	6	1859				D-loop				
69	G	1	1864				D-loop				
70	G	1	1864				D-loop				
72	T				71	1794	D-loop				
73	A	309	1555		1		D-loop				
75	G	2	1863				D-loop				
84	A	1864			1		D-loop				
89	T				11	1854	D-loop				
92	G	6	1859				D-loop				
93	A	1824	41				D-loop				
94	G	18	1847				D-loop				
95	A	1858			6	1	D-loop				
103	G	16	1848		1		D-loop				

Figure 67 : schématisation de la présentation des variants dans mtDB, tableau à double entrée : position génomique et base nucléotidiques (ATCG)

L'interrogation de mtDB en ligne est limitée à 10 variants par consultation nécessitant de répéter l'opération plusieurs fois pour chaque patient (figure). Le but est de connaître le nombre de fois où le variant a été observé dans la population étudiée (considère comme polymorphique si ≥ 4 occurrences $\sim 0.2\%$). Cette base fournit également des informations sur l'origine et l'haplogroupe des échantillons qui présentaient le variant (Figure 68)

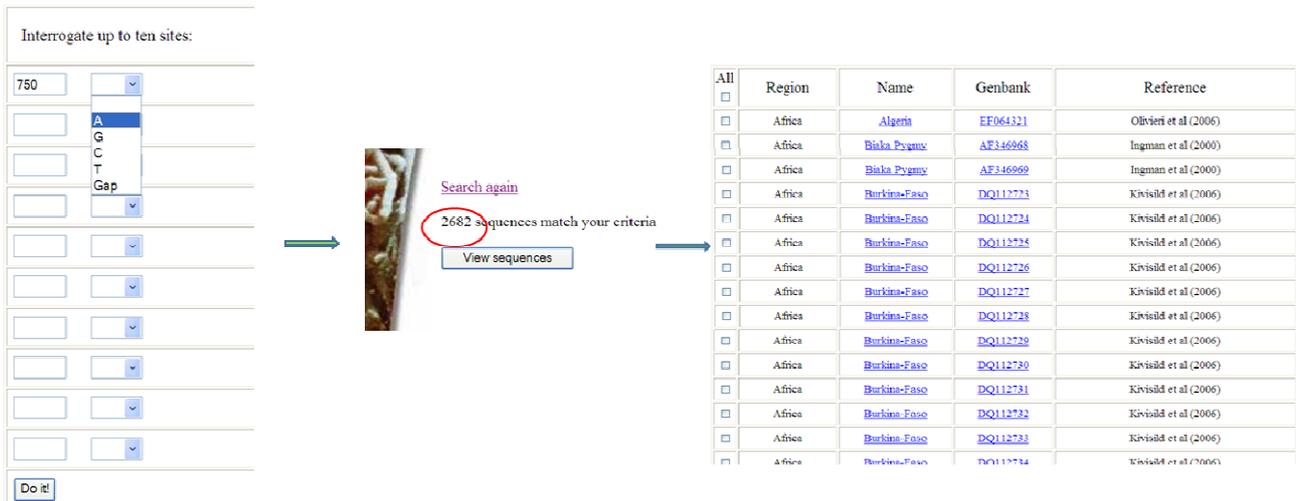


Figure 68 : requête puis visualisation d'un variant sur mtDB permettant d'avoir le nombre d'individus porteurs dans la base

La base « mitoamp » est consultable directement, elle permet la visualisation de toutes les variations présentes dans une séquence donnée (taille maximale = 100 pb à la fois), classées en polymorphismes, mutations constitutionnelles, mutations somatiques retrouvées dans les tissus tumoraux (figure 69).

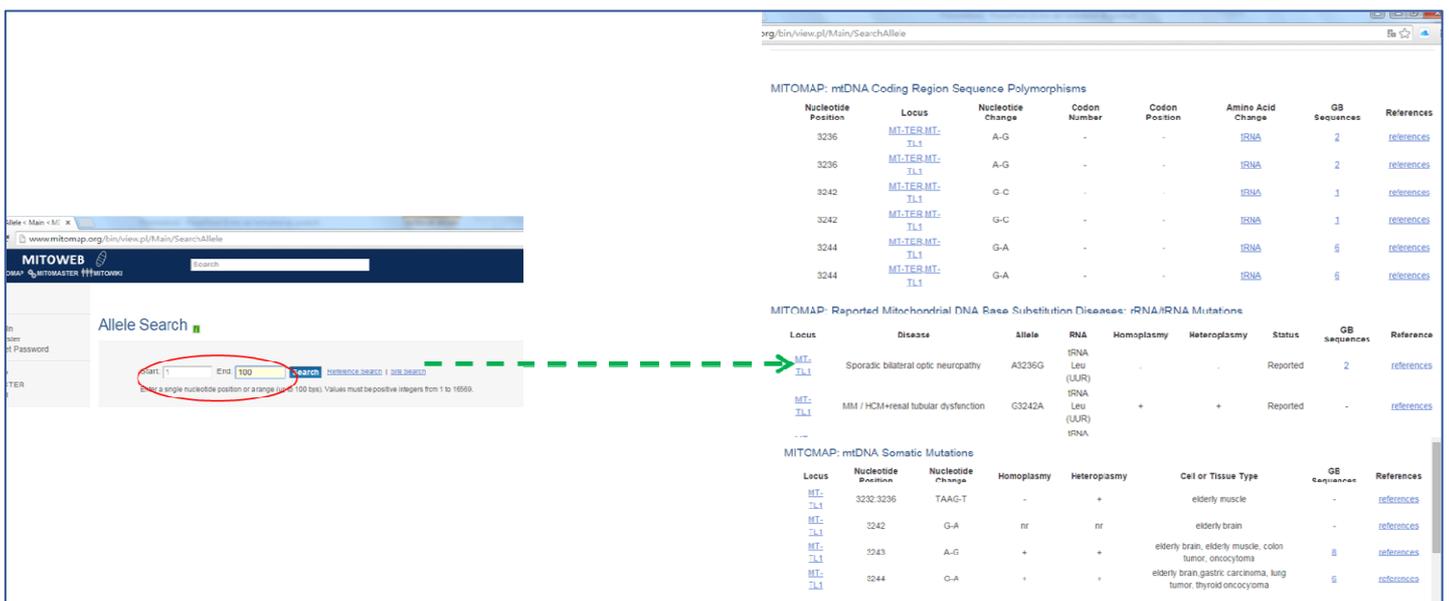


Figure 69 : schématisation de la requête puis visualisation des variants dans mitomap pour une région génomique donnée (max 100 pb)

Le site « mitomap » dispose également d'un module de recherche « mitomaster » qui est plus pratique car permet de consulter directement les données qui concernent une liste de variants d'intérêt, après importation dans un champ dédié (figure 70)

The screenshot shows the MITOMASTER web interface. At the top, there is a navigation bar with links for Mitomaster, Sequence Query, GenBank Query, SNV Query, Species, and Use Web API. Below this, a section titled 'Submit variant list as file upload or paste into text area' contains two options: 'Option 1' (Copy-paste) and 'Option 2' (Select a variant file). A text area under 'Option 1' contains a list of variants, with a green arrow pointing to it and the text 'Liste des variants d'intérêt importés'. A 'Submit' button is located below the text area. To the right of the text area, there are instructions on the format of the variant list. Below the 'Submit' button, there is a 'Processing sequences' status bar. At the bottom, a table displays the search results. The table has the following columns: Query, rCRS Position, Query Position, rCRS NT, Query NT, Mutation, Locus, tx Effect, GB Freq, Refs, Conservation, and Patient Report. The mutation 'm.3243G>A' is highlighted with a red circle, and the corresponding 'Patient Report' text is also circled in red.

Query	rCRS Position	Query Position	rCRS NT	Query NT	Mutation	Locus	tx Effect	GB Freq	Refs	Conservation	Patient Report
query	6461	0	A	G	transition	COI	syn:W=>W	21 (0.07%)	4 refs	100.00%	
query	10398	0	A	G	transition	ND3	non-syn:T=>A	13400 (43.81%)	212 refs 2 unpub	51.11%	PD protective factor / longevity / altered cell pH / metabolic syndrome / breast cancer risk
query	3954	0	C	T	transition	ND1	syn:A=>A	17 (0.06%)	2 refs	91.11%	
query	489	0	T	C	transition	ATT/D-Loop	non-coding	7627 (24.93%)	83 refs 1 unpub		
query	8860	0	A	G	transition	ATPase6	non-syn:T=>A	30265 (96.94%)	120 refs 5 unpub	71.11%	
query	10873	0	T	C	transition	ND4	syn:P=>P	10091 (32.99%)	83 refs 3 unpub	20.00%	
query	1888	0	G	A	transition	16S	rRNA	1713 (5.60%)	25 refs 1 unpub	6.67%	
query	12477	0	T	C	transition	ND5	syn:S=>S	208 (0.68%)	13 refs 1 unpub	97.78%	possible HCM susceptibility
query	3243	0	A	G	transition	L(UUA/G)TER	tRNA	8 (0.03%)	325 refs 1 unpub	97.78%	MELAS / LS / DMDF / MIDD / SNHL / CPEO / MM / FSGS / ASD / Cardiac+multi-organ dysfunction
query	4916	0	A	G	transition	ND2	syn:L=>L	43	4 refs	28.89%	

Figure 70 : étapes des requêtes puis de visualisation du tableau des résultats via mitomaster. Notez l'absence de hiérarchisation (entouré en rouge la mutation m.3243G>A)

Les résultats dans « mitomaster » se présentent sous la forme d'un tableau récapitulant les principales informations disponibles pour chaque variant. Cette présentation fournit une somme importante de données mais elle a pour inconvénient l'absence de hiérarchisation et une mauvaise visualisation des mutations. Ainsi dans notre exemple, la mutation m.3243G>A (cf figure) est noyée au milieu de la liste. Il s'agit donc de données brutes qu'il faudra investiguer/détailler variant par variant. Il s'agit d'une base de données très large incluant des sources différentes qui ne sont pas toujours affinées (BANDELT *et al.* 2009).

Tout ceci illustre bien la difficulté du tri de la quantité importante de variants générés par l'utilisation des technologies à haut débit plus exhaustives par rapport aux techniques conventionnelles ciblées. Cette procédure comporte beaucoup d'étapes de manipulations manuelles et peut être longue d'autant plus que les interfaces internet sont souvent assez lentes.

2.3.3. Outil de tri « maison »

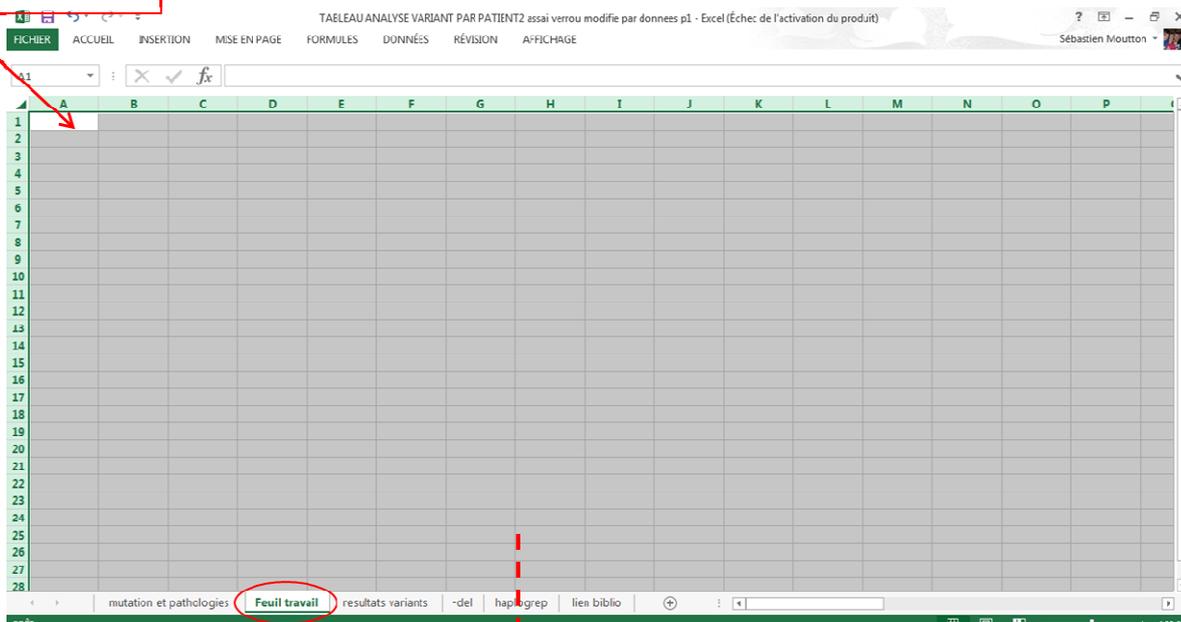
Nous avons donc décidé de mettre au point un outil de tri « maison » via Excel, grâce à différentes formules. Pour cela nous avons tout d'abord constitué une base de données locale en important les données du tableau « mtDB », ainsi que les mutations (« confirmed » et « reported ») de « mitomap ». Nous avons ensuite créé une interface pour consulter la base, toujours via Excel.

→ Premier filtre des polymorphismes / mutations connus

Un classeur verrouillé de travail a été créé, comportant une feuille d'entrée, servant à importer le fichier Excel des variants, à la sortie du logiciel Torrent Suite®. Ceci permet de générer automatiquement une feuille de résultats (figure 71) qui comprend :

- ✓ La liste des variants triés par surbrillance en :
 - VERT : les variants fréquents $\geq 0,2\%$ dans « mtDB » (benins)
 - ROUGE : les mutations confirmées dans « mitomap »
 - ORANGE : Les variants présents dans « mtDB » mais $< 0,2\%$
 - ROSE : les mutations reportées (non confirmées)
 - NOIR : les variants n'appartenant à aucune des catégories précédentes
- ✓ La fréquence, la couverture, le biais de brins et le score qualité pour chaque variant, permettant de résumer le tableau Excel de sortie, et d'alerter sur les scores ou les couvertures trop bas.

Import fichier Excel de sortie du serveur Torrent Suite®



Position	Ref	Variant	mutations	Frequence	Qualité	couverture	biais de brin
73	A	G	A73G	100	15023,5	949	0,5
263	A	G	A263G	100	11902,9	745	0,5
292	T	A	T292A	5	26,63	502	0,005
310	T	C	T310C	27,2	2461,63	228	0,5
311	-	C	-311C	72,8	2461,63	228	0,5
489	T	C	T489C	100	9685,76	614	0,5
514	-	CA	-514CA	85,9	8473,49	711	0,8795
709	G	A	G709A	100	31938,2	1993	0,5
750	A	G	A750G	100	31772	1991	0,5
1303	G	A	G1303A	100	31604,8	1987	0,5
1438	A	G	A1438G	100	31943,9	1998	0,5
1888	G	A	G1888A	100	29750,2	1871	0,5
2706	A	G	A2706G	100	31838,6	1994	0,5
3243	A	G	A3243G	87,3	22584,1	1830	0,5074
3921	C	T	C3921T	98,7	31078,6	1999	0,5042
3954	C	T	C3954T	100	31604,6	1987	0,5
4604	C	A	C4604A	6,5	18,09	77	0,9976
4769	A	G	A4769G	100	7785,92	489	0,5
4916	A	G	A4916G	99,1	31267,8	1998	0,5008
6461	A	G	A6461G	100	31848,2	1994	0,5
7028	C	T	C7028T	100	19152,7	1196	0,5
8701	A	G	A8701G	100	14173,9	896	0,5
8860	A	G	A8860G	81,2	129,47	15	0,5242
9540	T	C	T9540C	99,9	25410,6	1606	0,5008
9833	T	C	T9833C	100	31663,4	1989	0,5
10398	A	G	A10398G	100	31774,2	1993	0,5

codes couleurs:
 vert polymorphisme retrouvé >0,2% mtDB
 rouge mutation confirmée dans mitomap
 orange variant retrouvé <0,2% vuiejamais dans mtDB
 rose variant rapporté comme possiblement pathogène mitomap reported
 NB :tjs vérifier l'absence de redondance dans la liste des variats du patient

Feuille de résultats

Figure 71 : démonstration des variants tels que visualisés ans la feuille de résultats dédié après import du fichier Excel de sortie des variants

Le but de cet outil « maison » est d'obtenir un tri rapide des variants avec une visualisation claire et une interface simple d'utilisation en routine. Cette opération permet de détecter d'emblée les mutations pathogènes connues et de filtrer en quelques secondes en moyenne 70% des polymorphismes (en vert).

De plus, et afin d'identifier les variations fréquemment rencontrées dans le laboratoire, notamment les erreurs de séquençage du PGM®, nous avons constitué une base de données implémentée au fur et à mesure avec les variants issus des différents « runs ».

Pour chaque patient, les variants non étiquetés (ROSE, ORANGE, NOIR) sont importés dans ce tableau permettant d'obtenir en un clic le nombre de fois où chaque variant a été observé dans nos séries et d'éliminer les erreurs récurrentes. Figure 72

O	P	Q	R
P4,3	freq	P5,3	freq
-311C	4	-311C	4
A1438G	13	T794C	1
A3721G	2	A1438G	13
-4599A	1	C4604A	9
C4604A	9	C5024T	1
C8734A	2	C7098T	1
A8860C	11	A8860C	11
-12418A	4	-13046C	6
-13046C	6	C14766T	10
C14766T	10	A15992T	1
	0		0

Figure 72 : tableau Excel des variant « in house » pour chaque patients les variants inconnus sot rentré dans le tableau permettant de connaître le nombre de fois où ils ont été détectés dans le laboratoire (en orange les variants redondants)

→La recherche de l'haplogroupe du patient dans Haplogrep

Au final grâce aux opérations précédentes et en moins de dix minutes, ils ne persistaient que 0 à 5 variants à investiguer par patient, parmi lesquels une majorité de variants rares (<0,2%) dans « mtDB » (ORANGE).

Dans un grand nombre de cas, les variants rares « mtDB » présents chez un patient donné semblaient tous être présents à la même fréquence, et fait intéressant, étaient généralement portés par des individus qui avaient la même origine ethnique voire le même haplogroupe. Il est donc très probable que ces derniers soient des polymorphismes rares en lien avec l'haplogroupe (polymorphismes rares dans la population générale mais fréquents dans un haplogroupe donné). Nous avons donc décidé d'inclure la notion d'haplogroupe dans la démarche d'interprétation des variants.

Le logiciel « Haplogrep », consultable en ligne permet de reconstituer les haplogroupes à partir d'une liste de variants, en se basant sur les données de la dernière version de l'arbre phylogénétique (*PhyloTree Build.16*) grâce à un algorithme dédié (KLOSS-BRANDSTATTER *et al.* 2011).

L'utilisation de ce logiciel nécessite cependant des fichiers .hsd. Ce sont des fichiers de texte tabulé, comprenant 4 colonnes :

- ✓ L'identifiant du patient
- ✓ La région séquencée (tout l'ADNmt ou régions HV seules)
- ✓ L'haplogroupe attendu (optionnel) si l'origine ethnique est connue
- ✓ La liste des variants nommés avec la position génomique suivie du nucléotide observé (ex 8860 G)

Pour chaque patient, un fichier .hsd a donc été généré automatiquement à partir de notre classeur Excel de travail, permettant une consultation rapide du site Haplogrep (figures 73) qui classe les variant en : attendus ou non dans l'haplogroupe obtenu.

Les variations non attendues étaient ensuite classées soit en « hospot » : polymorphismes fréquents dans toutes les populations, « private mutation » : variants fréquents dans d'autres haplogroupes et « global mutation » pour les variants inconnus.

g29 TABLEAU ANALYSE VIKINGOT PAK PRESENTAQA - MICOLOU Excel

Avancement de résultat La mise à jour automatique des liens a été désactivée. Afficher le contenu

	A	B	C	D	E	F	G	H	I	J
1	test1	"1-16569"		16051G 16182C 16183C 16189C 16519C 73G	152C 263G 315.1C 356.1C 8281d 8282d 8283					
2	test2	"1-16569"		16224C 16311C 16519C 73G 146C 152C 263G 309.1C 315.1C 573.1C 573.2C 573.3C						
3	test3	"1-16569"		16183C 16189C 16193.1C 16223T 16278T 16362C 73G 260A 263G 309.1C 315.1C 489C						
4	Patient	"1-16569"						00 00 00 00 00 00 00 00 00 00		

Feuille de génération du fichier .hsd

PhyloTree.org - mtDNA tree Build 16 (19 Feb 2014)

Importation du fichier .hsd du patient

https://haplogrep.uibk.ac.at

HaploGrep 2.0 - Division of Genetic Epidemiology - Medical University Innsbruck

Organiser

- Diisques durs (4)
 - Windows (C:) 339 Go libres sur 450 Go
 - HP_RECOVERY (D:) 1,35 Go libres sur 12,3 Go
 - HP_TOOLS (E:) 3,9h Go libres sur 1,99 Go
 - Toshiba 110D (F:) 450 Go libres sur 465 Go
- Peripherals utilisant des supports de stockage amovibles (1)
 - ADATA UFD (G)

Nom du fichier: biblio

HaploGrep

Open Load Testdata Save Find haplogroups Apply recommendations

ID	Range	Haplogroup	Quality (%)	Status	Polymorphisms
test3	1-16569	O2a1d	62.7	top rank	16183C 16189C 16193.1C 16223T 16278T 16362C 73G 260A 263G 309.1C 315.1C 489C 4833G 10400T 13563G 14569A 523d 524d
test2	1-16569	H2a2a	67.7	top rank	16224C 16311C 16519C 73G 146C 152C 263G 309.1C 315.1C 573.1C 573.2C 573.3C
Patient3	1-16569	U2e1	67.3	top rank	73G 152C 310C 340T 508G 750G 1438G 1811G 2706G 3291C 4789G 5225G 5428C 6045T 6932G 7028T 8860C 8860G 11467G

change Haplogroup Legend

Pos	Haplogroup	Rank	Rank in HG	Rank in Sample
1	U2e1	0.673	0.458	0.888

Liste des variants classés

Expected	Found	Remaining	Reason
152C	yes		
340T	yes	310C	local priv mutation
508G	yes	3291C	global priv mutati...
750G	yes	5225G	global priv mutati...
1438G	yes	6932G	local priv mutation
1811G	yes	8860C	global priv mutati...
2706G	yes		
4789G	yes		
5428C	yes		
6045T	yes		
7028T	yes		

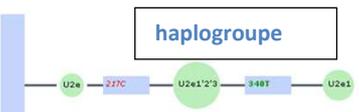


Figure 73 : Etapes de génération du haplogroupe pour chaque patient (images haplogrep)

→Investigation des variants non étiquetés dans « mitomap » et les logiciels de prédiction

Après ces étapes de tri, les variants restant non étiquetés ont été analysés via « mitomap», par analyses de:

- la fréquence dans « mitomap »
- la conservation nucléotidique
- l'effet sur un ARN ou une protéine, synonyme ou non
- l'exploration des références bibliographiques.
- La prédiction *in silico*, pour les variants protéiques non synonymes.

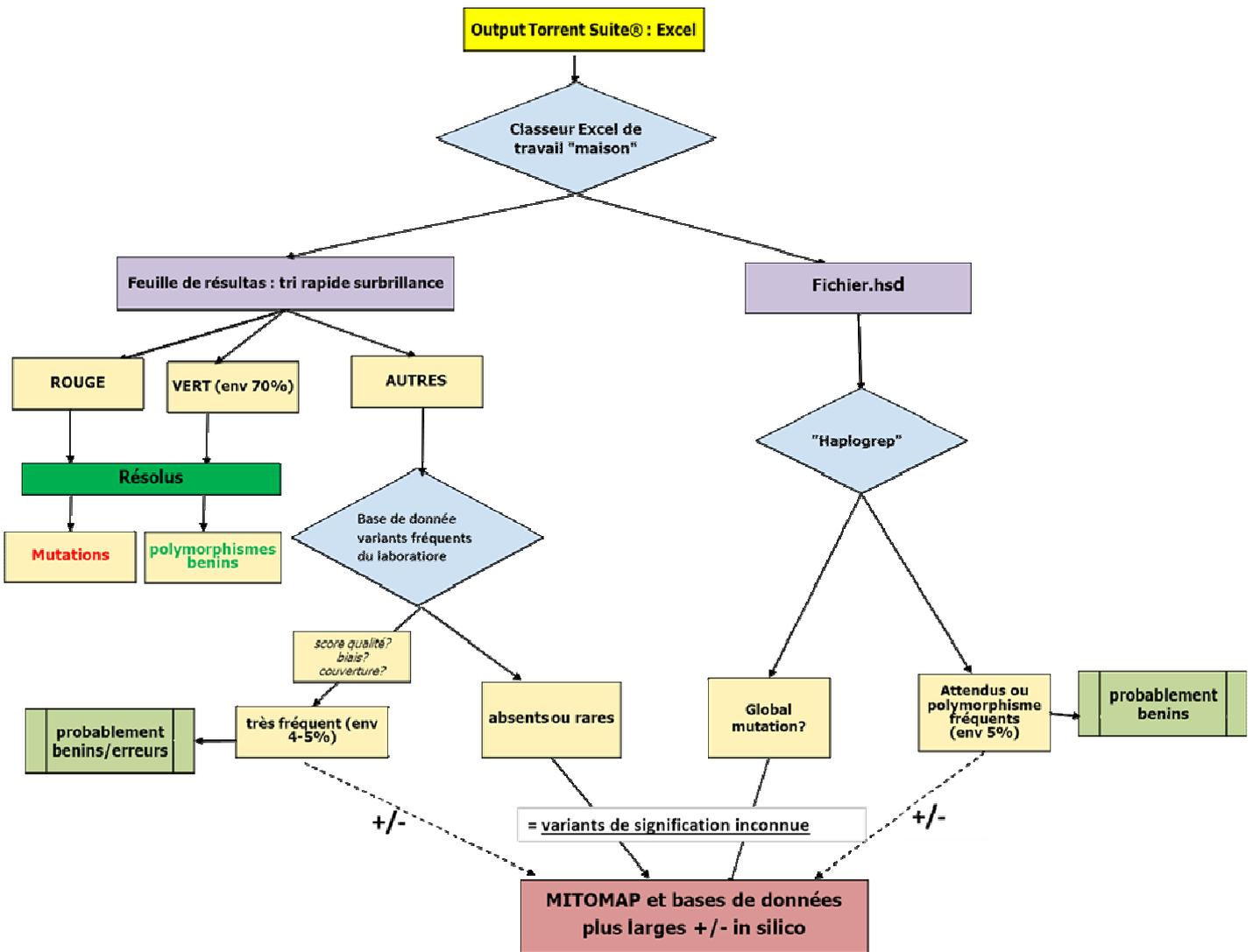


Figure 74 : algorithme de traitement des variants et par l'outil maison

2.3.4. Les mutations et nouveaux variants détectés dans notre série

Des mutations connues dont l'implication a été confirmée dans les cytopathies mitochondriales ont été retrouvées (tableau 17 ci-dessous)

Patient	mutation	taux
P13	m.3902_3908inv7	67%
p15	m.12315G>A	67%
p17	m.3291T>C	77%
p26	m.13513G>A	5,50%

Tableau 17 : Les mutations peu communes détectées

La mutation m. 3902_3908inv7 est une inversion impliquant une séquence de 7 pb flanquée par deux régions homologues. Elle a pour conséquence la substitution de trois acides aminés, respectant le cadre de lecture de la protéine *MT-ND1* : p.Asp199Gly, p.Leu200Lys et p.Ala201Val. Cette mutation a déjà été rapportée dans au moins deux cas de cytopathies mitochondriales avec déficit du complexe I (MUSUMECI *et al.* 2000) (BLAKELY *et al.* 2006). Le 1^{er} patient était un adulte de sexe masculin se plaignant d'intolérance à l'exercice et de myalgies et le 2^{ème} patient était une petite fille née avec une cardiomyopathie hypertrophique, une encéphalopathie et une acidose lactique. Le patient P13 porteur de cette mutation dans notre série se présentait avec une atteinte multisystémique associant un diabète, une surdité de perception bilatérale appareillée, une déficience intellectuelle légère et une néphropathie glomérulaire ayant évolué vers une insuffisance rénale terminale en attente de greffe.

La mutation m.12315G>A dans le gène *MT-TL2* codant pour l'ARNt (Leu) a été impliquée à l'état hétéroplasmique dans des tableaux d'ophtalmoplégie externe progressive avec myopathie et dans le syndrome de Kearns-Sayre (SHOUBRIDGE *et al.* 1997) (FU *et al.* 1996) (VALENTE *et al.* 2009). Cette mutation s'est exprimée chez le patient P15 par une ophtalmoplégie externe et une myopathie des membres en plus d'une atteinte du SNC à type d'ataxie cérébelleuse, de syndrome extra- pyramidal et d'une surdité.

La mutation m.3291T>C, dans le gène *MT-TL1* codant pour l'ARNt Leu a été largement décrite, toujours à l'état hétéroplasmique (GOTO *et al.* 1994) (GOTO 1995) . Il s'agit d'une des mutations responsables du syndrome de MELAS et plus généralement de phénotypes d'encéphalomyopathie mitochondriale. La patiente P 17, porteuse de cette mutation a été adressée dans notre centre dans le cadre d'une encéphalopathie évoluant depuis l'enfance, associée à une épilepsie partielle et une hyperlactacidémie

persistante. L'IRM objectivait une atrophie cérébrale diffuse sans lésions focales ni épisodes évoquant des pseudo-AVC.

Enfin, la mutation m.13513G>A du gène *MT-ND5* a été décrite aussi à l'état hétéroplasmique dans des phénotypes de MELAS, de Leigh et dans des phénotypes plus hétérogènes avec entre autres, une atteinte du SNC, une cardiomyopathie et une atrophie optique (SHANSKE *et al.* 2008). Le patient P26 porteur de cette mutation présentait un syndrome parkinsonien atypique fluctuant, évoluant depuis l'âge de 67 ans associé à une paralysie supra-nucléaire progressive et une atteinte neurovégétative (malaises, troubles vésico-sphinctériens).

Au total, dans notre série 4 patients sur les 22 étudiés (soit 18%) étaient porteurs de mutations de l'ADNmt confirmées pathogènes, ce qui est un bon rendement diagnostique pour les pathologies mitochondriales. Les quatre patients étaient des adultes mais deux d'entre eux avaient une présentation avec début précoce dans l'enfance (P13 et P17). En plus de ces mutations confirmées, des variants de signification inconnue ont été trouvés lors de cette étude, il s'agit principalement de variants absents ou rares dans « mtDB », synonymes ou non, portés par des gènes codant pour des ARNt ou des protéines. Ces variants sont résumés dans le tableau 18 ci-dessous, ils sont encore en cours d'investigation. En effet, l'interprétation de ce type de variants est très difficile en l'absence de tests de validation fiables *in silico* ou *in vitro*, applicables en routine diagnostique. Une démarche de recherche bibliographique large et approfondie est indispensable et le raisonnement sur la pathogénicité de ces variants inclut divers paramètres dont l'effet sur la protéine et les prédictions *in silico* lorsqu'elles sont disponibles, la conservation nucléotidique entre espèces, la ségrégation familiale, l'appartenance ou pas à l'haplogroupe du patient et le degré de fiabilité de ce dernier, le caractère hétéroplasmique ou homoplasmique du variant. Le plus souvent, les variations pathogènes sont hétéroplasmiques alors que les polymorphismes rares (ou privés) sont plutôt homoplasmiques. Néanmoins, aucun de ces arguments ne peut à lui seul trancher quant à la pathogénicité des nouveaux variants et aucune règle n'est immuable. Par exemple, des mutations pathogènes homoplasmiques existent et des sites non conservés entre espèces peuvent être le siège de mutations. Tout ceci rend compte de la complexité et de l'énormité du travail d'interprétation des données générées.

mutation	patients	taux	score qualité	gène	effet	freq mitomap	conservation	in silico
m.13922A>T	P10	3.6%	24	<i>MT-ND5</i>	non-syn: Y=>F	0 (0.00%)	40.00%	<i>Bénin</i>
m.8490T>C	P11	91%	7881	<i>MT-ATP8</i>	non-syn: M=>T	17 (0.06%)	20.00%	probablement pathogène Polyphen 0.78
m.15498G>A	P11	100%	31900	<i>MT-Cytb</i>	non-syn: G=>D	10 (0.03%)	88.89%	probablement pathogène Polyphen 0.7
m.794T>A	P19	100%	31540	<i>MT-RNR1</i>	ARNr12S?	42 (0.14%)	53.33%	
m.3079G>A	P20	10.3%	642	<i>MT-RNR2</i>	ARNr16S	0 (0.00%)	97.78%	
m.578T>C	P25	4.8%	78	<i>MT-TF</i>	tRNA Phe?	0 (0.00%)	91.11%	
m.6120A>G	P28	100%	25548	<i>MT-COI</i>	non-syn: I=>V	6 (0.02%)	93.33%	probablement pathogène Polyphen 0.9
m.9276G>A	P29	100%	9286	<i>MT-COIII</i>	non-syn: A=>T	5 (0.02%)	100.00%	<i>bénin</i>
m.13942A>G	P29	100%	27093	<i>MT-ND5</i>	non-syn: T=>A	32 (0.10%)	11.11%	<i>bénin</i>
m.9490C>T	P30	27%	3685	<i>MT-COIII</i>	non-syn: A=>V	18 (0.06%)	48.89%	<i>bénin</i>
m.7560T>C	P32	41.8%	7655	<i>MT-TD</i>	tRNA Asp ?	0 (0.00%)	95.56%	

Tableau 18 : nouveaux variants détectés dans la série

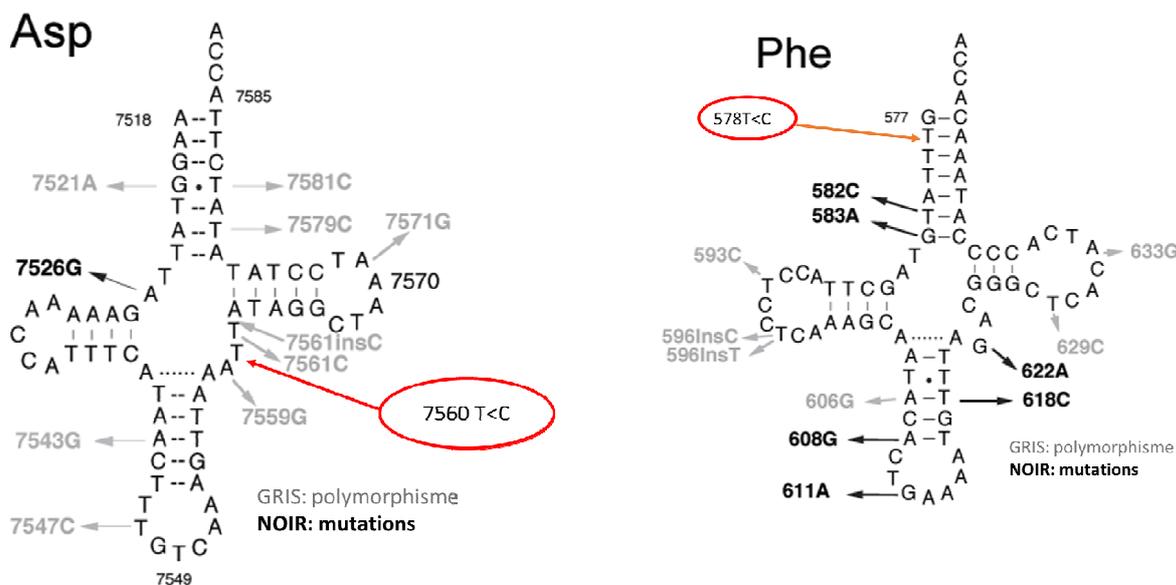


Figure 75 : localisation des deux variants sur les ARNt, source « Mamit-tRNA 8

Quatrième section : Conclusions et perspectives

La mise au point de la technique de NGS n'est pas complexe en soi quand on dispose du matériel adéquat. Cette technique est robuste et s'intègre bien dans la routine de travail du plateau technique de biologie moléculaire. Le traitement des données générées est également simple grâce au logiciel fourni Torrent Suite® qui est rapide et possède une interface facile, adaptée à l'utilisation de routine. Ce système a cependant l'inconvénient d'être fermé, supportant peu de modifications et d'ajustements (boîte noire). Il est difficile de décortiquer et de maîtriser tous ses paramètres, rendant les actions sur le logiciel très limitées et donc difficile à adapter en fonction des besoins. Ceci a notamment été bloquant dans la compréhension de la non détection du variant à 3%, ce qui souligne l'importance des collaborations biologistes-bioinformaticiens, depuis l'avènement du NGS qui dépasse les champs de la biologie médicale seule.

Depuis sa mise sur le marché, l'efficacité du séquenceur PGM® a été largement évaluée (BRAGG *et al.* 2013) (QUAIL *et al.* 2012) (LIU *et al.* 2012), avec une justesse (accuracy) atteignant les 99% pour les substitutions nucléotidiques, mais celle-ci concerne surtout les variations ayant des fréquences élevées. Le problème que pose le séquençage de l'ADNmt est la détection des variants hétéroplasmiques faibles. Plusieurs équipes ont rapporté une détection de variants qui pouvait aller jusqu'à 1% (SEO *et al.* 2015), mais celle-ci semblait aléatoire d'un « run » à l'autre dans les conditions de ce travail. Toutefois, il est indispensable d'être au moins aussi performant que la méthode de RFLP pour la détection des mutations communes recherchées habituellement. Pour cela, l'utilisation des logiciels de visualisation est intéressante car permet de vérifier directement sur les « reads » l'absence de ces dernières même à des taux faibles. Une étude plus approfondie du processus et des algorithmes bio-informatiques de tri et de détection des variants reste nécessaire, afin de bien situer les limites et éventuellement les corriger par l'utilisation de paramètres plus adaptés à l'étude de l'ADNmt

Concernant l'annotation et le tri des variants à partir des données du système Ion Torrent®, nous avons opté pour la mise au point d'un outil « maison ». En effet, la recherche de mutations pathogènes dans le génome mitochondrial est une application récente du NGS, d'où le manque d'outils dédiés. Ce dernier a été utilisé en premier lieu dans le domaine de la génétique des populations pour établir les haplogroupes. Ainsi, des logiciels publics existent dans ce cadre-là, comme « haplogrep » mais ne sont pas destinés à l'identification et l'annotation des mutations. De plus, établir les haplogroupes a des exigences de séquençage moindre que celles du diagnostic en pathologie humaine (importance moindre de la détection des hétéroplasmies faibles). De ce fait, les outils de traitement des variants rapides et faciles d'utilisation, adaptés à l'étude de l'ADNmt par NGS en routine n'existent pas encore à notre connaissance. Les outils qui émergent tel que mtTOOLBOX (CALABRESE *et al.* 2014), servant à trier directement les variants à partir des fichiers .BAM ou VCF seraient très intéressants s'ils n'étaient pas incompatibles avec l'explorateur Windows et ne nécessitaient pas de maîtriser les langages bio-informatiques (ici le langage Java).

Par ailleurs et même si le NGS a une bonne justesse, il n'est pas encore considéré comme une technique de référence, les mutations retrouvées par NGS nécessitent une confirmation par les techniques conventionnelles, classiquement par Sanger, mais ce dernier est limité dans la détection de mosaïques /hétéroplasmies à des seuils faibles (10-20%). Dans ces cas-là, la confirmation devra se faire par RFLP quand c'est possible, notamment pour les mutations communes, sinon par technique de Surveyor®. Ainsi, ces techniques plus anciennes ne serviraient plus d'outil diagnostique de première intention mais resteraient indispensables à la confirmation moléculaire des mutations.

Enfin, l'interprétation des nouveaux variants restent la part la plus longue et difficile du travail, le séquençage de l'ADNmt peut générer plus de 50 variations par patient, parmi lesquelles certaines sont rares voir absentes des bases de données, ne permettant pas de conclure quant à leur pathogénicité. Toute la difficulté réside dans la validation de ces nouvelles variations et leur tri en mutations privées versus polymorphismes rares. La mise en place de nouvelles stratégies et d'outils de validation *in vitro* ou *in silico* plus fiables et utilisables en routine devient une véritable nécessité.

L'analyse du génome mitochondrial est d'autant plus difficile que ce dernier code en majorité pour des ARN sur lesquels il est moins facile de faire des prédictions *in silico*.

Concernant les cytopathies mitochondriales, celles-ci possèdent l'avantage au même titre que toutes les maladies métaboliques d'avoir un marqueur biochimique: l'activité enzymatique de la chaîne respiratoire qui peut être étudiée *in vitro*.

Les cybrides restent un très bon modèle pour la validation des nouveaux variants de l'ADNmt, en analysant les conséquences de l'introduction de la variation à étudier sur la chaîne respiratoire (EL MEZIANE *et al.* 1998).

La PCR sur fibre unique à partir de biopsies musculaires a déjà été utilisée par des équipes pour rechercher une corrélation entre le taux de mutation et l'altération de la fibre. Cette technique peut être fastidieuse mais pourrait être réactualisée, elle trouverait son intérêt dans la validation des mutations hétéroplasmiques (SILVESTRI *et al.* 2000).

Il va donc falloir adapter ces techniques d'études fonctionnelles déjà existantes à une utilisation plus routinière via des plateformes dédiées par exemple. Ceci concerne de manière générale toutes les applications de re-séquencage par NGS en pathologie humaine, que ce soit mitochondrial ou nucléaire du fait de la masse importante de nouveaux variants générés en un temps court.

Au final, le NGS malgré ses limites reste un très bon et rapide outil de diagnostic, il permet de détecter les mutations peu communes, avec un bon rendement, 18%, dans ce travail. Ce rendement est supérieur à ce qui a été rapporté précédemment par d'autres équipes (de l'ordre de 7 %) (BANNWARTH *et al.* 2013). Ceci peut s'expliquer par la petite taille de l'échantillon et la sélection sur des critères histologiques conduisant à

l'inclusion d'une majorité d'adultes qui sont plus souvent porteurs de mutations de l'ADNmt que les populations pédiatriques.

Les mutations peu communes n'étaient pas systématiquement recherchées par technique de Sanger en routine à cause de la taille importante de l'ADNmt (16 kb). En cela, le NGS est d'une grande aide diagnostique, on peut discuter son utilisation en première ou en deuxième intention, après élimination de certaines mutations communes. En effet, si on considère par exemple le phénotype de MELAS il est dû dans 80% des cas à une seule mutation, la m.3243G<A et 90% des Leber sont expliqués par les 3 mutations les plus fréquentes. Néanmoins, le séquençage systématique chez les patients de l'ensemble de leur ADNmt pourrait modifier ces chiffres, en identifiant de nouvelles mutations.

Le NGS est aussi utile pour le diagnostic des délétions simples en précisant les points de cassures en un temps court, ce qui peut être utile en cas de demande de DPN, même si les risques de transmission sont faibles.

En conclusion, la technique de NGS est parfaitement adaptée et devient indispensable à l'étude de l'ADNmt en diagnostic de routine des cytopathies mitochondriales. Néanmoins, cela nécessite un travail d'interprétation plus long et plus fastidieux, ainsi qu'une expertise dans le domaine pour une bonne interprétation biologique et la validation des variations retrouvées.

Annexes

Annexe 1 : autres fonctions mitochondriales

L'Apoptose

L'apoptose ou mort cellulaire programmée est le processus par lequel les cellules déclenchent leur autodestruction. Elle a été mise en évidence par John Kerr, Andrew Wyllie et Alastair Currie en 1972 (KERR *et al.* 1972). Elle correspond à une mort cellulaire physiologique ou programmée indispensable au maintien de l'homéostasie tissulaire. Elle permet en effet l'élimination des cellules endommagées ou infectées et intervient dans le bon fonctionnement du système immunitaire mais aussi dans le développement embryonnaire et le renouvellement tissulaire. Contrairement à la nécrose, aucun processus inflammatoire n'est mis en jeu et l'intégrité des membranes est conservée. L'apoptose peut être induite par de nombreux stimuli (privation en facteur de croissance, exposition aux ultra-violets ou facteurs exogènes comme les anticancéreux) ou à l'inverse par l'absence de signaux de survie. Selon ces différents signaux, on pourra distinguer deux voies de signalisation : la voie des récepteurs de mort ou extrinsèque et la voie intrinsèque qui fait intervenir la mitochondrie. Il a été néanmoins démontré que les deux voies n'étaient complètement indépendantes et communiquaient entre elles via des intervenants communs. La plupart des signaux de mort convergent vers la mitochondrie qui tient le rôle d'un centre intégrateur et exécuteur pour l'apoptose (FIGURE 1) (ESTEVE *et al.* 2007)

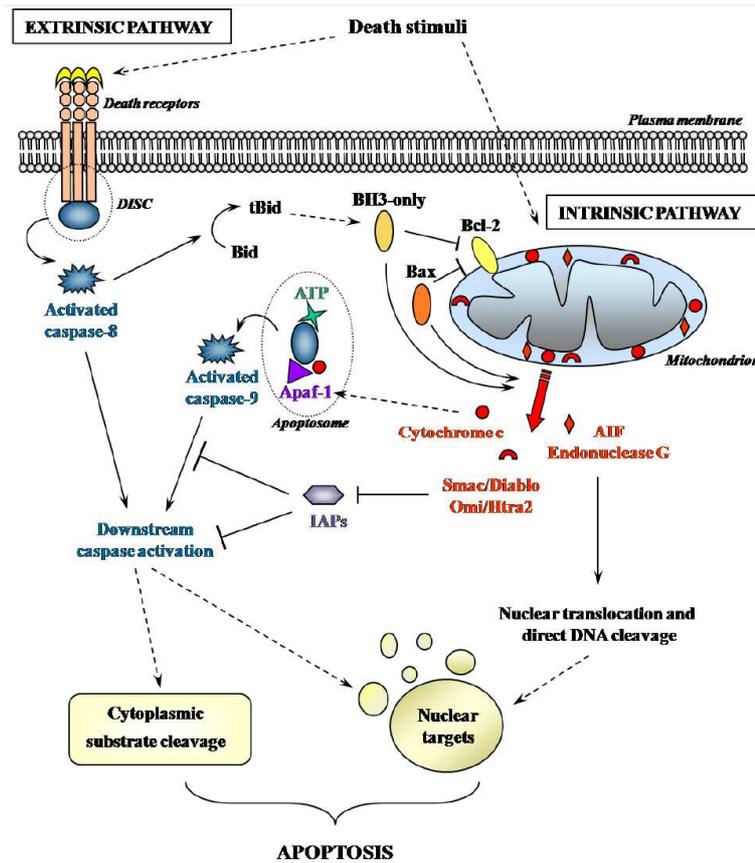


Figure 1 : les deux voies de signalisation apoptotiques et mitochondries (ESTEVE *et al.* 2007)

L'activation de la voie extrinsèque implique des protéines transmembranaires positionnées à la surface de la cellule: les récepteurs de mort cellulaire. Ils appartiennent à la superfamille du TNFR (Tumor Necrosis Factor Receptor) et sont capables d'induire la cascade apoptotique lorsqu'ils fixent les ligands appropriés menant à l'activation des caspases (THORBURN *et al.* 2004) via le recrutement cytoplasmique d'une protéine dite "adaptatrice" (FADD pour Fas-Associated Death Domain), et d'une pro-caspase initiatrice (pro-caspase 8 ou 10), pour former le DISC (Death Inducing Signaling Complex). La pro-caspase initiatrice est alors activée au sein de ce complexe par autoprotéolyse, et pourra à son tour activer les caspases effectrices 3, 6 et 7.

La voie intrinsèque mitochondriale ou voie du stress est activée en réponse à des signaux intracellulaires tels que les dommages à l'ADN, les signaux de stress (hypoxie, privation en facteurs sériques et cytokines), ou par la dérégulation de la transduction de signaux contrôlant la prolifération cellulaire (stimuli oncogéniques). Au cours de celle-ci, la plupart des signaux vont converger vers la mitochondrie. Suite à un stimulus de stress qui fait basculer la balance mort / survie vers une issue fatale, de nombreux changements interviennent sur la mitochondrie (ULIVIERI 2010). Selon l'intensité et la persistance du stimulus, ces

modifications mitochondriales peuvent conduire la cellule à un point de non-retour dans son chemin de mort, pour aboutir finalement à la libération de divers facteurs pro-apoptotiques, comme le cytochrome c, AIF (apoptosis inducing factor), l'endonucléase G et Smac/Diablo (RASOLA and BERNARDI 2007). Les mécanismes de libération de ces différents facteurs passent par une perméabilisation de la membrane mitochondriale et impliqueraient des protéines de la famille Bcl-2.

La perméabilisation transitoire mitochondriale :

Plusieurs modèles ont été proposés afin d'expliquer ce processus (RASOLA and BERNARDI 2007). Deux d'entre eux se sont distingués récemment : le modèle du PTP (permeability transition pore) et le modèle de perméabilisation directe de la membrane externe. Ces modèles ne sont pas exclusifs et il est fort probable qu'ils coexistent. Le modèle du PTP repose sur l'existence d'un complexe multiprotéique localisé aux sites de contact entre les membranes mitochondriales internes et externes (CROMPTON *et al.* 2002). Durant le fonctionnement mitochondrial normal, le PTP est fermé, ce qui permet le maintien du potentiel membranaire mitochondrial essentiel à la synthèse d'ATP, donc à la survie cellulaire. Dans des conditions apoptotiques, son ouverture prolongée, en provoquant la dépolarisation de la membrane mitochondriale interne et le gonflement de la matrice mitochondriale, conduirait au déroulement des crêtes et par la suite à la rupture de la membrane mitochondriale externe. Les mécanismes sous-jacents de ce phénomène ne sont pas encore clairement élucidés. Le deuxième modèle proposé est celui de la perméabilisation directe de la membrane mitochondriale en perdant leur intégrité structurale et fonctionnelle provoquant « une déchirure ». Cette déchirure serait due à un autre type de pore que le PTP et impliqueraient le groupe de protéines de la famille Bcl-2.

La famille Bcl-2 (B cell lymphoma 2)

Ces protéines de localisation essentiellement mitochondriale sont importantes pour la régulation des processus de mort ou de survie cellulaire. Cette superfamille est composée de près de 30 protéines, possédant de 1 à 4 domaines d'homologie BH (Bcl-2 homology), et subdivisées en trois catégories : un groupe de facteurs de survie (Bcl-2-like), et deux groupes de facteurs pro-apoptotiques (Bax-like et BH3-only). En fonction des stimuli de survie ou de mort, ces protéines vont agir de concert au niveau de la membrane mitochondriale pour inhiber ou bien autoriser l'ouverture des canaux et la libération des facteurs pro-apoptotiques (ARBEL and SHOSHAN-BARMATZ 2010).

Implication en pathologie humaine tumorale

A cause de son implication dans les voies de l'apoptose la mitochondrie possède un rôle clé dans la pathologie tumorale. Plusieurs observations ont montré la dysfonction des mitochondries dans les cellules cancéreuses, avec l'accumulation mutations de l'ADNmt, une production excessive de ROS, et une résistance aux deux voies intrinsèque et extrinsèque de l'induction de l'apoptose. La question de savoir si ces marqueurs sont à l'origine ou la conséquence de la transformation tumorale n'est cependant pas encore

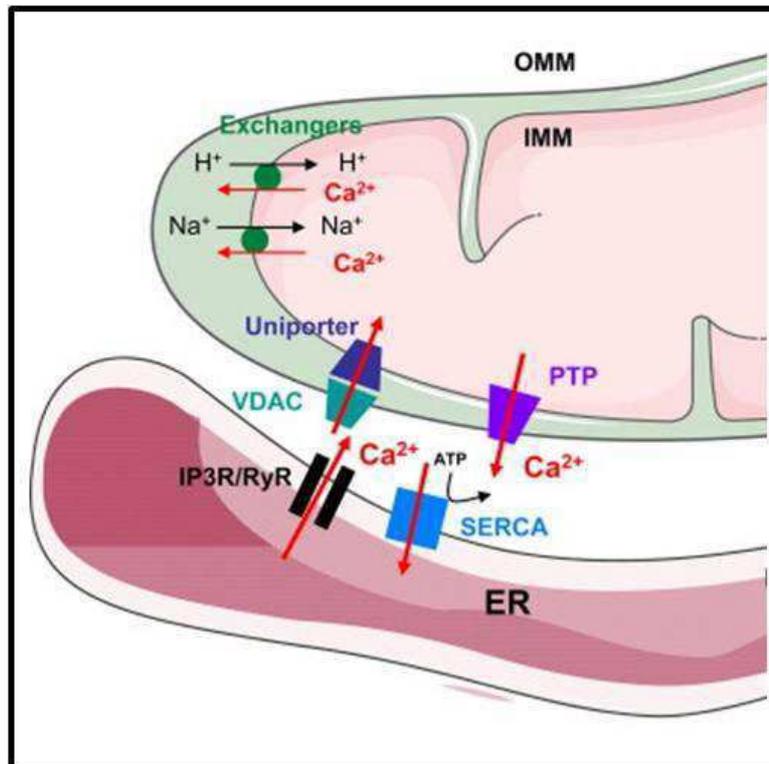
élucidée.. La mitochondrie intervient également dans les processus cancéreux via son rôle de carrefour métabolique, les cellules cancéreuses possèdent des propriétés métaboliques qui la différencient de la cellule normale. En effet, celles-ci semblent privilégier la voie de la glycolyse anaérobie au dépens de la production d'ATP par phosphorylation oxydative au niveau de la chaîne respiratoire mitochondriale, c'est l'effet Warburg (WARBURG 1956). Ceci aurait un avantage sélectif pour les cellules en s'affranchissant des conditions environnementales (concentrations en O₂) pour la production d'ATP et permettrait la mise à disposition des métabolites intermédiaires de la glycolyse pour les voies de la biosynthèse, très actives dans les tumeurs.

Tout ceci fait des mitochondries une cible thérapeutique intéressante et plusieurs études ont montré que la sensibilité aux agents anticancéreux dépendait du fonctionnement mitochondrial (MARTINEZ-OUTSCHOORN *et al.* 2011). Plusieurs stratégies sont aujourd'hui explorées pour adapter les chimiothérapies au profil métabolique des cellules tumorales et/ou cibler des acteurs mitochondriaux de l'apoptose, comme par exemple les protéines Bcl-2 ou l'inhibition de l'hexokinase, enzyme de la glycolyse et ce dans le but de neutraliser l'effet Warburg (MILANE *et al.* 2015).

Régulation de l'homéostasie du calcium :

La mitochondrie des cellules de mammifères joue un rôle important dans le stockage et l'homéostasie du calcium. En effet, la prise de calcium par des mitochondries isolées a été mise en évidence dès les années 1950 (WATANABE and WILLIAMS 1953). Cependant, l'accumulation de calcium dans les mitochondries a longtemps été considérée comme négligeable dans les conditions physiologiques par rapport à celle du réticulum endoplasmique.. Le rôle prépondérant de la mitochondrie dans la régulation des concentrations intracellulaires du calcium n'a été évoqué que plus récemment grâce à la mesure du calcium dans les mitochondries, rendue possible par la mise au point de marqueurs bioluminescents .Ces marqueurs sont dérivés de molécules chargées qui s'accumulent dans la matrice mitochondriale. Il apparaît qu'une élévation importante des ions calciques cytosoliques peut se produire dans plusieurs processus physiologiques (la sécrétion, l'apoptose, la fécondation, la contraction...) et ce par différentes voies de signalisations. L'élévation cytosolique du calcium peut aussi être causée par diverses circonstances anormales : ischémie, trauma, etc... activant diverses lipases et phosphatases pouvant provoquer la mort de la cellule. Parmi les « dissipateurs » de calcium cytosolique existe le stockage temporaire du cation dans les mitochondries, qui sera suivi d'un lent retour dans le cytoplasme (phénomène de *cycling*, (FALCKE *et al.* 1999). Il y a cependant un seuil de concentration calcique intra-mitochondrial qui ne peut être dépassé sans

mette en route de nouveaux mécanismes pathologiques, dont la production de ROS et la perméabilisation de la membrane mitochondriale (ouverture du PTP) conduisant à la mort cellulaire (KORGE *et al.* 2001).



(RIEUSSET 2011):

Figure 2 : homéostasie calcique via les mécanismes de transports mitochondriaux

La régulation du calcium intracellulaire mitochondrial passe par divers transporteurs de la mitochondrie (figure 2). Au niveau de la membrane externe le transport du calcium se fait librement grâce au VADC qui laisse passer toutes les molécules de poids moléculaire inférieur à 10 KDa. Au niveau de la membrane interne, le passage du calcium se fait par des transporteurs spécifiques qui ne sont pas encore bien caractérisés, on distingue :

Les transporteurs d'entrée mitochondriale du calcium :

L'uniport calcique mitochondrial : utilise le gradient électrochimique établi par le potentiel de membrane négatif (-180mV) généré par la force protonmotrice et conduit le calcium à entrer dans la matrice mitochondriale par le transporteur. La vitesse de transport du calcium requiert un certain seuil de concentration en calcium et en dessous de 200 à 300 nM, le calcium n'est pas transloqué dans la mitochondrie par ce mécanisme.

-Le RaM (Rapid Mode) : Il se différencie du précédent par le fait qu'il transporte le calcium environ 300 fois plus vite que le système uniport et pendant un temps très court en réponse à une accumulation de calcium (rôle de tampon).

Les transporteurs de sortie du calcium mitochondrial :

- Efflux de calcium sodium-dépendant (NCE) : ce mécanisme est prédomine dans les cellules non excitables .Cet échangeur sodium/calcium électrogénique, transporte 3 Na⁺ pour 1 Ca²⁺
- Efflux de calcium sodium-indépendant (NICE) ou système Ca²⁺/2H⁺ : ce mécanisme est dominant dans le foie, le rein, le poumon et le muscle lisse. Il transporte les ions Ca²⁺ depuis la matrice vers l'espace inter membranaire contre le gradient électrochimique calcique.
- Le système PTP (pore de perméabilité transitoire) peut aussi être considéré comme un mécanisme permettant la sortie de calcium en condition de stress oxydant et va permuter le mode d'import de calcium d'un régulateur physiologique vers un signal de mort cellulaire. De plus, une accumulation anormale d'ions calcium dans la matrice peut elle-même déclencher l'ouverture du PTP devenant agent pro-apoptotique

La régulation de l'homéostasie calcique passe également par les zones de contact de la mitochondrie avec le réticulum endoplasmique appelées MAMs pour Mitochondria Associated ER Membranes. En effet, Le réticulum endoplasmique, constitue le réservoir principal d'ions calciques avec une concentration de calcium de l'ordre de 0,5 à 1 mM au repos(Hannaert-Merah,Combettes et al. 1995; Marhl, Haberichter et al. 2000). Les MAMs permettraient entre autre de stabiliser les liaisons entre le RE et la mitochondrie et de favoriser les échanges directs de Ca²⁺ entre les deux réseaux. De plus, le regroupement spatial des mitochondries dans certains domaines de la cellule représente vraisemblablement un « tampon physiologique » qui prévient et/ou retarde la propagation de la vague calcique vers d'autres domaines et qui pourraient déclencher des réponses cellulaires non appropriées.

Il est à noter qu'en plus de son rôle dans la régulation de la concentration cytoplasmique en calcium, le transport calcique mitochondrial est aussi important pour réguler de nombreuses fonctions au sein même de la matrice mitochondriale et notamment dans le métabolisme énergétique. En effet, le calcium mitochondrial est impliqué dans l'activation de trois déshydrogénases du cycle de Krebs (pyruvate déshydrogénase, isocitrate déshydrogénase et α -cétoglutarate déshydrogénase).

La thermogénèse

Dans toutes les mitochondries la respiration cellulaire est couplée à la synthèse d'ATP. Les mitochondries du tissu adipeux brun constituent cependant un cas particulier car elles auront au sein de ce tissu une fonction thermogénique. Ce tissu est présent chez tous les Mammifères à la naissance (localisation majoritairement interscapulaire) puis devient inexistant quelques semaines plus tard, sauf chez les rongeurs et les hibernants où il persiste tout au long de la vie adulte. Dans des conditions de basse température, le système sympathique innervant le tissu brun est fortement activé. La lipolyse intracellulaire est augmentée ainsi que l'oxydation des acides gras. Paradoxalement l'énergie des oxydations n'est pas utilisée pour synthétiser de l'ATP, elle est totalement dissipée sous forme de chaleur. Il n'y a plus de couplage entre respiration et synthèse d'ATP. L'identification du mécanisme de découplage de la phosphorylation oxydative a été élucidé lorsque la protéine responsable du phénomène a été isolée : l'UCP (uncoupling protein). L'UCP est insérée dans la membrane interne de la mitochondrie et elle permet la rentrée des protons de l'espace intermembranaire vers la matrice. Ils court-circuitent ainsi l'ATPase et le flux de protons entrant dans la matrice, via l'UCP, produit de la chaleur. C'est la thermogénèse adaptative qui n'est détectable que dans les adipocytes du tissu adipeux brun chez les rongeurs exposés au froid et chez les hibernants lors du réveil. On a cependant identifié chez les tous mammifères des protéines découplantes mitochondriales UCP2 et UCP3 (la première UCP étant nommée UCP1), l'UCP2 étant exprimée dans de nombreux organes a des niveaux variables (poumon, tube digestifs...) alors que l'expression de l'UCP3 est plus restreinte (muscles squelettiques et tissu adipeux brun). Leur rôle n'a pas été clairement établi. Elles pourraient être activées lors d'un stress oxydant.

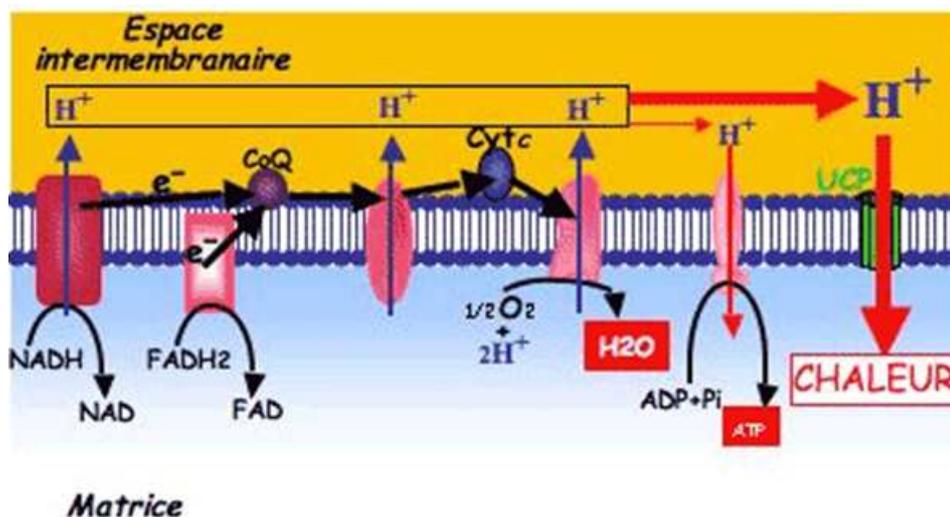


Figure 3 : découplage de la phosphorylation oxydative et thermogénèse

Annexes 2 : expression clinique des cytopathies mitochondriales

1 : tableau des principales atteintes par système dans les cytopathies mitochondriales

Atteinte neuromusculaire	<p>Hypotonie et grande acidose lactique néonatale (décès généralement avant 1 an) Encéphalomyopathie nécrosante subaiguë (débutant généralement avant 1 an, parfois plus tard) Atteinte cognitive : régression psychomotrice, retard mental, démence juvénile Encéphalopathies épisodiques (déficit sensitivomoteur avec troubles de la conscience et/ou de la vigilance) ou fixées Épilepsie partielle ou généralisée (souvent sévère et pharmacorésistante), état de mal épileptique, myoclonies Ataxie cérébelleuse Pseudo-accidents vasculaires Syndrome extrapyramidal et mouvements anormaux (dystonie) Atteinte médullaire : syndrome cordonal postérieur, syndrome pyramidal Migraine Myopathie débutant parfois en période néonatale, rhabdomyolyse, myoglobulinurie récurrente, faiblesse musculaire, myalgie, fatigabilité à l'effort, ptosis, ophthalmoparésie Neuropathie périphérique Dysautonomie</p>
Atteinte neurosensorielle	<p>Baisse d'acuité visuelle, hémianopsie latérale homonyme ou cécité Atrophie optique, rétinite pigmentaire Hypoacousie</p>
Atteinte cardiaque	<p>Cardiomyopathie hypertrophique (plus rarement dilatée) pouvant débuter dès la période néonatale, parfois associée à une non-compaction du ventricule gauche Bloc de conduction Troubles du rythme (Wolf Parkinson White)</p>
Atteinte hépato-gastrointestinale	<p>dysfonctionnement hépatocellulaire (syndrome d'Alpers) Hépatomégalie, cirrhose Insuffisance pancréatique exocrine Vomissements, diarrhée Pseudo-obstruction intestinale chronique idiopathique (adolescents > nourrissons)</p>
Atteinte rénale	<p>Tubulopathie proximale pouvant débuter dès la période néonatale Néphropathie tubulo-interstitielle (plus rare)</p>
Atteinte hématologique	<p>Anémie sidéroblastique Thrombopénie Neutropénie</p>
Atteinte endocrinologique	<p>Retard de croissance intra-utérin, retard staturopondéral Diabète insulino et non insulino-dépendant Hypoglycémie Déficit en GH, hypoparathyroïdisme, hyperaldostérionisme</p>

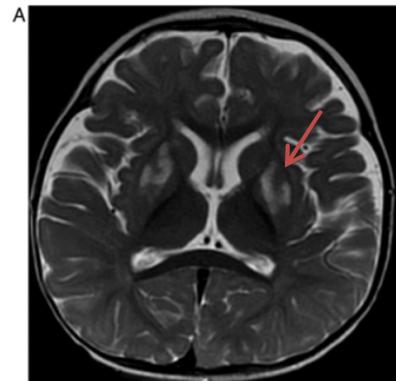
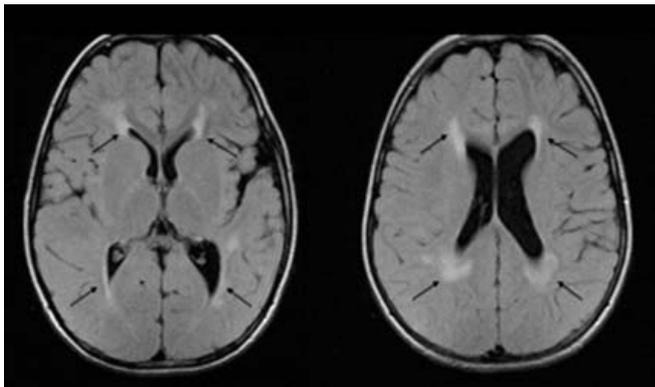
2 : les principales associations syndromiques décrites en lien avec une cytopathie mitochondriale

Disorder	Primary Features	Additional Features
<ul style="list-style-type: none"> Alpers-Huttenlocher syndrome 	<ul style="list-style-type: none"> Hypotonia Seizures Liver failure 	<ul style="list-style-type: none"> Renal tubulopathy
<ul style="list-style-type: none"> Ataxia neuropathy syndromes: Including MIRAS, SCAE, SANDO, MEMSA 	<ul style="list-style-type: none"> SANDO Other ANS: Sensory axonal neuropathy w/variable sensory & cerebellar ataxia 	<ul style="list-style-type: none"> Epilepsy Dysarthria, and/or Myopathy
<ul style="list-style-type: none"> CPEO 	<ul style="list-style-type: none"> External ophthalmoplegia Bilateral ptosis 	<ul style="list-style-type: none"> Mild proximal myopathy
<ul style="list-style-type: none"> KSS 	<ul style="list-style-type: none"> PEO onset at age <20 years Pigmentary retinopathy One of the following: CSF protein >1g/L, cerebellar ataxia, heart block 	<ul style="list-style-type: none"> Bilateral deafness Myopathy Dysphagia Diabetes mellitus Hypoparathyroidism Dementia
<ul style="list-style-type: none"> Pearson syndrome 	<ul style="list-style-type: none"> Sideroblastic anemia of childhood Pancytopenia Exocrine pancreatic failure 	<ul style="list-style-type: none"> Renal tubular defects
<ul style="list-style-type: none"> Infantile myopathy and lactic acidosis (fatal & non-fatal forms) 	<ul style="list-style-type: none"> Hypotonia in 1st year of life Feeding & respiratory difficulties 	<ul style="list-style-type: none"> Fatal form may be associated with a cardiomyopathy and/or the Toni-Fanconi-Debre syndrome
<ul style="list-style-type: none"> Leigh syndrome 	<ul style="list-style-type: none"> Subacute relapsing encephalopathy Cerebellar and brain stem signs Infantile onset 	<ul style="list-style-type: none"> Basal ganglia lucencies Maternal history of neurologic disease or Leigh syndrome
<ul style="list-style-type: none"> NARP 	<ul style="list-style-type: none"> Late-childhood or adult-onset peripheral neuropathy Ataxia Pigmentary retinopathy 	<ul style="list-style-type: none"> Basal ganglia lucencies Abnormal electroretinogram Sensorimotor neuropathy

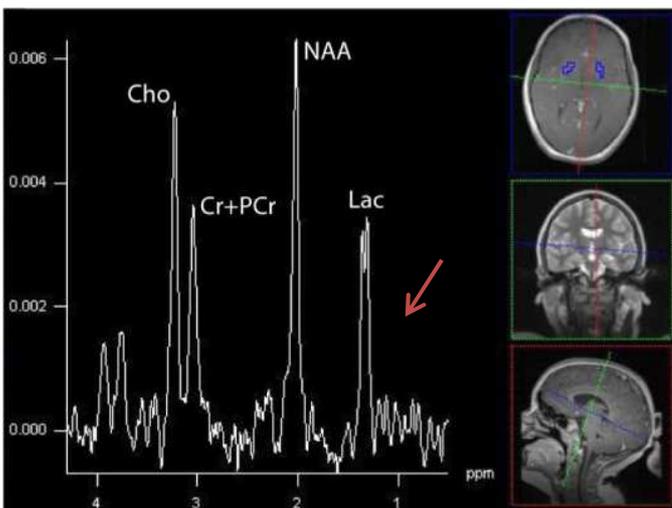
<ul style="list-style-type: none"> • MELAS 	<ul style="list-style-type: none"> • Stroke-like episodes at age <40 years • Seizures and/or dementia • Ragged-red fibers and/or lactic acidosis 	<ul style="list-style-type: none"> • Diabetes mellitus • Cardiomyopathy (initially hypertrophic; later dilated) • Bilateral deafness • Pigmentary retinopathy • Cerebellar ataxia
<ul style="list-style-type: none"> • MEMSA 	<ul style="list-style-type: none"> • Myopathy • Seizures • Cerebellar ataxia 	<ul style="list-style-type: none"> • Dementia • Peripheral neuropathy • Spasticity
<ul style="list-style-type: none"> • MERRF 	<ul style="list-style-type: none"> • Myoclonus • Seizures • Cerebellar ataxia • Myopathy 	<ul style="list-style-type: none"> • Dementia • Optic atrophy • Bilateral deafness • Peripheral neuropathy • Spasticity • Multiple lipomata
<ul style="list-style-type: none"> • LHON 	<ul style="list-style-type: none"> • Subacute painless bilateral visual failure • Males:females ~4:1 • Median age of onset 24 years 	<ul style="list-style-type: none"> • Dystonia • Cardiac pre-excitation syndromes

- CPEO = chronic progressive external ophthalmoplegia
- KSS = Kearns-Sayre syndrome
- LHON = Leber hereditary optic neuropathy
- MELAS = mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes
- MEMSA = myoclonic epilepsy myopathy sensory ataxia
- MERRF = myoclonic epilepsy with ragged-red fibers
- MIRAS = mitochondrial recessive ataxia syndrome
- NARP = neurogenic weakness with ataxia and retinitis pigmentosa
- SANDO = sensory ataxia neuropathy, dysarthria, ophthalmoplegia
- SCAE = spinocerebellar ataxia with epilepsy

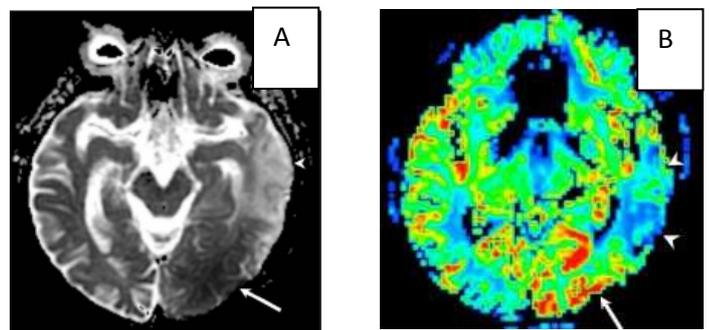
Annexe 3 : imagerie cérébrale



Annexe 3A : Hypersignaux de la substance blanche en coupe axiale FLAIR



Annexe 3E : Spectroscopie cérébrale : doublet de lactate



Annexe 3C :

A : restriction de diffusion en région postérieure dans un pseudo-AVC aigu (hypersignal antérieur correspond à un pseudo-AVC ancien)

B : ADC augmenté en postérieure (phase aigüe du pseudo-AVC)

Annexe 4 : déficits génétiques de cytopathies mitochondriales

1 : les déficits nucléaires (source GeneReviews®)

	Gènes	Mode de transmission
Gènes de structure de la CR		
	Complexe I	
Syndrome de Leigh	<i>NDUFS1, NDUFS3, NDUFS4, NDUFS7, NDUFS8, NDUFV1, NDUF2 NDUF3</i>	AR
Encéphalopathie et cardiomyopathie	<i>NDUFS2, NDUFV2</i>	AR
Pathologie néonatale létale	<i>NDUFS6</i>	AR
Encéphalomyopathie	<i>NDUFA1</i>	Liée à l’X
Cardioencéphalomyopathie, forme infantile létale	<i>NDUFA11</i>	AR
	Complexe II	
Syndrome de Leigh	<i>SDHA</i>	AR
Phéochromocytomes et paragangliomes familiaux	<i>SDHB, SDHD</i>	AD
Paragangliomes familiaux	<i>SDHC</i>	AD
	Complexe III	
Hypoglycémie, acidose lactique	<i>UQCRB</i>	AR
Retard psychomoteur sévère et signes extrapyramidaux	<i>UQCRC</i>	AR
	Complexe IV	
Encéphalomyopathie infantile	<i>COX6B1</i>	AR
	Complexe V	
Encéphalomyopathie néonatale	<i>ATP5E</i>	
Gènes impliqués dans l’assemblage de la CR		
	Complexe I	
Encéphalopathie progressive de début précoce	<i>NDUFAF2 (B17.2L)</i>	AR
Cardioencéphalomyopathie	<i>NDUFAF1 (CIA30)</i>	AR
Encéphalomyopathie	<i>C6orf66 (HRPAP20)</i>	AR
Pathologie néonatale létale	<i>C20orf7</i>	AR
	Complexe II	
Leucodystrophie mitochondriale	<i>SDHAF1</i>	AR
	Complexe III	
Encéphalopathie, tubulopathie et insuffisance rénale, Syndrome de Leigh, syndrome GRACILE	<i>BCS1L</i>	AR
	Complexe IV	
Syndrome de Leigh	<i>SURF1</i>	AR
Insuffisance hépatique et encéphalopathie néonatales	<i>SCO1</i>	AR
Cardioencéphalomyopathie néonatale	<i>SCO2</i>	AR
Tubulopathie et encéphalopathie néonatales, Syndrome de Leigh, cardiomyopathie	<i>COX10</i>	AR
Cardiomyopathie hypertrophique, Syndrome de Leigh	<i>COX15</i>	AR
Syndrome de Leigh	<i>LRPPRC</i>	AR
Encéphalomyopathie	<i>FASTKD2</i>	AR
	Complexe V	
Encéphalopathie de début précoce, dysmorphie, acidurie méthylglutaconique	<i>ATP12</i>	AR
Encéphalocardiomyopathie	<i>TMEM70</i>	AR
Encéphalomyopathie	<i>AIFM1</i>	Liée à l’X
Gènes impliqués dans la traduction mitochondriale		
	Mitoribosome	

Agénésie du corps calleux, dysmorphie et acidose lactique néonatale fatale	<i>MRPS16</i>	AR
Oedème cutané, cardiomyopathie et tubulopathie	<i>MRPS22</i>	AR
	Maturation des ARNt mitochondriaux	
Myopathie, acidose lactique et anémie sidérolastique	<i>PUS1</i>	AR
Atteinte hépatique	<i>TRMU</i>	AR
	Facteurs d'élongation	
Leucodystrophie infantile macrokystique avec micropolygyrie	<i>EFTu</i>	AR
Encéphalomyopathie, cardiomyopathie hypertrophique	<i>EFTs</i>	AR
Encéphalopathie sévère avec acidose lactique/atteinte hépatique	<i>EFG1</i>	AR
Gènes impliqués dans la maintenance de l'ADNmt		
AdPEO	<i>POLG1, POLG2, PEO1, ANTI, OPA1, RRM2B</i>	AD
ArPEO	<i>POLG1</i>	AR
MNGIE	<i>TP</i>	AR
Atteinte hépatocérébrale, syndrome d'Alpers	<i>POLG1, PEO1, DGUOK, MPV17</i>	AR
Myopathie	<i>TK2</i>	AR
Myopathie avec déplétion très sévère	<i>RRM2B</i>	AR
Encéphalopathie néonatale avec acidémie méthylmalonique	<i>SUCLG1, SUCLA2</i>	AR

2 : Les mutations de l'ADNmt confirmées (source *mitomap*)

<i>Spectre</i>	<i>gène</i>	<i>Syndrome</i>	<i>Mutations</i>
DYSTONIA/LEIGH SYNDROME	MTATP6	LS / NARP	T8993C
	MTATP6	NARP	T8993G
	MTATP6	LS	T9176G
	MTATP6	LS/FBSN	T9176C
	MTATP6	LS/Ataxia/ NARP-like disease	T9185C
	MTND3	LS	T10158C
	MTND3	LS / LS-like Disease / ESOC	T10191C
	MTND3	LS / Dystonia / Stroke	G10197A
	MTND4	LS	C11777A
	MTND5	LS	T12706C
	MTND6	LS	G14459A
	MTND6	LS/Dystonia/Ataxia	T14487C
ENCEPHALOMYOPATHY	MTND1	MELAS/LS	G3697A
	MTND5	MELAS/LS	G13513A
	MTND5	MELAS	A13514G
CA	MTTL1	MMC	A3260G

	MTTL1	MMC	C3303T
	MTTI	MICM	A4300G
DEAFNESS/ SENSORINEURAL HEARING LOSS	MTRNR1	DEAF	C1494T
	MTRNR1	DEAF	A1555G
	MTTS1 precursor	SNHL	A7445G
	MTTS1	SNHL	T7511C
DIABETES MELLITUS	MTTL1	DM/ DMDF / MIDD / SNHL / FSGS / Cardiac + multi-organ dysfunction	A3243G
ENCEPHALOMYOPATHY	MTTF	MELAS/ MM	G583A
	MTTV	AMDF	G1606A
	MTTL1	MELAS / LS	A3243G
	MTTL1	MELAS	C3256T
	MTTL1	MELAS	T3271C
	MTTL1	MELAS / Myopathy / Deafness + Cognitive Impairment	T3291C
	MTTQ	MELAS/ encephalopathy	G4332A
	MTTW	LS	A5537insT
	MTTS1	PEM / AMDF / Motor neuron disease-like	C7472insC C7471CC (='7472insC')
	MTTK	MERRF	A8344G
	MTTK	MERRF	T8356C
	MTTK	MERRF / MICM? DEAF/ Autism / LS/ Ataxia + Lipomas	G8363A
	MTTG	PEM	T10010C
	MTTH	MERRF-MELAS/ cerebral edema	G12147A
MITOCHONDRIAL MYOPATHY	MTTL1	MM/CPEO	A3243G
	MTTL1	MM	A3302G
	MTTI	CPEO/MS	G4298A
	MTTI	CPEO	G4308A
	MTTA	Myopathy	G5650A
	MTTN	CPEO/MM	G5703A
	MTTS1	MM / EXIT	G7497A
	MTTL2	CPEO / KSS	G12315A
	MTTE	Reversible COX deficiency myopathy	T14674C
	MTTE	MM+DM Encephalomyopathy	T14709C

<i>Leber</i>	MTND4	Leber	G11778A
	MTND1	Leber	G3460A
	MTND6	Leber	T14484C
	MTND1	Leber	G3635A
	MTND1	Leber	G3700A
	MTND1	Leber	G3733A
	MTND1	Leber	C4171A
	MTND4L	Leber	T10663C
	MTND6*	Leber	G14459A
	MTND6	Leber	C14482A
	MTND6	Leber	C14482G
	MTND6	Leber	C14495G
	MTND6	Leber	T14502C
	MTND6	Leber	C14568T

Abbreviations:

AMDF = Ataxia, Myopathy, and Deafness;

CPEO = Chronic Progressive Ophthalmoplegia; DEAF/SNHL = Deafness/Sensorineural Hearing Loss;

DM = Diabetes Mellitus;

D MDF = Diabetes Mellitus and Deafness;

FSGS = Focal Segmental Glomerulosclerosis;

MERRF = Myoclonic Epilepsy and Ragged Red Fiber disease;

MELAS = Mitochondrial Encephalomyopathy, Lactic acidosis and Stroke-like episodes;

MICM = Maternally Inherited Cardiomyopathy;

MIDD = Maternally Inherited Diabetes and Deafness;

LS = Leigh Syndrome;

MMC = Mitochondrial Myopathy and Cardiomyopathy;

MM = Mitochondrial Myopathy

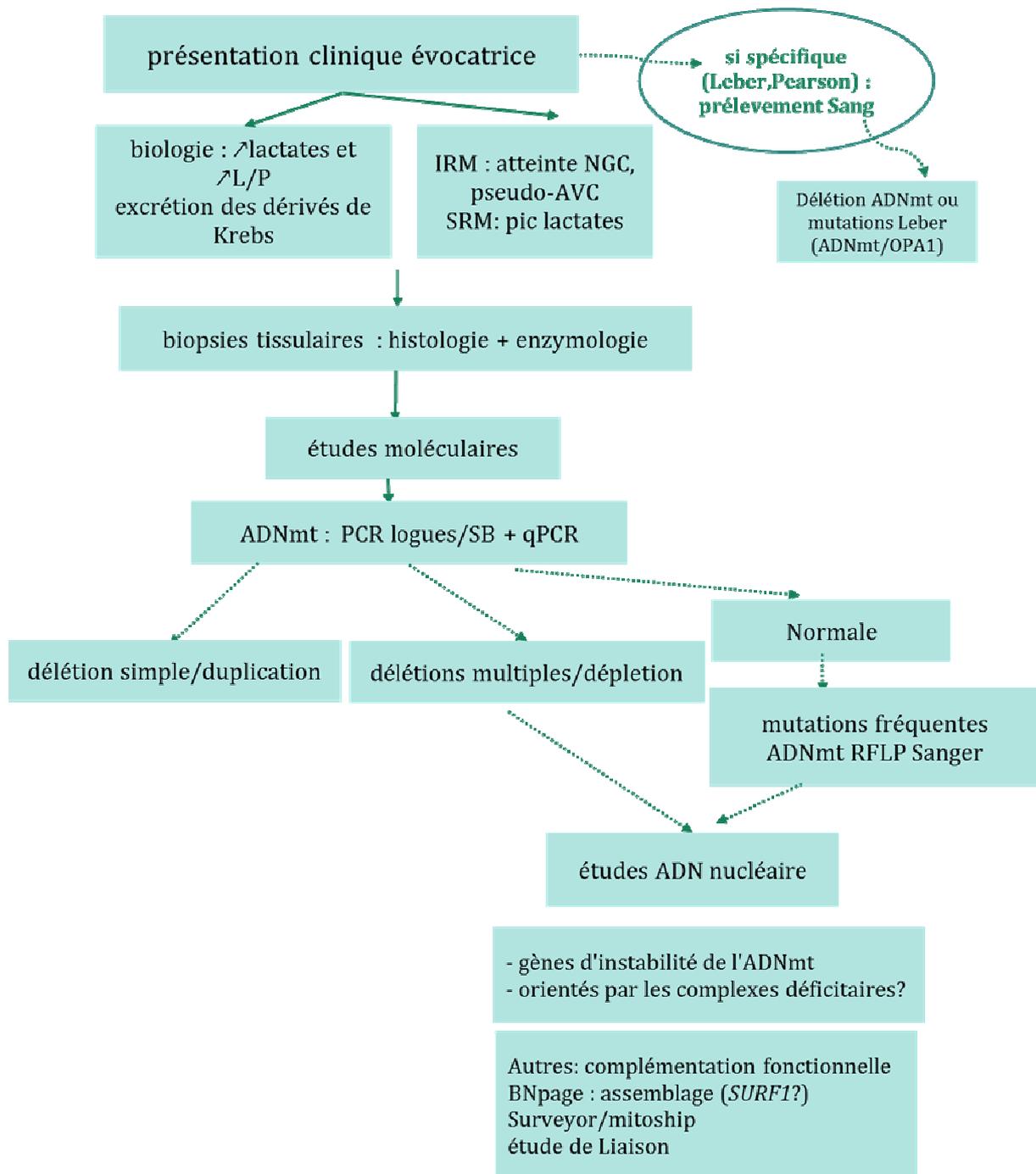
MNGIE = Mitochondrial NeuroGastroIntestinal Encephalopathy;

PEM = Progressive Encephalomyopathy;

SNHL = Sensorineural Hearing Loss.

NARP

Annexe 5 : stratégies diagnostique dans les cytopathies mitochondriales



Annexe 6 : table ASCII

0	<NUL>	32	<SPC>	64	@	96	`	128	Ä	160	†	192	ˆ	224	‡
1	<SOH>	33	!	65	A	97	a	129	Å	161	°	193	ı	225	•
2	<STX>	34	"	66	B	98	b	130	Ç	162	¢	194	¬	226	,
3	<ETX>	35	#	67	C	99	c	131	É	163	£	195	√	227	ˆ
4	<EOT>	36	\$	68	D	100	d	132	Ë	164	§	196	f	228	‰
5	<ENQ>	37	%	69	E	101	e	133	Ö	165	•	197	≈	229	À
6	<ACK>	38	&	70	F	102	f	134	Ü	166	¶	198	Δ	230	Ê
7	<BEL>	39	'	71	G	103	g	135	á	167	ß	199	«	231	Á
8	<BS>	40	(72	H	104	h	136	à	168	®	200	»	232	È
9	<TAB>	41)	73	I	105	i	137	â	169	©	201	...	233	É
10	<LF>	42	*	74	J	106	j	138	ã	170	™	202		234	Í
11	<VT>	43	+	75	K	107	k	139	ä	171	'	203	À	235	Î
12	<FF>	44	,	76	L	108	l	140	å	172	"	204	Ã	236	Ï
13	<CR>	45	-	77	M	109	m	141	ç	173	≠	205	Ö	237	Ì
14	<SO>	46	.	78	N	110	n	142	é	174	Æ	206	Œ	238	Ó
15	<SI>	47	/	79	O	111	o	143	è	175	Ø	207	œ	239	Ô
16	<DLE>	48	0	80	P	112	p	144	ê	176	∞	208	-	240	•
17	<DC1>	49	1	81	Q	113	q	145	ë	177	±	209	—	241	Ò
18	<DC2>	50	2	82	R	114	r	146	ì	178	≤	210	"	242	Ú
19	<DC3>	51	3	83	S	115	s	147	í	179	≥	211	"	243	Û
20	<DC4>	52	4	84	T	116	t	148	î	180	¥	212	'	244	Ü
21	<NAK>	53	5	85	U	117	u	149	ï	181	µ	213	'	245	ı
22	<SYN>	54	6	86	V	118	v	150	ñ	182	ð	214	÷	246	ˆ
23	<ETB>	55	7	87	W	119	w	151	ó	183	Σ	215	◊	247	ˆ
24	<CAN>	56	8	88	X	120	x	152	ò	184	Π	216	ÿ	248	—
25		57	9	89	Y	121	y	153	ô	185	π	217	ÿ	249	ˆ
26	<SUB>	58	:	90	Z	122	z	154	õ	186	ƒ	218	/	250	ˆ
27	<ESC>	59	;	91	[123	{	155	ö	187	ª	219	€	251	ˆ
28	<FS>	60	<	92	\	124		156	ú	188	º	220	<	252	ˆ
29	<GS>	61	=	93]	125	}	157	ù	189	Ω	221	>	253	ˆ
30	<RS>	62	>	94	^	126	~	158	û	190	æ	222	fi	254	ˆ
31	<US>	63	?	95	_	127		159	ü	191	ø	223	fi	255	ˆ

Annexe 7 : Matériels et méthodes

i. Amorces et conditions des PCR longues:

Premier couple 15774pb: **446F**: ACA TTA TTT TCC CCT CCC ACT CCC ATA C Tm 65.1

16220R: TTG ATT GCT GTA CTT GCT TGT AAG CAT G Tm 62.2

Deuxième couple 16204 pb: **9391F**: CAC GAG AAA GCA CAT ACC AAG Tm57.9

9027R: GCC TGC AGT AAT GTT AGC GG Tm 59.4

Mix Takara	Par échantillon (µl)
Tampon 10X	5
Mgcl2 25mM	4
Dntp 2,5 mM chacun	8
Amorce L1 F 10pmol/µl	1
Amorce L1 R 10 pmol/µl	1
Eau dist	20.5
Takara LA Taq 5U/µl	0.5
ADN	10µl (100ng)

Thermocycleur : Verity®

Programme:

94° 5min 1 cycle

94°30 sec

62° 30 sec

68° 13min

72° 10 min 1 cycle

Fin 4° infini

30 cycles

ii. Protocole de préparation de la librairie et séquençage

1. Digestion : Kit Ion shear® : digestion d'un total de 100ng d'ADN (suivre le MO du kit)

Vortexer et centrifuger les tubes de réactif : Buffer 10X et enzyme et les laisser dans de la glace (stockés normalement à 20°)

Préparer un tube ependorf par patient :

Y déposer un volume X (QSP 50ng) de chaque PCR 1 et 2

Rajouter 5µl de Buffer 10 x

Xµl eau (QSP35 avec vol ADN)

10µl d'enzyme

ADN PCR1 50ng	Vol X1(cf feuille de calcul)
ADN PCR2 50ng	Vol X2(cf feuille de calcul)
Buffer 10x	5µl
Enzyme	10µl
eau nucléase free	Vol N (35-X1-X2)µl
Total	50µl

NB : ne pas vortexer car mousse mais mélanger à la pipete

Mettre à 37° pdt 11 min PILE !! Au bain-marie

Stopper dans de la glace directement après

Rajouter 5µl de la solution STOP

Vortexer 5sec

Conserver les tubes dans la glace

2. Purification de la digestion : AMPure XP reagent®

Sortir les billes AMPure® du frigo +4° pièce séquenceur 30min avant la manip

Préparer une solution d'Ethanol à 70% extemporané (qsp 20ml)

Dans les tubes ependorfs :

- Rajouter 99 µl de billes AMPure® (1.8 x du volume à purifier) dans chaque tube : mélanger x5 Aller/Retour
- Laisser incubé 5 min
- Placer sur le rack magnétique 3min (jusqu'à ce que le milieu devienne limpide)
- Retirer le surnageant et le jeter
- Ajouter 500µl d'éthanol 70% puis l'éliminer
- Répéter une deuxième fois puis laisser sécher 5 min sur le rack
- Retirer du rack
- Eluer avec 30µl low TE (fourni avec Ion Xpress plus®) puis homogénéiser et laisser à TA 5min
- Placer sur la plaque 3 min (jusqu'à ce que le milieu devienne limpide)
- Récupérer le surnageant (25µl) dans un tube 0.2ml

3. Ligation des adaptateurs et barres codes (kit Ion Xplus®)

Se fait sur la quantité d'ADN récupéré (les 25µl) dans les tubes 0.2ml

Préparer le mix :

- 25µl d'ADN (récupéré)
- 10µl 10x ligase buffer (jaune)
- 2µl Ion P1 adapter (violet)
- 2µl dNTP mix
- 49µl H₂O
- 2µl DNA ligase (bleu)
- 8 µl Nick repair polymérase

Rajouter les 73µl du mix dans chaque tube 0.2ml

Rajouter ensuite individuellement : 2µl Ion Xpress barre code par tube (attention changer de gants en manipulant car risque de contamination dans l'identification)

Lancer la réaction au thermocycleur Verity®: programme ligation adaptateur shear : 25° 15min - 72°5min puis 4° <1h

Volume total de la réaction : 100µl

4. Puification puis amplification de la librairie digérée et identifiée : AMPure XP reagent®

Transférer les 100µl de la réaction dans un ependorf identifié

Ajouter 140 µl de solution de billes AMPure® dans chaque tube

Laisser incubé 5min TA

Mettre sur le rack magnétique 3min (jusqu'à ce que le milieu devienne limpide)

Éliminer le surnageant et rajouter 500 µl d'Eth 70% puis éliminer, répéter une 2^{ème} fois

Laisser sécher 5min puis sortir du rack

Pendant ce temps et pour chaque échantillon préparer un mix PCR *Platinum high fidelity* fourni avec le kit Ion Xplus®:

- PCR super mix high fidelity : 50 µl
- Library Amplification Primer Mix : 2µl

Homogénéiser puis éluer les billes avec les 52µl du mix

Laisser à TA 5min

Remettre sur le rack magnétique 3min (jusqu'à ce que le milieu devienne limpide)

Puis récupérer 50µl de l'éluat dans un tube 0.2ml

Lancer la réaction d'amplification fidèle : thermocycleur Verity®, programme amplification librairies : 95° 5min

5cycles de : 95°15sec+58°15sec+70°1min

4°pause

5. Purification de la librairie amplifiée pour récupérer les fragments d'intérêt (150-380pb) : AMPure XP reagent®

Ajouter 25µl de solution de billes AMPure XP reagent® dans chaque tube 0.2ml

Laisser incubé 5min à TA,

Placer 3 min sur la plaque magnétique

Récupérer 50 µl du surnageant (contient la librairie) donc attention ne pas le jeter mais le transférer dans un tube 0.2ml

Remettre 60µl de billes dans les nouveaux tubes 0.2ml pour une nouvelle purification

Éliminer le surnageant une fois la séparation faite sur la plaque magnétique

Laver avec 150µl Ethanol 70% 2 fois et laisser sécher

Éluer avec 50µl low TE et laisser incubé 3-5min à TA puis remettre sur la plaque magnétique

Récupérer 40µl de l'éluât (contient fragments d'intérêt) une fois la séparation faite dans un tube 0.2ml

iii. Paramètres du filtre des variants : indels et substitutions nucléotidiques

	SNP	indels
Minimum allele frequency min_allele_freq	0.02	0.1
Minimum quality min_variant_score	6	6
Minimum coverage min_coverage	6	15
Minimum coverage on either strand min_cov_each_strand	0	2
Maximum strand bias strand_bias	0.95	0.9
Minimum relative read quality data_quality_stringency	6.5	
Maximum common signal shift filter_unusual_predictions	0.3	

Références bibliographiques (Annexes) :

- ARBEL, N., and V. SHOSHAN-BARMATZ, 2010 Voltage-dependent anion channel 1-based peptides interact with Bcl-2 to prevent antiapoptotic activity. *J Biol Chem* **285**: 6053-6062.
- CROMPTON, M., E. BARKSBY, N. JOHNSON and M. CAPANO, 2002 Mitochondrial intermembrane junctional complexes and their involvement in cell death. *Biochimie* **84**: 143-152.
- DEVAUX-BRICOUT, M., D. GREVENT, A. S. LEBRE, M. RIO, I. DESGUERRE *et al.*, 2014 [Aspect of brain MRI in mitochondrial respiratory chain deficiency. A diagnostic algorithm of the most common mitochondrial genetic mutations]. *Rev Neurol (Paris)* **170**: 381-389.
- ESTEVE, M. A., M. CARRE and D. BRAGUER, 2007 Microtubules in apoptosis induction: are they necessary? *Curr Cancer Drug Targets* **7**: 713-729.
- FALCKE, M., J. L. HUDSON, P. CAMACHO and J. D. LECHLEITER, 1999 Impact of mitochondrial Ca²⁺ cycling on pattern formation and stability. *Biophys J* **77**: 37-44.
- KERR, J. F., A. H. WYLLIE and A. R. CURRIE, 1972 Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer* **26**: 239-257.
- KORGE, P., H. M. HONDA and J. N. WEISS, 2001 Regulation of the mitochondrial permeability transition by matrix Ca(2+) and voltage during anoxia/reoxygenation. *Am J Physiol Cell Physiol* **280**: C517-526.
- MARTINEZ-OUTSCHOORN, U. E., Z. LIN, C. TRIMMER, N. FLOMENBERG, C. WANG *et al.*, 2011 Cancer cells metabolically "fertilize" the tumor microenvironment with hydrogen peroxide, driving the Warburg effect: implications for PET imaging of human tumors. *Cell Cycle* **10**: 2504-2520.
- MILANE, L., M. TRIVEDI, A. SINGH, M. TALEKAR and M. AMIJI, 2015 Mitochondrial biology, targets, and drug delivery. *J Control Release* **207**: 40-58.
- RASOLA, A., and P. BERNARDI, 2007 The mitochondrial permeability transition pore and its involvement in cell death and in disease pathogenesis. *Apoptosis* **12**: 815-833.
- RIEUSSET, J., 2011 Mitochondria and endoplasmic reticulum: mitochondria-endoplasmic reticulum interplay in type 2 diabetes pathophysiology. *Int J Biochem Cell Biol* **43**: 1257-1262.
- THORBURN, A., J. THORBURN and A. E. FRANKEL, 2004 Induction of apoptosis by tumor cell-targeted toxins. *Apoptosis* **9**: 19-25.
- ULIVIERI, C., 2010 Cell death: insights into the ultrastructure of mitochondria. *Tissue Cell* **42**: 339-347.
- WARBURG, O., 1956 On respiratory impairment in cancer cells. *Science* **124**: 269-270.
- WATANABE, M. I., and C. M. WILLIAMS, 1953 Mitochondria in the flight muscles of insects. II. Effects of the medium on the size form, and organization of isolated sarcosomes. *J Gen Physiol* **37**: 71-90.

Références bibliographiques :

- Abramson, J., M. Svensson-Ek, B. Byrne and S. Iwata, 2001 Structure of cytochrome c oxidase: a comparison of the bacterial and mitochondrial enzymes. *Biochim Biophys Acta* 1544: 1-9.
- Alexander, C., M. Votruba, U. E. Pesch, D. L. Thiselton, S. Mayer et al., 2000 OPA1, encoding a dynamin-related GTPase, is mutated in autosomal dominant optic atrophy linked to chromosome 3q28. *Nat Genet* 26: 211-215.
- Amchenkova, A. A., L. E. Bakeeva, Y. S. Chentsov, V. P. Skulachev and D. B. Zorov, 1988 Coupling membranes as energy-transmitting cables. I. Filamentous mitochondria in fibroblasts and mitochondrial clusters in cardiomyocytes. *J Cell Biol* 107: 481-495.
- Anderson, S., A. T. Bankier, B. G. Barrell, M. H. de Bruijn, A. R. Coulson et al., 1981 Sequence and organization of the human mitochondrial genome. *Nature* 290: 457-465.
- Anderson, T. D., J. P. Ross, R. K. Roby, D. A. Lee and M. M. Holland, 1999 A validation study for the extraction and analysis of DNA from human nail material and its application to forensic casework. *J Forensic Sci* 44: 1053-1056.
- Andrews, R. M., I. Kubacka, P. F. Chinnery, R. N. Lightowlers, D. M. Turnbull et al., 1999 Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23: 147.
- Anthony, G., A. Reimann and B. Kadenbach, 1993 Tissue-specific regulation of bovine heart cytochrome-c oxidase activity by ADP via interaction with subunit VIa. *Proc Natl Acad Sci U S A* 90: 1652-1656.
- Arnold, S., and B. Kadenbach, 1997 Cell respiration is controlled by ATP, an allosteric inhibitor of cytochrome-c oxidase. *Eur J Biochem* 249: 350-354.
- Arpa, J., A. Cruz-Martinez, Y. Campos, M. Gutierrez-Molina, F. Garcia-Rio et al., 2003 Prevalence and progression of mitochondrial diseases: a study of 50 patients. *Muscle Nerve* 28: 690-695.
- Bakeeva, L. E., S. Chentsov Yu and V. P. Skulachev, 1978 Mitochondrial framework (reticulum mitochondriale) in rat diaphragm muscle. *Biochim Biophys Acta* 501: 349-369.
- Bandelt, H. J., A. Salas, R. W. Taylor and Y. G. Yao, 2009 Exaggerated status of "novel" and "pathogenic" mtDNA sequence variants due to inadequate database searches. *Hum Mutat* 30: 191-196.
- Bannwarth, S., V. Procaccio, A. S. Lebre, C. Jardel, A. Chausseot et al., 2013 Prevalence of rare mitochondrial DNA mutations in mitochondrial disorders. *J Med Genet* 50: 704-714.
- Bannwarth, S., V. Procaccio and V. Paquis-Flucklinger, 2005 Surveyor Nuclease: a new strategy for a rapid identification of heteroplasmic mitochondrial DNA mutations in patients with respiratory chain defects. *Hum Mutat* 25: 575-582.
- Barron, J. T., L. Gu and J. E. Parrillo, 1998 Malate-aspartate shuttle, cytoplasmic NADH redox potential, and energetics in vascular smooth muscle. *J Mol Cell Cardiol* 30: 1571-1579.
- Bartenhagen, C., and M. Dugas, 2015 Robust and exact structural variation detection with paired-end and soft-clipped alignments: SoftSV compared with eight algorithms. *Brief Bioinform.*
- Beau, I., M. Mehrpour and P. Codogno, 2011 Autophagosomes and human diseases. *Int J Biochem Cell Biol* 43: 460-464.
- Behin, A., C. Jardel, K. G. Claeys, J. Fagart, M. Louha et al., 2012 Adult cases of mitochondrial DNA depletion due to TK2 defect: an expanding spectrum. *Neurology* 78: 644-648.
- Bentley, D. R., S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton et al., 2008 Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456: 53-59.
- Blakely, E. L., K. J. Rennie, L. Jones, M. Elstner, Z. M. Chrzanowska-Lightowlers et al., 2006 Sporadic intragenic inversion of the mitochondrial DNA MTND1 gene causing fatal infantile lactic acidosis. *Pediatr Res* 59: 440-444.
- Boekema, E. J., and H. P. Braun, 2007 Supramolecular structure of the mitochondrial oxidative phosphorylation system. *J Biol Chem* 282: 1-4.

- Boesch, P., N. Ibrahim, A. Dietrich and R. N. Lightowers, 2010 Membrane association of mitochondrial DNA facilitates base excision repair in mammalian mitochondria. *Nucleic Acids Res* 38: 1478-1488.
- Boesch, P., F. Weber-Lotfi, N. Ibrahim, V. Tarasenko, A. Cosset et al., 2011 DNA repair in organelles: Pathways, organization, regulation, relevance in disease and aging. *Biochim Biophys Acta* 1813: 186-200.
- Bogehagen, D. F., 2012 Mitochondrial DNA nucleoid structure. *Biochim Biophys Acta* 1819: 914-920.
- Borisov, V. B., 2002 Defects in mitochondrial respiratory complexes III and IV, and human pathologies. *Mol Aspects Med* 23: 385-412.
- Boumans, H., L. A. Grivell and J. A. Berden, 1998 The respiratory chain in yeast behaves as a single functional unit. *J Biol Chem* 273: 4872-4877.
- Bragg, L. M., G. Stone, M. K. Butler, P. Hugenholtz and G. W. Tyson, 2013 Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data. *PLoS Comput Biol* 9: e1003031.
- Briere, J. J., J. Favier, V. El Ghouzzi, F. Djouadi, P. Benit et al., 2005 Succinate dehydrogenase deficiency in human. *Cell Mol Life Sci* 62: 2317-2324.
- Brun, N., Y. Robitaille, A. Grignon, B. H. Robinson, G. A. Mitchell et al., 1999 Pyruvate carboxylase deficiency: prenatal onset of ischemia-like brain lesions in two sibs with the acute neonatal form. *Am J Med Genet* 84: 94-101.
- Buermans, H. P., and J. T. den Dunnen, 2014 Next generation sequencing technology: Advances and applications. *Biochim Biophys Acta* 1842: 1932-1941.
- Calabrese, C., D. Simone, M. A. Diroma, M. Santorsola, C. Gutta et al., 2014 MToolBox: a highly automated pipeline for heteroplasmy annotation and prioritization analysis of human mitochondrial variants in high-throughput sequencing. *Bioinformatics* 30: 3115-3117.
- Capitanio, N., G. Capitanio, D. A. Demarinis, E. De Nitto, S. Massari et al., 1996 Factors affecting the H⁺/e⁻ stoichiometry in mitochondrial cytochrome c oxidase: influence of the rate of electron flow and transmembrane delta pH. *Biochemistry* 35: 10800-10806.
- Carelli, V., A. Baracca, S. Barogi, F. Pallotti, M. L. Valentino et al., 2002 Biochemical-clinical correlation in patients with different loads of the mitochondrial DNA T8993G mutation. *Arch Neurol* 59: 264-270.
- Carelli, V., F. N. Ross-Cisneros and A. A. Sadun, 2004 Mitochondrial dysfunction as a cause of optic neuropathies. *Prog Retin Eye Res* 23: 53-89.
- Carroll, J., I. M. Fearnley, J. M. Skehel, R. J. Shannon, J. Hirst et al., 2006 Bovine complex I is a complex of 45 different subunits. *J Biol Chem* 281: 32724-32727.
- Casari, G., M. De Fusco, S. Ciarmatori, M. Zeviani, M. Mora et al., 1998 Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. *Cell* 93: 973-983.
- Castellana, S., J. Ronai and T. Mazza, 2015 MitImpact: an exhaustive collection of pre-computed pathogenicity predictions of human mitochondrial non-synonymous variants. *Hum Mutat* 36: E2413-2422.
- Castro-Gago, M., V. Gonzalez-Conde, M. J. Fernandez-Seara, E. Rodrigo-Saez, S. Fernandez-Cebrian et al., 1999 [Early mitochondrial encephalomyopathy due to complex IV deficiency consistent with Alpers-Huttenlocher syndrome: report of two cases]. *Rev Neurol* 29: 912-917.
- Cavanagh, J. B., and B. N. Harding, 1994 Pathogenic factors underlying the lesions in Leigh's disease. Tissue responses to cellular energy deprivation and their clinico-pathological consequences. *Brain* 117 (Pt 6): 1357-1376.
- Chakrapani, A., L. Heptinstall and J. Walter, 1998 A family with Leigh syndrome caused by the rarer T8993C mutation. *J Inher Metab Dis* 21: 685-686.
- Chen, H., and D. C. Chan, 2006 Critical dependence of neurons on mitochondrial dynamics. *Curr Opin Cell Biol* 18: 453-459.
- Chen, X., R. Prosser, S. Simonetti, J. Sadlock, G. Jagiello et al., 1995 Rearranged mitochondrial genomes are present in human oocytes. *Am J Hum Genet* 57: 239-247.

- Chinnery, P. F., R. M. Andrews, D. M. Turnbull and N. N. Howell, 2001 Leber hereditary optic neuropathy: Does heteroplasmy influence the inheritance and expression of the G11778A mitochondrial DNA mutation? *Am J Med Genet* 98: 235-243.
- Chinnery, P. F., and G. Hudson, 2013 Mitochondrial genetics. *Br Med Bull* 106: 135-159.
- Chinnery, P. F., M. A. Johnson, T. M. Wardell, R. Singh-Kler, C. Hayes et al., 2000 The epidemiology of pathogenic mitochondrial DNA mutations. *Ann Neurol* 48: 188-193.
- Chomova, M., and P. Racay, 2010 Mitochondrial complex I in the network of known and unknown facts. *Gen Physiol Biophys* 29: 3-11.
- Chow, S. L., Z. J. Rooney, M. A. Cleary, P. T. Clayton and J. V. Leonard, 2005 The significance of elevated CSF lactate. *Arch Dis Child* 90: 1188-1189.
- Cipolat, S., O. Martins de Brito, B. Dal Zilio and L. Scorrano, 2004 OPA1 requires mitofusin 1 to promote mitochondrial fusion. *Proc Natl Acad Sci U S A* 101: 15927-15932.
- Clayton, D. A., 1982 Replication of animal mitochondrial DNA. *Cell* 28: 693-705.
- Clayton, D. A., 2003 Mitochondrial DNA replication: what we know. *IUBMB Life* 55: 213-217.
- Collu-Marchese, M., M. Shuen, M. Pauly, A. Saleem and D. A. Hood, 2015 The regulation of mitochondrial transcription factor A (Tfam) expression during skeletal muscle cell differentiation. *Biosci Rep* 35.
- Copeland, W. C., 2012 Defects in mitochondrial DNA replication and human disease. *Crit Rev Biochem Mol Biol* 47: 64-74.
- Damian, M. S., P. Seibel, W. Schachenmayr, H. Reichmann and W. Dorndorf, 1996 VACTERL with the mitochondrial np 3243 point mutation. *Am J Med Genet* 62: 398-403.
- Darin, N., A. Oldfors, A. R. Moslemi, E. Holme and M. Tulinius, 2001 The incidence of mitochondrial encephalomyopathies in childhood: clinical features and morphological, biochemical, and DNA abnormalities. *Ann Neurol* 49: 377-383.
- Das, S., N. Hajnoczky, A. N. Antony, G. Csordas, L. D. Gaspers et al., 2012 Mitochondrial morphology and dynamics in hepatocytes from normal and ethanol-fed rats. *Pflugers Arch* 464: 101-109.
- De Lonlay, P., C. Mugnier, D. Sanlaville, K. Chantrel-Groussard, P. Benit et al., 2002 Cell complementation using Genebridge 4 human:rodent hybrids for physical mapping of novel mitochondrial respiratory chain deficiency genes. *Hum Mol Genet* 11: 3273-3281.
- Debray, F. G., G. A. Mitchell, P. Allard, B. H. Robinson, J. A. Hanley et al., 2007 Diagnostic accuracy of blood lactate-to-pyruvate molar ratio in the differential diagnosis of congenital lactic acidosis. *Clin Chem* 53: 916-921.
- Delatycki, M. B., R. Williamson and S. M. Forrest, 2000 Friedreich ataxia: an overview. *J Med Genet* 37: 1-8.
- Delettre, C., J. M. Griffoin, J. Kaplan, H. Dollfus, B. Lorenz et al., 2001 Mutation spectrum and splicing variants in the OPA1 gene. *Hum Genet* 109: 584-591.
- Delettre, C., G. Lenaers, J. M. Griffoin, N. Gigarel, C. Lorenzo et al., 2000 Nuclear gene OPA1, encoding a mitochondrial dynamin-related protein, is mutated in dominant optic atrophy. *Nat Genet* 26: 207-210.
- Devaux-Bricout, M., D. Grevent, A. S. Lebre, M. Rio, I. Desguerre et al., 2014 [Aspect of brain MRI in mitochondrial respiratory chain deficiency. A diagnostic algorithm of the most common mitochondrial genetic mutations]. *Rev Neurol (Paris)* 170: 381-389.
- DiMauro, S., 2013 Mitochondrial encephalomyopathies--fifty years on: the Robert Wartenberg Lecture. *Neurology* 81: 281-291.
- DiMauro, S., and M. Hirano, 1993 Merrf.
- DiMauro, S., and E. A. Schon, 2001 Mitochondrial DNA mutations in human disease. *Am J Med Genet* 106: 18-26.
- Distelmaier, F., W. J. Koopman, L. P. van den Heuvel, R. J. Rodenburg, E. Mayatepek et al., 2009 Mitochondrial complex I deficiency: from organelle dysfunction to clinical disease. *Brain* 132: 833-842.
- Dowling, D. K., 2014 Evolutionary perspectives on the links between mitochondrial genotype and disease phenotype. *Biochim Biophys Acta* 1840: 1393-1403.

- Dudkina, N. V., H. Eubel, W. Keegstra, E. J. Boekema and H. P. Braun, 2005 Structure of a mitochondrial supercomplex formed by respiratory-chain complexes I and III. *Proc Natl Acad Sci U S A* 102: 3225-3229.
- Durcan, T. M., and E. A. Fon, 2015 The three 'P's of mitophagy: PARKIN, PINK1, and post-translational modifications. *Genes Dev* 29: 989-999.
- El-Hattab, A. W., A. M. Adesina, J. Jones and F. Scaglia, 2015 MELAS syndrome: Clinical manifestations, pathogenesis, and treatment options. *Mol Genet Metab* 116: 4-12.
- El Meziane, A., S. K. Lehtinen, I. J. Holt and H. T. Jacobs, 1998 Mitochondrial tRNA^{Leu} isoforms in lung carcinoma cybrid cells containing the np 3243 mtDNA mutation. *Hum Mol Genet* 7: 2141-2147.
- Filippi, M., A. Charil, M. Rovaris, M. Absinta and M. A. Rocca, 2014 Insights from magnetic resonance imaging. *Handb Clin Neurol* 122: 115-149.
- Fragouli, E., K. Spath, S. Alfarawati, F. Kaper, A. Craig et al., 2015 Altered levels of mitochondrial DNA are associated with female age, aneuploidy, and provide an independent measure of embryonic implantation potential. *PLoS Genet* 11: e1005241.
- Fu, K., R. Hartlen, T. Johns, A. Genge, G. Karpati et al., 1996 A novel heteroplasmic tRNA^{Leu}(CUN) mtDNA point mutation in a sporadic patient with mitochondrial encephalomyopathy segregates rapidly in skeletal muscle and suggests an approach to therapy. *Hum Mol Genet* 5: 1835-1840.
- Fuste, J. M., S. Wanrooij, E. Jemt, C. E. Granycome, T. J. Cluett et al., 2010 Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. *Mol Cell* 37: 67-78.
- Gegg, M. E., J. M. Cooper, K. Y. Chau, M. Rojo, A. H. Schapira et al., 2010 Mitofusin 1 and mitofusin 2 are ubiquitinated in a PINK1/parkin-dependent manner upon induction of mitophagy. *Hum Mol Genet* 19: 4861-4870.
- Genova, M. L., and G. Lenaz, 2014 Functional role of mitochondrial respiratory supercomplexes. *Biochim Biophys Acta* 1837: 427-443.
- Ghezzi, D., P. Goffrini, G. Uziel, R. Horvath, T. Klopstock et al., 2009 SDHAF1, encoding a LYR complex-II specific assembly factor, is mutated in SDH-defective infantile leukoencephalopathy. *Nat Genet* 41: 654-656.
- Glater, E. E., L. J. Megeath, R. S. Stowers and T. L. Schwarz, 2006 Axonal transport of mitochondria requires Milton to recruit kinesin heavy chain and is light chain independent. *J Cell Biol* 173: 545-557.
- Golan, D., and P. Medvedev, 2013 Using state machines to model the Ion Torrent sequencing process and to improve read error rates. *Bioinformatics* 29: i344-351.
- Goto, Y., 1995 Clinical features of MELAS and mitochondrial DNA mutations. *Muscle Nerve Suppl* 3: S107-112.
- Goto, Y., K. Tsugane, Y. Tanabe, I. Nonaka and S. Horai, 1994 A new point mutation at nucleotide pair 3291 of the mitochondrial tRNA(Leu(UUR)) gene in a patient with mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes (MELAS). *Biochem Biophys Res Commun* 202: 1624-1630.
- Griparic, L., N. N. van der Wel, I. J. Orozco, P. J. Peters and A. M. van der Bliek, 2004 Loss of the intermembrane space protein Mgm1/OPA1 induces swelling and localized constrictions along the lengths of mitochondria. *J Biol Chem* 279: 18792-18798.
- Guan, M. X., 2011 Mitochondrial 12S rRNA mutations associated with aminoglycoside ototoxicity. *Mitochondrion* 11: 237-245.
- Haas, R. H., 1988 Thiamin and the brain. *Annu Rev Nutr* 8: 483-515.
- Haas, R. H., S. Parikh, M. J. Falk, R. P. Saneto, N. I. Wolf et al., 2008 The in-depth evaluation of suspected mitochondrial disease. *Mol Genet Metab* 94: 16-37.
- Habersetzer, J., W. Ziani, I. Larrieu, C. Stines-Chaumeil, M. F. Giraud et al., 2013 ATP synthase oligomerization: from the enzyme models to the mitochondrial morphology. *Int J Biochem Cell Biol* 45: 99-105.
- Henley, C. M., 3rd, and J. Schacht, 1988 Pharmacokinetics of aminoglycoside antibiotics in blood, inner-ear fluids and tissues and their relationship to ototoxicity. *Audiology* 27: 137-146.

- Holt, I. J., J. M. Cooper, J. A. Morgan-Hughes and A. E. Harding, 1988 Deletions of muscle mitochondrial DNA. *Lancet* 1: 1462.
- Hood, D. A., I. Irrcher, V. Ljubicic and A. M. Joseph, 2006 Coordination of metabolic plasticity in skeletal muscle. *J Exp Biol* 209: 2265-2275.
- Horsefield, R., V. Yankovskaya, G. Sexton, W. Whittingham, K. Shiomi et al., 2006 Structural and computational analysis of the quinone-binding site of complex II (succinate-ubiquinone oxidoreductase): a mechanism of electron transfer and proton conduction during ubiquinone reduction. *J Biol Chem* 281: 7309-7316.
- Huttemann, M., P. Pecina, M. Rainbolt, T. H. Sanderson, V. E. Kagan et al., 2011 The multiple functions of cytochrome c and their regulation in life and death decisions of the mammalian cell: From respiration to apoptosis. *Mitochondrion* 11: 369-381.
- Iizuka, T., and F. Sakai, 2005 Pathogenesis of stroke-like episodes in MELAS: analysis of neurovascular cellular mechanisms. *Curr Neurovasc Res* 2: 29-45.
- Ingman, M., and U. Gyllensten, 2006 mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences. *Nucleic Acids Res* 34: D749-751.
- Janssen, R. J., L. G. Nijtmans, L. P. van den Heuvel and J. A. Smeitink, 2006 Mitochondrial complex I: structure, function and pathology. *J Inherit Metab Dis* 29: 499-515.
- Kadenbach, B., R. Ramzan, L. Wen and S. Vogt, 2010 New extension of the Mitchell Theory for oxidative phosphorylation in mitochondria of living organisms. *Biochim Biophys Acta* 1800: 205-212.
- Kajander, O. A., T. A. Kunnas, M. Perola, S. K. Lehtinen, P. J. Karhunen et al., 1999 Long-extension PCR to detect deleted mitochondrial DNA molecules is compromised by technical artefacts. *Biochem Biophys Res Commun* 254: 507-514.
- Kearns, T. P., 1965 External Ophthalmoplegia, Pigmentary Degeneration of the Retina, and Cardiomyopathy: A Newly Recognized Syndrome. *Trans Am Ophthalmol Soc* 63: 559-625.
- Kerr, D. S., 2013 Review of clinical trials for mitochondrial disorders: 1997-2012. *Neurotherapeutics* 10: 307-319.
- Kishi, M., Y. Yamamura, T. Kurihara, N. Fukuhara, K. Tsuruta et al., 1988 An autopsy case of mitochondrial encephalomyopathy: biochemical and electron microscopic studies of the brain. *J Neurol Sci* 86: 31-40.
- Kloss-Brandstatter, A., D. Pacher, S. Schonherr, H. Weissensteiner, R. Binna et al., 2011 HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Hum Mutat* 32: 25-32.
- Koga, Y., N. Povalko, J. Nishioka, K. Katayama, N. Kakimoto et al., 2010 MELAS and L-arginine therapy: pathophysiology of stroke-like episodes. *Ann N Y Acad Sci* 1201: 104-110.
- Kogelnik, A. M., M. T. Lott, M. D. Brown, S. B. Navathe and D. C. Wallace, 1998 MITOMAP: a human mitochondrial genome database--1998 update. *Nucleic Acids Res* 26: 112-115.
- Kussmaul, L., and J. Hirst, 2006 The mechanism of superoxide production by NADH:ubiquinone oxidoreductase (complex I) from bovine heart mitochondria. *Proc Natl Acad Sci U S A* 103: 7607-7612.
- Lagouge, M., C. Argmann, Z. Gerhart-Hines, H. Meziane, C. Lerin et al., 2006 Resveratrol improves mitochondrial function and protects against metabolic disease by activating SIRT1 and PGC-1alpha. *Cell* 127: 1109-1122.
- Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: R25.
- Lebovitz, R. M., 1996 Ownership of human genes. *Nature* 382: 17.
- Lee, Y. J., S. Y. Jeong, M. Karbowski, C. L. Smith and R. J. Youle, 2004 Roles of the mammalian mitochondrial fission and fusion mediators Fis1, Drp1, and Opa1 in apoptosis. *Mol Biol Cell* 15: 5001-5011.
- Li, H. Z., R. Y. Li and M. Li, 2014 A review of maternally inherited diabetes and deafness. *Front Biosci (Landmark Ed)* 19: 777-782.
- Lincke, C. R., C. van den Bogert, L. G. Nijtmans, R. J. Wanders, P. Tamminga et al., 1996 Cerebellar hypoplasia in respiratory chain dysfunction. *Neuropediatrics* 27: 216-218.

- Liu, L., Y. Li, S. Li, N. Hu, Y. He et al., 2012 Comparison of next-generation sequencing systems. *J Biomed Biotechnol* 2012: 251364.
- Lopez, M. F., B. S. Kristal, E. Chernokalskaya, A. Lazarev, A. I. Shestopalov et al., 2000 High-throughput profiling of the mitochondrial proteome using affinity fractionation and automation. *Electrophoresis* 21: 3427-3440.
- Maitra, A., Y. Cohen, S. E. Gillespie, E. Mambo, N. Fukushima et al., 2004 The Human MitoChip: a high-throughput sequencing microarray for mitochondrial mutation detection. *Genome Res* 14: 812-819.
- Malfatti, E., M. Bugiani, F. Invernizzi, C. F. de Souza, L. Farina et al., 2007 Novel mutations of ND genes in complex I deficiency associated with mitochondrial encephalopathy. *Brain* 130: 1894-1904.
- Man, P. Y., P. G. Griffiths, D. T. Brown, N. Howell, D. M. Turnbull et al., 2003 The epidemiology of Leber hereditary optic neuropathy in the North East of England. *Am J Hum Genet* 72: 333-339.
- Mansfield, P., and A. A. Maudsley, 1977 Medical imaging by NMR. *Br J Radiol* 50: 188-194.
- Martin Negrier, M. L., M. Coquet, B. T. Moretto, J. Y. Lacut, M. Dupon et al., 1998 Partial triplication of mtDNA in maternally transmitted diabetes mellitus and deafness. *Am J Hum Genet* 63: 1227-1232.
- McElhoe, J. A., M. M. Holland, K. D. Makova, M. S. Su, I. M. Paul et al., 2014 Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq. *Forensic Sci Int Genet* 13: 20-29.
- Milane, L., M. Trivedi, A. Singh, M. Talekar and M. Amiji, 2015 Mitochondrial biology, targets, and drug delivery. *J Control Release* 207: 40-58.
- Morato, L., E. Bertini, D. Verrigni, A. Ardisson, M. Ruiz et al., 2014 Mitochondrial dysfunction in central nervous system white matter disorders. *Glia* 62: 1878-1894.
- Murphy, A. N., 2009 In a flurry of PINK, mitochondrial bioenergetics takes a leading role in Parkinson's disease. *EMBO Mol Med* 1: 81-84.
- Musumeci, O., A. L. Andreu, S. Shanske, N. Bresolin, G. P. Comi et al., 2000 Intragenic inversion of mtDNA: a new type of pathogenic mutation in a patient with mitochondrial myopathy. *Am J Hum Genet* 66: 1900-1904.
- Ohama, E., S. Ohara, F. Ikuta, K. Tanaka, M. Nishizawa et al., 1987 Mitochondrial angiopathy in cerebral blood vessels of mitochondrial encephalomyopathy. *Acta Neuropathol* 74: 226-233.
- Olichon, A., L. J. Emorine, E. Descoins, L. Pelloquin, L. Bricchese et al., 2002 The human dynamin-related protein OPA1 is anchored to the mitochondrial inner membrane facing the inter-membrane space. *FEBS Lett* 523: 171-176.
- Paradies, G., G. Petrosillo, M. Pistolese and F. M. Ruggiero, 2001 Reactive oxygen species generated by the mitochondrial respiratory chain affect the complex III activity via cardiolipin peroxidation in beef-heart submitochondrial particles. *Mitochondrion* 1: 151-159.
- Park, M. K., M. C. Ashby, G. Erdemli, O. H. Petersen and A. V. Tepikin, 2001 Perinuclear, perigranular and sub-plasmalemmal mitochondria have distinct functions in the regulation of cellular calcium transport. *EMBO J* 20: 1863-1874.
- Park, Y. Y., S. Lee, M. Karbowski, A. Neutzner, R. J. Youle et al., 2010 Loss of MARCH5 mitochondrial E3 ubiquitin ligase induces cellular senescence through dynamin-related protein 1 and mitofusin 1. *J Cell Sci* 123: 619-626.
- Parson, W., C. Strobl, G. Huber, B. Zimmermann, S. M. Gomes et al., 2013 Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). *Forensic Sci Int Genet* 7: 543-549.
- Pavlakis, S. G., P. C. Phillips, S. DiMauro, D. C. De Vivo and L. P. Rowland, 1984 Mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes: a distinctive clinical syndrome. *Ann Neurol* 16: 481-488.
- Pearson, H. A., J. S. Lobel, S. A. Kocoshis, J. L. Naiman, J. Windmiller et al., 1979 A new syndrome of refractory sideroblastic anemia with vacuolization of marrow precursors and exocrine pancreatic dysfunction. *J Pediatr* 95: 976-984.

- Petruzzella, V., C. T. Moraes, M. C. Sano, E. Bonilla, S. DiMauro et al., 1994 Extremely high levels of mutant mtDNAs co-localize with cytochrome c oxidase-negative ragged-red fibers in patients harboring a point mutation at nt 3243. *Hum Mol Genet* 3: 449-454.
- Phadke, M., M. R. Lokeshwar, S. Bhutada, C. Tampi, R. Saxena et al., 2012 Kearns Sayre Syndrome--case report with review of literature. *Indian J Pediatr* 79: 650-654.
- Poulton, J., M. E. Deadman and R. M. Gardiner, 1989 Duplications of mitochondrial DNA in mitochondrial myopathy. *Lancet* 1: 236-240.
- Prezant, T. R., J. V. Agopian, M. C. Bohlman, X. Bu, S. Oztas et al., 1993 Mitochondrial ribosomal RNA mutation associated with both antibiotic-induced and non-syndromic deafness. *Nat Genet* 4: 289-294.
- Priault, M., B. Salin, J. Schaeffer, F. M. Vallette, J. P. di Rago et al., 2005 Impairing the bioenergetic status and the biogenesis of mitochondria triggers mitophagy in yeast. *Cell Death Differ* 12: 1613-1621.
- Quail, M. A., M. Smith, P. Coupland, T. D. Otto, S. R. Harris et al., 2012 A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13: 341.
- Quinzii, C. M., and M. Hirano, 2011 Primary and secondary CoQ(10) deficiencies in humans. *Biofactors* 37: 361-365.
- Rodriguez-Hernandez, A., M. D. Cordero, L. Salviati, R. Artuch, M. Pineda et al., 2009 Coenzyme Q deficiency triggers mitochondria degradation by mitophagy. *Autophagy* 5: 19-32.
- Rolfe, D. F., and G. C. Brown, 1997 Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiol Rev* 77: 731-758.
- Ronaghi, M., S. Karamohamed, B. Pettersson, M. Uhlen and P. Nyren, 1996 Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem* 242: 84-89.
- Rusk, N., 2011 Torrents of sequence. *Nat Meth* 8: 44-44.
- Russell, O., and D. Turnbull, 2014 Mitochondrial DNA disease-molecular insights and potential routes to a cure. *Exp Cell Res* 325: 38-43.
- Rustin, P., D. Chretien, T. Bourgeron, B. Gerard, A. Rotig et al., 1994 Biochemical and molecular investigations in respiratory chain deficiencies. *Clin Chim Acta* 228: 35-51.
- Samson, J. F., P. G. Barth, J. I. de Vries, F. H. Menko, W. Ruitenbeek et al., 1994 Familial mitochondrial encephalopathy with fetal ultrasonographic ventriculomegaly and intracerebral calcifications. *Eur J Pediatr* 153: 510-516.
- Sanger, F., S. Nicklen and A. R. Coulson, 1977 DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74: 5463-5467.
- Saraste, M., 1999 Oxidative phosphorylation at the fin de siecle. *Science* 283: 1488-1493.
- Sato, M., and K. Sato, 2013 Maternal inheritance of mitochondrial DNA by diverse mechanisms to eliminate paternal mitochondrial DNA. *Biochim Biophys Acta* 1833: 1979-1984.
- Sazanov, L. A., 2007 Respiratory complex I: mechanistic and structural insights provided by the crystal structure of the hydrophilic domain. *Biochemistry* 46: 2275-2288.
- Sazanov, L. A., 2015 A giant molecular proton pump: structure and mechanism of respiratory complex I. *Nat Rev Mol Cell Biol* 16: 375-388.
- Schagger, H., and G. von Jagow, 1991 Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem* 199: 223-231.
- Scherz-Shouval, R., and Z. Elazar, 2007 ROS, mitochondria and the regulation of autophagy. *Trends Cell Biol* 17: 422-427.
- Scherz-Shouval, R., E. Shvets and Z. Elazar, 2007 Oxidation as a post-translational modification that regulates autophagy. *Autophagy* 3: 371-373.
- Schlame, M., and M. Ren, 2006 Barth syndrome, a human disorder of cardiolipin metabolism. *FEBS Lett* 580: 5450-5455.
- Seneca, S., K. Vancampenhout, R. Van Coster, J. Smet, W. Lissens et al., 2015 Analysis of the whole mitochondrial genome: translation of the Ion Torrent Personal Genome Machine system to the diagnostic bench? *Eur J Hum Genet* 23: 41-48.

- Seo, S. B., X. Zeng, J. L. King, B. L. Larue, M. Assidi et al., 2015 Underlying Data for Sequencing the Mitochondrial Genome with the Massively Parallel Sequencing Platform Ion Torrent PGM. *BMC Genomics* 16 Suppl 1: S4.
- Shanske, S., J. Coku, J. Lu, J. Ganesh, S. Krishna et al., 2008 The G13513A mutation in the ND5 gene of mitochondrial DNA as a common cause of MELAS or Leigh syndrome: evidence from 12 cases. *Arch Neurol* 65: 368-372.
- Shoubridge, E. A., T. Johns and G. Karpati, 1997 Complete restoration of a wild-type mtDNA genotype in regenerating muscle fibres in a patient with a tRNA point mutation and mitochondrial encephalomyopathy. *Hum Mol Genet* 6: 2239-2242.
- Silvestri, G., M. Rana, F. Odoardi, A. Modoni, E. Paris et al., 2000 Single-fiber PCR in MELAS(3243) patients: correlations between intratissue distribution and phenotypic expression of the mtDNA(A3243G) genotype. *Am J Med Genet* 94: 201-206.
- Simone, D., F. M. Calabrese, M. Lang, G. Gasparre and M. Attimonelli, 2011 The reference human nuclear mitochondrial sequences compilation validated and implemented on the UCSC genome browser. *BMC Genomics* 12: 517.
- Smirnova, E., L. Griparic, D. L. Shurland and A. M. van der Bliek, 2001 Dynamin-related protein Drp1 is required for mitochondrial division in mammalian cells. *Mol Biol Cell* 12: 2245-2256.
- Song, W., J. Chen, A. Petrilli, G. Liot, E. Klinglmayr et al., 2011 Mutant huntingtin binds the mitochondrial fission GTPase dynamin-related protein-1 and increases its enzymatic activity. *Nat Med* 17: 377-382.
- Stoneking, M., 2000 Hypervariable sites in the mtDNA control region are mutational hotspots. *Am J Hum Genet* 67: 1029-1032.
- Szabadkai, G., A. M. Simoni and R. Rizzuto, 2003 Mitochondrial Ca²⁺ uptake requires sustained Ca²⁺ release from the endoplasmic reticulum. *J Biol Chem* 278: 15153-15161.
- Taanman, J. W., 1999 The mitochondrial genome: structure, transcription, translation and replication. *Biochim Biophys Acta* 1410: 103-123.
- Triepels, R. H., L. P. van den Heuvel, J. L. Loeffen, C. A. Buskens, R. J. Smeets et al., 1999 Leigh syndrome associated with a mutation in the NDUFS7 (PSST) nuclear encoded subunit of complex I. *Ann Neurol* 45: 787-790.
- Turnbull, D. M., and P. Rustin, 2015 Genetic and biochemical intricacy shapes mitochondrial cytopathies. *Neurobiol Dis*.
- Valanne, L., L. Ketonen, A. Majander, A. Suomalainen and H. Pihko, 1998 Neuroradiologic findings in children with mitochondrial disorders. *AJNR Am J Neuroradiol* 19: 369-377.
- Valente, L., D. Piga, E. Lamantea, F. Carrara, G. Uziel et al., 2009 Identification of novel mutations in five patients with mitochondrial encephalomyopathy. *Biochim Biophys Acta* 1787: 491-501.
- Vithayathil, S. A., Y. Ma and B. A. Kaiparettu, 2012 Transmitochondrial cybrids: tools for functional studies of mutant mitochondria. *Methods Mol Biol* 837: 219-230.
- von Kleist-Retzow, J. C., V. Cormier-Daire, P. de Lonlay, B. Parfait, D. Chretien et al., 1998 A high rate (20%-30%) of parental consanguinity in cytochrome-oxidase deficiency. *Am J Hum Genet* 63: 428-435.
- Wallace, D. C., G. Singh, M. T. Lott, J. A. Hodge, T. G. Schurr et al., 1988 Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science* 242: 1427-1430.
- Wang, J., E. S. Schmitt, M. L. Landsverk, V. W. Zhang, F. Y. Li et al., 2012 An integrated approach for classifying mitochondrial DNA variants: one clinical diagnostic laboratory's experience. *Genet Med* 14: 620-626.
- Waterham, H. R., J. Koster, C. W. van Roermund, P. A. Mooyer, R. J. Wanders et al., 2007 A lethal defect of mitochondrial and peroxisomal fission. *N Engl J Med* 356: 1736-1741.
- Wolf, N. I., and J. A. Smeitink, 2002 Mitochondrial disorders: a proposal for consensus diagnostic criteria in infants and children. *Neurology* 59: 1402-1405.
- Yoshii, S. R., C. Kishi, N. Ishihara and N. Mizushima, 2011 Parkin mediates proteasome-dependent protein degradation and rupture of the outer mitochondrial membrane. *J Biol Chem* 286: 19630-19640.

- Zhang, M., E. Mileykovskaya and W. Dowhan, 2005 Cardiolipin is essential for organization of complexes III and IV into a supercomplex in intact yeast mitochondria. *J Biol Chem* 280: 29403-29408.
- Zhu, X., X. Peng, M. X. Guan and Q. Yan, 2009 Pathogenic mutations of nuclear genes associated with mitochondrial disorders. *Acta Biochim Biophys Sin (Shanghai)* 41: 179-187.
- Zorzano, A., and M. Claret, 2015 Implications of mitochondrial dynamics on neurodegeneration and on hypothalamic dysfunction. *Front Aging Neurosci* 7: 101.
- Zuchner, S., I. V. Mersiyanova, M. Muglia, N. Bissar-Tadmouri, J. Rochelle et al., 2004 Mutations in the mitochondrial GTPase mitofusin 2 cause Charcot-Marie-Tooth neuropathy type 2A. *Nat Genet* 36: 449-451.

Summary :

Diagnosis of mitochondrial diseases by sequencing the entire mitochondrial genome using "new generation sequencing" technology

Mitochondrial diseases are a group of genetic disorders that arise from a defect of the mitochondrial respiratory chain. The role of the mitochondrial respiratory chain is to generate ATP molecules, the main cellular source of energy. It is composed of five protein complexes (I,II,III,IV,V), encoded by either nuclear DNA or mitochondrial DNA. Mitochondrial diseases are a group of clinically heterogeneous disorders, commonly associated with the elevation of the lactic acid in the plasma or in the cerebral spinal fluid. It is noteworthy that the activity of mitochondrial complexes can be measured in vitro, in tissues or cellular cultures. However, the absence of metabolic derangement and normal complexes activity do not exclude the diagnosis. Molecular analyses are the most efficient way to confirm mitochondrial disease. However, molecular diagnosis is challenging because of the extreme molecular heterogeneity. Indeed, more than 250 nuclear genes have been implicated in mitochondrial disease and this number is still expanding. Moreover, sequencing of the entire mitochondrial DNA (16.5kb) is not realized in routine laboratory work by Sanger method, which is a very expensive and long process. Thus, only few mutations and large scale rearrangements are commonly investigated. In this project we aimed to develop the new generation sequencing (NGS), a rapid and less expensive method, to study mitochondrial DNA in Bordeaux hospital. We have also evaluated the yield of this new technique to identify new mutations in a group of patients, presented with clinical suspicion of mitochondrial disease.

Résumé:

Les cytopathies mitochondriales sont un groupe de maladies génétiques dont l'origine commune est un déficit de la chaîne respiratoire mitochondriale. Cette dernière a pour rôle essentiel de fournir l'énergie nécessaire à un fonctionnement cellulaire normal sous forme d'ATP. Elle est constituée d'un ensemble de complexes multiprotéiques dont les sous-unités sont codées pour une petite partie (13 sous unités) par le génome mitochondrial (ADNmt), les autres par le génome nucléaire. Les cytopathies mitochondriales ont une expression clinique très hétérogène, le plus souvent multi-systémique, s'associant à des altérations du bilan métabolique telle que l'hyper-lactacidémie. Il est également possible de mettre en évidence le déficit des complexes de la chaîne respiratoire par dosage de leurs activités dans les tissus ou les cellules en cultures. Ces anomalies métaboliques sont cependant loin d'être constantes et des bilans normaux n'éliminent pas une cytopathie mitochondriale. Ceci rend compte de la complexité diagnostique de ces pathologies, seules les explorations moléculaires posent formellement le diagnostic. Les études de l'ADNmt étaient jusqu'à très récemment longues, coûteuses et loin d'être exhaustives. Seuls les grands réarrangements et quelques mutations ponctuelles, dites communes étaient recherchés en routine diagnostique. L'objectif de ce travail était de mettre en place au CHU de Bordeaux, la technique de séquençage de nouvelle génération (NGS) pour l'étude rapide de l'ADNmt dans son ensemble, et ce pour chaque patient suspect de cytopathie mitochondriale. Nous avons dans un second temps évalué l'apport diagnostique de cette nouvelle technique dans une série de patients, ayant une présentation clinique très évocatrice de cytopathie mitochondriale, et pour lesquels les techniques conventionnelles n'avaient pas retrouvé d'anomalies de l'ADNmt

Discipline: génétique médicale

Mots clés : Cytopathies mitochondriales, génome mitochondrial, séquençage de nouvelle génération
NGS