





**La généalogie comme facteur prédictif de la progression des atteintes en dystrophie myotonique de type 1: une démonstration de la puissance de l'intersectorialité**

**par Joanie Bouchard**

**Mémoire présenté à l'Université du Québec à Chicoutimi en vue de l'obtention du grade de Maître ès sciences (M. Sc.) en Sciences cliniques et biomédicales**

Québec, Canada

© Joanie Bouchard, 2024

## RÉSUMÉ

La dystrophie myotonique de type 1 (DM1) est une maladie autosomique dominante. La DM1 est causée par une répétition anormale du triplet de nucléotide CTG situé sur le gène DMPK. Il existe une grande variabilité entre les individus autant en lien avec la gravité des atteintes que le rythme de progression, et ce, même à l'intérieur d'une même famille. Ainsi, à l'heure actuelle, un manque de connaissances flagrant réside dans la prédiction de l'évolution des divers signes et symptômes associés à la DM1. L'objectif de cette étude était d'évaluer si des regroupements généalogiques pouvaient être utilisés pour mieux comprendre la gravité et la progression des symptômes dans la DM1. La cohorte était composée de 200 patients du Saguenay–Lac-Saint-Jean qui ont participé à une étude longitudinale sur 20 ans. Plusieurs mesures cliniques pertinentes à la progression de la maladie ont été recueillies. La longueur des répétitions CTG a été comparée avec différents phénotypes et mesures. De plus, le fichier de population BALSAC a été utilisé pour reconstituer les généalogies. Des regroupements généalogiques ont été effectués à l'aide du coefficient d'apparentement, et ce, sur plusieurs générations. Le coefficient d'apparentement a été utilisé afin de faire des comparaisons avec différentes mesures cliniques et voir si cette mesure pourrait devenir un outil pour évaluer la progression de la maladie. La conclusion a été que le coefficient d'apparentement est une mesure trop large pour évaluer l'effet d'un seul gène. Bien que les analyses n'aient pas permis de déceler une relation claire entre le coefficient d'apparentement et différentes variables cliniques, les résultats ont permis de développer une méthode pour former de grands regroupements qui pourra être utilisée pour former de plus grandes familles dans d'autres analyses jumelant la génétique et la généalogie pour mieux comprendre une maladie.

## TABLE DES MATIÈRES

RÉSUMÉ.....	ii
TABLE DES MATIÈRES .....	iii
LISTE DES TABLEAUX.....	v
LISTE DES FIGURES.....	vi
LISTE DES ABRÉVIATIONS.....	viii
REMERCIEMENTS .....	0
CHAPITRE 1 .....	1
INTRODUCTION.....	1
1.1 ATTEINTES MULTISYSTÉMIQUES ET VARIABILITÉS PHÉNOTYPIQUES EN DM1 .....	1
1.2 CLASSIFICATION DE LA MALADIE.....	3
1.3 TRANSMISSION AUTOSOMIQUE DOMINANTE .....	5
1.4 LA MUTATION DE COURTE RÉPÉTITION EN TANDEM CTG.....	6
1.5 INSTABILITÉ GERMINALE OU PHÉNOMÈNE D’ANTICIPATION.....	7
1.6 MOSAÏCISME.....	8
1.7 GÈNES IMPLIQUÉS DANS LA DM1.....	9
1.8 PRÉVALENCE DE LA DM1 .....	11
1.9 EFFET FONDATEUR .....	12
1.9.1 EFFET FONDATEUR AU SAGUENAY-LAC-SAINT-JEAN .....	13
1.10 LES GÉNÉALOGIES.....	15
1.11 LE FICHER DE POPULATION BALSAC.....	15
1.11.1 LA COMPLÉTUDE .....	16
1.11.2 LE COEFFICIENT D’APPARENTEMENT .....	16
1.12 PROBLÉMATIQUE.....	18
1.12.1 BUT ET HYPOTHÈSES .....	18
1.12.2 OBJECTIFS .....	18
CHAPITRE 2 .....	19
MÉTHODOLOGIE.....	19
2.1 POPULATION À L’ÉTUDE.....	19
2.1.1 CRITÈRES D’INCLUSION.....	19
2.1.2 CRITÈRES D’EXCLUSION.....	19
2.1.3 CONSIDÉRATIONS ÉTHIQUES .....	19
2.2 VARIABLES DE L’ÉTUDE.....	20
2.2.1 ÉTUDE CORRÉLATIONNELLE DES DONNÉES .....	22
2.2.2 INTERVALLES DES RÉPÉTITIONS CTG .....	22
2.3 CRÉATION DU FICHER GÉNÉALOGIQUE.....	23
2.3.1 COLLECTE DES DONNÉES GÉNÉALOGIQUES .....	23

2.3.2 INCORPORATION DE LA BASE DE DONNÉES DANS LE FICHER BALSAC...	23
2.3.3 VALIDATION DES GÉNÉALOGIES À L'AIDE DE GENLIB .....	23
2.4 REGROUPEMENTS BASÉS SUR LE COEFFICIENT D'APPARENTEMENT .....	24
CHAPITRE 3 .....	26
RÉSULTATS .....	26
3.1 ÉTUDE CORRÉLATIONNELLE DES DONNÉES .....	26
3.2 ANALYSE PHÉNOTYPIQUE .....	27
3.3 ANALYSE DES REGROUPEMENTS PAR INTERVALLE DES RÉPÉTITIONS CTG .....	29
3.4 VALIDATIONS DES GÉNÉALOGIES .....	32
3.5 CONCEPTION DES RÉSEAUX À L'AIDE DU COEFFICIENT D'APPARENTEMENT. .....	33
3.6 ANALYSE DES REGROUPEMENTS PAR COEFFICIENT D'APPARENTEMENT..	34
3.6.1 RÉPARTITION EN 8 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	35
3.6.2 LES FAMILLES ET LES PHÉNOTYPES EN FONCTION DES REGROUPEMENTS .....	39
CHAPITRE 4 .....	41
DISCUSSION .....	41
4.1 RETOUR SUR LES OBJECTIFS .....	41
4.2 PORTRAIT DU NOMBRE DE RÉPÉTITIONS CTG EN FONCTION DES PHÉNOTYPES ET DES SYMPTÔMES .....	41
4.3 NOMBRE DE RÉPÉTITIONS CTG ET PHÉNOMÈNE D'ANTICIPATION.....	43
4.4 LE COEFFICIENT D'APPARENTEMENT .....	44
4.5 PERSPECTIVES .....	45
4.6 RETOMBÉES.....	46
4.7 FORCES ET LIMITES DE L'ÉTUDE .....	47
CONCLUSION .....	49
LISTE DE RÉFÉRENCES .....	50
CERTIFICATION ÉTHIQUE .....	56
ANNEXE 1 .....	57

## LISTE DES TABLEAUX

TABLEAU 1 – PRÉVALENCE DE LA DMI DANS DIFFÉRENTS LIEUX .....	11
TABLEAU 2 – CHI CARRÉ POUR LES PHÉNOTYPES ET LA GÉNÉRATION EN FONCTION DES INTERVALLES DU NOMBRE DE CTG .....	32
TABLEAU 3 – DÉNOMBREMENT DES 3 PHÉNOTYPES DE DMI EN FONCTION DANS LES 8 SOUS REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT SELON LES TROIS VARIABLES UTILISÉES.....	35
TABLEAU 4 – CHI CARRÉ DES FAMILLES EN FONCTIONS DE DIFFÉRENTS NOMBRES DE SOUS- DIVISIONS DE REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	39
TABLEAU 5 – CHI CARRÉ DES PHÉNOTYPES EN FONCTIONS DE DIFFÉRENTS NOMBRES DE SOUS-DIVISIONS DU COEFFICIENT D'APPARENTEMENT. ....	40

## LISTE DES FIGURES

FIGURE 1 – DIFFÉRENTS SYSTÈMES TOUCHÉS PAR LA DM1 [8] .....	2
FIGURE 2 – REPRÉSENTATION DES 5 PHÉNOTYPES DE LA MALADIE SUR UNE ÉCHELLE DE TEMPS [8, 13] .....	3
FIGURE 3 – REPRÉSENTATION DE LA TRANSMISSION AUTOSOMIQUE DOMINANTE .....	5
FIGURE 4 – APERÇU DU COEFFICIENT D'APPARENTEMENT SELON UN ARBRE GÉNÉALOGIQUE. ....	17
FIGURE 5 – CORRÉLATIONS ENTRE LES VARIABLES UTILISÉES DANS CETTE ÉTUDE MESURÉES CHEZ LES PATIENTS DM1 .....	26
FIGURE 6 – NOMBRE DE RÉPÉTITIONS CTG EN FONCTION DES DIFFÉRENTS PHÉNOTYPES DE LA MALADIE. ....	27
FIGURE 7 – LA FORCE DE PRÉHENSION DE LA PINCE EN FONCTION DES DIFFÉRENTS PHÉNOTYPES DE LA MALADIE.....	28
FIGURE 8 – SÉVÉRITÉ DE LA MALADIE EN FONCTION DE LA MESURE DE FORCE DU DORSIFLÉCHISSEUR DE LA CHEVILLE.....	29
FIGURE 9 – LA FORCE DE PINCEMENT SELON 3 REGROUPEMENTS PAR INTERVALLES DU NOMBRE DE CTG. ....	30
FIGURE 10 – FORCE DU DORSIFLÉCHISSEUR DE LA CHEVILLE SELON 3 REGROUPEMENTS PAR INTERVALLES DU NOMBRE DE CTG .....	31
FIGURE 11 – COMPLÉTUDE MOYENNE DE 200 PATIENTS ATTEINTS DE DM1 ORIGINAIRES DU SAGUENAY-LAC-SAINT-JEAN.....	32
FIGURE 12 – RÉSEAUX BASÉS SUR LE COEFFICIENT D'APPARENTEMENT.....	34
FIGURE 13 – LE NOMBRE DE RÉPÉTITIONS CTG EN FONCTION DE 8 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	37
FIGURE 14 – LA FORCE DE PINCEMENT EN FONCTION DE 8 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	38
FIGURE 15 – LA FORCE DU DORSIFLÉCHISSEUR DE LA CHEVILLE EN FONCTION DE 8 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	39
FIGURE 16 – NOMBRE DE RÉPÉTITION CTG EN FONCTION DE 3 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	57
FIGURE 17 – NOMBRE DE RÉPÉTITION CTG EN FONCTION DE 5 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	57
FIGURE 18 – NOMBRE DE RÉPÉTITION CTG EN FONCTION DE 8 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	58
FIGURE 19 – NOMBRE DE RÉPÉTITION CTG EN FONCTION DE 10 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	58

FIGURE 20 – NOMBRE DE RÉPÉTITION CTG EN FONCTION DE 15 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	59
FIGURE 21 – NOMBRE DE RÉPÉTITION CTG EN FONCTION DE 20 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	59
FIGURE 22 – FORCE DE PINCEMENT EN FONCTION DE 3 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	60
FIGURE 23 – FORCE DE PINCEMENT EN FONCTION DE 5 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	60
FIGURE 24 – FORCE DE PINCEMENT EN FONCTION DE 8 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	61
FIGURE 25 – FORCE DE PINCEMENT EN FONCTION DE 10 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	61
FIGURE 26 – FORCE DE PINCEMENT EN FONCTION DE 15 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	62
FIGURE 27 – FORCE DE PINCEMENT EN FONCTION DE 20 SOUS-REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	62
FIGURE 28 – FORCE DU DORSIFLÉCHISSEUR DE LA CHEVILLE EN FONCTION DE 3 SOUS- REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	63
FIGURE 29 – FORCE DU DORSIFLÉCHISSEUR DE LA CHEVILLE EN FONCTION DE 4 SOUS- REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	63
FIGURE 30 – FORCE DU DORSIFLÉCHISSEUR DE LA CHEVILLE EN FONCTION DE 8 SOUS- REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	64
FIGURE 31 – FORCE DU DORSIFLÉCHISSEUR DE LA CHEVILLE EN FONCTION DE 10 SOUS- REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	64
FIGURE 32 – FORCE DU DORSIFLÉCHISSEUR DE LA CHEVILLE EN FONCTION DE 15 SOUS- REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	65
FIGURE 33 – FORCE DU DORSIFLÉCHISSEUR DE LA CHEVILLE EN FONCTION DE 20 SOUS- REGROUPEMENTS DU COEFFICIENT D'APPARENTEMENT.....	65



## LISTE DES ABRÉVIATIONS

ADN	Acide désoxyribonucléique
CMNMJ	Clinique des maladies neuromusculaires de Jonquière
CORAMH	Corporation de recherche et d'action sur les maladies héréditaires.
DM	Dystrophie musculaire
DM1	Dystrophie myotonique de type 1
DM2	Dystrophie myotonique de type 2
DMPK	Dystrophia Mytonica Protein Kinase
FRQS	Fonds de recherche du Québec – Santé
GRIMN	Groupe de recherche interdisciplinaire en maladies neuromusculaires
IDPATIENT	Identification du participant
STR	<i>Short tandem repeat</i>
UQAC	Université du Québec à Chicoutimi
FUQAC	Fondation de l'Université du Québec à Chicoutimi
SP-PCR	Small-Pool Polymerase Chain Reaction

## REMERCIEMENTS

Je voudrais tout d'abord remercier mon directeur Simon Girard de m'avoir initiée en bio-informatique. Après 6 années à travailler dans son laboratoire, il a toujours été humain et il a toujours su m'encourager à continuer dans la bonne direction. Merci de m'avoir confié un projet qui était nouveau et peu connu dans le domaine scientifique, mais qui m'a tellement appris.

J'aimerais également remercier Claudia Moreau qui a pu être présente pour répondre à mes innombrables questions et qui a toujours été présente pour un soutien aux analyses. Sans toi, Claudia, le laboratoire ne serait pas ce qu'il est devenu aujourd'hui. Un merci également à toute l'équipe du laboratoire de Genopop qui m'a épaulée et conseillée lors de mes analyses. J'aimerais également remercier Élise Duchesne et Cynthia Gagnon de m'avoir permis de travailler avec leurs données et de m'avoir guidée dans les différentes analyses. Je tiens également à remercier le Groupe de recherche interdisciplinaire en maladies neuromusculaires (GRIMN) de m'avoir accueillie à la clinique et pour la révision des dossiers.

Je tiens également à remercier les organismes subventionnaires pour leur aide financière lors de mon parcours, soit les Fonds de recherche du Québec - Santé (FRQS), la Corporation de recherche et d'action sur les maladies héréditaires (CORAMH), le Centre intersectoriel en santé durable (CISD) ainsi que la Fondation de l'Université du Québec à Chicoutimi (FUQAC).

Un merci particulier à mon copain de m'avoir endurée dans les pires moments et d'avoir su me redonner le courage afin de compléter ce que je croyais sans fin. Également un gros merci à calcul Québec et calcul Canada de nous permettre de stocker nos données et de faire nos analyses sur des supercalculateurs.

## **CHAPITRE 1 INTRODUCTION**

La dystrophie myotonique est une maladie neurodégénérative faisant partie des dystrophies musculaires regroupant 160 maladies affectant les muscles, le système nerveux périphérique ou la jonction neuromusculaire [1, 2]. La dystrophie musculaire (DM) est caractérisée par une détérioration progressive de la force musculaire, c'est-à-dire qu'il y a un affaiblissement des muscles avec l'évolution de la maladie [2]. Certaines DM peuvent se développer dès la naissance tandis que d'autres vont se développer plus tardivement à l'âge adulte [1]. Les DM sont étudiées depuis de nombreuses années, les premières recherches sur ce sujet ont été faites dans les années 1850 par le chercheur Edward Meryon. Il est l'un des premiers chercheurs à avoir effectué des analyses sur des groupes familiaux de patients atteints de dystrophie musculaire [3]. Dans ses recherches, Edward Meryon décrit en détail l'histologie des atteintes musculaires qui sera associée plus tard à la dystrophie musculaire de Duchenne. La différenciation des différentes dystrophies musculaires n'était pas encore très claire à cette époque [4].

La dystrophie myotonique se divise en deux types, soit la dystrophie myotonique de type 1 (DM1) et la dystrophie myotonique de type 2 (DM2). Les deux types ont des symptômes similaires, mais les atteintes de la DM2 sont souvent moins sévères que la DM1 [1]. La principale différence est sur le plan génétique, la maladie ne touche pas les mêmes régions génomiques. Dans l'étude qui suit, la DM1 sera davantage abordée.

### **1.1 ATTEINTES MULTISYSTÉMIQUES ET VARIABILITÉS PHÉNOTYPIQUES**

#### **EN DM1**

La DM1, également appelée maladie de Steinert, est une maladie génétique caractérisée par une détérioration progressive du système musculaire. La première description de la DM1 a été faite par Steinert en 1909 [5]. La maladie qui peut affecter autant les femmes que les hommes est multisystémique, car elle affecte plusieurs systèmes et organes à la fois, ce qui ne facilite pas son

diagnostic. En effet, la maladie peut toucher les systèmes nerveux, respiratoire, musculaire, visuel, cardiovasculaire, digestif et/ou reproducteur [6, 7] (Figure 1).

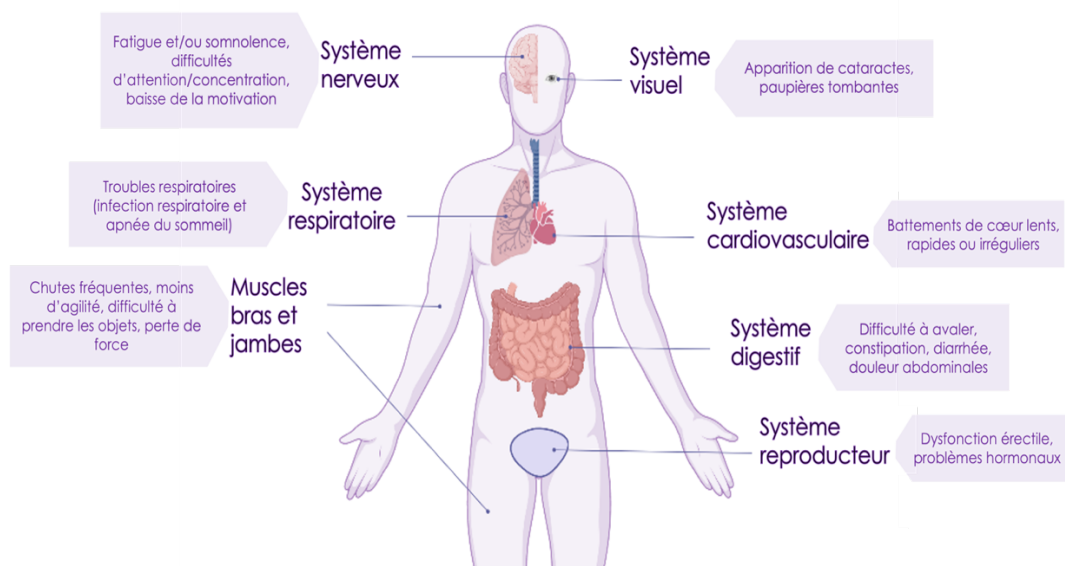


Figure 1 – Différents systèmes touchés par la DM1.  
Tiré de CORAMH, Maladies héréditaires dominantes, 2023.

Les principaux symptômes présents chez les patients atteints de DM1 sont la faiblesse musculaire et la myotonie [5]. La myotonie est un trouble du tonus musculaire caractérisé par un délai dans le relâchement musculaire suite à une contraction [8, 9]. La faiblesse musculaire est souvent associée à un déficit moteur ou une fatigue musculaire [5]. Également, les muscles distaux sont les parties les plus affectées, mais plus la maladie évolue, plus les muscles proximaux sont touchés [8, 10, 11]. Les atteintes peuvent être nombreuses et leur gravité varie d'une personne à l'autre [5]. La DM1 est une maladie neuromusculaire avec un phénotype très large et imprévisible [10]. Tout d'abord, l'âge d'apparition des premiers symptômes est très variable, la maladie peut apparaître autant chez un nouveau-né que chez une personne âgée [1, 5]. Les symptômes vont également évoluer avec le temps et être différents pour chaque patient. La sévérité des systèmes touchés variera d'une personne à l'autre. Certains vont naître avec de nombreux symptômes très sévères tandis que d'autres vont vivre leur vie normalement et découvrir à un âge avancé qu'ils sont atteints de la maladie [12].

Il est difficile de prédire la progression de la maladie étant donné sa grande diversité individuelle. Les chercheurs ont trouvé différents moyens de classifier la maladie.

## 1.2 CLASSIFICATION DE LA MALADIE

Les chercheurs n'ont pas trouvé de consensus de classification, il existe donc plusieurs manières de classifier la maladie selon certains critères distincts. Pour l'étude qui suit, la classification s'est faite selon cinq phénotypes (congénitale, infantile, juvénile, adulte et tardif) (Figure 2)[13, 14]. La frontière entre les différentes «formes» n'est pas rigide, il est parfois difficile d'émettre le diagnostic phénotypique d'un individu.

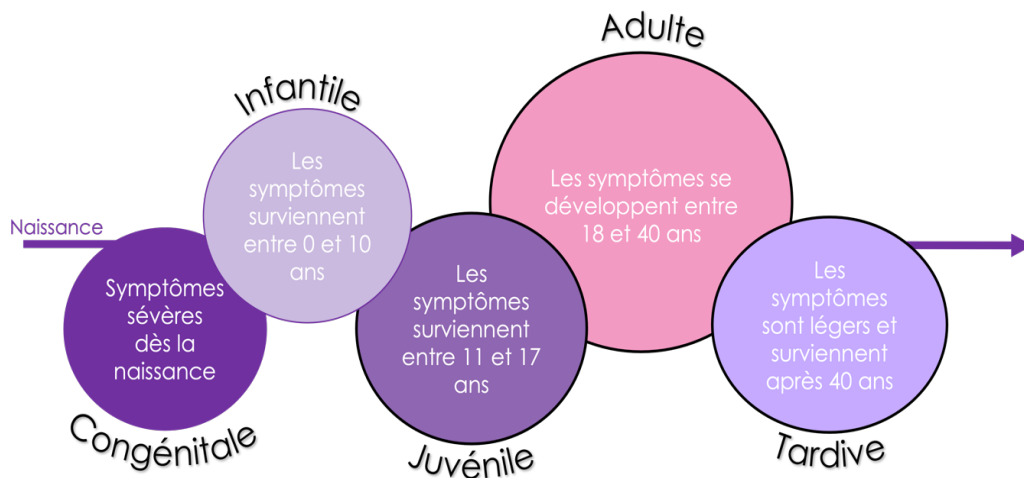


Figure 2 – Représentation des cinq phénotypes de la maladie selon l'âge d'apparition des symptômes.

Tiré de CORAMH, Maladies héréditaires dominantes, 2023.

Cette division se base sur une classification de 2016 élaborée par des chercheurs ayant créé la base de données des patients atteints de DM1 de la Clinique des maladies neuromusculaires de Jonquièrre (CMNMJ). Les patients ont été divisés en cinq grands phénotypes caractérisés selon l'âge d'apparition de la maladie, le nombre de répétitions CTG et le résultat obtenu avec l'échelle de mesure *Muscular Impairment Rating Scale* (MIRS) [6, 15, 16]. C'est à l'aide de cette échelle qu'il est possible de mesurer l'atteinte musculaire chez les patients atteints de DM1. Elle se divise en cinq niveaux (1)

Pas de faiblesse, 2) Faiblesse légère, 3) Faiblesse distale 4) Faiblesse proximale faible à modérée et 5) Faiblesse distale sévère) [17].

Les cinq phénotypes, soit le congénital, l'infantile, le juvénile, l'adulte et le tardif, seront décrits dans les prochains paragraphes dans l'ordre du plus sévère au moins sévère (Figure 2).

Premièrement, la forme congénitale est la plus sévère. Pour ce phénotype, les symptômes vont être présents dans la phase prénatale ou dans les premiers mois de vie de l'enfant. Le phénotype congénital est reconnu dans la plupart des études. Il y a de grandes atteintes liées aux systèmes nerveux central et respiratoire. Ces atteintes vont souvent être accompagnées d'une déficience intellectuelle. Dès la naissance, il est possible de voir certaines déformations squelettiques montrant la gravité de ce phénotype. C'est une forme très sévère entraînant la mort dans 15 % des cas. Les femmes portant un enfant atteint de ce phénotype présentent un risque accru de fausse couche, d'accouchement prématuré ou compliqué [4].

Le phénotype infantile survient entre 1 mois et 10 ans. Dans cette forme, il est possible de voir différentes atteintes du système nerveux central telles que des retards de langage ou des problèmes de sommeil. Généralement, les symptômes clés de la DM1 tels que la myotonie et les problèmes musculaires arrivent plus tard dans le développement de la maladie.

Le troisième phénotype est le juvénile qui a été identifié par des chercheurs afin de séparer le phénotype adulte et infantile [18]. Les symptômes surviennent lors de l'adolescence, soit entre 11 et 17 ans et sont similaires à ceux retrouvés dans le phénotype adulte. C'est principalement l'âge d'apparition qui permet de distinguer ces deux phénotypes.

Le quatrième phénotype est l'adulte pour lequel les symptômes surviennent entre 18 et 40 ans, mais la maladie est souvent diagnostiquée vers 40 ans lorsque les symptômes s'aggravent [19]. Chez la femme, une grossesse peut déclencher les premiers symptômes ou aggraver ceux déjà présents [20].

Finalement, le cinquième phénotype est le tardif qui est souvent considéré comme une forme asymptomatique de la maladie [2]. Les personnes atteintes ne s'aperçoivent pas des symptômes tout au long de leur vie. Les symptômes, qui sont légers, surviennent après l'âge de 40 ans [7]. Les plus

récurrents sont la calvitie et les cataractes. Ces symptômes sont souvent associés au vieillissement normal d'une personne et non à une pathologie. Ce phénotype est souvent diagnostiqué lors de recherche sur l'origine génétique de la maladie d'un membre de la famille. La classification des différents phénotypes permet de regrouper les individus en fonction de leurs symptômes et de la sévérité de la maladie.

### 1.3 TRANSMISSION AUTOSOMIQUE DOMINANTE

La maladie est transmise de façon autosomique dominante [21, 22]. Cela signifie que pour un couple où l'un des deux parents est atteint, il y a 50 % de probabilité de transmettre la maladie à la descendance, et ce pour chaque grossesse (Figure 3).

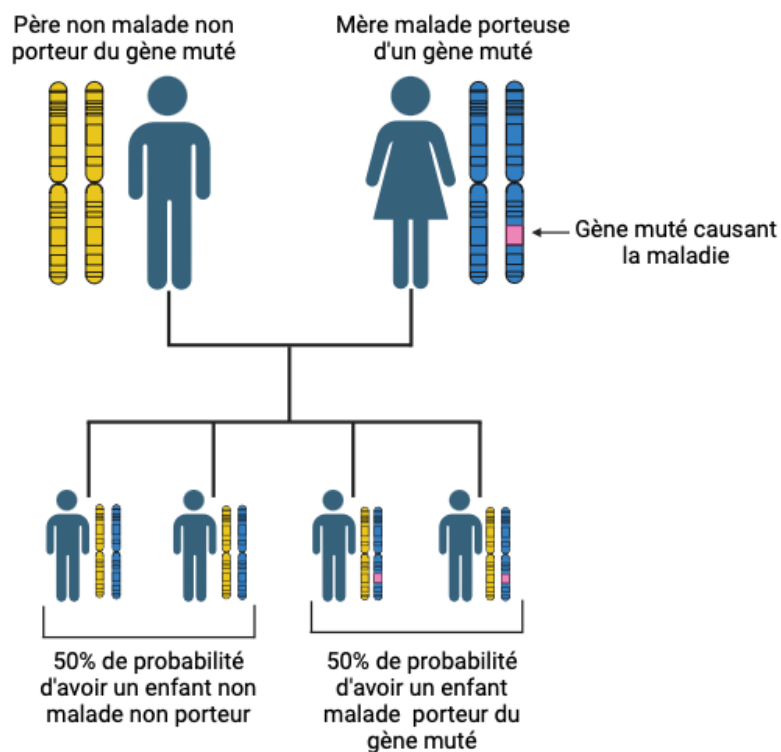


Figure 3 – Représentation de la transmission autosomique dominante  
Tiré de CORAMH, Maladies héréditaires dominantes, 2023.

Il suffit donc d'une seule copie du gène pour avoir la maladie. Lorsqu'un membre d'une famille apprend le diagnostic, les autres membres peuvent par prévention faire un test afin de voir

s'ils ont le variant causant la maladie. Les symptômes peuvent être tardifs, donc chaque membre peut passer un test afin de valider le diagnostic. Pour faire ce diagnostic, il faut faire une analyse génétique et estimer la longueur des répétitions CTG qui a un impact sur la gravité des symptômes.

#### **1.4 LA MUTATION DE COURTE RÉPÉTITION EN TANDEM CTG**

C'est en 1992 que des chercheurs découvrent que la DM1 est une maladie à triplet. Ce type de maladie est causée par une expansion instable d'un triplet de nucléotides dans l'ADN qui peut également être appelé courte répétition en tandem [4].

La courte répétition en tandem, ou *Short tandem repeat* (STR), est une mutation causée par la répétition consécutive de 1 à 6 nucléotides [23]. Ces répétitions sont fréquentes dans le génome, avec souvent peu ou pas de conséquence. Cependant, le nombre de répétitions est parfois d'une longueur inhabituel et se situe sur une région génomique essentielle à certaines fonctions [23]. Cela peut entraîner des conséquences au niveau fonctionnel telles que l'apparition d'une pathologie. Les maladies à triplets sont causées par une expansion de trois nucléotides. Pour la DM1, les trois nucléotides causant la maladie sont une cytosine, une thymine et une guanine (CTG). Ce triplet est situé dans la région 3' non traduite du gène *Dystrophia Mytonica Protein Kinase* (DMPK) sur le locus q13.3 du chromosome 19 [24]. Les 3 nucléotides peuvent se répéter entre 5 et 37 fois chez les personnes saines et peuvent atteindre plus de 1000 répétitions chez les patients atteints de DM1 [25]. Il semble que le nombre de répétitions concorde avec les différents phénotypes et la gravité des symptômes [16, 26]. Le phénotype tardif est celui qui a le moins de répétition CTG tandis que le phénotype congénital est celui qui a le nombre de répétitions CTG le plus élevé [16]. En 2019, une étude a été réalisée pour évaluer la longueur du nombre de répétitions CTG en fonction de différents symptômes. Ils ont utilisé neuf mesures de force musculaire et respiratoire pour faire des corrélations avec le nombre de CTG [16]. Les corrélations qui étaient les plus fortes étaient la force de pincement et la force de la cheville. Les variables de force pourraient aider à suivre le rythme de progression de



la maladie due à leur corrélation avec la sévérité de la maladie [16]. Cependant, les intervalles du nombre de répétitions CTG pour chacun des phénotypes ne sont pas clairement définis.

Certains chercheurs croient qu'il existe un seuil où le nombre de répétitions n'a plus d'importance sur la gravité de la maladie. Ce nombre de répétitions serait en quelque sorte un niveau de saturation pour la maladie [27]. En effet, certains chercheurs ont récemment effectué des travaux pour mieux comprendre ce phénomène de génotype/phénotype et croient qu'il y aurait deux seuils de nombre de répétitions pour la maladie. Le premier seuil serait à ~50 répétitions où les premiers symptômes pourraient apparaître. Ce seuil serait le seuil le plus faible pour que les symptômes apparaissent. Le second seuil se situerait entre 200 et 300 répétitions CTG et ce seuil quant à lui représente un plafond au regard de la sévérité des symptômes. Les chercheurs ont remarqué qu'au-delà de ce nombre de répétitions, la gravité des symptômes n'était pas exponentielle [27]. Cependant, il manque encore de l'information et de la recherche sur ce deuxième seuil afin de bien comprendre le phénomène. Des chercheurs ont également mis en place des collectes de données plus précises telles que l'âge d'apparition des symptômes et ajouté des mesures de forces musculaires afin d'avoir une meilleure vision de l'évolution de la maladie [28-30].

### **1.5 INSTABILITÉ GERMINALE OU PHÉNOMÈNE D'ANTICIPATION**

Il semble que la répétition sur le gène DMPK présente une instabilité importante. En effet, il existe un phénomène d'anticipation chez les familles de patients atteints de DM1, la maladie devient donc de plus en plus sévère au fil des générations [31, 32]. Ce phénomène est appelé instabilité germinale et se produit lorsque l'on observe aux générations subséquentes des stades plus précoces de la maladie avec plus de symptômes ainsi qu'un plus grand nombre de répétitions CTG [33]. Il a été remarqué que la cascade est hautement accélérée lorsque la mutation est transmise par la mère [34-36]. En effet, une grande expansion de CTG causant le phénotype congénital est presque exclusivement transmise par une mère atteinte [34]. Il est important de mentionner que le phénomène d'anticipation n'est pas toujours présent, parfois les parents peuvent même transmettre moins de

répétitions à leurs enfants, cependant cela est très peu fréquent [37]. Ces exceptions complexifient les recherches sur ce phénomène.

Également, certains individus vont avoir ce qui est appelé une prémutation, c'est-à-dire un nombre de répétitions CTG compris entre 35 et 50 [38]. Avec cette prémutation, les individus n'auront aucun symptôme, mais il est possible qu'avec l'anticipation qu'ils transmettent un nombre plus élevé de répétitions à leur enfant. En héritant d'un nombre supérieur de répétitions CTG, il est possible que leur descendance développe la maladie sans qu'eux-mêmes soient atteints [39]. De plus, il existe un phénomène d'instabilité somatique qui se manifeste par une augmentation du nombre de répétitions CTG avec l'âge, et ce dans les différents tissus [28, 40]. Donc, une prémutation peut évoluer dans le temps et amener le développement d'un phénotype tardif avec des symptômes légers [38]. Étant donné qu'un nombre de répétitions plus élevé est associé à des symptômes plus sévères de la maladie, ces phénomènes pourraient expliquer l'aggravation des symptômes dans le temps [28].

## **1.6 MOSAÏCISME**

Le mosaïcisme est un autre phénomène observé chez les patients DM1 et pouvant complexifier grandement les études sur la maladie. Le mosaïcisme fait référence à une constitution génétique distincte dans les différentes cellules du corps, et ce en provenance d'un seul zygote [16, 28, 32]. Le mosaïcisme germinale se produit au niveau des cellules gonadiques, ovule ou spermatozoïde, qui développent une nouvelle mutation lors de la division cellulaire. Cette mutation, maintenant acquise dans les cellules de l'individu, se multipliera lors de la division cellulaire et il deviendra plus propice de la transmettre à la descendance. Tandis que le mosaïcisme somatique correspond à une mutation qui se développe de façon postzygotique, qui peut survenir à n'importe quel stade la vie et ce dans n'importe quel tissu. Cette mutation est spontanée et survient également lors de la division cellulaire [16]. Ce qui complexifie les recherches, c'est qu'il existe une différence du nombre de répétitions entre les différents tissus, mais également entre les différentes cellules d'un même tissu [16]. La

majorité des études se focalisent sur les différentes sources retrouvées dans le tissu sanguin, puisqu'il est facilement accessible.

Le mosaïsme entraîne des difficultés dans les recherches sur la DM1 pour obtenir une mesure précise du nombre de répétitions chez un individu étant donné la grande variabilité dans les tissus [41, 42]. Les symptômes peuvent varier selon le nombre de répétitions dans les différentes cellules [28, 32]. Donc si les répétitions sont plus élevées dans les muscles, les symptômes seront plus au niveau musculaire et si les répétitions sont plus élevées dans les cellules de la matière blanche, les symptômes seront plus au niveau cognitif et ainsi de suite [32].

Le sang est le tissu avec lequel il est plus facile de travailler puisque cela demande moins d'implication de la part du patient. Cependant la longueur des répétitions CTG est la plupart du temps inférieure dans le sang que dans les muscles [28, 32]. Lopez Castek et al. ont démontré le profil du nombre de répétitions CTG dans différents tissus, et ce chez le fœtus et chez l'adulte atteint de DM1 [42]. Ils ont montré que le nombre le plus élevé de répétitions chez un adulte atteint de DM1 se trouvait dans le cœur et que le nombre le plus bas était dans le cervelet. Ils ont également remarqué que chez le fœtus le nombre de répétitions était plus élevé dans le cœur et le plus bas dans le foie [42]. Le nombre de répétitions CTG peut également varier au cours de la vie d'une personne entraînant par le fait même une augmentation de la gravité des symptômes [32]. Plusieurs phénomènes sont encore étudiés aujourd'hui en DM1 car certaines fonctions des gènes impliqués dans la pathologie ne sont pas encore bien connues.

### **1.7 GÈNES IMPLIQUÉS DANS LA DM1**

Il y aurait respectivement deux gènes impliqués dans la maladie, le gène *Dystrophia myotonica protein kinase* (DMPK) et son promoteur, le gène *SIX homeobox 5*. Tout d'abord, le gène DMPK est responsable de la fabrication d'une protéine appelée protéine kinase de la dystrophie myotonique. Cette protéine porte le nom de la maladie qu'elle cause à la suite de son insuffisance. Plusieurs études ont été faites afin de démontrer la relation entre le gène DMPK et la DM1. Il a d'abord été démontré

que la protéine DMPK a une fonction inhibitrice pour la protéine appelée myosine phosphatase [43]. Cette dernière jouerait un rôle majeur pour la contraction et la relaxation musculaire. [41]. Ces études ont permis de mieux comprendre la protéine DMPK et de constater qu'elle était moins abondante dans plusieurs tissus, notamment le muscle squelettique et cardiaque [41].

Il a été démontré que suite à un épissage alternatif, plusieurs isoformes de la protéine DMPK pourraient être présentes dans la cellule [44]. Son isoforme la plus abondante, d'une longueur de 80kb est principalement exprimée dans le muscle lisse, squelettique et cardiaque [45]. Un déséquilibre dans les isoformes pourrait entraîner des variations dans les symptômes et amener un trouble multisystémique sévère [46]. Des chercheurs ont travaillé sur l'isoforme A de la DMPK et ont découvert que celui-ci protégerait les cellules de la mort cellulaire en empêchant le stress oxydatif. L'isoforme A étant moins fréquent chez les personnes atteintes, n'est pas présent pour inhiber la mort cellulaire, et entraînerait ainsi les symptômes de la maladie. [47].

Par la suite, le deuxième gène connu dans la génétique de la maladie est le gène *SIX homeobox 5*, codant pour une protéine du nom de SIX5. Ce gène est, tout comme le gène DMPK, situé sur le chromosome 19 (19q13.32) et est composé de 3 exons. Situé en aval du gène DMPK, c'est principalement le promoteur du gène *SIX homeobox 5* qui est touché par la répétitions CTG. Le promoteur est la séquence qui initie la transcription, il est donc primordial pour l'expression d'un gène [48]. L'expansion CTG entraînerait donc une diminution de l'expression du gène *SIX homeobox 5* [49]. Ce gène code pour un facteur de transcription. Il est important de mentionner que ce gène n'est pas une cause de la maladie, mais un facteur qui peut amener des fluctuations des symptômes chez les patients. Kirby et al. (2001) ont démontré à l'aide de drosophiles que le gène SIX5 aurait un impact sur le développement des muscles et des gonades [50]. Chez l'humain, si l'homologue D-six4 est défectueux, ceci peut engendrer une réduction de la masse testiculaire [50]. Il est possible qu'une faible expression du gène SIX5 puisse expliquer les faiblesses musculaires ainsi que les problèmes d'hypogonadismes qui sont présents chez les patients atteints de DM1.

Des études faites sur des souris *knock-out* ont comparé la fonction des deux gènes. Des souris pour lesquelles le gène DMPK a été inactivé ont exprimé de la myopathie et des malformations cardiaques [51, 52]. Tandis que des souris invalidées pour le gène SIX5 ont développé des problèmes d’infertilité et des cataractes [45, 50, 53]. Le processus cellulaire de la DM1 demeure un processus complexe qui est encore grandement étudié aujourd’hui.

### 1.8 PRÉVALENCE DE LA DM1

La DM1 est la forme de dystrophie musculaire la plus fréquente chez l’adulte [21]. Mondialement, elle est considérée comme une maladie rare, car elle touche une personne sur 12,000. Cependant, sa fréquence varie selon les régions. Au Saguenay-Lac-Saint-Jean elle atteint un sommet avec 1 personne sur 632 atteintes de la maladie (voir tableau 1). La prévalence de la maladie y est donc dix à trente fois plus élevée que dans les autres populations.

**Tableau 1 – Prévalence de la DM1 dans différents lieux**

Auteur	Lieu	Fréquence par 100 000 habitants
Chung et al. 2003 [54]	Chine	0,37
Hsiao et al. 2006 [55]	Taiwan	0,46
Darin et Tulinius, 2000 [56]	Suède	5
Mladenovic et al. 2006 [57]	Serbie	5,3
Mah et al. 2016 [58]	Mondiale	8,26
Nakagawa et al. 1991 [59]	Japon	9,13
Siciliano et al. 2001 [60]	Italie	9,31
Norwood et al. 2009 [61]	Angleterre	10,6
Ford et al. 2006 [62]	Nouvelle Zélande	11,6
Magee et Nevin. 1999 [63]	Irlande du Nord	11,95

Segel et al. 2003 [64]	Israël	15,7
Medica et al. 1997 [65]	Croatie	18,1
Lopez de Munain et al. 1993 [66]	Espagne	26,5
Mathieu et Prevost 2012 [67]	Saguenay-Lac-Saint-Jean	158

Plusieurs s'étonnent de voir une prévalence toujours aussi élevée de cette maladie au Saguenay-Lac-Saint-Jean. Cette maladie monogénique devrait être vouée à s'éliminer par elle-même pour plusieurs raisons. Tout d'abord, c'est une maladie qui amène une diminution de la fécondité chez l'homme et une baisse de la valeur sélective générale [68]. Il existe un phénomène d'anticipation augmentant la longueur des répétitions au fil des générations, ce qui devrait conduire à une diminution de la prévalence de la maladie avec le temps [68]. De plus, lorsque c'est la mère qui est atteinte, le phénomène d'anticipation est omniprésent et sa descendance décèdera plus souvent en bas âge. Donc plusieurs chercheurs se questionnent à savoir pourquoi cette maladie ne va pas en diminuant, mais reste toujours aussi élevée dans cette population [67, 68]. Par ailleurs, au Saguenay-Lac-Saint-Jean, la maladie est étudiée depuis de nombreuses années et la cause de la forte prévalence de la DM1 serait un effet fondateur [69, 70].

### **1.9 EFFET FONDATEUR**

Un effet fondateur est la formation d'une nouvelle population sur un nouveau territoire à partir d'un groupe restreint d'individus quittant une population d'origine [12]. Lorsque la nouvelle population se forme, cela crée un goulot d'étranglement, diminuant ainsi le bagage allélique qui formera la prochaine génération. Le goulot d'étranglement est une diminution de la taille d'une population, et ce sur plusieurs générations consécutives [71]. À la suite de ce goulot d'étranglement, il y a régulièrement un isolement de la population, créant une distance envers les autres populations, et ce sur plusieurs générations. En plus de l'isolement, il peut y avoir également une forte natalité. Tous ces phénomènes peuvent entraîner une variation des fréquences alléliques dans la nouvelle

population [12]. Pour ce qui est des variants rares, l'effet fondateur pourrait entraîner une augmentation de la fréquence de certains variants dans la population [72].

### **1.9.1 EFFET FONDATEUR AU SAGUENAY-LAC-SAINT-JEAN**

Dans la littérature actuelle, il a été possible de caractériser trois mouvements de population qui ont eu lieu à différentes périodes et qui auraient contribué à l'effet fondateur du Saguenay-Lac-Saint-Jean. Le premier mouvement migratoire a lieu au début du 17<sup>e</sup> siècle alors qu'une petite partie de la population de la France est venue s'installer sur le territoire de la vallée du Saint-Laurent [12, 73]. Environ 10,000 immigrants ont quitté la France et sont pour la plupart, les ancêtres des Canadiens français du Québec [14].

Le deuxième mouvement migratoire s'est fait à la fin du 17<sup>e</sup> siècle lorsque des habitants de la vallée du Saint-Laurent sont allés s'installer dans la région de Charlevoix [12, 73]. Environ 600 immigrants s'installent dans cette région.

C'est vers 1838 que le troisième mouvement de population a lieu colonisant la région du Saguenay-Lac-Saint-Jean. Certains des habitants de Charlevoix étaient à la recherche de nouvelles terres, ils se sont donc déplacés vers le Saguenay-Lac-Saint-Jean [12]. Entre 1838 et 1911, environ 30 000 individus ont migré vers la région du Saguenay-Lac-Saint-Jean, 70 % étaient originaire de Charlevoix et 30 % d'autres régions comme la vallée du Saint-Laurent. Au début des années 1900, ceux-ci représentent pour la plupart les fondateurs de cette région [14, 74].

De plus, à cette époque, deux autres aspects ont augmenté l'intensité de l'effet fondateur au Saguenay-Lac-Saint-Jean. Tout d'abord, il y avait une forte natalité, les gens avaient plusieurs enfants en raison de la forte présence de l'église [12]. Le taux de natalité d'autrefois était très élevé et les femmes mettaient au monde plus de 10 enfants [75, 76]. Pendant que le Québec a multiplié par 5 sa population, le Saguenay l'a multipliée par 25 [77].

Le second aspect qui a augmenté l'intensité de l'effet fondateur est que les habitants du Saguenay-Lac-Saint-Jean avaient des déplacements restreints entre les autres régions. Le Saguenay

est une région éloignée, ses voies de transports mal desservies et sa distance par rapport aux grandes villes ont favorisé son isolement. Tous ces aspects ont favorisé la hausse en fréquence de certaines maladies rares dans cette région.[78, 79]

La DM1 n'est pas la seule maladie qui a une fréquence élevée dans la région du Saguenay-Lac-Saint-Jean. En effet, plusieurs maladies dominantes et récessives sont actuellement répertoriées [12, 14]. Il est vrai que plusieurs maladies sont plus fréquentes dans la région du Saguenay-Lac-Saint-Jean, mais il existe également certaines maladies qui sont pratiquement absentes dans la région, mais qui sont fréquentes ailleurs dans le monde. Par exemple, l'hémophilie est une maladie fréquente en France, mais cette maladie est plus rare au Québec [12].

Dans la région du Saguenay-Lac-Saint-Jean, il existe de fausses croyances quant à la consanguinité proche de la population. D'ailleurs, une étude a démontré que la région du Saguenay n'est pas la région du Québec qui a le plus haut taux de consanguinité [69, 80]. Cette étude utilise des bases de données de généalogie allant jusqu'aux années 1600 et compare 26 régions du Québec. La région du Saguenay-Lac-Saint-Jean arrive au 19<sup>e</sup> rang sur les 26 sous-régions du Québec. Cette faible consanguinité proche n'était pas ce qui était attendu dû à l'effet fondateur qu'a connu la région. La cause de ce niveau faible de consanguinité n'est pas encore bien connue, mais il serait possible que ce soit dû à la forte expansion de population ou aux différents processus de colonisation [80].

Pour ce qui est de la transmission de certaines maladies dominantes, celles-ci suite à l'effet fondateur, se sont retrouvées en fréquence plus élevée dans la région. L'augmentation de la fréquence allélique de la répétition CTG causant la DM1 dans la région de Charlevoix et du Saguenay Lac-Saint-Jean n'est pas encore bien comprise. Certains auteurs croient qu'une prémutation s'est introduite dans la région et qu'à la suite d'un effet fondateur et d'un phénomène d'anticipation, la maladie aurait augmenté en fréquence [21, 81, 82].

Des études ont essayé de retracer les ancêtres responsables de la DM1 au Saguenay-Lac-Saint-Jean. Il semblerait que deux haplotypes seraient à l'origine du gène muté de la DM1 au Saguenay-Lac-Saint-Jean. Ces deux haplotypes contribueraient de 70 à 80 % des chromosomes affectés dans la



région [21, 22, 83-86]. Ce serait deux ancêtres qui auraient introduit la maladie dans Charlevoix, puis au Saguenay-Lac-Saint-Jean [21].

Les populations se différencient par des variants génétiques qui peuvent résulter de plusieurs facteurs. Au Québec, l'accès à des données généalogiques très complètes permet de faire des analyses sur la génétique des populations du Québec. Le Saguenay-Lac-Saint-Jean a été grandement étudié sous un aspect généalogique pour mieux comprendre la démographie et ses conséquences sur la génétique de la population.

### **1.10 LES GÉNÉALOGIES**

Les généalogies sont construites à l'aide des liens familiaux existants entre un individu et l'ensemble de ses ancêtres [87]. Les généalogies peuvent offrir une grande puissance à plusieurs domaines de recherche. Elles sont très peu utilisées dans les recherches génétiques puisqu'il existe très peu de banques de données regroupant toutes les informations. Avoir la généalogie complète sur plusieurs générations permet de faire des liens entre chacun des individus de la cohorte. Au Saguenay-Lac-Saint-Jean, les chercheurs peuvent avoir l'opportunité de travailler avec la généalogie de leur cohorte grâce au fichier de population BALSAC.

### **1.11 LE FICHER DE POPULATION BALSAC**

Le fichier BALSAC est une infrastructure de recherche initiée et dirigée par l'Université du Québec à Chicoutimi [88]. Il est un fichier de reconstitution généalogique créé à partir des données provenant des actes de l'état civil du Québec [88]. Ces actes sont en majorité des contrats de mariage, des actes de naissance et de décès. Dans tous ces documents, il y a des dates, des noms et des lieux qui permettent au chercheur de faire des liens et de relier les individus entre eux [89]. Les données couvrent la population du Québec depuis le début du 17<sup>e</sup> siècle jusqu'à la période actuelle [89]. Toutes les données ont été informatisées et interconnectées pour créer de grandes lignées généalogiques. C'est à l'aide de ces informations qu'il est possible de retracer la généalogie des participants [88]. Ce projet amène des possibilités de recherche permettant de mettre en lumière les caractéristiques d'une

population en retraçant la trajectoire individuelle ainsi que l'histoire généalogique et familiale de l'ensemble d'une population. Ce fichier a été utilisé dans de nombreuses études multidisciplinaires et est encore utilisé aujourd'hui dans de nombreuses recherches autant en sciences sociales, en sciences de la santé, en sciences humaines ou en sciences médicales [80, 90-94]. Grâce au fichier BALSAC, il a été possible d'étudier l'histoire démographique des populations du Québec [70, 80, 92]. C'est en partie grâce au fichier de population BALSAC que l'effet fondateur qui est observé au Québec est bien documenté.

### **1.11.1 LA COMPLÉTUDE**

Cette mesure équivaut au nombre d'ancêtres présents dans une généalogie par rapport au nombre d'ancêtres attendus à chaque génération [95]. Par exemple, pour la troisième génération d'un individu, le nombre d'ancêtres attendus est de 4, équivalent à ses 2 grands-parents du côté paternel et ses 2 grands-parents du côté maternel. Si on retrace seulement 3 des 4 grands-parents, la complétude sera de 75 % à la 3<sup>e</sup> génération pour cet individu. Lorsqu'il y a un grand nombre d'ancêtres qui sont absents dès les premières générations, cela influe grandement sur la complétude totale de la généalogie. Donc, la complétude est une bonne façon de visualiser la qualité des généalogies.

### **1.11.2 LE COEFFICIENT D'APPARENTEMENT**

Le coefficient d'apparement est un concept en généalogie qui se calcule par la probabilité que deux individus partagent le même allèle provenant d'un ancêtre commun [78]. Selon Prost et al. (2022), le coefficient d'apparement entre deux personnes A et B se définit « [...] comme la probabilité pour qu'un gène pris au hasard chez A et un autre pris dans les mêmes conditions chez B, soient identiques par descendance » [96].

Il ne faut pas se méprendre avec le terme parenté qui se traduit par le lien de parenté entre deux individus. Le coefficient d'apparement est une probabilité de partage de gène entre deux individus. Avec le coefficient d'apparement, l'individu A a 25 % de probabilité de partager un allèle commun

avec sa mère (Figure 4). Plus le coefficient d'apparement est élevé, plus les individus ont de probabilités de partager des allèles communs par descendance [77].

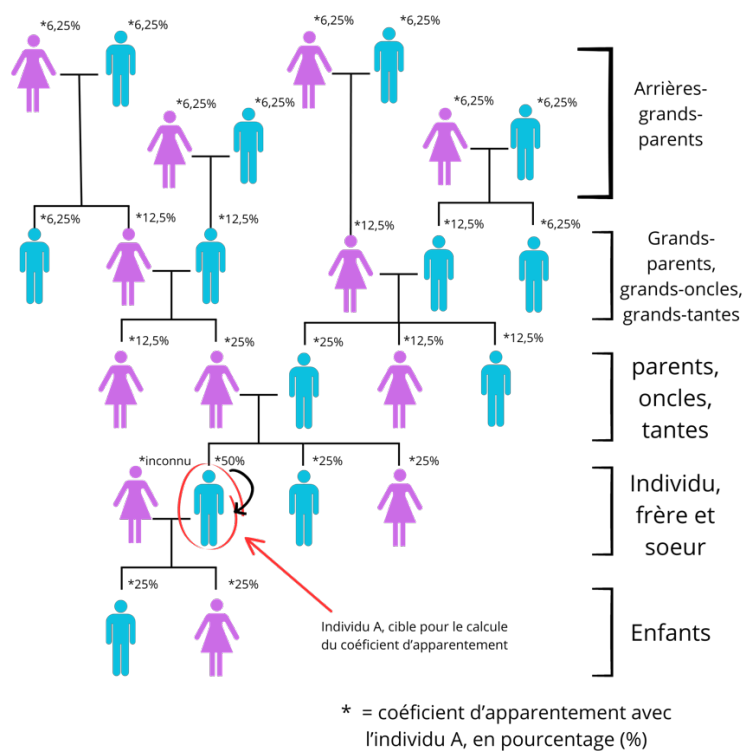


Figure 4 – Aperçu du coefficient d'apparement selon un arbre généalogique. Le coefficient d'apparement est calculé par paires pour l'individu A.

## **1.12 PROBLÉMATIQUE**

Comme mentionné, il a été démontré que la DM1 est causée par une expansion anormale d'un triplet de nucléotides CTG dans le gène DMPK, situé sur le chromosome 19 [5, 97]. En général, le nombre de répétitions CTG peut être un déterminant de la sévérité de la maladie. Cependant, il y a un manque de connaissances flagrant dans la prédiction de l'évolution des divers signes et symptômes associés à la DM1. Il y a une grande diversité des symptômes interindividuelle à l'intérieur d'un même phénotype, il est donc difficile de prédire de façon précise l'atteinte du patient avec seulement la longueur des répétitions CTG.

### **1.12.1 BUT ET HYPOTHÈSES**

L'étude avait comme but de dresser un portrait généalogique des patients DM1 de la CMNJ et voir si des corrélations étaient présentes entre les regroupements généalogiques avec un fort apparentement et différents phénotypes de la maladie.

1. Les regroupements généalogiques avec un fort apparentement auront des phénotypes similaires;
2. Il existe une corrélation entre le coefficient d'apparentement et les longueurs de répétition CTG;

### **1.12.2 OBJECTIFS**

Objectif 1: Dresser un portrait de la longueur des répétitions CTG en fonction de différents symptômes.

Objectif 2: Intégration et reconstruction généalogique de l'ensemble des patients de l'étude longitudinale sur la DM1 à l'aide du fichier de population BALSAC.

Objectif 3: Créer des regroupements généalogiques basés sur le coefficient d'apparentement pour lesquels les symptômes concordent.

## **CHAPITRE 2 MÉTHODOLOGIE**

### **2.1 POPULATION À L'ÉTUDE**

Ce projet s'appuie sur une étude longitudinale ayant débuté en 2001 à laquelle 200 participants atteints de différents phénotypes de DM1 ont participé [98]. La cohorte est composée de patients originaires du Saguenay-Lac-Saint-Jean, une population bien caractérisée généalogiquement et génétiquement [14, 68, 69, 99]. Les patients sont suivis à la CMNMJ depuis plusieurs années et une base de données a été construite en se basant sur plusieurs facteurs et variables de chaque patient.

#### **2.1.1 CRITÈRES D'INCLUSION**

Pour être admis dans cette étude, les patients devaient avoir un diagnostic de DM1 (juvénile adulte ou tardif) confirmé par analyse génétique et être âgés de 18 ans ou plus.

#### **2.1.2 CRITÈRES D'EXCLUSION**

Des patients avec le phénotype infantile ou congénital de DM1 ou avec d'autres conditions influençant la force des muscles ont été exclus de cette étude.

#### **2.1.3 CONSIDÉRATIONS ÉTHIQUES**

Il convient de mentionner que les participants ont donné leur consentement à l'utilisation secondaire des données recueillies dans le cadre de projets de recherche auxquels ils ont participé (données cliniques et moléculaires) de même qu'à la consultation de leur dossier médical (pour la collecte de données généalogiques). Seuls les renseignements nécessaires pour répondre aux objectifs scientifiques de ce projet ont été extraits. Tous les renseignements extraits sont demeurés strictement confidentiels dans les limites prévues par la loi. Afin de préserver l'identité des participants et la confidentialité des renseignements, ceux-ci n'ont été identifiés que par un numéro

de code. La clé du code reliant le nom du participant à son dossier de recherche a été conservée dans l'ordinateur protégé par un mot de passe d'un chercheur de l'équipe.

De même, la clé du code reliant le nom du participant à sa généalogie a été conservée dans l'ordinateur protégé par un mot de passe du professionnel de recherche de BALSAC. À la suite de la reconstruction des généalogies, l'ensemble des données a été codé. Seuls les membres de l'équipe de recherche ont eu accès aux données afin de réaliser l'étude en lien avec les objectifs scientifiques du projet. Également, dans le cadre de cette étude, les membres de l'équipe de recherche ne présentaient aucun conflit d'intérêts. Le projet a été approuvé par un comité éthique du Centre intégré universitaire de santé et de services sociaux (CIUSS) du Chicoutimi et a été reconnu par un comité éthique de l'Université du Québec à Chicoutimi (UQAC).

## **2.2 VARIABLES DE L'ÉTUDE**

### **2.2.1 VARIABLES CLINIQUES**

Les patients atteints de DM1 de la CMNMJ sont suivis depuis de nombreuses années, ce qui a permis d'amasser une large base de données cliniques pour mesurer la force ou la stabilité des patients [17, 67, 98-100]. Ces données cliniques comprennent 45 outils de mesures qui évaluent les atteintes principales en DM1 incluant les performances motrices et cognitives, les caractéristiques démographiques et des variables moléculaires. La base de données a été collectés lors d'études antérieures et le projet se base sur les données qui ont été collectées en 2019 [98]. Dans cette étude, aucun échantillon test pour des individus de même âge mais non atteint n'a été collecté. Les résultats se concentre sur les patients qui ont ou non certains critères phénotypiques en fonction des variables observés.

Les variables cliniques qui ont été utilisées dans cette étude étaient la mesure de la force du dorsifléchisseur de la cheville, la force de préhension de la pince, le nombre de répétitions CTG et les différents phénotypes de sévérité [16, 98]. Les deux variables de forces ont été sélectionnées, car ce

sont les mesures des muscles distaux (main et cheville) qui sont le plus affectés par la maladie. De plus, ces deux mesures sont celles qui semblent avoir une plus forte corrélation avec le nombre de répétitions CTG [16]. Les CTG ont également été sélectionnés pour les analyses puisqu'elles sont un élément important dans la caractérisation des différents phénotypes. Un échantillon de 96 individus était disponible pour la force du dorsifléchisseur de la cheville et un échantillon de 190 individus était disponible pour la force de pincement et le nombre de CTG. C'est à l'aide du Microfet-2 que la force du dorsifléchisseur de la cheville a été mesurée. Par la suite, c'est à l'aide d'une jauge de pincement qu'il a été possible de mesurer la force de préhension de la pince des participants. Finalement, c'est à l'aide d'une SP-PCR (*Small Pool-Polymerase Chain Reaction*) et d'une hybridation par Southern blot que la longueur des répétitions a été mesurée. Les mesures des répétitions ont été prises par SP-PCR afin d'évaluer l'instabilité somatique. Grâce à cette méthode, il est possible d'estimer la longueur de l'allèle progéniteur lorsque le patient est né et d'estimer la longueur de l'allèle modale lors de la prise de l'échantillon [16, 25]. Lors de cette étude, c'est l'allèle progéniteur qui a été utilisé afin d'avoir la mesure à la naissance et non celle qui peut être variable dans le temps. Les différents phénotypes ont également été utilisés pour les analyses soit le phénotype tardif, adulte et juvénile. Pour plus ample renseignement sur la collecte de donnée et les différentes prises de mesure, se référer à Légaré et al. [98].

### **2.2.2 VARIABLES GÉNÉALOGIQUES**

L'identification familiale et la génération des patients, identifiés avec les dossiers de la CMNMJ, ont été utilisées comme autres covariables dans les analyses. Les familles ont été utilisées pour voir la diversité dans les regroupements du coefficient d'apparentement. La génération, c'est-à-dire à quel niveau générationnel de transmission le patient se situe par rapport à un ancêtre commun, a été utilisée afin de valider le phénomène d'anticipation dans la cohorte. Le niveau générationnel est représenté par un chiffre entre 1 et 4 : 1 étant l'individu le plus prêt de l'ancêtre commun et 4 celui le plus éloigné de ce même ancêtre. La génération 1 est donc la plus ancienne et la génération 4 est plus

près de la génération actuelle. Les autres données récoltées lors de l'étude n'ont pas été utilisées dans les analyses.

### **2.2.1 ÉTUDE CORRÉLATIONNELLE DES DONNÉES**

Tout d'abord, l'ensemble des analyses a été effectué à l'aide du logiciel R. Les premières analyses faites sur le logiciel ont été des études corrélacionnelles des données. C'est à l'aide d'une carte de chaleur qu'il a été possible de voir les relations entre les variables sélectionnées. Par la suite, des diagrammes à moustaches ont été réalisés selon les différents degrés de sévérité de la maladie et les 3 variables : la force de la pince, la force du dorsifléchisseur de la cheville et le nombre de répétitions CTG.

### **2.2.2 INTERVALLES DES RÉPÉTITIONS CTG**

Le nombre de répétitions CTG a également été utilisé pour faire des regroupements à l'aide d'intervalles. Ceci a été fait dans le but de caractériser le nombre de CTG en fonction des différentes variables et voir si la tendance avec les différents phénotypes concordait. Les nombres de répétitions CTG ont été divisés en 3 groupes soit [50-300], [300-800] et [800 et +]. La division s'est basée sur la classification de plusieurs études utilisant quatre ou cinq divisions phénotypiques [6, 13, 26, 29, 101]. Les phénotypes moins sévères sont caractérisés par moins de 300 répétitions, ceux de sévérité médiane par 300 à 800 répétitions, et les plus sévères par 800 répétitions ou plus. Des diagrammes à moustache ont été faits pour visualiser les regroupements en fonction des deux variables de force, soit celle de pincement et celle du dorsifléchisseur de la cheville.

Des tests de Shapiro-Wilk ont été réalisés pour les répétitions CTG, la force de pincement et la force du dorsifléchisseur de la cheville et les résultats indiquent que les trois variables ne suivent pas une distribution normale. Des tests non paramétriques de Kruskal-Wallis ont donc été effectués pour comparer les différents groupes. Après avoir fait les tests de Kruskal-Wallis, des tests de comparaisons multiples Wilcoxon ont été effectués afin de voir quels étaient les groupes différents.



Un chi carré a également été effectué pour voir la relation entre les phénotypes et la génération en fonction des différents groupes d'intervalles du nombre de répétitions CTG.

## **2.3 CRÉATION DU FICHER GÉNÉALOGIQUE**

### **2.3.1 COLLECTE DES DONNÉES GÉNÉALOGIQUES**

La récolte de données s'est faite à l'aide des dossiers des patients atteints de DM1 suivis à la CMNMJ. Les dossiers des patients sont classés par famille et chacune des familles a un arbre généalogique qui a été dessiné à la main par des conseillers génétiques [99]. Les arbres généalogiques ont été utilisés pour récolter les données familiales des patients et comptaient entre 3 et 4 générations partiellement complètes. Chaque arbre a été ouvert et analysé par deux étudiantes qui validaient chacune des données transcrites. Il a été possible de récolter des noms, des années de naissance, des années de décès, des années de mariage, etc. Toutes ces données ont été inscrites dans un tableau pour attribuer un numéro à chaque famille et à chaque patient avec un code. Chaque membre de la famille du patient avait un numéro et devenait un nouvel individu pour notre base de données. Au total, 1,437 individus étaient dans la liste finale et chaque individu avait un code unique et un code associé à sa famille. Au total, 56 familles ont été reconstruites à l'aide des arbres généalogiques.

### **2.3.2 INCORPORATION DE LA BASE DE DONNÉES DANS LE FICHER BALSAC**

La base de données contenant les informations familiales des patients a ensuite été envoyée à BALSAC pour la reconstruction des généalogies complètes. Grâce au fichier BALSAC, il a été possible de reconstruire les généalogies des patients étendues jusqu'à 18 générations. Une fois le corpus généalogique reconstruit par BALSAC, il a été possible d'exporter les données et d'avoir un fichier généalogique complet composé de tous les ancêtres des patients.

### **2.3.3 VALIDATION DES GÉNÉALOGIES À L'AIDE DE GENLIB**

À la suite de l'exportation des données codées de BALSAC sur des serveurs, l'information généalogique a été jointe la base de données de la CMNM. Cette base de données est celle qui contient

les données cliniques des patients qui ont été utilisées pour les analyses. Après toutes les validations, un total de 200 patients avec leur généalogie a été pris pour les analyses.

Les données généalogiques demandent des outils adaptés afin de pouvoir les utiliser dans les analyses. GENLIB est une librairie R qui a été conçue initialement pour travailler les données avec du fichier de population BALSAC et faciliter le travail avec les généalogies. La librairie a été implantée dans le logiciel R afin de la rendre libre d'accès. GENLIB permet d'obtenir des mesures généalogiques précises telles que la complétude, l'apparentement, la consanguinité, la profondeur généalogique moyenne, etc. Afin de vérifier si les généalogies avaient une bonne profondeur, la complétude a été calculée à l'aide de GENLIB.

#### **2.4 REGROUPEMENTS BASÉS SUR LE COEFFICIENT D'APPARENTEMENT**

GENLIB a été utilisé pour calculer le coefficient d'apparentement. Le logiciel compare la probabilité de partage d'allèle pour chaque paire d'individus en formant une matrice. Plusieurs matrices ont été créées, la première matrice prenait en compte l'ensemble des généalogies soit de 0 à 18 générations. Dans cette matrice, les patients étaient tous liés de loin ou de près. Par conséquent, d'autres matrices ont été créées en prenant chaque génération. Ceci a pu être fait en coupant la généalogie à plusieurs profondeurs. Lorsqu'une génération est isolée, toutes les générations inférieures sont prises en compte, mais cela exclut les générations supérieures. Par exemple, si une matrice du coefficient d'apparentement est faite pour la génération 3, seuls le patient, les parents et les grands-parents sont pris en compte dans le calcul du coefficient d'apparentement. De cette façon, les autres générations plus anciennes ne viennent pas influencer le calcul de l'apparentement.

La représentation graphique du coefficient d'apparentement a été faite à l'aide de la librairie 'igraph' sur le logiciel R. Le coefficient d'apparentement se calcule par paires d'individus, il est donc difficile de le représenter avec un graphique standard. Igraph forme un graphique en forme de réseau. Le réseau représente chaque individu par un point et les lignes représentent une distance, ici la distance est le coefficient d'apparentement. Donc, plus les points sont proches, plus le coefficient

d'apparement est grand entre les deux individus tandis que plus les points sont éloignés, plus le coefficient d'apparement est faible. S'il n'y a pas de ligne, le coefficient d'apparement est inexistant. Chaque ligne représente une relation entre les individus. Les igrph ont été faits jusqu'à la génération 6 puisque les regroupements ne démontraient aucune variation sur le plan des générations subséquentes. L'origine de la cohorte fait en sort que des ancêtres communs arrivent très tôt dans les généalogies créant par conséquent de nombreux liens entre les points. Les figures ont donc été faites pour les générations 3, 4, 5 et 6.

À la suite de la création des réseaux, des regroupements basés sur le coefficient d'apparement ont été formés selon la méthode de Louvain. Cette méthode se base sur l'optimisation de la modularité, une mesure permettant d'avoir des groupes avec la configuration la plus adéquate selon la dispersion des points des igrph [102]. Un nombre variable de regroupements basé sur le coefficient d'apparement (3, 5, 8, 10, 15, 20) a été formé afin d'évaluer leur influence sur les différentes variables cliniques.

Des diagrammes à moustache ont été créés pour visualiser les différentes variables cliniques, soit la force du dorsifléchisseur de la cheville, la force de préhension de la pince et le nombre de répétitions CTG en fonction des différents regroupements.

## CHAPITRE 3 RÉSULTATS

### 3.1 ÉTUDE CORRÉLATIONNELLE DES DONNÉES

Le premier résultat montre les corrélations entre les différentes variables utilisées pour les analyses (Figure 5). La carte de chaleur montre les corrélations positives en orangé et les corrélations négatives en bleu clair. Plus la couleur est foncée, plus la corrélation est puissante. Il est possible de voir une corrélation modérée entre la force du dorsifléchisseur de la cheville et la force de pincement (0,57). Ces deux mesures de force ont également une corrélation avec le nombre de répétitions CTG (-0,48 et -0,65 respectivement)). Finalement, la sévérité corrèle également avec l'âge du début de la maladie (-0,58), le nombre de CTG (0,6) ainsi que les deux variables de force (-0,49 et -0,61). L'ensemble de corrélations est conforme à ce qui était attendu, mais les corrélations ne sont pas très fortes.



Figure 5 – Corrélations entre les variables utilisées dans cette étude mesurées chez les patients DM1

Note. La force du dorsifléchisseur de la cheville est en % de la valeur prédite moyenne pour les côtés droit et gauche, Sévérité = sévérité de la DM1 (phénotypes tardif, adulte ou juvénile)

### 3.2 ANALYSE PHÉNOTYPIQUE

Par la suite, la relation entre les différents degrés de sévérité de la maladie avec le nombre de CTG montre que les groupes phénotypiques sont différents avec un *p-value* de  $9,677e-20$  au test de Kruskal Wallis (Figure 6). Le test de comparaison multiple de Wilcoxon apporte plus de précision sur la différence entre les groupes. Le nombre de CTG est largement inférieur pour le phénotype tardif comparativement aux deux autres phénotypes. Tandis qu'il n'y a pas de différence entre le phénotype juvénile et adulte.

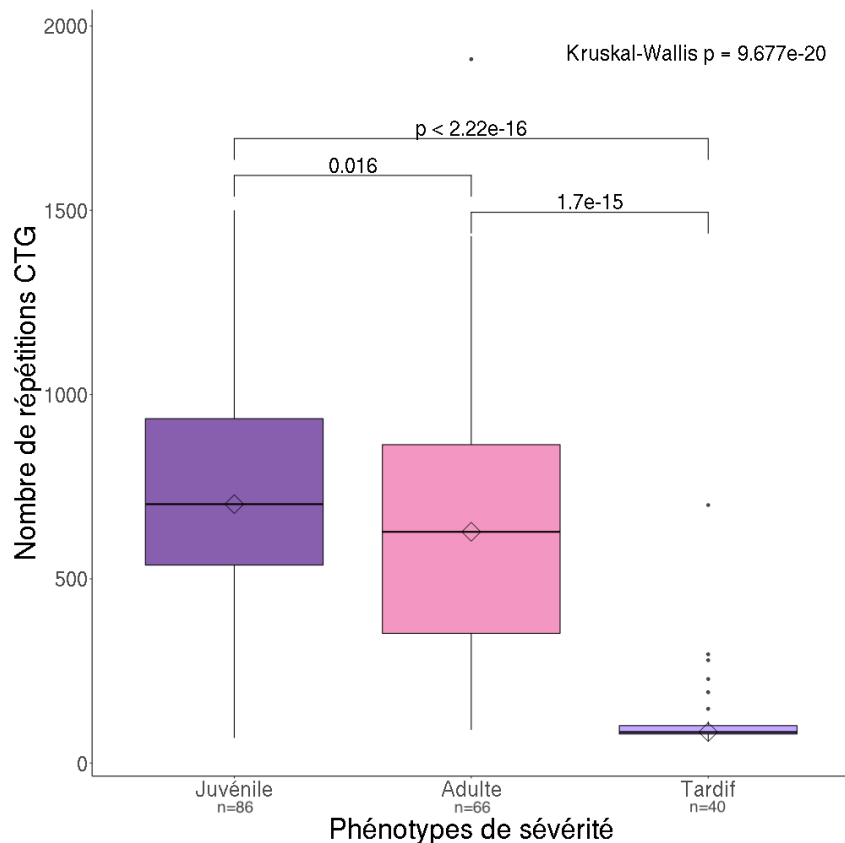


Figure 6 – Nombre de répétitions CTG en fonction des différents phénotypes de la maladie.

Par la suite, la relation des différents degrés de sévérité de la maladie avec la force de pincement montre que les groupes phénotypiques sont différents avec un *p-value* de  $3,657e-15$  au test de Kruskal Walli (Figure 7). Avec le test de comparaison multiple de Wilcoxon, il est possible de voir que le

groupe Tardif est différent des deux autres phénotypes avec une force de pincement largement plus élevé. Les groupes Juvénile et Adulte ne sont pas différents l'un de l'autre.

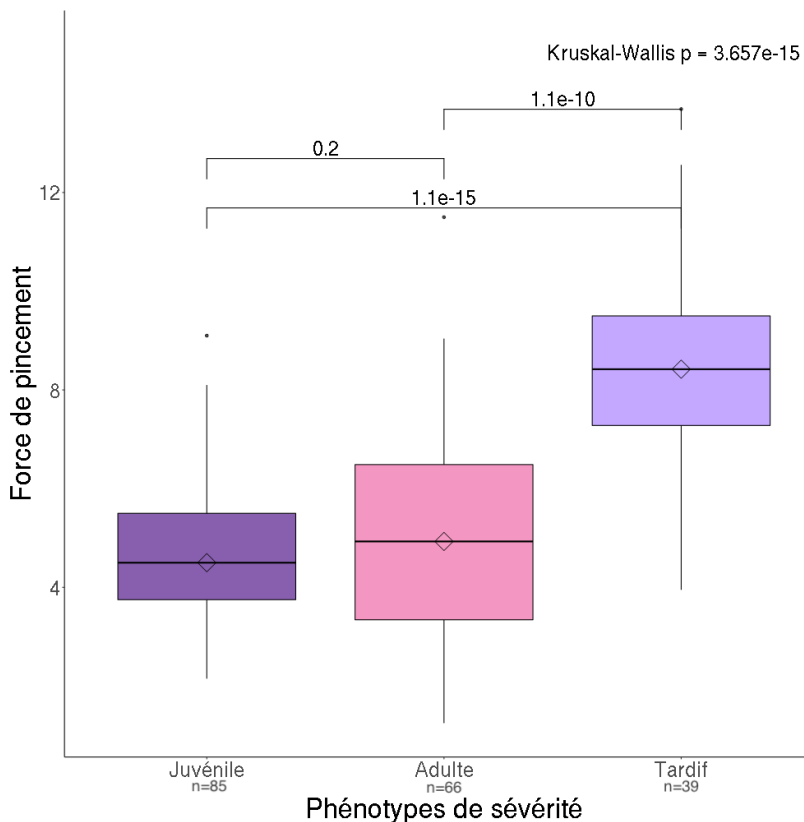


Figure 7 – La force de préhension de la pince en fonction des différents phénotypes de la maladie.

Finalement, la relation entre les différents degrés de sévérité de la maladie avec la force du dosifléchisseur de la cheville montre que les groupes phénotypiques sont également différents avec un *p-value* de  $1,38 \times 10^{-6}$  au test de Kruskal Wallis (Figure 8). Avec le test de comparaison multiple de Wilcoxon, il est possible de voir que le phénotype tardif est celui qui est le plus différent des 2 autres groupes avec une force du dorsifléchisseur de la cheville plus élevée. Les groupes Juvénile et Adulte ne sont pas différents l'un de l'autre. Ces deux phénotypes ne peuvent pas être distingués en utilisant le nombre de CTG, la force de pincement ou la force du dorsifléchisseur de la cheville. Pourtant ces deux phénotypes sont distincts dans la classification de la maladie.

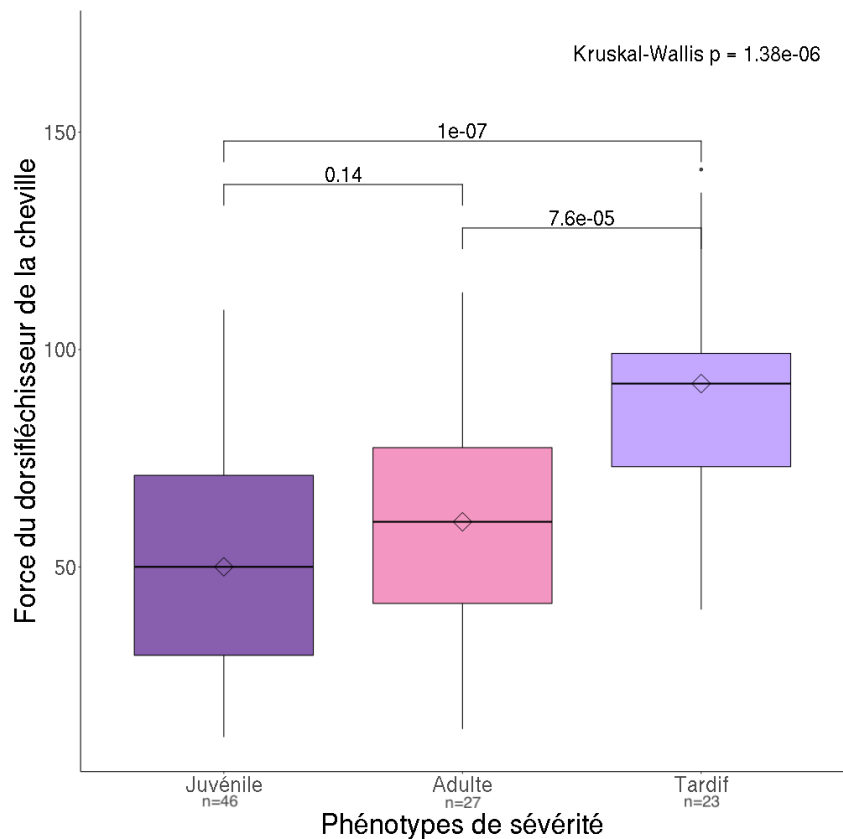


Figure 8 – Sévérité de la maladie en fonction de la mesure de force du dorsifléchisseur de la cheville.

### 3.3 ANALYSE DES REGROUPEMENTS PAR INTERVALLE DES RÉPÉTITIONS

#### CTG

La force de pincement et la force du dorsifléchisseur de la cheville en fonction du nombre de CTG sont illustrées aux figures 10 et 11. Les intervalles du nombre de CTG sont représentés sur l'axe des abscisses et les variables de force sont représentées sur l'axe des ordonnées.

Il est possible de voir que la majorité des patients ayant un phénotype adulte et juvénile sont dans les regroupements [300-800] et [800 et +] et que la majorité du phénotype tardif est dans le regroupement [50-300]. Les regroupements avec un nombre de CTG plus bas ont une force de pincement plus grande. Le test de Kruskal-Wallis indique que les groupes sont différents pour la force de pincement avec un *p-value* de 4,645e-18 (Figure 9). Lorsque chacun des groupes est comparé à

l'aide du test de comparaison multiple de Wilcoxon, il est possible de voir que les intervalles sont tous différents l'un de l'autre.

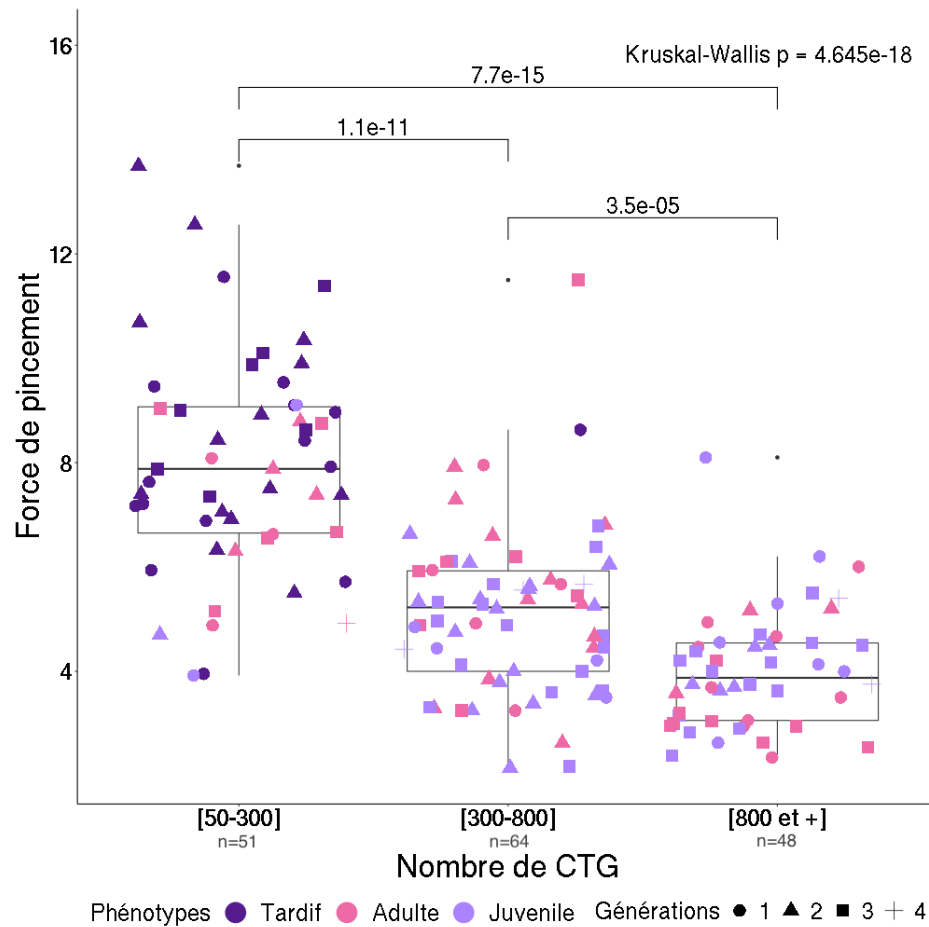


Figure 9 – La force de pincement selon 3 regroupements par intervalles du nombre de CTG.

Le test de Kruskal-Wallis indique également que les groupes sont différents pour la force du dorsifléchisseur de la cheville avec un p-value de  $2,415 \times 10^{-5}$  (Figure 10). Lorsque chacun des groupes est comparé à l'aide du test de comparaison multiple de Wilcoxon il est possible de voir que les regroupements par intervalle du nombre de CTG sont tous différents, excepté les intervalles [300-800] et [800 et +]. Finalement, un test de chi carré a été fait pour évaluer la relation des phénotypes et de la génération en fonction des intervalles du nombre de CTG. À titre de rappel, la génération est



le niveau générationnel de transmission du patient par rapport à un ancêtre commun, 1 étant le plus près de l'ancêtre commun et 4 le plus éloigné.

Celui-ci montre une différence significative pour les différentes générations et les phénotypes dans les groupes des différents intervalles du nombre de CTG (Tableau 2). Les résultats étaient également attendus, cependant le nombre de répétitions CTG n'explique pas les variables de force de manière individuelle puisqu'il est possible de voir de nombreux chevauchements au niveau des phénotypes entre les intervalles pour les phénotypes adulte et juvénile. Le phénotype tardif est bien concentré dans son intervalle.

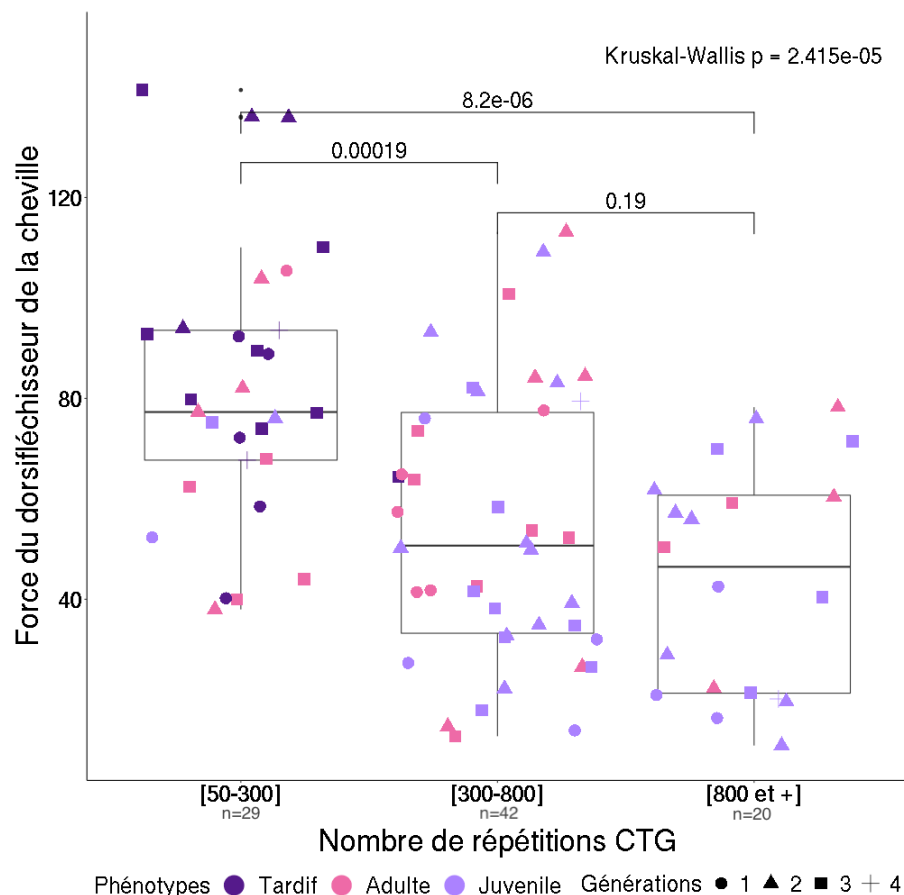


Figure 10 – Force du dorsifléchisseur de la cheville selon 3 regroupements par intervalles du nombre de CTG

Tableau 2 – Chi carré pour les phénotypes et la génération en fonction des intervalles du nombre de CTG

	X-squared	df	p-value
Phénotypes	98,654	4	< 2.2e-16
Génération	16,015	6	0.01368

### 3.4 VALIDATIONS DES GÉNÉALOGIES

Un graphique a été réalisé pour visualiser la complétude de l'ensemble de la cohorte (Figure 11). Il est possible de voir que le pourcentage d'ancêtres pour les patients atteints de DM1 est près de 100 % jusqu'à la 9<sup>e</sup> génération. Ce résultat démontre une belle qualité des généalogies pour la cohorte de DM1.

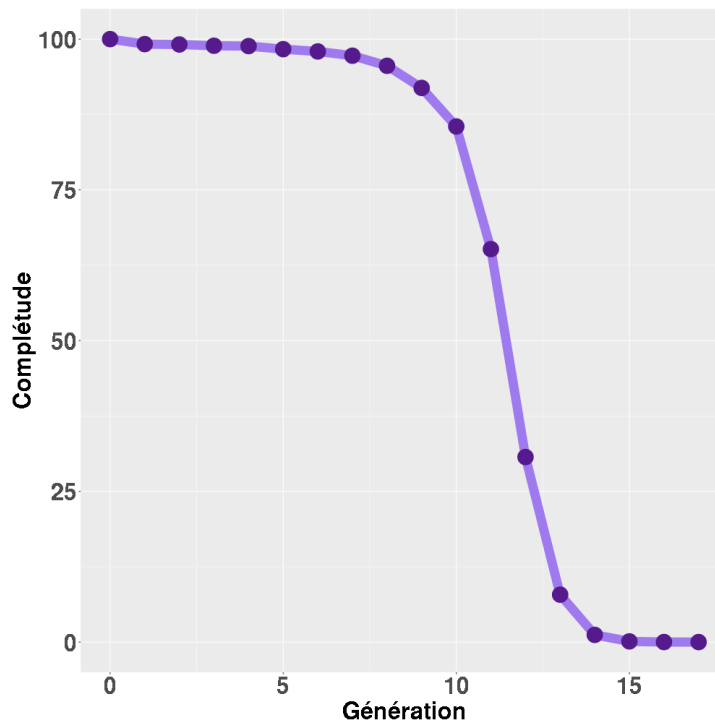


Figure 11 – Complétude moyenne de 200 patients atteints de DM1 originaires du Saguenay-Lac-Saint-Jean.

### **3.5 CONCEPTION DES RÉSEAUX À L'AIDE DU COEFFICIENT**

#### **D'APPARENTEMENT.**

Puisque le nombre de répétitions CTG n'explique pas toutes les variabilités individuelles des symptômes, le coefficient d'apparement a été utilisé pour voir s'il pouvait apporter plus d'explications.

Les réseaux faits à l'aide de la librairie igraph à une profondeur maximale de 3-4-5 et 6 générations ont permis de voir la répartition des différentes familles touchées par la DM1 selon le coefficient d'apparement (Figure 12). Chaque point est un individu, chaque couleur est un regroupement fait à l'aide du modèle de Louvain et chaque liaison correspond à un lien d'apparement avec une longueur équivalant au coefficient d'apparement. Il est possible de voir que les familles se relient de plus en plus lorsque l'on monte dans les générations. Cela est expliqué par l'origine de la cohorte provenant d'une jeune population; les ancêtres communs arrivent très tôt dans les généalogies créant par conséquent de nombreux liens entre les points.

Plus la profondeur généalogique augmente, plus les points sont rapprochés. À la génération 3 et 4, il est possible de voir encore certaines distances entre les points. La génération 3 représente les familles nucléaires (enfants, parents et grands-parents) donc les familles sont toutes séparées et n'ont pas de lien entre elles. Lorsque la profondeur généalogique augmente, les groupes deviennent plus près et les familles se regroupent entre elles. À partir de la génération 5 et 6, la distance entre les différentes familles devient très rapprochée puisque le nombre d'ancêtres communs rencontrés en montant dans les généalogies augmente.

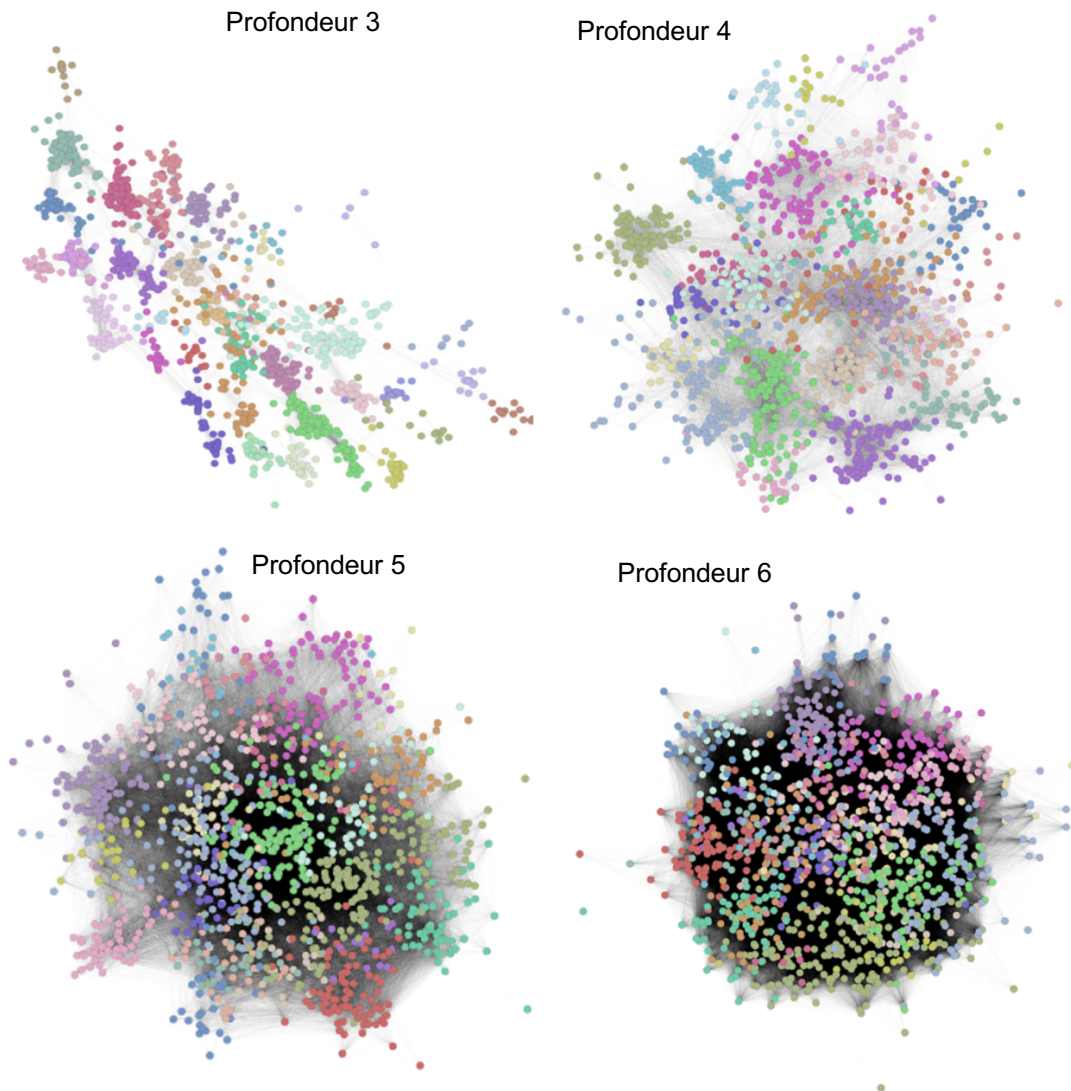


Figure 12 – Réseaux basés sur le coefficient d'apparentement.

Note. Les couleurs représentent les familles et la longueur des liens est le degré d'apparentement. Le premier graphique (en haut à gauche) est la profondeur de 3 générations, le deuxième graphique (en haut à droite) la profondeur de 4 générations, le troisième graphique (en bas à gauche) la profondeur de 5 générations et le dernier graphique (en bas à droite) la profondeur de 6 générations.

### 3.6 ANALYSE DES REGROUPEMENTS PAR COEFFICIENT

#### D'APPARENTEMENT

Par la suite, le nombre de CTG, la force de pincement et la force du dorsifléchisseur de la cheville ont été visualisés sous forme de diagramme à moustache en fonction de différents nombres de regroupements du coefficient d'apparentement. Plusieurs graphiques ont été créés pour chaque

variable clinique selon différents nombres de regroupements du coefficient d'apparement disponible (Annexe 1).

Les sous-divisions de regroupement du coefficient d'apparement sont disposées sur l'axe des x et différentes variables cliniques sont sur l'axe des y. Les différentes couleurs des points sur les graphiques sont les 54 familles. Les regroupements ont été mis en ordre croissant selon l'axe des y pour une meilleure visualisation.

### **3.6.1 RÉPARTITION EN 8 SOUS-REGROUPEMENTS DU COEFFICIENT**

#### **D'APPAREMENT**

C'est la répartition en huit sous-divisions du coefficient d'apparement qui a été sélectionnée pour plus amples explications de cette section. Cette sous-division était celle qui représentait le mieux l'ensemble des résultats. Il est possible de voir une distribution de chacun des phénotypes selon les huit différentes sous-divisions (voir tableau 3).

La division en 8 regroupements du coefficient d'apparement montre un résultat significatif de force faible au test de Kruskal Wallis (Figure 13). Le test de comparaison multiple de Wilcoxon montre que les regroupements du coefficient d'apparement 7 et 5 sont différents. Ceci voudrait dire que le nombre de répétitions CTG pourrait être expliqué par certains regroupements. Cependant, ce résultat significatif entre les regroupements du coefficient d'apparement 7 et 5 n'est pas suffisant pour tirer des conclusions. Le test ne trouve pas de différence significative pour les autres sous-divisions.

**Tableau 3 – Dénombrement des 3 phénotypes de DM1 en fonction dans les 8 sous regroupements du coefficient d'apparement selon les trois variables utilisées.**

<b>Variables</b>	<b>Regroupements</b>	<b>Phénotype juvénile</b>	<b>Phénotype adulte</b>	<b>Phénotype tardif</b>
<b>Nombre de répétitions CTG</b>	<b>1</b>	1	4	0
	<b>2</b>	5	2	3

	<b>3</b>	15	12	8
	<b>4</b>	7	9	5
	<b>5</b>	10	7	14
	<b>6</b>	17	7	5
	<b>7</b>	18	12	5
	<b>8</b>	1	4	0
<b>Force de pincement</b>	<b>1</b>	1	3	0
	<b>2</b>	5	2	3
	<b>3</b>	15	12	8
	<b>4</b>	7	9	5
	<b>5</b>	10	7	13
	<b>6</b>	17	7	5
	<b>7</b>	19	13	4
	<b>8</b>	2	4	0
<b>Force du dorsifléchisseur de la cheville</b>	<b>1</b>	2	0	1
	<b>2</b>	0	0	1
	<b>3</b>	8	7	2
	<b>4</b>	8	2	2
	<b>5</b>	11	7	4
	<b>6</b>	6	7	4
	<b>7</b>	8	7	3
	<b>8</b>	1	1	1

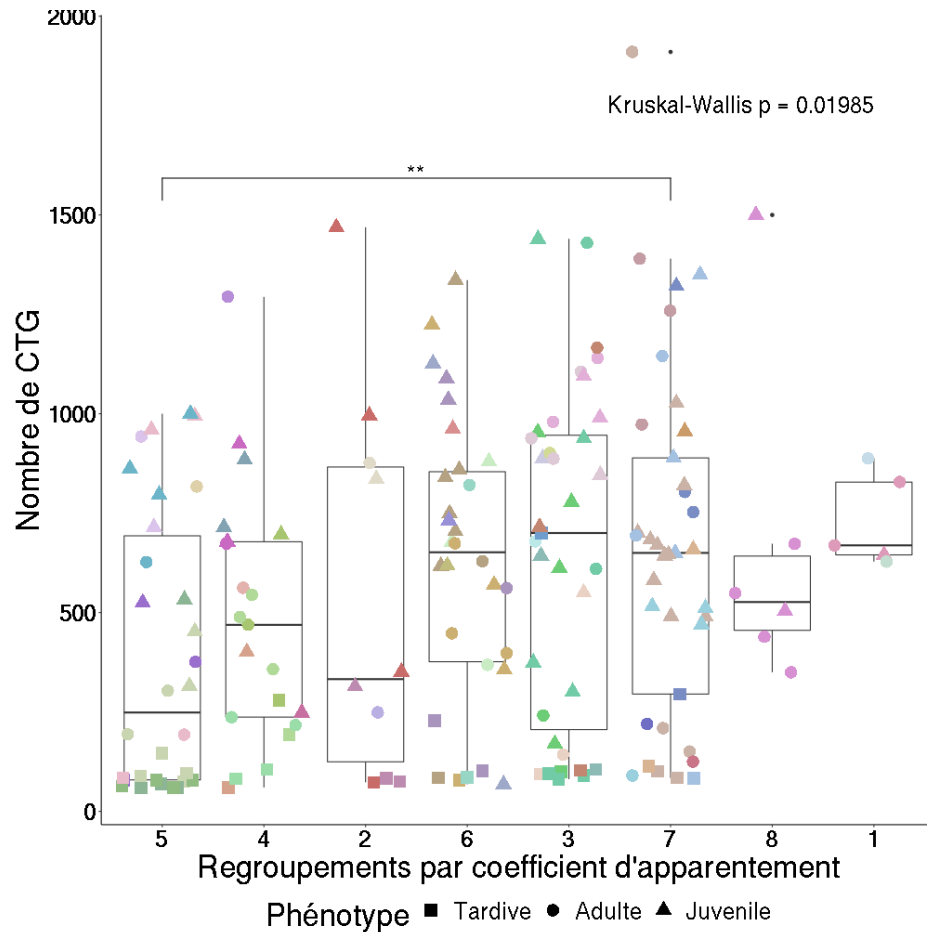


Figure 13 – Le nombre de répétitions CTG en fonction de 8 sous-regroupements du coefficient d'appurtenance.

Note. Les couleurs représentent les différentes familles.

Également, la force de pincement montre une valeur significative de force faible au test de Kruskal Wallis pour la division en huit regroupements du coefficient d'appurtenance (Figure 14). Le test de comparaison multiple de Wilcoxon a été en mesure de détecter une différence entre les regroupements 5 et 3, 5 et 6 ainsi que 5 et 7. Ceci porte à croire que la force de pincement pourrait être expliquée par certains regroupements du coefficient d'appurtenance. Le coefficient d'appurtenance pourrait donc amener certaines réponses à la variation des symptômes. Cependant, les résultats ne sont pas suffisants pour tirer de conclusion. Le test n'a pas été en mesure de trouver une différence significative entre les autres divisions.

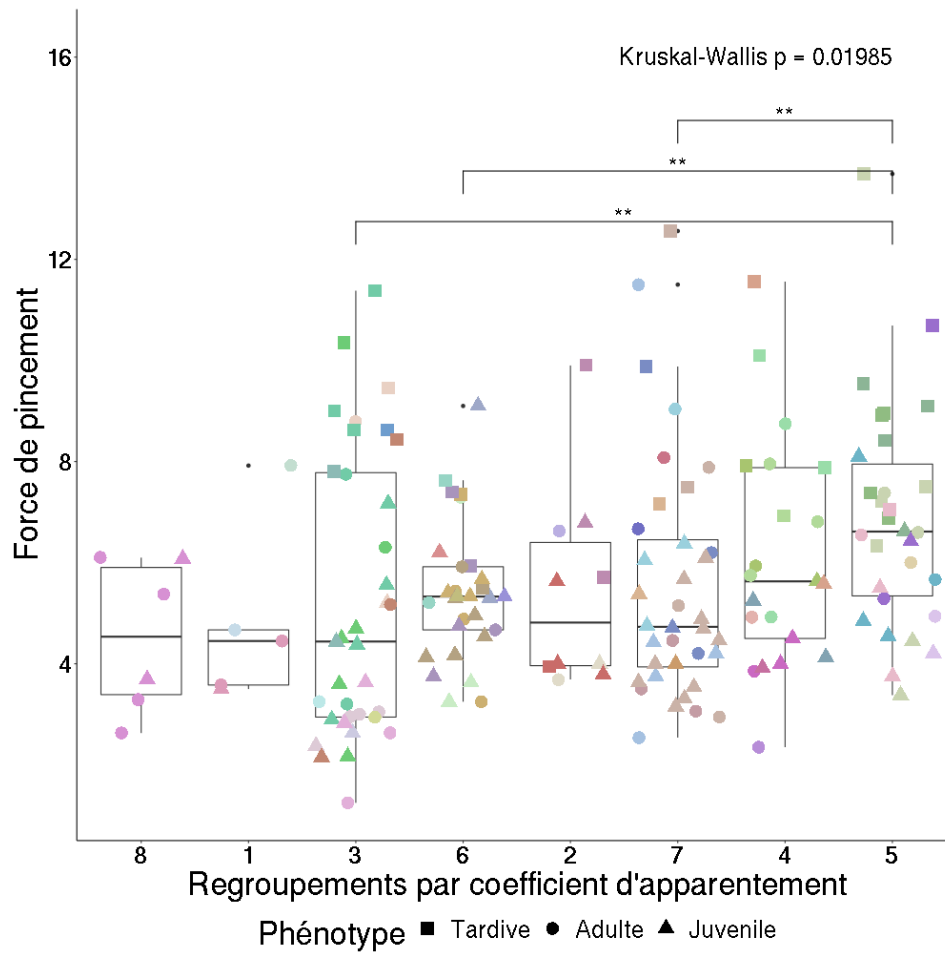


Figure 14 – La force de pincement en fonction de huit sous-regroupements du coefficient d'apparement. Note. Les couleurs représentent les différentes familles.

Finalement, le test de Kruskal Wallis pour les résultats concernant la force du dorsifléchisseur de la cheville en fonction des différentes sous-divisions de regroupements du coefficient d'apparement ne montre aucune donnée significative (Figure 15).



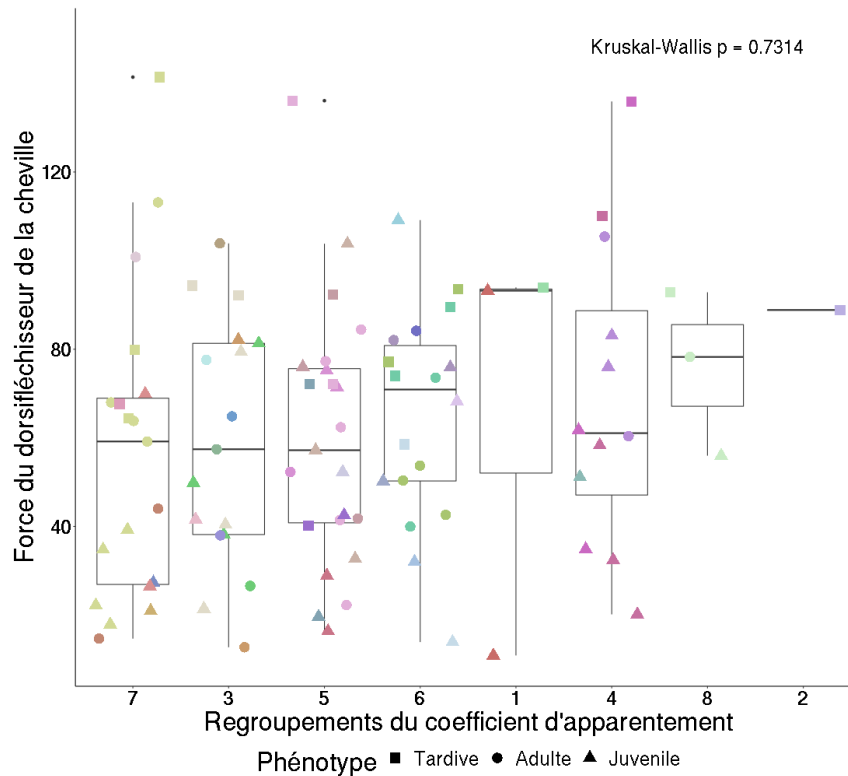


Figure 15 – La force du dorsifléchisseur de la cheville en fonction de 8 sous-regroupements du coefficient d'apparement.

Note. Les couleurs représentent les différentes familles.

### 3.6.2 LES FAMILLES ET LES PHÉNOTYPES EN FONCTION DES REGROUPEMENTS

Comme il était attendu, les familles se regroupent dans une seule sous-division et ce peu importe le nombre de regroupements. En effet, le test de chi carré démontre des valeurs significatives de p-value pour les familles en fonction des différentes sous-division du coefficient d'apparement (Tableau 4).

Tableau 4 – Chi carré des familles en fonctions de différents nombres de sous-division de regroupements du coefficient d'apparement.

Sous-division du coefficient d'apparement	X-squared	df	Chi carré (p-value)
4	344	106	< 2,2e-16
6	860	265	< 2,2e-16

8	1204	371	< 2,2e-16
11	1548	477	< 2,2e-16
15	2236	689	< 2,2e-16
20	3440	1060	< 2,2e-16

Également, les trois différents phénotypes, soit juvénile, adulte et tardif, sont représentés sous les trois différentes formes. Le test du chi carré est non significatif quant aux différents phénotypes en fonction des sous-divisions du coefficient d'apparement (Tableau 5). En effet, les phénotypes sont répartis de manière aléatoire dans chacun des regroupements.

**Tableau 5 – Chi carré des phénotypes en fonctions de différents nombres de sous-divisions du coefficient d'apparement.**

Sous-division du coefficient d'apparement	X-squared	df	p-value
4	6,5594	4	0,1611
6	15,469	10	0,1159
8	23,391	14	0,0542
11	25,223	18	0,1189
15	34,548	26	0,1218
20	51,008	40	0,1139

## **CHAPITRE 4 DISCUSSION**

### **4.1 RETOUR SUR LES OBJECTIFS**

Ce projet de recherche avait comme objectifs de dresser un portrait global du nombre de répétitions CTG en fonction de différents symptômes, de reconstruire la généalogie des patients atteints de DM1 de l'étude longitudinale et de vérifier s'il existait des regroupements généalogiques pour lesquels les symptômes étaient similaires. Le premier objectif a été atteint puisqu'un portrait global du nombre de répétitions CTG a été fait en fonction des différents phénotypes et certains symptômes. Le portrait des CTG et des différents symptômes avait déjà été réalisé auparavant, cependant l'étude ciblait des objectifs différents [16]. Par la suite, le deuxième objectif a été atteint puisque la généalogie complète des patients atteints de DM1 de l'étude longitudinale a été construite. Cette base de données pourra être utilisée pour de futurs projets. Le dernier objectif a également été atteint, cependant il n'a pas été possible de trouver une relation claire entre les symptômes de la DM1 et les regroupements du coefficient d'apparement. Certains résultats significatifs ont été démontrés, mais les résultats n'étaient pas suffisants pour être en mesure de démontrer une relation claire. L'ensemble des résultats sera discuté dans les prochains paragraphes.

### **4.2 PORTRAIT DU NOMBRE DE RÉPÉTITIONS CTG EN FONCTION DES PHÉNOTYPES ET DES SYMPTÔMES**

Les résultats pour les différentes variables en fonction du nombre de CTG montrent que les corrélations avec les différents symptômes ne sont pas parfaites. La variation du nombre de répétitions CTG ne permet pas d'expliquer toutes les variations des symptômes. Les deux diagrammes des mesures de force en fonction des différents intervalles du nombre de CTG montrent que les mesures de forces peuvent être influencées par le nombre de répétitions CTG. Cependant, les corrélations ne sont pas très fortes, il n'est pas possible d'expliquer la force en fonction du nombre de répétitions CTG. Le nombre de répétitions CTG n'est pas la seule variable qui module la faiblesse musculaire

de la maladie. Les phénotypes ont tous une mesure de force qui est variable et celle-ci peut être expliquée par les symptômes multisystémiques de la maladie. Les muscles ne seront pas affectés de manière égale à l'intérieur d'un même phénotype puisqu'il y a une grande variation interindividuelle. Bien que les tests statistiques démontrent une différence, il n'est pas possible de prédire de façon précise l'atteinte du patient avec seulement la longueur des répétitions CTG. Ceci explique en partie l'aspect complexe du diagnostic de la maladie.

Également, il est possible de voir que la corrélation du nombre de CTG et de la sévérité de la maladie n'est pas très élevée. Plusieurs recherches montrent que les CTG corrélaient avec la sévérité de la maladie [16, 98]. Cette corrélation imparfaite entre le nombre de CTG et la sévérité de la maladie laisse croire qu'il y aurait autre chose qui pourrait intervenir dans la génétique de la maladie. C'est en partie en raison de cette corrélation imparfaite que ce projet a pris forme, il faut aller voir plus loin et trouver de nouveaux outils pour mieux comprendre la maladie. D'autres recherches mettent en avant cette corrélation, mais il ne faut pas oublier que le nombre de CTG est utilisé pour faire la classification de la maladie. Cette corrélation était attendue, mais elle reste peu élevée. Cette faible corrélation pourrait être expliquée par la présence d'autres modificateurs génétiques qui pourraient également venir moduler la sévérité de la maladie. Des modificateurs connus en DM1 sont les interruptions au niveau de l'expansion CTG tels que CCG, CCG, CGC ou GCG [33, 103-106]. Ceux-ci vont parfois permettre de stabiliser l'expansion de la répétition en amoindrissant la gravité des symptômes. Les interruptions ne sont pas toujours prises en compte lors du calcul de la longueur de la répétition. Ces interruptions se retrouvent chez environ 5 % des patients et peuvent influencer l'aspect phénotypique de la maladie. Il ne faut pas seulement regarder le nombre de répétitions CTG, mais l'ensemble des modificateurs qui influence le phénotype de la maladie.

Dans le même ordre d'idée, quelques chercheurs se sont posé des questions sur la prévalence normale de l'expansion CTG dans la population en général. Habituellement, les recherches vont regarder le nombre de répétitions CTG chez les gens atteints par la maladie et non chez les gens sains. Il est donc difficile de connaître la longueur normale de ce STR dans la population. Une étude récente

a voulu mieux cerner la longueur de cette expansion dans la population [107]. Les chercheurs ont calculé la prévalence à l'aide de 50 382 échantillons de sang séché anonymisé provenant du programme de dépistage néonatal de l'État de New York [107]. À la suite des analyses, 24 échantillons avaient un taux supérieur à 50 répétitions. Ce taux augmenterait la prévalence à 1 sur 100 000, équivaut à 5 fois plus que la prévalence mondiale actuelle [107]. Il serait donc possible que la maladie soit grandement sous-diagnostiquée ou que notre compréhension de cette répétition en tandem ne soit pas complète.

#### **4.3 NOMBRE DE RÉPÉTITIONS CTG ET PHÉNOMÈNE D'ANTICIPATION**

Par la suite, les générations des patients ne démontrent pas de tendance quant à leur distribution dans les différents intervalles de CTG. Ceci apporte comme conclusion que la génération de l'individu n'aurait pas d'impact sur le spectre du nombre de CTG. Cette conclusion n'est pas ce qui est attendu en vue du phénomène d'anticipation qui existe sur le plan génétique de la DM1. Les générations 3 ou 4 auraient dû avoir un nombre de CTG moins élevé avec des variables de forces plus grandes que les générations 1 et 2 ou du moins en grande partie. Ceci apporte un questionnement sur le phénomène d'anticipation de la maladie. Les familles ne semblent pas avoir la même évolution au niveau de la transmission des répétitions CTG. Il est possible qu'il y ait des différences interfamiliales pour l'évolution du nombre de répétitions dans le temps. Cependant, mettre toutes les familles dans la même analyse a pu influencer les résultats. Les différences interfamiliales sont importantes à prendre en compte. Les individus avec les mêmes niveaux générationnels entre les familles n'ont pas tous les mêmes phénotypes. Il est possible que pour la génération 1 d'une famille le phénotype soit tardif, mais que pour une autre famille le phénotype soit juvénile pour cette même génération. Si un individu à la génération 1 d'une famille était plus atteint qu'un autre individu à la même génération d'une autre famille, cela a pu amener une confusion dans l'analyse. Ces résultats pourraient être expliqués par le fait que nous n'avons pas les phénotypes les plus sévères, congénitaux et infantiles, dans l'échantillon utilisé. L'ajout de ces deux phénotypes aurait pu apporter une meilleure distinction entre les

générations et les phénotypes puisqu'ils se distinguent particulièrement quant à la sévérité des symptômes [13, 108]. Cependant, le risque de décès pour les phénotypes congénitales et infantiles est plus grand; beaucoup décèdent avant la naissance. De ce fait, les échantillons sont beaucoup plus faibles pour ces phénotypes comparativement aux trois autres [36]. C'est pour cette raison que les phénotypes sont absents dans cette étude. Également, le nombre d'individus pour les différentes générations est bien réparti, excepté pour la génération 4 comptant seulement 6 individus. Ce petit nombre a pu biaiser les résultats. Toutes ces raisons peuvent expliquer que le phénomène d'anticipation n'ait pas été observé.

#### **4.4 LE COEFFICIENT D'APPARENTEMENT**

Les réseaux ont été réalisés pour visualiser le coefficient d'apparentement de la cohorte, et ce, à différentes profondeurs généalogiques. Il est possible de voir que les individus se regroupent rapidement dans les généalogies. Ceci peut être facilement expliqué par les origines de la cohorte, provenant de la région du Québec. Les événements démographiques qui ont façonné la population font en sorte que les habitants originaires du Saguenay-Lac-Saint-Jean ont de nombreux ancêtres communs qui sont les fondateurs de la région [99]. Donc les ancêtres communs entre les individus apparaîtront très tôt dans les généalogies.

Par la suite, les graphiques des variables de force en fonction des sous-divisions du coefficient d'apparentement montrent que les membres d'une même famille se regroupent, peu importe le nombre de sous-divisions. Ce résultat était attendu puisque les membres d'une même famille ont un plus grand coefficient d'apparentement par rapport aux autres individus. Également, certaines familles se regroupent dans les sous-divisions signifiant qu'elles ont des ancêtres communs plus rapprochés dans leur généalogie et ont ainsi un plus grand coefficient d'apparentement. Bien que certaines familles se regroupent entre elles, les variables de force et le nombre de CTG sont très variables dans les différentes sous-divisions. Ceci était également attendu puisque les symptômes sont très variables entre les différentes familles et à l'intérieur d'une même famille.

Les résultats générés pour évaluer la relation entre les regroupements du coefficient d'apparement et les différentes variables cliniques n'ont pas été très concluants. Les résultats significatifs pour la force de pincement en fonction des différents sous-regroupements du coefficient d'apparement laissent croire que la force de la pince pourrait être en partie expliquée par le coefficient d'apparement. Évidemment, ces résultats significatifs ne sont pas suffisants pour permettre de tirer des conclusions quant à la relation entre les variables cliniques et le coefficient d'apparement. Il y a tout de même une piste à suivre au regard de la généalogie qui pourrait peut-être apporter des réponses sur la variabilité des symptômes. Les résultats ont permis de comprendre que le coefficient d'apparement est une mesure trop générale pour évaluer l'effet d'un seul loci. Cette mesure se base sur l'ensemble des ancêtres et sur une probabilité pour l'ensemble du génome tandis que les répétitions CTG sont situées sur un seul gène. D'ailleurs, l'expansion CTG causant la DM1 au Saguenay-Lac-Saint-Jean aurait pour origine deux ancêtres communs; regarder l'ensemble de la généalogie était trop vaste pour cette cohorte [21]. Il faudrait cibler ces deux ancêtres et suivre la transmission, prendre l'ensemble des ancêtres était une mesure trop vaste pour les analyses. C'est pourquoi le coefficient d'apparement était une mesure trop large pour former des regroupements et les comparer à des variables cliniques dans le but d'avoir un effet potentiel.

Il est tout de même important de mentionner que ce projet a exploré pour la première fois les regroupements basés sur le coefficient d'apparement. Bien que les résultats n'aient pas été concluants, la généalogie demeure une piste à explorer pour mieux comprendre la DM1.

#### **4.5 PERSPECTIVES**

Il serait judicieux de génotyper ou de séquencer l'ADN des patients et de jumeler ces données avec les données généalogiques pour mieux comprendre la transmission de la répétition CTG. Avec l'haplotype causant la maladie, il serait possible de faire plusieurs autres analyses pour trouver les ancêtres à l'origine de la DM1 au Saguenay-Lac-Saint-Jean. Yotova et al. (2005) avaient été en mesure de trouver que l'haplotype A était responsable d'environ 75% de la fréquence de la DM1 au

Saguenay-Lac-Saint-Jean [21]. Avoir cet haplotype à l'aide de la génétique pourrait grandement augmenter la puissance des analyses. Il serait possible de mieux comprendre son cheminement à travers les générations. Le gène SIX5 amenant une fluctuation dans les symptômes de la maladie, il serait bien dans une prochaine étude de revisiter le nombre de répétitions CTG au niveau de ce gène. Ce gène pourra peut-être amener une compréhension plus large sur l'aspect multisystémique de la maladie.

Il serait également pertinent d'aller voir l'ensemble du génome des patients pour voir si d'autres gènes pourraient influencer la gravité des symptômes de la maladie. La longueur de l'expansion CTG est probablement influencée par d'autres gènes qui pourraient être un facteur clé dans la variabilité des symptômes. Quelques gènes ont déjà été ciblés comme modificateurs de la maladie [28, 29, 40]. Avec la généalogie et la génétique, il serait possible de voir si certaines de ces mutations sont uniques à la région et voir la transmission dans les familles.

Enfin, les regroupements à l'aide du coefficient d'apparentement pourraient être une bonne méthode pour regrouper des familles entre elles et aller voir au niveau génétique ce qui unit ces familles. Si ces familles sont liées par un ancêtre commun, il est possible que leur profil génétique causant la maladie soit similaire et ceci pourrait apporter de nouvelles pistes sur la cause génétique de la maladie. Bien que la maladie soit principalement causée par la mutation d'un gène, il est possible que d'autres gènes viennent influencer la gravité des symptômes.

#### **4.6 RETOMBÉES**

Ce projet a permis d'apporter de nouvelles connaissances dans les domaines de la DM1 et de la généalogie. Le corpus généalogique qui a été construit pourra être réutilisé dans d'autres projets et amener d'autres connaissances sur la transmission de la DM1 dans la région du SLSJ. Grâce aux recherches faites dans le cadre de ce projet cela a permis à l'équipe de faire une demande pour génotyper la cohorte et apporter de nouvelles connaissances au niveau génétique. Ces données ont



permis d'appuyer la demande afin d'aller plus en profondeur avec la génétique et ainsi mieux comprendre la fréquence élevée dans la région du SLSJ.

L'ensemble de ces résultats ont ajouté un rayonnement supplémentaire sur les recherches en DM1 fait aux SLSJ en construisant un corpus généalogique. Les données pourront être utilisées dans plusieurs champs de recherche autant en santé, en histoire, en sociologie, en génétique, etc; démontrant la puissance de l'intersectorialité dans l'avancement des connaissances. Enfin, suite aux recherches découlant de ce projet, il serait pertinent que des articles soient rédigés afin de propager l'information sur la pertinence de l'utilisation des généalogies en recherche.

#### **4.7 FORCES ET LIMITES DE L'ÉTUDE**

Une grande force de l'étude est que cela a permis de réaliser une reconstruction complète de la généalogie des patients atteints de DM1 de la CMNMJ. Cette base de données peut maintenant être utilisée dans d'autres recherches sur la maladie et augmenter ainsi la puissance des analyses. En plus de pouvoir étudier la DM1, la généalogie offrira une opportunité unique d'explorer la transmission des maladies rares en général. La population qui était à l'étude était un grand avantage dû à son origine et son histoire. L'effet fondateur qu'a connu la population apporte un aspect important pour la recherche. La cohorte utilisée était formée de familles et elles sont toutes liées de près ou de loin, il est donc plus facile de trouver ce qui les lie à la DM1. Les études travaillant avec des patients atteints de DM1 ont souvent des échantillons populationnels faibles puisque la DM1 est une maladie rare. Avoir un échantillon de 200 participants est une force dans ce domaine.

Néanmoins, il est important de mentionner que le nombre de participants n'était pas le même pour chaque échantillon de variables. La force du dorsifléchisseur de la cheville comprenait seulement 96 données. Ce nombre inférieur a pu amener un manque de puissance dans les analyses. De plus, c'était l'une des premières études faites avec des regroupements du coefficient d'apparentement dans une population à effet fondateur. L'une des difficultés rencontrées a été le peu de documentation dans le domaine, très peu de recherches ont été faites dans le même sens. Plusieurs recherches ont été faites

sur la population du Saguenay-Lac-Saint-Jean, mais les regroupements généalogiques basés sur le coefficient d'apparentement n'ont jamais été tentés auparavant.

## CONCLUSION

La présente étude a permis une reconstruction complète de la généalogie des patients DM1 à l'aide du fichier de population BALSAC. Ces généalogies ont été utilisées pour explorer les relations entre différentes variables cliniques de la DM1 et les regroupements à l'aide du coefficient d'apparentement. Bien que les analyses n'aient pas permis de déceler une relation entre le coefficient d'apparentement et différentes variables cliniques, les résultats ont permis de comprendre que le coefficient d'apparentement était une mesure trop large pour étudier l'expansion CTG. La méthode de regroupement à l'aide du coefficient d'apparentement pourrait être une bonne manière pour faire de plus grands regroupements familiaux et étudier les transmissions.

## LISTE DE RÉFÉRENCES

1. Dystrophie Musculaire Canada. *Site officiel de l'Association*. 2022; Available from: <https://muscle.ca/fr/>.
2. AFMTéléthon, I.p.g., *Avancées dans la maladie de Steinert*. Savoir et comprendre, Avancées de la recherche, 2022.
3. Emery, A.E. and M.L. Emery, *Edward Meryon (1809-1880) and muscular dystrophy*. Journal of Medical Genetics, 1993. **30**(6): p. 506-511.
4. Harper, P.S., *Myotonic Dystrophy : The Fact*, ed. O.x.U. Press. 2009, Unitd State.
5. Harper, P.S., et al., *Myotonic dystrophy : present management, future therapy*. 2004, Oxford New York: Oxford University Press.
6. De Antonio, M., et al., *Unravelling the myotonic dystrophy type 1 clinical spectrum: A systematic registry-based study with implications for disease classification*. Revue Neurologique, 2016. **172**(10): p. 572-580.
7. Gagnon, C., et al., *Participation restriction in childhood phenotype of myotonic dystrophy type 1: a systematic retrospective chart review*. Developmental medicine and child neurology, 2017. **59**(3): p. 291-296.
8. Kaliman, P. and E. Llagostera, *Myotonic dystrophy protein kinase (DMPK) and its role in the pathogenesis of myotonic dystrophy 1* Cellular Signalling, 2008. **20**(11): p. 1935-1941.
9. Victor, D., *Muscle biopsy*, ed. S. Elsevier. 2013, Chine.
10. Thornton, C., *The myotonic dystrophies*. Seminars in Neurology, 1999. **19**(1): p. 25-33.
11. Fayssol, A. and O. Nardi, *Heart and Steinert's disease*, in *Annales De Cardiologie Et D'angéiologie*. 2011. p. 225-229.
12. Coramh. *Maladies héréditaires dominantes*. 2023 [cited 2023 19 janvier]; Available from: <https://coramh.org/maladies-hereditaires-dominantes/#anchor>.
13. De Antonio, M., et al., *Unravelling the myotonic dystrophy type 1 clinical spectrum: A systematic registry-based study with implications for disease classification*. Revue neurologique, 2016. **172**(10): p. 572-580.
14. Bchetnia, M., et al., *Genetic burden linked to founder effects in Saguenay–Lac-Saint-Jean illustrates the importance of genetic screening test availability*. 2021.
15. Gagnon, C., J. Mathieu, and L. Noreau, *Life habits in myotonic dystrophy type 1*. Journal of Rehabilitation Medicine, 2007. **39**(7): p. 560-566.
16. Overend, G., et al., *Allele length of the DMPK CTG repeat is a predictor of progressive myotonic dystrophy type 1 phenotypes*. Human molecular genetics, 2019. **28**(13): p. 2245-2254.
17. Mathieu, J., et al., *A 10-year study of mortality in a cohort of patients with myotonic dystrophy*. Neurology, 1999. **12**(8): p. 1658-62.
18. Steyaert, J., et al., *A study of the cognitive and psychological profile in 16 children with congenital or juvenile myotonic dystrophy*. Clinical genetics, 1997. **52**(3): p. 135-41.
19. Jean, S., et al., *Comparisons of intellectual capacities between mild and classic adult-onset phenotypes of myotonic dystrophy type 1 (DMI)*. Orphanet journal of rare diseases, 2014. **9**: p. 186.

20. CIUSSS, S.L.-S.-J. *GUIDE À L'INTENTION DES OMNIPRATICIENS : DYSTROPHIE MYOTONIQUE DE TYPE 1*. 2019 [1 octobre 2023]; Available from: [https://santesaglac.gouv.qc.ca/medias/2019/02/OMNI\\_DM1\\_Final\\_fr.pdf](https://santesaglac.gouv.qc.ca/medias/2019/02/OMNI_DM1_Final_fr.pdf).
21. Yotova, V., et al., *Anatomy of a founder effect: myotonic dystrophy in Northeastern Quebec*. Human Genetics, 2005. **117**(2-3): p. 177-187.
22. Laing, N., et al., *An overview of the genetic basis of cardiovascular disease*. South African Medical Journal, 2019. **109**(6): p. 364-370.
23. Fan, H. and J.-Y. Chu, *A brief review of short tandem repeat mutation*. Genomics, proteomics & bioinformatics, 2007. **5**(1): p. 7-14.
24. Brook, J., et al., *Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member*. Cell, 1992. **68**(4): p. 799:808.
25. Eyleen, C., V. Melissa, and F. Morales, *Analysis of mutational dynamics at the DMPK (CTG)<sub>n</sub> locus identifies saliva as a suitable DNA sample source for genetic analysis in myotonic dystrophy type 1*. PLoS ONE, 2019. **14**(5).
26. Savić, D.a., et al., *250 CTG repeats in DMPK is a threshold for correlation of expansion size and age at onset of juvenile-adult DM1*. Human Mutation, 2002. **19**(2): p. 131-139.
27. Morales, F., et al., *Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity*. Human molecular genetics, 2012. **21**(16): p. 3558-67.
28. Morales, F., et al., *Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity*. Human Molecular Genetics, 2012. **21**(16): p. 3558-3567.
29. Tomé, S. and G. Gourdon, *DM1 Phenotype Variability and Triplet Repeat Instability: Challenges in the Development of New Therapies*. International journal of molecular sciences, 2020. **21**(2).
30. Raymond, K., et al., *A 9-year follow-up study of the natural progression of upper limb performance in myotonic dystrophy type 1: A similar decline for phenotypes but not for gender*. Neuromuscular Disorders, 2017. **27**(7): p. 673-682.
31. Darren, G.M., et al., *Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses*. Human Molecular Genetics, 1995. **4**(1): p. 1-8.
32. Cumming, S.A., et al., *De novo repeat interruptions are associated with reduced somatic instability and mild or absent clinical features in myotonic dystrophy type 1*. European journal of human genetics : EJHG, 2018. **26**(11): p. 1635-1647.
33. Tsai, Y.-C., et al., *Identification of a CCG-Enriched Expanded Allele in Patients with Myotonic Dystrophy Type 1 Using Amplification-Free Long-Read Sequencing*. The Journal of molecular diagnostics : JMD, 2022. **24**(11): p. 1143-1154.
34. López de Munain, A., et al., *Influence of the sex of the transmitting grandparent in congenital myotonic dystrophy*. Journal of medical genetics, 1995. **32**(9): p. 689-91.
35. Geifman-Holtzman, O. and K. Fay, *Prenatal diagnosis of congenital myotonic dystrophy and counseling of the pregnant mother: Case report and literature review*. American Journal of Medical Genetics, 1998. **78**(3): p. 250-253.
36. Barbé, L., et al., *CpG Methylation, a Parent-of-Origin Effect for Maternal-Biased Transmission of Congenital Myotonic Dystrophy*. American journal of human genetics, 2017. **100**(3): p. 488-505.

37. Leeflang, E.P., M.S. McPeck, and N. Arnheim, *Analysis of meiotic segregation, using single-sperm typing: meiotic drive at the myotonic dystrophy locus*. American journal of human genetics, 1996. **59**(4): p. 896-904.
38. Jan, R., et al. *Characterisation of Non-Pathogenic Premutation-Range Myotonic Dystrophy Type 2 Alleles*. 2021. **10**, 3934 DOI: 10.3390/jcm10173934.
39. Pratte, A., et al., *Anticipation in myotonic dystrophy type 1 parents with small CTG expansions*. American journal of medical genetics. Part A, 2015. **167A**(4): p. 708-14.
40. Flower, M., et al., *MSH3 modifies somatic instability and disease severity in Huntington's and myotonic dystrophy type 1*. Brain : a journal of neurology, 2019.
41. Maeda, M., et al., *Identification, tissue-specific expression, and subcellular localization of the 80- and 71-kDa forms of myotonic dystrophy kinase protein*. The Journal of biological chemistry, 1995. **270**(35): p. 20246-9.
42. López Castel, A., et al., *Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues*. Human molecular genetics, 2011. **20**(1): p. 1-15.
43. Murányi, A., et al., *Myotonic dystrophy protein kinase phosphorylates the myosin phosphatase targeting subunit and inhibits myosin phosphatase activity*. FEBS Letters, 2001. **493**(2-3): p. 80-84.
44. Groenen, P.J., et al., *Constitutive and regulated modes of splicing produce six major myotonic dystrophy protein kinase (DMPK) isoforms with distinct properties*. Human molecular genetics, 2000. **9**(4): p. 605-16.
45. Kaliman, P. and E. Llagostera, *Myotonic dystrophy protein kinase (DMPK) and its role in the pathogenesis of myotonic dystrophy 1*. cellular signaling, 2008. **20**: p. 1935-1941.
46. Wansink, D.G., et al., *Alternative splicing controls myotonic dystrophy protein kinase structure, enzymatic activity, and subcellular localization*. Molecular and cellular biology, 2003. **23**(16): p. 5489-501.
47. Pantic, B., et al., *Myotonic dystrophy protein kinase (DMPK) prevents ROS-induced cell death by assembling a hexokinase II-Src complex on the mitochondrial surface*. Cell death & disease, 2013. **4**: p. e858.
48. Institute, N.H.G.R. *Promoter*. 2024 [cited 2024 4 avril]; Available from: <https://www.genome.gov/>.
49. Boucher, C.A., et al., *A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic dystrophy unstable (CTG) n repeat*. Human Molecular Genetics, 1995. **4**(10): p. 1919-1925.
50. Kirby, R.J., et al., *Drosophila homolog of the myotonic dystrophy-associated gene, SIX5, is required for muscle and gonad development*. Current biology : CB, 2001. **11**(13): p. 1044-9.
51. Harmon, E.B., et al., *Myotonic dystrophy protein kinase is expressed in embryonic myocytes and is required for myotube formation*. Developmental Dynamics, 2008. **237**(9): p. 2353-2366.
52. Kaliman, P., et al., *Myotonic dystrophy protein kinase phosphorylates phospholamban and regulates calcium uptake in cardiomyocyte sarcoplasmic reticulum*. The Journal of biological chemistry, 2005. **280**(9): p. 8016-21.
53. Klesert, T.R., et al., *Mice deficient in Six5 develop cataracts: implications for myotonic dystrophy*. Nature genetics, 2000. **25**(1): p. 105-9.

54. Chung, B.M., V.M.M. Wong, and P.M. Ip, *Prevalence of Neuromuscular Diseases in Chinese Children: A Study in Southern China*. Journal of Child Neurology, 2003. **18**(3): p. 217-219.
55. Hsiao KM, et al., *Epidemiological and genetic studies of myotonic dystrophy type 1 in Taiwan*. Neuroepidemiology, 2003. **22**(5): p. 283-9.
56. Darin, N. and M. Tulinius, *Neuromuscular disorders in childhood: a descriptive epidemiological study from western Sweden*. Neuromuscular Disorders, 2000. **10**(1): p. 1-9.
57. Mladenovic, J., et al., *Survival and mortality of myotonic dystrophy type 1 (Steinert's disease) in the population of Belgrade*. European Journal of Neurology, 2006. **13**(5): p. 451-454.
58. Mah, J.K., et al., *A Systematic Review and Meta-analysis on the Epidemiology of the Muscular Dystrophies*. Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques, 2016. **43**(1): p. 163-177.
59. Nakagawa M, et al., *Epidemiology of progressive muscular dystrophy in Okinawa, Japan. Classification with molecular biological techniques*. Neuroepidemiology. Neuroépidémiologie, 1991. **10**(4): p. 185-91.
60. Siciliano, G., et al., *Epidemiology of myotonic dystrophy in Italy: re-appraisal after genetic diagnosis*. Clinical Genetics, 2001. **59**(5): p. 344-349.
61. Norwood, F.L.M., et al., *Prevalence of genetic muscle disease in Northern England: in-depth analysis of a muscle clinic population*. Brain : a journal of neurology, 2009. **132**(Pt 11): p. 3175-86.
62. Ford C, Kidd A, and H.-T. G, *Myotonic dystrophy in Otago, New Zealand*. Nouvelle-Zélande Med J, 2006. **119**(1241).
63. Magee A and N. NC, *The epidemiology of myotonic dystrophy in Northern Ireland*. Community Genet., 1999. **2**(4): p. 179-83.
64. Segel, R., et al., *Prevalence of myotonic dystrophy in Israeli Jewish communities: Inter-community variation and founder premutations*. American Journal of Medical Genetics Part A, 2003. **119A**(3): p. 273-278.
65. Medica, I., D. Markovi, and B. Peterlin, *Genetic epidemiology of myotonic dystrophy in Istria, Croatia*. Acta Neurologica Scandinavica, 1997. **95**(3): p. 164-166.
66. López de Munain A, et al., *Prevalence of myotonic dystrophy in Guipúzcoa (Basque Country, Spain)*. Neurology, 1993. **43**(8): p. 1573-6.
67. Mathieu, J. and C. Prévost, *Epidemiological surveillance of myotonic dystrophy type 1: a 25-year population-based study*. Neuromuscular disorders : NMD, 2012. **22**(11): p. 974-9.
68. Bouchard, G.r. and M. De Braekeleer, *Pourquoi des maladies héréditaires? : Population et génétique au Saguenay-Lac-Saint-Jean*, ed. L.é.d. Septentrion. 1992. 184.
69. Vézina, H., M. Tremblay, and L. Houde, *Mesures de l'apparentement biologique au Saguenay-Lac-St-Jean (Québec, Canada) à partir de reconstitutions généalogiques*. 2004. p. 67-83.
70. Gauvin, H., et al., *Genome-wide patterns of identity-by-descent sharing in the French Canadian founder population*. European Journal of Human Genetics, 2014. **22**(6): p. 814-821.
71. Breaking the Bottleneck of Genomes: Understanding Gene Function Across, T., B. United States. Department of Energy. Office of, and R. Environmental, *Breaking the*

- Bottleneck of Genomes: Understanding Gene Function Across Taxa : workshop report*. 2019, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research: [Washington, D.C.].
72. Gagnon, L., et al., *Fine-scale genetic structure and rare variant frequencies*. bioRxiv, 2024: p. 2024.02.02.578687.
  73. Michael, H.R. and S.H. Richard, *A population history of North America*, C.U. Press, Editor. 2000: Cambridge, United Kingdom. p. 736.
  74. Gauvreau, D. and M. Bourque *Mouvements migratoires et familles : le peuplement du Saguenay avant 1911*. Revue d'histoire de l'Amérique française, 1988. **42**, 167-192 DOI: 10.7202/304677ar.
  75. Juneau, C.I., I.e. Québec . Ministère de la famille et de, and Q.b. Bureau de la statistique du, *Un portrait statistique des familles et des enfants au Québec*. 1999, Ministère de la famille et de l'enfance: Québec.
  76. Canada, S. *Fécondité : moins d'enfants, mères plus âgées*. 2018 2018-05-17 [cited 2023 1 février]; Available from: <https://www150.statcan.gc.ca/n1/pub/11-630-x/11-630-x2014002-fra.htm>.
  77. Prost, M., M.F. Mayer, and M. Boivert, *Le coefficient de parenté et ses composantes proche et éloignée. Estimation à partir du registre de population «Vallouise en Briançonnais»*. Antropo, 2022. **48**: p. 1-14.
  78. Templeton, A.R., *Human population genetics and genome*, ed. A. Press. 2019: Elsevier. 485.
  79. Maia R T and Araujo M C, *Population genetics*, ed. IntechOpen. 2022, London.
  80. Gagnon, L., et al., *Deciphering the genetic structure of the Quebec founder population using genealogies*. 2023.
  81. Jan, R., et al. *Characterisation of Non-Pathogenic Premutation-Range Myotonic Dystrophy Type 2 Alleles*. Journal of Clinical Medicine, 2021. **10**, 3934 DOI: 10.3390/jcm10173934.
  82. Joosten, I.B.T., et al., *Parental repeat length instability in myotonic dystrophy type 1 pre- and protomutations*. European journal of human genetics : EJHG, 2020. **28**(7): p. 956-962.
  83. Abbruzzese, C., et al., *Instability of a premutation allele in homozygous patients with myotonic dystrophy type 1*. Annals of neurology, 2002. **52**(4): p. 435-41.
  84. Brunner, H.G., et al., *Myotonic dystrophy is closely linked to the gene for muscle-type creatine kinase (CKMM)*. Human Genetics, 1989. **81**(4): p. 308-310.
  85. Goldman, A., M. Ramsay, and T. Jenkins, *New founder haplotypes at the myotonic dystrophy locus in southern Africa*. American journal of human genetics, 1995. **56**(6): p. 1373-8.
  86. Goldman, A., et al., *Founder effect and prevalence of myotonic dystrophy in South Africans: molecular studies*. American journal of human genetics, 1996. **59**(2): p. 445-52.
  87. Yang, P.Q., *Transnationalism and genealogy*. 2020, MDPI: Basel, Switzerland.
  88. BALSAC, *Projet BALSAC : bibliographie*. 2004, Université du Québec à Chicoutimi.
  89. Bouchard, G.r., et al., *Le fichier de population BALSAC. Situation et perspectives*. Annales de démographie historique, 1999. **1998**(2): p. 187-196.
  90. Anderson-Trocmé, L., et al., *On the genes, genealogies, and geographies of Quebec*. Science, 2023. **380**(6647): p. 849-855.



91. Moreau, C., et al., *Native American Admixture in the Quebec Founder Population*. PLoS ONE, 2013. **8**(6).
92. Moreau, C., et al., *When genetics and genealogies tell different stories-maternal lineages in Gaspesia*. Annals of human genetics, 2011. **75**(2): p. 247-54.
93. Roy-Gagnon, M.-H.I.n., et al., *Genomic and genealogical investigation of the French Canadian founder population structure*. Human Genetics, 2011. **129**(5): p. 521-531.
94. Nelson, D., et al., *Inferring Transmission Histories of Rare Alleles in Population-Scale Genealogies*. The American Journal of Human Genetics, 2018. **103**(6): p. 893-906.
95. Gauvin, H., et al. *GENLIB: an R package for the analysis of genealogical data*. BMC Bioinformatics, 2015. **16**, 1-10 DOI: 10.1186/s12859-015-0581-5.
96. Prost, M., F. Mayer, and M. Boivert, *The kinship coefficient and its close and distant components. Estimation based on the "Vallouise en Briançonnais" population register*. Antropo, 2022. **48**: p. 1-14.
97. Fu, Y.H., et al., *An unstable triplet repeat in a gene related to myotonic muscular dystrophy*. Science (New York, N.Y.), 1992. **255**(5049): p. 1256-8.
98. Légaré, C., et al., *DMPK gene DNA methylation levels are associated with muscular and respiratory profiles in DMI*. Neurol Genet, 2019. **5**(3): p. e338.
99. Mathieu, J., M. De Braekeleer, and C. Prévost, *Genealogical reconstruction of myotonic dystrophy in the Saguenay-Lac-Saint- Jean area (Quebec, Canada)*. Neurology, 1990. **40**(5).
100. Petitclerc, E.m., et al., *Lower limb muscle strength impairment in late-onset and adult myotonic dystrophy type 1 phenotypes*. Muscle & Nerve, 2017. **56**(1): p. 57-63.
101. Gagnon, C., et al., *Health supervision and anticipatory guidance in adult myotonic dystrophy type 1*. Neuromuscular disorders : NMD, 2010. **20**(12): p. 847-51.
102. De Meo, P., et al., *Generalized Louvain method for community detection in large networks*, in *2011 11th International Conference on Intelligent Systems Design and Applications*. 2011, IEEE. p. 88-93.
103. Cumming, S.A., et al., *De novo repeat interruptions are associated with reduced somatic instability and mild or absent clinical features in myotonic dystrophy type 1*. European journal of human genetics : EJHG, 2018. **26**(11): p. 1635-1647.
104. Rasmussen, A., et al., *High Resolution Analysis of DMPK Hypermethylation and Repeat Interruptions in Myotonic Dystrophy Type 1*. Genes, 2022. **13**(6).
105. Antoine, M., et al. *Robust Detection of Somatic Mosaicism and Repeat Interruptions by Long-Read Targeted Sequencing in Myotonic Dystrophy Type 1*. International Journal of Molecular Sciences, 2021. **22**, 2616 DOI: 10.3390/ijms22052616.
106. Santoro, M., et al., *Expansion size and presence of CCG/CTC/CGG sequence interruptions in the expanded CTG array are independently associated to hypermethylation at the <b>DMPK</b> locus in myotonic dystrophy type 1 (DMI)*. BBA - Molecular Basis of Disease, 2015. **1852**(12): p. 2645-2652.
107. Johnson, N.E., et al., *Population-Based Prevalence of Myotonic Dystrophy Type 1 Using Genetic Analysis of Statewide Blood Screening Program*. Neurology, 2021. **96**(7): p. e1045-e1053.
108. Gagnon, C., J. Mathieu, and L. Noreau, *Life habits in myotonic dystrophy type 1*. Journal of rehabilitation medicine, 2007. **39**(7): p. 560-6.

## **CERTIFICATION ÉTHIQUE**

Ce mémoirea fait l'objet d'une certification éthique auprès du CER-UQAC. Le numéro du certificat est 2024-1393 - Évaluation de la fonction motrice, des atteintes multisystémiques et de la participation sociale dans la dystrophie myotonique de type 1 : Une étude longitudinale sur une période de 9 ans.

### ANNEXE 1

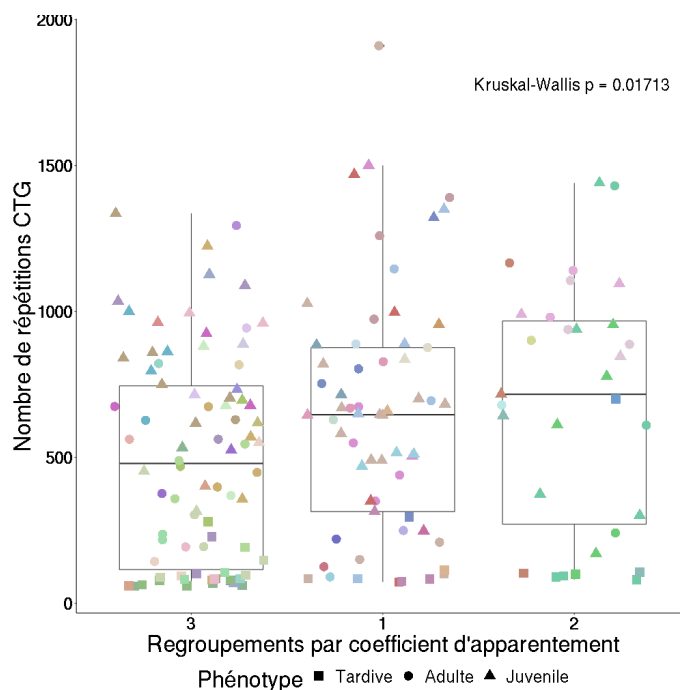


Figure 16 – Nombre de répétition CTG en fonction de 3 sous-regroupements du coefficient d'apparement.

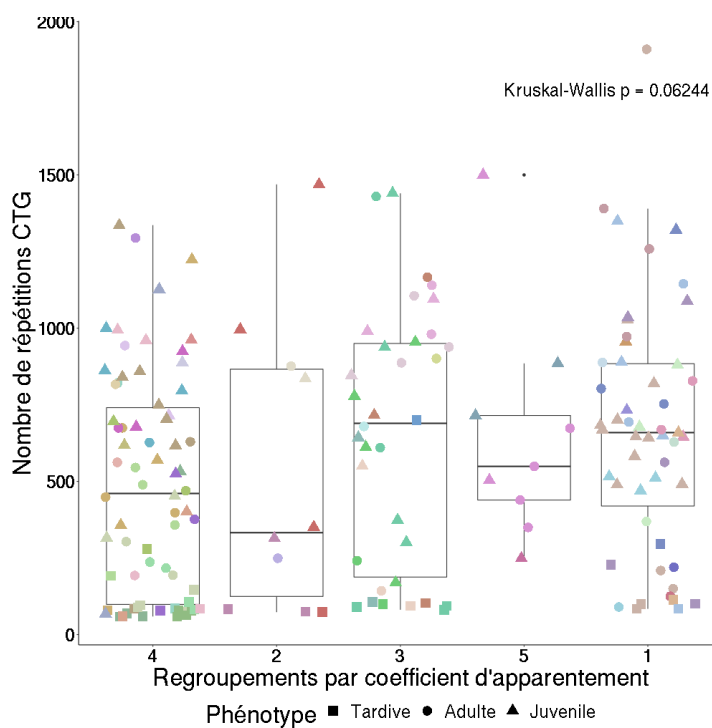


Figure 17 – Nombre de répétition CTG en fonction de 5 sous-regroupements du coefficient d'apparement.

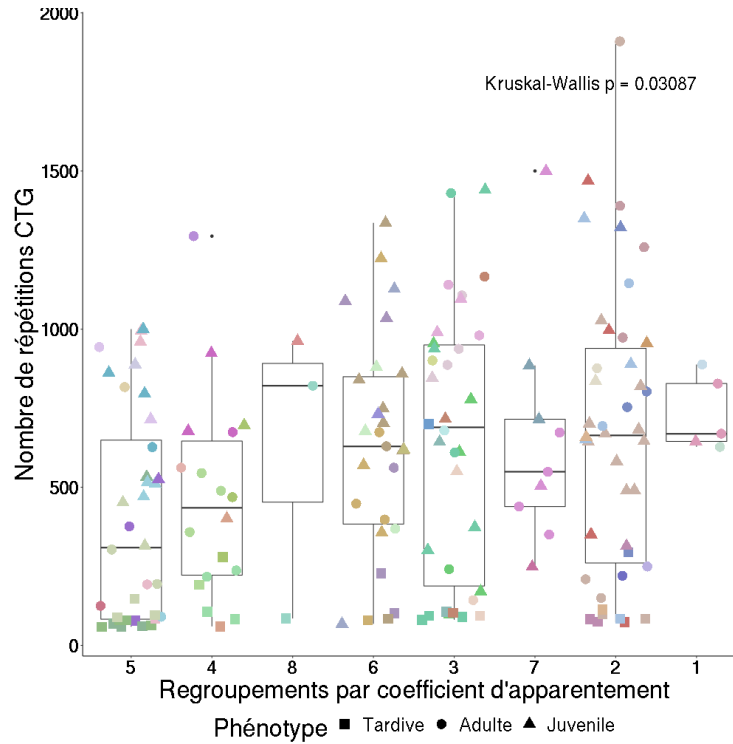


Figure 18 – Nombre de répétition CTG en fonction de 8 sous-regroupements du coefficient d'apparement.

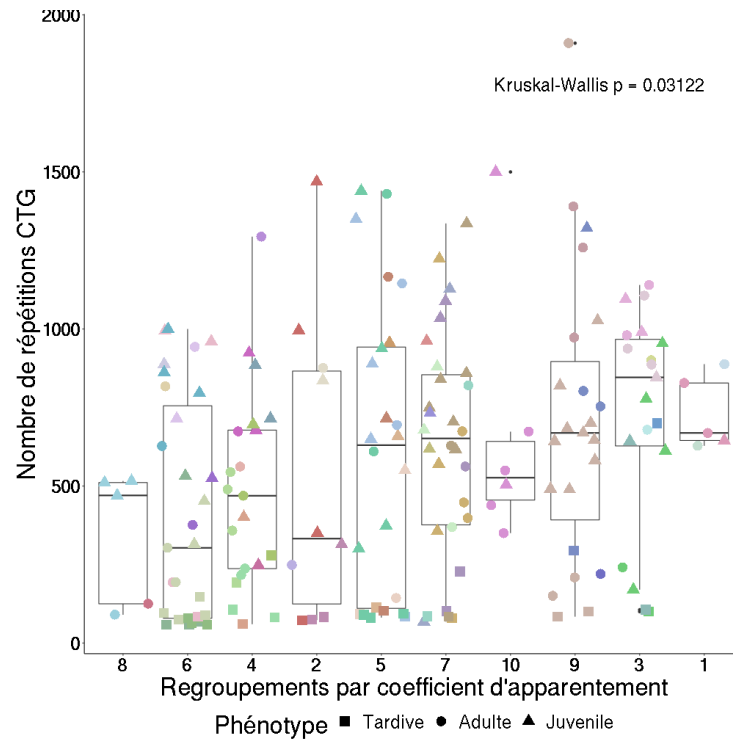


Figure 19 – Nombre de répétition CTG en fonction de 10 sous-regroupements du coefficient d'apparement.

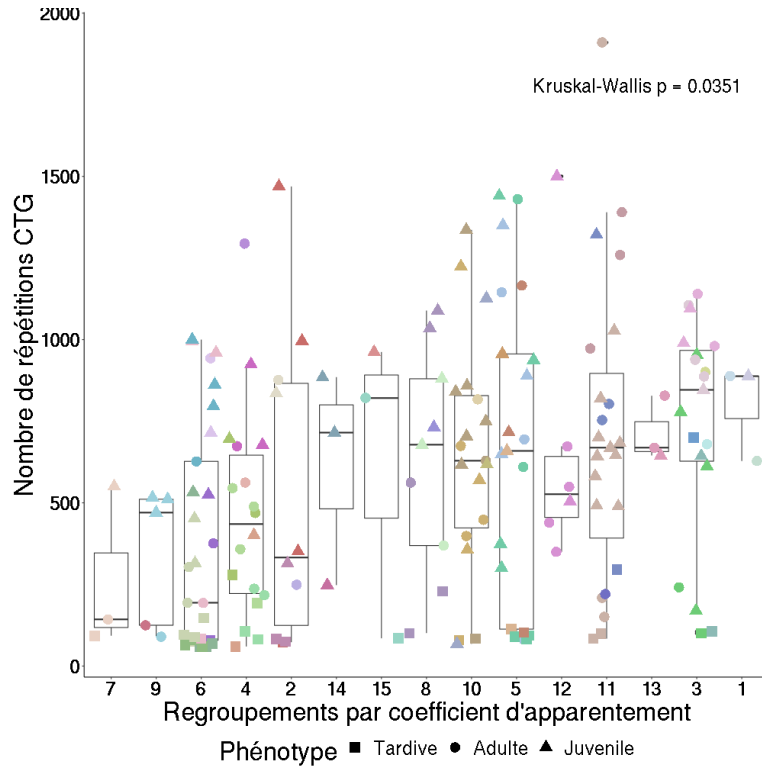


Figure 20 – Nombre de répétition CTG en fonction de 15 sous-regroupements du coefficient d'apparement.

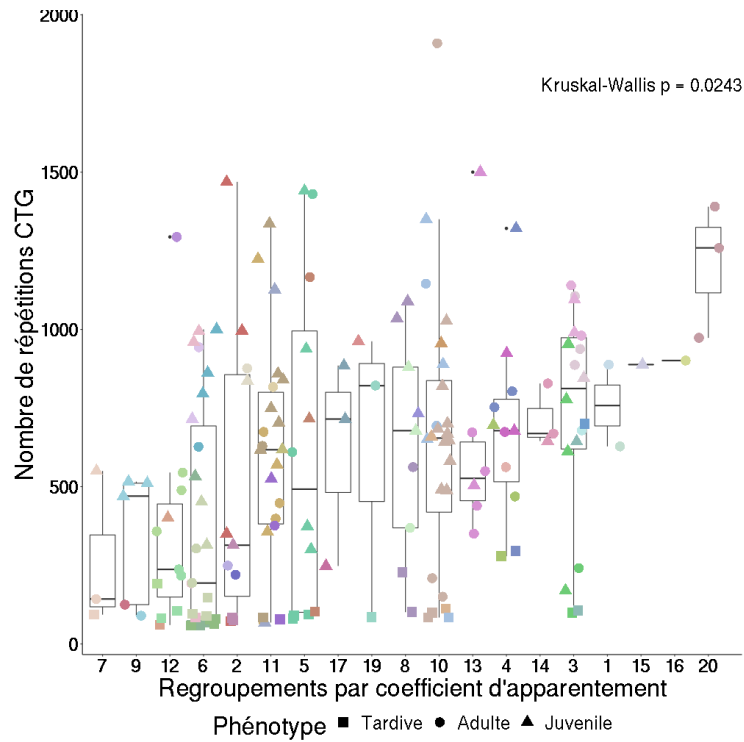


Figure 21 – Nombre de répétition CTG en fonction de 20 sous-regroupements du coefficient d'apparement.

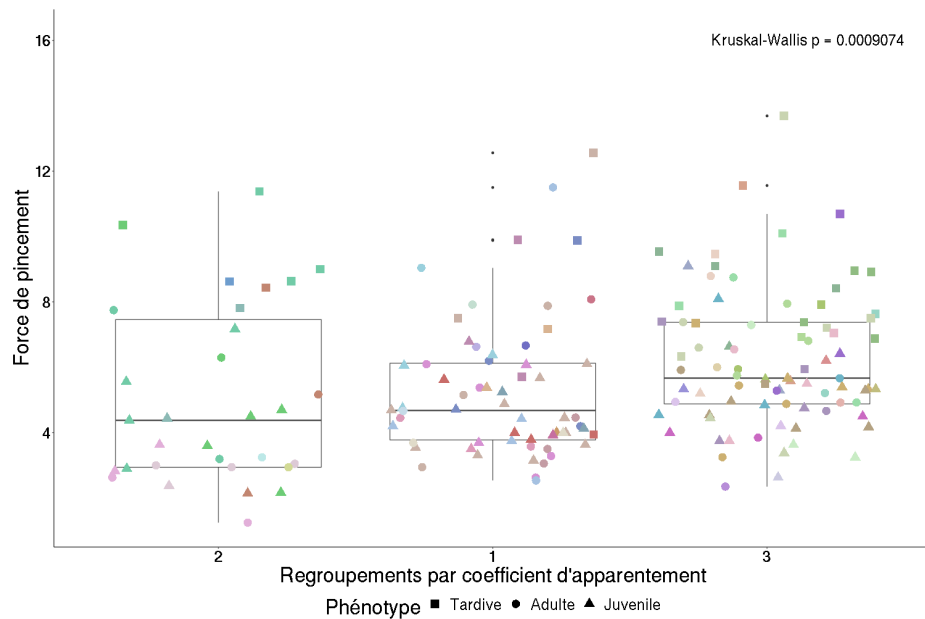


Figure 22 – Force de pincement en fonction de 3 sous-regroupements du coefficient d'apparement.

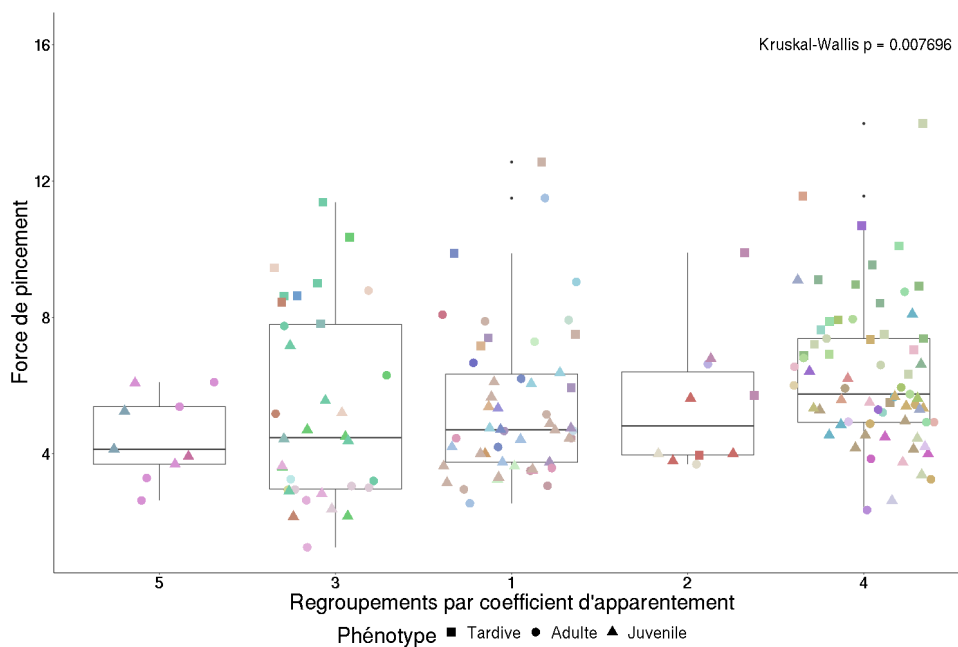


Figure 23 – Force de pincement en fonction de 5 sous-regroupements du coefficient d'apparement.

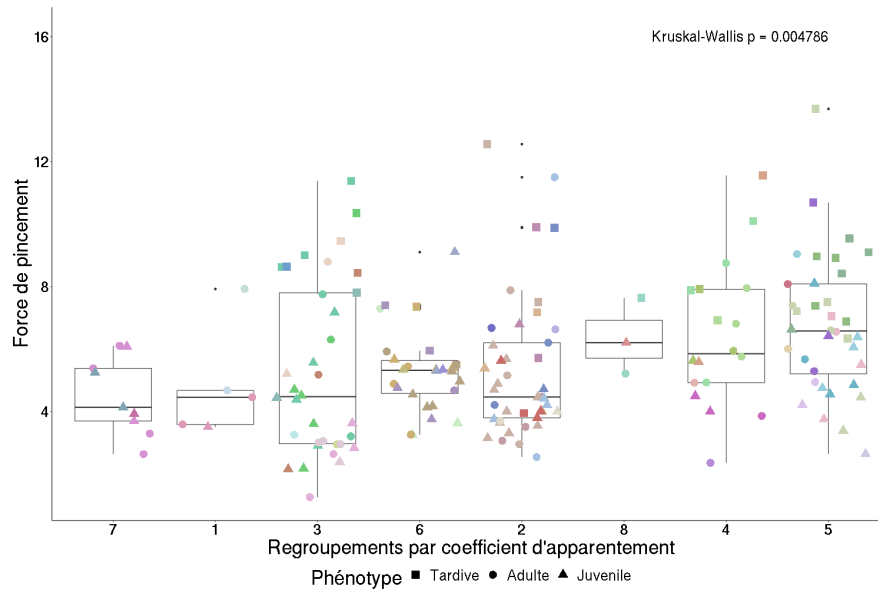


Figure 24 – Force de pincement en fonction de 8 sous-regroupements du coefficient d'apparement.

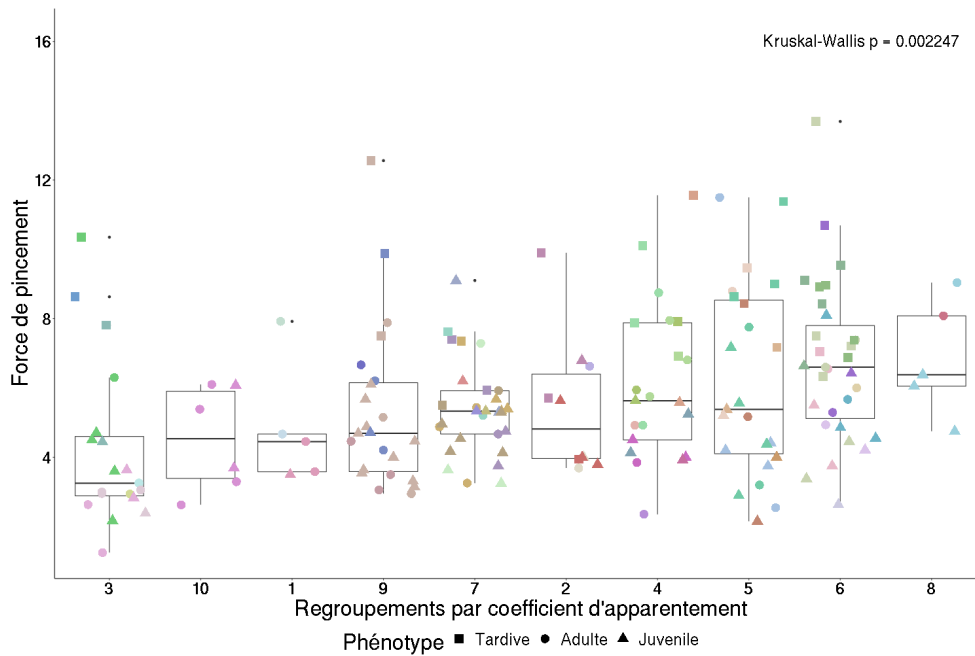


Figure 25 – Force de pincement en fonction de 10 sous-regroupements du coefficient d'apparement.

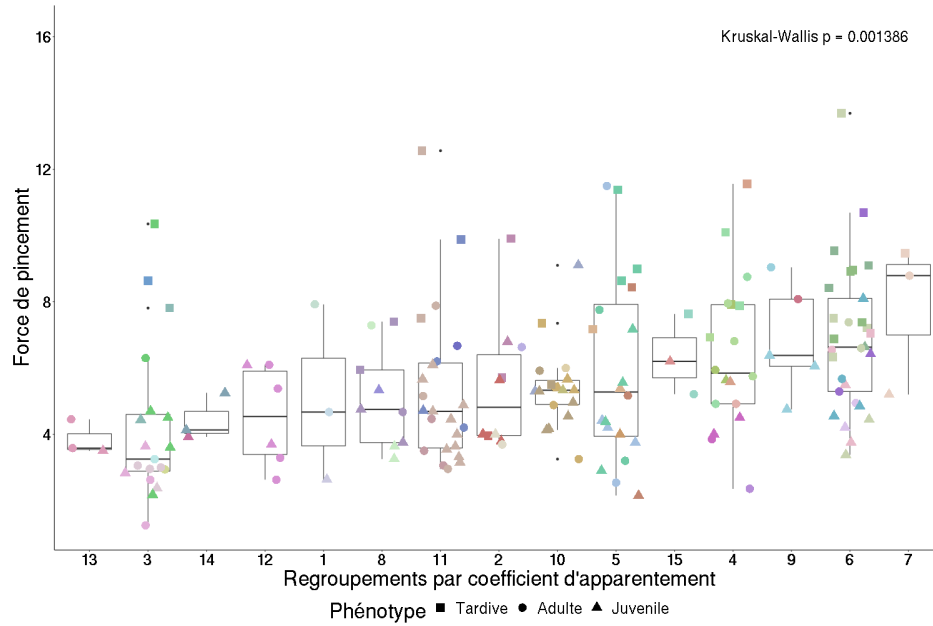


Figure 26 – Force de pincement en fonction de 15 sous-regroupements du coefficient d'apparement.

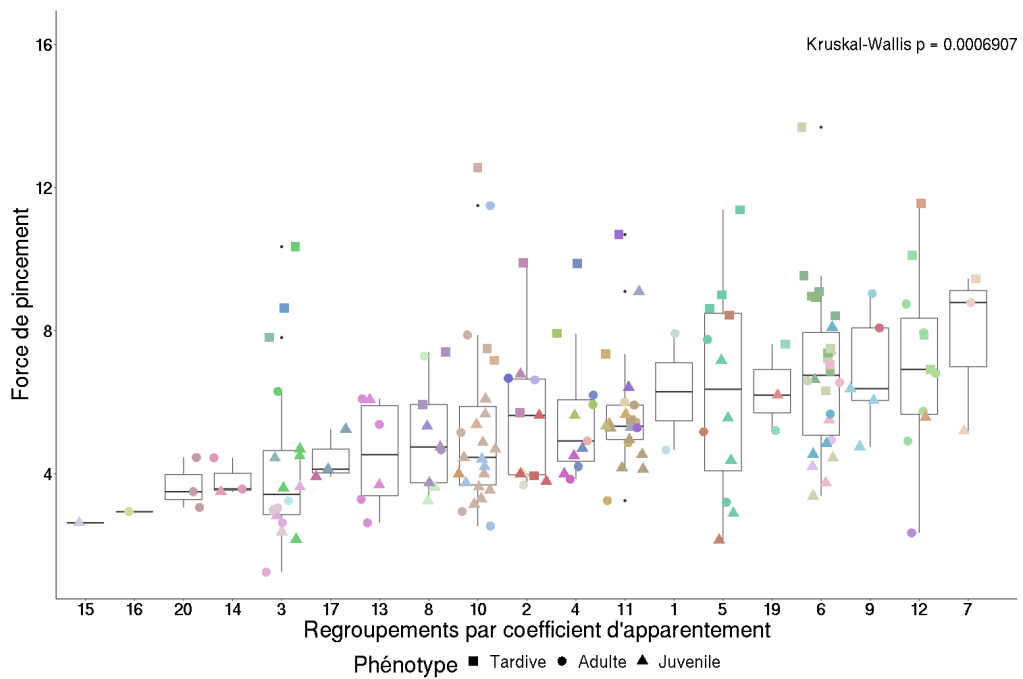


Figure 27 – Force de pincement en fonction de 20 sous-regroupements du coefficient d'apparement.



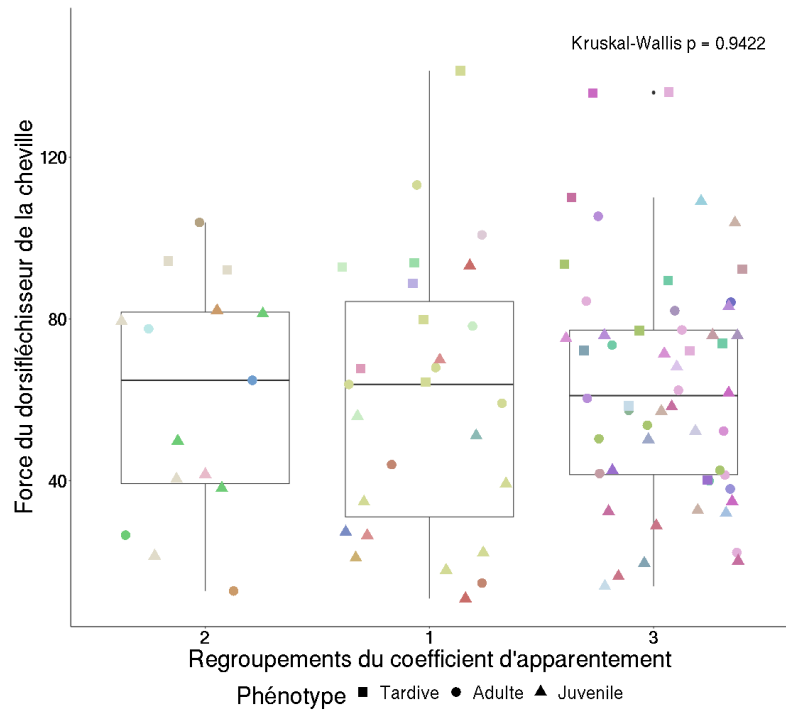


Figure 28 – Force du dorsifléchisseur de la cheville en fonction de 3 sous-regroupements du coefficient d'apparement.

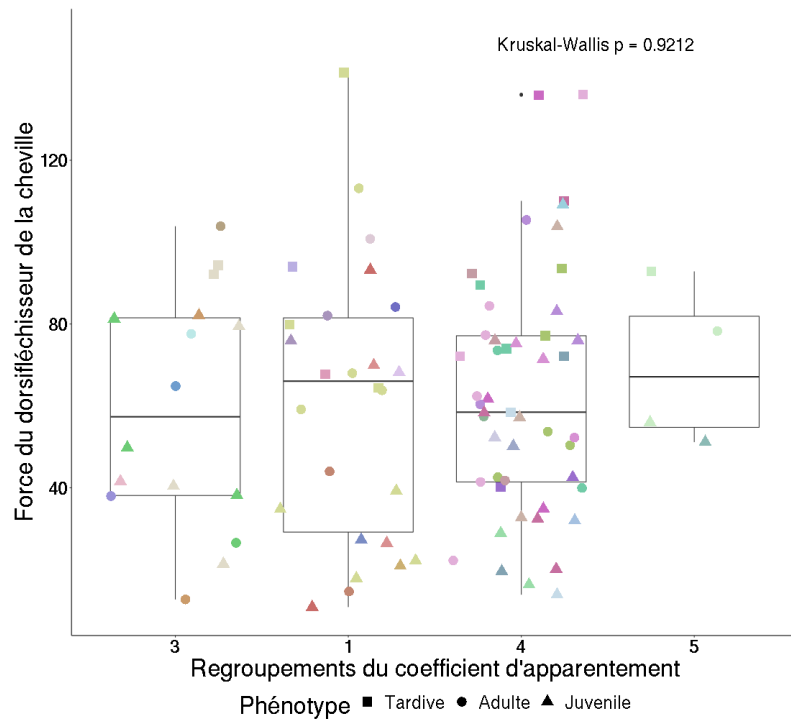


Figure 29 – Force du dorsifléchisseur de la cheville en fonction de 4 sous-regroupements du coefficient d'apparement.

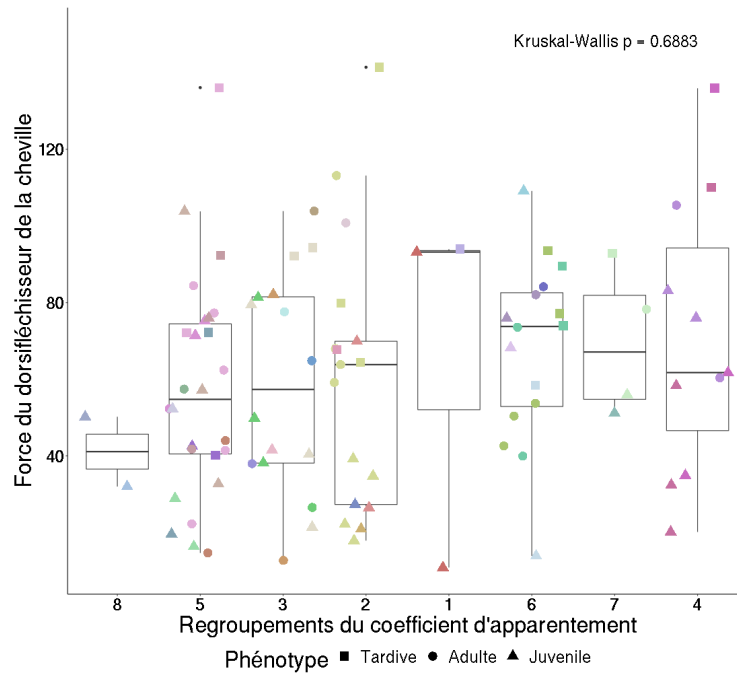


Figure 30 – Force du dorsifléchisseur de la cheville en fonction de 8 sous-regroupements du coefficient d'apparement.

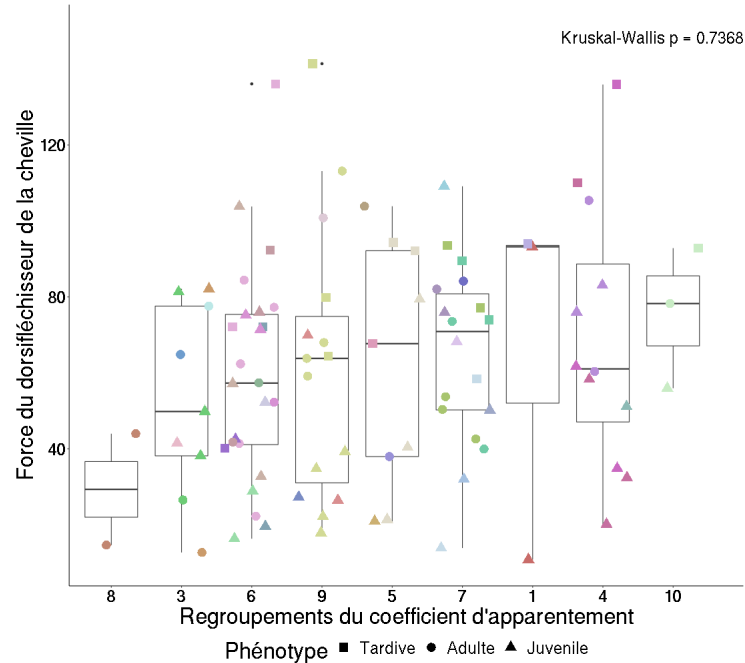


Figure 31 – Force du dorsifléchisseur de la cheville en fonction de 10 sous-regroupements du coefficient d'apparement.

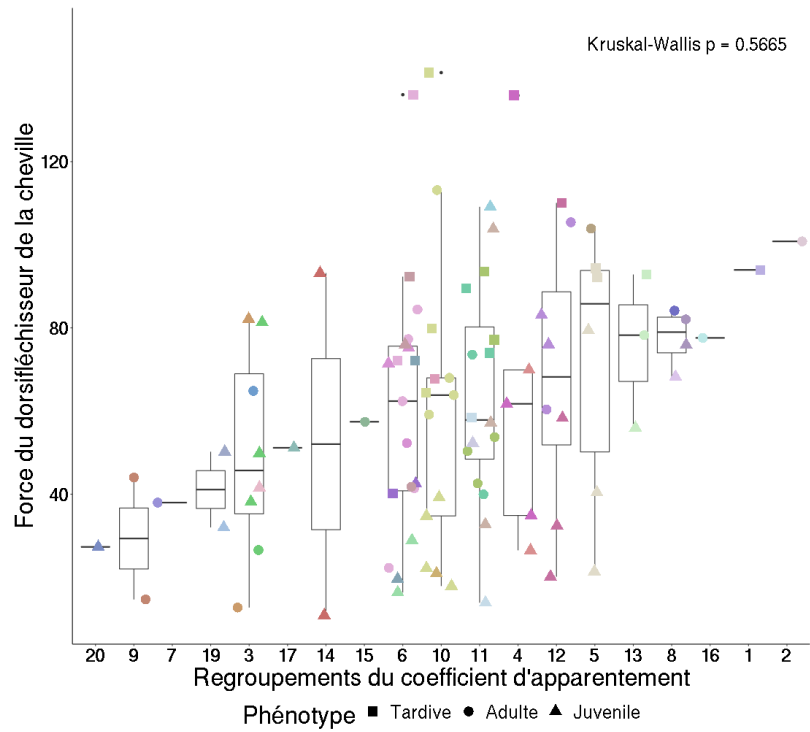


Figure 32 – Force du dorsifléchisseur de la cheville en fonction de 15 sous-regroupements du coefficient d'apparement.

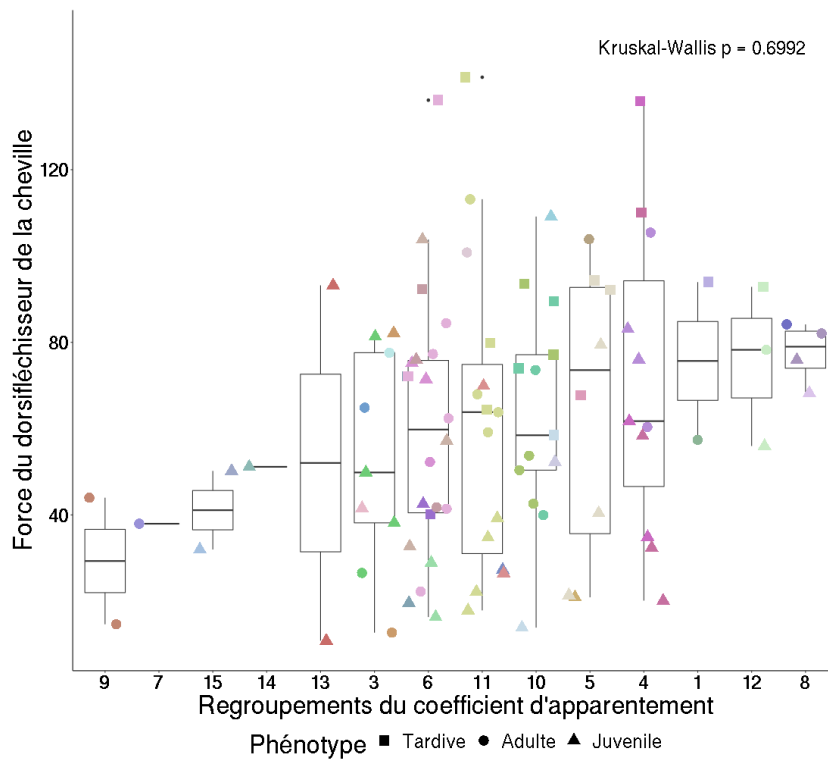


Figure 33 – Force du dorsifléchisseur de la cheville en fonction de 20 sous-regroupements du coefficient d'apparement.

