

Université de Montréal

**Étude de nouvelles maladies neurogénétiques chez les  
Canadien français**

par

Martine Tétreault

Biologie Moléculaire

Faculté de Médecine

Thèse présentée à la Faculté Médecine  
en vue de l'obtention du grade de doctorat  
En biologie moléculaire

avril, 2013

© Martine Tétreault, 2013

## Résumé

Depuis déjà plusieurs décennies, nous sommes en mesure d'identifier les mutations responsable de diverses maladies mendéliennes. La découverte des gènes responsables de ces maladies permet non seulement un meilleur diagnostic clinique pour ces familles, mais aussi de mieux comprendre les mécanismes physiopathologiques de ces maladies ainsi que mieux définir la fonction normale des gènes causales. Ultimement, ces découvertes mènent à l'identification de cibles thérapeutiques pour le traitement de ces maladies. Les progrès technologiques sont depuis toujours un facteur très important dans la découverte de ces gènes mutés. De l'approche traditionnelle de clonage positionnel en passant par la première séquence du génome humain et maintenant les technologies de séquençage à grande échelle, de plus en plus de maladies ont maintenant une entité génétique. Dans le cadre de ce projet de doctorat, nous avons utilisé tant les approches traditionnelles (leucodystrophies) que les nouvelles technologies de séquençage (polyneuropathie douloureuse) qui ont mené à l'identification du gène causal pour plusieurs de nos familles. L'efficacité de ces deux approches n'est plus à démontrer, chacune d'entre elles possèdent des avantages et des inconvénients. Dans le cadre de ces projets, nous avons utilisé la population canadienne-française connue pour ces effets fondateurs et la présence, encore aujourd'hui, de grandes familles. Les différents projets ont permis d'établir certains avantages et inconvénients quant à l'utilisation de ces techniques et de la population canadienne-française. Dans le cadre d'un phénotype assez homogène et bien défini comme celui du projet leucodystrophie, l'approche traditionnel par gène candidat nous a permis d'identifier le gène causal, *POLR3B*, sans trop de difficulté. Par contre, pour les autres projets où nous sommes en présence d'une hétérogénéité clinique et génétique une approche non-biaisée utilisant le séquençage exomique a obtenu un plus grand succès. La présence de grandes familles est un grand avantage dans les deux approches. Dans le projet polyneuropathie douloureuse, une grande famille originaire du Saguenay-Lac-St-Jean nous a permis d'identifier le gène *NAGLU* comme responsable suite à l'exclusion des autres variants candidats par analyse de ségrégation. Comme *NAGLU* était déjà associé à un phénotype qui diffère sur plusieurs points à celui de

notre famille, une approche traditionnelle n'aurait pas été en mesure d'identifier *NAGLU* comme le gène causal. Dans l'analyse de nos données de séquençage exomique, nous avons observé que plusieurs variants rares, absents des bases de données, étaient partagés entre les différents individus Canadiens français. Ceci est probablement dû à la démographie génétique particulière observée chez les Canadiens français. En conclusion, les technologies de séquençage à grande échelle sont avantageuses dans l'étude de maladies hétérogènes au niveau clinique et génétique. Ces technologies sont en voie de modifier l'approche d'identification de gènes en permettant une analyse de génétique inversée, c'est-à-dire de la génétique vers la clinique.

**Mots-clés** : Clonage positionnel, séquençage exomique, polyneuropathie, douleur, leucodystrophies, Canadien-Français.

## Abstract

Since many decades we are able to identify mutations responsible for Mendelian diseases. The identification of the causative gene not only allows a better diagnostic to these families, but also allows a better understanding of the pathophysiological mechanisms of these disorders and of the normal function of a gene. Ultimately, those discoveries lead to the identification of therapeutic targets that will enable clinicians to treat these diseases. Technological progress has forever driven gene identification. Starting with traditional approaches like positional cloning passing through the first sequence of the human genome and finally the recent high-throughput sequencing technologies, more and more diseases are now linked to a genetic cause. During my PhD, I had the opportunity to use traditional approaches (leukodystrophy) as well as new sequencing technologies (painful sensory polyneuropathy) which both led to gene identification for some of our families. The efficiency of these approaches is well known, each of them has advantages and disadvantages. In these projects, we used the french-canadian population well known for its founder effect and the presence, still today, of large families. Working on these projects allowed us to establish advantages and disadvantages concerning the use of those technologies and the french-canadian population. In the presence of a homogenous and well defined phenotype, like the leukodystrophy project, the traditional approach enables us to rapidly identify the causative gene (*POLR3B*). On the other hand, the unbiased exome sequencing approach has had more success for diseases characterized with clinical and genetic heterogeneity. Large families are a great advantage for both methods. In the painful sensory polyneuropathy project, a large family originated from the Saguenay-Lac-St-Jean region allowed us to identify the *NAGLU* gene as responsible for the disease after exclusion of the candidate variants by segregation analysis. *NAGLU* has already been associated with a phenotype that differs in many points with the clinical features observed in our family. In this case a traditional approach would have failed to identify *NAGLU* as the causative gene. In the analysis of our exome sequencing results, we observed many rare variants absent from databases but shared between french-canadian individuals. This enrichment in rare variants is probably due to the particular

genetic demography of Quebec. In conclusion, high-throughput sequencing technologies are advantageous in the study of clinically and genetically heterogeneous diseases. These technologies are changing the gene identification approach towards reverse genetics, meaning genetics towards clinic.

**Keywords:** Positional cloning, exome sequencing, neuropathy, pain, leukodystrophy, French-Canadian.

# Table des matières

Résumé.....	i
Abstracts .....	ii
Liste des tableaux.....	vi
Liste des figures .....	vii
Liste des abréviations .....	viii
Remerciements .....	x

<b>Chapitre1:Avancement en génétique de l'étude des maladies mendéliennes .....</b>	<b>1</b>
1.1. Approches traditionnelles.....	2
1.2. Séquençage de nouvelle génération .....	5
1.3. Séquençage de nouvelle génération et l'étude des maladies neurologiques.....	6
1.4. Séquençage de nouvelle génération et diagnostic moléculaire.....	9
1.5. Effet fondateur Québécois.....	11
1.6. Mes projets de thèse dans le contexte des progrès technologiques .....	13
<b>Chapitre 2: Leucodystrophies .....</b>	<b>14</b>
2.1 Résultats préliminaire: Description d'une nouvelle forme de leucodystrophie avec effet fondateur québécois et mutations dans le gène <i>POLR3A</i> .....	15
2.2. Recessive mutation in <i>POLR3B</i> encoding the second largest subunit of POL III, cause a rare hypomyelinating leukodystrophie .....	18
2.3. Mutations in <i>POLR3A</i> and <i>POLR3B</i> are a major cause of hypomyelinating leukodystrophies with or whitout dental abnormalities and/or hypogonadotropic hypogonadism.....	27
2.4. Les leucodystrophies liées à l'enzyme POL III représente un groupe important de leucodystrophies hypomyélinisantes .....	40

<b>Chapitre 3: Polynévrite et ataxie sensitive tardive causé par une mutation dominante dans le gène <i>NAGLU</i></b> .....	42
3.1. Adult onset painful sensory neuropathy caused by a dominant <i>NAGLU</i> mutation .....	45
3.2. Phénotype intermédiaire chez les porteurs de mutations associées aux maladies de l'entreposage lysosomale .....	54
Discussion: Approches traditionnelles et approches de nouvelle génération pour l'identification de gènes de maladies mendéliennes .....	56
4.1. Des approches traditionnelles au séquençage de nouvelle génération .....	57
4.2. Utilisation des familles CF pour l'identification de gènes candidats .....	61
4.3. Impact des technologies NGS sur la clinique et la médecine personnalisée.....	62
<b>ANNEXE 1. Curriculum vitae</b> .....	I
.....	XVI

## Liste des tableaux

Table I: Utilisation des techniques NGS dans l'étude des maladies du désordre du mouvement.

Table II: Liste des tests cliniques disponible aux États-Unis utilisant le séquençage ciblé.

Table III: Maladies neurologiques avec effet fondateur au Québec.

Table IV: Leucodystrophies hypomyélinisantes.

Table V: Table clinique des patients atteint du syndrome TACH.

Table VI: Mutations identifiées dans le gène *POLR3A*.

Table VII: Clinical and MRI features of the patients.

Table VIII: Mutations identified in patients with hypomyelinating leukodystrophies.

Table IX: Genotypes of SNPs in patients with the p.V523E mutation in *POLR3B*.

Table X: Neuropathies affectant les petites fibres acquises ou de causes génétiques.

Table XI: Clinical features of sensory ataxia cases.

Table XII: Statistical analysis of exome sequencing results.

Table XIII: Candidate variants list uncovered by WES analysis.

Table XIV: NAGLU activity in patients' leucocytes.



## Liste des figures

Figure 1: Illustration chronologique des avancements en génétique humaine des quatre dernières décennies.

Figure 2: IRM des patients atteints de TACH.

Figure 3: *POLR3B* mutations.

Figure 4: Teeth abnormalities.

Figure 5: MRI characteristics

Figure 6: Magnetic resonance imaging of patient P3.

Figure 7: Multi-species alignment of *POLR3A* and *POLR3B*.

Figure 8: Sequence traces of *POLR3B* splicing variants.

Figure 9: Pedigree of a large FC family affected by a late painful sensory polyneuropathy.

Figure 10: Protein location of p.Ile403Thr *NAGLU* mutation.

Figure 11: Variants identifiés par séquençage exonique et partagés par 10 individus CF non-apparentés, atteints de maladies différentes et originaires de différentes régions de la province.

## Liste des abréviations

4H: Hypomyélinisation, hypodontie, hypogonadisme et hypogonadotrope  
ACMG: Collège américain de génétique médicale  
AD: Autosomal dominant  
ADN: acide désoxyribonucléique  
ALS: Sclérose latérale amyotrophique  
AR: Autosomal récessif  
ARN: Acide ribonucléique  
ARNT: Acide ribonucléique de transfert  
BWA: Burrows wheeler alignment  
CCDS: Séquence codante consensus  
CDS: Région codante  
CF: Canadienne-Française  
CMT: Charcot-Marie-Tooth  
CNV: Variation de nombre de copies  
DARS2: Aspartyl tARN synthétases mitochondriale  
DRG: Ganglion spinal  
EVS: Exomevariants server  
GBA: Acide glucosidase beta  
GLA: Allpha-galactosidase  
HCAHC: Hypomyélinisation avec atrophie cérébelleuse et hypoplasie du corps calleux  
HD: Maladie Huntington  
HGP: Projet séquençage humain  
HS: Heparan sulfate  
HTT: Huntingtine  
IRM: Résonance magnétique  
LO: Leucodystrophie avec oligodontie  
MD: Désordre du mouvement  
MPS IIIB: Mucopolysaccharidose IIIB  
mtDNA: ADN mitochondriale  
NASDO: Neuropathie sensitive avec ataxie, dysarthrie et ophtalmoplégie  
NGS: Séquençage de nouvelle génération  
NHSA: Neuropathie héréditaire sensitive et autonome  
NIH: Institut national en santé des États-Unis  
NPF: Neuropathie de petite fibre  
NTRK1: Récepteur de type I tyrosine kinase neurotrophique  
PAF: Polyneuropathie amyloïdes familiale  
PD: Maladie de Parkinson  
PMD: Pelizaeus-Merzbacher  
POL III: Polymérase III

POLG: Polymérase gamma  
POLR3A: Sous-unité A de l'enzyme polymérase III  
POLR3B: Sous-unité B de l'enzyme polymérase III  
POLR3D: Sous-unité D de l'enzyme polymérase III  
POLR3E: Sous-unité de l'enzyme polymérase III  
RAB7A: Protéine liée à RAS Rab-7a  
RFLP: Polymorphisme de longueur par fragment de restriction  
SCA: Ataxie spinocérébelleuse  
SLSJ: Saguenay Lac-St-Jean  
SNC: Système nerveux central  
SNP: Polymorphisme de simple nucléotide  
SPG: Paraplégie spastique  
SPTLC1: Sérine palmoyltransférase longue chaîne sous-unité A  
STS: microsatellite  
TR: Séquençage ciblé  
TTR: Transthyrérine  
WB: Immunobuvardage de type Western  
WES: Séquençage exonique  
WGS: Séquençage du génome entier  
XL: Lié au chromosome X

## Remerciements

Je voudrais tout d'abord remercier mon superviseur Dr. Bernard Brais pour les dix années passé dans son laboratoire. Bernard m'a accueilli dans son laboratoire et m'a confié un projet très intéressant et stimulant. Je remercie Bernard pour la confiance qu'il a eu en moi depuis les dix dernières années. D'étudiante à la maîtrise, assistante de recherche et finalement étudiante au doctorat j'ai beaucoup appris et acquis d'autonomie grâce à cet environnement de recherche. Les prochaines étapes de ma formation nous permettra de collaborer et j'espère mèneront à de belles découvertes.

Un très gros merci à Dr. Geneviève Bernard qui m'a donné l'opportunité de collaborer avec elle sur le projet leucodystrophie qui a mené à de très belles découvertes. Dans cette collaboration s'est développée une relation d'amitié qui a résulté en beaucoup de discussions scientifique accompagné d'une coupe de vin. Geneviève je te souhaite beaucoup de succès dans tes futures recherches.

Je voudrais remercier Dr. Hussein Daoud. Une amitié importante s'est développée de cette relation de travail. Ses connaissances scientifiques, son écoute et son support ont été très important pour moi dans les derniers miles de cette thèse. Je te souhaite tout le succès que tu mérites pour ta carrière.

Je voudrais aussi remercier tous les membres du laboratoire et tous les anciens membres que j'ai eu l'opportunité de côtoyer pendant mes dix années passé au laboratoire.

Un merci spécial aux familles et collaborateurs. Sans eux ces découvertes ne serait pas possible.

Finalement, je voudrais remercier tous mes amis et ma famille pour leur support et compréhension. Leur soutien fut très important tout au long de ce processus. Un merci

spécial à Matthieu Boucher qui a été de bons conseils lors de ma décision de poursuivre au doctorat. Aucun mot n'est assez fort pour exprimer la reconnaissance que j'ai pour mes parents, Donna et Michel. Leur support, encouragements et amour inconditionnel m'ont donné la force de continuer à travailler fort même dans les moments difficile.

## **Chapitre 1:**

Avancements en génétique de l'étude  
des maladies mendéliennes

L'identification des causes génétiques de maladies mendéliennes est depuis longtemps un grand défi de la médecine. L'impact le plus immédiat d'une telle découverte est la possibilité d'offrir un test diagnostique aux familles atteintes et dans le cas de maladies congénitales, un diagnostic prénatal<sup>1</sup>. De plus, l'identification d'un gène causal permet de mieux comprendre les mécanismes physiopathologiques qui causent la maladie ainsi que la fonction normale de ce gène<sup>2</sup>. La découverte de gènes mutés est le premier chaînon à l'identification de cibles thérapeutiques menant ultimement à un traitement de ces maladies.

Il est important d'étudier les maladies mendéliennes puisque même si la majorité sont rares, ensemble elles contribuent significativement aux maladies présentes dans la société. On estime qu'il existe environ 7 000 maladies mendéliennes dont plus de 3 000 ont déjà une cause génétique connue (Online Mendelian Inheritance in Man; OMIM)<sup>3</sup>. Les avancements technologiques des dernières décennies ont largement contribué à la découverte de ces gènes. Plusieurs projets d'envergure ont contribué de façon majeure à modifier les stratégies d'identification de gènes pour ces maladies.

### **1.1 Approches traditionnelles**

Les approches traditionnelles d'identification de gènes (karyotypage, clonage positionnel, analyse de liaison, cartographie par homozygotie et analyse de variation de nombre de copies (CNV)) ont largement contribué à l'identification de gènes causant des maladies mendéliennes dans les dernières décennies<sup>1</sup>. L'un des premiers avancements technologiques à avoir influencé de façon importante la recherche en génétique est la technique de séquençage de type Sanger. Cette technique où l'on séquence l'ADN par l'incorporation sélective de didésoxyribonucléotides par un ADN polymérase a été développée par Frederick Sanger en 1977<sup>4</sup>. Encore aujourd'hui, cette technique est la méthode de choix pour l'identification de mutations en recherche et en clinique (Figure 1). Les travaux de Frederick Sanger ont ouvert la voie à plusieurs développements technologiques servant à améliorer la capacité de lier une maladie à une cause génétique.

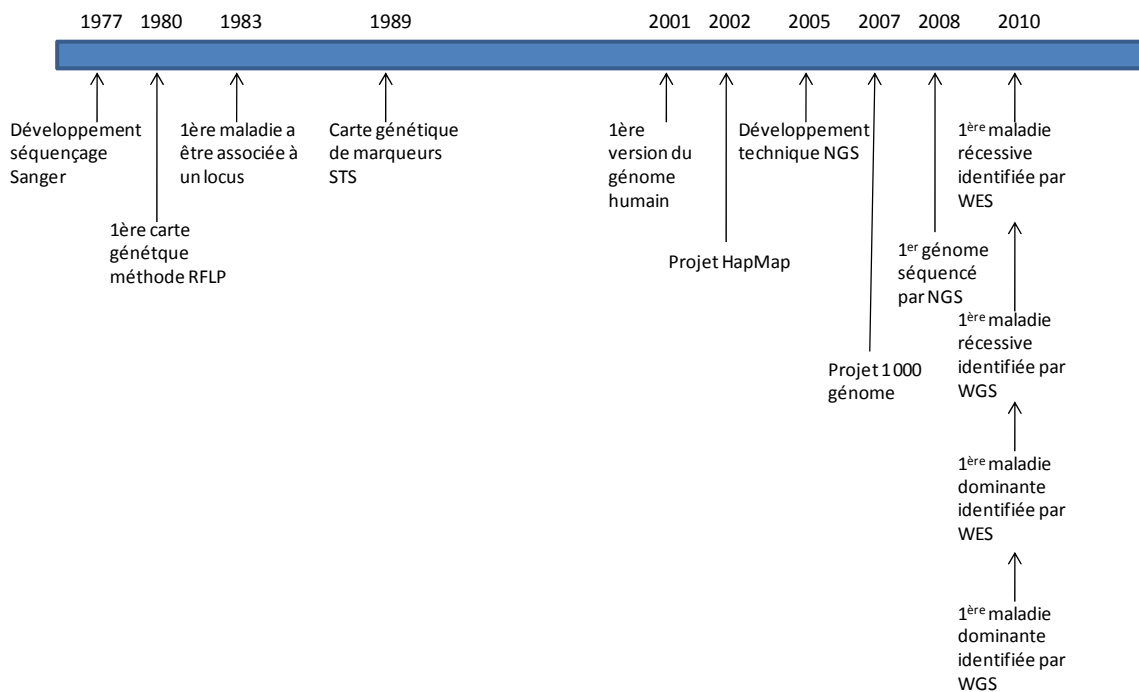
La première carte génétique a été décrite il y a plus de 30 ans, en 1980, en utilisant la technique de polymorphisme de longueur par fragments de restriction (RFLP)<sup>1; 5</sup> (Figure 1). Ceci permettait d'étudier la ségrégation de certains marqueurs avec un trait et ainsi localiser le gène causal. En 1983, une maladie a été associée à un locus chromosomique pour la première fois. Il s'agit d'une liaison de la maladie d'Huntington (HD; MIM 143100), une maladie neurodégénérative à transmission dominante, au chromosome 4<sup>6</sup>. Il a fallu 10 ans, soit en 1993, pour que les expansions de triplet CAG dans le gène Huntingtine (*HTT*) soient découvertes dans la maladie d'Huntington. La production en 1989 d'une carte de marqueurs microsatellites (STS) dispersés sur l'ensemble du génome à grandement contribué à améliorer la capacité de lier une maladie à une région chromosomique<sup>1; 7</sup> (Figure 1).

C'est dans les années 80 que la communauté scientifique a entamé des discussions sur les avantages qui découleraient de la connaissance de la séquence complète du génome humain pour la compréhension de la biologie humaine et des maladies<sup>2</sup>. C'est en 2001 que le projet d'envergure de séquençage du génome humain (HGP) a offert à la communauté scientifique la première séquence du génome humain (Figure 1). La possibilité d'utiliser une séquence de référence lors de la recherche de mutations a grandement accéléré la découverte de gènes associés à des maladies mendéliennes. Le projet HGP a été précurseur à d'autres projets internationaux à grande échelle comme le projet HapMap et le projet 1000 génomes. Dans le cadre du projet HapMap qui a vu le jour en 2002 (Figure 1), plusieurs laboratoires ont joint leurs efforts afin d'identifier la position chromosomique des polymorphismes de simple-nucléotide (SNP) et d'étudier la diversité des haplotypes présents dans le génome humain<sup>8</sup>. L'identification de plus de huit millions de SNP communs (fréquence allélique de plus de 5%) a permis le développement de puces de génotypage SNP de grande densité. Cet outil de recherche est très utile dans l'étude des maladies complexes, mais il est aussi utilisé pour des analyses de liaison classiques et de cartographie par homozygotie dans l'étude des maladies mendéliennes. Les puces SNP ont graduellement



remplacé l'utilisation des marqueurs STS, modifiant ainsi les analyses de génotypage classique.

Les approches traditionnelles menant à l'identification de loci (analyse de liaison, cartographie par homozygotie) combinées au séquençage Sanger ont eu un grand succès dans l'identification de mutations causant les maladies mendéliennes. Par contre, les limites de ces techniques font en sorte que plusieurs maladies demeuraient sans diagnostic moléculaire. D'autres progrès technologiques ont donc été nécessaires pour permettre de continuer ces recherches.



**Figure 1:** Illustration chronologique des avancements en génétique humaine des quatre dernières décennies (Adaptation de Gilissen C. et al)<sup>1</sup>.

## 1.2 Séquençage de nouvelle génération

Même si les approches traditionnelles ont grandement contribué à l'avancement des connaissances en génétique humaine, ces techniques ne sont pas en mesure de permettre l'identification de tous les types de mutations. En 2005, l'apparition des techniques de séquençage de nouvelle génération (NGS) a eu un impact majeur sur la recherche en génomique<sup>1</sup> (Figure 1). Il y a deux approches de séquençage non-biaisées pour l'identification de mutations: le séquençage exomique (WES) et le séquençage du génome entier (WGS). Bien que ces nouvelles technologies aient aussi des limites, elles répondent à un besoin que les approches traditionnelles ne peuvent combler. L'utilisation de WES permet de séquencer en une seule étape toutes les régions codantes du génome. Bien que les exons représentent seulement une petite portion du génome (1%), c'est dans ces régions que la majorité des variants pathogènes présentement identifiés se retrouvent<sup>1</sup>. Le WES est rapidement devenu l'outil principal pour l'analyse génétique des maladies mendéliennes. L'approche ultime pour détecter toutes les variations génomiques reste tout de même le WGS, puisque autant les régions exoniques que les introniques sont analysées.

Les premiers instruments NGS ont principalement été utilisés pour des projets à grande échelle comme le projet 1000 génomes (Figure 1). Ce projet collaboratif, qui a débuté en 2007 et qui regroupait des laboratoires de plusieurs pays, avait pour but de produire un catalogue de variants afin de mieux comprendre la diversité du génome<sup>8</sup>. Pour ce faire, les populations étudiées lors du projet HapMap ainsi que de nouvelles populations ont été étudiées par WGS. C'est en 2008 que la technologie NGS nous a donné la première séquence d'un génome humain. Il s'agit du génome de James Watson, le codécouvreur de la structure de l'ADN. La banque de variants 1000 génomes ainsi que le génome de Watson sont maintenant disponibles au public et sont utilisés comme référence dans la recherche de variants rares causant des maladies mendéliennes.

Il a fallu attendre l'année 2009 pour que ces technologies soient disponibles à l'ensemble de la communauté scientifique. Ceci a grandement facilité l'identification de

gènes causant des maladies mendéliennes rares. En 2010, les premiers gènes causant des maladies récessives et dominantes ont été identifiés par WES et WGS<sup>1</sup> (Figure 1). En fait, après seulement deux années, plus de 30 nouveaux gènes associés à des maladies mendéliennes avaient été identifiés par WES. Non seulement les techniques NGS ont permis une expansion dans l'identification de gènes causals, mais elles ont aussi fourni un nouvel aperçu de la génétique des maladies puisque les limites des approches traditionnelles ont biaisé nos connaissances.

Les projets d'envergure et les avancements technologiques des dernières décennies ont directement influencé les découvertes en génétique humaine. La découverte de gènes causant les maladies mendéliennes a augmenté significativement avec chacun de ces avancements technologiques.

### **1.3 Séquençage de nouvelle génération et l'étude des maladies neurologiques**

Obtenir une meilleure compréhension des mécanismes physiopathologiques des maladies neurologiques est depuis longtemps la priorité de plusieurs études génétiques<sup>9</sup>. L'utilisation des approches traditionnelles a obtenu beaucoup de succès dans les dernières années, menant à l'identification de plusieurs gènes et facteurs de risque associés à des maladies mendéliennes touchant le système nerveux. Dues aux limites de ces techniques et à l'hétérogénéité des maladies neurologiques, plusieurs patients demeurent sans un diagnostic moléculaire. Le projet pilote "UndiagnosedDiseases" de l'Institut national en santé des États-Unis (NIH) démontre bien cette réalité des maladies impliquant le système nerveux. Dans ce projet, des individus chez lesquels les techniques de diagnostic moléculaire actuelles ont échoué à l'identification du gène causal ont été recrutés dans le but de démontrer l'efficacité de la technologie NGS pour le diagnostic clinique. Dans la première année, 160 individus ont été recrutés, dont 53% étaient atteints d'une maladie neurologique<sup>9</sup>.

L'une des problématiques dans l'étude génétique des maladies neurologiques est la grande hétérogénéité clinique et génétique. Les maladies associées à un désordre du mouvement (MD) sont un groupe de maladies démontrant cette hétérogénéité. Ce groupe de maladies est caractérisé par une incapacité à produire ou contrôler un mouvement et est souvent accompagné d'autres présentations cliniques (crise épileptique, retard mental, déficience immunitaire, symptômes psychiatriques, etc.)<sup>10</sup>. Les MD regroupent plusieurs syndromes dont l'ataxie, la dystonie, le tremblement essentiel, la maladie d'Huntington, l'atrophie multisystème, le Parkinson, le syndrome des jambes sans repos et le syndrome de Tourette. Dans le cas des leucodystrophies, l'atteinte de la matière blanche est souvent associée à des symptômes MD. L'hétérogénéité clinique est observée dans plusieurs sous-groupes de MD où un seul gène est associé à plusieurs phénotypes. Dans le cas de l'hétérogénéité génétique, les principales caractéristiques cliniques d'un patient sont associées à plusieurs gènes. Par exemple, la maladie Charcot-Marie-Tooth (CMT), caractérisée par une atteinte sensitive distale et une atrophie musculaire, a été associée à plus de 40 loci et une grande variabilité phénotypique intrafamiliale est observée, ce qui rend le diagnostic avec les méthodes traditionnelles difficile et laborieux<sup>10</sup>. À ce jour, les gènes associés aux MD expliquent seulement un faible pourcentage des cas indiquant que plusieurs gènes restent à être identifiés. La technologie NGS semble être l'approche idéale pour identifier les gènes liés aux MD héréditaires. Ceci aura pour effet d'améliorer le diagnostic et la compréhension du processus pathologique de ce groupe de maladies neurologiques. Les technologies NGS ont déjà contribué à l'expansion des gènes candidats associés aux MD en identifiant des nouveaux gènes et en associant des nouveaux phénotypes neurologiques avec des gènes déjà connus (Table I).

**Table I:** Utilisation des techniques de NGS dans l'étude des maladies du désordre du mouvement.

<b>NOUVEAUX GÈNES DE MD IDENTIFIÉS PAR NGS</b>				
<b>Disease</b>	<b>Mode</b>	<b>Method</b>	<b>Gene</b>	<b>OMIM</b>
Sclérose latérale amyotrophique (ALS) avec démence frontotemporale	AD	TR	c9orf72	614260
Epilepsie infantile familiale bénigne et convulsion infantile avec syndrome choreoathetosis	AD	TR	PRRT2	614386
CMT type II	AD	WES	DYNC1H1	600112
Dyskinésie paroxysmal kinésigénique	AD	WES TR	PRRT2	614386
Parkinson	AD	WES	VPS35	601501
Ataxie spinocérébelleuse (SPG)	AD	WES	TGM6	613900
Tourette/ tic chronique	AD	WES	DNAJC13	614334
			OFCC1	614287
Calcification idiopathique des ganglions basal	AD	WES	PDGFRB	615007
Dystonie de torsion	AD	WES	GNAL	615073
Dystonie craniocervical	AD	WES	ANO3	615034

<b>GÈNES CONNUS DE MD ASSOCIÉS AVEC UN NOUVEAU PHÉNOTYPE</b>					
<b>Disease</b>	<b>Phénotype précédent</b>	<b>Mode</b>	<b>Method</b>	<b>Gene</b>	<b>OMIM</b>
Myopathie avec inclusion de corps associé à un syndrome Paget précoce	ALS	AD	WES	VCP	601023
Paraplégie spastique (SP)	neurodégénération associée à l'hydroxylase des acides gras	AR	WES	FA2H	611026
Développement de tumeurs	leucoencéphalopathie diffuse avec sphéroïde héréditaire	AD	WES	CSF1R	164770
Retard mental (ID) et autisme	SP héréditaire type 30/ Neuropathie sensitive héréditaire type II	AR	WES	KIF1A	601255
Paraplégie spastique/Parkinson	ALS juvénile	AR	WES	SPG11	610844
ID lié au chromosome X/ syndrome Allan-Herndon-Dudley/PMD-like	Pelizaues-Merzbacher (PMD)	XL	WES	MCT8	300095
Acidurie glutarique type I	Dystonie progressive généralisée	AR	WES	GCDH	608801
SPG type 28	Ataxie spastique et neuropathie	AR	WES	AFG3L2	604581
Paralysie cérébrale tétraplégie spastique avec ID	SP héréditaire type 47	AR	WES	AP4B1	607245
SP avec ID sévère	phénotype neurologique multiple	AR	WES	AP4S1	607243
Atrophie cérébrale, macrocéphalie, crise, retard développemental	SPG	AR	WES	SYT14	610949
Épilepsie myoclinique progressive	Dystrophie musculaire congénitale	AR	WES	COL6A2	N/A
Ataxie épisodique/ migraine hémiplégique	Épilepsie infantile familiale bénigne et convulsion infantile avec syndrome choreoathetosis	AD	WES	PRRT2	N/A
tremblement essentiel	ALS	AD	WES	FUS	614782

adaptation de Krebs CE. Et al

AD: autosomal dominant

TR: séquençage ciblé

AR: autosomal récessif

#### 1.4 Séquençage de nouvelle génération et diagnostic moléculaire

Bien que le séquençage de type Sanger est encore aujourd'hui la méthode de choix dans le diagnostic moléculaire, les technologies de séquençage NGS font rapidement leur place<sup>11; 12</sup>. L'une des raisons pour cet engouement de l'utilisation des technologies NGS en clinique est qu'elles ont le potentiel d'identifier plusieurs types de mutations dans le génome entier d'un patient en une seule expérience. L'un des effets bénéfiques à court terme de l'utilisation des technologies NGS en clinique est la possibilité d'offrir un diagnostic précis, ce qui a beaucoup de valeur pour les patients, même en l'absence de thérapie disponible<sup>2</sup>.

Toutes les technologies de séquençage NGS ont le potentiel d'être utilisées dans un contexte de diagnostic moléculaire. En ce moment, le WES et le séquençage ciblé sont utilisés plus fréquemment. Le premier pas des technologies NGS en diagnostic moléculaire a été le séquençage ciblé. Il était logique de passer d'un test évaluant un gène à la fois à un test permettant d'analyser une multitude de gènes. Le séquençage ciblé est très utile en présence de maladies hétérogènes puisqu'il permet d'analyser un panel de gènes causant des maladies avec des phénotypes qui se chevauchent. Le WES permet aussi de faire une analyse ciblée des gènes, mais parfois une meilleure couverture des gènes d'intérêt est obtenue avec le séquençage ciblé. De plus, le WES expose le laboratoire diagnostique à d'éventuelles découvertes fortuites. Bien que le collège américain de génétique médicale (ACMG) aient effectué des recommandations, aucun consensus au niveau éthique n'est disponible en ce moment à savoir si ces découvertes devraient être divulguées aux patients<sup>3</sup>. Certains groupes pensent que dans un futur proche, l'application du séquençage ciblé ainsi que du WES sera obsolète et remplacée par le WGS.

L'utilisation des technologies NGS en diagnostic moléculaire est déjà implantée dans plusieurs laboratoires cliniques. En date de septembre 2012, environ 15 laboratoires de diagnostic clinique aux États-Unis offraient des tests NGS pour plus de 50 maladies (Table II). Chaque test utilise la technique de séquençage ciblé et regroupe dans un panel des dizaines à

des centaines de gènes. Plusieurs tests pour le diagnostic de maladies affectant les systèmes neurologique, neuromusculaire et métabolique ont été développés.

Les technologies NGS font rapidement leur place en diagnostic moléculaire. Ils offrent beaucoup d'avantages en comparaison avec le séquençage Sanger, comme la possibilité d'analyser un grand nombre de gènes rapidement et à plus faible coût. Par contre, plusieurs défis restent à être surmontés avant que les technologies NGS deviennent les méthodes de choix en diagnostic clinique, comme l'entreposage et la gestion de la grande quantité de données et l'interprétation pathologique des variants identifiés<sup>13</sup>.

**Table II:** Liste des tests cliniques disponibles aux États-Unis utilisant le séquençage ciblé.

Groupe de maladies	Maladies	Nombre de gènes testés
Cancer	Cancer héréditaire (exemple: sein, colon, ovarien)	10-50
Maladies cardiaque	Cardiomyopathies	50-70
	Arrythmies (exemple: syndrome long QT)	30-10
	Aortopathies (exemple: syndrome Marfan)	10
Maladies du système immunitaire	Syndrome d'immunodéficience sévère	18
	Fièvre périodique	7
Maladies neurologique, neuromusculaire et métabolique	Ataxie	40
	Énergétique cellulaire, métabolique	656
	Maladies congénitale de glycosylation	23-28
	Démence (exemple: Parkinson, Alzheimer)	32
	Retard développementale, autisme, ID	30-150
	Épilepsie	53-130
	Neuropathie héréditaire	34
	Microcéphalie	11
	Maladies mitochondriale	37-450
	Dystrophie musculaire	12-45
Maladies sensitive	Maladies des yeux (exemple: retinis pigmentosa)	66-140
	Surdité et syndrome relié	23--72
Autres	Rasopathies (exemple: Noonan's syndrome)	10
	Maladies pulmonaire (exemple: fibrose kystic)	12-40
	Petite stature	12

adaptation de Rehm et al

## 1.5 Effet fondateur Québécois

La population canadienne-française (CF), bien connue pour ses effets fondateurs régionaux, est depuis longtemps utilisée pour des études génétiques. Grâce à la plus grande homogénéité génétique régionale et la présence de grandes familles, la population CF a participé à l'identification de plusieurs loci et gènes. C'est la démographie génétique particulière au Québec qui a créé ces effets fondateurs<sup>14</sup>.

Entre les années 1608 (fondation de la ville de Québec) et 1759 (prise de possession de la colonie par les Anglais), la Nouvelle-France a reçu 25 000 colons. Parmi ceux-ci, seulement de l'ordre de 8 500, dont 1 600 femmes, s'y sont installés de manière permanente. La majorité des immigrants Français étaient originaires des villes portuaires longeant l'Atlantique ou de la région de Paris. Les deux-tiers du pool génétique de la population CF actuelle prend ses origines des 2 600 colons qui sont arrivés en Nouvelle-France avant 1680<sup>14</sup>. Donc, l'homogénéité génétique relative des québécois est principalement due au faible nombre de colons ayant contribué plutôt qu'à la consanguinité. Sur les siècles on peut parler d'un apparemment important surtout dans chacune des régions et non de consanguinité au sens strict.

Au moment de la conquête des Anglais en 1759, les Canadiens français étaient surtout installés sur des terres agricoles longeant le fleuve St-Laurent et le Richelieu. Au XIX<sup>e</sup> siècle, les Canadiens français ont migré à l'intérieur des terres peuplant des régions plus éloignées comme le Saguenay Lac-St-Jean (SLSJ)<sup>14</sup>. Les barrières religieuses et linguistiques combinées au très petit nombre de nouveaux colons d'origines autres ont fait en sorte qu'il n'y a pas eu de mélange entre les nouveaux arrivants et la population déjà installée. Cette absence de métissage entre les populations et la tendance lourde des sociétés agricoles régionales de rester enclavées sur la terre patrimoniale ont limité la mobilité surtout dans les régions éloignées



et ont contribué à établir un plus haut niveau d'apparentement et donc à l'homogénéité génétique plus grande observée aujourd'hui entre les Canadiens français.

Due à cette homogénéité génétique, nous observons dans la population CF un enrichissement de certains allèles rares; c'est-à-dire des allèles qui ne sont pas observés dans les autres populations ou avec une fréquence allélique très basse. Cet enrichissement d'allèle peut être spécifique à une région du Québec donnant ainsi lieu à un effet fondateur micro-régional. Parmi les allèles ayant subi un enrichissement certains ont été associés à des maladies mendéliennes rares, faisant en sorte que ces maladies soient plus fréquentes au Québec ou même plus fréquentes dans une région en particulier<sup>14</sup>. Dans ces maladies à effet fondateur micro-régional, une mutation explique la majorité des cas et le taux de porteur pour cette mutation est très élevé, pouvant atteindre une personne sur 20. Plusieurs maladies à effet fondateur au Québec sont en fait des maladies affectant le système nerveux, ce qui démontre l'importance d'étudier les maladies neurologiques dans cette population (Table III). Cette homogénéité génétique dans la population CF est un avantage pour les analyses de liaison et de clonage positionnel et a ainsi permis l'identification de plusieurs gènes causant des maladies mendéliennes.

**Table III:** Maladies neurologiques avec effet fondateur au Québec.

Maladie	Mode	Gène	OMIM	Mutations
Neuropathie sensitive et autonome héréditaire de type II	AR	HSN2	201300	c.943C-T; c.981_919insA
Neuropathie liée au chromosome X	X	GJB1	302800	p.Ser26Trp
Neuropathie optique de Leber	mtDNA	ND6	535000	mt.T14484C; mtG11778A
CMT4C	AR	SH3TC2	601596	p.Arg954*
Ataxie spastique de Charlevoix-Saguenay	AR	SACS	270550	c.6594delT
Ataxie de Friedreich	AR	FRDA	229300	expansion GAA
Dystrophie musculaire oculopharyngée	AD	PABPN1	164300	expansion GCN
Dystrophie myotonique	AD	DMPK	160900	expansion CTG
Syndrome du X-fragile	X	FMR1	309550	expansion CGG
Dystrophie musculaire des ceintures 2D	AR	SGCA	608099	Arg77Cys
Syndrome de Joubert	AR	TMEM231	614949	p.Tyr4*; Asp209Asn
Syndrome de Joubert	AR	C5orf42	614615	p.Arg1336Trp; p.Ala1564Thr; c.7400+1G>A; p.Pro2136Hisfs* 31; p.Arg1602*; p.Arg2493*

Adaptation de Laberge AM, et al  
mtDNA: ADN mitochondriale

## 1.6 Les deux projets dans le contexte des progrès technologiques

Ces deux projets ont en commun l'objectif d'identifier les causes génétiques des maladies neurologiques par l'étude de cohorte CF. Dans le cadre du projet sur les leucodystrophies, les résultats présentés ne sont pas une étude avec des patients CF mais découlent directement d'une étude précédente dans laquelle nous avons identifié un haplotype partagé et une mutation avec un effet fondateur micro-régional à la région Beauce-Bellechasse. Dans ce projet, nous avons pour hypothèse que les patients négatifs pour des mutations dans le gène *POLR3A* possédaient des mutations dans une autre sous-unité de POL III et qu'ensemble représentait une importante cause de leucodystrophie hypomyélinisante. Comme ce projet a débuté avant l'accessibilité des nouvelles technologies, nous avons utilisé une approche traditionnelle gène candidat. Dans le cadre du projet polyneuropathie douloureuse tardive,

l'étude était concentrée sur une grande famille CF originaire de la région du Saguenay Lac-St-Jean (SLSJ). L'enrichissement de variants rares dans la population CF et plus particulièrement au SLSJ était un paramètre à tenir compte lors de la recherche de la mutation. Pour un grand nombre de patients atteint d'une neuropathie douloureuse la cause est à ce jour inconnue. Des facteurs génétiques pourraient être l'explication pour certains de ces patients. L'objectif de ce projet de recherche était d'identifier la cause génétique responsable de cette neuropathie douloureuse. Compte tenu la grande variabilité clinique et génétique associée aux neuropathies douloureuse et l'accessibilité de technologies de séquençage, nous avons opté pour une approche non-biaisé de séquençage exomique afin d'identifier le gène causal.

Bien qu'il s'agit de deux maladies neurodégénératives, il n'y a pas de lien directe entre les pathologies et elles consistent toute deux en groupe de maladies très hétérogène cliniquement et génétiquement. Pour cette raison, nous avons décidé de traiter des deux pathologies séparément dans leur chapitre respectif. Nous avons plutôt décidé de discuter des avantages et limites des deux approches utilisés dans le cadre de ces projets et de l'impact des nouvelles technologies de séquençage dans le diagnostic clinique.

# **Chapitre 2:** Leucodystrophies

Le terme leucodystrophie se réfère à une atteinte de la matière blanche au niveau du cerveau. Les leucodystrophies sont des maladies génétiques associées à une dégénérescence ou un développement anormal de la myéline dans le système nerveux central (SNC) et parfois dans les nerfs périphériques<sup>15</sup>. La majorité des leucodystrophies se présente durant l'enfance ou l'adolescence, mais certaines formes adultes ont été décrites. Le parcours de la maladie est progressif et mène habituellement à une mort prématurée. Les leucodystrophies peuvent être séparées en deux catégories, soit démyélinisantes ou hypomyélinisantes, selon les caractéristiques observées à la résonance magnétique (IRM)<sup>16</sup>. Les leucodystrophies démyélinisantes sont causées par un problème au niveau de la maintenance de la myéline tandis que les leucodystrophies hypomyélinisantes sont causées par un problème dans la formation de la myéline<sup>33, 38</sup>. On estime que 30 à 40% des leucodystrophies sont toujours sans diagnostic moléculaire. Combiné au fait qu'à ce jour les leucodystrophies sont presque toutes incurables, cela démontre l'importance d'identifier les gènes causals et de mieux comprendre les mécanismes physiopathologiques sous-jacents.

Plusieurs formes de leucodystrophies hypomyélinisantes ont été décrites (table IV). Parmi eux, la leucodystrophie Pelizaeus-Merzbacher, causé par des mutations dans le gène *PLP1* sur le chromosome X, est la forme prototypique<sup>33,38</sup>. Il s'agit d'une maladie hétérogène avec un sous-type néonatal très sévère ou une forme plus légère adulte. Les résultats qui seront discuté dans cette thèse portent sur les leucodystrophies causées par des mutations dans des sous-unités de l'enzyme Pol III.

**Table IV: Leucodystrophies hypomyélinisantes.**

Group	Disorder	Inheritance	Gene(s)	
<b>Hypomyelinating leukodystrophies</b>	PMD (Pelizaeus-Merzbacher disease)	X-linked	<i>PLP1</i>	
	PMD-like or HLD2, SPG44	AR	<i>GJC2 (GJA12)</i>	
	Hypomyelination with congenital cataracts (HCC) or HLD5	AR	<i>FAM126A (DRCTNNB1A, HCC)</i>	
	HABC (Hypomyelination with atrophy of basal ganglia and cerebellum) or HLD6	AR	<i>Unknown</i>	
	Poi III-related leukodystrophies (HLD 7 and HLD8) [4H syndrome, ADDH, TACH, LO, HCAHC]	AR	<i>POLR3A and POLR3B</i>	
	18q- syndrome	Sporadic	18q22.3-q23 ( <i>MBP</i> )	
	Free sialic acid storage disease	AR	<i>SLC17A5</i>	
	Oculodentodigital dysplasia	AD (occasionally AR)	<i>GJA1</i>	
	Cockayne syndrome and Trichotiodystrophy with hypersensitivity to sunlight	AR	<i>ERCC8, ERCC6 and ERCC3, GTF2H5, ERCC2 (respectively)</i>	
	Fucosidosis	AR	<i>FUCA1</i>	
	Galactosemia	AR	<i>GALT</i>	
	SOX10-associated disorders	AD	<i>SOX10</i>	
	Hypomyelination of Early Myelinating Structures (HEMS)	Presumed X-linked	<i>Unknown</i>	
	<b>Prominent cerebral atrophy and slowly progressing myelination / early-onset neuronal degenerative disorders</b>	Hypomyelinating leukodystrophy-3 (HLD3)	AR	<i>AIMP1</i>
		Mitochondrial hsp60 chaperonopathy (HLD4)	AR	<i>HSPD1</i>
Global Cerebral Hypomyelination		AR	<i>SLC25A12</i>	
Early-onset GM1 and GM2		AR	<i>GLB1, HEXA</i>	
Infantile NCL		AR	<i>PPT1</i>	
Serine synthesis defects		AR	<i>PHGDH, PSAT1, PSPH</i>	
<b>Delayed myelination</b>		Allen-Herdon-Dudley syndrome	X-linked	<i>SLC16A2 (MCT8)</i>
	Others			

Adaptation de Perlman SJ et al 2012

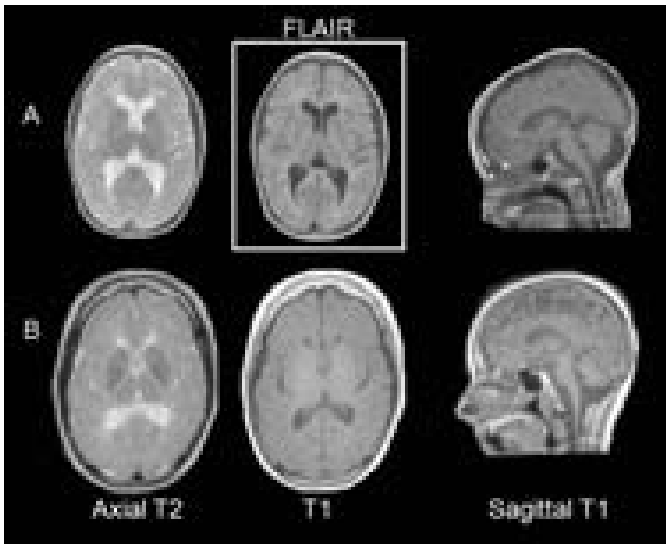
## 2.1 Résultats préliminaires: Description d'une nouvelle forme de leucodystrophie avec effet fondateur québécois et mutations dans le gène *POLR3A*.

L'étude génétique des leucodystrophies au laboratoire a débuté avec le recrutement de sept enfants, appartenant à cinq familles différentes, atteints d'une même forme de leucodystrophie. Parmi les cinq familles, deux d'entre elles étaient consanguines et originaires de la région de Beauce-Bellechasse<sup>17</sup>. Tous les patients présentaient un retard

développemental, de l'ataxie et du tremblement (table V). Avec la progression de la maladie, l'atteinte motrice se détériore et mène à l'utilisation d'un fauteuil roulant pour la majorité des cas. D'autres composantes cliniques sont aussi observées chez certains patients comme de la dysarthrie, dysphagie, crise partielle complexe, hypodontie et hypogonadisme hypogonadotrope (table V). Au niveau de l'IRM, tous les patients démontraient une hypomyélinisation ainsi qu'un amincissement du corps calleux (Figure 2).

**Table V:** Table clinique des patients atteints du syndrome TACH.

Patient #	Family #	Gender	Developmental delay	Age of onset (y)	Age in 2009	Ataxia	Dysmetria	Spasticity	Hyperreflexia	Babinski	Dysarthria	Dysphagia	Drooling	Tremor	Gaze-evoked nystagmus	Optic atrophy	Bladder dysfunction	Seizures	Wheelchair	Age - wheelchair (y)	Neuropathy	Cognitive regression (mild)	Hypodontia	Hypogonadotropic hypogonadism	Age of death
1	I	M	+	1	6	+	+	+	+	+	+	-	+	+	N/A	+	-	+	4	-	-	-	-	-	-
2	II	M	-	2,5	14	+	+	+	+	+	+	+	+	+	+	+	+	+	7	N/A	+	-	-	-	-
3	III	M	+	2	10	+	+	+	+	+	+	-	-	+	+	-	+	+	5	-	+	-	-	-	-
4	IV	M	+	1	14	+	+	+	+	+	+	+	+	+	+	+	+	+	8	-	+	+	-	-	-
5	IV	M	+	5	12	+	+	+	+	-	+	-	-	+	-	-	-	-	-	-	-	+	+	-	-
6	IV	F		3	N/A	+	+	+	+	+	+	+	+	+	-	+	+	-	+	8	-	+	-	-	21
7	V	M		3	28	+	+	+	+	+	+	+	+	+	+	+	+	-	+	12	-	+	-	+	
Mean				2,5																			Mean	7,3	



**Figure 2:** IRM de patients atteints du syndrome TACH. A) IRM du patient 7 (âgé de 11 ans, atteinte sévère) B) IRM du patient 7 (âgé de 21 ans, atteinte légère). Les IRM démontrent une hypomyélinisation et un amincissement du corps calleux.

Une analyse de cartographie par homozygotie a été effectuée en utilisant des données de génotypage SNP obtenues pour les deux familles consanguines<sup>17</sup>. Une région homozygote partagée de 12,6Mb sur le chromosome 10 (10q22.3-q23.31) a été identifiée. Au même moment, une troisième famille consanguine originaire de la région Beauce-Bellechasse a été recrutée. Cette famille partageait les symptômes cliniques et d'IRM avec les familles précédentes. Puisque cette troisième famille était aussi homozygote pour le même haplotype, les données de génotypage des trois familles ont été combinées afin de réduire l'intervalle candidat à 4,3Mb<sup>18</sup>. Une famille syrienne atteinte d'une leucodystrophie associée à une oligodontie (LO) a aussi été liée au chromosome 10q22<sup>19</sup>. Compte tenu du chevauchement clinique entre les patients CF et la famille syrienne, nous avons réduit la grandeur de l'intervalle candidat à 2,99Mb en combinant les deux loci.

Par une approche gène candidat, des mutations dans le gène codant pour la sous-unité A de l'enzyme polymérase III (*POLR3A*) ont été identifiées<sup>20</sup>. Des mutations ont été identifiées dans la cohorte de patients TACH ainsi que LO. Il a été noté que ces deux phénotypes (TACH et LO) partageaient des signes cliniques avec le syndrome



d'hypomyélinisation, hypodontie, hypogonadisme et hypogonadotrope (4H). Bien que le syndrome 4H a été décrit il y a plusieurs années, aucune cause génétique ne lui avait été associée<sup>21</sup>. Des patients 4H ont été recrutés et séquencés pour *POLR3A*. Des mutations ont aussi été identifiées dans ce groupe de patients. Au total, 14 mutations chez 19 patients ont été identifiées (Table VI)<sup>20</sup>. L'analyse protéique par immunobuvardage de type Western (WB) a démontré une diminution significative du niveau de POLR3A (23.4%) dans les fibroblastes de patients en comparaison avec des contrôles.

L'enzyme polymérase III (POL III) est constituée de 17 sous-unités. La sous-unité A constitue avec la sous-unité B le pore catalytique de l'enzyme. Elle est responsable de la transcription de plus de 200 petits ARN non-codant, tels que les ARNt, 5S, 7SK et U6<sup>22</sup>. Une perturbation dans le complexe enzymatique pourrait mener à une importante déficience dans la synthèse des protéines.

**Table VI:** Mutations identifiées dans le gène *POLR3A*

Individual Number	Family Number	Ethnicity	Diagnosis	Mutation(s)	
				DNA	Protein
1	I	FC	TACH	c.1674C>G	p.Phe558Leu
				c.3742insACC	p.1248insThr
2	II	FC	TACH	c.2015G>A	p.Gly672Glu
3	II	FC	TACH	c.2015G>A	p.Gly672Glu
4	II	FC	TACH	c.2015G>A	p.Gly672Glu
5	III	FC	TACH	c.2015G>A	p.Gly672Glu
6	IV	FC	TACH	c.2015G>A	p.Gly672Glu
7	V	W USA	4H	c.2554A>G	p.Met852Val
				c.2711-1G>A	p.Arg873AlafsX878
8	VI	AA USA	4H	c.2324A>T	p.Asn775Ile
				c.1114G>A	p.Asp372Asn
9	VII	W USA	4H	c.2830G>T	p.Glu944X
				c.3013C>T	p.Arg1005Cys
10	VIII	W USA	4H	c.2554A>G	p.Met852Val
				c.2711-1G>A	p.Arg873AlafsX878
11	IX	France	4H	c.4006C>T	p.Gln1336X
				c.1907C>A	p.Ser636Tyr
12	X	Syria	LO	c.2003p18G>A	p.Tyr637CysfsX650
13	X	Syria	LO	c.2003p18G>A	p.Tyr637CysfsX650
14	X	Syria	LO	c.2003p18G>A	p.Tyr637CysfsX650
15	X	Syria	LO	c.2003p18G>A	p.Tyr637CysfsX650
16	X	Syria	LO	c.2003p18G>A	p.Tyr637CysfsX650
17	X	Syria	LO	c.2003p18G>A	p.Tyr637CysfsX650
18	XI	W Europe	LO	c.418C>T	p.Arg140X
				c.2554A>G	p.Met852Val
19	XI	Guatemala	LO	c.2171G>A	p.Cys724Tyr

## **2.2 Recessive mutations in POLR3B, encoding the second largest subunit of POL III, cause a rare hypomyelinating leukodystrophy.**

L'approche traditionnelle combinant la cartographie par homozygotie avec une sélection de gènes candidats a mené à l'identification de mutations dans le gène *POLR3A* chez des patients atteints de leucodystrophie hypomyélinisante (TACH, LO et 4H). Certains patients ayant un phénotype correspondant au syndrome 4H se sont avérés être négatifs pour des mutations dans *POLR3A*. Un des objectifs de ce projet de thèse était d'identifier la cause génétique chez ces patients. Puisqu'il était impossible de distinguer sur une base clinique les patients 4H avec mutations dans *POLR3A* de ceux sans mutation, nous avons émis l'hypothèse qu'une autre sous-unité de POL III pourrait aussi causer une leucodystrophie hypomyélinisante. Comme la sous-unité B de l'enzyme (*POLR3B*) constitue le pore catalytique avec *POLR3A*, le gène *POLR3B* nous a semblé le meilleur candidat. Nous avons donc séquencé ce gène chez trois patients pour lesquels aucune mutation dans *POLR3A* n'avait été identifiée. Ceci a mené à une publication de type rapport dans le journal "American Journal of Human Genetics".

### **2.2.1 Contributions des auteurs**

1. Martine Tétréault\*: Participation à l'hypothèse de recherche. Analyse des séquences des exons et jonctions exon-intron de *POLR3B*. Analyse de la pathogénicité des variants avec des outils bio-informatiques. Rédaction du manuscrit.
2. Karine Choquet\*: Amplification PCR des fragments *POLR3B* et participation à l'analyse des séquences. Analyse de la conservation des nucléotides mutés dans les autres espèces. Participation aux figures du manuscrit et révision du manuscrit.
3. Simona Orcesi: Évaluation clinique des patients. Révision du manuscrit.
4. Davide Tonducci: Évaluation clinique des patients. Révision du manuscrit.
5. Umberto Balottin: Évaluation clinique des patients. Révision du manuscrit.

6. Martin Teichmann: Révision du manuscrit.
7. Sébastien Fribourg: Structure 3D de la protéine. Révision du manuscrit.
8. Raphael Schiffmann: Évaluation clinique des patients. Révision du manuscrit.
9. Bernard Brais: Participation à l'hypothèse de recherche. Évaluation clinique des patients. Révision du manuscrit.
10. Adeline Vanderver: Évaluation clinique des patients. Révision du manuscrit.
11. Geneviève Bernard: Participation à l'hypothèse de recherche. Évaluation clinique des patients. Participation à l'écriture du manuscrit.

\*Co-premier auteur du papier.

**2.2.2 Recessive mutations in *POLR3B*, encoding the second largest subunit of POL III, cause a rare hypomyelinating leukodystrophy. *American Journal of Human Genetics* 2011 Nov 11; 89(5): 652-5.**

#### **ABSTRACT**

Mutations in *POLR3A* encoding the largest subunit of RNA polymerase III (Pol III) were found to be responsible for the majority of cases presenting with three clinically overlapping hypomyelinating leukodystrophy phenotypes. We uncovered in three cases without *POLR3A* mutation, recessive mutations in *POLR3B* which codes for the second largest subunit of Pol III. Mutations in genes coding for Pol III subunits are a major cause of childhood-onset hypomyelinating leukodystrophies typically characterized by prominent cerebellar dysfunction, oligodontia and hypogonadotropichypogonadism.

#### **RESULTS AND DISCUSSION**

Leukodystrophies are a heterogeneous group of neurodegenerative disorders characterized by abnormal central nervous system white matter<sup>23</sup>. It is estimated that at least 30% to 40% of patients with leukodystrophies remain without a precise diagnosis, despite extensive investigations<sup>24</sup>. We recently demonstrated that the majority of cases

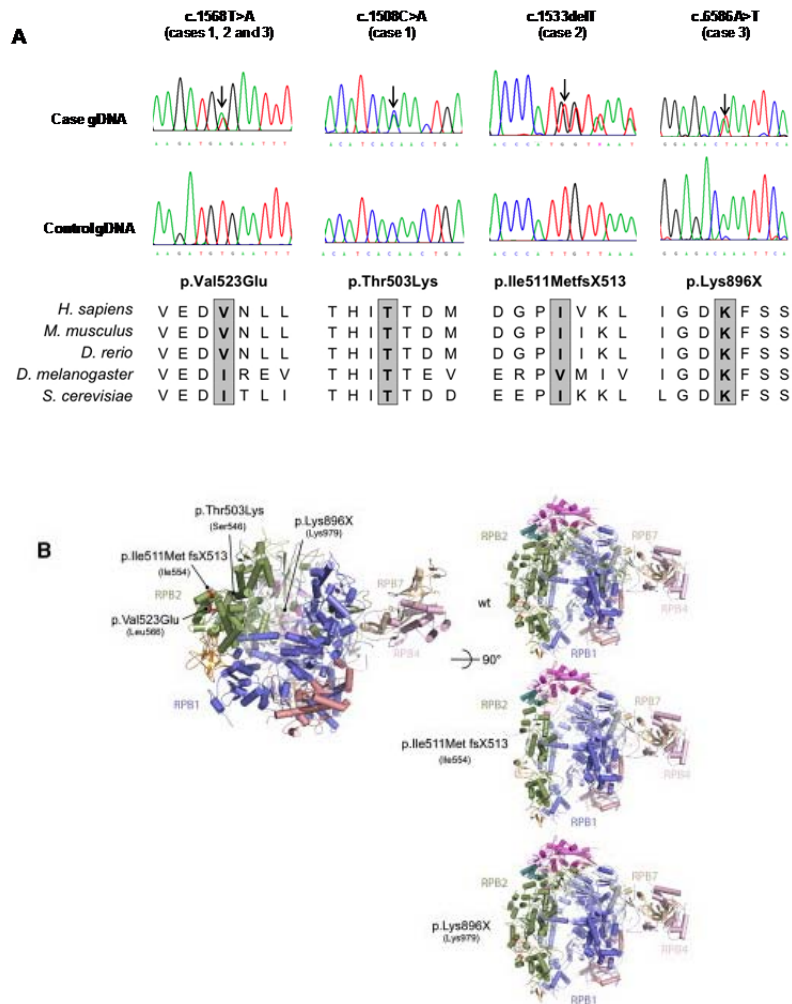
affected by Tremor-Ataxia with Central Hypomyelination (TACH)<sup>17; 18</sup>, Hypomyelination, Hypodontia and HypogonadotropicHypogonadism (4H syndrome, MIM 612440)<sup>25-28</sup> and Leukodystrophy with Oligodontia (LO, MIM 607694)<sup>19; 29</sup> are caused by mutations in *POLR3A* (MIM 610210)<sup>20</sup>. *POLR3A* is the largest of the 17 subunits that constitute RNA polymerase III (Pol III). It forms, together with the second largest subunit (*POLR3B*), the catalytic center of the enzyme. Pol III transcribes small untranslated RNAs (e.g. tRNAs, 5S RNA, 7SK RNA, U6 RNA) involved in the regulation of essential cellular processes, including transcription, RNA processing and translation<sup>22</sup>. We suggested that these allelic conditions be referred to as Pol III-related hypomyelinating leukodystrophies. Having not identified *POLR3A* mutations in four of nine 4H syndrome cases<sup>20</sup>, we searched for mutations in *POLR3B* which codes for the other subunit that forms Pol III's catalytic site.

The research project was approved by the institutional Ethics Committee of the Centre de Recherche du CHUM (CRCHUM), Montreal, Canada, the Institutional Review Board of the National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland, USA and the Children's National Medical Center, Washington D.C., USA. Informed consent was obtained from all participants. We sequenced all exons, exon-intron boundaries and 3' and 5' UTR of *POLR3B* ([NM\\_018082](#), hg19) on available genomic DNA from two of the four 4H cases not found previously to carry *POLR3A* mutations (cases 1-2)<sup>17</sup> and one new case of 4H syndrome (case 3).

All three cases were found to be compound heterozygote for mutations in *POLR3B* (Figure 3A). One missense mutation in exon 15, c.1568T>A (p.Val523Glu), is common to all three individuals from different ethnic background; two US residing persons of mixed European descent (cases 1 and 2) and one Italian individual (case 3). This common mutation could result from an ancestral haplotype shared by all three individuals. However, we do not have access to genotyping data on these individuals to validate this hypothesis. Case 1 is carrying another missense mutation also in exon 15 changing a threonine for a lysine at position 503 (c.1508C>A, p.Thr503Lys). The impact of these two missense mutations on

*POLR3B* was assessed using AlignGVGD, PhD-SNP, PolyPhen and SIFT. All bioinformatics programs predicted the mutations to be pathogenic. In addition to the common mutation, case 2 was carrying a deletion of one base pair (c.1533delT, p.Ile511MetfsX513) in exon 15 causing a frameshift and leading to a premature stop codon at position 513. The second mutation in case 3 is a nonsense variant in exon 11 (c.2686A>T) creating a stop codon at position 896 (p.Lys896X). As seen in *POLR3A* mutations, none of the participants is homozygous for null mutations<sup>17</sup>. All parents were unaffected and found to carry one of their affected child's mutations. All variants appear to be conserved across species (Figure 3A) and were absent in more than 340 control chromosomes except for the common mutation, c.1568T>A, which was found in 2/374 (0.5%) control chromosomes. All control individuals were of European descent. This variant has never been reported in SNPs database (dbSNP132).

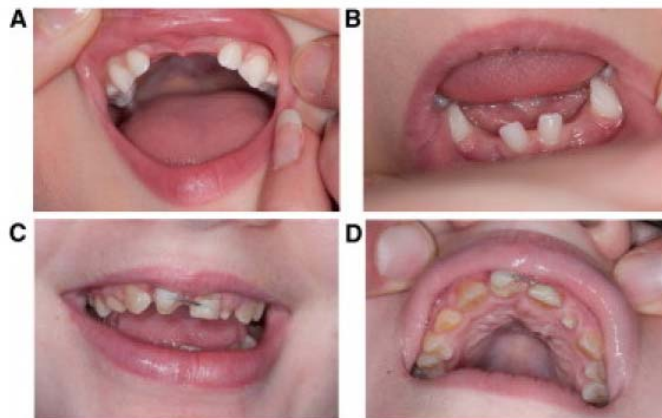
The locations of the different mutations are shown in Figure 3B. *POLR3B* (RPC2) and *POLR2B* (RPB2) subunits of RNA polymerase III and RNA polymerase II, respectively, are highly conserved proteins with a similar structure<sup>30</sup>. Extrapolating *POLR3B* (RPC2) from the yeast Pol II structure, the common mutation c.1568T>A (p.Val523Glu) corresponds to Leu566 in *POLR2B* (RPB2) which is, across species, always an aliphatic or aromatic residue. The mutation c.1508C>A (p.Thr503Lys) corresponds to S546 in RPB2, which is highly conserved across species. The null mutation c.2686A>T (p.Lys896X) corresponding to Lys979 in *POLR2B* (RPB2) and the deletion mutation leading to a frameshift and to a premature stop codon c.1533delT (p.Ile511MetfsX513) corresponding to Ile554 in *POLR2B* (RPB2) are predicted to generate, if the messenger RNA (mRNA) is stable enough, a truncated *POLR3B* subunit leading to a non-functional polymerase since the active site is expected to be largely affected in its function. From the Pol III electron microscopy structure and photocross-linking experiments<sup>31; 32</sup>, the p.Val523Glu and p.Thr503Lys residues are located near the "jaw", an area of Pol III where *POLR3D* (RPC4) and *POLR3E* (RPC5) are localized (counterpart of yeast RPC53 and RPC37). These two point mutations are predicted to affect locally the *POLR3B* structure and thus impair the proper function of RPC4 and RPC5 in transcription.



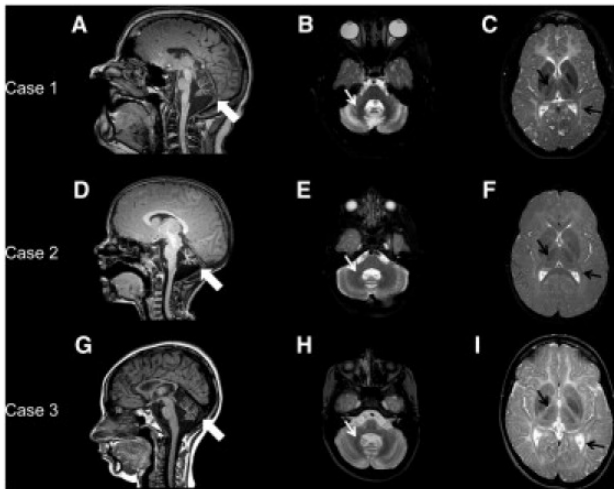
**Figure 3: POLR3B mutations. A)** Genomic sequence chromatograms and amino acid conservation across species is shown for all missense and nonsense *POLR3B* mutations. Case 1 is compound heterozygote for mutations c.1508C>A and c.1568T>A, case 2 is compound heterozygote for mutations c.1533delT and c.1568T>A, and case 3 is compound heterozygote for mutations c.1568T>A and c.2686A>T. The amino acid conservation across species is shown for all mutations. Abbreviations used: *D. melanogaster*: *Drosophila melanogaster*, *D. rerio*: *Daniorerio*, *H. sapiens*: *Homo sapiens*, *M. musculus*: *Musmusculus*, *S. cerevisiae*: *Saccharomyces cerevisiae*. **B)** Three dimensional representations of *POLR3B* missense, deletion (frameshift) and nonsense mutations. Point mutations identified in *POLR3B* are displayed according to their equivalent positions in the yeast RNA Polymerase II RPB2 subunit. Numbers in brackets refer to the positions in yeast RPB2. Images have been

generated with PyMol software (Schrödinger). RPB1 is shown in blue, RPB2 in green, RPB7 in brown, RPB4 in pink, RPB8 in orange, and RPB11 in magenta.

All three cases had characteristic clinical phenotypes of 4H syndrome<sup>26-28</sup>. They all presented in early childhood with developmental delay and developed dysarthria as well as progressive motor difficulties including cerebellar ataxia and in case of individuals 1 and 2, progressive spasticity. Individuals 1 and 3 both developed hypogonadotrophic hypogonadism while individual 2 was too young to evaluate for endocrine dysfunction. All cases had teeth abnormalities, which are presented in Figure 4 for cases 1 and 2 and consisted of neonatal upper incisors, delayed eruption of deciduous teeth and permanent teeth, abnormal sequence of eruption and malposition in case 3. The clinical features of case 3 have been published previously<sup>21</sup>. MRI features of the three patients were also characteristic of previous descriptions of 4H syndrome<sup>33</sup> (Figure 3).



**Figure 4:** Teeth abnormalities. **A)** and **B)** demonstrate the teeth abnormalities seen in case 1: unspecified age at tooth eruption, absence of permanent mandibular second premolars with abnormal shape of the permanent maxillary central and lateral incisors. **C)** and **D)** represent the teeth abnormalities of case 2: delayed tooth eruption with complete retention of the primary maxillary central incisors and the mandibular lateral incisors.



**Figure 5:** MRI characteristics. **A), D), G)** Sagittal T1-weighted images of the brain for case 1 (A), 2 (D) and 3 (G) at the level of the midline demonstrating the thin corpus callosum and vermian cerebellar atrophy (thick white arrows). **B), E), H)** Axial T2-weighted images of the brain for case 1 (B), 2 (E), and 3 (H) at the level of the pons and the dentate nucleus demonstrating a mild T2 hyperintensity of the cerebellar white matter associated with the typical hypointensity of the dentate nucleus (small white arrows) (Steenweg ME. et al). **C), F), I)** Axial T2-weighted images of the brain for case 1 (C), 2 (F), and 3 (I) at the level of the basal ganglia showing the mild hyperintense white matter characteristic of hypomyelination, with relative sparing of the optic radiations and anterolateral nucleus of the thalamus (small black arrows), which appear hypointense (Steenweg ME. et al).

Pol III-related hypomyelinating leukodystrophies are a genetically and clinically heterogeneous group of disorders. We demonstrated previously that *POLR3A* mutations account for the majority of our relatively small group of affected individuals and are reporting that *POLR3B* mutations account for the remaining cases for which DNA was available (2/4), as well as one other case of 4H syndrome (case 3). It is however likely that genes encoding for the other Pol III subunits are mutated in cases of hypomyelinating leukodystrophies. As previously suggested, since Pol III is responsible for the transcription of transfer RNAs (tRNAs) and other essential small RNAs, it is possible that mutations in *POLR3A* and *POLR3B* lead to an abnormal Pol III function which may affect the levels of certain (t)RNAs important for the development of the central nervous system white matter<sup>34</sup>, leading to abnormal protein production. This pathophysiological mechanism is also



attributed to Leukoencephalopathy with Brainstem and Spinal cord involvement and Lactate elevation (LBSL, MIM 611105) caused by mutations in the nuclear coded mitochondrial Aspartyl-tRNA synthetase (*DARS2*, MIM 610956) and hypomyelinating leukodystrophy-3 (MIM 260600) caused by mutation in Aminoacyl-tRNA synthetase complex-Interacting-Multifunctional Protein 1 (*AIMP1/p43*, MIM 603605)<sup>34; 35</sup>. We<sup>35; 36</sup> believe that genes coding for other Pol III subunits and possibly genes coding for proteins interacting with Pol III could be mutated in cases of genetically uncharacterized hypomyelinating leukodystrophies.

## **ACKNOWLEDGEMENTS**

We would like to thank all participants and the clinicians who referred patients to us. We would like to thank the “Fondationsur les Leucodystrophies” and the European Leukodystrophy Association (ELA) for financing this research project. Dr G Bernard received fellowships from the RMGA (Réseau de Médecine Génétique Appliquée) and FRSQ (Fonds de Recherche en Santé du Québec). Martine Tétréault received the Frederick Banting and Charles Best Doctoral scholarship from the CIHR (Canadian Institute of Health Research). Dr A Vanderver’s contribution is supported in part by the Intramural Research Program of the National Human Genome Research Institute and by the Myelin Disorders Bioregistry Project. Dr M Teichmann has received grants from the Agence Nationale de Recherche (ANR), the National Cancer Institute (INCa), the regional government of Aquitaine and the Ligue Contre le Cancer (équipe labellisée). Dr S Fribourg has received grants from La Ligue Contre le Cancer (Comité Dordogne), the regional government of Aquitaine and the Inserm. The authors wish to thank all collaborators (including Dr William McClintock) and patients participating in the Myelin Bioregistry Project.

**2.3 Mutations in POLR3A and POLR3B are a major cause of hypomyelinating leukodystrophies with or without dental abnormalities and/or hypogonadotropic hypogonadism.**

Avec l'identification de mutations dans les gènes *POLR3A* et *POLR3B*, une nouvelle forme de leucodystrophie hypomyélinisante associée à l'enzyme POL III a été décrite. Malgré une variabilité phénotypique entre les patients avec quatre phénotypes liés à POL III (TACH, LO, 4H et HCAHC), certaines caractéristiques cliniques se chevauchent entre les différents phénotypes<sup>20; 37</sup>. Nous avons donc émis l'hypothèse que la combinaison de certaines caractéristiques cliniques était spécifique à des mutations dans *POLR3A* ou *POLR3B*. Pour cette étude, nous avons sélectionné des patients selon des caractéristiques cliniques et radiologiques très précises et analysé par séquençage la présence de mutations dans l'un des gènes associés aux leucodystrophies liées à POL III. Cette étude a mené à une publication de type rapport dans le journal "Journal of Medical Genetics".

### **2.3.1 Contributions des co-auteurs**

1. Hussein Daoud\*: Participation à l'hypothèse de recherche. Participation à l'analyse des séquences ADN et cDNA. Analyse de la pathogénicité et la conservation des variants par des outils bio-informatiques. Participation à l'écriture du manuscrit.
2. Martine Tétreault\*: Participation à l'hypothèse de recherche. Participation à l'analyse des séquences ADN et cDNA. Analyse de la pathogénicité des variants par des outils bio-informatiques. Analyse d'haplotypes pour la mutation commune de *POLR3B*. Participation à l'écriture du manuscrit.
3. William Gibson: Évaluation des patients. Révision du manuscrit.
4. KetherGuerero: Aide technique.
5. Ana Cohen: Évaluation des patients. Révision du manuscrit.
6. JaninaGburek-Augustat: Évaluation des patients. Révision du manuscrit.
7. MatthisSynofzik: Évaluation des patients. Révision du manuscrit.
8. Bernard Brais: Évaluation des patients. Révision du manuscrit.
9. Cathy Stevens: Évaluation des patients. Révision du manuscrit.
10. Rocio Sanchez-Carpintero: Évaluation des patients. Révision du manuscrit.

11. Cyril Goizet: Évaluation des patients. Révision du manuscrit.
  12. Sakkubai Naidu: Évaluation des patients. Révision du manuscrit.
  13. Adeline Vandever: Évaluation des patients. Révision du manuscrit.
  14. Geneviève Bernard: Participation à l'hypothèse de recherche. Évaluation clinique des patients. Participation à l'écriture du manuscrit.
- \* Co-premier auteur du papier.

**2.3.2 Mutations in POLR3A and POLR3B are a major cause of hypomyelinating leukodystrophies with or without dental abnormalities and/or hypogonadotropic hypogonadism. *Journal of Medical Genetics* 2013 Mar; 50(3): 194-7.**

#### **ABSTRACT**

**Background:** Leukodystrophies are a heterogeneous group of inherited neurodegenerative disorders characterized by abnormal central nervous system white matter. Mutations in *POLR3A* and *POLR3B* genes were recently reported to cause four clinically overlapping hypomyelinating leukodystrophy phenotypes. Our aim was to investigate the presence and frequency of *POLR3A* and *POLR3B* mutations in patients with genetically unexplained hypomyelinating leukodystrophies with typical clinical and/or radiologic features of Pol III-related leukodystrophies. **Methods:** The entire coding region and the flanking exon/intron boundaries of *POLR3A* and/or *POLR3B* genes were amplified and sequenced in fourteen patients. **Results:** Recessive mutations in *POLR3A* or *POLR3B* were uncovered in all fourteen patients. Eight novel mutations were identified in *POLR3A*: six missenses, one nonsense, and one frameshift mutation. Seven patients carried compound heterozygous mutations in *POLR3B*, of whom six shared the common mutation in exon 15 (p.V523E). Seven novel mutations were identified in *POLR3B*: four missenses, two splice sites and one intronic mutation. **Conclusions:** To date, our group has described 37 patients, of whom 27 have mutations in *POLR3A* and 10 in *POLR3B*, respectively. Altogether, our results

further support that *POLR3A* and *POLR3B* mutations are a major cause of hypomyelinating leukodystrophies and suggest that *POLR3A* mutations are more frequent.

## INTRODUCTION

Leukodystrophies are a heterogeneous group of inherited neurodegenerative disorders characterized by abnormal central nervous system white matter, affecting mainly the synthesis and maintenance of cerebral myelin<sup>15; 38</sup>. It is estimated that despite extensive investigations, at least 30-40% of patients remain without a molecular diagnostic. Leukodystrophies are further classified into demyelinating and hypomyelinating according to their MRI characteristics, and several of them have now been genetically described<sup>23</sup>. Recently a group of four clinically overlapping hypomyelinating leukodystrophies (MIM 607694, MIM 614381) has been associated with mutations in polymerase III enzyme (Pol III) subunits<sup>20; 37; 39-41</sup>. Clinically, the patients affected by Pol-III related leukodystrophies present between early childhood to adolescence, have motor regression with upper motor neuron signs, cerebellar features and some milder degree of cognitive dysfunction and/or regression. In most cases dental abnormalities and/or hypogonadotropic hypogonadism are also present<sup>42</sup>. The MRI features show partial deficits in the deposition of myelin in the brain referred to as hypomyelination<sup>33</sup>.

Pol III is an enzyme responsible for the transcription of a group of more than 200 small untranslated RNAs (e.g. tRNAs, 5S, 7SK RNA and U6 RNA) involved in the regulation of essential cellular processes, such as transcription, RNA processing and translation<sup>22</sup>. The enzyme is composed of 17 subunits. *POLR3A* (MIM 614258) and *POLR3B* (MIM 614366) are the two largest and together form the enzyme's catalytic center. It has been suggested that mutations in these two subunits cause hypomyelinating leukodystrophy by disturbing the assembly of the Pol III complex or by leading to an abnormal enzymatic activity, thus affecting the level of small RNAs important for the development of white matter<sup>41</sup>.

In this study, we report a mutational analysis of a group of fourteen patients presenting clinical and imaging characteristics of Pol III-related leukodystrophies<sup>33</sup>. *POLR3A* and/or *POLR3B* were sequenced in all cases and mutations were uncovered in all fourteen patients. These results further support *POLR3A* and *POLR3B* as a major cause of hypomyelinating leukodystrophies.

## **METHODS**

### **Patients**

All patients included in this study were selected according to a combination of clinical findings and brain MRI features supporting the existence of a Pol III-related hypomyelinating leukodystrophies<sup>33</sup>. All patients and family members signed an informed consent from at least one of the following four institutions who had approved this research: the Centre de Recherche du CHUM (Montreal, Quebec, Canada), the Montreal Children's Hospital / McGill University Health Center Research Institute (Montreal, Quebec, Canada), the National Institute of Neurological Disorders and Stroke (NIH Bethesda, Maryland, USA) and the Children's National Medical Center (Washington D.C., USA).

### **Clinical and MRI selection criteria**

Patients were selected for sequencing if they had compatible clinical and/or radiological features of Pol III-related leukodystrophies<sup>41;43</sup>. More specifically, compatible clinical features were as follow: motor delay/regression, pyramidal and cerebellar features, dental abnormalities and hypogonadotropichypogonadism. The patients did not need to have all the features. These clinical features are summarized in Bernard and Vanderver 2012<sup>43</sup>. The table VII summarizes the clinical features of the patients. As for the MRI features, hypomyelination was a necessary criteria<sup>23</sup>, i.e. hyperintense signal of the white matter compared with grey matter structures in T2-weighted sequences and variable signal of the white matter (i.e. hypo-, iso- or hyperintense) compared to grey matter structures in T1 weighted images. Compatible MRI features included those reported by

Steenweg et al<sup>33</sup> and summarized in Bernard and Vanderver<sup>43</sup>: T2 hypointensities (relative preservation of the myelin) of the following structures: dentate nuclei, anterolateral nuclei of the thalami, globipallidi, opticradiations and/or corticospinal tracts at the level of the internal capsules, with or without the following: cerebellar atrophy, thin corpus callosum, white matter atrophy most prominent posteriorly. The figure 6 illustrates the MRI features of patient 3. Of note, MRI was not available for patient P1 as she passed away before MRI became available. Her brain pathology was however compatible with hypomyelination.

**Table VII:** This table summarizes the clinical and MRI features of the patients. Some of the patients are too young for the hypogonadotropic hypogonadism to manifest. The MRI features were reviewed by GB and AV.

Patients	Ethnicity	Compatible MRI	Onset	Dental abnormalities	Hypogonadotropic hypogonadism	Upper motor neuron signs	Cerebellar signs	Intellectual disability	Wheelchair	Age of death
<b>POLR3A</b>										
P1	French-Canadian	N/A but brain biopsy compatible with hypomyelination (before MRI era)	Early childhood	N/A	N/A	Yes	Yes	Yes	N/A	14 and 19 years
P2	Tunisian	Yes	6 years	Yes	No	No	Yes	Yes	N/A	N/A
P3	Tunisian	Yes	6 years	Yes	No	No	Yes	Yes	N/A	N/A
P4	Caucasian	Yes	14 months	Yes	No	Yes	Yes	Yes	Yes	N/A
P5	Caucasian	Yes	5 years	Yes	No	Yes	Yes	No	No	N/A
P6	Malaysian / Caucasian	Yes	4.5 years	Yes	No	No	No	No	No	N/A
P7	Spanish	Yes	5 years	Yes	No	No	No	Yes	No	N/A
<b>POLR3B</b>										
P8	N/A	Yes	15 years	Yes	Yes	Yes	Yes	N/A	No	N/A
P9	North European Caucasian	Yes	< 1 year	Yes	No	Yes	Yes	Yes	No	N/A
P10	German	Yes	< 1 year	Yes	Yes	No	Yes	Yes	No	N/A
P11	North European Caucasian	Yes	15 years	Yes	Yes	Yes	Yes	No	No	N/A
P12	Caucasian	Yes	Early childhood	Yes	No	N/A	Yes	N/A	Yes	N/A
P13	Caucasian	Yes	5 years	No	Yes	Yes	Yes	Yes	No	N/A
P14	North European Caucasian	Yes	< 1 year	Yes	No	Yes	Yes	Yes	No	N/A

MRI: Magnetic Resonance Imaging; N/A: not available.

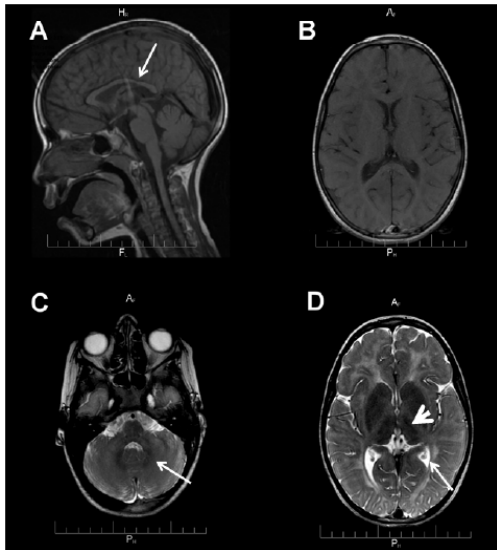
## DNA Screening

Genomic DNA was extracted from peripheral blood cells using standard methods. The entire coding sequence and intron-exon boundaries of *POLR3A* (NM\_007055.3) and/or *POLR3B* (NM\_018082) genes were amplified by polymerase chain reaction as

previously described<sup>20; 41</sup>. Oligonucleotide primers were designed using the exon primer software from the UCSC genome browser (<http://genome.ucsc.edu>) or the Primer 3 software (<http://frodo.wi.mit.edu/primer3/>). PCR products were sequenced at the Genome Quebec Innovation Center and McGill University (Montreal, Quebec, Canada) using an 3730XL Genetic Analyzer (Applied Biosystems, Foster City, California). All mutations were confirmed in independent reactions by sequencing both strands. SeqMan v4.03 (DNASTAR, Inc., Wisconsin, USA) and Mutation Surveyor v4.0.6 (SoftGenetics, Pennsylvania, USA) softwares were used for mutation detection analyses. Co-segregation analysis of the identified mutations was performed by sequencing the corresponding amplicons in available family members.

### **RNA Analysis**

Total RNA was extracted from lymphoblastoid cells derived from the patients with the predicted splice sites mutations. cDNA was generated from RNA using the SuperscriptIII enzyme (Invitrogen). Reverse transcription PCR (RT-PCR) was performed using *POLR3B*-specific primers and RT-PCR products were subsequently purified and sequenced.



**Figure 6:** Magnetic resonance imaging of patient P3. Selected MRI images of patient P3. The MRI demonstrate hypomyelination<sup>23</sup> i.e. slightly hyperintense white matter on T1-weighted images **(B)** and hyperintense white matter on T2-weighted images **(D)** compared to grey matter structures. On sagittal T1 image, at the level of the midline **(A)**, the corpus callosum is slightly thin (arrow). Of note, this patient does not have cerebellar atrophy. **(C)** On T2-weighted image at the level of the cerebellum, there is hypointensity of the dentate nuclei (arrow). **(D)** On T2-weighted images at the level of the basal ganglia, there are hypointensities of the optic radiations (arrow) and the anterolateral nuclei of the thalami (arrow head), but in this patient, not of the globipallidi and corticospinal tracts at the level of the internal capsules<sup>33</sup>

## RESULTS

The entire coding region and splice junctions of *POLR3A* and/or *POLR3B* genes were sequenced in fourteen patients with a hypomyelinating leukodystrophy suggestive of Pol III-related leukodystrophies according to typical clinical and MRI features described previously<sup>22; 33</sup>. All fourteen patients were found to carry mutations in one of the two genes: seven had mutations in *POLR3A* and seven in *POLR3B* (Table VIII). In *POLR3A*, ten mutations were identified, eight of which are novel (Table VIII). Three patients had homozygous missense mutation: p.Val166Ile was found in patient P1 and p.Trp671Arg in two siblings (P2 and P3). The remaining four patients were compound heterozygote. Among these, one patient (P4)



carried an insertion of one base pair (c.1741insA) in exon 13 leading to a frameshift mutation and a premature stop codon (p.Lys581Serfs\*28) in addition to a missense mutation p.Ser602Arg; one patient (P6) had an insertion of an adenine at position 1302 leading to a premature stop codon (p.Tyr434\*) and a missense mutation p.Met852Val; the two remaining patients carried two missense mutations each; P5 had p.Glu1261Lys and p.Ala387Gly mutations and P7 had p.Pro91Leu and p.Arg1005Cys mutations. As previously reported<sup>20</sup> none of these patients are homozygote for null mutations. Segregation analyses in families for which DNA was available in parents showed that the unaffected parents carry one heterozygous mutation each. None of these mutations were found in the Single Nucleotide Polymorphism database (dbSNP 135). Only the p.Glu1261Lys and p.Val166Ile missense mutations were found each in one individual out of 6503 from the NHLBI Exome Variant Server in the heterozygous state. Finally, all the missense mutations identified in *POLR3A* affect evolutionarily conserved amino acids (Figure 7A).

**Table VIII:** Mutations identified in patients with hypomyelinating leukodystrophies.

Patients	Nucleotide change	Amino acid change	Location
<i>POLR3A (NM_007055.3)</i>			
P1	c.[496G>A];[496G>A]*	p.[Val166Ile];[Val166Ile]*	Exon 5*
P2	c.[2011T>C];[2011T>C]	p.[Trp671Arg];[Trp671Arg]	Exon 15
P3	c.[2011T>C];[2011T>C]	p.[Trp671Arg];[Trp671Arg]	Exon 15
P4	c.[1741insA];[=]	p.[Lys581Serfs*28];[=]	Exon 13
	c.[1804A>C];[=]	p.[Ser602Arg];[=]	Exon 14
P5	c.[3781G>A];[=]	p.[Glu1261Lys];[=]	Exon 29
	c.[1160C>G];[=]	p.[Ala387Gly];[=]	Exon 8
P6	c.[2554A>G];[=]	p.[Met852Val];[=]	Exon 19
	c.[1302insA];[=]	p.[Tyr434*];[=]	Exon 10
P7	c.[272C>T];[=]	p.[Pro91Leu];[=]	Exon 3
	c.[3013C>T];[=]	p.[Arg1005Cys];[=]	Exon 23
<i>POLR3B (NM_018082)</i>			
P8	c.[1568T>A];[=]	p.[Val523Glu];[=]	Exon 15
	c.[2084-6A>G];[=]	p.[Gly695Valfs*5];[=]	Intron 19
P9	c.[1568T>A];[=]	p.[Val523Glu];[=]	Exon 15
	c.[2084-6A>G];[=]	p.[Gly695Valfs*5];[=]	Intron 19
P10	c.[1568T>A];[=]	p.[Val523Glu];[=]	Exon 15
	c.[1579T>C];[=]	p.[Cys527Arg];[=]	Exon 15
P11	c.[1568T>A];[=]	p.[Val523Glu];[=]	Exon 15
	c.[2817+30T>A];[=]	p.[?];[=]	Intron 24
P12	c.[1324C>T];[=]	p.[Arg442Cys];[=]	Exon 14
	c.[1568T>A];[=]	p.[Val523Glu];[=]	Exon 15
P13	c.[312G>T];[=]	p.[Leu104Phe];[=]	Exon 6
	c.[2570+1G>A];[=]	p.[Gly818Alafs*13];[=]	Intron 22
P14	c.[802A>G];[=]	p.[Ser268Gly];[=]	Exon 10
	c.[1568T>A];[=]	p.[Val523Glu];[=]	Exon 15

Novel mutations in bold.  
\*Presumed mutation, DNA not available as patient deceased; unaffected family members heterozygotes for the mutation.

A		p.P91L	p.V166I	p.A387G	p.S602R	p.M852V	p.R1005C	p.E1261K
H.sapiens	LELPCFH	NGTVKKC	VHVAKIL	QIFSVIL	FHTMAGR	QLDRITP	GGIEAAR	
M.musculus	LELPCFH	NGTVKKC	VHVAKIL	QIFSVIL	FHTMAGR	QLDRITP	GGIEAAR	
D.erio	LELPCFH	NGSVKKC	VHVAKIL	QIFSLIL	FHTMAGR	QLDRITP	GGIEAAR	
C.elegans	LDVPVFH	NGTVKKA	IHVAMTL	QLIELLI	FHTMGR	ANGCLTK	GGIEAAR	
D.melanogaster	LALPVFH	NGSVKKG	VRVAKIL	QMFSLIM	FHTMAGR	QIECLIT	GGIEAAR	
S.cerevisiae	LALPVFH	NGVVKKA	DRVAKVL	QVFSLLI	FHAISGR	QLYRISE	GGIEAAR	

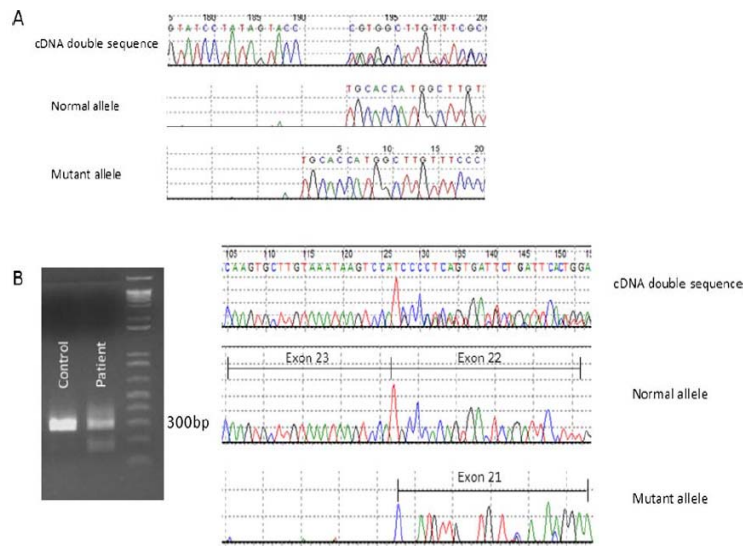
  

B		p.L104F	p.S268G	p.R442C	p.V523E	p.C527R
H.sapiens	ECRLRDM	FGPSLEE	VLCRLGY	VEDVNLLCGEE		
M.musculus	ECRLRDM	FGPSLEE	VLSRLSY	VEDVNLLCGEE		
D.erio	ECRLRDM	FAPSLIE	VLSRLSF	VEDVNLLCGEE		
C.elegans	ECRLRDM	FAPSLIE	VLCRLGY	VEDLNINVIHPCII		
D.melanogaster	ECRLRDT	FGASLID	VLSRLSY	VEDIREVSGNP		
S.cerevisiae	ECRLRDM	FAVNLEE	VLSRLSY	VEDITLIDSAS		

**Figure 7:** Multi-species alignment of *POLR3A* (A) and *POLR3B* (B) proteins showing the conservation of mutated amino acids across species. Sequence alignment was done using the Clustal W server (<http://www.ch.embnet.org/software/ClustalW.html>). The mutated residues are shown in red. Orthologous sequences include: Homo sapiens, Mus musculus, Danio rerio, Caenorhabditis elegans, Drosophila melanogaster, and Saccharomyces cerevisiae.

Eight mutations were identified in *POLR3B*, including seven that had not been previously described (Table IX). All patients were found to be compound heterozygotes. Six patients carried the common missense mutation p.Val523Glu in exon 15, in addition to another mutation (Table IX). Haplotype of the carrier chromosome for this shared mutation were extracted from sequencing results. Five surrounding SNPs in these six patients as well as three previously-reported patients carrying this common missense mutation<sup>41</sup> shows that all carriers share a common haplotype, suggesting that the p.Val523Glu mutation derives from a single ancestral mutation (Table IX). This haplotype was only observed in individuals carrying the p.Val523Glu mutation. In addition to the common mutation, two unrelated patients P8 and P9 carried a substitution of an adenine to a guanine (c.2084-6A>G) in intron 19, six base pairs before exon 20. To determine the impact of this mutation on splicing, we performed RT-PCR targeting exons flanking exon 20. We demonstrated that the c.2084-6A>G mutation caused the creation a cryptic acceptor splice site leading to a frameshift and a premature stop codon (p.Gly695Valfs\*5) five amino acids downstream (Figure 8A). One

patient (P13) had a missense mutation p.Leu104Phe in exon 6 in addition to a mutation affecting the consensus donor splice site of intron 22 (c.2570+1G>A). Similarly, RT-PCR analysis using total RNA extracted from this patient showed that this mutation caused the skipping of exon 22 leading to a frameshift and a premature truncated protein (p.Gly818Alafs\*13) (Figure 8B). One patient (P11) carried the common mutation in exon 15 in addition to a one base pair substitution (c.2817+30T>A) in intron 24, predicted to create a cryptic acceptor splice site in intron 24 by the BDGP Splice Site Prediction by Neural Network ([http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)). The remaining four patients carried the common missense mutation p.Val523Glu in addition to another missense mutation each, as p.Cys527Arg was found in P10, p.Arg442Cys in P12 and p.Ser268Gly in P14. As for *POLR3A*, none of these mutations were found in the Single Nucleotide Polymorphism database (dbSNP 135), except the common mutation (p.Val523Glu) which was found in 2/374 (0.5%) of control chromosomes and the intronic variant (c.2817+30T>A) which was observed in the 1000 genome project with a minor allele frequency of 0.001. Finally, all missense mutations appear to be located at regions conserved across species (Figure 7B).



**Figure 8: (A)** Sequence traces showing the effect of c.2084-6A>G mutation in *POLR3B*. The mutant allele subtracted from the double sequence by Mutation Surveyor softwareshow a 5bp shift due to the retention of 5 nucleotides from the intron. **(B)** Electrophoresis and sequence traces showing the effect of c.2570+1G>A mutation in *POLR3B*. Gel electrophoresis shows two separate bands in the patient. The size of the lower band corresponds to the deletion of exon 22. The mutant allele subtracted from the double sequence by Mutation Surveyor shows the junction between exon 21 and 23 on one allele

**Table IX:** Genotypes of SNPs in patients with the p.V523E mutation in *POLR3B*.

Cases	rs ID					
	rs7134881	rs7135054	rs138249161	rs34234806	rs17038460	rs13561
	Intron 4	Intron 5	Exon 15 (p.V523E)	Intron 19	Exon 20	Exon 28
	106763013	106763268	106826199	106838456	106848414	106903321
P8	C	A	A	T	G	T
	C	A	T	T	A	C
P9	C	A	A	T	G	T
	C	A	T	T	A	T
P10	C	A	A	T	G	T
	C	A	T	T	A	C
P11	C	A	A	T	G	T
	C	A	T	T	A	T
P12	C	A	A	T	G	T
	C	A	T	T	A	C
P14	C	A	A	T	G	T
	C	A	T	T	A	C
Tétreault et al, 2011	C	A	A	T	G	T
	C	A	T	C	A	T
	C	A	A	T	G	T
	T	G	T	T	A	T
	C	A	A	T	G	T
	T	G	T	C	A	T
Shared haplotype						

## DISCUSSION

In this study, we undertook a mutational analysis of *POLR3A* and *POLR3B* genes to evaluate the type and frequency of mutations in a cohort of fourteen patients with hypomyelinating leukodystrophies carefully selected according to a combination of diagnostic clinical and brain MRI features as described by Steenweg *et al* in 2010<sup>33</sup> and reviewed by Bernard and Vanderver in 2012<sup>43</sup>, respectively. Recessive mutations were uncovered in all fourteen patients. Six unrelated patients carried ten mutations in *POLR3A*, eight of which were novel. These mutations included eight missenses, one nonsense, and one frameshift mutation (Table VIII) and they are distributed throughout the gene and not clustered in a specific region or domain of the encoded protein. In *POLR3B*, seven novel mutations as well as one previously described mutation were identified in seven patients. Among these, six patients carried the common missense mutation p.Val523Glu in exon 15. While these patients are of different ethnical backgrounds, we show for the first time that this common mutation results from an ancestral haplotype shared by these six individuals as well as three additional patients previously reported to carry this mutation (Table IX)<sup>41</sup>. The seven novel mutations in *POLR3B* included four missenses, two splice site mutations, and one intronic mutation (Table VIII). The splice site mutations cause a frameshift leading to predicted truncated proteins of different sizes. To the best of our knowledge, this is the first description of *POLR3B* splicing mutations leading to premature stop codons. On the other hand, the intronic mutation in patient P11 is located 30 base pairs after the intron 24 donor splice site. Although quite far from the consensus canonical splice sites, this mutation was predicted by bioinformatic analysis to create a cryptic acceptor splice site in intron 24 that would probably lead to a frameshift in the encoded protein. Unfortunately, we could not verify this hypothesis as cell lines from this patient were not available.

Here, we show that mutations in *POLR3A* and *POLR3B* genes may occur at any given coding exon or splicing junctions of these genes highlighting the high degree of mutational heterogeneity of Pol III-related hypomyelinating leukodystrophies. Yet, *POLR3A* sequencing

should be prioritized for genetic testing since 73% (27/37) of all patients in our cohort carry mutations in *POLR3A* and 27% (10/37) of patients have mutations in *POLR3B*<sup>20; 39; 41</sup>. On the other hand, when mutations in *POLR3A* are ruled out, *POLR3B* should be sequenced. The exon 15 of *POLR3B* appears to be a “hotspot” for mutations as four different mutations (the common missense mutation as well as three other mutations) were identified in this specific exon, which correspond to a third of all the mutations uncovered by our group so far in this gene. Consequently, the exon 15 of *POLR3B* could be prioritized for sequencing in patients for whom no mutation is identified in *POLR3A*.

Taking together, mutations in either *POLR3A* or *POLR3B* genes were uncovered in fourteen patients with a hypomyelinating leukodystrophy phenotype in this study. Combining these numbers with our previous reports<sup>20; 41</sup> brings the number of patients with mutations in these two genes to 37, 27 in *POLR3A* and 10 in *POLR3B*, respectively. This very high yield of mutation detection is likely due to our stringent selection criteria but is strongly suggesting that mutations in these two genes represent a major cause of hypomyelinating leukodystrophies with or without dental abnormalities and/or hypogonadotropic hypogonadism.

## **ACKNOWLEDGEMENTS**

We wish to thank all family members for their participation. We would also like to thank for their services the Genome Quebec Innovation Center and McGill University. Dr. Bernard wishes to thank the Montreal Children’s Hospital Foundation, the MSSA (Medical Staff Service Association), the Montreal Children’s Hospital Associates in Neurology. The project was funded by Fondation sur les Leucodystrophies, the Montreal Children’s Hospital and McGill University Health Center Research Institutes and the Myelin Disorders Bioregistry project with special thanks to Johanna L Schmidt. HD received a postdoctoral fellowship from the Canadian Institutes of Health Research (CIHR). MT received the Frederick Banting and Charles Best Doctoral scholarship from the CIHR. GB received a ResearchScholar Junior 1 of the Fonds de recherche en santé du Québec (FRQS).

## **2.4 Les leucodystrophies liées à l'enzyme POL III représentent un groupe important des leucodystrophies hypomyélinisantes.**

Dans ce projet, nous avons identifié, par une approche gène candidat, des mutations dans le gène *POLR3B* chez des patients atteints du syndrome 4H. Dans l'analyse mutationnelle effectuée sur 14 patients, nous avons démontré la présence d'une entité clinique propre aux leucodystrophies liées à POL III. Ces deux études combinées à l'étude préliminaire décrivant les premières mutations dans le gène *POLR3A*, ainsi que les études démontrant la présence de mutations dans les gènes *POLR3A* et *POLR3B* chez des patients japonais<sup>37; 40</sup>, démontrent que les leucodystrophies liées à l'enzyme POL III sont un groupe important de leucodystrophies hypomyélinisante, même probablement la plus commune de toutes les formes.

Des études fonctionnelles sont nécessaires afin de mieux comprendre comment des mutations dans les sous-unités de l'enzyme POL III mènent à ces maladies neurologiques. Il serait intéressant d'évaluer l'impact des mutations sur l'activité transcriptionnelle de l'enzyme. Il est possible d'isoler le complexe POL III du noyau des cellules et ensuite, par un essai luciférase, étudier la capacité du complexe muté de transcrire le gène adénoviral VA1<sup>44</sup>. Certaines mutations se retrouvent dans le pore catalytique de l'enzyme et peuvent donc affecter directement l'efficacité enzymatique. D'autres mutations se retrouvent dans les autres domaines et pourraient plutôt avoir un effet indirect, par un assemblage anormal du complexe. Une immunoprécipitation du complexe POL III préalablement isolé suivit d'un marquage à l'argent et d'une analyse de spectrométrie de masse permettraient de vérifier la composition du complexe chez les patients<sup>45</sup>. Comme l'enzyme POL III est essentiel à la viabilité, il est possible que d'autres sous-unités viennent compenser pour le manque de *POLR3A* ou *POLR3B*. De plus, une perturbation dans l'efficacité enzymatique de POL III doit nécessairement avoir un impact sur la synthèse protéique. Il serait intéressant d'étudier cet

impact sur les petits ARN directement transcrits par POL III ainsi que sur l'ensemble des protéines puisque les petits ARN, comme les ARNt, sont essentiels à la synthèse protéique.



## **Chapitre 3:**

Polynévrite et ataxie sensitive tardive  
causé par une mutation dominante dans  
le gène *NAGLU*

Les neuropathies douloureuses sont causées par une dégénérescence des nerfs sensitifs. Les nerfs sensitifs sont l'extension des neurones des ganglions spinaux (DRG) et sont divisés en deux catégories: de grand diamètre et de petit diamètre<sup>46</sup>. Les nerfs sensitifs de grand diamètre ont pour terminaison les muscles, les jointures ainsi que les tendons et sont principalement responsable des sensations de proprioception, faiblesse musculaire, sensation de vibration, hypoflexie ou areflexie<sup>46; 47</sup>. Les nerfs sensitifs de petit diamètre ont pour terminaison la peau ainsi que les organes viscéraux et ont pour rôle la transmission des stimuli thermonociceptifs. Les neuropathies douloureuses peuvent être causées par une atteinte des grandes fibres, des petites fibres et même par une combinaison d'une atteinte des deux types de fibres<sup>47</sup>. Le diagnostic des neuropathies affectant les petites fibres (NPF) est un défi pour les cliniciens puisque les dommages causés aux fibres nerveuses non-myélinisées ou finement myélinisées ne sont pas détectable par une étude de conduction des nerfs. En fait, les limitations de la technique permettent seulement de déceler une atteinte des grands axones myélinisés ou des feuilles de myélines<sup>48</sup>.

La douleur ressentie par les patients atteints d'une NPF se présente de diverse manière, principalement par une sensation de brûlement, d'élançement, de picotement ou de démangeaison<sup>47</sup>. La présence d'une intolérance au toucher des draps ou des chaussettes ou un syndrome de jambes sans repos peuvent aggraver la sensation de douleur au niveau des pieds. Des crampes et des picotements peuvent être observés suivant une dysfonction des nocicepteurs musculaires. Des symptômes négatifs, tel qu'une perte de sensation thermal, perte de sensation des piqûres et des engourdissements, sont aussi observés. Les symptômes se présentent habituellement de manière symétrique débutant au niveau des pieds et se propageant au niveau proximale.

Les NPF peuvent être acquise ou transmise héréditairement (table X)<sup>49</sup>. Les neuropathies causées par le diabète, par des syndromes métaboliques ou immunitaires semblent être plus commune<sup>46; 48</sup>. Les neuropathies héréditaires représentent tout de même

un groupe important des neuropathies sans diagnostic. Les formes prédominantes sont les neuropathies sensitive et autonome (NHTA), les polyneuropathies amyloïdes familiale (PAF), certaines maladies lysosomales et les neuropathies associées à des syndromes mitochondriaux<sup>49; 50</sup>. Il y a trois types de NHTA pour lesquelles les petites fibres sont majoritairement atteintes; type I, IV et V. Le type le plus fréquent est NHTA I, qui est constitué de quatre sous types (A-D) tous transmis selon un mode dominant. Un gène causal a été identifié pour seulement deux sous-types A (*SPTLC1*) et C (*RAB7A*). NHTA IV est une forme congénitale à transmission récessive causée par des mutations dans le gène *NTRK1*. Ce sont les fibres non-myélinisées qui sont affectées dans NHTA I et IV. NHTA V est similaire au type IV mais se distingue par une absence sélective des petites fibres myélinisées. Ce type est causé par des mutations dans les gènes *NTRK1* et *NGFβ*<sup>51-54</sup>. Les PAF sont causées par une atteinte des nerfs sensitifs de grand et petit diamètre. La forme la plus commune est causée par des mutations dans le gène *TTR*. À ce jour, deux maladies lysosomales ont été associées avec une atteinte des petites fibres: la maladie de Fabry et la gangliosidose GM2<sup>50</sup>. La maladie de Fabry est une maladie de transmission récessive liée à l'X et due à un déficit de l'enzyme  $\alpha$ -galactosidase (*GLA*). On observe chez les patients une accumulation pathogène de céramide trihexosides. La gangliosidose GM2 est due à un déficit de l'enzyme hexosaminidase constituée de deux sous-unités A (*HEXA*) et B (*HEXB*). On observe chez les patients une accumulation de GM2 dans les tissus. La neuropathie ataxie sensitive avec dysarthrie et ophthalmoplégie (NASDO) est un syndrome mitochondriale causé par des mutations dans *POLG*. Les phénotypes associés à des mutations dans *POLG* sont très variable.

**Table X:** Neuropathies affectant les petites fibres acquises ou de causes génétiques.

	Cause	Length-dependent	Non-length-dependent	References
Metabolic	Diabetes	+		[5,6,9,17-19]
	Impaired glucose tolerance	+		[5,6,9,17-19]
	Rapid glycaemia control in diabetes	+		[20]
	Hypothyroidism	+	+	[6,21-23]
	Vitamin B <sub>12</sub> deficiency		+	[22]
Infectious disease	HIV	+		[24,25]
	Influenza		+	[26]
	Hepatitis C	+	+	[27]
Drugs and toxics	Antiretroviral drugs	+		[25]
	Metronidazole	+		[28]
	Bortezomib	+		[29,30]
	Statin	+		[31]
	Nitrofurantoin		+	[32]
	Flecainide	+		[33]
	Linezolid	+		[34,35]
	Chronic alcohol abuse	+		[36,37]
Immune-mediated	Celiac disease		+	[38]
	Sarcoidosis	+		[15,39]
	Sjögren's syndrome	+	+	[6,11,40-42]
	Systemic lupus erythematosus	+		[43,44]
	Rheumatoid arthritis		+	[10,42]
	Inflammatory bowel diseases	+	+	[10,45]
	Paraneoplastic syndrome	+	+	[46**]
	Monoclonal gammopathy	+	+	[6,11,47-49]
	Complex regional pain syndrome type 1		+	[50]
Genetic	HSAN type IV		+	[51]
	Fabry's disease	+	+	[52]
	Familial amyloidosis	+		[53**]
	Hemochromatosis	+		[54]
	Familial burning feet syndrome	+		[55,56]
	Sodium channel mutations [SCN9A]	+		[16**,57*]
Idiopathic	Unknown	+	+	[6,10]

HSAN, hereditary sensory neuropathy.

Lauria G. Curr Opin Neurol 2012

Comme le démontre le syndrome NASDO, les ataxies sensibles font aussi parties des maladies avec un désordre sensible affectant le système périphérique ou central<sup>55</sup>. En présence d'une neuropathie périphérique, on note de la faiblesse, des changements autonomiques et sensitifs. Au niveau des changements sensitifs, les patients démontrent des symptômes "négatifs", tel qu'une perte de sensation, et des symptômes "positifs", comme de la douleur.

La majorité des formes de neuropathie douloureuse ont été associée avec le diabète de type 2. Malgré tout, les neuropathies douloureuses causées par un défaut génétique représentent un groupe important et un grand nombre de patients sont toujours sans diagnostic moléculaire. Il est donc important de continuer la caractérisation génétique de ce groupe de maladie.

### **3.1 Adult onset painful sensory polyneuropathy caused by a dominant *NAGLU* mutation. Article en préparation, à soumettre sous forme de rapport dans *American Journal of Human Genetics*.**

Dans le cadre de ce projet, nous avons recruté une grande famille canadienne-française (FC) originaire du Saguenay-Lac-St-Jean (SLSJ). Les patients sont atteints d'une forme de polyneuropathie sensitive douloureuse associée à une ataxie avec début encore plus tardif. Compte tenu de la grande hétérogénéité clinique et génétique observée dans le groupe des polyneuropathies, nous avons opté pour une approche de séquençage NGS afin d'identifier le gène responsable.

#### **3.1.1 Contribution des co-auteurs**

1. Martine Tétreault: Participation à l'hypothèse de recherche. Analyse des données de séquençage exomique. Analyse de la pathogénicité des variants par des outils bioinformatiques. Validation des variants et analyse de ségrégation. Participation à la structure 3D de la protéine. Écriture du manuscrit.
2. Marie-Josée Dicaire: Aide technique.
3. Pierre Allard: Essai enzymatique. Révision du manuscrit.
4. Diane Leblanc: Essai enzymatique. Révision du manuscrit.
5. KalleGehring: Structure 3D de la protéine. Révision du manuscrit.
6. Nadine Leclerc: Recrutement des patients.
7. Jean Mathieu: Évaluation clinique des patients. Révision du manuscrit.

8. Bernard Brais: Participation à l'hypothèse de recherche. Participation à l'interprétation des résultats. Évaluation clinique des patients. Participation à l'écriture du manuscrit.

### **3.1.2 Adult onset painful sensory polyneuropathy caused by a dominant *NAGLU* mutation**

#### **ABSTRACT**

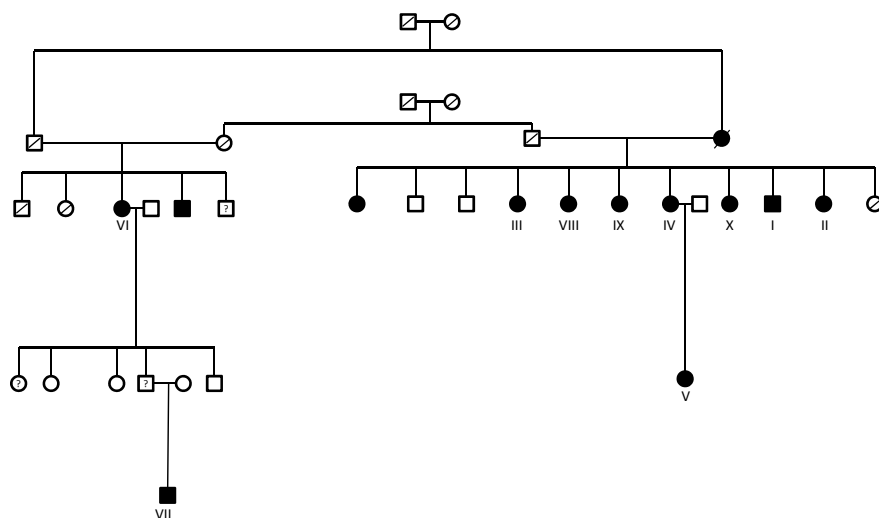
Late-onset painful sensory neuropathies are usually considered to be acquired and most commonly associated with common diseases such as diabetes. Adult presentations of known hereditary forms such as Amyloidosis and Fabry's are accompanied by other organ involvement. We performed exome sequencing in a large French-Canadian family with a dominantly inherited late onset painful sensory neuropathy to uncover the underlying genetic mutation. The main clinical features of this dominant painful sensory polyneuropathy are the appearance of excessive cramps as early as 20 years old, constant painful paresthesias in the feet and later the hands appearing on average around age 55 (30-65) that interfere with sleep and progresses into a mild sensory ataxia. Electrophysiological studies are normal until late in the course of the disease where it documents a sensory polyneuropathy. By exome sequencing we uncovered a heterozygous missense mutation p.I403T in the  $\alpha$ -N-acetyl-glucosaminidase (*NAGLU*) gene segregating in the family with the disease. Recessive *NAGLU* mutations cause the severe childhood lysosomal disease mucopolysaccharidosis IIIB (MPS-IIIB). Family members carrying the mutation showed a decrease of 10-43% of the enzymatic function. The late onset and variable severity of the symptoms may have precluded the description of such symptoms in parents of mucopolysaccharidosis IIIB cases since they usually are lost to follow-up. The identification of a dominant phenotype associated with a *NAGLU* mutation supports that some carriers of lysosomal enzyme mutations may develop later in life much milder phenotypes.

## RESULTS AND DISCUSSION

Late-onset sensory polyneuropathies, and in particular painful neuropathies, are a heterogeneous group of disorders often accompanying other illnesses such as Diabetes and are rarely found to have a genetic basis<sup>47</sup>. Most dominant and recessive sensory hereditary neuropathies have an early onset, are associated with autonomic dysfunction and classically are painless (HSAN)<sup>56</sup>. Sensory ataxias are another group of sensory disorders where a peripheral or central sensory dysfunction will lead or contribute to an ataxic gait<sup>55</sup>. For most of these sensory ataxias other neurological systems and organs are involved, as in Friedreich ataxia, and pain is not a constant or even a frequent manifestation. The genetic work up of late forms of painful adult onset sensory neuropathy is still limited. The advent of next-generation sequencing (NGS) technologies such as whole exome sequencing (WES) has opened the way to a rapid and cost-effective search for mutations for Mendelian disorders<sup>1</sup>. In the case of consanguineous or very large families, exome sequencing has been particularly rewarding. The identification of a large family with many members affected by a dominant late-onset painful neuropathy prompted the use of exome sequencing to identify the mutated gene.

We recruited a large French-Canadian (FC) family from the Saguenay region of Québec (15 individuals; 12 affected cases) presenting a late-onset sensory painful neuropathy (Figure 9). The pedigree is suggestive of a dominant mode of transmission with intra-familial variability in age of onset and severity. All participants signed an informed consent approved by the ethics board of the Centre de Recherche du CHUM (CRCHUM). The main clinical features are summarized in table XI. The first symptoms are lower extremity cramps that start between the age of 30 and 65 (mean 55) that evolve over a decade or more into continuous painful feet and then distal upper extremity paresthesias that are the usual symptoms that lead to medical evaluation. Poor sleep related to cramps and paresthesias increases with age. Patients have no dysautonomic symptoms and have stable blood pressure. On examination weakness is never observed in the oldest cases. Deep tendon

reflexes are preserved initially but are lost only distally with age. The most dramatic sensory finding on examination is decreased vibration sense that evolves into a complete loss at the level of the toes. Position sense is largely preserved. There is a striking glove decreased sensation to touch and pinprick. There is no nystagmus or cerebellar signs on examination. Tandem walk is impossible in the oldest cases and difficult in some of the younger cases. EMGs demonstrated a pure axonal sensory neuropathy only in the two eldest cases, while only normal values were observed in younger participants.



**Figure 9:** Pedigree of a large FC family affected by a late-onset painful sensory neuropathy.

**Table XI:** Clinical features of sensory painful neuropathie.

Clinical features/patients' ID	3915	3917	3994	3995	4827	4519	4710	3996	3993	3992
	I	II	III	IV	V	VI	VII	VIII	IX	X
Status	Affected	Affected	Affected	Affected	Affected	Affected	Possible	Possible	Affected	Affected
Sex	Male	Female	Female	Female	Female	Female	Male	Female	Female	Female
Year of birth	1947	1948	1939	1944	1972	1935	1986	1941	1942	1946
Age exam	61	60	74	67	41	75	24	71	70	67
Onset first symptoms (years)	52	48	65	60	30	60	20		61	64
Sleep disturbances	yes	yes	yes	yes	yes	yes	yes	no	yes	yes
Sensory ataxia	yes	yes	yes	yes	no	yes	no	no	yes	yes
Pain	yes	yes		yes	yes	yes	yes	no	yes	yes
Loss of DTR	no	no	yes	yes	no	no	no	yes	yes	yes
Decreased vibration sense	yes	yes	yes	yes	yes	yes	no	yes	yes	yes
Pinprick level	yes	yes	yes	no	yes	yes	no	no	yes	yes
muscle weakness	no	no	no	no	no	no	no	no	no	no
EMG	abnormal	abnormal	NA	NA	Normal	NA	Normal			NA



We selected four affected individuals (cases I, II, VI and VII) for WES analysis. WES was performed at Perkin Elmer sequencing facilities using the Agilent SureSelect whole exome 50Mb capture kit (v4) and Illumina HiSeq2000. We obtain high quality sequencing results, with an average of 97.5% of the reads mapping to the reference sequence and 72% of the nucleotides covered at 30X or more (Table XII). Sequence alignment and base calling was performed using Burrows Wheeler Alignment (BWA) and GATK v.2.9 via the Perkin Elmer Webportal. No mutations were uncovered in the genes for the hereditary sensory and autonomic neuropathies I-V (HSANI-V), in the Fabry's  $\alpha$ -galactosidase gene nor the hexoaminidase A and B of GM2 gangliosidosis. We compared the sequencing results of all four affected and kept only the variants shared by all and inherited in a dominant manner. Since familial hereditary painful neuropathies can be considered rare diseases, we have filtered against variants databases (dbSNP132, 1000genomes) and in-house collection of 10 FC controls exomes to exclude common variants. For the next step of filtering, we selected variants present in coding sequences (CDS) and donor or acceptor splice-sites resulting in 21 candidate variants. Finally, we manually filtered against dbSNP137 and the exome variant server database to further eliminate other common variants. This last step of filtering led to a short list of four candidate variants (Table XIII). Among them, only the missense variant (I403T) in *NAGLU* was predicted to be pathogenic by the bioinformatics programs Polyphen and Mutation Taster. Because of the small number of candidate variants, we performed a segregation analysis for all four variants in the family. Only the p.Ile403Thr mutation in exon 6 of *NAGLU* segregated with the disease, with all 12 affected cases carrying the missense. This variant has not been previously observed in any available databases or in over 50 FC controls. Screening of the exons and intron-exon boundaries of *NAGLU* in an additional 50 cases affected by a FC sensory ataxia or a painful neuropathy did not detect any other pathological changes.

**Table XII:** Statistical analysis of exome sequencing results.

INDIVIDUAL ID	TOTAL READS	% READ MAPPED	% READS PROPERLY PAIRED	MEAN COVERAGE	%COVERAGE > 30X
I	52 232 644	98.4	99.1	54.5	67.3
II	89 795 734	98.2	94.7	90.4	79.3
III	90 570 462	98	95.1	90	79.9
IV	81 995 146	95.5	92.1	73.3	61.5

**Table XIII:** Candidate variants listuncovered by WES analysis.

chromosome	position	gene	region	type	dbsnp	1000genomes	refbase	Affected	dbSNP135	EVS	Change	Polyphen	Disease
chr13	24243247	TNFRSF19	CDS	snv		0	G	R	rs145685720	MAF:0,39%	R419Q	benign	
chr13	50129660	RCBTB1	CDS	snv		0	G	S	rs13971493	C:0,18%	E198D	benign	Alcohol dependance susceptibility
chr15	40660708	DISP2	CDS	snv		0	C	S			P799A	benign	
chr17	40695232	NAGLU	CDS	snv		0	T	Y			I403T	damaging	Mucopolysaccharidosis IIIB (recessive)

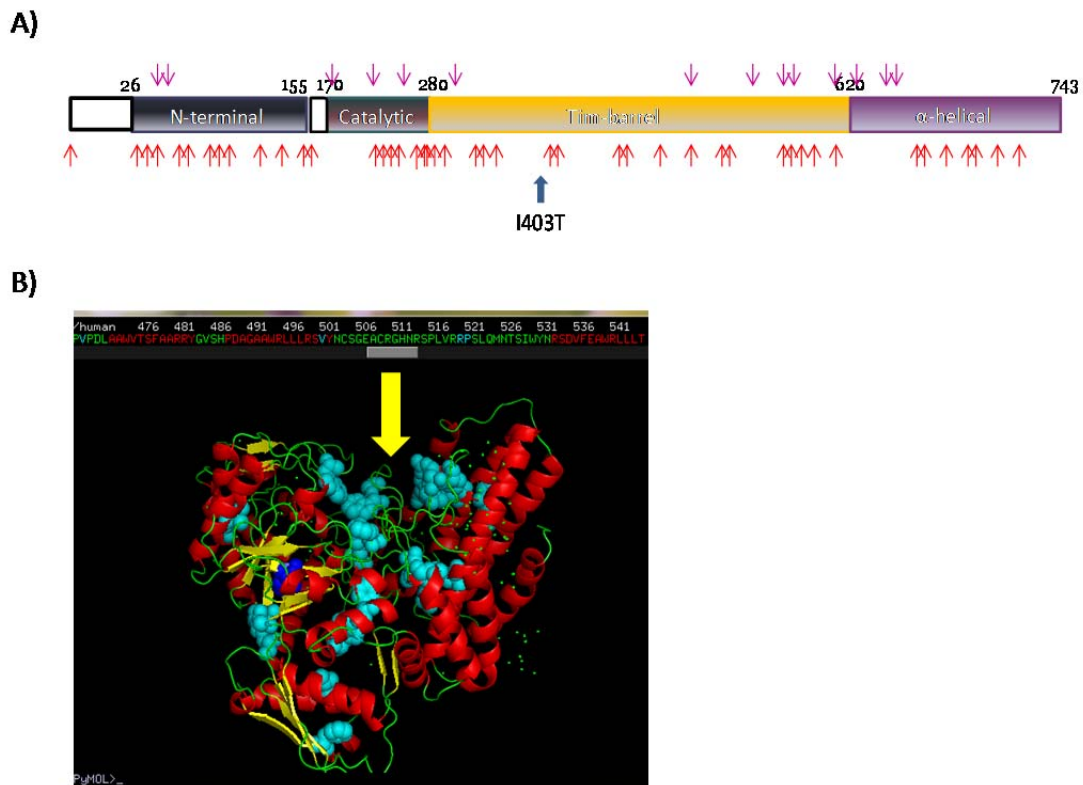
The *NAGLU* gene encodes the  $\alpha$ -N-acetyl-glucosaminidase enzyme, involved in the degradation of the glycosaminoglycan compound heparan sulfate (HS)<sup>57</sup>. More than 100 recessive mutations in this gene have previously been associated with Mucopolysaccharidosis IIIB (MPS IIIB) disorder (MIM: 252920), a severe childhood onset neurodegenerative disease. Most of patients affected by MPS IIIB have a complete loss of enzyme activity. Impaired function of the enzyme leads to the accumulation of partially degraded mucopolysaccharides inside lysosomes, and increased excretion of HS in urine<sup>57-59</sup>. Because of the association of *NAGLU* mutations with a recessive disease, we first searched for a second mutation in our family. Sequencing of the entire cDNA and 93% of the intronic regions has not uncovered any putative pathological changes. To confirm the biological impact of the p.Ile403Thr variant we performed an enzymatic assay using a previously described method<sup>60</sup> standardized in the Clinical Medical Genetics laboratory of the CHU-Sainte-Justine Hospital on isolated leucocytes of cases I, VI and VII. A 10-43% decrease enzyme activity was observed for all of them (Table XIV) when compared with a control individual. A decrease in the obligate carrier range previously described for

NAGLU heterozygote mutation carriers<sup>61</sup> was observed for cases I and VII, confirming the biological impact of the p.Ile403Thr variant.

**Table XIV:** NAGLU activity measurement from patients' leucocytes.

Patients' ID	NAGLU activity (nmol/h/mg)
I	0.99
VI	1.4
VII	0.88
control	1.92

Mutations causing MPS IIIB have been usually associated with severe childhood phenotypes though a growing number of attenuated phenotypes are being described<sup>62</sup>. Though peripheral nerve involvement has not been described in cases or in the few published autopsies<sup>63</sup>, sensory changes, ataxia and sleep disturbance have been described in some of the milder heterozygote cases in a Finnish family<sup>57</sup>. It has been suggested that mutations causing severe phenotype are present in clusters according to the protein 3D structure<sup>64</sup>. The location of the p.Ile403Thr mutation on a linear protein representation and the human NAGLU 3D protein structure show that it maps to a cluster of mutations that are associated with severe MPS IIIB phenotypes (Figure 10). Four distinct domains are present in the NAGLU protein. The amino acid at position 403 is not part of the catalytic domain, but is present in the Tim-barrel domain. This domain is important for the formation of the catalytic site as well as the formation of the NAGLU homotrimer<sup>64; 65</sup>. The location of the p.Ile403Thr mutation on the protein structure (Figure 10B), as well as bioinformatics analysis (IMutant and MuPro), support an indirect impact of the mutation on enzyme activity by disturbing the protein stability. The more recent evidence that NAGLU forms a functional trimer<sup>64</sup> may explain why the mutated protein has an even greater impact on function in vitro.



**Figure 10:** Protein location of p.Ile403Thr NAGLU mutation. **A)** Linear representation of the four domains of the NAGLU protein showing the p.Ile403Thr variant, and mutations associated with attenuated phenotypes (pink arrows) and severe phenotypes (red arrows). **B)** 3D structure of the human NAGLU protein showing the p.Ile403Thr mutation (dark blue) and severe mutations clusters (light blue). The yellow arrow represents the catalytic site.

In this study, only the p.Ile403Thr variant in the gene *NAGLU* was identified in association with the dominant painful sensory polyneuropathy. We demonstrated by an enzymatic assay a decrease activity of NAGLU in patients' leucocytes. According to MuPro and IMutant bioinformatics tools and the protein domain location, the missense variant is predicted to lead to an unstable protein that is either degraded in the lysosome or fail to exit the endoplasmic reticulum<sup>64</sup>.

Recessive mutations in *NAGLU* have usually been associated with a very severe lysosomal storage disorder, MPS IIIB, though milder forms with prolonged walking and life expectancy are increasingly being described<sup>57; 62</sup>. Only in the mildest cases has a sensory polyneuropathy with ataxia has been documented<sup>57</sup>. Since deep tendon reflexes were preserved late in the course, the more severe neuropsychiatric symptoms may have not allowed the exploration of neuropathy that is, at least early in the disease, very subtle. Could a pain syndrome contribute to the explosive behaviour of some patients? In the case of parents or other carriers in families it is more likely that they were lost to follow-up and never examined for such mild late onset symptoms with no cognitive decline. It seems that dominant mutations in *NAGLU* could be responsible for late-onset phenotypes. Two lines of evidence suggest that *NAGLU* mutations could be implicated in late neurodegenerative diseases. First, an association study with common SNPs present in *NAGLU* suggested an association between rs2071046 and rs2676533 and Parkinson's disease (PD)<sup>66</sup>. However, no Parkinson symptoms or signs were observed in our cases. Secondly, it has been observed that *NAGLU* KO transgenic mice accumulate hyperphosphorylated Tau in neurons supporting that it may also be a tauopathy<sup>67</sup>. A secondary accumulation of gangliosides GM2 and GM3 in visceral organs and brain has also been documented<sup>68</sup>. A deficit in the hexosaminidase enzyme A and B cause GM2 gangliosidosis disease. This disease is associated with many neurological features such as painful sensory polyneuropathy<sup>50</sup>. Though late milder phenotypes have been observed in other lysosomal diseases, most cases were carriers of two mutations. Phenotypes in heterozygote carriers have been described in only two lysosomal storage disorders. In Pompe disease, caused by mutation in *GAA*, it was suggested that a milder phenotypes was observed in carriers<sup>69</sup>. Heterozygote carriers of mutations in *GBA*, responsible for the recessive Gaucher disease, have a higher risk of developing PD<sup>9</sup>. This study suggest that carriers of severe pathogenic mutation in *NAGLU*, and potentially other lysosomal diseases, may develop late milder phenotypes caused by slow neurotoxicity that may clinically by their milder nature not resemble the recessive condition. With the growing use of exome sequencing this phenomenon may be documented more frequently. Furthermore, our observation underlines that late onset painful neuropathies may be caused

by gene mutation, and that *NAGLU* should be sequenced in all such late cases with no other obvious cause.

## **ACKNOWLEDGMENTS**

We would like to thank all participants who partake in this study. We would also like to thank Dr. Edward Szekeres, at Perkin Elmer, for the bioinformatics support. This work was supported by the Fondation GO, the Association de la neuropathie sensitive and the Monaco Family Fondation. MT received the Frederick Banting and Charles Best Doctoral Fellowship from the Canadian Institute of Health Research (CIHR).

### **3.2 Phénotype intermédiaire chez les porteurs de mutations associées aux maladies de surcharge lysosomale.**

L'avancement technologique des dernières années m'a donné l'opportunité d'utiliser une approche de séquençage NGS pour l'identification de mutations dans un groupe de maladies hétérogènes, tel que les neuropathies. Dans ce contexte, nous avons identifié une mutation dans le gène *NAGLU* par séquençage exomique dans une grande famille CF atteinte d'une polyneuropathie sensitive douloureuse avec début tardif. Nous avons démontré par un test enzymatique une diminution de l'activité enzymatique chez nos patients. Ces résultats soulèvent la question de la possibilité que les porteurs de mutations dans le gène *NAGLU*, ou dans d'autres gènes associés à des maladies lysosomales, développent un phénotype intermédiaire moins sévère et plus tardif. Bien que cette possibilité ait été suggérée pour d'autres maladies lysosomales, aucune étude ne démontre clairement cette liaison. D'autres analyses sont nécessaires afin de confirmer cette relation de cause à effet.

À ce stade de l'étude, nous n'avons pas été en mesure d'effectuer le test d'efficacité de l'enzyme *NAGLU* sur plus de trois individus. Nous observons une diminution de l'activité chez tous les patients testés en comparaison avec un contrôle prélevé au même moment. Par

contre, pour l'individu VI cette diminution n'est pas significative et se trouve dans l'intervalle des valeurs pour individus normaux. Par contre, un certain chevauchement des valeurs d'activité enzymatique existe entre les porteurs et les individus normaux et l'étendu de ce chevauchement n'est pas bien connue<sup>61</sup>. Pour cette raison, nous ne pouvons pas exclure que cette valeur corresponde aussi à une activité enzymatique déficiente pouvant mener à un phénotype intermédiaire. Nous voulons refaire une analyse d'efficacité enzymatique sur la totalité des individus recrutés (atteint et non-atteint) afin d'évaluer l'intervalle d'activité enzymatique pathogène dans cette famille. De plus, comme une variabilité phénotypique est présente dans cette famille les données enzymatiques pourraient nous permettre de faire une étude de corrélation phénotype-génotype.

Nos résultats génétiques ainsi qu'enzymatiques sont très suggestif de la présence d'un phénotype intermédiaire chez les porteurs de mutations *NAGLU*. Afin de valider notre hypothèse d'autres familles avec des symptômes cliniques chevauchant ceux de MPS IIIB devraient être recrutées. Nous concentrons nos efforts à identifier des familles MPS IIIB au Québec, dans lesquelles nous pourrions évaluer le statut clinique des parents et peut être des autres porteurs. Comme il s'agit d'une maladie à début tardif et tout de même légère, nous pensons que l'association entre le phénotype des porteurs et le gène *NAGLU* n'a pas été fait. Le recrutement d'une grande famille Québécoise nous a permis de faire cette association.

Lorsqu'il s'agit de maladie de surcharge lysosomale, le traitement le plus évident est le remplacement enzymatique (Van Gelder CM. Exp Opin Pharmacol 2013). Par contre, comme le phénotype des patients atteints de MPS IIIB est principalement dû à une atteinte du SNC et que l'enzyme injecté ne traverse pas la barrière hémato-encéphalique, ce type de traitement n'a pas obtenu de succès, Le phénotype intermédiaire observé chez nos patients semble plutôt être une gangliopathie, donc une atteinte du système nerveux périphérique, le remplacement enzymatique pourrait être une thérapie envisageable chez nos patients.

## **Chapitre 4:**

Discussion: Approches traditionnelles et  
approches de nouvelle génération pour  
l'identification de gènes de maladies  
mendéliennes



La recherche de mutations de gènes causant des maladies mendéliennes est depuis plusieurs décennies une grande priorité de la médecine. En plus de permettre un meilleur diagnostic et conseil génétique, l'identification du gène responsable en ouvrant la porte à une meilleure compréhension pathogénique peut mener à de nouveaux traitements<sup>1</sup>. Ce type de recherche de gènes fut grandement influencée par les avancements technologiques et les grands projets d'envergures, tel que le séquençage du génome humain<sup>2</sup> et le projet Hapmap<sup>8</sup>. Dans le cadre de ma formation graduée, j'ai été témoin de plusieurs de ces avancements en génétique humaine dont certains, comme les méthodes de séquençage de nouvelle génération (NGS), ont directement influencé la recherche de gènes.

#### **4.1 Des approches traditionnelles au séquençage de nouvelle génération**

Autant les approches traditionnelles de clonage positionnelle que les techniques de séquençage de nouvelle génération ont obtenu du succès dans l'identification des gènes causant des maladies mendéliennes<sup>1; 3; 11</sup>. Malgré ce succès incontesté, des limitations de ces approches font en sorte que certaines maladies ou groupe de maladies demeurent sans diagnostic moléculaire.

L'utilisation des approches traditionnelles par cartographie de gènes se sont avérées très efficace dans l'identification de gènes causant des maladies mendéliennes. Ces approches impliquent principalement une combinaison d'analyse de liaison à l'échelle du génome suivit par un séquençage de gènes candidats dans la région génomique d'intérêt. L'approche d'analyse de liaison classique, ou de cartographie par homozygotie dans le cas de famille consanguine, nécessite des données génétiques sur un grand nombre d'individus d'une même famille ou de plusieurs familles consanguines ou provenant d'une même région. Bien que ces approches aient identifié plus de 3 000 gènes causant des maladies mendéliennes<sup>3</sup> certaines maladies ne peuvent être étudiées par ces méthodes conventionnelles et par conséquent le gène causal reste à être identifié. Parmi ces maladies, il y a bien sûr les maladies dites complexes ou multigéniques. Il existe aussi des

maladies ou les mutations dans le même gène donnent des phénotypes plus hétérogènes ainsi que des conditions très rares pour lesquelles on retrouve qu'un petit nombre d'individus atteint par famille ou pour lesquelles on retrouve seulement des individus non-connus comme apparenté. Dans ces cas l'approche d'identification d'une région candidate peut être très difficile si elle repose que sur des analyses de partage de marqueurs neutres. De plus, les approches par cartographies vont souvent identifier une région génomique de grande taille ce qui rend l'identification du gène causal laborieuse et coûteuse car elle nécessite souvent le séquençage Sanger d'un grand nombre des gènes de la région<sup>1</sup>.

Les technologies NGS ont apporté des solutions à plusieurs des limitations des approches traditionnelles. Bien qu'une analyse de liaison puisse être effectuée préalablement, cette étape n'est plus une nécessité. Le WES permet le séquençage des régions codantes du génome et le WGS permet le séquençage du génome entier en une seule expérience, éliminant la nécessité de sélectionner des gènes candidats. La production de données est devenue par le fait même moins laborieuse et grandement accélérée ce qui a fait chuter les coûts. Le séquençage NGS offre une plus grande flexibilité dans l'hypothèse qui sera retenu au niveau des analyses, puisqu'une approche utilisant des individus d'une ou plusieurs familles ou utilisant des individus non-apparentés peut être utilisée. Malgré tous les avantages des méthodes NGS en comparaison aux approches traditionnelles, elles ne sont pas sans leurs propres limites qui peuvent empêcher l'identification de variants pathogènes.

L'une des limites du WES est la couverture incomplète des régions exoniques. Les sondes utilisées dans les méthodes de capture sont générées à partir de séquences provenant des bases de données d'annotation de gènes, comme la base de donnée de séquences consensus codante (CCDS) ou la base de donnée RefSeq, donc les exons non-décrits ou non-annotés ne sont pas capturés<sup>3</sup>. Les kits de capture ont une couverture limitée des régions régulatrices (promoteurs, microARN, etc.) limitant ainsi l'identification de mutations pathogènes dans ces régions. Au niveau expérimental, la capture et le séquençage ne sont jamais complètement efficace faisant en sorte que certains exons ne sont pas

couverts. Bien que le WGS partage plusieurs limites avec le WES<sup>9</sup>, il procure l'avantage d'étudier un plus grand nombre de variants et non seulement ceux présent dans les régions codantes. Cette approche produit une très grande quantité de données, l'entreposage et la gestion d'une telle quantité peut être un facteur limitant pour plusieurs laboratoires. De plus, l'interprétation de la pathogénicité des variants non-codant peut être très difficile. Les méthodes NGS semblent produire un taux d'erreurs plus élevés que le séquençage Sanger, mais ceci peut être facilement remédié s'il y a une couverture adéquate des nucléotides<sup>3</sup>. Une bonne couverture est aussi essentielle pour la détection de variants hétérozygotes. Il y a constamment des améliorations dans la capture et dans les approches informatiques afin de remédier à ces limites. Par exemple, des algorithmes informatiques sont maintenant disponibles afin de détecter des CNV à partir de résultats NGS. Malgré ceci, l'identification de mutations d'expansion de régions répétées, comme dans le cas des maladies de type polyglutamine, sont encore difficile à détecter par les NGS. Ceci est principalement dû à la petite taille des fragments séquencés (100-150bp)<sup>9</sup>.

Afin d'obtenir une liste de variants candidats, diverses stratégies de filtrage sont utilisées. Dans la majorité des études deux hypothèses principales sont avancées: 1) Les mutations causant des maladies mendéliennes sont des variants rares ou très rares et donc peu probable d'être présent dans les bases de données; 2) Il est très peu probable que des variants synonymes soient responsable de ces maladies mendéliennes<sup>13</sup>. Cette approche rigoureuse à l'avantage de rapidement diminuer le nombre de variants candidats, mais appelle à la prudence. De plus en plus de variants rares et un taux élevé de variants faux-positifs (estimé à 15-17%) furent introduits dans les bases de données telles que dbSNP. L'approche optimale permettrait d'utiliser une fréquence allélique limite permettant de distinguer les variants rares des variants communs. Cette valeur limite serait tout de même difficile à déterminer et serait différente selon les études dépendant du mode de transmission de la maladie, sa pénétrance et sa prévalence variable dans différentes populations.

Dans le cadre de mon projet de doctorat, j'ai eu l'opportunité de travailler sur des projets utilisant différentes approches. L'approche traditionnelle par gène candidat utilisée pour le projet leucodystrophie a permis l'identification de mutations dans le gène *POLR3B*. La raison principale pour laquelle cette approche conventionnelle a obtenu du succès est la présence d'homogénéité clinique observée entre les patients de l'étude. L'utilisation d'une approche WES de patients non-apparentés aurait probablement mené au même résultat. Dans cette éventualité, nous aurions comparé les résultats de séquençage des différents patients et observé que la totalité des patients possédaient deux variants dans le gène *POLR3A* comme ce fut le cas pour un groupe japonais<sup>37</sup>.

L'hétérogénéité clinique et génétique présente dans les polyneuropathies héréditaires rend l'identification du gène causal par les approches traditionnelles plus laborieux. Dans l'étude génétique d'une famille CF atteinte d'une forme de polyneuropathie sensitive douloureuse, nous avons opté pour une approche par WES afin d'accélérer le processus de découverte. Cette approche a mené à l'identification d'un variant candidat dans le gène *NAGLU*. La combinaison des données de séquençage nouvelle génération à une analyse familiale est l'une des raisons principales du succès de cette approche. En comparant les données de séquences de quatre individus apparentés mais provenant de différentes branches de la même famille nous avons été en mesure de significativement réduire le nombre de variants candidats. Les résultats obtenus pour ce projet sont très suggestifs qu'une mutation hétérozygote dans *NAGLU* soit une des premières mutations identifiées responsable d'une forme tardive de neuropathie douloureuse. Afin de vraiment établir ce lien de cause à effet, d'autres familles devront être identifiées. Dans l'éventualité où nous n'arrivons pas à confirmer la relation entre le variant *NAGLU* et le phénotype sensitif, d'autres approches pourraient être utilisés afin d'identifier le gène causal. Nous pourrions refaire une analyse de nos données d'exomes avec un algorithme détectant les CNV<sup>70</sup>, afin de s'assurer que nous n'avons pas manqué une délétion ou duplication. Comme la mutation pourrait être située dans un intron<sup>10</sup>, nous pourrions envisager un séquençage du génome entier afin de couvrir ces régions. Dans l'éventualité où nous devons aussi chercher pour une

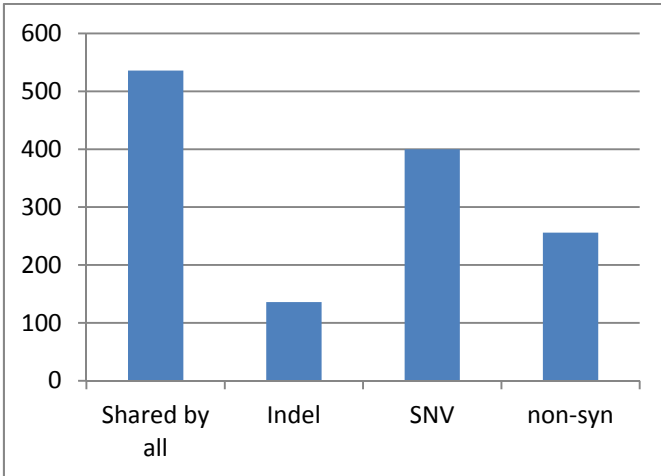
expansion de triplet répétés, il serait peut-être plus avantageux d'utiliser une technologie comme celle de Roche 454 permettant le séquençage de plus gros fragments (environ 400bp) puisque le séquençage de petits fragments comme la méthode d'Illumina ou bien SoliD d'AppliedBiosystem ont de la difficulté à détecter ce type de variation.

#### **4.2 Utilisation de familles CF pour l'identification de gènes candidats.**

L'histoire du peuplement du Québec a eu une grande influence sur la prévalence et la distribution régionale variable des maladies génétiques<sup>14; 71</sup>. La petite taille relative de la population CF en Nouvelle-France ainsi que l'endogamie observée suite à la conquête des Anglais a contribué à moduler la prévalence de certains allèles. Ceci a donné lieu à des variations génétiques, pouvant être associées à des maladies mendéliennes, plus fréquentes dans certaines régions du Québec. Ainsi une plus grande homogénéité des pools génétiques régionaux québécois fut constituée.

Dans le cadre du projet leucodystrophie, l'utilisation de familles CF originaire de la région de Beauce-Bellechasse a grandement aidé à l'identification d'un locus candidat et finalement à des mutations dans le gène *POLR3A*. La recherche d'un haplotype partagé entre les cinq familles démontrant un même phénotype clinique a permis d'identifier une région sur le chromosome 10. Pour ce projet, l'homogénéité génétique observée dans la population CF nous a permis de rapidement identifier un locus et un gène candidat. Sur plusieurs projets dont j'ai eu l'opportunité de travailler pendant mon doctorat, nous n'avons pas eu autant de chance. Malgré l'augmentation constante de la densité des marqueurs SNP sur les puces, certaines régions chromosomiques possèdent un haut taux de marqueurs non-informatifs. Qui plus est, avec ces puces, on détecte souvent plusieurs régions d'homozygoties et d'haplotypes partagées par plusieurs participants qui ne correspondent qu'à des régions communes ancestrales partagées par un grand nombre de CF. Il est donc essentiel de toujours comparer les régions d'intérêts avec des données de génotypage sur des individus CF non atteint de la maladie à l'étude.

Malgré le partage aléatoire plus grand, les particularités des pools génétiques régionaux québécois peut être d'une grande utilité dans l'analyse de données générées par des technologies NGS. Il faut tout de fois être prudent, puisque notre expérience sur plusieurs projets a démontré que certains variants jamais déposés sur des banques de données sont plus fréquent dans la population CF. En comparant les données d'exomes d'une dizaine d'individus CF non-apparentés, atteints de maladies différentes et originaires de différentes régions de la province, nous avons observé un enrichissement de certains allèles (Figure 11). Plusieurs de ces variants étaient absents des bases de données telles que dbSNP, 1000genomes et EVS. Il est donc important de filtrer les résultats de séquençage exomique contre des individus de même population afin d'éliminer les variants communs à cette population. Cette étape appel aussi à une grande vigilance, puisque plusieurs maladies plus fréquentes au Québec ont un taux de porteurs élevé (par exemple 1/22 pour LGMD2D dans les Îles de la Madeleines)<sup>72</sup>. Il faut donc, surtout dans l'étude de maladies à transmission récessive, ne pas utiliser des critères d'exclusion trop stricte qui pourraient éliminer des variants d'intérêts. En développant une bonne stratégie d'analyse selon la population, le mode de transmission, l'âge de début des symptômes ainsi que la prévalence de la maladie, il est possible d'obtenir une petite liste relativement courte de variants candidats tout en limitant la possibilité d'avoir éliminé les variants d'intérêts par les étapes de filtrage.



**Figure 11** : Variants identifiés par séquençage exonique et partagés par 10 individus CF non-apparentés, atteint de maladies différentes et originaires de différentes régions de la province.

#### 4.3 Impact des technologies NGS sur la clinique et la médecine personnalisée

Le développement des technologies NGS ont révolutionné la recherche de gènes causant des maladies Mendéliennes. Ces techniques ont aussi rapidement fait leur place dans le monde du diagnostic moléculaire<sup>73</sup>. Avant l'intégration de WES et WGS dans la gamme de services des laboratoires diagnostics, plusieurs défis techniques et éthiques restent à être surmontés<sup>73</sup>.

Un défi important pour les laboratoires est l'analyse et la gestion des données de séquençage d'exomique. Il est important pour les laboratoires de développer une méthode d'analyse simple et efficace et obtenir l'équipement et le support informatique nécessaire. Une autre difficulté à surmonter est d'obtenir la même sensibilité et spécificité que les méthodes conventionnelles de diagnostic. La couverture incomplète est aussi un problème que devront affronter les laboratoires diagnostic et les cliniciens traitant. Certains gènes associés à des maladies ne sont pas complètement couverts et ainsi des variants pathogènes pourraient être manqués. Il est important de compléter le séquençage en utilisant la méthode Sanger avant de conclure à un diagnostic négatif. Afin de répondre à cette problématique, plusieurs laboratoires proposent une approche de séquençage ciblé. Comme

une validation génétique et parfois biologique sont nécessaires afin de confirmer la pathogénicité d'un variant, les délais d'analyse pourraient être une limitation pour les maladies génétiquement homogènes. Par contre, en présence d'hétérogénéité génétique, où plusieurs gènes peuvent causer le phénotype, les tests conventionnels sont laborieux et coûteux et nécessitent souvent plusieurs années avant qu'un diagnostic soit obtenu<sup>11</sup>. Les tests diagnostiques sont sujets à des normes et une standardisation des procédures. L'implantation d'un tel protocole couvrant l'extraction d'ADN au rapport clinique est un grand défi des laboratoires cliniques. Finalement, l'interprétation des données reste le plus grand défi<sup>73</sup>. Les cliniciens devront développer une expertise afin d'interpréter les rapports et offrir un diagnostic adéquat aux patients. De plus, un consensus éthique devra être établi afin de déterminer si les découvertes fortuites doivent être divulguées ou non.

Les technologies NGS ont le potentiel d'améliorer le conseil génétique offert aux patients atteints de diverses maladies en offrant un diagnostic plus précis. Selon plusieurs experts, cet impact sera plus important dans le diagnostic des maladies hétérogènes où plusieurs nouveaux phénotypes seront associés à des gènes connus.

Pendant ma formation doctorale, j'ai travaillé sur plusieurs projets touchant des maladies différentes. Pour certains de ces projets j'ai utilisé une approche traditionnelle et pour d'autres une approche de séquençage exomique. Certains projets ont eu du succès et pour d'autres l'hétérogénéité clinique et génétique est encore aujourd'hui un obstacle. Une grande leçon que j'ai apprise est que chaque projet, maladie sont différents et développer une hypothèse de recherche et une stratégie d'analyse propre à chacun est la clé du succès.



## Bibliographie

1. Gilissen, C., Hoischen, A., Brunner, H.G., and Veltman, J.A. (2011). Unlocking Mendelian disease using exome sequencing. *Genome biology* 12, 228.
2. Green, E.D., Guyer, M.S., and National Human Genome Research, I. (2011). Charting a course for genomic medicine from base pairs to bedside. *Nature* 470, 204-213.
3. Ku, C.S., Cooper, D.N., Polychronakos, C., Naidoo, N., Wu, M., and Soong, R. (2012). Exome sequencing: dual role as a discovery and diagnostic tool. *Annals of neurology* 71, 5-14.
4. Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 74, 5463-5467.
5. Botstein, D., White, R.L., Skolnick, M., and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics* 32, 314-331.
6. Gusella, J.F., Wexler, N.S., Conneally, P.M., Naylor, S.L., Anderson, M.A., Tanzi, R.E., Watkins, P.C., Ottina, K., Wallace, M.R., Sakaguchi, A.Y., et al. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* 306, 234-238.
7. Olson, M., Hood, L., Cantor, C., and Botstein, D. (1989). A common language for physical mapping of the human genome. *Science* 245, 1434-1435.
8. Mardis, E.R. (2011). A decade's perspective on DNA sequencing technology. *Nature* 470, 198-203.
9. Bras, J., Guerreiro, R., and Hardy, J. (2012). Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease. *Nature reviews Neuroscience* 13, 453-464.
10. Krebs, C.E., and Paisan-Ruiz, C. (2012). The use of next-generation sequencing in movement disorders. *Frontiers in genetics* 3, 75.

11. Nelen, M., and Veltman, J.A. (2012). Genome and exome sequencing in the clinic: unbiased genomic approaches with a high diagnostic yield. *Pharmacogenomics* 13, 511-514.
12. Rehm, H.L. (2013). Disease-targeted sequencing: a cornerstone in the clinic. *Nature reviews Genetics* 14, 295-300.
13. Ku, C.S., Naidoo, N., and Pawitan, Y. (2011). Revisiting Mendelian disorders through exome sequencing. *Human genetics* 129, 351-370.
14. Laberge, A.M., Michaud, J., Richter, A., Lemyre, E., Lambert, M., Brais, B., and Mitchell, G.A. (2005). Population history and its impact on medical genetics in Quebec. *Clinical genetics* 68, 287-301.
15. Kohlschutter, A., and Eichler, F. (2011). Childhood leukodystrophies: a clinical perspective. *Expert review of neurotherapeutics* 11, 1485-1496.
16. Vanderver, A., Tonduti, D., Bernard, G., Lai, J., Rossi, C., Carosso, G., Quezado, M., Wong, K., and Schiffmann, R. (2013). More than hypomyelination in Pol-III disorder. *Journal of neuropathology and experimental neurology* 72, 67-75.
17. Bernard, G., Thiffault, I., Tetreault, M., Putorti, M.L., Bouchard, I., Sylvain, M., Melancon, S., Laframboise, R., Langevin, P., Bouchard, J.P., et al. (2010). Tremor-ataxia with central hypomyelination (TACH) leukodystrophy maps to chromosome 10q22.3-10q23.31. *Neurogenetics* 11, 457-464.
18. Tetreault, M., Putorti, M.L., Thiffault, I., Sylvain, M., Vanderver, A., Schiffmann, R., Brais, B., and Bernard, G. (2012). TACH leukodystrophy: locus refinement to chromosome 10q22.3-23.1. *The Canadian journal of neurological sciences Le journal canadien des sciences neurologiques* 39, 122-123.
19. Chouery, E., Delague, V., Jalkh, N., Salem, N., Kfoury, J., Rodriguez, D., Chabrol, B., Boespflug-Tanguy, O., Levy, N., Serre, J.L., et al. (2011). A whole-genome scan in a large family with leukodystrophy and oligodontia reveals linkage to 10q22. *Neurogenetics* 12, 73-78.
20. Bernard, G., Chouery, E., Putorti, M.L., Tetreault, M., Takanohashi, A., Carosso, G., Clement, I., Boespflug-Tanguy, O., Rodriguez, D., Delague, V., et al. (2011). Mutations of POLR3A encoding a catalytic subunit of RNA polymerase Pol III cause a

- recessive hypomyelinating leukodystrophy. *American journal of human genetics* 89, 415-423.
21. Orcesi, S., Tonduti, D., Uggetti, C., Larizza, D., Fazzi, E., and Balottin, U. (2010). New case of 4H syndrome and a review of the literature. *Pediatric neurology* 42, 359-364.
  22. Dumay-Odelot, H., Durrieu-Gaillard, S., Da Silva, D., Roeder, R.G., and Teichmann, M. (2010). Cell growth- and differentiation-dependent regulation of RNA polymerase III transcription. *Cell cycle* 9, 3687-3699.
  23. Schiffmann, R., and van der Knaap, M.S. (2009). Invited article: an MRI-based approach to the diagnosis of white matter disorders. *Neurology* 72, 750-759.
  24. Schiffmann, R., and van der Knaap, M.S. (2004). The latest on leukodystrophies. *Current opinion in neurology* 17, 187-192.
  25. Bekiesinska-Figatowska, M., Mierzewska, H., Kuczynska-Zardzewialy, A., Szczepanik, E., and Obersztyn, E. (2010). Hypomyelination, hypogonadotropic hypogonadism, hypodontia - First Polish patient. *Brain & development* 32, 574-578.
  26. Timmons, M., Tsokos, M., Asab, M.A., Seminara, S.B., Zirzow, G.C., Kaneski, C.R., Heiss, J.D., van der Knaap, M.S., Vanier, M.T., Schiffmann, R., et al. (2006). Peripheral and central hypomyelination with hypogonadotropic hypogonadism and hypodontia. *Neurology* 67, 2066-2069.
  27. Wolf, N.I., Harting, I., Boltshauser, E., Wiegand, G., Koch, M.J., Schmitt-Mechelke, T., Martin, E., Zschocke, J., Uhlenberg, B., Hoffmann, G.F., et al. (2005). Leukoencephalopathy with ataxia, hypodontia, and hypomyelination. *Neurology* 64, 1461-1464.
  28. Wolf, N.I., Harting, I., Innes, A.M., Patzer, S., Zeitler, P., Schneider, A., Wolff, A., Baier, K., Zschocke, J., Ebinger, F., et al. (2007). Ataxia, delayed dentition and hypomyelination: a novel leukoencephalopathy. *Neuropediatrics* 38, 64-70.
  29. Atrouni, S., Daraze, A., Tamraz, J., Cassia, A., Caillaud, C., and Megarbane, A. (2003). Leukodystrophy associated with oligodontia in a large inbred family: fortuitous association or new entity? *American journal of medical genetics Part A* 118A, 76-81.
  30. Jasiak, A.J., Armache, K.J., Martens, B., Jansen, R.P., and Cramer, P. (2006). Structural biology of RNA polymerase III: subcomplex C17/25 X-ray structure and 11 subunit enzyme model. *Molecular cell* 23, 71-81.

31. Fernandez-Tornero, C., Bottcher, B., Rashid, U.J., Steuerwald, U., Florchinger, B., Devos, D.P., Lindner, D., and Muller, C.W. (2010). Conformational flexibility of RNA polymerase III during transcriptional elongation. *The EMBO journal* 29, 3762-3772.
32. Wu, C.C., Lin, Y.C., and Chen, H.T. (2011). The TFIIF-like Rpc37/53 dimer lies at the center of a protein network to connect TFIIC, Bdp1, and the RNA polymerase III active center. *Molecular and cellular biology* 31, 2715-2728.
33. Steenweg, M.E., Vanderver, A., Blaser, S., Bizzi, A., de Koning, T.J., Mancini, G.M., van Wieringen, W.N., Barkhof, F., Wolf, N.I., and van der Knaap, M.S. (2010). Magnetic resonance imaging pattern recognition in hypomyelinating disorders. *Brain : a journal of neurology* 133, 2971-2982.
34. Dittmar, K.A., Goodenbour, J.M., and Pan, T. (2006). Tissue-specific differences in human transfer RNA expression. *PLoS genetics* 2, e221.
35. Feinstein, M., Markus, B., Noyman, I., Shalev, H., Flusser, H., Shelef, I., Liani-Leibson, K., Shorer, Z., Cohen, I., Khateeb, S., et al. (2010). Pelizaeus-Merzbacher-like disease caused by AIMP1/p43 homozygous mutation. *American journal of human genetics* 87, 820-828.
36. Boespflug-Tanguy, O., Aubourg, P., Dorboz, I., Begou, M., Giraud, G., Sarret, C., and Vaurs-Barriere, C. (2011). Neurodegenerative disorder related to AIMP1/p43 mutation is not a PMLD. *American journal of human genetics* 88, 392-393; author reply 393-395.
37. Saitsu, H., Osaka, H., Sasaki, M., Takanashi, J., Hamada, K., Yamashita, A., Shibayama, H., Shiina, M., Kondo, Y., Nishiyama, K., et al. (2011). Mutations in POLR3A and POLR3B encoding RNA Polymerase III subunits cause an autosomal-recessive hypomyelinating leukoencephalopathy. *American journal of human genetics* 89, 644-651.
38. Perlman, S.J., and Mar, S. (2012). Leukodystrophies. *Advances in experimental medicine and biology* 724, 154-171.
39. Potic, A., Brais, B., Choquet, K., Schiffmann, R., and Bernard, G. (2012). 4H syndrome with late-onset growth hormone deficiency caused by POLR3A mutations. *Archives of neurology* 69, 920-923.

40. Terao, Y., Saitsu, H., Segawa, M., Kondo, Y., Sakamoto, K., Matsumoto, N., Tsuji, S., and Nomura, Y. (2012). Diffuse central hypomyelination presenting as 4H syndrome caused by compound heterozygous mutations in POLR3A encoding the catalytic subunit of polymerase III. *Journal of the neurological sciences* 320, 102-105.
41. Tetreault, M., Choquet, K., Orcesi, S., Tonduti, D., Balottin, U., Teichmann, M., Fribourg, S., Schiffmann, R., Brais, B., Vanderver, A., et al. (2011). Recessive mutations in POLR3B, encoding the second largest subunit of Pol III, cause a rare hypomyelinating leukodystrophy. *American journal of human genetics* 89, 652-655.
42. Wolff, A., Koch, M.J., Benzinger, S., van Waes, H., Wolf, N.I., Boltshauser, E., and Luder, H.U. (2010). Rare dental peculiarities associated with the hypomyelinating leukoencephalopathy 4H syndrome/ADDH. *Pediatric dentistry* 32, 386-392.
43. Bernard, G., and Vanderver, A. (1993). Pol III-Related Leukodystrophies. In *GeneReviews*, R.A. Pagon, T.D. Bird, C.R. Dolan, K. Stephens, and M.P. Adam, eds. (Seattle (WA)).
44. Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., and Pagano, A. (2007). The expanding RNA polymerase III transcriptome. *Trends in genetics : TIG* 23, 614-622.
45. Chiu, Y.H., Macmillan, J.B., and Chen, Z.J. (2009). RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. *Cell* 138, 576-591.
46. Hsieh, S.T. (2010). Pathology and functional diagnosis of small-fiber painful neuropathy. *Acta neurologica Taiwanica* 19, 82-89.
47. Hoeijmakers, J.G., Faber, C.G., Lauria, G., Merkies, I.S., and Waxman, S.G. (2012). Small-fibre neuropathies--advances in diagnosis, pathophysiology and management. *Nature reviews Neurology* 8, 369-379.
48. Lauria, G., and Lombardi, R. (2012). Small fiber neuropathy: is skin biopsy the holy grail? *Current diabetes reports* 12, 384-392.
49. Botez, S.A., and Herrmann, D.N. (2010). Sensory neuropathies, from symptoms to treatment. *Current opinion in neurology* 23, 502-508.
50. Sedel, F., Turpin, J.C., and Baumann, N. (2007). [Neurological presentations of lysosomal diseases in adult patients]. *Revue neurologique* 163, 919-929.
51. Auer-Grumbach, M. (2008). Hereditary sensory neuropathy type I. *Orphanet journal of rare diseases* 3, 7.

52. Houlden, H., King, R., Blake, J., Groves, M., Love, S., Woodward, C., Hammans, S., Nicoll, J., Lennox, G., O'Donovan, D.G., et al. (2006). Clinical, pathological and genetic characterization of hereditary sensory and autonomic neuropathy type 1 (HSAN I). *Brain : a journal of neurology* 129, 411-425.
53. Einarsdottir, E., Carlsson, A., Minde, J., Toolanen, G., Svensson, O., Solders, G., Holmgren, G., Holmberg, D., and Holmberg, M. (2004). A mutation in the nerve growth factor beta gene (NGFB) causes loss of pain perception. *Human molecular genetics* 13, 799-805.
54. Minde, J., Toolanen, G., Andersson, T., Nennesmo, I., Remahl, I.N., Svensson, O., and Solders, G. (2004). Familial insensitivity to pain (HSAN V) and a mutation in the NGFB gene. A neurophysiological and pathological study. *Muscle & nerve* 30, 752-760.
55. Spinazzi, M., Angelini, C., and Patrini, C. (2010). Subacute sensory ataxia and optic neuropathy with thiamine deficiency. *Nature reviews Neurology* 6, 288-293.
56. Axelrod, F.B., and Gold-von Simson, G. (2007). Hereditary sensory and autonomic neuropathies: types II, III, and IV. *Orphanet journal of rare diseases* 2, 39.
57. Selmer, K.K., Gilfillan, G.D., Stromme, P., Lyle, R., Hughes, T., Hjorthaug, H.S., Brandal, K., Nakken, S., Misceo, D., Egeland, T., et al. (2012). A mild form of Mucopolysaccharidosis IIIB diagnosed with targeted next-generation sequencing of linked genomic regions. *European journal of human genetics : EJHG* 20, 58-63.
58. Yogalingam, G., and Hopwood, J.J. (2001). Molecular genetics of mucopolysaccharidosis type IIIA and IIIB: Diagnostic, clinical, and biological implications. *Human mutation* 18, 264-281.
59. Beesley, C.E., Jackson, M., Young, E.P., Vellodi, A., and Winchester, B.G. (2005). Molecular defects in Sanfilippo syndrome type B (mucopolysaccharidosis IIIB). *Journal of inherited metabolic disease* 28, 759-767.
60. Pollard, L.M., Jones, J.R., and Wood, T.C. (2013). Molecular characterization of 355 mucopolysaccharidosis patients reveals 104 novel mutations. *Journal of inherited metabolic disease* 36, 179-187.

61. Vance, J.M., Conneally, P.M., Wappner, R.S., Yu, P.L., Brandt, I.K., and Pericak-Vance, M.A. (1981). Carrier detection in Sanfilippo syndrome type B: report of six families. *Clinical genetics* 20, 135-140.
62. Valstar, M.J., Bruggenwirth, H.T., Olmer, R., Wevers, R.A., Verheijen, F.W., Poorthuis, B.J., Halley, D.J., and Wijburg, F.A. (2010). Mucopolysaccharidosis type IIIB may predominantly present with an attenuated clinical phenotype. *Journal of inherited metabolic disease* 33, 759-767.
63. Hadfield, M.G., Ghatak, N.R., Nakoneczna, I., Lippman, H.R., Myer, E.C., Constantopoulos, G., and Bradley, R.M. (1980). Pathologic findings in mucopolysaccharidosis type IIIB (Sanfilippo's syndrome B). *Archives of neurology* 37, 645-650.
64. Meiyappan M, C.M., Norton AW. (2012). Crystal Structure of Human Alpha-N-acetylglucosaminidase. In United States Patent Application publication. (United States, Shire Human Genetics Therapies).
65. Ochoa-Leyva, A., Barona-Gomez, F., Saab-Rincon, G., Verdel-Aranda, K., Sanchez, F., and Soberon, X. (2011). Exploring the Structure-Function Loop Adaptability of a (beta/alpha)(8)-Barrel Enzyme through Loop Swapping and Hinge Variability. *Journal of molecular biology* 411, 143-157.
66. Winder-Rhodes, S.E., Garcia-Reitbock, P., Ban, M., Evans, J.R., Jacques, T.S., Kemppinen, A., Foltynie, T., Williams-Gray, C.H., Chinnery, P.F., Hudson, G., et al. (2012). Genetic and pathological links between Parkinson's disease and the lysosomal disorder Sanfilippo syndrome. *Movement disorders : official journal of the Movement Disorder Society* 27, 312-315.
67. Ohmi, K., Kudo, L.C., Ryazantsev, S., Zhao, H.Z., Karsten, S.L., and Neufeld, E.F. (2009). Sanfilippo syndrome type B, a lysosomal storage disease, is also a tauopathy. *Proceedings of the National Academy of Sciences of the United States of America* 106, 8332-8337.
68. Heldermon, C.D., Hennig, A.K., Ohlemiller, K.K., Ogilvie, J.M., Herzog, E.D., Breidenbach, A., Vogler, C., Wozniak, D.F., and Sands, M.S. (2007). Development of sensory, motor and behavioral deficits in the murine model of Sanfilippo syndrome type B. *PloS one* 2, e772.

69. Kroos, M., Hoogeveen-Westerveld, M., van der Ploeg, A., and Reuser, A.J. (2012). The genotype-phenotype correlation in Pompe disease. *American journal of medical genetics Part C, Seminars in medical genetics* 160, 59-68.
70. Fromer, M., Moran, J.L., Chambert, K., Banks, E., Bergen, S.E., Ruderfer, D.M., Handsaker, R.E., McCarroll, S.A., O'Donovan, M.C., Owen, M.J., et al. (2012). Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *American journal of human genetics* 91, 597-607.
71. Laberge, A.M. (2007). [Prevalence and distribution of genetic diseases in Quebec: impact of the past on the present]. *Medecine sciences : M/S* 23, 997-1001.
72. Tetreault, M., Srour, M., Allyson, J., Thiffault, I., Loisel, L., Robitaille, Y., Bouchard, J.P., and Brais, B. (2011). Founder mutation for alpha-sarcoglycan-LGMD2D in a Magdalen Islands Acadian cluster. *The Canadian journal of neurological sciences Le journal canadien des sciences neurologiques* 38, 747-752.
73. Ku, C.S., Cooper, D.N., and Roukos, D.H. (2012). The 'sequence everything' approach and personalized clinical decision challenges. *Expert review of molecular diagnostics* 12, 319-322.

## Web source

Align GVGD: [agvgd.iarc.fr/](http://agvgd.iarc.fr/)

BDGP splice site prediction by Neural Network: [www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)

Clustal W: [www.ch.embnet.org/software/clustalw.html](http://www.ch.embnet.org/software/clustalw.html)

EVS: [.evs.gs.washington.edu](http://.evs.gs.washington.edu)

IMutant: [gpcr2.biocomp.unibo.it/~emidio/I-Mutant/I-Mutant\\_help.html](http://gpcr2.biocomp.unibo.it/~emidio/I-Mutant/I-Mutant_help.html)

Mutations Taster: [www.mutationtaster.org](http://www.mutationtaster.org)

Mupro: [www.ics.uci.edu/~baldig/mutation.html](http://www.ics.uci.edu/~baldig/mutation.html)

PhD-SNP: [snps.biofold.org/phd-snp/phd-snp.html](http://snps.biofold.org/phd-snp/phd-snp.html)

Polyphen: [genetics.bwh.harvard.edu/pphz](http://genetics.bwh.harvard.edu/pphz)

Primer 3: [Frodo.mit.edu/Primer3/](http://Frodo.mit.edu/Primer3/)

SIFT: [sift.jcvi.org/](http://sift.jcvi.org/)



UCSC: [genome.ucsc.edu](http://genome.ucsc.edu)

## **Annexe 1: Curriculum Vitae**

### D E G R E E   Q U A L I F I C A T I O N S

- **PhD. Molecular Biology** , Université de Montréal 2009-  
D<sup>r</sup> Bernard Brais
- **MSc. Molecular Biology**, Université de Montréal 2003-2006  
D<sup>r</sup> Bernard Brais
- **BSc. Biochemistry**, Université de Montréal 2000-2003

### P R O F E S S I O N A L   E X P E R I E N C E

- **PhD candidate Molecular Biology;** 2009-  
Dr. Bernard Brais, CRCHUM  
Thesis: Genetic characterization of new neurogenetic diseases in the French Canadian population.
- **Research Assistant** 2006-2008  
Laboratoire Neurogenetique, Dr. Bernard Brais, CRCHUM
- **MSc student** 2003-2006  
Dr. Bernard Brais, CRCHUM  
Thesis: Characterisation of a new recessive form of congenital muscular dystrophy with hyperlaxity in the French Canadian population.
- **Research Assistant** 2002-2003  
Molecular diagnostic laboratory, Dr. George Chong, MJGH
- **Stagiaire** Summer2002  
Dr. George Chong, MJGH  
Research project: Mutation identification in French-Canadian families affected by hereditary colon cancer.

### F E L L O W S H I P S

- **Doctoral Fellowship Frederick Banting et Charles Best** 2010-2013  
Canadian Institutes of Health Research
- **Doctoral Fellowship** DECLINE  
Association Française contre les myopathies
- **Bourse d'excellence** 2009  
Programme biologie moléculaire, Université de Montréal
- **CIHR-short term research grant** 2005  
Allowing participation to CGDN Bioinformatic Workshop

### A W A R D S

- **Oral presentation** 2012  
Programme de Biologie moléculaire, Université de Montréal
- **Poster presentation** 2005  
CR-CHUM annual conference
- **Travel Award** 2004  
World Muscle Society

## INVITED SPEAKER

- **Munster University**, Institute for evolution and biodiversity, Germany 2012  
Title: Hypomyelinating Leukodystrophy: Identifying the causative genes by combining regional founder effect and traditional genetic approaches.
- **Hôpital Sainte-Justine**, Neurology department, Montreal 2011  
Title: Genetic Heterogeneity in Congenital Muscular Dystrophies.

## PUBLICATIONS

**Published referred papers:** 14 peer review articles, 5 as a first author.

(\* co-first author)

Daoud H\*, **Tetreault M\***, Gibson W, Guerero K, Ana Cohen, Janina Gburek-Augustat, Matthis Synofzik, Bernard Brais, Cathy Stevens, Rocio Sanchez-Carpintero, Cyril Goizet, Sakkubai Naidu, Adeline Vanderver and Geneviève Bernard. Mutations in *POLR3A* and *POLR3B* are a major cause of hypomyelinating leukodystrophy. ACCEPTED *Journal of Medical Genetic*

Thiffault I, Dicaire MJ, **Tetreault M**, Huang KN, Demers-Lamarche J, Bernard G, Duquette A, Larivière R, Gehring K, Montpetit A, McPherson P, Richter A, Mitchell G, Dupre N, Prevost C, Bouchard JP, Mathieu J, Brais B. Diversity of ARSACS mutations in French-Canadians. *Canadian Journal of Neurological Sciences* ACCEPTED #1622

Bayat V\*, Thiffault I\*, Jaiswal M, **Tetreault M**, Donti T, Sasarman F, Bernard G, Demers-Lamarche J, Dicaire MJ, Mathieu J, Vanasse M, Bouchard JP, Rioux MF, Lesage J, Lourenco CM, Li Z, Haueter C, Shoubridge E, Graham BH, Brais B, Bellen H. Mutations in the Mitochondrial Methionyl-tRNA Synthetase Cause a Neurodegenerative Phenotype in Flies and a Recessive Ataxia (ARSAL) in humans. *Plos Biology*. 2012; 10 (3)

**Tetreault M**, Putorti ML, Thiffault I, Sylvain M, Vanderver A, Schiffmann R, Brais B, Bernard G. TACH leukodystrophy: locus refinement to chromosome 10q22.3-23.1. *Canadian Journal of Neurological Sciences* 2012 Jan; 39(1): 122-3

**Tetreault M\***, Choquet K\*, Orcesi S, Tonditu D, Balottin U, Teichmann M, Fribourg S, Schiffmann R, Brais B, Vanderver A, Bernard G. Recessive mutations in *POLR3B*, encoding the second largest subunit of pol III, cause a rare hypomyelinating leukodystrophy. *American Journal of Human Genetic* 2011 Nov 11; 89(5): 652-5

Bernard G\*, Chouery E\*, Putorti ML, **Tetreault M**, Takanohashi A, Carosso G, Clement I, Boespflug-Tanguy O, Fribourg S, Teichmann M, Mergabane A, Schiffmann R, Vanderver A, Brais B. Mutations of the RNA polymerase pol III subunit *POLR3A* cause a recessive hypomyelinating leukodystrophy. *American Journal of Human Genetic* 2011 Sep 9; 89(3): 415-23

**Tetreault M**, Srour M, Allyson J, Thiffault I, Loisel L, Robitaille Y, Bouchard JP, Brais B. Founder mutation for  $\alpha$ -sarcoglycan-LGMD2D in a Magdalen Islands Acadian cluster. *Canadian Journal of Neurological Sciences* 2011 Sep; 38(5): 747-52

Bernard G, Thiffault I, **Tetreault M**, Putorti ML, Bouchard I, Sylvain M, Melancon S, Laframboise R, Langevin P, Bouchard JP, Vanasse M, Vanderver A, Sebire G, Brais B. Tremor ataxia with central

hypomyelination (TACH) leukodystrophy maps to chromosome 10q22.3-10q23.31. *Neurogenetics* 2010 Oct; 11(4): 457-64

Loggia ML., Bushnell C., **Tetreault M.**, Thiffault I, Bherer C, Mohammed NK, Kuchinad AA, Laferriere A, Dicaire MJ, Loisel L, Mogil JS, Brais B. Carrier of recessive *WKN1/HSN2* mutations for hereditary sensory and autonomic neuropathy type 2 (HSAN2) are more sensitive to thermal stimuli. *Journal of Neuroscience* 2009 Feb; 29 (7): 2162-6

Gosselin I, Thiffault I., **Tetreault M.**, Chau V, Dicaire MJ, Loisel L, Emond Monique, Senderek J, Mathieu J, Dupre M, Vanasse M, Puymirat J, Brais B. Founder SH3TC2 mutations are responsible for a CMT4C French-Canadians cluster. *Neuromuscular Disorder* 2008 Jun; 18 (6): 483-92

Rossignol E, Mathieu J, Thiffault I, Tetreault M, Dicaire MJ, Chrestian N, Dupre N, Puymirat J, Brais B. A novel founder SCN4A mutation causes painful cold-induced myotonia in French-Canadians. *Neurology* 2007 Nov 13; 69(20): 1937-41

Jarry J, Rioux MF, Bolduc V, Robitaille Y, Khoury V, Thiffault I, **Tetreault M.**, Loisel L, Bouchard JP, Brais B. A novel autosomal recessive limb-girdle muscular dystrophy with quadriceps atrophy maps to 11p13-p12. *Brain* 2007 Feb; 130(Pt2): 368-80; Epub 2006 Sep 28

**Tetreault M.**, Duquette A., Thiffault I., Bherer C., Jarry J., Loisel L., Banwell B., D'Anjou G., Mathieu J., Robitaille Y., Vanasse M., Brais B. A new form of congenital muscular dystrophy with joint hyperlaxity maps to 3p23-21. *Brain* 2006 Aug; 129(Pt 8):2077-84. Epub 2006 Jun 7.

Thiffault I., Rioux M.F., **Tetreault M.**, Jarry J., Loisel L., Poirier J., Gros-louis F., Mathieu J., Vanasse M., Rouleau G.A., Bouchard J.P., Lesage J., Brais B. A new form of autosomal recessive spastic ataxia with intrafamilial variability associated with a leukodystrophy maps to 2q33-34. *Brain* 2006 Sep; 129 (Pt 9):2332-40. Epub 2006 May 3.

**Published abstracts:** 36 published abstracts, (5 oral and 31 poster presentations); 12 as a first author.