



**HAL**  
open science

# Évaluation d'un haplotype associé à la duplication en cis du gène SMN1 : utilité pour le conseil génétique de l'amyotrophie spinale proximale

Ferdi Kundul

## ► To cite this version:

Ferdi Kundul. Évaluation d'un haplotype associé à la duplication en cis du gène SMN1 : utilité pour le conseil génétique de l'amyotrophie spinale proximale. Médecine humaine et pathologie. 2020. dumas-02949211

**HAL Id: dumas-02949211**

**<https://dumas.ccsd.cnrs.fr/dumas-02949211>**

Submitted on 25 Sep 2020

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

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

Université de Rouen

Faculté de Médecine et de Pharmacie

Evaluation d'un haplotype associé à la  
duplication en *cis* du gène *SMN1* : utilité  
pour le conseil génétique de  
l'amyotrophie spinale proximale

Par Ferdi Kundul

Thèse de doctorat de Médecine

Présentée et soutenue publiquement le -- avril 2020

Devant un jury composé de :

- Professeur Thierry Frébourg, PU-PH, Président du Jury

- Professeur Loïc Favennec, PU-PH

- Docteur Gaël Nicolas, MCU-PH

- Docteur Pascale Saugier-Veber, MCU-PH, Directrice de thèse

**ANNEE UNIVERSITAIRE 2019 - 2020**

**U.F.R. SANTÉ DE ROUEN**

-----

DOYEN : **Professeur Benoît VEBER**

ASSESSEURS : **Professeur Michel GUERBET**  
**Professeur Agnès LIARD**  
**Professeur Guillaume SAVOYE**

**I - MEDECINE**

**PROFESSEURS DES UNIVERSITES – PRATICIENS HOSPITALIERS**

Mr Frédéric <b>ANSELME</b>	HCN	Cardiologie
Mme Gisèle <b>APTER</b>	Havre	Pédopsychiatrie
Mme Isabelle <b>AUQUIT AUCKBUR</b>	HCN	Chirurgie plastique
Mr Jean-Marc <b>BASTE</b>	HCN	Chirurgie Thoracique
Mr Fabrice <b>BAUER</b>	HCN	Cardiologie
Mme Soumeya <b>BEKRI</b>	HCN	Biochimie et biologie moléculaire
Mr Ygal <b>BENHAMOU</b>	HCN	Médecine interne
Mr Jacques <b>BENICHOU</b>	HCN	Bio statistiques et informatique médicale
Mr Olivier <b>BOYER</b>	UFR	Immunologie
Mme Sophie <b>CANDON</b>	HCN	Immunologie
Mr François <b>CARON</b>	HCN	Maladies infectieuses et tropicales
Mr Philippe <b>CHASSAGNE</b>	HCN	Médecine interne (gériatrie)
Mr Vincent <b>COMPERE</b>	HCN	Anesthésiologie et réanimation chirurgicale
Mr Jean-Nicolas <b>CORNU</b>	HCN	Urologie
Mr Antoine <b>CUVELIER</b>	HB	Pneumologie
Mr Jean-Nicolas <b>DACHER</b>	HCN	Radiologie et imagerie médicale
Mr Stéfan <b>DARMONI</b>	HCN	Informatique médicale et techniques de communication
Mr Pierre <b>DECHELOTTE</b>	HCN	Nutrition

Mr Stéphane <b>DERREY</b>	HCN	Neurochirurgie
Mr Frédéric <b>DI FIORE</b>	CHB	Cancérologie
Mr Fabien <b>DOGUET</b>	HCN	Chirurgie Cardio Vasculaire
Mr Jean <b>DOUCET</b>	SJ	T thérapeutique - Médecine interne et gériatrie
Mr Bernard <b>DUBRAY</b>	CHB	Radiothérapie
Mr Frank <b>DUJARDIN</b>	HCN	Chirurgie orthopédique - T traumatologique
Mr Fabrice <b>DUPARC</b>	HCN	Anatomie - Chirurgie orthopédique et traumatologique
Mr Eric <b>DURAND</b>	HCN	Cardiologie
Mr Bertrand <b>DUREUIL</b>	HCN	Anesthésiologie et réanimation chirurgicale
Mme Hélène <b>ELTCHANINOFF</b>	HCN	Cardiologie
Mr Manuel <b>ETIENNE</b>	HCN	Maladies infectieuses et tropicales
Mr Thierry <b>FREBOURG</b>	UFR	Génétique
Mr Pierre <b>FREGER</b> ( <i>sumombre</i> )	HCN	Anatomie - Neurochirurgie
Mr Jean François <b>GEHANNO</b>	HCN	Médecine et santé au travail
Mr Emmanuel <b>GERARDIN</b>	HCN	Imagerie médicale
Mme Priscille <b>GERARDIN</b>	HCN	Pédopsychiatrie
M. Guillaume <b>GOURCEROL</b>	HCN	Physiologie
Mr Dominique <b>GUERROT</b>	HCN	Néphrologie
Mr Olivier <b>GUILLIN</b>	HCN	Psychiatrie Adultes
Mr Claude <b>HOUDAYER</b>	HCN	Génétique
Mr Fabrice <b>JARDIN</b>	CHB	Hématologie
Mr Luc-Marie <b>JOLY</b>	HCN	Médecine d'urgence
Mr Pascal <b>JOLY</b>	HCN	Dermato –Vénéréologie
Mme Bouchra <b>LAMIA</b>	Havre	Pneumologie
Mme Annie <b>LAQUERRIERE</b>	HCN	Anatomie et cytologie pathologiques
Mr Vincent <b>LAUDENBACH</b>	HCN	Anesthésie et réanimation chirurgicale
Mr Joël <b>LECHEVALLIER</b>	HCN	Chirurgie infantile
Mr Hervé <b>LEFEBVRE</b>	HB	Endocrinologie et maladies métaboliques
Mr Thierry <b>LEQUERRE</b>	HCN	Rhumatologie
Mme Anne-Marie <b>LEROI</b>	HCN	Physiologie
Mr Hervé <b>LEVESQUE</b>	HCN	Médecine interne
Mme Agnès <b>LIARD-ZMUDA</b>	HCN	Chirurgie Infantile
Mr Pierre Yves <b>LITZLER</b>	HCN	Chirurgie cardiaque
Mr Bertrand <b>MACE</b>	HCN	Histologie, embryologie, cytogénétique
M. David <b>MALTETE</b>	HCN	Neurologie

Mr Christophe <b>MARGUET</b>	HCN	Pédiatrie
Mme Isabelle <b>MARIE</b>	HCN	Médecine interne
Mr Jean-Paul <b>MARIE</b>	HCN	Oto-rhino-laryngologie
Mr Loïc <b>MARPEAU</b>	HCN	Gynécologie - Obstétrique
Mr Stéphane <b>MARRET</b>	HCN	Pédiatrie
Mme Véronique <b>MERLE</b>	HCN	Epidémiologie
Mr Pierre <b>MICHEL</b>	HCN	Hépatogastro-entérologie
M. Benoit <b>MISSET</b> ( <i>détachement</i> )	HCN	Réanimation Médicale
Mr Marc <b>MURAINÉ</b>	HCN	Ophtalmologie
Mr Christophe <b>PEILLON</b>	HCN	Chirurgie générale
Mr Christian <b>PFISTER</b>	HCN	Urologie
Mr Jean-Christophe <b>PLANTIER</b>	HCN	Bactériologie - Virologie
Mr Didier <b>PLISSONNIER</b>	HCN	Chirurgie vasculaire
Mr Gaëtan <b>PREVOST</b>	HCN	Endocrinologie
Mr Jean-Christophe <b>RICHARD</b> ( <i>détachement</i> )	HCN	Réanimation médicale - Médecine d'urgence
Mr Vincent <b>RICHARD</b>	UFR	Pharmacologie
Mme Nathalie <b>RIVES</b>	HCN	Biologie du développement et de la reproduction
Mr Horace <b>ROMAN</b> ( <i>disponibilité</i> )	HCN	Gynécologie - Obstétrique
Mr Jean-Christophe <b>SABOURIN</b>	HCN	Anatomie – Pathologie
Mr Mathieu <b>SALAUN</b>	HCN	Pneumologie
Mr Guillaume <b>SAVOYE</b>	HCN	Hépatogastrologie
Mme Céline <b>SAVOYE-COLLET</b>	HCN	Imagerie médicale
Mme Pascale <b>SCHNEIDER</b>	HCN	Pédiatrie
Mr Lilian <b>SCHWARZ</b>	HCN	Chirurgie Viscérale et Digestive
Mr Michel <b>SCOTTE</b>	HCN	Chirurgie digestive
Mme Fabienne <b>TAMION</b>	HCN	Thérapeutique
Mr Luc <b>THIBERVILLE</b>	HCN	Pneumologie
Mr Hervé <b>TILLY</b> ( <i>surnombre</i> )	CHB	Hématologie et transfusion
M. Gilles <b>TOURNEL</b>	HCN	Médecine Légale
Mr Olivier <b>TROST</b>	HCN	Anatomie -Chirurgie Maxillo-Faciale
Mr Jean-Jacques <b>TUECH</b>	HCN	Chirurgie digestive
Mr Benoît <b>VEBER</b>	HCN	Anesthésiologie - Réanimation chirurgicale
Mr Pierre <b>VERA</b>	CHB	Biophysique et traitement de l'image
Mr Eric <b>VERIN</b>	Les Herbiers	Médecine Physique et de Réadaptation
Mr Eric <b>VERSPYCK</b>	HCN	Gynécologie obstétrique

Mr Olivier <b>MITTECOQ</b>	HC	Rhumatologie
Mme Marie-Laure <b>WELTER</b>	HCN	Physiologie

### **MAITRES DE CONFERENCES DES UNIVERSITES – PRATICIENS HOSPITALIERS**

Mme Najate <b>ACHAMRAH</b>	HCN	Nutrition
Mme Noëlle <b>BARBIER-FREBOURG</b>	HCN	Bactériologie – Virologie
Mr Emmanuel <b>BESNIER</b>	HCN	Anesthésiologie - Réanimation
Mme Carole <b>BRASSE LAGNEL</b>	HCN	Biochimie
Mme Valérie <b>BRIDOUX HUYBRECHTS</b>	HCN	Chirurgie Vasculaire
Mr Gérard <b>BUCHONNET</b>	HCN	Hématologie
Mme Mireille <b>CASTANET</b>	HCN	Pédiatrie
Mme Nathalie <b>CHASTAN</b>	HCN	Neurophysiologie
Mr Moïse <b>COEFFIER</b>	HCN	Nutrition
Mr Serge <b>JACQUOT</b>	UFR	Immunologie
Mr Joël <b>LADNER</b>	HCN	Epidémiologie, économie de la santé
Mr Jean-Baptiste <b>LATOUCHE</b>	UFR	Biologie cellulaire
M. Florent <b>MARGUET</b>	HCN	Histologie
Mme Chloé <b>MELCHIOR</b>	HCN	Gastroentérologie
Mr Thomas <b>MOUREZ</b> ( <i>détachement</i> )	HCN	Virologie
Mr Gaël <b>NICOLAS</b>	UFR	Génétique
Mme Muriel <b>QUILLARD</b>	HCN	Biochimie et biologie moléculaire
Mme Laëtitia <b>ROLLIN</b>	HCN	Médecine du Travail
Mme Pascale <b>SAUGIER-VEBER</b>	HCN	Génétique
Mme Anne-Claire <b>TOBENAS-DUJARDIN</b>	HCN	Anatomie
Mr David <b>WALLON</b>	HCN	Neurologie
Mr Julien <b>WILS</b>	HCN	Pharmacologie

### **PROFESSEUR AGREGE OU CERTIFIE**

Mr Thierry <b>WABLE</b>	UFR	Communication
Mme Mélanie <b>AUVRAY-HAMEL</b>	UFR	Anglais

## II - PHARMACIE

### PROFESSEURS DES UNIVERSITES

Mr Jérémy <b>BELLIEN</b> (PU-PH)	Pharmacologie
Mr Thierry <b>BESSON</b>	Chimie Thérapeutique
Mr Jean <b>COSTENTIN</b> (Professeur émérite)	Pharmacologie
Mme Isabelle <b>DUBUS</b>	Biochimie
Mr Abdelhakim <b>EL OMRI</b>	Pharmacognosie
Mr François <b>ESTOUR</b>	Chimie Organique
Mr Loïc <b>FAVENNEC</b> (PU-PH)	Parasitologie
Mr Jean Pierre <b>GOULLE</b> (Professeur émérite)	Toxicologie
Mr Michel <b>GUERBET</b>	Toxicologie
Mme Christelle <b>MONTEIL</b>	Toxicologie
Mme Martine <b>PESTEL-CARON</b> (PU-PH)	Microbiologie
Mr Rémi <b>VARIN</b> (PU-PH)	Pharmacie clinique
Mr Jean-Marie <b>VAUGEOIS</b>	Pharmacologie
Mr Philippe <b>VERITE</b>	Chimie analytique

### MAITRES DE CONFERENCES DES UNIVERSITES

Mme Cécile <b>BARBOT</b>	Chimie Générale et Minérale
Mr Frédéric <b>BOUNOURE</b>	Pharmacie Galénique
Mr Thomas <b>CASTANHEIRO MATIAS</b>	Chimie Organique
Mr Abdeslam <b>CHAGRAOUI</b>	Physiologie
Mme Camille <b>CHARBONNIER (LE CLEZIO)</b>	Statistiques
Mme Elizabeth <b>CHOSSON</b>	Botanique
Mme Marie Catherine <b>CONCE-CHEMTOB</b>	Législation pharmaceutique et économie de la santé
Mme Cécile <b>CORBIERE</b>	Biochimie
Mme Nathalie <b>DOURMAP</b>	Pharmacologie
Mme Isabelle <b>DUBUC</b>	Pharmacologie
Mme Dominique <b>DUTERTE- BOUCHER</b>	Pharmacologie
Mr Gilles <b>GARGALA</b> (MCU-PH)	Parasitologie

Mme Nejla EL <b>GHARBI-HAMZA</b>	Chimie analytique
Mme Marie-Laure <b>GROULT</b>	Botanique
Mr Hervé <b>HUE</b>	Biophysique et mathématiques
Mme Hong <b>LU</b>	Biologie
Mme Marine <b>MALLETER</b>	Toxicologie
M. Jérémie <b>MARTINET</b> (MCU-PH)	Immunologie
Mme Tiphaine <b>ROGEZ-FLORENT</b>	Chimie analytique
Mr Mohamed <b>SKIBA</b>	Pharmacie galénique
Mme Malika <b>SKIBA</b>	Pharmacie galénique
Mme Christine <b>THARASSE</b>	Chimie thérapeutique
Mr Frédéric <b>ZIEGLER</b>	Biochimie

### **PROFESSEURS ASSOCIES**

Mme Cécile <b>GUERARD-DETUNCQ</b>	Pharmacie officinale
Mme Caroline <b>BERTOUX</b>	Pharmacie

### **PAU-PH**

M. Mikaël **DAOUPHARS**

### **PROFESSEUR CERTIFIE**

Mme Mathilde <b>GUERIN</b>	Anglais
----------------------------	---------

### **ASSISTANTS HOSPITALO-UNIVERSITAIRES**

Mme Alice <b>MOISAN</b>	Virologie
M. Henri <b>GONDÉ</b>	Pharmacie

### **ATTACHES TEMPORAIRES D'ENSEIGNEMENT ET DE RECHERCHE**

M. Abdel <b>MOUHAJIR</b>	Parasitologie
M. Maxime <b>GRAND</b>	Bactériologie

### **ATTACHE TEMPORAIRE D'ENSEIGNEMENT**

Mme Ramla <b>SALHI</b>	Pharmacognosie
------------------------	----------------



## LISTE DES RESPONSABLES DES DISCIPLINES PHARMACEUTIQUES

Mme Cécile <b>BARBOT</b>	Chimie Générale et minérale
Mr Thierry <b>BESSON</b>	Chimie thérapeutique
Mr Abdeslam <b>CHAGRAOUI</b>	Physiologie
Mme Elisabeth <b>CHOSSON</b>	Botanique
Mme Marie-Catherine <b>CONCE-CHEMTOB</b>	Législation et économie de la santé
Mme Isabelle <b>DUBUS</b>	Biochimie
Mr Abdelhakim <b>ELOMRI</b>	Pharmacognosie
Mr François <b>ESTOUR</b>	Chimie organique
Mr Loïc <b>FAVENNEC</b>	Parasitologie
Mr Michel <b>GUERBET</b>	Toxicologie
Mme Martine <b>PESTEL-CARON</b>	Microbiologie
Mr Mohamed <b>SKIBA</b>	Pharmacie galénique
Mr Rémi <b>VARIN</b>	Pharmacie clinique
M. Jean-Marie <b>VAUGEOIS</b>	Pharmacologie
Mr Philippe <b>VERITE</b>	Chimie analytique



## ENSEIGNANTS MONO-APPARTENANTS

### PROFESSEURS

Mr Paul <b>MULDER</b> (phar)	Sciences du Médicament
Mme Su <b>RUAN</b> (med)	Génie Informatique

### MAITRES DE CONFERENCES

Mr Sahil <b>ADRIOUCH</b> (med)	Biochimie et biologie moléculaire(Unité Inserm 905)
Mme Gaëlle <b>BOUGEARD-DENOYELLE</b> (med)	Biochimie et biologie moléculaire(UMR 1079)
Mme Carine <b>CLEREN</b> (med)	Neurosciences (Néovasc)
M. Sylvain <b>FRAINEAU</b> (med)	Physiologie (Inserm U 1096)
Mme Pascaline <b>GAILDRAT</b> (med)	Génétique moléculaire humaine (UMR 1079)
Mr Nicolas <b>GUEROUT</b> (med)	Chirurgie Expérimentale
Mme Rachel <b>LETELLIER</b> (med)	Physiologie
Mr Antoine <b>OVRARD-PASCAUD</b> (med)	Physiologie (Unité Inserm 1076)
Mr Frédéric <b>PASQUET</b>	Sciences du langage, orthophonie
Mme Christine <b>RONDANINO</b> (med)	Physiologie de la reproduction
Mr Youssan Var <b>TAN</b>	Immunologie
Mme Isabelle <b>TOURNIER</b> (med)	Biochimie (UMR 1079)

### **CHEF DES SERVICES ADMINISTRATIFS : Mme Véronique DELAFONTAINE**

*HCN - Hôpital Charles Nicolle*

*HB - Hôpital de BOIS GUILLAUME*

*CB - Centre Henri Becquerel*

*CHS - Centre Hospitalier Spécialisé du Rouvray*

*CRMPR - Centre Régional de Médecine Physique et de Réadaptation SJ – Saint Julien Rouen*

**L'université de Rouen et l'UFR de médecine et de pharmacie de Rouen n'entendent donner aucune approbation ni improbation aux opinions émises dans cette thèse. Ces opinions sont propres à leurs auteurs.**

# Evaluation d'un haplotype associé à la duplication en *cis* du gène *SMN1*: utilité pour le conseil génétique de l'amyotrophie spinale proximale

## Table des matières

Evaluation d'un haplotype associé à la duplication en <i>cis</i> du gène <i>SMN1</i> : utilité pour le conseil génétique de l'amyotrophie spinale proximale.....	12
Liste des figures.....	14
Liste des tableaux.....	16
Liste des abréviations.....	17
Introduction.....	18
Contexte de l'étude.....	19
Généralités sur l'amyotrophie spinale proximale.....	19
Le conseil génétique de la SMA.....	29
Identification d'un haplotype associé à la duplication en <i>cis</i> du gène <i>SMN1</i> .....	33
Objectif de l'étude.....	37
Matériel et méthodes.....	38
Constitution des sous-groupes analysés.....	38
Méthodes.....	41
Résultats.....	59
Proportion de variants dans chaque sous-groupe.....	59
Absence d'association des variants recherchés à la duplication en <i>cis</i> du gène <i>SMN2</i> .....	60
Performances du test.....	60
Discussion et perspectives.....	62

Technologies alternatives à la recherche de l'haplotype.....	65
Conclusion.....	76
Bibliographie.....	77
SERMENT D'HIPPOCRATE.....	82
Résumé.....	83
<i>Abstract</i> .....	84

## Liste des figures

FIGURE 1 : ÉLÉMENT DE 500 KB DUPLIQUÉ-INVERSÉ .....	24
FIGURE 2 : RÔLE DE LA PROTÉINE SMN DANS LA FORMATION DES COMPLEXES SNRNP .....	26
FIGURE 3 : GÉNOTYPES DES PATIENTS ATTEINTS DE SMA ET DES SUJETS HÉTÉROZYGOTES (AZAD ET AL., 2019) .....	28
FIGURE 4 : EXEMPLE D'UN PARENT DE CAS-INDEX 2T AVEC GRANDS-PARENTS 1T ET 3T.....	32
FIGURE 5 : EXEMPLE D'UN SUJET 2+0 AYANT DES ENFANTS 3T ET 1T.....	32
FIGURE 6 : EXEMPLE D'UN PARENT DE CAS-INDEX 2T AVEC GRANDS-PARENTS 1T ET 3T.....	39
FIGURE 7 : EXEMPLE D'UN SUJET 2+0 AYANT DES ENFANTS 3T ET 1T.....	39
FIGURE 8 : PRINCIPE DE LA QMPSF .....	42
FIGURE 9 : LIMITATION DU NOMBRE DE CYCLES DE PCR.....	42
FIGURE 10 : ETAPE DE DIGESTION PAR L'ENZYME <i>DRAI</i> .....	45
FIGURE 11 : ETAPE DE DIGESTION PAR L'ENZYME <i>HINF</i> I.....	45
FIGURE 12 : EXEMPLES D'ÉLECTROPHÉROGRAMMES DE QMPSF <i>DRAI</i> ET <i>HINF</i> I .....	47
FIGURE 13 : BORNES DE L'AMPLICON DE SMN1 GÉNÉRÉ ET LOCALISATION DES VARIANTS D'INTÉRÊT .....	48
FIGURE 14 : AMPLICONS GÉNÉRÉS.....	50
FIGURE 15 : EXTENSION DES AMORCES DE SNAPSHOT.....	53
FIGURE 16 : EXEMPLES D'ÉLECTROPHÉROGRAMMES DE SNAPSHOT.....	54
FIGURE 17 : EXEMPLES DE RÉSULTATS ISSU DU KIT AMPLIDEX PCR/CE <i>SMN1/2</i> .....	56
FIGURE 18 : ASPECT D'UNE LAME DE PEIGNAGE MOLÉCULAIRE.....	67
FIGURE 19 : EXEMPLE D'IMAGE DE MORSE CODE OBTENUE SUR LE GÈNE <i>BRCA1</i> .....	68
FIGURE 20 : INTERPRÉTATION DES MORSE CODES OBTENUS.....	69
FIGURE 21 : LES 5 ÉTAPES DU PROTOCOLE DE PEIGNAGE MOLÉCULAIRE.....	69
FIGURE 22 : ETAPE DE PRÉPARATION DES ÉCHANTILLONS AVANT L'ANALYSE OPTIQUE DES DONNÉES (BARSEGHYAN, ET AL., 2018).....	70
FIGURE 23 : PUCE SAPHYR CHIP.....	71
FIGURE 24 : LINÉARISATION DES FRAGMENTS D'ADNG.....	71
FIGURE 25 : ETAPE D'ALIGNEMENT ET DE RECONSTITUTION D'UNE CARTOGRAPHIE CONSENSUS (CHAN ET AL., 2018).....	72

FIGURE 26 : REMANIEMENTS STRUCTURAUX DÉTECTABLES PAR CARTOGRAPHIE OPTIQUE BIONANO (CHAN <i>ET AL.</i> , 2018).....	73
FIGURE 27 : PRINCIPE DU SÉQUENÇAGE NANOPORE (DEAMER <i>ET AL.</i> , 2016) .....	74
FIGURE 28 : <i>FLOW-CELL</i> NANOPORE.....	74



## Liste des tableaux

TABLEAU 1 : CLASSIFICATION CLINIQUE DES AMYOTROPHIES SPINALES PROXIMALES (WANG <i>ET AL.</i> , 2007; GROTTO <i>ET AL.</i> , 2016).....	20
TABLEAU 2 : GÉNOTYPES AU LOCUS <i>SMN1</i> DES PATIENTS ATTEINTS DE SMA .....	27
TABLEAU 3 : FRÉQUENCES ALLÉLIQUES DU NOMBRE DE COPIES DU GÈNE <i>SMN1</i> PAR ALLÈLE DANS DIFFÉRENTES SOUS-POPULATIONS (HENDRICKSON <i>ET AL.</i> , 2009).....	30
TABLEAU 4 : FRÉQUENCE DES GÉNOTYPES DANS DIFFÉRENTES SOUS-POPULATIONS (HENDRICKSON <i>ET AL.</i> , 2009).....	31
TABLEAU 5 : HAPLOTYPES DE MICROSATELLITES PAR GÉNOTYPE AU LOCUS <i>SMN1</i> (LUO <i>ET AL.</i> , 2014) .....	34
TABLEAU 6 : DU VARIANT C.*3+80T>G DANS D'AUTRES POPULATIONS (LUO <i>ET AL.</i> , 2014).....	35
TABLEAU 7 : PRÉSENCE DES VARIANTS D'INTÉRÊT SELON LE GÉNOTYPE (ALÍAS <i>ET AL.</i> , 2018) .....	36
TABLEAU 8 : QMPSF <i>DRAI</i> .....	44
TABLEAU 9 : QMPSF <i>HINFI</i> .....	44
TABLEAU 10 : AMORCES DE PCR DE SNAPSHOT .....	49
TABLEAU 11 : CARACTÉRISTIQUES DES AMORCES DE SNAPSHOT CONÇUES .....	52
TABLEAU 12 : RÉSULTATS DE RECHERCHE DES VARIANTS SELON LE GÉNOTYPE .....	59
TABLEAU 13 : RÉSULTATS DE RECHERCHE DES VARIANTS SELON LA PRÉSENCE OU L'ABSENCE D'UNE DUPLICATION EN <i>CIS</i> DU GÈNE <i>SMN1</i> .....	60
TABLEAU 14 : SPÉCIFICITÉ DE L'ASSOCIATION DES VARIANTS RECHERCHÉS À LA DUPLICATION EN <i>CIS</i> DU GÈNE <i>SMN1</i> .....	60
TABLEAU 15 : AMÉLIORATION DU TAUX DE DÉTECTION ET DIMINUTION DU RISQUE RÉSIDUEL EN CAS D'HAPLOTYPE NÉGATIF .....	61

## Liste des abréviations

ADN Acide désoxyribonucléique

ARN Acide ribonucléique

CNV *Copy Number Variation*

CPK Créatine phosphokinase

ddNTP Didésoxyribonucléotide

dNTP Désoxyribonucléotide

ENMG Électroneuromyogramme

FN Faux négatif

FP Faux positif

Kb Kilobases

MLPA *Multiplex Ligation-Dependent Probe Amplification*

NGS *Next Generation Sequencing*

nt nucléotides

pb Paires de bases

QMPSF *Quantitative Multiplex PCR of Short Fragments*

Se Sensibilité

SMA *Spinal Muscular Atrophy*

SNV Single Nucleotide Variant

Sp Spécificité

VN Vrai négatif

VP Vrai positif

VPN Valeur prédictive négative

VPP Valeur prédictive positive

## Introduction

L'amyotrophie spinale proximale (SMA) est l'une des maladies de transmission autosomique récessive les plus fréquentes, avec une incidence de 1/6000 à 1/10 000 naissances vivantes et une prévalence des porteurs hétérozygotes de 1/35 à 1/117 selon la population considérée (Alías *et al.*, 2014; Hendrickson *et al.*, 2009; Sugarman *et al.*, 2012). Cette maladie se caractérise par une dégénérescence progressive des motoneurones alpha de la corne antérieure de la moelle épinière, à l'origine d'une faiblesse musculaire et d'une atrophie symétrique et progressive touchant préférentiellement les muscles des membres inférieurs et les muscles ventilatoires. Les SMA sont classées en 5 sous-types cliniques, du type 0 au type IV, en fonction de l'âge de début et de la sévérité de la maladie, les formes les plus graves débutant *in utero* (type 0, récemment décrit) (Grotto *et al.*, 2016), alors que les formes de meilleur pronostic (type IV) débutent plus tardivement à l'âge adulte et sont d'évolution plus lente.

La cause la plus fréquente de la SMA est la délétion à l'état homozygote du gène *SMN1*, situé en 5q11q13 (Lefebvre *et al.*, 1995). Chez 95 % des patients, les deux copies de l'exon 7 du gène *SMN1* sont délétées ou rendues non fonctionnelles par une conversion génique *SMN1-SMN2* (Alías *et al.*, 2009; Cuscó *et al.*, 2001; Hahnen *et al.*, 1996; Wirth 2000; Wirth *et al.*, 1997, 1999). Chez 5 % des patients, une délétion de l'exon 7 du gène *SMN1* est associée à une mutation ponctuelle en *trans* (Alías *et al.*, 2009; Chen *et al.*, 1999; Wirth 2000). Un gène copie hautement homologue, *SMN2*, a été décrit et n'est que partiellement fonctionnel.

Le conseil génétique dans cette maladie repose sur la mise en évidence, chez des apparentés, de variations délétères présentes à l'état hétérozygote, comme une délétion du gène *SMN1* ou une mutation ponctuelle. Toutefois, le nombre de copies du gène *SMN1* sur un même chromosome 5 est variable selon les individus, avec notamment des allèles porteurs de deux copies en *cis* du gène *SMN1*, soit sur le même chromosome 5. La ségrégation au sein d'une même famille d'une duplication en *cis* et d'une délétion entraîne un risque de génotypes dits "2+0", c'est à dire avec 2 copies du gène *SMN1* en *cis* et 0 copie sur le deuxième chromosome, avec par conséquent un risque de transmettre

la maladie. Etant donné que les techniques de dépistage des hétérozygotes permettent de quantifier le nombre de copies du gène *SMN1* mais pas de connaître la phase des copies de ce gène, le génotype 2+0 constitue un piège majeur du conseil génétique dans la SMA puisqu'il n'est pas distinguable du génotype 1+1, c'est à dire avec une copie du gène *SMN1* sur chaque chromosome (2 copies en *trans*).

Deux variations ont été récemment rapportées comme associées à la duplication en *cis* du gène *SMN1* dans la population juive ashkénaze (Luo *et al.*, 2014). Cette association dans la population juive ashkénaze a été confirmée par une seconde étude menée dans la population espagnole (Alías *et al.*, 2018). La sensibilité de cette association était évaluée à 20 % et la spécificité 99 %.

L'objet de ce travail est d'évaluer la pertinence de l'analyse de ces 2 variations en complément de la quantification du gène *SMN1* pour ajouter de la fiabilité au conseil génétique en permettant de dépister ces génotypes 2+0.

## **Contexte de l'étude**

### **Généralités sur l'amyotrophie spinale proximale**

L'amyotrophie spinale proximale (*Spinal Muscular Atrophy* en anglais, ou SMA) constitue la deuxième maladie autosomique récessive létale la plus fréquente après la mucoviscidose, avec une incidence de 1/6000 à 1/10000 nouveaux cas par an en France. La SMA se caractérise par une dégénérescence des motoneurones alpha de la corne antérieure de la moelle épinière. La fréquence des hétérozygotes dans la population générale est de 1/35 à 1/117 selon la population considérée (Alías *et al.*, 2014; Hendrickson *et al.*, 2009; Ogino *et al.*, 2002; Prior *et al.*, 2010; Sugarman *et al.*, 2012). La SMA a été décrite pour la première fois dans les années 1890 par Guido Werdnig et Johann Hoffmann (Hoffmann, 1900, 1893; Werdnig, 1894, 1891). À quelques années d'intervalle, ces neurologues autrichien et allemand ont observé, chez plusieurs enfants de mêmes fratries, une atteinte neurologique associant une hypotonie et une faiblesse musculaire progressives, débutant

avant l'âge d'un an et entraînant un décès précoce dans l'enfance suite à des complications respiratoires.

Par la suite, le spectre clinique de la SMA a été étendu aux formes modérées de l'adulte (Kugelberg and Welander, 1956).

### *Présentations cliniques*

La SMA est classiquement divisée en cinq types en fonction de l'âge de début et de la sévérité de la maladie. La première classification, établie au début des années 1990 par Munsat et Davis, était basée sur l'âge d'apparition des premiers signes de la maladie (Munsat et Davies, 1992). Une seconde classification, basée sur des caractères fonctionnels, a été proposée par Wang *et al.* en 2007. Cette classification, plus objective, repose sur les dernières acquisitions motrices de l'enfant (Wang *et al.*, 2007). Une forme particulièrement sévère fœtale se compliquant de décès en période néonatale, a été décrite en 2016 (Grotto *et al.*, 2016). Le type I et le type II représentent les formes les plus fréquentes de SMA (Ogino *et al.*, 2004) (Tableau 1).

Type de SMA	Âge d'apparition des symptômes	Motricité du patient	Espérance de vie
Type 0	anténatal	Diminution des mouvements actifs fœtaux	< 1 mois
Type I Maladie de Werdnig-Hoffmann	0-6 mois	Incapable de s'asseoir	< 2 ans
Type II	7-18 mois	Incapable de se mettre debout	> 2 ans
Type III Maladie de Kugelberg-Welander	> 18 mois	Capable de se mettre debout et de marcher	Adulte
Type IV	2ème-3ème décade	Marche préservée	Adulte

Tableau 1 : Classification clinique des amyotrophies spinales proximales (Wang *et al.*, 2007; Grotto *et al.*, 2016)

En réalité, la présentation clinique associée à la SMA est un continuum et il existe des chevauchements phénotypiques entre différents types. La rapidité d'évolution et l'espérance de vie sont corrélées à l'âge de début : plus la maladie débute à un âge précoce, plus l'évolution est rapide (Bertini *et al.*, 2005).

La SMA de type 0 se caractérise par un début *in utero* marqué par une diminution des mouvements actifs fœtaux, une augmentation de la clarté nucale, des malformations cardiaques congénitales, un polyhydramnios et des contractures articulaires, s'intégrant dans une séquence malformative d'akinésie fœtale. Un tableau clinique sévère se poursuit en période néonatale avec une hypotonie majeure, une détresse ventilatoire immédiate, une atteinte des paires crâniennes et des épisodes de bradycardie récurrente. Le décès survient généralement dans le premier mois de vie.

La SMA de type I ou Maladie de Werdnig-Hoffmann est souvent évoquée devant une hypotonie du tronc et une absence d'acquisition ou une perte précoce de la tenue de la tête, secondaire à une atteinte des muscles paraspinaux. Les premiers signes, plus discrets, débutent avant l'âge de 6 mois par une hypotonie, initialement au niveau des racines et prédominant aux membres inférieurs. Cette atteinte, symétrique, évolue vers une quadriparésie flasque. Les mouvements des extrémités sont longtemps préservés. L'examen met en évidence une abolition des réflexes ostéotendineux. Des fasciculations des mains et de la langue peuvent être observées. L'atteinte des muscles intercostaux entraîne une insuffisance respiratoire progressive. La toux est inefficace. Le diaphragme étant préservé, une respiration abdominale paradoxale et une déformation du thorax en carène sont observées. Il peut exister une atteinte du tronc cérébral, principalement des paires crâniennes et du bulbe. L'atteinte des nerfs crâniens est à l'origine de troubles de la déglutition, de troubles de la succion, d'une faiblesse du cri. La motricité faciale et l'oculomotricité sont longtemps préservées. L'atteinte des centres bulbaires est à l'origine de troubles des fonctions respiratoire et cardiaque prenant souvent l'aspect de troubles dysautonomiques, pouvant aller jusqu'à la mort subite. Il n'y a pas de trouble sensitif, ni de trouble sphinctérien, et classiquement pas d'atteinte des fonctions supérieures. L'évolution est marquée par une aggravation progressive. L'insuffisance respiratoire chronique, aggravée par les infections respiratoires et les fausses routes, entraînent un décès avant 2 ans, généralement avant 12 à 18 mois, en l'absence de traitement.

La SMA de type II est définie par une absence d'acquisition de la marche. Le développement de l'enfant est normal au cours des premiers mois de vie, avec une acquisition possible de la station

assise. La maladie débute entre 6 et 18 mois. Les premières manifestations peuvent être des difficultés à la station assise libérée, une absence d'acquisition de la station debout avec appui ou l'absence d'acquisition de la marche. Les réflexes ostéotendineux peuvent être initialement conservés, mais disparaissent au cours de l'évolution. La motricité des membres supérieurs est initialement conservée. Il existe des fasciculations des mains et de la langue. Les fonctions cognitives sont habituellement très bonnes, avec une acquisition précoce du langage. L'évolution est marquée par une aggravation globale du déficit musculaire et par la survenue de complications orthopédiques (scoliose sévère, rétractions articulaires) et respiratoires sévères (syndrome restrictif secondaire à l'atteinte des muscles intercostaux et abdominaux, difficultés à la toux secondaire à l'atteinte des muscles expirateurs et à une possible composante centrale par atteinte des centres bulbaires), des troubles de la déglutition, des difficultés de mastication, des troubles digestifs (dilatation gastrique aiguë, reflux gastro-œsophagien, constipation), des complications nutritionnelles et des troubles vasomoteurs. Avec une prise en charge intensive multidisciplinaire, la survie de ces enfants peut être prolongée jusqu'à l'âge adulte (Bertini *et al.*, 2005; Cuisset *et al.*, 2012)

La SMA de type III ou Maladie de Kugelberg-Welander débute après l'acquisition de la marche par une présentation d'allure myopathique. Les premiers signes peuvent se manifester par une gêne à la marche dès son acquisition, ou plus tardivement par des difficultés à courir ou à monter les escaliers et des chutes, ou encore par une gêne lors des épreuves sportives au cours de l'enfance. Le déficit et l'amyotrophie prédominent au niveau des racines et aux membres inférieurs. Les réflexes ostéotendineux sont conservés aux membres supérieurs. L'évolution peut être marquée par la survenue de complications orthopédiques (rétractions articulaires, scoliose) et respiratoires. La progression est plus lente que dans les types I et II. La sévérité est variable et le maintien de la marche à l'âge adulte est corrélé à l'âge de début de la maladie. Un meilleur pronostic est observé dans les formes débutant après l'âge de trois ans. L'espérance de vie n'est pas significativement différente de celle de la population générale (Bertini *et al.*, 2005; D'Amico *et al.*, 2011).

La SMA de type IV débute au cours de la troisième ou de la quatrième décennie et se manifeste par une gêne motrice proximale. La marche est conservée à l'âge adulte, il n'y a pas d'atteinte respiratoire (Bertini *et al.*, 2005; D'Amico *et al.*, 2011).

Le diagnostic de SMA est le plus souvent fortement évoqué sur la présentation clinique. Chez le petit enfant (types I et II), la symptomatologie fait évoquer une atteinte de la corne antérieure de la moelle. Chez le plus grand enfant ou l'adulte (type III et IV), la présentation est plutôt pseudo-myopathique. Parfois, les CPK modérément augmentées associées à ce tableau clinique peuvent faire évoquer une dystrophie musculaire. En cas de doute diagnostique, la confirmation, avant l'analyse moléculaire, peut être apportée par l'électroneuromyogramme (ENMG), montrant : un tracé neurogène, des vitesses de conduction motrices et sensitives normales ou légèrement diminuées (Hausmanowa-Petrusewicz, 1988), voire pour les types III par la biopsie musculaire, qui, demandée dans l'hypothèse d'une affection musculaire, objective des altérations neurogènes : atrophie des fibres de type 1 et 2, hypertrophie des fibres de type 1, phénomène de *grouping* neurogène (Castiglioni *et al.*, 2011).

### ***Bases moléculaires de la SMA***

Les présentations cliniques de la SMA sont très hétérogènes au regard de la sévérité et de l'évolution de la maladie. Il était difficile de concevoir qu'un seul défaut génétique puisse être à l'origine des différents types de SMA. Pourtant, en 1995, des altérations dans le gène *SMN1* (*Survival Motor Neuron 1*) ont été identifiées dans les 4 types de SMA (Lefebvre *et al.*, 1995).

Le gène *SMN1* est localisé en 5q11q13 dans une région de 500 Kb qui s'est dupliquée et inversée au cours de l'évolution. Situé dans l'élément dupliqué télomérique, il possède donc un gène copie dans l'élément dupliqué centromérique : le gène *SMN2*, partiellement fonctionnel (Figure 1). Le gène *SMN1* code pour la protéine SMN (*Survival Motor Neuron*).



La SMA résulte d'une inactivation bi-allélique du gène *SMN1*. La sévérité de la SMA est corrélée au taux de protéine SMN fonctionnelle produite, les formes les plus sévères étant associées à de faibles taux de protéine (Lefebvre *et al.*, 1997).



Figure 1 : Élément de 500 Kb dupliqué-inversé

Le gène *SMN2* constitue un gène-copie du gène *SMN1* au sein d'un élément de 500 Kb dupliqué-inversé situé à un locus plus proche du centromère que *SMN1*. Contrairement au gène *SMN1* qui code pour le transcrite pleine longueur, le gène *SMN2* code majoritairement pour le transcrite  $\Delta 7$  (90 %), le transcrite pleine longueur représentant une part minoritaire de son expression (10 %).

Les gènes *SMN1* et *SMN2* ont une forte homologie, suggérant un événement de duplication – inversion relativement récent dans l'évolution, et ne diffèrent que par 5 nucléotides. Il existe en particulier une différence localisée dans l'exon 7 du gène *SMN2*, c.840C>T, qui est une variation synonyme (p.Phe280Phe) entraînant une modification de l'épissage : le gène *SMN1* produit un transcrite pleine longueur (transcrite *full length*) incluant l'exon 7, tandis que le gène *SMN2* est à l'origine essentiellement d'un transcrite dénué d'exon 7 (transcrite  $\Delta 7$ , 90 %) mais également d'une petite quantité (10 %) de transcrite pleine longueur (Cartegni et Krainer, 2002; Kashima et Manley, 2003; Lorson *et al.*, 1999; Monani *et al.*, 1999). Le transcrite  $\Delta 7$  génère une protéine instable car ne s'oligomérisant pas, contrairement au transcrite pleine longueur. La protéine SMN procède donc en grande partie de l'expression du gène *SMN1* mais aussi dans une moindre mesure du gène *SMN2*. Le nombre de copies du gène *SMN2* est variable d'un sujet à l'autre. L'absence totale de copies des gènes *SMN1* et *SMN2* (génotype 0 *SMN1*, 0 *SMN2*) n'est pas viable. Les patients présentent au moins 1 copie du gène *SMN2* générant une petite quantité de transcrite pleine longueur compatible avec la vie. Des travaux ont montré une corrélation entre le nombre de copies du gène *SMN2* et le type de SMA (Feldkötter *et al.*, 2002). Les quantités de transcrits et de protéines pleine longueur sont d'autant plus grandes qu'il y a de copies du gène *SMN2*, améliorant le pronostic de la SMA (Lefebvre

*et al.*, 1995; Lorson *et al.*, 1999; Monani *et al.*, 1999). Le gène *SMN2* est le principal facteur modificateur de la sévérité de la SMA.

La protéine SMN s'exprime de manière ubiquitaire et se localise aussi bien dans le cytoplasme que dans le noyau des cellules. Dans le noyau, la protéine SMN est localisée au sein de structures particulières appelées gems pour *gemini of coiled bodies*. Les gems semblent être étroitement associés au corps de Cajal. Ces derniers contiennent un niveau élevé de *small nucleolar ribonucleoproteins* (SnRNP) qui jouent un rôle clé dans le mécanisme de l'épissage (Liu and Dreyfuss, 1996).

La protéine SMN appartient à un complexe multi-protéique composé de SMN sous forme oligomérisée, de gemines, d'Unrip (*Unr-interacting protein*) et de protéines Sm (*RNA binding proteins*). Le complexe SMN intervient dans l'assemblage des snRNP du spliceosome dans le cytoplasme. Le complexe SMN permet la liaison de 7 protéines Sm à un snRNA. Les snRNP ainsi formés pourront subir une hyperméthylation de la coiffe et une maturation de l'extrémité 3' et être importées dans le noyau, où elles participent à l'épissage des ARN pré-messagers (Paushkin *et al.*, 2002) (Figure 2). Dans le noyau, les snRNP sont assemblés à partir des snRNA au sein des corps de Cajal en présence d'autres protéines dont la coïline (Singh *et al.*, 2017). Les gems correspondent à des corps de Cajal sans coïline. Les snRNP jouent un rôle clé dans le mécanisme de l'épissage (Liu and Dreyfuss, 1996). Les patients atteints de SMA ont en effet des anomalies de l'épissage de certains transcrits. Outre sa fonction dans la formation des snRNP, la protéine SMN est impliquée dans de nombreuses fonctions telles que la régulation de la biogenèse de petits RNP nucléaires ou nucléolaires, la régulation des particules de reconnaissance du signal et de la télomérase, le transport de certains ARNm, la réparation de l'ADN, la transcription, la traduction, la synthèse de sélénoprotéines, le trafic macromoléculaire, la formation de granules de stress, la signalisation cellulaire et la maintenance du cytosquelette (Paushkin *et al.*, 2002; Singh *et al.*, 2017). Les différents tissus ont des besoins spécifiques en protéine SMN, du fait de la variété et du nombre de partenaires avec lesquels SMN entre en interaction. Deux hypothèses ont été émises pour tenter d'expliquer

l'atteinte neuromusculaire prédominante de la SMA. Selon la première hypothèse, l'absence de la protéine SMN entrainerait une altération des complexes snRNP et par conséquent des anomalies d'épissage sur certains transcrits spécifiques ayant un rôle central dans le motoneurone, expliquant par là même l'atteinte préférentielle des motoneurons  $\alpha$ . Selon la deuxième hypothèse, ce serait d'une perturbation du transport de certains ARNm dans les neurones, que procéderait une réduction des taux de certains transcrits dans l'axone distal (Burghes and Beattie, 2009).

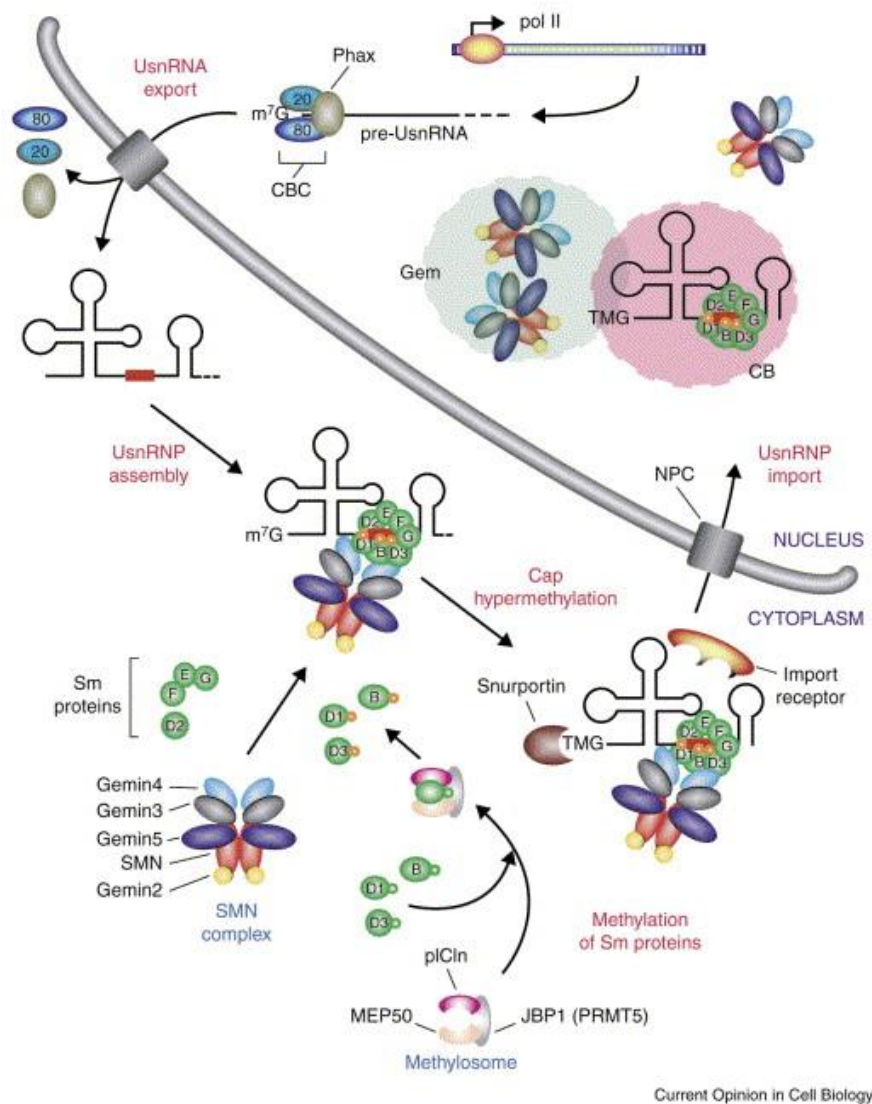


Figure 2 : Rôle de la protéine SMN dans la formation des complexes snRNP

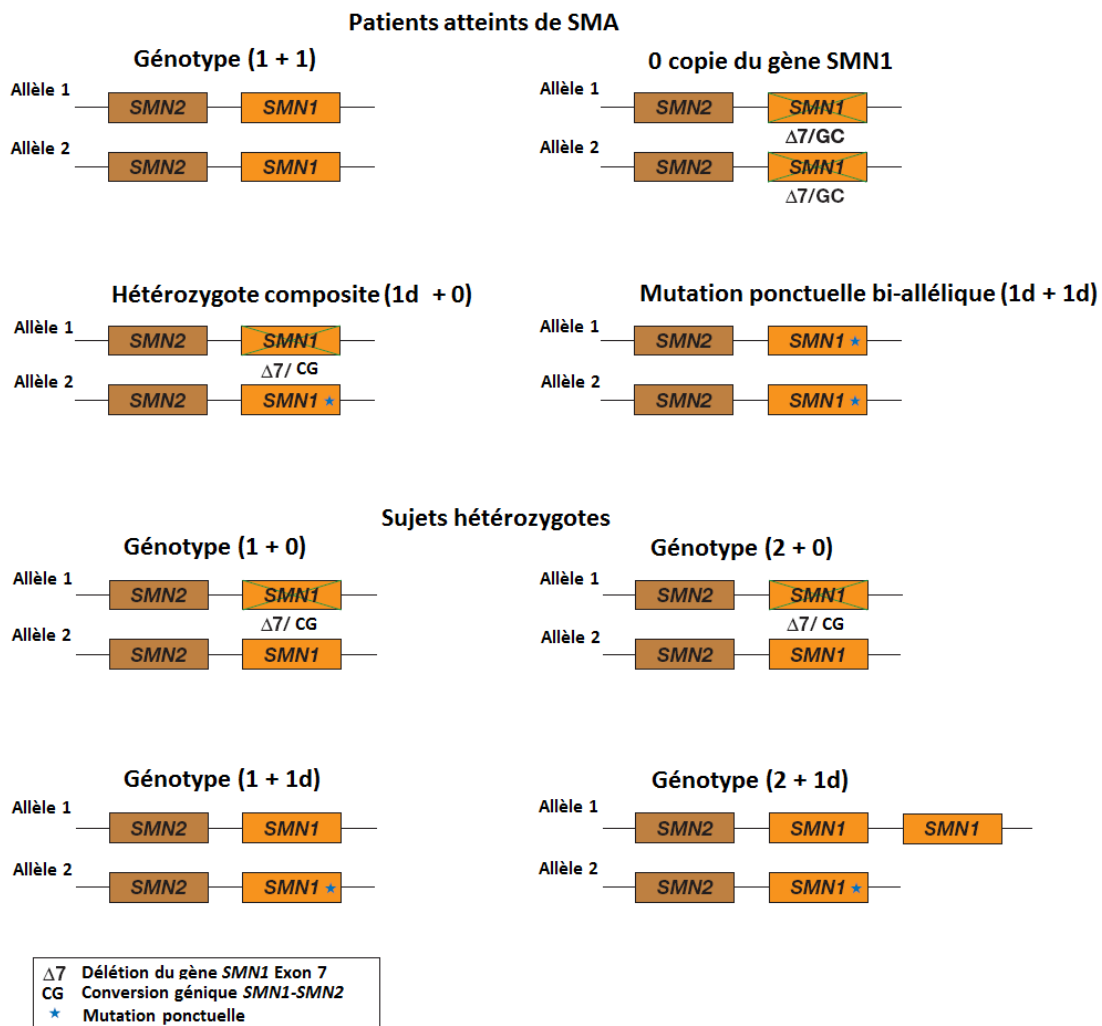
snRNA : *small nuclear RNA* ; CB : *Cajal body* ; NPC : *Nuclear Pore Complex* (Paushkin et al., 2002)

La confirmation du diagnostic de SMA repose sur l'identification de deux évènements mutationnels en *trans* chez le patient (Scheffer et al., 2001). Ainsi, 95 % des patients atteints de SMA présentent

une délétion à l'état homozygote du gène *SMN1* ou une conversion génique partielle de *SMN1* en *SMN2* qui correspond au remplacement d'une séquence du gène *SMN1* par celle du gène *SMN2*. Chez 5 % des patients, une mutation ponctuelle intragénique sur un allèle est associée à une délétion ou une conversion génique sur le second allèle en *trans* (Alías *et al.*, 2009; Lefebvre *et al.*, 1995; Wirth *et al.*, 1999). Enfin, d'exceptionnels cas de mutations ponctuelles à l'état homozygote ont été décrits dans des familles consanguines (Thauvin-Robinet *et al.*, 2012) (Tableau 2).

Génotype <i>SMN1</i> observé	Fréquence estimée
Délétion homozygote	95 %
Délétion hétérozygote + mutation ponctuelle	5 %
Mutation ponctuelle + mutation ponctuelle	<< 1 %

Tableau 2 : Génotypes au locus *SMN1* des patients atteints de SMA



**Figure 3 : Génotypes des patients atteints de SMA et des sujets hétérozygotes (Azad et al., 2019)**

En haut sont représentés les génotypes possibles pour des patients atteints de SMA : délétion homozygote (95 %), hétérozygotes composites (délétion du gène *SMN1* + mutation ponctuelle en *trans*, 5 %), mutation bi-allélique du gène *SMN1* (exceptionnel). En bas sont représentés plusieurs génotypes possibles pour des hétérozygotes : génotype (1+0), (2+0), (1+1d), (2+1d). L'allèle 1d correspond à l'allèle porteur d'une mutation ponctuelle.

La majorité des variations causales de la SMA étant des délétions du gène *SMN1*, le dépistage des porteurs hétérozygotes fait appel à des méthodes quantitatives d'analyse distinguant le gène *SMN1* du gène *SMN2*, telles que la QMPSF, la MLPA, la PCR digitale ou la technologie TaqMan quantitative. Il existe toutefois des allèles porteurs de 2 copies du gène *SMN1* en *cis*, soit sur le même chromosome 5, à une fréquence non négligeable dans la population générale (4 %) (Ogino et al., 2002). Or, les méthodes d'analyses quantitatives ne permettent pas de déterminer la phase des différentes copies du gène *SMN1*. De ce fait, les sujets porteurs de 2 copies en *cis* (génotype 2+0)

accompagnées d'une délétion de ce gène sur le second allèle en *trans* (génotype 2+0) et donc à risque de transmettre la maladie, ne sont pas distinguables des sujets porteurs de 2 copies en *trans* (génotype 1+1) qui sont, quant à eux, beaucoup moins à risque de transmettre la maladie (au risque près de survenue d'une délétion *de novo*, d'une conversion génique ou d'une mutation ponctuelle) (Mailman *et al.*, 2002).

### **Le conseil génétique de la SMA**

Le conseil génétique dans la SMA repose sur la mise en évidence, chez des apparentés de patients atteints de la maladie, de variations délétères présentes à l'état hétérozygote, comme une délétion du gène *SMN1* ou une mutation ponctuelle. Compte tenu de la fréquence de la délétion à l'état hétérozygote dans la population générale, un conseil génétique est indiqué chez les apparentés d'un cas index atteint de SMA. Au cours de la consultation, le généticien évalue le risque *a priori* de cet apparenté d'avoir un enfant atteint de SMA sur la base de l'arbre généalogique. Il peut prescrire une analyse pour rechercher la délétion à l'état hétérozygote et calculera un risque *a posteriori* qui prendra en compte le résultat de cette analyse. La recherche de ce statut d'hétérozygote n'est réalisée que chez les personnes majeures. Même si l'apparenté est porteur de 2 copies du gène *SMN1*, il persiste un risque qu'il soit hétérozygote de l'ordre de 10 %. Ce risque résiduel résulte du risque de duplication en *cis* du gène *SMN1* masquant une délétion sur le second allèle, de mutation ponctuelle et de délétion *de novo*.

### **Notion de risque résiduel et existence de duplications en cis du gène SMN1**

Les analyses quantitatives mettant en évidence le nombre de copies du gène *SMN1* ne rendent pas compte des mutations ponctuelles (0,02 % chez les porteurs de 2 copies de *SMN1* ; 1,3% des patients atteints de SMA), du risque de délétion *de novo* de *SMN1* (0,02 %) et des génotypes 2+0 (8 %) (Ogino *et al.*, 2002). La probabilité conditionnelle chez un apparenté porteur de 2 copies du gène *SMN1*

d'être tout de même « hétérozygote » ou à risque de transmettre une amyotrophie spinale infantile à sa descendance est dénommée « risque résiduel ». Celui-ci est au total estimé à environ 10 % dans la population française, mais est variable selon les populations. La variabilité du risque résiduel résulte de la variabilité de fréquence des duplications en *cis* du gène *SMN1* dans les différentes sous-populations. C'est dans la population africaine que la prévalence de sujets porteurs d'une duplication en *cis* du gène *SMN1* est la plus élevée, avec une fréquence allélique de l'allèle "2 copies du gène *SMN1*" de 26,91 %, et de ce fait une fréquence également élevée de génotypes 2+0, de sorte que la sensibilité de détection des hétérozygotes dans cette population n'est que de 71 % et le risque résiduel pour un individu de la population africaine après la mise en évidence de 2 copies de 1/121 au lieu de 1/391 dans la population française (Hendrickson *et al.*, 2009; Luo *et al.*, 2014) (Tableaux 3 et 4).

Ethnie	Fréquence allélique (%)			
	0 copie	1 copie	2 copies	1 copie*
<b>Caucasiens</b>	1,42	95,32	<b>3,18</b>	0,03
<b>Juifs ashkénazes</b>	1,21	90,72	<b>8,25</b>	0,02
<b>Asiatiques</b>	0,96	93,38	<b>5,71</b>	0,02
<b>Afro-américains</b>	0,77	71,88	<b>26,91</b>	0,01
<b>Hispaniques</b>	0,44	91,88	<b>8,04</b>	0,01

Tableau 3 : Fréquences alléliques du nombre de copies du gène *SMN1* par allèle dans différentes sous-populations (Hendrickson *et al.*, 2009)

Fréquence des allèles "0 copie", "1 copie", "2 copies" du gène *SMN1* et "1 copie avec mutation ponctuelle" (1 copie\*). L'allèle "2 copies" est plus fréquent dans la population africaine (> 26 %) comparativement aux autres populations (entre 3 et 9 %).

	Génotype	Fréquence génotypique (%)				
		Caucasiens	Juifs ashkénazes	Asiatiques	Afro-américains	Hispaniques
Non porteurs	2 + 2	0,11	0,65	0,32	7,48	0,59
	2 + 1	6,21	14,62	10,58	39,31	14,16
	1 + 1	90,81	82,30	87,20	51,69	84,39
Hétérozygotes	2 + 1*	0,0017	0,0036	0,002	0,0077	0,0012
	<b>2 + 0</b>	<b>0,09</b>	<b>0,19</b>	<b>0,11</b>	<b>0,41</b>	<b>0,07</b>
	1 + 1*	0,05	0,04	0,03	0,02	0,01
	1 + 0	2,72	2,20	1,75	1,08	0,86
Patients atteints de SMA	1* + 1*	$7,1 \cdot 10^{-6}$	$5,1 \cdot 10^{-6}$	$2,0 \cdot 10^{-6}$	$3,1 \cdot 10^{-6}$	$6,2 \cdot 10^{-7}$
	1* + 0	0,00076	0,00055	0,00021	0,00033	0,000067
	0 + 0	0,021	0,015	0,0059	0,0091	0,0019

Tableau 4 : Fréquence des génotypes dans différentes sous-populations (Hendrickson *et al.*, 2009)

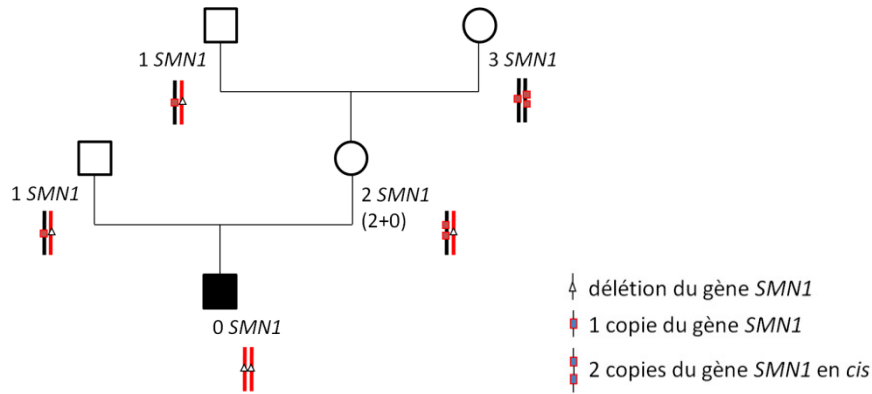
La proportion de génotypes 2+0 en population générale, varie de 0,09 % dans la population caucasienne nord-américaine, à 0,41 % dans la population afro-américaine aux Etats-Unis (Hendrickson *et al.* 2009; Ogino *et al.* 2002). Dans le cas d'un apparenté ayant un risque à priori de  $\frac{1}{2}$  d'être hétérozygote pour la délétion du gène *SMN1* (par exemple, l'oncle d'un patient atteint de SMA), son risque à posteriori d'être hétérozygote sachant qu'il a 2 copies du gène *SMN1* est de  $\frac{1}{11}$  s'il est issu de la population caucasienne et de  $\frac{3}{13}$  s'il est issu de la population africaine (d'après Luo *et al.*, 2014). Cela explique le choix d'analyser simultanément les deux conjoints dans le cadre d'un conseil génétique de couple.

### **Modalités d'identification d'un génotype 2+0**

Les duplications en *cis* du gène *SMN1* masquant une délétion sur le second allèle sont décelées par la mise en évidence d'une ségrégation familiale particulière des allèles de ce gène, impliquant la présence d'un génotype 2+0 chez un apparenté à un cas-index atteint de SMA. Pour mettre en évidence une ségrégation révélatrice d'une duplication en *cis* du gène *SMN1* chez l'un des apparentés, il est nécessaire de prendre en considération le nombre de copies de *SMN1* sur au moins 3 générations consécutives.

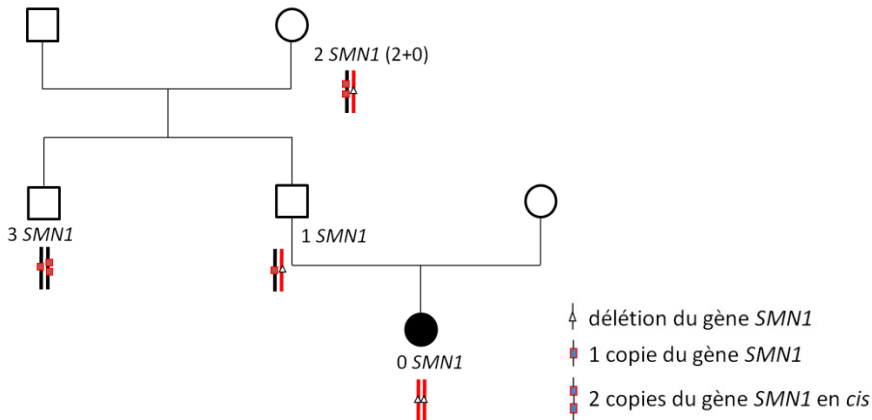


La reconstitution des haplotypes, issus de l'étude des microsatellites au locus *SMN1* (C212, C272, I105 et VS19), peut s'avérer utile dans la mesure où un apparenté porteur de 2 copies du gène *SMN1* et porteur par ailleurs de l'haplotype correspondant à la délétion de *SMN1*, tel que chez le cas-index, a nécessairement un génotype 2+0 (figures 4 et 5).



**Figure 4 : Exemple d'un parent de cas-index 2T avec grands-parents 1T et 3T**

Dans cet exemple, la mère du cas-index possède 2 copies du gène *SMN1* et ses propres parents ont 1 copie et 3 copies de ce gène, impliquant la présence des 2 copies en *cis* chez la mère du cas-index.



**Figure 5 : Exemple d'un sujet 2+0 ayant des enfants 3T et 1T**

Dans cet autre exemple, la grand-mère paternelle du cas-index possède 2 copies du gène *SMN1* et a deux enfants ayant 3 copies et 1 copie de ce gène. Par ailleurs, une analyse des microsatellites dans cette famille a montré que la grand-mère du cas-index était porteuse de l'allèle porteur de la délétion du gène *SMN1*. Elle a donc un génotype 2+0.

### **Identification d'un haplotype associé à la duplication en *cis* du gène *SMN1***

Dans l'optique d'améliorer la détection des génotypes 2+0, l'équipe américaine de Luo *et al.* a cherché en 2014 à mettre en évidence des allèles fondateurs porteurs de 0 copie du gène *SMN1* d'une part et de deux copies du gène *SMN1* d'autre part, dans un échantillon de population juive ashkénaze. Les profils des microsatellites au locus *SMN1* (D5S681, D5S435, MS1 et D5S610) ont été analysés chez 23 sujets de génotype 1+0 et 72 sujets de génotype 2+1, soit respectivement porteurs des allèles "0 copie" et "2 copies" (Luo *et al.*, 2014). La fréquence des haplotypes retrouvés dans la population étudiée a été comparée à celle d'un groupe contrôle constitué de 78 sujets de génotype 1+1. Ainsi, plusieurs haplotypes de microsatellites sont significativement enrichis dans les groupes 1+0 et 2+1 par rapport au groupe contrôle ( $p < 0.05$ ) (Tableau 5).

Génotype	D5S681	D5S435	MS1	D5S610	Fréq. C., % (n = 78)	Fréq. Gr., %	P valeur
<b>1+0</b> (n = 20)	2	1	4	4	0	7,5	1
	2	1	4	5	0	10	0,001
	2	5	4	5	0,6	22,5	<0,0001
	2	5	7	5	1,4	5	0,19
<b>Dup en cis</b> <b>3 SMN1/0</b> <b>SMN2 et</b> <b>3 SMN1/1</b> <b>SMN2</b> (n = 42)	2	3	7	5	0	4,7	0,01
	2	5	3	4	2,6	3,6	0,12
	<b>2</b>	<b>5</b>	<b>6</b>	<b>10</b>	<b>0</b>	<b>19</b>	<b>&lt;0,0001</b>
	2	5	9	2	1,8	4,8	0,19
	4	3	4	2	0	3,4	0,03
	4	5	4	2	0,8	3,6	0,003
<b>Dup en cis</b> <b>3 SMN1/2</b> <b>SMN2</b> (n = 9)	2	4	9	5	0	5,6	0,035
	2	5	4	5	0,6	11,1	<0,0001
	2	5	7	2	0	5,6	1
	2	5	9	4	0	16,7	<0,0001
	4	1	4	10	0	5,6	1
	4	5	4	9	1,8	11,1	0,02
	6	4	4	5	0	5,6	0,025
	7	3	2	5	0	5,6	0,003

Tableau 5 : Haplotypes de microsatellites par génotype au locus *SMN1* (Luo *et al.*, 2014)

Fréq. C. : Fréquence dans le groupe contrôle (1+1); Fréq. Gr. : Fréquence dans le groupe considéré. L'étude des microsatellites au locus *SMN1* chez des sujets (1+0) d'une part, et (2+1) d'autre part, a permis aux auteurs de mettre en évidence un enrichissement significatif en certains haplotypes par rapport à un groupe contrôle constitué de sujets 1+1. L'association la plus statistiquement significative ( $P$  valeur < 0.0001) et absente chez les sujets contrôle est en gras.

L'haplotype 2, 5, 6, 10 (pour les répétitions des microsatellites D5S681, D5S435, MS1 et D5S610 respectivement) s'est révélé être statistiquement enrichi dans le groupe "duplications en *cis*", avec 19/42 sujets porteurs dans ce groupe contre 0/78 dans le groupe contrôle ( $P$  < 0,0001). Ces premiers résultats étaient donc compatibles avec l'hypothèse initiale d'allèles fondateurs de duplications en *cis* du gène *SMN1*. Les auteurs ont ensuite amplifié par PCR le gène *SMN1* chez les sujets porteurs de cet haplotype en utilisant un couple d'amorce hybridant de manière préférentielle sur ce gène, *SMN1*-E7F et *SMN1*-E8R. Un séquençage Sanger des exons 7 à 8 de ces allèles a montré que tous les sujets porteurs de l'haplotype 2, 5, 6, 10 étaient également porteurs de 2 variants : NM\_000344.3(*SMN1*):c.\*3+80T>G et NM\_000344.3(*SMN1*):c.\*211\_\*212del.

Une combinaison d'analyse par PCR long range et par TA Cloning a permis de phaser ces 2 polymorphismes et de confirmer leur localisation en *cis*, sur le même gène *SMN1*. Pour déterminer si ces deux variants se retrouvaient dans d'autres populations, les auteurs ont déterminé le nombre de copies du gène *SMN1* (par MLPA) et la présence du variant c.\*3+80T>G chez 276 sujets afro-américains, 250 sujets asiatiques, 262 sujets hispaniques et 458 sujets caucasiens (Tableau 6).

Ethnie	Copies du gène <i>SMN1</i>	c.*3+80T>G absent	c.*3+80T>G hétérozygote	c.*3+80T>G homozygote	Total
<b>Juifs ashkénazes</b>	2	315	0	0	315
<b>Afro-américains</b>	1	4	0	1	5
	2	108	26	2	136
	3	21	86	4	111
	4	1	23	0	24
<b>Asiatiques</b>	1	2	0	0	2
	2	222	0	0	222
	3	20	2	0	22
	4	1	1	0	2
<b>Hispaniques</b>	1	1	0	0	1
	2	203	12	1	219
	3	20	20	0	40
	4	0	2	0	2
<b>Caucasiens</b>	1	12	0	0	12
	2	413	2	0	415
	3	23	4	0	27
	4	2	2	0	4

Tableau 6 : du variant c.\*3+80T>G dans d'autres populations (Luo *et al.*, 2014)

Dans la population afro-américaine, 86 des 111 sujets ayant 3 copies du gène *SMN1* (probables génotypes 2+1, mais génotype 3+0 non exclu) présentent le variant, mais ce variant est également présent chez des sujets ne présentant pas de duplication en *cis* du gène *SMN1* : 1 sujet du groupe 1 *SMN1* (n = 5), 28 sujets du groupe 2 *SMN1* (génotypes 1+1 et 2+0, (n = 136)). A l'inverse, dans la population asiatique, aucun sujet sans duplication du gène *SMN1* n'est porteur du variant. Dans la population hispanique, 1 variant au moins se trouve sur un allèle sans duplication en *cis* du gène *SMN1*. Dans la population caucasienne, 4 sujets du groupe 3 *SMN1* (n = 27) et 2 sujets du groupe 4 *SMN1* (n = 4) présentent le variant.

Ces résultats révèlent une association moins spécifique du variant c.\*3+80T>G, dans la population africaine que dans les autres populations dans lesquelles les allèles non porteurs d'une duplication en *cis* du gène *SMN1* sont quasi systématiquement également non porteurs du variant recherché (Luo *et al.*, 2014).

Cette étude a par la suite été répliquée en 2018 par Alías *et al.*, chez 270 sujets non apparentés issus de la population espagnole (Alías *et al.*, 2018). Ces sujets ont été répartis en différents sous-groupes selon leur génotype, établi par l'étude de la ségrégation familiale des gènes *SMN1* et *SMN2*, éventuellement couplée à une analyse de microsatellites. Les 270 sujets de l'étude comprenaient à la fois des patients atteints de SMA (génotype 0 *SMN1*, (n = 16) et conversions géniques *SMN1-SMN2* (n = 20), des sujets hétérozygotes (génotypes 1+0 (n = 41), 2+0 (n = 32) et 3+0 (n = 1)) et de sujets non porteurs de la délétion (génotypes 1+1 (n = 99), 2+1 (n = 58), 2+2 (n = 3)). Des amorces R111 (intron 6) et 541C1120 (exon 8) ont été utilisées pour amplifier par PCR les gènes *SMN1* et *SMN2* de manière non spécifique. Les deux variants précédemment mis en évidence, ont été recherchés par séquençage Sanger (Tableau 7).

	Conversions							
	0 <i>SMN1</i>	<i>SMN1-SMN2</i>	1+0	2+0	3+0	1+1	2+1	2+2
<b>c.*3+80T&gt;G</b>	0	0	1	7	0	0	11	0
<b>c.*211_*212del</b>	0	4	1	6	0	0	11	0
<b>Nombre de sujets</b>	16	20	41	32	1	99	58	3

Tableau 7 : Présence des variants d'intérêt selon le génotype (Alías *et al.*, 2018)

Les deux variants d'intérêt ont été recherché par séquençage Sanger chez les 270 sujets de l'étude, répartis en sous-groupes selon leur génotype. Les sujets porteurs de conversions géniques *SMN1-SMN2* ont un gène présentant l'intron 7 du gène *SMN2* et l'exon 8 du gène *SMN1*.

Parmi les patients atteints de SMA avec un génotype 0 *SMN1*, aucun ne présentait les variants recherchés (0/16). En revanche parmi les patients présentant 0 copie du gène *SMN1*, avec une conversion génique *SMN1-SMN2*, certains présentaient, l'un des deux variants (4/20, 20 %). Un seul sujet du groupe 1+0 présentait les variants, alors que 7 des 32 sujets 2+0 présentaient ces variants (21,8 %). Dans ce groupe un sujet n'était porteur que du variant c.\*3+80T>G. Le seul sujet 3+0 ne présentait pas les variants. Concernant les sujets non hétérozygotes, aucun parmi les groupes 1+1 et 2+2 ne présentait les variants, alors que 11 sujets en étaient porteurs dans le groupe 2+0 (11/58, 19,3 %,  $P < 0.001$ ). Afin de démontrer l'association des variants au gène *SMN1* et non au gène *SMN2*,

les auteurs ont montré que ces variants n'étaient pas présents chez les sujets porteurs de 0 copie du gène *SMN1* et que chez les sujets présentant des conversions géniques (hybrides *SMN1-SMN2*), certains ne présentaient que le variant c.\*211\_\*212del, qui se trouve dans l'exon 8 du gène, c'est à dire la "partie *SMN1*" de ce gène hybride.

Les deux variants d'intérêt se sont révélés être associés aux duplications en *cis* du gène *SMN1* avec une sensibilité de l'association de 20 %, une spécificité de 99 %, une VPP de 61,9 % et une VPN de 93,8 % (Alías *et al.*, 2018).

La forte spécificité de cette association a conduit, en 2019, à une troisième étude sur 198 sujets 2+0 dans laquelle les auteurs ont décrit une méthode de détection du variant NM\_000344.3(*SMN1*):c.\*3+80T>G par sonde Taqman, à la fois robuste et répétable (Azad *et al.*, 2019), dans l'optique d'une intégration au conseil génétique.

### **Objectif de l'étude**

L'objectif principal de cette étude est de déterminer si l'association des deux variants précédemment mis en évidence dans les populations juive ashkénaze (Luo *et al.*, 2014) et espagnole (Alías *et al.*, 2018), NM\_000344.3(*SMN1*):c.\*3+80T>G (variant 1) et NM\_000344.3(*SMN1*):c.\*211\_\*212del (variant 2), avec des duplications en *cis* du gène *SMN1* (génotypes 2+0, 3 copies, 4 copies, 3+0), est extrapolable à la population française analysée au CHU de Rouen, en faisant appel à des techniques de SNaPshot et le kit AmpliDeX PCR-CE *SMN1/2* (Asuragen, Austin, Etats-Unis), à la fois rapides et ciblées.

Un objectif secondaire est de confirmer la spécificité de l'association de ces variants aux duplications en *cis* de *SMN1* et non de *SMN2*.

## Matériel et méthodes

### Constitution des sous-groupes analysés

199 sujets de la population française ont été analysés pour déterminer la présence ou l'absence des variants 1 et 2, nommés plus haut, dans le gène *SMN1*. Ces sujets étaient répartis en plusieurs sous-groupes selon leur génotype :

- 46 sujets étaient porteurs de 2 copies en *trans* du gène. Ils n'étaient pas apparentés à un cas-index. En effet, les apparentés à un cas-index atteint de SMA qui auraient 2 copies du gène *SMN1* auraient un risque plus élevé d'avoir deux copies en *cis* que des sujets issus de la population générale. Nous avons donc minimisé ce risque en ne sélectionnant que des conjoints d'apparentés, analysés dans le cadre du dépistage des hétérozygotes en vue d'estimer le risque pour ce couple d'avoir un enfant atteint de SMA.
- 59 sujets étaient porteurs d'une seule copie du gène,
- 56 sujets étaient porteurs de 3 copies du gène avec une duplication en *cis* du gène et un allèle porteur d'une copie en *trans*.
- 1 sujet était porteur d'une triplication en *cis* du gène (génotype 3+0). Il s'agissait d'un parent de cas-index.
- 12 sujets étaient porteurs de 4 copies du gène,
- 25 sujets étaient porteurs de 2 copies en *cis* du gène *SMN1* masquant une délétion sur le second allèle (génotype 2+0). Il s'agissait de mères, de pères ou d'autres apparentés à un cas-index, porteurs de 2 copies du gène *SMN1* et pour lesquels l'étude de ségrégation familiale des gènes *SMN1* et *SMN2* affirmait ce génotype. Ils étaient issus d'une collection de 18 ans de diagnostic moléculaire de la SMA au CHU de Rouen.

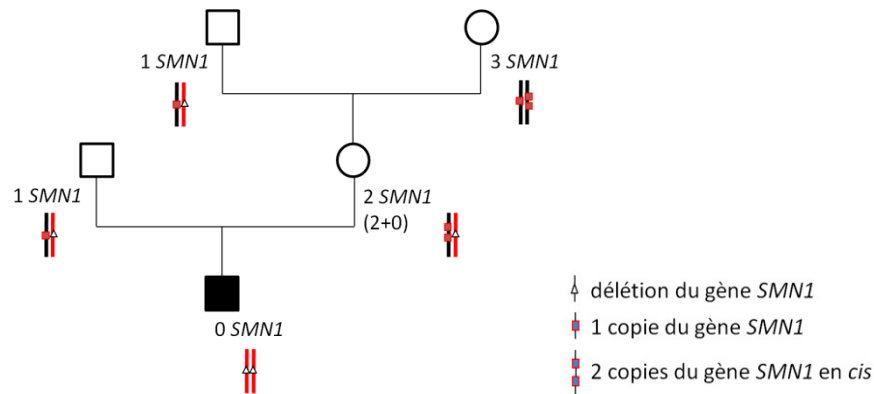


Figure 6 : Exemple d'un parent de cas-index 2T avec grands-parents 1T et 3T

Dans cet exemple, la mère du cas-index possède 2 copies du gène *SMN1* et ses propres parents ont 1 copie et 3 copies de ce gène, impliquant la présence des 2 copies en *cis* chez la mère du cas-index.

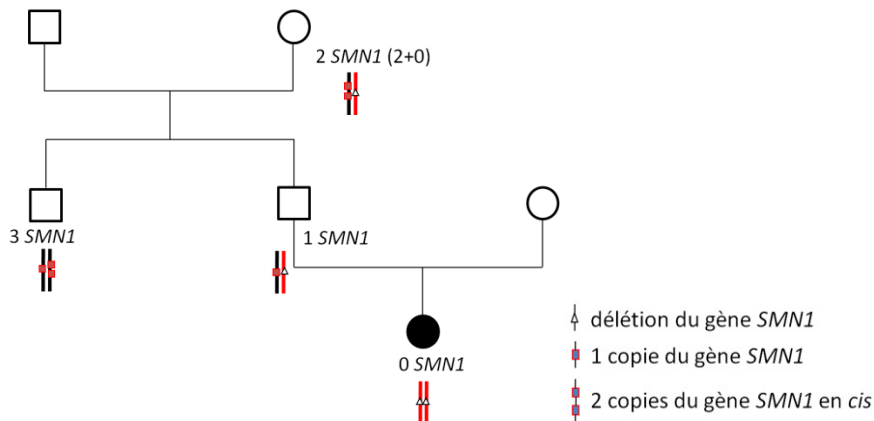


Figure 7 : Exemple d'un sujet 2+0 ayant des enfants 3T et 1T

Dans cet autre exemple, la grand-mère paternelle du cas-index possède 2 copies du gène *SMN1* et a deux enfants ayant 3 copies et 1 copie de ce gène. Par ailleurs, une analyse des microsatellites dans cette famille a montré que la grand-mère du cas-index était porteuse de l'allèle porteur de la délétion du gène *SMN1*. Elle possède donc 2 copies en *cis* du gène *SMN1*.

Dans chaque sous-groupe, si plusieurs sujets étaient apparentés, un seul sujet était inclus par tirage au sort.

Nous avons distingué deux groupes de sujets :

- d'une part les sujets présentant une duplication ou une triplication en *cis* du gène *SMN1* (groupe "dup en *cis*"), à savoir ceux ayant un génotype 2+0, 3 copies, 4 copies ou 3+0



- d'autre part les sujets ne présentant pas de duplication en *cis* du gène (groupe "absence de dup en *cis*"), à savoir ceux ayant un génotype 1 copie ou 2 copies en *trans*.

Tous ces sujets avaient déjà fait l'objet d'une analyse du nombre de copies du gène *SMN1* par une technique de QMPSF au CHU de Rouen.

## Méthodes

### QMPSF

La QMPSF (*Quantitative Multiplex PCR of Short Fluorescent fragments*) est la technique que nous avons utilisé pour déterminer le nombre de copies des gènes *SMN1* et *SMN2* (Charbonnier *et al.*, 2000).

### Principe

Il s'agit d'une méthode qualitative basée sur des données quantitatives. Elle permet une détection simple et rapide de remaniements génomiques (délétions ou duplications). Le principe de cette méthode semi-quantitative repose sur 4 points :

- Des amplicons de petite taille (<500 pb) sont générés pour chaque région d'intérêt grâce à 2 amorces dont l'une est marquée par un fluorochrome;
- la PCR est réalisée avec un nombre limité de cycles pour se situer dans la phase exponentielle de la PCR;
- Plusieurs amplicons sont simultanément amplifiés dans le même tube en présence d'un amplicon contrôle C localisé dans une région indépendante;
- Après migration des produits de PCR sur séquenceur automatique, l'analyse des électrophérogrammes est basée sur la comparaison de l'aire sous le même pic entre 2 échantillons distincts après superposition informatique des chromatogrammes. Ceci permet de s'affranchir des différences d'efficacité de chaque amplification, inéluctables malgré le choix d'amplicons courts et d'amorces de Tm voisin. L'analyse des électrophérogrammes repose sur la comparaison de l'aire des pics après normalisation selon la formule suivante (*Dosage Quotient*; DQ) :

$$DQ = \frac{(A_{\text{patient}}/A_{\text{témoin normal}})}{(A_{\text{patient}}/A_{\text{témoin normal}})}$$

La valeur du coefficient de dosage (DQ) d'échantillons de référence ayant 2 copies du gène *SMN1* et 2 copies du gène *SMN2* doit être comprise entre 0,8 et 1,2. On appelle facteur de correction la valeur du rapport entre l'aire sous la courbe de l'amplicon contrôle C de patient ( $A_{C \text{ patient}}$ ) par l'aire sous la courbe de l'amplicon contrôle C du témoin normal ( $A_{C \text{ témoin normal}}$ ). Cette valeur correspond au dénominateur de la formule citée ci-dessus.

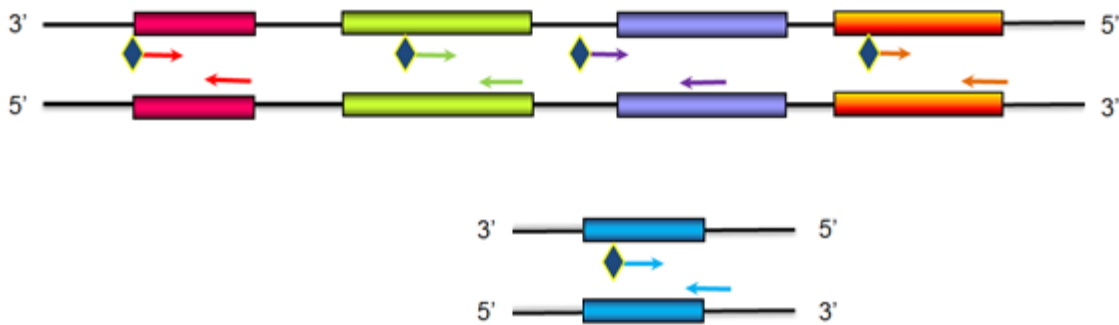


Figure 8 : Principe de la QMPSF

Cette technique débute par l'amplification en multiplex de plusieurs (15 au maximum) petits fragments (<500 pb). L'une des deux amorces de chaque paire est marquée par un fluorochrome 6-fam, représenté par un losange. L'amplicon contrôle est représenté en bleu.

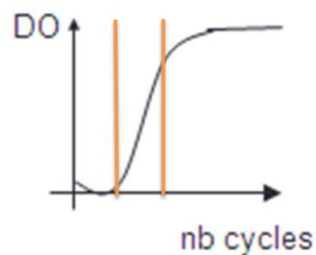


Figure 9 : Limitation du nombre de cycles de PCR

Afin de permettre une quantification initiale des fragments amplifiés, il est impératif de rester dans la phase exponentielle de la réaction de PCR. Le nombre de cycles doit donc être limité à 23 à 25 cycles au maximum. DO : Densité optique (reflète la quantité d'amplicons générés).

## Etape d'amplification par PCR

Pour chaque échantillon, deux PCR en multiplex ont été réalisées, l'une en vue d'une digestion enzymatique par *DraI* (New England Biolabs, Ipswich, États-Unis) l'autre en vue d'une digestion enzymatique par *HinfI* (New England Biolabs, Ipswich, États-Unis), les gènes *SMN2* et *SMN1* comprenant des séquences de restriction pour *DraI* et *HinfI*, respectivement. Cette digestion a pour but de pouvoir distinguer les amplicons des gènes copies *SMN1* et *SMN2* par une troncature soit de l'amplicon de *SMN1* uniquement (*HinfI*), soit de l'amplicon de *SMN2* uniquement (*DraI*).

Les PCR *DraI* et *HinfI* amplifient une séquence commune aux gènes *SMN* (ne différant que par la présence du site de restriction) ainsi qu'un témoin de PCR et un témoin de digestion. Pour la PCR *DraI*, les témoins de digestion et de PCR sont respectivement l'exon 18 et l'exon 11 du gène *MLH1*. Pour la PCR *HinfI*, les témoins de digestion et de PCR sont respectivement l'exon 3 et l'exon 4 du gène *MLH1*. L'une des deux amorces de chaque paire est couplée à un fluorochrome 6-fam (6-Carboxyfluorescéine) qui émet un pic de fluorescence à 488 nm après excitation par un laser argon-ion, permettant via une détection et une quantification de la fluorescence émise à un temps de migration donné, de quantifier chaque amplicon obtenu. Ces amorces sont commandées via l'interface en ligne *Integrated DNA Technologies* (Coralville, États-Unis) (<https://www.idtdna.com/pages>).

Amplicon	Nom de l'amorce	Séquence de l'amorce
<b>T SMA <i>DraI</i></b>	X7-Dra	5'-CCTTCCTCTTTTTGATTTTGTTC-3'
	R111 fam	5'-AGACTATCAACTTAATTTCTGATCA-3'
<b>T dig <i>DraI</i></b>	<i>MLH1</i> ex18F	5'-GTAGTCTGTGATCTCCGTTT-3'
	<i>MLH1</i> ex18Rfam-V2	5'-TGTATGTAGGTCCTGTCTCA-3'
<b>T PCR <i>DraI</i></b>	<i>MLH1</i> ex11Ffam	5'-TTTGACCACTGTGTCATCTG-3'
	<i>MLH1</i> ex11R	5'-ACCTGGGTGAAGTACATCCT-3'

Tableau 8 : QMPSF *DraI*

La QMPSF suivie d'une digestion par l'enzyme de restriction *DraI* génère 3 amplicons dont un correspondant aux gènes *SMN1* et *SMN2* de manière non spécifique, un correspondant à l'exon 18 du gène *MLH1* et un correspondant à l'exon 11 de ce même gène. Pour chaque paire, l'amorce *reverse* (R) est couplée à un fluorochrome 6-fam (6-Carboxyfluorescéine) émettant un pic de fluorescence à 488 nm (Integrated DNA Technologies, Coralville, États-Unis), permettant la détection et la quantification ultérieure de ces amplicons sur un séquenceur capillaire.

Amplicon	Nom de l'amorce	Séquence de l'amorce
<b>T SMA <i>HinfI</i></b>	<i>SMN HinfI</i> F	5'-CGTTAGATAGCTTCCTTTATTTTCCTTACAGGGATT-3'
	<i>SMN HinfI</i> R fam	5'-GATAGGGTTACTTTCATAATGCTGGCAGACT-3'
<b>T dig <i>HinfI</i></b>	<i>MLH1</i> ex3 F fam	5'-CGTTAGATAGATCTGGATATTGTATGTGAAAGGTT-3'
	<i>MLH1</i> ex3 R	5'-GATAGGGTTAACAGACAATGTCATCACAGGAG-3'
<b>T PCR <i>HinfI</i></b>	<i>MLH1</i> ex4 F fam	5'-CGTTAGATAGCAGTTCAGATAACCTTCCCTTTG-3'
	<i>MLH1</i> ex4 R	5'-GATAGGGTTAAGCTGTTTCGTTGTAATAGTAACATG-3'

Tableau 9 : QMPSF *HinfI*

La QMPSF suivie d'une digestion par l'enzyme de restriction *HinfI* génère également 3 amplicons dont un correspondant aux gènes *SMN1* et *SMN2*, un correspondant à l'exon 3 du gène *MLH1* et un correspondant à l'exon 4 de ce même gène. Là encore, pour chaque paire, l'une des deux amorces est couplée à un fluorochrome 6-fam permettant sa détection sur un séquenceur capillaire.

Les QMPSF *DraI* et *HinfI* sont réalisées simultanément dans les conditions décrites dans la littérature (Saugier-veber *et al.*, 2001; Vezain *et al.*, 2010).

Les mix sont répartis dans des microtubes et 100 ng d'ADN sont ajoutés dans chaque microtube. L'amplification par PCR est réalisée au moyen de la Taq *diamond Taq polymerase* (Eurogentec, Liège, Belgique). La PCR *DraI* comporte 24 cycles et la PCR *HinfI* 23 cycles. Ces PCR génèrent des amplicons

respectivement de 188 pb et 135 pb, à partir du gène *SMN1* et du gène *SMN2* de façon non spécifique.

### Etape de digestion

La totalité des produits de PCR ainsi obtenus subissent une digestion enzymatique par les enzymes *DraI* ou *HinfI*, d'au moins 4h30 dans un bain-marie à 37°C. La différence de taille des fragments d'amplicons des gènes *SMN1* et *SMN2* obtenus après digestion respective par les enzymes *DraI* et *HinfI* permettra de les départager lors de la migration (Figures 10 et 11).

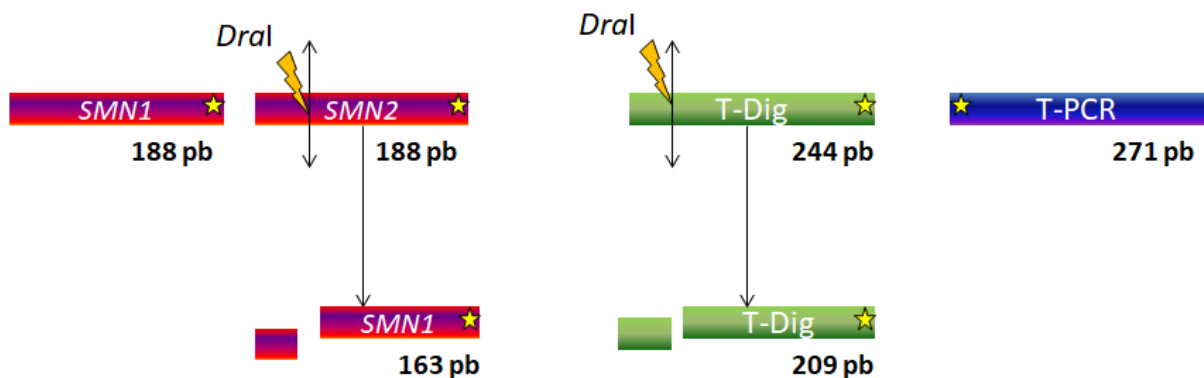


Figure 10 : Etape de digestion par l'enzyme *DraI*

L'étoile représente la position du fluorochrome 6-fam. L'amplicon commun *SMN* (188 pb) est scindé, après digestion par l'enzyme *DraI*, en un fragment *SMN2* fluorescent (163 pb) et un petit fragment non fluorescent (25 pb). L'amplicon *SMN1* n'est pas digéré et conserve sa taille initiale (188 pb), de même que le témoin de PCR (271 pb). Le témoin de digestion (244 pb) est scindé en un fragment fluorescent (209 pb) et un petit fragment non fluorescent (35 pb).

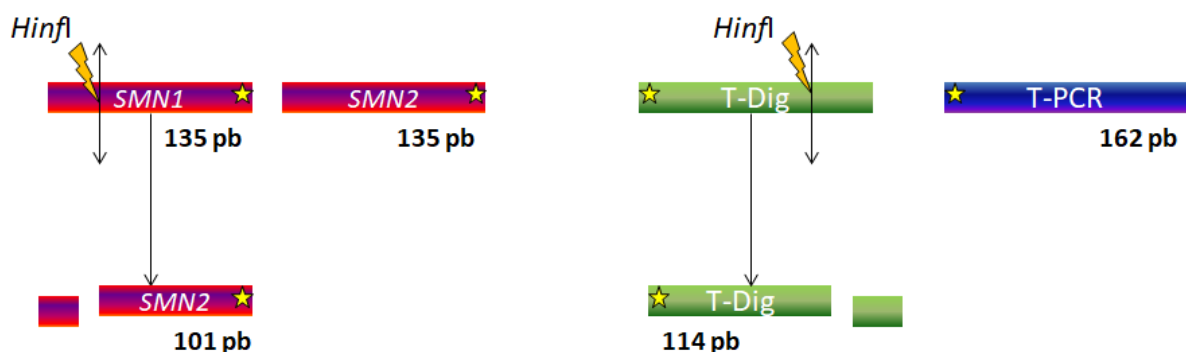


Figure 11 : Etape de digestion par l'enzyme *HinfI*

L'étoile représente la position du fluorochrome 6-fam. L'amplicon commun *SMN* (135 pb) est scindé, après digestion par l'enzyme *HinfI*, en un fragment *SMN2* fluorescent (101 pb) et un petit fragment non fluorescent (34 pb). L'amplicon *SMN2* n'est pas digéré et conserve sa taille initiale (135 pb), de même que le témoin de PCR (162 pb). Le témoin de digestion est scindé en un fragment fluorescent (114 pb) et un petit fragment non fluorescent.

### Etape de migration et analyse des données

Après purification par Sephadex G-100 (Sigma-Aldrich, Saint-Louis, États-Unis), les fragments fluorescents sont séparés par électrophorèse capillaire sur le séquenceur Genetic Analyzers 3130xl (Thermo Fisher Scientific, Waltham, États-Unis) en présence du marqueur de taille GS120 liz de 120 pb (Applied Biosystem, Foster City, États-Unis), permettant la séparation et l'analyse des quatre pics d'amplicons obtenus (*SMN1*, *SMN2*, témoin de digestion et témoin de PCR) sur le logiciel Gene Mapper 4.1 (Thermo Fisher Scientific, Waltham, États-Unis) (Figure 12).

## Exemples d'électrophérogrammes

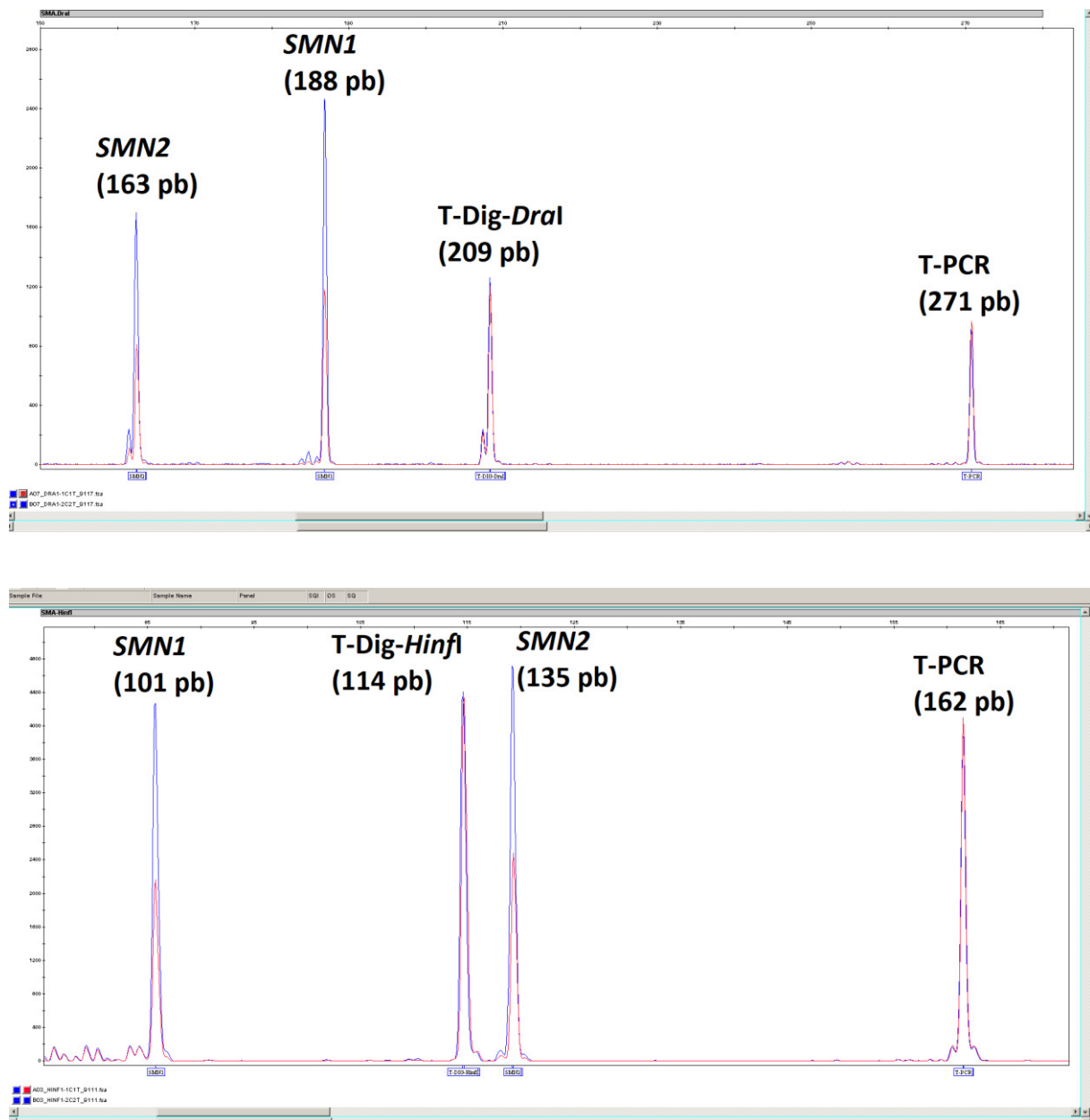


Figure 12 : Exemples d'électrophérogrammes de QMPSF *Dral* et *HinfI*

En haut, QMPSF *Dral*. En bas, QMPSF *HinfI*. Sujet porteur d'une copie du gène *SMN1* et d'une copie du gène *SMN2* (en rouge), dont les pics de fluorescence ont été alignés sur le pic témoin de digestion d'un contrôleur porteur de 2 copies du gène *SMN1* et 2 copies du gène *SMN2* (en bleu). En QMPSF *Dral*, l'amplicon des gènes *SMN* (188 pb) génère, après digestion, un fragment d'amplicon du gène *SMN2* fluorescent (163 pb), alors que l'amplicon du gène *SMN1* n'est pas scindé, tout comme le témoin de PCR (271 pb). En QMPSF *HinfI*, l'amplicon des gènes *SMN* (135 pb) génère, après digestion, un fragment d'amplicon du gène *SMN1* fluorescent (101 pb), alors que l'amplicon du gène *SMN2* n'est pas scindé, tout comme le témoin de PCR (162 pb).

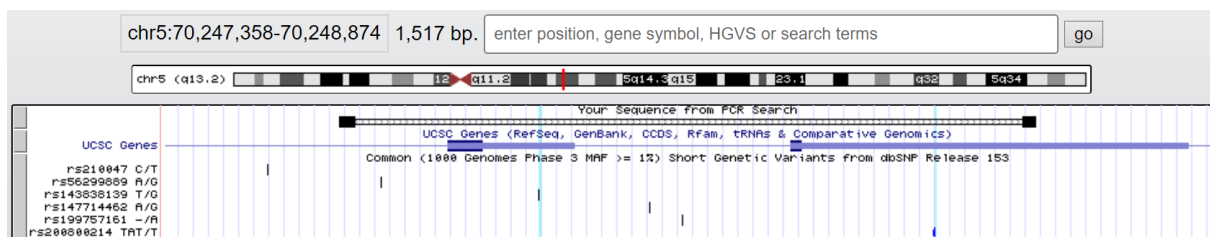


## SNapShot

Nous avons utilisé une technique de SNaPshot pour rechercher les variants 1 et 2 chez 154 des 199 sujets de l'étude.

### Principe

Une séquence commune aux gènes *SMN1* et *SMN2* incluant l'exon 7 et l'exon 8 UTR et comprenant les loci des deux variants recherchés a été amplifiée par PCR (Figure 13). Une réaction d'extension d'amorce est ensuite réalisée sur le produit de PCR purifiée. Les amorces de réaction d'extension d'amorce sont spécifiques d'une séquence située en 3' du variant recherché. L'extension d'amorce consiste en l'élongation d'une seule base à l'aide de ddNTP fluorescents, aboutissant ainsi à un oligonucléotide dont la dernière base en 3' correspond au génotype recherché. Après une étape de purification enzymatique, le génotype est obtenu par migration des amorces étendues sur un séquenceur capillaire 3130xl Applied Biosystem (Foster City, États-Unis), puis visualisé sur le logiciel Gene Mapper 4.1 (Thermo Fisher Scientific, Waltham, États-Unis).



**Figure 13 : Bornes de l'amplicon de *SMN1* généré et localisation des variants d'intérêt**

Cette représentation d'un des deux amplicons, celui s'alignant sur le gène *SMN1*, provient de l'outil en ligne : PCR *in silico* de UCSC Genome Browser (<https://genome.ucsc.edu>). Le couple d'amorces est représenté en haut par des rectangles noirs encadrant l'amplicon, représenté par un double trait. Les bornes de cet amplicon se situent au niveau de l'intron 6 et de la région flanquant le gène *SMN1* en 3'. Les deux variants recherchés, rs143838139 et rs200800214, sont surlignés en bleu.

## Etape de PCR

La PCR a été réalisée au moyen d'un kit de réaction Type-it 2X (Qiagen, Hilden, Allemagne), et de la paire d'amorce décrites dans la littérature (Alías *et al.*, 2018), avec un programme de PCR au plateau, comprenant 35 cycles (tableau 10).

Nom de l'amorce	Séquence de l'amorce	Tm
R111 (F)	5'-AGACTATCAACTTAATTTCTGATCA-3'	53.5°C
541C1120 (R)	5'-CTACAACACCCTTCTCACAG-3'	53.2°C

Tableau 10 : Amorces de PCR de SNaPShot

Le couple d'amorces de PCR de SNaPShot issues de la littérature (Alías *et al.*, 2018) inclue les deux variants recherchés dans un seul amplicon.

Cette réaction de PCR génère deux amplicons :

## UCSC In-Silico PCR

```
>chr5:70247611+70248621 1011bp AGACTATCAACTTAATTTCTGATCA CTACAACACCCTTCTCACAG
AGACTATCAACTTAATTTCTGATCAatattttggttgaataaaataagtaaa
atgtcttgtgaaacaaaatgctttttaacatccatataaaagctatctata
tatagctatctatgtctatataagctatatttttttaacttcctttattttc
cttacaggggtttcagacaaaatcaaaaagaaggaaggtgctcacattcct
taaattaaggagtaagtctgccagcattatgaaagtgaatcttacttttg
taaaactttatgggttgggaaacaaaatgtttttgaacatttaaaaagt
tcagatggttaaaaagttgaaaggttaatgtaaaacaatcaatattaaga
attttgatgcaaaaactattagataaaaagggttaatctacatccctactag
aattctcatacttaactgggtggttatgtggaagaaacatactttcacia
taaagagcttttaggatatgatgccattttatatcactagtaggcagacca
gcagactttttttatgtgatatgggataacctaggcatactgcactgt
acactctgacatatgaagtgtctagtcaagtttaactgggtgtccacaga
ggacatgggttaactggaattcgtcaagcctctgggttctaatttctcatt
tgcaggaaatgctggcatagagcagcactaaatgacaccactaaagaaac
gatcagacagatctggaatgtgaagcgttatagaagataactggcctcat
ttcttcaaaaatcaagtggtgggaaagaaaaaggaagtggaatgggta
actcttcttgattaaaagttatgtaataacaaaatgcaatgtgaaatatt
tactggactctattttgaaaaacatctgtaaaagactgggggtgggggt
gggaggccagcacggtggtgaggcagttgagaaaatttgaatgtggatta
gattttgaaatgatattggataattattggtaattttatgagCTGTGAGAA
GGGTGTTGTAG
>chr5:69372191+69373201 1011bp AGACTATCAACTTAATTTCTGATCA CTACAACACCCTTCTCACAG
AGACTATCAACTTAATTTCTGATCAatattttggttgaataaaataagtaaa
atgtcttgtgaaacaaaatgctttttaacatccatataaaagctatctata
tatagctatctatctatataagctatatttttttaacttcctttattttc
cttacaggggttttagacaaaatcaaaaagaaggaaggtgctcacattcct
taaattaaggagtaagtctgccagcattatgaaagtgaatcttacttttg
taaaactttatgggttgggaaacaaaatgtttttgaacatttaaaaagt
tcagatggttagaaaagttgaaaggttaatgtaaaacaatcaatattaaga
attttgatgcaaaaactattagataaaaagggttaatctacatccctactag
aattctcatacttaactgggtggttgggtggtggaagaaacatactttcacia
taaagagcttttaggatatgatgccattttatatcactagtaggcagacca
gcagactttttttatgtgatatgggataacctaggcatactgcactgt
acactctgacatatgaagtgtctagtcaagtttaactgggtgtccacaga
ggacatgggttaactggaattcgtcaagcctctgggttctaatttctcatt
tgcaggaaatgctggcatagagcagcactaaatgacaccactaaagaaac
gatcagacagatctggaatgtgaagcgttatagaagataactggcctcat
ttcttcaaaaatcaagtggtgggaaagaaaaaggaagtggaatgggta
actcttcttgattaaaagttatgtaataacaaaatgcaatgtgaaatatt
tactggactctattttgaaaaacatctgtaaaagactgagggtgggggt
gggaggccagcacggtggtgaggcagttgagaaaatttgaatgtggatta
gattttgaaatgatattggataattattggtaattttatgagCTGTGAGAA
GGGTGTTGTAG
```

Figure 14 : Amplicons générés

Les amorces de PCR de SNaPshot, reprises des données issues de la littérature, ne sont pas spécifiques du gène *SMN1* : comme illustré ci-dessus par la PCR *in silico* réalisée via l'outil en ligne d'UCSC Genome Browser (<http://genome-euro.ucsc.edu/cgi-bin/hgPcr>), le couple d'amorce utilisé s'aligne à la fois sur le gène *SMN1* et sur le gène *SMN2*, étant ainsi à l'origine de deux amplicons « copies ».

Les amplicons ainsi obtenus sont ensuite purifiés par un traitement enzymatique utilisant la SAP (Applied Biosystem, Foster City, États-Unis) et l'EXO1 (Roche, Bâle, Suisse). L'enzyme EXO1 permet la dégradation des amorces de PCR résiduelles. Quant à l'enzyme SAP, elle permet la déphosphorylation des dNTP résiduels. Cette technique de purification permet de procéder à l'étape d'extension d'amorce directement après.

#### Etape d'extension d'amorce de SNaPShot

Une amorce de SNaPShot a été conçue en amont de chaque variant recherché. De manière à pouvoir réaliser l'extension d'amorce en multiplex, les températures d'hybridation devaient être proches, c'est à dire avoir une différence de moins de 5°C. La température d'hybridation a été calculée pour cette étude via l'interface en ligne Primer 3. Des queues additionnelles sont ajoutées en 5' des amorces afin de les différencier au moment de la migration (Tableau 11).

Variant recherché	Séquence de la queue additionnelle 5'-3' (longueur en nt)	Séquence de l'amorce 5'-3' (longueur en nt)	Orientation (sens/antisens)	Tm (°C)	Séquence complète 5'-3'
Variant 1	AGGGTTAGATAGGGTTA (17 nt)	GGTTTGTGGAAAACAAATGTTTTGAACA (29 nt)	Sens	54,8	AGGGTTAGATAGGGTTAGGTTTGTGGAAAACAAATGTTTT TGAACA (46 nt)
Variant 2	ATAGGGTTAGATAGGGTTA (19 nt)	TATGTAATAACCAAATGAAATATTTTACTGGACTCT (36 nt)	Sens	55,1	ATAGGGTTAGATAGGGTTATATGTAATAACCAAATGAAAT ATTTTACTGGACTCT (55 nt)

Tableau 11 : Caractéristiques des amorces de SNaPshot conçues

La réaction de primer extension consiste en l'élongation de chaque amorce de SNaPshot d'une seule base couplée à un fluorochrome spécifique, permettant ainsi le génotypage de chacun des deux loci d'intérêt (Figure 15).

Les amorces ainsi étendues sont ensuite purifiées par un traitement enzymatique utilisant la SAP (Applied Biosystem, Foster City, États-Unis).

### Étape de migration

Après une étape de migration des amorces étendues sur le séquenceur Genetic Analyzers, leur temps de migration, la couleur du fluorochrome et l'intensité de la fluorescence sont représentées sur des électrophérogrammes par le logiciel Gene Mapper 4.1 (Thermo Fisher Scientific, Waltham, États-Unis) (Figure 16).

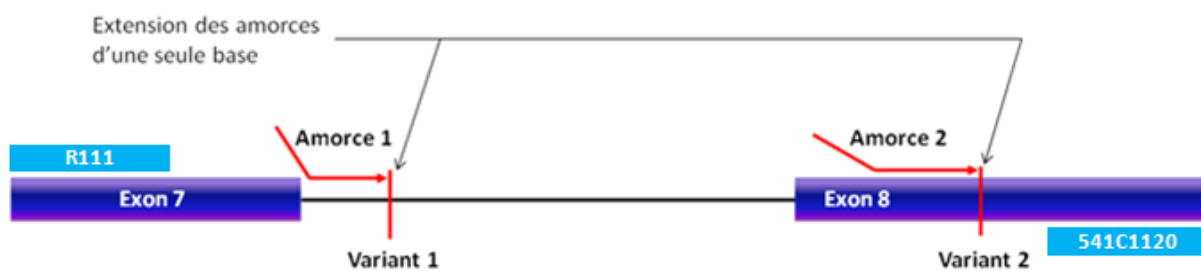


Figure 15 : Extension des amorces de SNaPshot

L'extension d'amorces de SNaPshot a été réalisée sur l'amplicon de 1011 pb incluant en la fin de l'exon 7, l'intron 7 et le début de l'exon 8. Ces amorces, conçues pour s'hybrider en amont de chacun des 2 variants, sont étendues d'une seule base, correspondant au génotype au locus de chaque variant. Les amorces de PCR utilisées pour l'amplification de ce fragment de 1011 pb sont schématisées en bleu clair : R111 (sens) et 541C1120 (antisens).

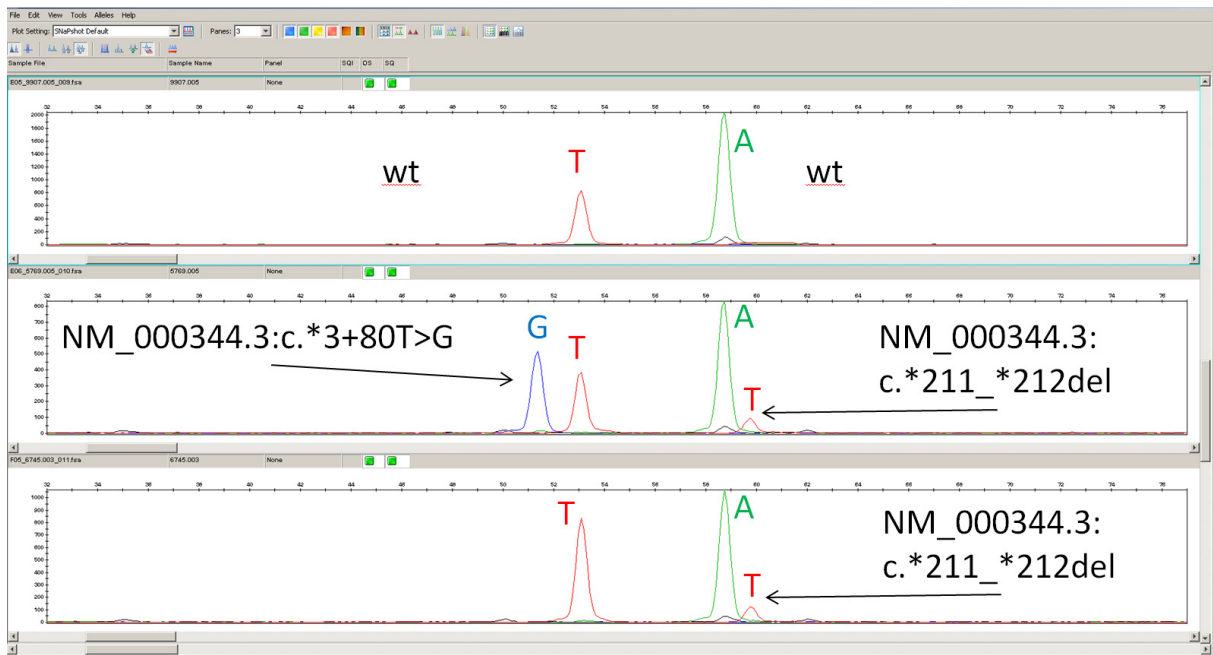


Figure 16 : Exemples d'électrophérogrammes de SNaPshot

Ci-dessus, quelques exemples d'électrophérogrammes de SNaPshot tels que nous les avons analysés. En haut : un sujet ayant un génotype sauvage, ne présentant aucun des deux variants recherchés ; Au milieu : un sujet présentant les deux variants recherchés ; En bas : un sujet présentant uniquement le variant 2 (seul sujet de l'étude ne présentant qu'un des deux variants).

### Kit AmplideX PCR/CE SMN1/2

Une partie des sujets (n = 45) a été explorée au moyen d'un kit AmplideX PCR/CE SMN1/2 (Asuragen, Austin, Etats-Unis).

### Principe

Le kit AmplideX PCR/CE SMN1/2 permet de quantifier les gènes *SMN1* et *SMN2*. Il s'agit d'une PCR multiplexe amplifiant l'exon 7 des gènes *SMN1* et *SMN2* de façon allèle spécifique, avec un gène de ménage (*EC*, *Endogenous Control*) à partir d'ADN génomique purifié. Le kit permet également de détecter certaines conversions de gènes *SMN1*-vers-*SMN2* et *SMN2*-vers-*SMN1* en comparant l'identité de séquence de l'exon 7 et de l'intron 7. Les éventuels pics de conversion génique entre *SMN1* et *SMN2* apparaissent comme des pics hybrides, identifiés par les génotypes spécifiques de l'exon 7, les pics hybrides *SMN1* indiquant la conversion du gène *SMN2* en *SMN1* et les pics hybrides *SMN2* indiquant la conversion du gène *SMN1* en *SMN2*.

De plus, le kit détermine la présence des variants 1 et 2, ainsi que le variant c.859G>C associé à une augmentation de l'inclusion de l'exon 7 des transcrits de *SMN2* (Prior *et al.*, 2009; Vezain *et al.*, 2011). Les amplicons marqués par fluorescence sont séparés par électrophorèse capillaire (CE, *Capillary Electrophoresis*), en fonction de leur taille, et classés comme gène de ménage (EC), *SMN1*, *SMN2*, *SMN1* hybride et/ou *SMN2* hybride, le gène étant déterminé par le statut de l'exon 7 (i.e. c.840C pour *SMN1* et c.840T pour *SMN2*), ainsi que les allèles de chacun des trois variants (variants 1 et 2, et c.859G>C). Le logiciel AmpliDeX PCR/CE Reporter (Asuragen, Austin, Etats-Unis) permet ensuite la visualisation des données. Cet outil logiciel effectue automatiquement un contrôle qualité par lots et échantillon, calcule les zones de pic, quantifie le nombre de copies d'exon 7 *SMN1* et *SMN2* et détermine la présence des variants recherchés directement à partir des fichiers d'électrophérogramme (Figure 17).

### Workflow

Les étapes de ce test comprennent la préparation du mix de PCR, l'amplification, puis l'analyse par électrophorèse capillaire. De l'ADN génomique purifié ou du sang total peut être ajouté dans un puits contenant le mix de PCR (PCR 2X) et les amorces (HEX SMN1/2 Plus) dans un volume de réaction final de 15 µL. Après environ 1 heure de PCR, les produits de PCR sont ajoutés à un deuxième mix composé de formamide Hi-Di™ et d'un marqueur de taille ROX 1000. Après une étape de dénaturation, les amplicons sont séparés sur un séquenceur capillaire.



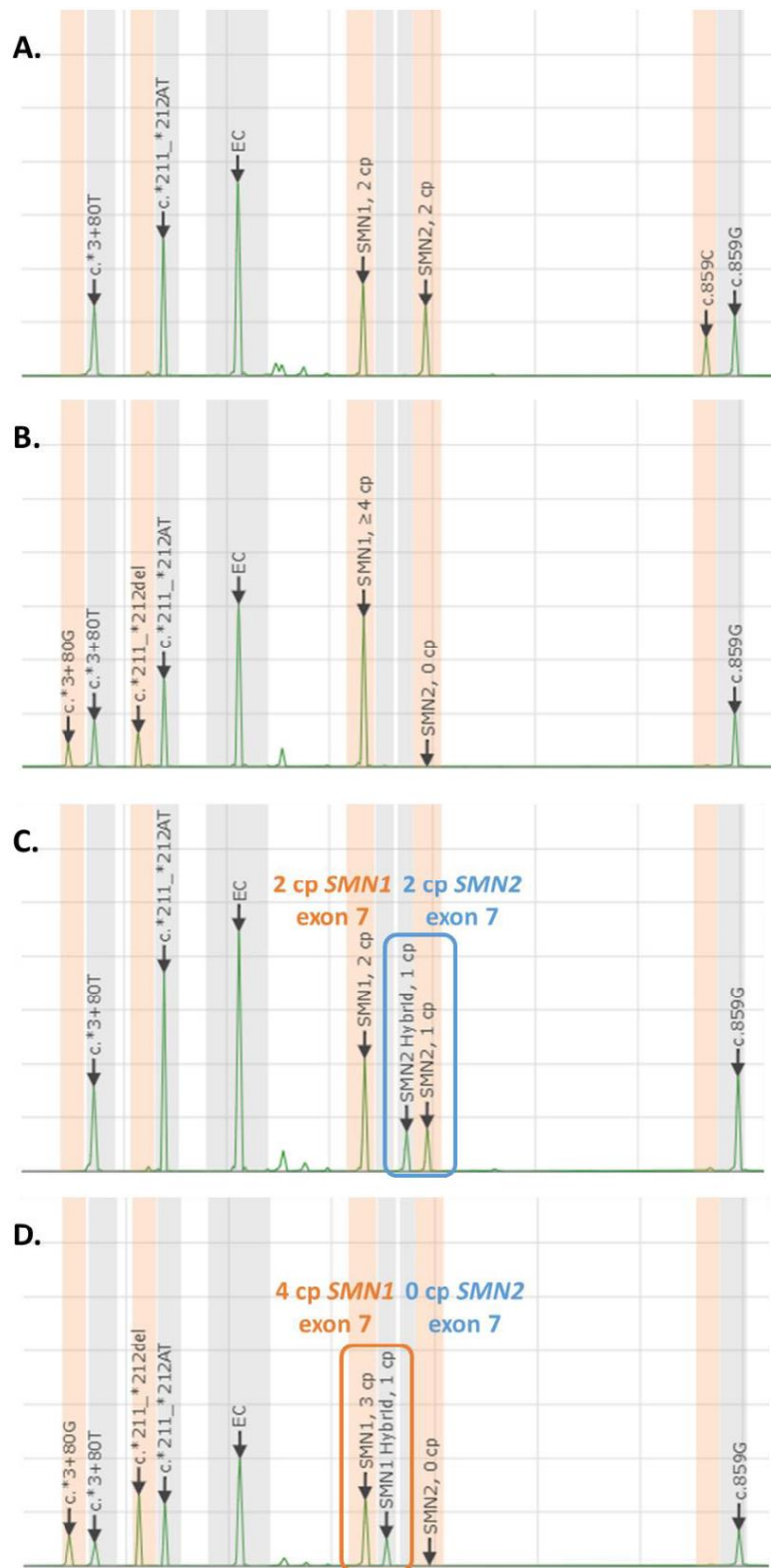


Figure 17 : Exemples de résultats issu du kit AmpliconX PCR/CE *SMN1/2*

A. Génotype 2 copies de *SMN1*, 2 copies de *SMN2*, avec le variant c.859C. B. Génotype 4 copies de *SMN1*, 0 copie de *SMN2*, avec les variants 1 et 2 présents. C. Génotype 2 copies de *SMN1*, 2 copies de *SMN2* qui comprend 1 copie de *SMN2* et 1 copie d'un hybride *SMN2*, indiquant une conversion du gène *SMN1* en *SMN2* dans l'exon 7. D. Génotype 4 copies de *SMN1*, 0 copie de *SMN2* qui comprend 3 copies de *SMN1* et 1 copie d'un hybride *SMN1*, indiquant une conversion du gène *SMN2* en *SMN1* dans l'exon 7 qui comprend également les variants 1 et 2.

## Analyse statistique

Calcul du nombre de sujets nécessaires. Le calcul du nombre de sujets nécessaires a été effectué via le site internet BiostaTGV (<https://biostatgv.sentiweb.fr>), en prenant en compte la proportion des porteurs des variants recherchés dans chacun des deux groupes - "dup en *cis*" et "absence de dup en *cis*" - observée après l'inclusion des premiers sujets, nous permettant ainsi d'avoir les proportions escomptées, ainsi que le risque de première espèce  $\alpha$  que nous avons fixé à 5 %, la puissance  $1-\beta$  fixée à 90 % et le caractère unilatéral du test.

Les expressions de « résultat positif » ou de « résultat négatif », habituelles en biostatistique, ont été utilisées pour désigner respectivement la « présence » ou « l'absence » des variants recherchés. Ainsi, le nombre de vrais positifs (VP), de faux positifs (FP), de vrais négatifs (VN) et de faux négatifs (FN) sont issus d'un tableau 2 x 2, indiquant en lignes, « présence d'au moins un des deux variants recherchés » ou « absence des variants recherchés », et indiquant en colonnes, « duplication en *cis* » ou « absence de duplication en *cis* ».

Nous avons, dans cette étude, considéré deux groupes qualitatifs et indépendants pour lesquels nous avons comparé deux proportions binomiales par un test exact de Fischer.

Les performances globales du test de recherche des variants ont ensuite été évaluées par :

- Ses caractéristiques intrinsèques que représentent sa sensibilité  $Se$  et sa spécificité  $Sp$ , calculées selon les formules suivantes :

$$Se = \frac{VP}{VP + FN}$$

$$Sp = \frac{VN}{FP + VN}$$

- Ses caractéristiques extrinsèques que représentent les valeurs prédictives positives et négatives - VPP et VPN - calculées selon les formules suivantes:

$$VPP = \frac{Se \times P}{Se \times P + (1 - P)(1 - Sp)}$$

$$VPN = \frac{Sp(1 - P)}{Sp(1 - P) + P(1 - Se)}$$

*Avec P : prévalence de la duplication en cis du gène SMN1 dans la population de l'étude*

La notion de « taux de détection » correspond à la probabilité pour un sujet d'être hétérozygote sachant qu'il possède deux copies du gène *SMN1*. Elle est déterminée par la capacité de la technique utilisée à déceler l'ensemble des variations causales de la SMA à l'état hétérozygote.

Le terme de « risque résiduel » correspond à la différence : (1 – taux de détection). Il peut être affiné grâce au théorème de Bayes, en prenant en compte le risque *a priori* pour un apparenté d'être hétérozygote compte tenu de son lien de parenté avec le cas index et les apparentés déjà analysés.

$$P(A|B) = \frac{P(B|A).P(A)}{P(B|A).P(A) + P(B|\bar{A}).P(\bar{A})}$$

Avec A : être hétérozygote

B : avoir 2 copies du gène *SMN1*

## Résultats

### Proportion de variants dans chaque sous-groupe

Les variants 1 et 2 étaient soit présents simultanément, soit absents. Seul un sujet était porteur du variant 1 et pas du variant 2. Ces deux variants n'ont pas été retrouvés chez les 59 sujets porteurs d'une seule copie du gène *SMN1* (génotype 1+0). Aucun des 46 sujets porteurs de 2 copies en *trans* du gène *SMN1* (génotype 1+1) ne présentait les variants. Ils étaient présents chez 5 des 25 sujets porteurs d'une duplication en *cis* du gène *SMN1* masquant une délétion sur le second allèle (génotype 2+0) (5/25 ; 20,0 %). Dans ce sous-groupe, notons qu'un sujet présentait le variant 1 mais pas le variant 2. Parmi les 56 sujets porteurs de 3 copies du gène *SMN1* (génotype 2+1), 6 présentaient les deux variants (6/56 ; 10,7 %). Le seul sujet de notre étude porteur d'une triplification en *cis* du gène *SMN1* (génotype 3+0) ne présentait aucun des deux variants. Cinq des 12 sujets porteurs de 4 copies du gène *SMN1* (génotype 2+2) étaient porteurs des variants (5/12 ; 41,7%) (Tableau 12).

	1+0 <i>SMN1</i>	2 <i>SMN1</i>	2+0 <i>SMN1</i>	3 <i>SMN1</i>	3+0 <i>SMN1</i>	4 <i>SMN1</i>
Variant 1	0	0	6	6	0	5
Variant 2	0	0	5	6	0	5
Total (N = 199)	59	46	25	56	1	12

Tableau 12 : Résultats de recherche des variants selon le génotype

Nous avons par la suite considéré le groupe dup en *cis* (n = 94 sujets). Ces sujets correspondent aux génotypes 2+0, 3 *SMN1*, 4 *SMN1* et 3+0. Dans ce groupe, 16 des 96 sujets porteurs d'une duplication ou d'une triplification en *cis* du gène *SMN1* présentaient les deux variants (16/94 ; 17,0 %). L'un d'entre eux ne présentait que le deuxième variant.

Parallèlement, nous avons considéré le groupe absence de dup en *cis* (n = 105 sujets). Ces sujets correspondent aux génotypes 1 *SMN1* et 2 *SMN1*. Aucun sujet n'était porteur des variants recherchés (0/105 ; 0,0 %) (Tableau 13).

	Dup en <i>cis</i>	Absence de dup en <i>cis</i>	Total
<b>Présence des variants</b>	16	0	16
<b>Absence des variants</b>	77	105	182
<b>Total</b>	93	105	199

Tableau 13 : Résultats de recherche des variants selon la présence ou l'absence d'une duplication en *cis* du gène *SMN1*

Le groupe dup en *cis* était donc enrichi en porteurs des variants 1 et 2 (  $P$  valeur =  $1,6.10^{-7}$ , test exact de Fischer).

### Absence d'association des variants recherchés à la duplication en *cis* du gène *SMN2*

Afin de nous assurer de la spécificité de l'association des variants 1 et 2 à la duplication en *cis* de *SMN1* et non à la duplication en *cis* de *SMN2* nous avons recherché leur présence chez 12 sujets qui présentaient une duplication en *cis* du gène *SMN2*, sans duplication en *cis* du gène *SMN1*. Parmi eux, 3 sujets étaient atteints de SMA avec 3 copies du gène *SMN2* et 0 copies du gène *SMN1*, 4 étaient porteurs de 3 copies du gène *SMN2* et d'une copie du gène *SMN1* et 5 sujets étaient porteurs de 3 copies du gène *SMN2* et de 2 copies en *trans* du gène *SMN1* (Tableau 14).

	3 <i>SMN2</i> , 0 <i>SMN1</i>	3 <i>SMN2</i> , 1 <i>SMN1</i>	3 <i>SMN2</i> , 2 <i>SMN1</i>
<b>Variant 1</b>	0	0	0
<b>Variant 2</b>	0	0	0
<b>Total (N = 12)</b>	3	4	5

Tableau 14 : Spécificité de l'association des variants recherchés à la duplication en *cis* du gène *SMN1*.

Ces résultats de recherche des variants par l'utilisation du kit AmpliDeX PCR-CE *SMN1/2* chez des sujets présentant une duplication en *cis* du gène *SMN2* mais pas de duplication en *cis* du gène *SMN1* confirment l'absence d'association entre les duplications en *cis* du gène *SMN2* et les deux variants de l'étude.

### Performances du test

La sensibilité est de 16/94 (17%), la spécificité est de 105/105 (100%) La valeur prédictive positive est de 100%, valeur résultant directement de celle de la Spécificité. La valeur prédictive négative est de

92% (avec une prévalence de la duplication en *cis* du gène *SMN1* de 7,5/10 000) à 94% (avec une prévalence de la duplication en *cis* du gène *SMN1* de 1/1000).

Dans une perspective d'utilisation des résultats de recherche des deux variants dans le cadre du conseil génétique, nous avons comparé les taux de détection des sujets hétérozygotes pour la délétion du gène *SMN1* sans et avec le résultat de recherche des variants. De manière similaire, nous avons calculé le risque résiduel sans et avec le résultat de recherche de ces deux variants. Le taux de détection des sujets hétérozygotes pour la délétion du gène *SMN1* est passé de 90 % avec une détermination du nombre de copies de ce gène par QMPSF à 94 % après exclusion des variants 1 et 2. Il résulte de ce nouveau taux de détection post-test une diminution du risque résiduel d'être hétérozygote pour la délétion du gène *SMN1* (Tableau 15).

Risque <i>a priori</i> d'hétérozygotie	Taux de détection	Risque résiduel	Taux de détection après haplotype négatif	Risque résiduel après haplotype négatif
<b>1 sur 2</b>	90 %	1 sur 11	94 %	1 sur 18
<b>1 sur 4</b>	90 %	1 sur 31	94 %	1 sur 51
<b>1 sur 8</b>	90 %	1 sur 71	94 %	1 sur 118
<b>1 sur 40</b>	90 %	1 sur 391	94 %	1 sur 651

**Tableau 15 : Amélioration du taux de détection et diminution du risque résiduel en cas d'haplotype négatif**

Le taux de détection est passé de 90 % à 94 % après exclusion des variants recherchés. Le risque résiduel s'en est trouvé diminué et est passé, par exemple pour un sujet dont le risque *a priori* d'être hétérozygote était de 1/2, à 1/18. Ce risque est passé à 1/51 pour un sujet ayant un risque à priori de 1/4, à 1/118 pour un sujet ayant un risque à priori de 1/8 et à 1/651 pour un sujet issu de la population générale avec un risque a priori de 1/40.

## Discussion et perspectives

L'arrivée des méthodes de quantification du nombre de copies du gène *SMN1* comme la QMPSF, a permis de révéler l'existence dans la population générale de sujets porteurs de trois copies du gène, c'est-à-dire ayant un chromosome 5 avec une duplication en *cis* de ce gène (Chen *et al.*, 1999; McAndrew *et al.*, 1997; Ogino *et al.*, 2001; Rochette *et al.*, 1997; Scheffer *et al.*, 2001; Wirth *et al.*, 1999). Cette considération implique que, dans les familles SMA où une délétion du gène *SMN1* est transmise aux apparentés, des sujets ayant 2 copies du gène *SMN1* peuvent être porteurs d'un génotype 2+0, non distinguable du génotype 1+1. L'existence de duplications en *cis* du gène *SMN1* dans la population constitue donc un piège majeur pour le conseil génétique de la SMA (Ogino *et al.*, 2002).

Notre étude fait suite aux travaux précédemment réalisés dans les populations juive ashkénaze et espagnole, qui ont révélé une association entre deux variants et des allèles ayant 2 copies du gène *SMN1* (Luo *et al.*, 2014; Alías *et al.*, 2018). Il s'agit d'une étude de réplique qui cherche à déterminer si ces résultats peuvent être extrapolés dans la population explorée dans les laboratoires de génétique moléculaire des CHU français : nous montrons que les deux variants y sont également associés à des duplications en *cis* du gène *SMN1* avec une spécificité élevée ( $Sp = 100\%$ ). De ce fait, c'est un haplotype dont la mise en évidence augmente le taux de détection des hétérozygotes pour la délétion du gène *SMN1*.

De manière à ne pas biaiser la spécificité du test, nous nous sommes assurés qu'aucun sujet assigné au groupe 1+1 n'ait en réalité un génotype 2+0. Une telle erreur d'assignation aurait en effet pu conduire à détecter les variants 1 et 2 chez des individus considérés à tort comme 1+1, ce qui aurait entraîné un effondrement de la spécificité du test, alors rendu inutilisable pour le conseil génétique. Ainsi, nous n'avons pas inclus au groupe 1+1 les apparentés à un cas-index. En effet, un apparenté à un cas-index atteint de SMA ayant 2 copies du gène *SMN1*, parce qu'il a un risque d'avoir hérité la délétion du gène *SMN1* familiale a un risque plus élevé d'avoir ses deux copies en *cis* que des sujets

issus de la population générale. Nous avons donc minimisé ce risque en ne sélectionnant que des conjoints d'apparentés, analysés dans le cadre d'un dépistage de couple.

Nous avons limité notre étude aux sujets explorés au CHU de Rouen, mais comme nous l'avons souligné précédemment, la fréquence des allèles porteurs de duplications en *cis* du gène *SMN1* varie de 4 à 8 % en fonction de la population étudiée. Ces allèles sont notamment plus fréquents dans la population africaine (Sugarman *et al.*, 2012), (Luo *et al.*, 2014). L'origine ethnique du patient a donc une incidence sur le taux de détection des sujets 2+0 et est donc à prendre en considération. Nous n'avons pas eu accès à l'origine ethnique des familles étudiées au titre du conseil génétique, ce qui constitue une limite de notre étude. Il serait utile dans des études rassemblant un nombre de sujets suffisamment grand, de stratifier les résultats en fonction de l'origine ethnique des sujets.

Nous avons donc montré que deux variants étaient associés dans la population française à des allèles ayant une duplication en *cis* du gène *SMN1* avec une sensibilité de 17 % et une spécificité de 100 %, que c'est sur ce gène que se situaient ces variants et qu'il n'existe aucune association avec les duplications en *cis* du gène *SMN2*. Ces résultats suggèrent l'existence d'un évènement fondateur dans la population africaine ayant conduit à une duplication en *cis* du gène *SMN1* (fréquence allélique 27 % en population africaine (Hendrickson *et al.*, 2009; Sugarman *et al.*, 2012)) sur un allèle porteur par ailleurs des 2 variants que nous avons explorés, NM\_000344.3(*SMN1*):c.\*3+80T>G et NM\_000344.3(*SMN1*):c.\*211\_\*212del (fréquences alléliques 27% et 29% respectivement dans la population africaine (<https://gnomad.broadinstitute.org/>)), constituant un haplotype. La duplication en *cis* du gène *SMN1* serait donc survenue après les variations 1 et 2. Plusieurs évènements à l'origine de duplications en *cis* du gène *SMN1*, certains associés aux variants 1 et 2 et d'autres non associés à ces variants, seraient survenus dans la population africaine, expliquant la moindre spécificité de cette association dans cette population. Cet haplotype se serait répandu dans d'autres populations, ayant une moindre diversité d'évènements originels de duplications en *cis* du gène *SMN1*. C'est l'importance du "groupe fondateur" qui aurait initialement déterminé la proportion de



ces haplotypes. Les prévalences différentes de ces haplotypes dans les différentes populations pourraient donc s'expliquer par l'importance de vagues migratoires.

Mis à part dans la population africaine, la recherche des variants chez des individus porteurs de deux copies ou plus du gène *SMN1* s'avère informatif et pourrait être intégré à l'analyse moléculaire de la SMA réalisée dans le cadre du conseil génétique. Idéalement, si la sensibilité du test avait été plus importante, il aurait été envisageable d'analyser uniquement l'apparenté sans son conjoint, ou bien l'apparenté puis le conjoint de manière séquentielle, dans le cadre d'un dépistage de couple. Toutefois, au vue de la valeur relativement basse de la sensibilité (17%), il ne semble pas prudent de se passer de l'analyse simultanée des deux conjoints. En effet, parmi les 25 sujets du groupe 2+0, seulement 5 étaient porteurs des variants. Il y a donc 80 % de faux négatifs. Par conséquent, il s'avère impossible d'utiliser cet haplotype pour trancher sur le statut d'hétérozygote. Cependant, cet haplotype pourrait être utilisé pour affiner le calcul du risque résiduel : chez un apparenté porteur de 2 copies du gène *SMN1*, le risque résiduel passerait de 10 %, sans effectuer le test de détection des variants, à 6 % si le test est réalisé et que le sujet n'est pas porteur des variants. Il ne serait possible d'analyser uniquement l'apparenté sans le conjoint qu'avec des techniques permettant de caractériser de manière plus certaine les duplications en *cis* du gène *SMN1*.

Du fait de la sensibilité insuffisante du test de détection des variants, nous n'envisageons le test de recherche des haplotypes que comme une étape transitoire avant l'arrivée de futures méthodes d'analyse qui permettraient un diagnostic direct de la présence d'une duplication. Pour visualiser ces duplications en *cis* sur un chromosome, une telle méthode devrait pouvoir permettre de caractériser des fragments d'ADN suffisamment longs pour contenir deux copies en *cis* du gène *SMN1*. De tels fragments peuvent être obtenus par extraction via un kit Chromium Genome reagent kit (10X Genomics, Pleasanton, États-Unis).

Après extraction, plusieurs techniques pourraient caractériser la présence de 2 copies de *SMN1* sur le même fragment : la PCR digitale, le peignage moléculaire, les technologies Nanopore/Bionano, voire le séquençage à haut débit de *reads* longs.

## Technologies alternatives à la recherche de l'haplotype

Il apparaît comme extrêmement utile de recourir à des techniques permettant de déterminer avec la plus grande certitude la phase des gènes *SMN1* pour chaque sujet dans le cadre du conseil génétique dans la famille d'un cas-index atteint de SMA. Il semble possible d'y parvenir grâce à l'utilisation de techniques qui séquentent ou caractérisent des fragments d'ADN « longs ». La longueur minimale de fragment d'ADN étudiée doit correspondre à celle séparant deux copies en *cis* du gène *SMN1*.

Toutefois, étant donné qu'il existe plusieurs événements fondateurs, il existe plusieurs architectures différentes de duplications en *cis* du gène *SMN1*. Nous n'avons à l'heure actuelle aucune donnée concernant l'architecture de ces remaniements et la distance séparant 2 copies en *cis* du gène *SMN1*.

Nous ne savons donc pas a priori si les techniques d'extraction de fragments d'ADN longs le seront suffisamment pour inclure 2 copies du gène sur un seul fragment. Mais ce n'est pas là le seul écueil, car il n'est pas certain que les différentes caractérisations ou « signatures » du gène *SMN1* soit par PCR digitale, peignage moléculaire, caractérisation Bionano ou Nanopore soient suffisamment résolutive pour discriminer les gènes *SMN1* et *SMN2* qui se retrouveraient sur les mêmes fragments.

Il apparaît donc nécessaire de déterminer si ces techniques peuvent permettre de distinguer les gènes *SMN1* et *SMN2*. Il faudrait par ailleurs cartographier l'architecture des différentes duplications en *cis* du gène *SMN1*, en vue de concevoir un test de dépistage simple des génotypes 2+0.

Il est donc essentiel d'explorer les techniques sus-citées en vue d'une amélioration du conseil génétique. Elles sont décrites ci-après.

### PCR digitale

La PCR digitale est une technique quantitative permettant de déterminer de façon absolue le nombre de copies d'une séquence donnée. Cette technique repose sur le principe de quantification par PCR

en utilisant la dilution limite et la loi statistique de Poisson. Ce principe a été décrit en 1992 par Sykes *et al.* (Sykes *et al.*, 1992). La loi de Poisson est une modélisation d'événements rares : on considère une variable aléatoire quantitative  $X$ , à valeurs dans  $\mathbb{N}$ .  $X$  se répartit selon une loi de Poisson de paramètre  $\lambda$  ( $\lambda > 0$ ), si et seulement si pour tout entier naturel,  $P(X = k) = e^{-\lambda} \left( \frac{\lambda^k}{k!} \right)$ . On note  $P(\lambda)$  cette loi, et on montre alors que son espérance, sa variance et son écart-type vérifient  $E(X) = \lambda$  ;  $\text{Var}(X) = \lambda$  et  $\sigma(X) = \sqrt{\lambda}$  (Bénazeth, 2004).

En divisant l'échantillon et le mélange réactionnel de qPCR dans un grand nombre de petits volumes distincts de telle sorte qu'il n'y ait que 0 ou 1 molécule cible par réaction individuelle (réacteur), la répartition des échantillons peut alors être modélisée par une loi de Poisson. Lorsqu'un réacteur contient au moins une molécule cible, alors cette molécule est amplifiée et révélée selon le principe de la qPCR. Chacun des réacteurs est analysé. Il est soit positif et prend la valeur 1, soit négatif et prend la valeur 0. À partir de la loi de Poisson, on peut déterminer la concentration initiale de molécules cibles :  $C \left[ \frac{\text{copies}}{\mu\text{L}} \right] = N \frac{\ln \frac{N}{N-X}}{V}$ , où  $N$  est le nombre total de réacteurs,  $X$  le nombre de réacteurs positifs et  $V$  le volume de la PCR digitale.

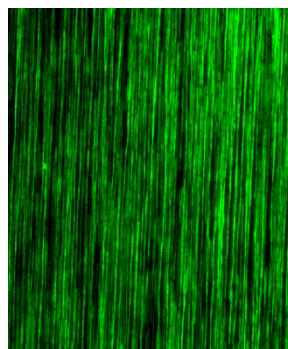
L'étape clé consiste donc à diviser l'échantillon et le mélange réactionnel dans un grand nombre de réacteurs. Deux méthodes différentes peuvent être utilisées (Vogelstein and Kinzler, 1999). La première est la PCR digitale sur puce, chaque puce contenant plusieurs milliers de puits qui interviennent chacun comme un réacteur. La seconde correspond à une PCR digitale en gouttelettes (ddPCR, *digital droplets PCR*). Ce système, dérivé des techniques de microfluidique, permet d'isoler l'échantillon dans des émulsions titrées, contenant de l'huile, de l'eau et un agent chimique stabilisateur, chaque goutte intervenant alors comme un réacteur. Cette technique a pour grand avantage de permettre une quantification absolue sans avoir à utiliser une gamme d'étalonnage.

Utilisée parallèlement sur des fragments d'ADN longs, à partir d'une première extraction, et courts, à partir d'une deuxième extraction, chez un sujet 2+0, la PCR digitale pourrait permettre de mettre en évidence un nombre de gouttelettes positives pour le gène *SMN1* deux fois plus important à partir

des fragments courts qu'à partir des fragments longs, signant un génotype 2+0. En effet, un fragment long comprenant la duplication en *cis* positiverait une gouttelette. Deux fragments courts, comprenant chacun une copie du gène *SMN1*, positiveraient deux gouttelettes. Cette même stratégie appliquée à un sujet 1+1 donnerait le même nombre de gouttelettes positives sur des fragments longs que sur des fragments courts.

### *Peignage moléculaire*

Le peignage moléculaire, mis au point par la société de biotechnologie Genomic Vision (Bagneux, France), aboutit à l'obtention, non pas d'une séquence nucléotidique, mais d'un pattern ou code morse génomique (*Genomic Morse Code, GMC*). En effet des sondes fluorescentes hybridées sur l'ADN génomique permettent de dresser la "cartographie en couleur" d'un locus. Cette technique consiste à baigner une lame de verre, préalablement traitée avec du silane vinylique, dans une solution d'ADN extrait, puis à la retirer à vitesse constante après une brève incubation permettant à l'ADN de se fixer à la surface du verre. Les forces de tension qui apparaissent alors au niveau du ménisque soit à l'interface air-liquide entraînent l'étirement et l'alignement des molécules d'ADN sur la lame (Figure 18). De cette façon, chaque molécule d'ADN est « peignée » avec un facteur d'étirement constant de 2 kb/ $\mu\text{m}$  permettant des mesures sur toute la longueur de la fibre d'ADN.



**Figure 18 : Aspect d'une lame de peignage moléculaire**

Aspect « peigné » des molécules d'ADN sur une lame silanisée après traction constante dans un ménisque air-liquide

L'ADN ainsi peigné est hybridé avec une combinaison de sondes fluorescentes de différentes couleurs et tailles, afin de reconnaître une certaine région d'intérêt. La séquence d'ADN à analyser est ainsi marquée avec la combinaison « de tirets et de points », créant un « *Morse Code* » spécifique, par exemple, à un gène cible et à ses régions flanquantes (Figure 19). La stratégie sous-jacente au GMC est d'utiliser la distribution spatiale des sondes pour fournir un complément d'informations au-delà des sondes elles-mêmes. La reconnaissance de différents motifs dans le *Morse Code* se base non seulement sur la taille et la couleur des sondes, mais aussi sur leur distribution et la distance entre elles. L'étirement identique de l'ADN permet des mesures de la longueur des sondes ainsi que des espaces qui les séparent. Une modification de la configuration observée par rapport au *Morse Code* de référence évoque la présence d'un réarrangement, éventuellement équilibré, dans la région cible (amplifications, délétions, répétitions, inversions et translocations) (Figure 20).

Il est possible que cette technique puisse permettre la détermination d'un GMC spécifique du gène *SMN1*, permettant de le distinguer du gène *SMN2*. Dans l'éventualité où cette distinction serait impossible, nous pourrions imaginer que la distance entre deux copies de gènes *SMN* puisse permettre de distinguer le gène *SMN2*, d'une deuxième copie en *cis* du gène *SMN1*. Mais cette stratégie se heurte au manque de connaissances sur l'architecture chromosomique des duplications en *cis* du gène *SMN1*.

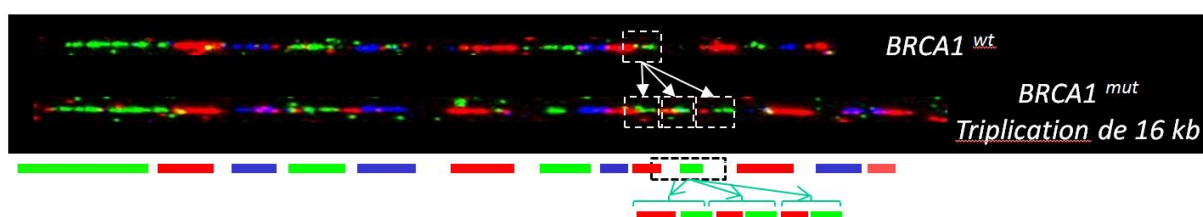


Figure 19 : Exemple d'image de Morse Code obtenue sur le gène *BRCA1*

Visualisation d'un pattern de morse code de référence sur le gène *BRCA1* (en haut) et d'un pattern de morse code correspondant à une triplification au sein de ce gène (en bas).

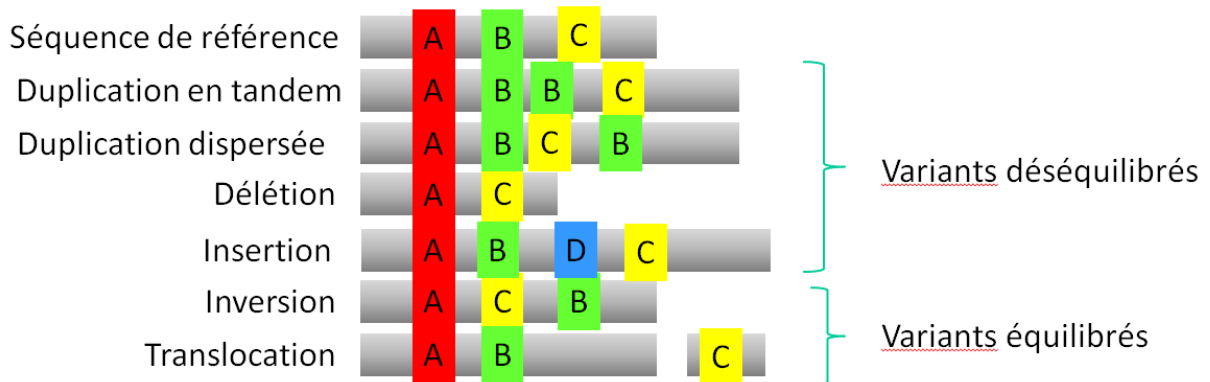


Figure 20 : Interprétation des Morse Codes obtenus

L'interprétation des *morse codes* permet la mise évidence des variants équilibrés comme déséquilibrés, de par la distance inter-sondes, la longueur des sondes, la couleur des sondes.

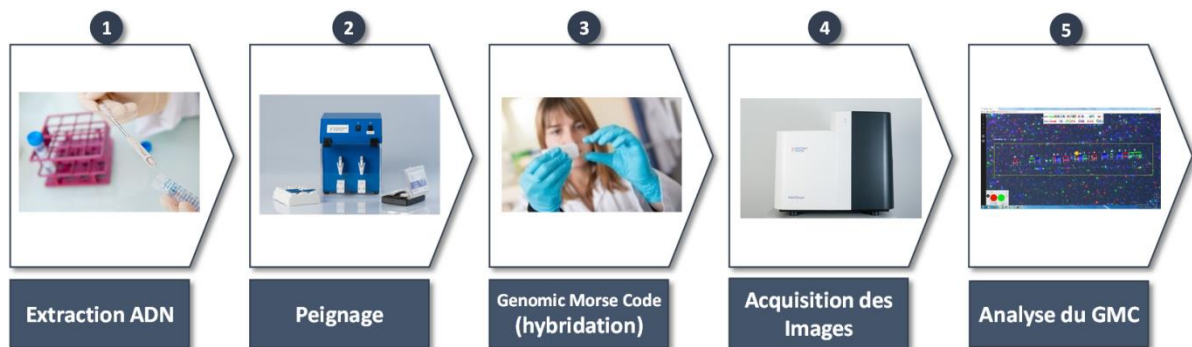


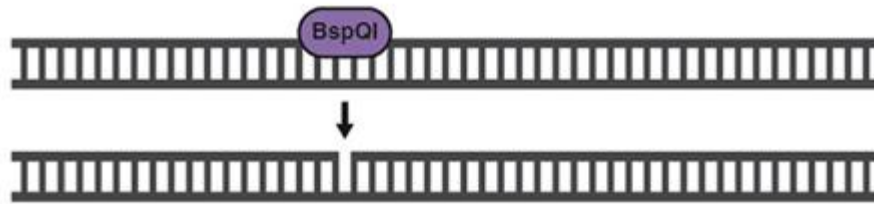
Figure 21 : Les 5 étapes du protocole de peignage moléculaire

### Technologie Bionano

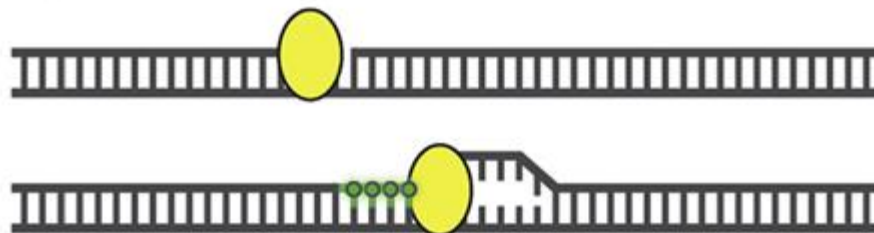
La technologie Bionano (Bionano Genomics, San Diego, États-Unis) permet l'obtention d'une cartographie optique de longs fragments d'ADN puis la reconstruction de la cartographie d'un chromosome entier. L'ADN est traité par des nickases (enzymes de restriction locus-spécifiques entraînant une coupure simple brin de l'ADN, ex : BspQI, BssSI) et l'élongation des brins excisés est réalisée par une Taq polymérase en présence de nucléotides fluorescents qui sont incorporés à la place des nucléotides excisés, puis une ligase répare les extrémités non liées du brin fluorescent nouvellement synthétisé, générant ainsi un pattern de fluorescence spécifique tout au long de la molécule d'ADNg (Figure 22). Ces réactifs sont issus du kit NLRs (*Nicks, Labels, Repairs and Stains*;

Bionano Genomics, San Diego, États-Unis). La longue molécule d'ADN issue de l'extraction tissulaire initiale reste donc intacte durant l'ensemble de ce processus.

1. Excision simple brin par une nickase (BspQI, BssSI)



2. Intégration de nucléotides fluorescents par la Taq polymérase



3. Réparation de l'extrémité 3' du brin fluorescent néosynthétisé



Figure 22 : Etape de préparation des échantillons avant l'analyse optique des données (Barseghyan, *et al.*, 2018)

Le long fragment d'ADNg, ainsi marqué aux niveaux des sites de restriction de la nickase par les dNTP fluorescents, est linéarisé dans une puce contenant un champ de nano-canaux intriqués (Saphyr Chip, Bionano) permettant de défaire la structure tertiaire de l'ADN et ne laissant passer qu'une double hélice à la fois (Figure 24).



Figure 23 : Puce Saphyr Chip

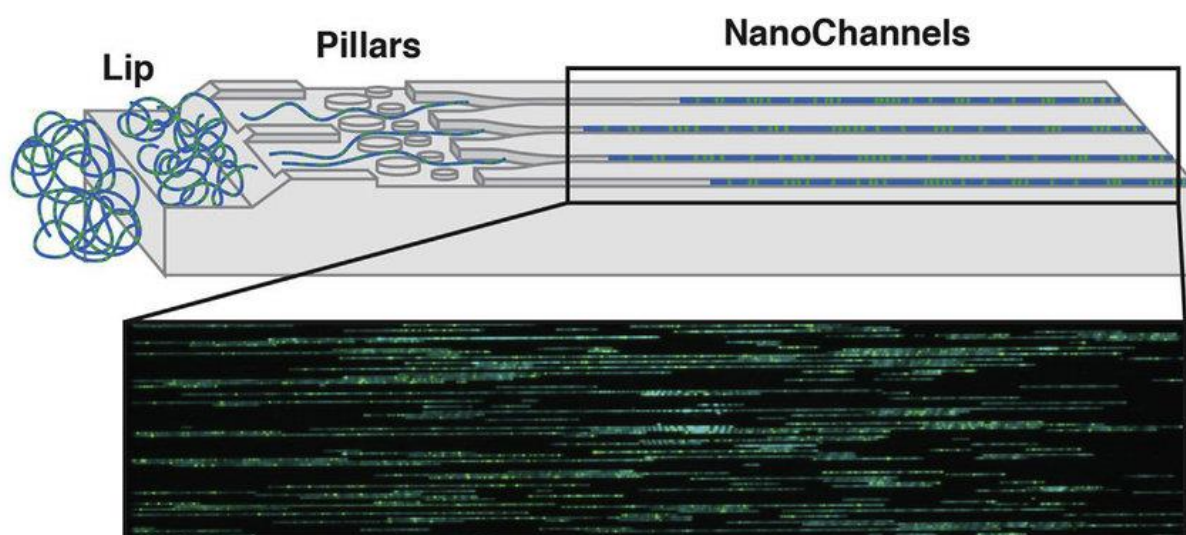


Figure 24 : Linéarisation des fragments d'ADNg

De longues molécules d'ADN de plusieurs kilobases sont marquées par des nucléotides fluorescents au niveau de sites reconnus par des nickases, créant des patterns optiques spécifiques de chaque locus. Les fragments d'ADN marqués sont linéarisés dans une puce (Saphyr Chip, Bionano) avant l'étape d'alignement sur une référence pour reconstituer la carte optique du chromosome, révélant d'éventuels variants structuraux.

Les patterns de fluorescence sont détectés et les régions chevauchantes entre les différents fragments d'ADN sont reconnues pour reconstruire la cartographie optique du chromosome entier (Figur 25). Cette cartographie est comparée, *in silico*, via le logiciel Bionano Access (Bionano Genomics, San Diego, États-Unis), à une référence, permettant la détection et le phasage de variants



structuraux (Chan *et al.*, 2018), tels que des insertions, des délétions, des expansions de séquences répétées, des duplications en tandem, ou des remaniements équilibrés (Figure 26).

Nous ne sommes pas certains que les duplications en *cis* du gène *SMN1* entraînent des configurations optiques spécifiques que nous pourrions distinguer des duplications en *cis* du gène *SMN2*. Il est probable qu'il faille coupler l'interprétation des configurations optiques aux résultats de la quantification des copies des gènes *SMN1* et *SMN2*. Dans l'éventualité où cette distinction serait impossible, nous pouvons imaginer que la distance entre deux copies de gènes *SMN* puisse permettre de distinguer le gène *SMN2*, d'une deuxième copie en *cis* du gène *SMN1*. Mais cette stratégie se heurte au manque de connaissances sur l'architecture chromosomique des duplications en *cis* du gène *SMN1*.

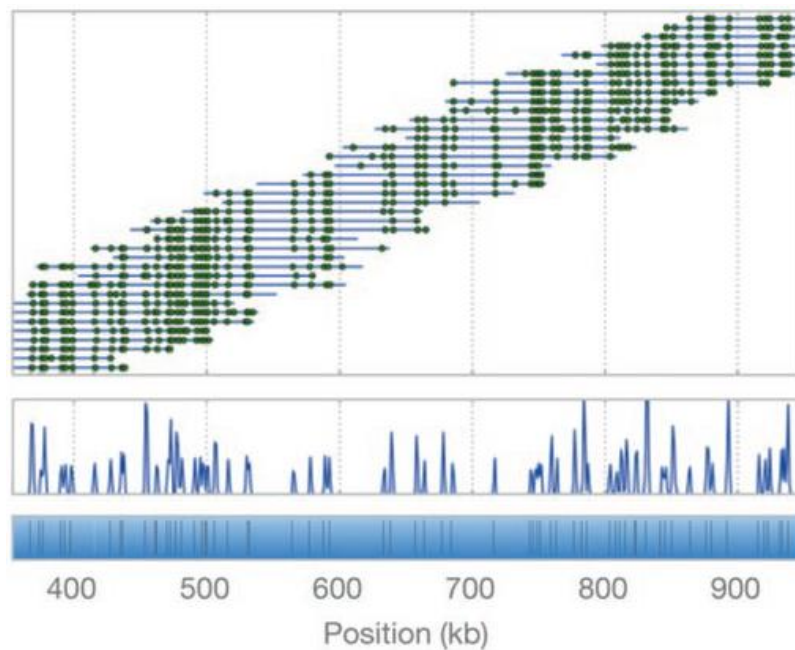


Figure 25 : Etape d'alignement et de reconstitution d'une cartographie consensus (Chan *et al.*, 2018)

Les fragments d'ADN marqués au niveau des sites de restriction de la nickase sont alignés sur une cartographie optique de référence via le logiciel Bionano Access, avant l'obtention d'une cartographie optique consensus et d'un phasage.

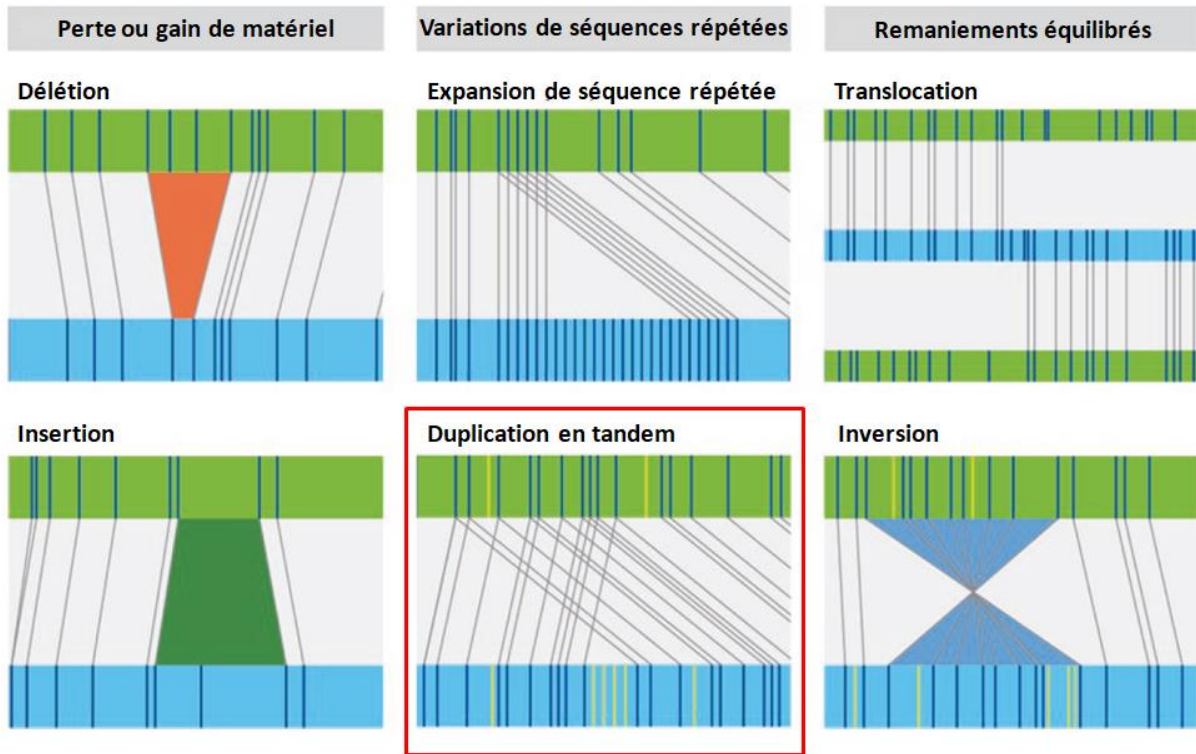


Figure 26 : Remaniements structuraux détectables par cartographie optique Bionano (Chan *et al.*, 2018)

L'algorithme utilisé par le logiciel Bionano Access permet le phasage et la détection des insertions, des délétions, des expansions de séquences répétées, des duplications en tandem, ou des remaniements équilibrés.

### Technologie Nanopore

La technologie Nanopore (Oxford Nanopore Technologies, Oxford, Royaume-Uni) repose sur la détermination d'un pattern de tension électrique de part et d'autre d'une membrane constituée de polymères de synthèse. La membrane sépare un compartiment *cis* et un compartiment *trans*, chacun rempli d'une solution ionique distincte, à l'origine d'une tension électrique de part et d'autre de la membrane. Des fragments d'ADNg, longs de 6000 à 48000 pb selon la méthode de fragmentation, sont déposés dans le compartiment *cis*.

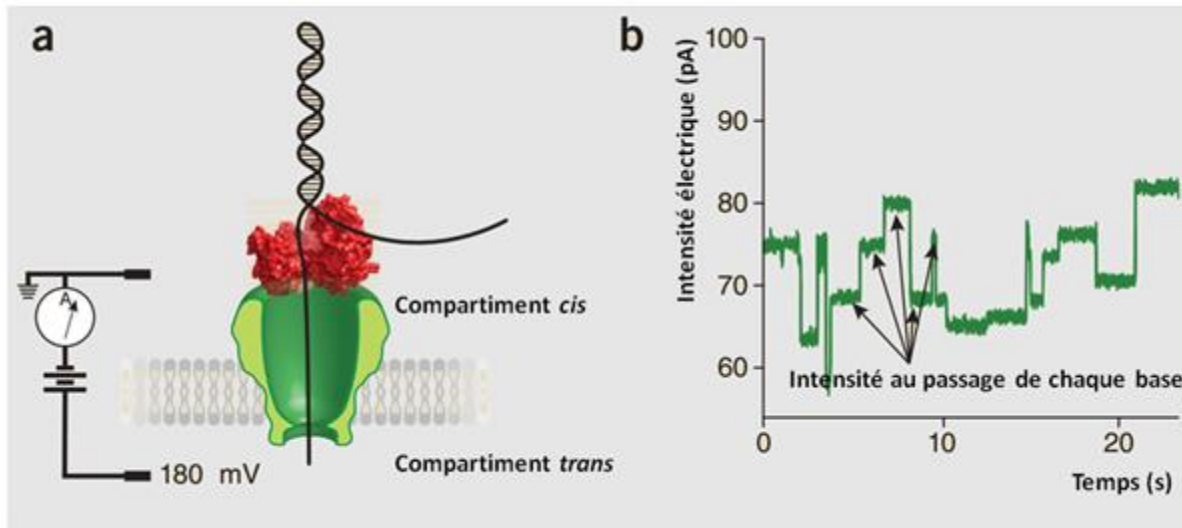


Figure 27 : Principe du séquençage Nanopore (Deamer *et al.*, 2016)

La technologie Nanopore permet le séquençage de longs fragments d'ADNg, grâce à une signature d'intensité de courant électrique spécifique à chaque base. a) L'ADNg extrait et fragmenté (6000 à 48000 pb selon la méthode de fragmentation utilisée) est placé dans le compartiment *cis* et migre dans le compartiment *trans* via un pore protéique (MspA ou  $\alpha$ -hemolysine, en vert sur le schéma) couplé à une hélicase (en rouge sur le schéma) qui linéarise et dénature l'ADN en une molécule simple brin. b) Chaque base génère à son passage un courant ionique entre les deux compartiments, se traduisant par un courant électrique mesurable. Il en résulte une signature d'intensité de courant électrique spécifique de chaque locus.



Figure 28 : Flow-cell Nanopore

La molécule d'ADN double brin migre du compartiment *cis* vers le compartiment *trans* dans un gradient de tension électrique. Elle est dénaturée en molécule simple brin et linéarisée par une hélicase couplée au pore (MspA,  $\alpha$ -hemolysine), lors de son passage (Figure 27). A son passage, chaque nucléotide entraîne une perturbation du courant ionique qui lui est caractéristique et par là même de l'intensité électrique de part et d'autre de la membrane. Les successions de perturbations

propres à chaque nucléotide permettent *in fine* un séquençage du brin d'ADN. Une *flow-cell* contient 2048 nanopores en parallèle individualisés les uns des autres (Deamer *et al.*, 2016).

La technologie Nanopore étant une technique de séquençage contrairement au peignage moléculaire et à la technologie Bionano, il est probable qu'elle permette de distinguer les gènes *SMN1* et *SMN2* de par leurs 5 différences nucléotidiques et ainsi de déterminer le nombre de copies de ces gènes sur des fragments suffisamment longs pour inclure 2 copies en *cis* du gène *SMN1*.

Ainsi, si tant est que les fragments d'ADNg étudiés soient suffisamment longs, les technologies de PCR digitale, de peignage moléculaire, Bionano ou Nanopore pourraient constituer des alternatives à la recherche des variants en mettant en évidence directement une duplication en *cis* du gène *SMN1*, soit par une discordance entre nombre de copies du gène *SMN1* entre des fragments courts et longs issus d'un même sujet (PCR digitale), soit par une cartographie optique ou électrique du locus *SMN1* comprenant les deux copies de la duplication en *cis* de ce gène (peignage moléculaire, Bionano, Nanopore). Ces technologies restent toutefois confrontées à l'homologie importante entre les gènes *SMN1* et *SMN2* et pourraient ne pas permettre de les distinguer.

Par ailleurs, ces techniques sont lourdes et coûteuses si bien qu'elles semblent difficilement utilisables en pratique diagnostique. Nous pourrions imaginer qu'après une étape de description des architectures des duplications en *cis*, il soit possible de convertir ces résultats en test simples et non coûteux.

## Conclusion

Nous avons montré l'association de deux variants, préalablement décrits dans la littérature, NM\_000344.3(*SMN1*):c.\*3+80T>G et NM\_000344.3(*SMN1*):c.\*211\_\*212del, à des duplications en *cis* du gène *SMN1*, chez des sujets analysés au laboratoire de génétique moléculaire du CHU de Rouen dans le cadre du conseil génétique autour d'un cas atteint d'amyotrophie spinale infantile. Cette association s'est avérée être très spécifique ( $Sp = 100\%$ ) mais relativement peu sensible ( $Se = 17\%$ ), ne permettant pas d'exclure de façon certaine la présence d'une duplication en *cis* du gène *SMN1* chez un apparenté porteur de 2 copies de ce gène. Nous pouvons toutefois intégrer les données de recherche des variants 1 et 2 au calcul du risque résiduel d'être hétérozygote pour un individu porteur de 2 copies du gène *SMN1*. La diminution du risque résiduel qui passe de 10 % à 6 % chez un sujet porteur de 2 copies du gène *SMN1* et porteur par ailleurs des variants 1 et 2 n'apparaît pas suffisamment importante pour se passer de l'analyse du conjoint en cas de projet parental.

Un diagnostic moléculaire plus direct de duplication en *cis* du gène *SMN1* pourrait être apporté par des techniques permettant de séquencer ou de cartographier des fragments d'ADN génomique suffisamment longs pour inclure deux copies en *cis* du gène *SMN1*, remaniement dont nous ignorons actuellement l'architecture précise. Toutefois, ces analyses de fragments longs restent confrontées à un possible manque de résolution, c'est à dire à une éventuelle impossibilité de distinguer les gènes *SMN1* et *SMN2*, hautement homologues. Ainsi, nous pourrions envisager de combiner ces techniques avec les méthodes de quantification actuelles du nombre de copies des gènes *SMN1* et *SMN2* pour mieux caractériser cette région complexe de notre génome et répondre le plus précisément possible aux apparentés à un cas-index atteint de SMA demandeurs d'un conseil génétique.

## **Bibliographie**

- Alías, L., Barceló, M.J., Bernal, S., Martínez-Hernández, R., Also-Rallo, E., Vázquez, C., Santana, A., Millán, J.M., Baiget, M., Tizzano, E.F., 2014. Improving detection and genetic counseling in carriers of spinal muscular atrophy with two copies of the SMN1 gene. *Clin. Genet.* 85, 470–475. <https://doi.org/10.1111/cge.12222>
- Alías, L., Bernal, S., Calucho, M., Martínez, E., March, F., Gallano, P., Fuentes-Prior, P., Abuli, A., Serra-Juhe, C., Tizzano, E.F., 2018. Utility of two SMN1 variants to improve spinal muscular atrophy carrier diagnosis and genetic counselling. *Eur. J. Hum. Genet.* 26, 1554–1557. <https://doi.org/10.1038/s41431-018-0193-4>
- Alías, L., Bernal, S., Fuentes-Prior, P., Barceló, M.J., Also, E., Martínez-Hernández, R., Rodríguez-Alvarez, F.J., Martín, Y., Aller, E., Grau, E., Peciña, A., Antiñolo, G., Galán, E., Rosa, A.L., Fernández-Burriel, M., Borrego, S., Millán, J.M., Hernández-Chico, C., Baiget, M., Tizzano, E.F., 2009. Mutation update of spinal muscular atrophy in Spain: molecular characterization of 745 unrelated patients and identification of four novel mutations in the SMN1 gene. *Hum. Genet.* 125, 29–39. <https://doi.org/10.1007/s00439-008-0598-1>
- Azad, A.K., Huang, C.-K., Jin, H., Zou, H., Yanakakis, L., Du, J., Fiddler, M., Naeem, R., Goldstein, Y., 2019. Enhanced Carrier Screening for Spinal Muscular Atrophy: Detection of Silent (SMN1: 2 + 0) Carriers Utilizing a Novel TaqMan Genotyping Method. *Lab Med.* <https://doi.org/10.1093/labmed/lmz088>
- Bénazeth, S., 2004. *Biomathématiques: pharmacie, médecine 1re & 2e années.* Masson, Paris.
- Bertini, E., Burghes, A., Bushby, K., Estournet-Mathiaud, B., Finkel, R.S., Hughes, R. a. C., Iannaccone, S.T., Melki, J., Mercuri, E., Muntoni, F., Voit, T., Reitter, B., Swoboda, K.J., Tiziano, D., Tizzano, E., Topaloglu, H., Wirth, B., Zerres, K., 2005. 134th ENMC International Workshop: Outcome Measures and Treatment of Spinal Muscular Atrophy, 11-13 February 2005, Naarden, The Netherlands. *Neuromuscul. Disord.* 15, 802–816. <https://doi.org/10.1016/j.nmd.2005.07.005>
- Burghes, A.H.M., Beattie, C.E., 2009. Spinal muscular atrophy: why do low levels of survival motor neuron protein make motor neurons sick? *Nat. Rev. Neurosci.* 10, 597–609. <https://doi.org/10.1038/nrn2670>
- Cartegni, L., Krainer, A.R., 2002. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. *Nat. Genet.* 30, 377–384. <https://doi.org/10.1038/ng854>
- Castiglioni, C., Levicán, J., Rodillo, E., Garmendia, M.A., Díaz, A., Pizarro, L., Contreras, L., 2011. [Clinical, electrophysiological and molecular study of 26 Chilean patients with spinal muscular atrophy]. *Rev Med Chil* 139, 197–204. <https://doi.org/S0034-98872011000200009>
- Chan, S., Lam, E., Saghbini, M., Bocklandt, S., Hastie, A., Cao, H., Holmlin, E., Borodkin, M., 2018. Structural Variation Detection and Analysis Using Bionano Optical Mapping. *Methods Mol. Biol.* 1833, 193–203. [https://doi.org/10.1007/978-1-4939-8666-8\\_16](https://doi.org/10.1007/978-1-4939-8666-8_16)
- Charbonnier, F., Raux, G., Wang, Q., Drouot, N., Cordier, F., Limacher, J.M., Saurin, J.C., Puisieux, A., Olschwang, S., Frebourg, T., 2000. Detection of exon deletions and duplications of the mismatch repair genes in hereditary nonpolyposis colorectal cancer families using multiplex polymerase chain reaction of short fluorescent fragments. *Cancer Res.* 60, 2760–2763.

- Chen, K.L., Wang, Y.L., Rennert, H., Joshi, I., Mills, J.K., Leonard, D.G., Wilson, R.B., 1999. Duplications and de novo deletions of the SMN1 gene demonstrated by fluorescence-based carrier testing for spinal muscular atrophy. *Am. J. Med. Genet.* 85, 463–469.
- Cuisset, J.-M., Estournet, B., French Ministry of Health, 2012. Recommendations for the diagnosis and management of typical childhood spinal muscular atrophy. *Rev. Neurol. (Paris)* 168, 902–909. <https://doi.org/10.1016/j.neurol.2012.07.020>
- Cuscó, I., Barceló, M.J., del Rio, E., Martín, Y., Hernández-Chico, C., Bussaglia, E., Baiget, M., Tizzano, E.F., 2001. Characterisation of SMN hybrid genes in Spanish SMA patients: de novo, homozygous and compound heterozygous cases. *Hum. Genet.* 108, 222–229. <https://doi.org/10.1007/s004390000452>
- D'Amico, A., Mercuri, E., Tizzano, F.D., Bertini, E., 2011. Spinal muscular atrophy. *Orphanet J Rare Dis* 6, 71. <https://doi.org/10.1186/1750-1172-6-71>
- Deamer, D., Akeson, M., Branton, D., 2016. Three decades of nanopore sequencing. *Nat. Biotechnol.* 34, 518–524. <https://doi.org/10.1038/nbt.3423>
- Feldkötter, M., Schwarzer, V., Wirth, R., Wienker, T.F., Wirth, B., 2002. Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *Am. J. Hum. Genet.* 70, 358–368. <https://doi.org/10.1086/338627>
- Grotto, S., Cuisset, J.-M., Marret, S., Drunat, S., Faure, P., Audebert-Bellanger, S., Desguerre, I., Flurin, V., Grebille, A.-G., Guerrot, A.-M., Journel, H., Morin, G., Plessis, G., Renolleau, S., Roume, J., Simon-Bouy, B., Touraine, R., Willems, M., Frébourg, T., Verspyck, E., Saugier-veber, P., 2016. Type 0 Spinal Muscular Atrophy: Further Delineation of Prenatal and Postnatal Features in 16 Patients. *J Neuromuscul Dis* 3, 487–495. <https://doi.org/10.3233/JND-160177>
- Hahnen, E., Schönling, J., Rudnik-Schöneborn, S., Zerres, K., Wirth, B., 1996. Hybrid survival motor neuron genes in patients with autosomal recessive spinal muscular atrophy: new insights into molecular mechanisms responsible for the disease. *Am. J. Hum. Genet.* 59, 1057–1065.
- Hausmanowa-Petrusewicz, I., 1988. Electrophysiological findings in childhood spinal muscular atrophies. *Rev. Neurol. (Paris)* 144, 716–720.
- Hendrickson, B.C., Donohoe, C., Akmaev, V.R., Sugarman, E.A., Labrousse, P., Boguslavskiy, L., Flynn, K., Rohlf, E.M., Walker, A., Allitto, B., Sears, C., Scholl, T., 2009. Differences in SMN1 allele frequencies among ethnic groups within North America. *J. Med. Genet.* 46, 641–644. <https://doi.org/10.1136/jmg.2009.066969>
- Hoffmann, J., 1900. Ueber die hereditäre progressive spinale Muskelatrophie im Kindesalter. *München Med. Wschr.* 47.
- Hoffmann, J., 1893. Ueber chronische spinale Muskelatrophie im Kindesalter, auf familiärer Basis. *Dtsch. Z Nervenheilk* 3, 427–470.
- Kashima, T., Manley, J.L., 2003. A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. *Nat. Genet.* 34, 460–463. <https://doi.org/10.1038/ng1207>

- Kugelberg, E., Welander, L., 1956. Heredofamilial juvenile muscular atrophy simulating muscular dystrophy. *AMA Arch Neurol Psychiatry* 75, 500–509. <https://doi.org/10.1001/archneurpsyc.1956.02330230050005>
- Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., 1995. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80, 155–165. [https://doi.org/10.1016/0092-8674\(95\)90460-3](https://doi.org/10.1016/0092-8674(95)90460-3)
- Lefebvre, S., Burlet, P., Liu, Q., Bertrand, S., Clermont, O., Munnich, A., Dreyfuss, G., Melki, J., 1997. Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat. Genet.* 16, 265–269. <https://doi.org/10.1038/ng0797-265>
- Liu, Q., Dreyfuss, G., 1996. A novel nuclear structure containing the survival of motor neurons protein. *EMBO J.* 15, 3555–3565.
- Lorson, C.L., Hahnen, E., Androphy, E.J., Wirth, B., 1999. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc. Natl. Acad. Sci. U.S.A.* 96, 6307–6311. <https://doi.org/10.1073/pnas.96.11.6307>
- Luo, M., Liu, L., Peter, I., Zhu, J., Scott, S.A., Zhao, G., Eversley, C., Kornreich, R., Desnick, R.J., Edelmann, L., 2014. An Ashkenazi Jewish SMN1 haplotype specific to duplication alleles improves pan-ethnic carrier screening for spinal muscular atrophy. *Genet. Med.* 16, 149–156. <https://doi.org/10.1038/gim.2013.84>
- Mailman, M.D., Heinz, J.W., Papp, A.C., Snyder, P.J., Sedra, M.S., Wirth, B., Burghes, A.H.M., Prior, T.W., 2002. Molecular analysis of spinal muscular atrophy and modification of the phenotype by SMN2. *Genet. Med.* 4, 20–26. <https://doi.org/10.1097/00125817-200201000-00004>
- McAndrew, P.E., Parsons, D.W., Simard, L.R., Rochette, C., Ray, P.N., Mendell, J.R., Prior, T.W., Burghes, A.H., 1997. Identification of proximal spinal muscular atrophy carriers and patients by analysis of SMN1 and SMN2 gene copy number. *Am. J. Hum. Genet.* 60, 1411–1422. <https://doi.org/10.1086/515465>
- Monani, U.R., Lorson, C.L., Parsons, D.W., Prior, T.W., Androphy, E.J., Burghes, A.H., McPherson, J.D., 1999. A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Hum. Mol. Genet.* 8, 1177–1183. <https://doi.org/10.1093/hmg/8.7.1177>
- Munsat, T.L., Davies, K.E., 1992. International SMA consortium meeting. (26-28 June 1992, Bonn, Germany). *Neuromuscul. Disord.* 2, 423–428. [https://doi.org/10.1016/s0960-8966\(06\)80015-5](https://doi.org/10.1016/s0960-8966(06)80015-5)
- Ogino, S., Leonard, D.G., Rennert, H., Gao, S., Wilson, R.B., 2001. Heteroduplex formation in SMN gene dosage analysis. *J Mol Diagn* 3, 150–157. [https://doi.org/10.1016/S1525-1578\(10\)60666-6](https://doi.org/10.1016/S1525-1578(10)60666-6)
- Ogino, S., Leonard, D.G.B., Rennert, H., Ewens, W.J., Wilson, R.B., 2002. Genetic risk assessment in carrier testing for spinal muscular atrophy. *Am. J. Med. Genet.* 110, 301–307. <https://doi.org/10.1002/ajmg.10425>
- Ogino, S., Wilson, R.B., Gold, B., 2004. New insights on the evolution of the SMN1 and SMN2 region: simulation and meta-analysis for allele and haplotype frequency calculations. *Eur. J. Hum. Genet.* 12, 1015–1023. <https://doi.org/10.1038/sj.ejhg.5201288>



- Paushkin, S., Gubitz, A.K., Massenet, S., Dreyfuss, G., 2002. The SMN complex, an assemblyosome of ribonucleoproteins. *Curr. Opin. Cell Biol.* 14, 305–312. [https://doi.org/10.1016/s0955-0674\(02\)00332-0](https://doi.org/10.1016/s0955-0674(02)00332-0)
- Prior, T.W., Krainer, A.R., Hua, Y., Swoboda, K.J., Snyder, P.C., Bridgeman, S.J., Burghes, A.H.M., Kissel, J.T., 2009. A positive modifier of spinal muscular atrophy in the SMN2 gene. *Am. J. Hum. Genet.* 85, 408–413. <https://doi.org/10.1016/j.ajhg.2009.08.002>
- Prior, T.W., Snyder, P.J., Rink, B.D., Pearl, D.K., Pyatt, R.E., Mihal, D.C., Conlan, T., Schmalz, B., Montgomery, L., Ziegler, K., Noonan, C., Hashimoto, S., Garner, S., 2010. Newborn and carrier screening for spinal muscular atrophy. *Am. J. Med. Genet. A* 152A, 1608–1616. <https://doi.org/10.1002/ajmg.a.33474>
- Rochette, C.F., Surh, L.C., Ray, P.N., McAndrew, P.E., Prior, T.W., Burghes, A.H., Vanasse, M., Simard, L.R., 1997. Molecular diagnosis of non-deletion SMA patients using quantitative PCR of SMN exon 7. *Neurogenetics* 1, 141–147. <https://doi.org/10.1007/s100480050021>
- Saugier-veber, P., Drouot, N., Lefebvre, S., Charbonnier, F., Vial, E., Munnich, A., Frébourg, T., 2001. Detection of heterozygous SMN1 deletions in SMA families using a simple fluorescent multiplex PCR method. *J. Med. Genet.* 38, 240–243. <https://doi.org/10.1136/jmg.38.4.240>
- Scheffer, H., Cobben, J.M., Matthijs, G., Wirth, B., 2001. Best practice guidelines for molecular analysis in spinal muscular atrophy. *Eur. J. Hum. Genet.* 9, 484–491. <https://doi.org/10.1038/sj.ejhg.5200667>
- Singh, R.N., Howell, M.D., Ottesen, E.W., Singh, N.N., 2017. Diverse role of survival motor neuron protein. *Biochim Biophys Acta Gene Regul Mech* 1860, 299–315. <https://doi.org/10.1016/j.bbagr.2016.12.008>
- Sugarman, E.A., Nagan, N., Zhu, H., Akmaev, V.R., Zhou, Z., Rohlf, E.M., Flynn, K., Hendrickson, B.C., Scholl, T., Sirko-Osada, D.A., Allitto, B.A., 2012. Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: clinical laboratory analysis of >72,400 specimens. *Eur. J. Hum. Genet.* 20, 27–32. <https://doi.org/10.1038/ejhg.2011.134>
- Sykes, P.J., Neoh, S.H., Brisco, M.J., Hughes, E., Condon, J., Morley, A.A., 1992. Quantitation of targets for PCR by use of limiting dilution. *BioTechniques* 13, 444–449.
- Thauvin-Robinet, C., Drunat, S., Saugier-veber, P., Chantereau, D., Cossée, M., Cassini, C., Soichot, P., Masurel-Paulet, A., De Monléon, J.V., Sagot, P., Huet, F., Antin, M., Calmels, N., Faivre, L., Gérard, B., “réseau français de génétique moléculaire,” 2012. Homozygous SMN1 exons 1-6 deletion: pitfalls in genetic counseling and general recommendations for spinal muscular atrophy molecular diagnosis. *Am. J. Med. Genet. A* 158A, 1735–1741. <https://doi.org/10.1002/ajmg.a.35402>
- Vezain, M., Gérard, B., Drunat, S., Funalot, B., Fehrenbach, S., N’Guyen-Viet, V., Vallat, J.-M., Frébourg, T., Tosi, M., Martins, A., Saugier-veber, P., 2011. A leaky splicing mutation affecting SMN1 exon 7 inclusion explains an unexpected mild case of spinal muscular atrophy. *Hum. Mutat.* 32, 989–994. <https://doi.org/10.1002/humu.21528>
- Vezain, M., Saugier-veber, P., Goïna, E., Touraine, R., Manel, V., Toutain, A., Fehrenbach, S., Frébourg, T., Pagani, F., Tosi, M., Martins, A., 2010. A rare SMN2 variant in a previously unrecognized composite splicing regulatory element induces exon 7 inclusion and reduces

- the clinical severity of spinal muscular atrophy. *Hum. Mutat.* 31, E1110-1125. <https://doi.org/10.1002/humu.21173>
- Vogelstein, B., Kinzler, K.W., 1999. Digital PCR. *Proc. Natl. Acad. Sci. U.S.A.* 96, 9236–9241. <https://doi.org/10.1073/pnas.96.16.9236>
- Wang, C.H., Finkel, R.S., Bertini, E.S., Schroth, M., Simonds, A., Wong, B., Aloysius, A., Morrison, L., Main, M., Crawford, T.O., Trela, A., Participants of the International Conference on SMA Standard of Care, 2007. Consensus Statement for Standard of Care in Spinal Muscular Atrophy. *J Child Neurol* 22, 1027–1049. <https://doi.org/10.1177/0883073807305788>
- Werdnig, G., 1894. Die frühinfantile progressivespinale Amyotrophie. *Arch. Psychiat. Nervenkr.* 26, 706–744.
- Werdnig, G., 1891. Zwei frühinfantile hereditäre Fälle von progressiver Muskelatrophie unter dem Bilde der Dystrophie, aber auf neurotischer Grundlage. *Arch. Psychiat. Nervenkr.* 22, 437–480.
- Wirth, B., 2000. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Hum. Mutat.* 15, 228–237. [https://doi.org/10.1002/\(SICI\)1098-1004\(200003\)15:3<228::AID-HUMU3>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1098-1004(200003)15:3<228::AID-HUMU3>3.0.CO;2-9)
- Wirth, B., Herz, M., Wetter, A., Moskau, S., Hahnen, E., Rudnik-Schöneborn, S., Wienker, T., Zerres, K., 1999. Quantitative analysis of survival motor neuron copies: identification of subtle SMN1 mutations in patients with spinal muscular atrophy, genotype-phenotype correlation, and implications for genetic counseling. *Am. J. Hum. Genet.* 64, 1340–1356. <https://doi.org/10.1086/302369>
- Wirth, B., Schmidt, T., Hahnen, E., Rudnik-Schöneborn, S., Krawczak, M., Müller-Myhsok, B., Schönling, J., Zerres, K., 1997. De novo rearrangements found in 2% of index patients with spinal muscular atrophy: mutational mechanisms, parental origin, mutation rate, and implications for genetic counseling. *Am. J. Hum. Genet.* 61, 1102–1111. <https://doi.org/10.1086/301608>

## SERMENT D'HIPPOCRATE

Au moment d'être admis à exercer la médecine, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité.

Mon premier souci sera de rétablir, de préserver ou de promouvoir la santé dans tous ses éléments, physiques et mentaux, individuels et sociaux. Je respecterai toutes les personnes, leur autonomie et leur volonté, sans aucune discrimination selon leur état ou leurs convictions. J'interviendrai pour les protéger si elles sont affaiblies, vulnérables ou menacées dans leur intégrité ou leur dignité. Même sous la contrainte, je ne ferai pas usage de mes connaissances contre les lois de l'humanité. J'informerai les patients des décisions envisagées, de leurs raisons et de leurs conséquences. Je ne tromperai jamais leur confiance et n'exploiterai pas le pouvoir hérité des circonstances pour forcer les consciences. Je donnerai mes soins à l'indigent et à quiconque me les demandera. Je ne me laisserai pas influencer par la soif du gain ou la recherche de la gloire.

Admis dans l'intimité des personnes, je tairai les secrets qui me seront confiés. Reçu à l'intérieur des maisons, je respecterai les secrets des foyers et ma conduite ne servira pas à corrompre les mœurs. Je ferai tout pour soulager les souffrances. Je ne prolongerai pas abusivement les agonies. Je ne provoquerai jamais la mort délibérément.

Je préserverai l'indépendance nécessaire à l'accomplissement de ma mission. Je n'entreprendrai rien qui dépasse mes compétences. Je les entretiendrai et les perfectionnerai pour assurer au mieux les services qui me seront demandés.

J'apporterai mon aide à mes confrères ainsi qu'à leurs familles dans l'adversité. Que les hommes et mes confrères m'accordent leur estime si je suis fidèle à mes promesses ; que je sois déshonoré et méprisé si j'y manque.

## Résumé

**Introduction :** L'amyotrophie spinale proximale (SMA) est une maladie neuromusculaire d'origine génétique de sévérité variable, résultant d'une inactivation homozygote du gène *SMN1*. Le dépistage des hétérozygotes repose sur la détermination du nombre de copies du gène *SMN1* par des techniques quantitatives telles que la QMPSF ou la MLPA, mais ces techniques sont confrontées à l'existence de duplications en *cis* du gène *SMN1* dans la population générale. Elles ne permettent pas de connaître la phase des gènes *SMN1*. Des variants associés à des duplications en *cis* de ce gène ont été mis en évidence dans la population juive ashkénaze et espagnole suggérant une possible utilité pour le conseil génétique.

**Objectifs :** Nous avons répliqué les études d'associations entre les variants et les duplications en *cis* du gène *SMN1* à un échantillon de population de sujets analysés au CHU de Rouen, afin de déterminer les paramètres de sensibilité et de spécificité de cette association dans la population de notre étude. Nous avons cherché à savoir si ces variants pouvaient également être associés aux duplications en *cis* du gène *SMN2*.

**Matériel et méthodes :** 199 sujets ont été classés en sous-groupes selon leur génotype : 1+0, 1+1, 2+0, 2+1, 3+0, 2+2. Nous avons établi leur génotype aux locus des deux variants d'intérêt, par SNaPshot par un kit Amplidex PCR/CE *SMN1/2*. Nous avons considéré 2 groupes : le groupe "dup en *cis*" (génotypes 2+0, 2+1, 3+0, 2+2) et "absence de dup en *cis*" (génotypes 1+0 et 1+1). Nous avons évalué la différence de proportions de variants entre les groupes "dup en *cis*" et "absence de dup en *cis*" par un test exact de Fisher et calculé les valeurs de Se, Sp, VPP, VPN du test de détection des variants.

**Résultats :** Le groupe "dup en *cis*" est significativement enrichi en variants recherchés par rapport au groupe "absence de dup en *cis*" ( $P = 1,6 \cdot 10^{-7}$ ). La Se est de 17 %, la Sp de 100 %, la VPP de 100 % et la VPN de 92 % à 94 %. Les variants recherchés ne sont pas associés aux duplications en *cis* du gène *SMN2*.

**Discussion et conclusion :** La faible valeur de Se ne permet pas d'éliminer de façon certaine une duplication en *cis* du gène *SMN1* chez un apparenté porteur de 2 copies de ce gène, analysé dans le cadre d'un conseil génétique. Il est donc nécessaire de continuer à analyser le conjoint en cas de projet parental d'un apparenté atteint de SMA. Nous pouvons toutefois intégrer la recherche des variants à l'analyse du nombre de copies du gène *SMN1* afin d'affiner le calcul du risque résiduel. Nous pouvons envisager des techniques de séquençage ou de cartographie sur des fragments d'ADNg longs (peignage moléculaire, PCR digitale, technologies Bionano et Nanopore) pour objectiver les duplications en *cis* du gène *SMN1* en anticipant toutefois leur possible manque de résolution pour distinguer les gènes *SMN1* et *SMN2*.

## **Abstract**

**Introduction :** *Spinal muscular atrophy (SMA) is a genetic neuromuscular disease of varying severity, resulting from homozygous inactivation of the SMN1 gene. The detection of heterozygotes is based on the determination of the copy number of the SMN1 gene by quantitative techniques, such as QMPSF or MLPA. However, these techniques are faced with the existence of duplications in cis of the SMN1 gene in the general population. Indeed quantitative techniques do not allow us to know the phase of the SMN1 genes. Variants associated with duplications in cis of this gene have been identified in the Ashkenazi Jewish as well as in the Spanish population, suggesting a possible utility for genetic counseling.*

**Objectives :** *We have replicated previous associations studies between two variants and duplications in cis of the SMN1 gene to a sample of the population of subjects analyzed at the Rouen University Hospital, in order to determine the parameters of sensitivity and specificity of this association in the population of our study. We investigated whether these variants could also be associated with duplications in cis of the SMN2 gene.*

**Materials and methods :** *199 subjects were classified into subgroups according to their genotype: 1+0, 1+1, 2+0, 2+1, 3+0, 2+2. We established their genotype at the loci of the two variants of interest, using a SNaPShot technique and an Amplidex PCR/CE SMN1/2 kit. We considered 2 groups : the group "dup in cis" (genotypes 2+0, 2+1, 3+0, 2+2) and "absence of dup in cis" (genotypes 1+0 and 1+1). We evaluated the difference in the proportions of variants between the "dup in cis" and "absence of dup in cis" groups by an exact Fisher test and calculated the values of Se, Sp, VPP, VPN of the variant detection test.*

**Results :** *The group "dup in cis" is significantly enriched in sought-after variants compared to the group "absence of dup in cis" ( $P = 1.6 \cdot 10^{-7}$ ). Se is 17 %, Sp is 100 %, VPP 100% and NPV 92% to 94%. The variants sought are not associated with cis duplications of the SMN2 gene.*

**Discussion and conclusion :** *The low value of Se does not make it possible to definitively eliminate a duplication in cis of the SMN1 gene in a relative carrying 2 copies of this gene, analyzed in the care of genetic counseling. It is therefore necessary to continue to analyze the spouse of a relative with SMA in the event of a parental plan. We could however integrate the search for variants into the analysis of the number of copies of the SMN1 gene in order to refine the calculation of the residual risk. We can consider sequencing or mapping techniques on long gDNA fragments (molecular combing, digital PCR, Bionano and Nanopore technologies) to assess the duplications in cis of the SMN1 gene while anticipating their possible lack of resolution to distinguish genes SMN1 and SMN2.*