



HAL
open science

La thérapie génique pour soigner les maladies génétiques

Eva Dor

► **To cite this version:**

Eva Dor. La thérapie génique pour soigner les maladies génétiques. Sciences de l'Homme et Société. 2023. dumas-04530380

HAL Id: dumas-04530380

<https://dumas.ccsd.cnrs.fr/dumas-04530380>

Submitted on 3 Apr 2024

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

ESIT - Université Sorbonne Nouvelle - Paris 3

La thérapie génique pour soigner les maladies génétiques

Mémoire de traduction | Sous la direction de Mme Susana Mauduit-Peix

Parcours Traduction éditoriale, économique et technique | Anglais - Français

Eva Dor

2022-2023



Remerciements

*Merci à Léonore, pour son apport graphique, à Anne, mon Bescherelle humain, et à Laurent,
sans qui ce mémoire n'aurait pas eu lieu.*

Merci à ma directrice de mémoire, Mme Mauduit-Peix.

Table des matières

Partie 1 : Exposé.....	5
Introduction	5
I- Les bases de la génétique	6
1. Le génome	7
2. Le code génétique, de l'ADN à la protéine	10
3. Le séquençage des génomes	16
II- Les maladies génétiques	20
1. La mutation génétique	20
2. Les types de maladie génétique	24
3. Le cas de la myopathie de Duchenne	31
III- Les méthodes de la thérapie génique	36
1. Les différentes approches	36
2. Vecteurs et ciblage	40
3. Le cas de CRISPR-Cas9	43
IV- Recherche, application et éthique	48
1. Comprendre et interpréter des résultats scientifiques.....	49
2. Recherche et applications concrètes	55
3. Questions éthiques.....	57
Conclusion.....	58
Partie 2 : TEXTE-SUPPORT ET SA TRADUCTION.....	59
Partie 3 : STRATÉGIE DE TRADUCTION.....	87
I- Présentation du texte-support.....	87
II- Méthode de travail et démarches de recherche.....	88

1. Déroulé	89
2. Corpus comparables	93
3. Corpus alignés	103
III- Repérage et analyse des principaux problèmes et difficultés de traduction	105
1. Justification des choix de traduction	105
2. Difficultés de traduction	108
Partie 4 : ANALYSE TERMINOLOGIQUE	123
Fiches terminologiques.....	123
Glossaire FR-EN.....	135
Lexique FR-EN.....	143
Lexique EN-FR.....	148
BIBLIOGRAPHIE CRITIQUE.....	153
Bibliographie anglaise	153
Ressources pédagogiques.....	153
Ressources académiques	154
Bibliographie française.....	155
Ressources pédagogiques.....	155
Ressources académiques	159
ANNEXE	162
INDEX	167

Partie 1 : Exposé

Introduction

Ces dernières années, les avancées technologiques en thérapie génique n'ont cessé de se développer, si bien que les méthodes de modification ou de réparation du génome sont nombreuses, qu'elles en soient encore au stade d'essai clinique et pré-clinique ou bien déjà appliquées, entre autres, en médecine. Les possibilités d'édition du génome sont infinies et posent donc potentiellement des problèmes éthiques dont la résolution semble s'attarder loin derrière le rythme effréné des trouvailles technologiques. Plus récemment, le développement de l'outil de CRISPR-Cas9, des ciseaux moléculaires mis au point en 2012 par deux chercheuses, a ouvert la voie vers de nouvelles possibilités de réparation des cellules pour soigner les maladies génétiques, dont certains cancers.

La thérapie génique permet de corriger des mutations responsables de certains désordres en introduisant dans les cellules un matériel génétique sain. Les premières thérapies consistaient à suppléer un gène défectueux dans le cas de maladies monogénétiques (mutation sur un seul gène). Mais avec la multiplication récente des stratégies thérapeutiques, il est désormais possible, comme grâce à CRISPR-Cas9, de venir directement réparer ou éliminer de l'ADN au sein de l'organisme touché.

Ce domaine en constante évolution, et de même relativement récent dans l'histoire des sciences médicales, suppose de nombreux questionnements linguistiques liés à sa traduction. La recherche essentiellement anglophone, bien que menée internationalement, construit progressivement autour de ce sujet un univers lexical et jargonnant qui s'agrandit sans cesse sans pour autant s'accompagner d'équivalences instaurées dans d'autres langues. Même si une immense partie de la phraséologie et de la terminologie préexiste à ce domaine (mêlant génétique, histologie, stratégie thérapeutique, médecine, imagerie, anatomie, etc.), il reste tout de même une part langagière difficilement accessible pour le traducteur non spécialisé.

Quand bien même une partie de la communauté scientifique française travaillerait quotidiennement ce sujet en français, ce qui est le cas, les publications restent toutes très majoritairement rédigées en anglais. Par conséquent, à moins d'avoir côtoyé plusieurs années des spécialistes de la thérapie génique, il faut se contenter la plupart du temps, pour traduire, de ressources anglophones.

- ❖ *En gras les termes faisant l'objet d'une fiche terminologique*
- ❖ *Apposés à un astérisque les termes présents dans le glossaire ou le lexique*
- ❖ *En notes de bas de page les références citées ou consultées pour développer une notion*

I- Les bases de la génétique

Afin de comprendre l'origine et les caractéristiques des maladies génétiques d'une part, et le fonctionnement des différentes méthodes de thérapie génique d'autre part, il convient de reprendre les bases de la génétique. La génétique*, sous-discipline de la biologie, joue un rôle central dans la construction et le fonctionnement de tous les individus.

Du grec *genno*, « donner naissance », c'est la science de l'hérédité, de la variation héréditaire et des gènes¹. L'hérédité, elle, correspond à « l'ensemble des caractères (aspects ou propriétés biologiques) que les êtres vivants se transmettent de génération en génération »². Donc, par là même, la génétique joue un rôle central dans l'aspect physique d'un individu, son intelligence, mais aussi sa propension à développer des maladies. Nous retracerons ici les différentes étapes de la traduction du code génétique, de l'échelle moléculaire à l'échelle somatique.

1. Le génome

Le génome* correspond à l'information génétique de tout organisme vivant. Il se constitue de grandes molécules d'ADN* (acide désoxyribonucléique) prenant la forme de chromosomes* dans le noyau de nos cellules (voir Fig.1). Chez l'être humain, chaque cellule compte 46 chromosomes, soit 23 paires de chromosomes, et pour chaque paire, un chromosome est hérité de la mère, l'autre du père. Seules les cellules reproductives, appelées gamètes*, ne possèdent que 23 chromosomes (une version de chaque paire).

Les chromosomes

Comme nous le disions, les chromosomes sont constitués d'une longue molécule de deux brins d'ADN enroulés en forme d'hélice. Chaque brin d'ADN est quant à lui composé d'un enchaînement de molécules organiques, appelées nucléotides*, et dont il existe quatre types, selon leur base azotée* : les nucléotides à adénine (A), les nucléotides à thymine (T), les nucléotides à cytosine (C) et les nucléotides à guanine (G). Les deux brins d'ADN s'appairellent entre eux selon une règle de complémentarité entre les nucléotides (A avec T et C avec G), et forment ainsi des paires de bases*.

¹ Bechkri, S. Cours de génétique, L2 Sciences biologiques. Constantine : Université Frères Mentouri, Faculté des sciences de la nature et de la vie, 2020. Tiré de Beaudry, JR. *Génétique générale*. Montréal : Décarie, 1985, 504 p.

² Ibid

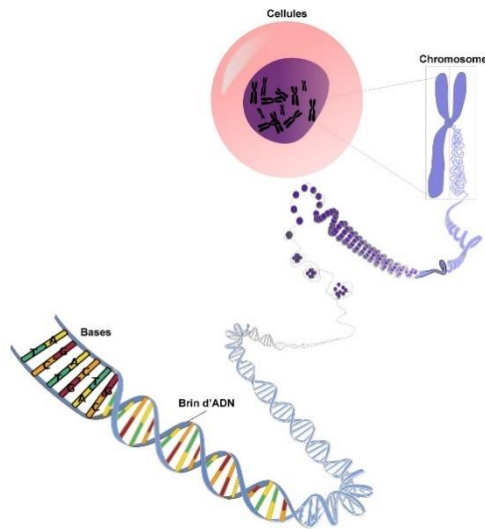


Figure 1 : De la cellule aux paires de bases. Tiré et adapté du National Human Genome Research Institute

L'enchaînement des paires suit un ordre et une longueur de séquence spécifique et unique pour chaque gène, ce qui permet d'assigner son code à des fonctions distinctes dans le corps.

Pourquoi dit-on qu'un brin d'ADN est orienté 5' vers 3'

Un nucléotide est composé de trois molécules : sa base azotée, que nous avons déjà mentionnée, un phosphate et un désoxyribose*. Ces molécules sont elles-mêmes composées d'éléments chimiques. Par exemple, la molécule de phosphate est composée de 5 carbones. Pour assembler les nucléotides entre eux, et ainsi créer la structure du brin d'ADN, des liaisons relient les molécules d'un nucléotide aux molécules du nucléotide d'à côté. Plus précisément, ces liaisons s'établissent entre le 3^e carbone (en position 3') de la molécule de phosphate du premier nucléotide, et le 5^e carbone (en position 5') du phosphate du nucléotide suivant dans la séquence. Par conséquent, dans la chaîne nucléotidique, une extrémité commence par la partie du phosphate liée au carbone 5' du premier nucléotide, et l'autre extrémité se termine par la partie du phosphate liée au carbone 3' du dernier nucléotide (voir fig.2).

Ainsi, par convention, on lit toujours une séquence d'ADN dans le sens 5' vers 3'.

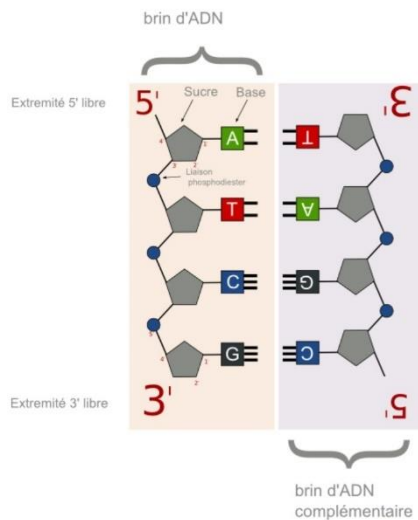


Figure 2 : L'assemblage des nucléotides entre eux, à l'origine de la convention de lecture de l'ADN. Tiré et adapté de S. Schutz

Les cellules

Toute l'information génétique est contenue dans les cellules*. Un organisme vivant, et disons ici le corps humain, est entièrement fait de cellules. Leur multiplication permet au fœtus de se développer et aux membres de se régénérer tout au long de la vie. Les cellules humaines sont des cellules eucaryotes*. Elles sont constituées d'un noyau contenant le matériel génétique et entouré par le cytoplasme*. Elles sont séparées de leur milieu par une membrane. Dans le cytoplasme se trouvent des mitochondries, organites responsables de fournir de l'énergie à la cellule. Par opposition, les cellules procaryotes, comme les virus*, ne possèdent pas de noyau et leur matériel génétique flotte librement dans le cytoplasme. Grâce au processus de division cellulaire, appelé mitose*, les cellules se démultiplient pour donner deux cellules identiques, et transmettent ainsi leurs attributs à chaque cellule fille*. Lors de la première étape de la mitose, les chromosomes présents dans le noyau des cellules sont dupliqués, séparés en deux au centre, puis réassemblés. Un exemplaire de chaque se retrouve ainsi dans deux cellules filles issues d'une même cellule ³.

³ La division cellulaire : la mitose [VIDEO]. Médiathèque du CEA, 2021. Commissariat à l'Énergie Atomique (CEA) [en ligne]. Disponible sur cea.fr/multimedia (Consulté en avril 2023)

Les cellules post-mitotiques* sont ainsi des cellules issues de la division cellulaire et qui ne se diviseront plus.

« Toutes les cellules de l'organisme sont issues de deux cellules, l'ovule de la mère et le spermatozoïde du père. Après l'union de l'ovule et du spermatozoïde (fécondation), l'ovule fécondé constitue une seule cellule. Cette cellule, le zygote, se divise de nombreuses fois et, au fur et à mesure qu'elle se divise, les cellules filles développent différentes caractéristiques et fonctions. Ces différentes cellules finissent par former les différents organes. »⁴

Il existe ainsi de nombreux types de cellules, dont la fonction et la structure varient (cellules sanguines, cellules musculaires, etc.).

Comprendre la structure d'une cellule et le processus de mitose permet de distinguer l'ADN nucléaire* de l'ADN mitochondrial*. En effet, l'ADN porteur du patrimoine génétique correspond à l'ADN qui se trouve dans le noyau des cellules, et qui est transmis à l'identique de cellule mère en cellule fille, soit l'ADN nucléaire. L'ADN mitochondrial, quant à lui, est un ADN circulaire présent dans les mitochondries (organites situés dans le cytoplasme de la cellule), et qui ne code que pour des fonctions propres aux mitochondries elles-mêmes. L'ADN dont il sera question par la suite est l'ADN nucléaire, hérité du père et de la mère, et déterminant pour toutes les fonctions du corps.

2. Le code génétique, de l'ADN à la protéine

La séquence codante

Un gène* correspond à un segment de l'ADN porteur d'un message à l'origine de l'expression de certains caractères (cheveux blonds ou bruns, groupe sanguin, etc.) mais aussi de toutes les fonctions biologiques (comme la fonction musculaire).

⁴ Villa-Forte, A. Cellules. *Le Manuel MSD*, 2022.

On peut dire qu'il est comme un manuel d'instruction pour fabriquer des protéines, et que ces dernières permettent le bon fonctionnement de notre corps. Sans la séquence d'ADN correcte, il n'y a pas de protéine (ou une protéine tronquée), et le corps dysfonctionne. Chez l'Homme, un chromosome est porteur d'environ 20 000 gènes, chacun situé à un endroit bien précis, qu'on appelle locus*. Dans chaque gène, on retrouve des parties codantes, les exons*, et non-codantes, les introns*, qui s'alternent.

La séquence codante d'un gène correspond donc à la partie des exons de l'ADN ou de l'ARN du gène, exons qui seront ensuite traduits en protéine.

La transcription de l'ADN

Pour qu'une séquence codante soit traduite en protéine fonctionnelle, il faut tout d'abord qu'ait lieu une transcription de l'ADN. Par définition, la transcription, qui a lieu dans le noyau des cellules eucaryotes, correspond à la synthèse d'une molécule d'ARN* (acide ribonucléique) à partir d'une matrice d'ADN⁵. De même qu'un brin d'ADN, la molécule d'ARN issue de la transcription se compose d'un enchaînement de nucléotides et elle porte la même information génétique que la région d'ADN transcrite. Comme nous le verrons, de nombreux types d'ARN sont utilisés dans la synthèse des protéines.

Lors de la transcription, le brin d'ADN utilisé, le « brin matrice » ou « non-sens », sert à produire un ARN dont l'enchaînement des nucléotides lui est complémentaire. Par cette règle de complémentarité, on obtient donc un brin d'ARN dont la séquence de bases est identique au brin complémentaire non transcrit, le « brin sens », sauf que les nucléotides à T sont remplacés par des nucléotides à uracile (U). La transcription est effectuée par une ARN polymérase*, assisté par plusieurs autres protéines, et l'ARN obtenu est synthétisé à partir de désoxyribonucléotides* (dNTP) présents dans le milieu, qui sont ajoutés un à un sur la chaîne en formation lors du processus de synthèse de l'ARN. Les nucléotides sont toujours ajoutés à l'extrémité 3' du brin d'ARN, donc le sens de transcription est de 5' vers 3'.

⁵ Bechkri, S. Cours de génétique, L2 Sciences biologiques. Constantine : Université Frères Mentouri, Faculté des sciences de la nature et de la vie, 2020. Tiré de Beaudry, JR. *Génétique générale*. Montréal : Décarie, 1985, 504 p.

Chaque séquence transcrite se compose de trois parties. D'abord, le promoteur* correspond à une séquence spécifique de l'ADN qui est reconnue par le mécanisme de transcription et qui sert de point de départ à cette transcription. Ensuite, la région codant l'ARN sert de matrice à la synthèse de l'ARN. Enfin, le terminateur est une autre séquence signalant l'arrêt de la transcription (voir fig.3). À partir de la transcription, on obtient un ARN pré-messager* (ARNpm).

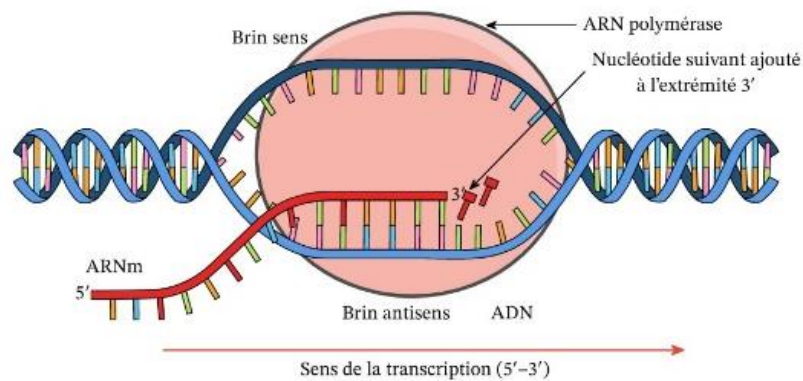


Figure 3 : schéma d'une transcription. Tiré et adapté du site Nagwa.com

Le rôle des ARN

À la suite de la transcription, l'ARNpm subit une maturation, toujours dans le noyau cellulaire, lors de laquelle seuls les exons (séquences codantes) sont conservés, et les introns sont retirés. Les exons sont liés bout à bout lors de la phase d'épissage*⁶ (voir fig.4). On obtient alors un ARN messager* (ARNm) mature.

L'ARNm est donc constitué d'une succession d'exons, dont seule la lecture selon un cadre précis aboutit à la fabrication d'une protéine⁷. On parle du cadre de lecture* de la séquence. Cette lecture des nucléotides se fait 3 par 3, c'est-à-dire de codon en codon*.

⁶ De l'ADN à l'ARNm. Travail pratique. Plateforme pédagogique e-enseignement de l'académie de Nice [en ligne]. Disponible sur e-enseignement.ac-nice.fr (consulté en avril 2023)

⁷=Cadre de lecture. Glossaire, 2016. AFM-Téléthon [en ligne]. Disponible sur afm-telethon.fr (consulté en

Le début et la fin de l'ARNm sont délimités par un codon d'initiation et un codon stop* respectivement, qui agissent comme signal de début et de fin de la retranscription du message.

Cependant, il existe plusieurs types d'ARN, exerçant chacun différentes fonctions. Par exemple, nous avons déjà mentionné l'ARN polymérase, l'enzyme* qui identifie les gènes sur l'ADN puis synthétise le brin d'ARN à partir de cette séquence. Il peut fabriquer un ARNm, mais aussi un ARN ribosomique (ARNr) ou un ARN de transfert (ARNt). Comme nous le verrons, ces deux derniers interviennent lors de la phase de synthèse des protéines. Chez les cellules eucaryotes, le type d'ARN polymérase détermine le type d'ARN synthétisé.

Quant à lui, à la fin de la retranscription, l'ARNm est exporté à l'extérieur du noyau et rejoint la fabrique de protéine, pour lui apporter l'information génétique du gène.

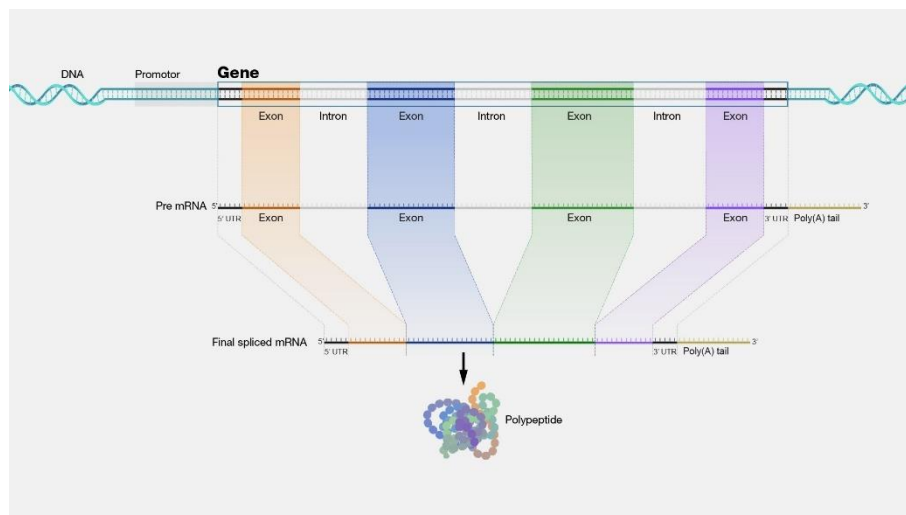


Figure 4 : Du gène à la protéine, en passant par la phase de transcription et la phase d'épissage.
Tirée et adaptée du National Institute of Health

La synthèse des protéines

Les protéines sont le matériau principal de construction des organismes vivants et permettent le bon fonctionnement des cellules⁸.

mai 2023)

⁸ Berrada, S. Les protéines : structure, propriétés et applications technologiques. Biochimie appliquée dans les filières SBSSA. Académie de Montpellier. Académie de Dijon. 2009, 13 p.

Il existe un très grand nombre de protéines différentes, ainsi, elles font bouger les muscles, contrôlent le métabolisme, servent d'anticorps, d'enzyme, etc. Chez l'être humain, on compte des dizaines de milliers de protéines. Leur structure principale est faite d'une séquence d'acides aminés, la chaîne polypeptidique, qui se replie sur elle-même pour donner lieu à des structures secondaires plus complexes. Il existe dans la nature 20 types d'acides aminés différents, et l'on considère qu'un enchaînement de plus de 50 acides aminés constitue une molécule de protéine. Cependant, la plupart des protéines en possèdent 100 à 1000. La propriété d'une protéine dépend de l'ordre séquentiel des acides aminés. Les 20 acides aminés sont comme des mots qui, selon leur agencement, formeraient une phrase déchiffrable.

Une fois dans le cytoplasme de la cellule, l'ARN messager se dirige vers les ribosomes*, organites responsables de la synthèse des protéines. Chaque ribosome est constitué d'une grande et d'une petite sous-unité, indépendantes mais qui s'assemblent lors du processus de synthèse des protéines. Un ribosome est composé de protéines et d'ARNr (ribosomiques).

Lors de la première phase de synthèse de la protéine, les codons de l'ARNm vont être traduits en acide aminés. Le code qui fait correspondre un codon à un acide aminé est connu des scientifiques. Par exemple, le codon CAG code pour un acide aminé appelé glutamine. On sait aussi que les codons UAA, UAG et UGA sont des codons stop.

Comme il y a 3 nucléotides par codon et 4 types de nucléotide différents (A, U, C, G), il y a 4^3 , soit 64 codons différents possibles. Parmi eux, trois sont des codons stop, qui signalent la fin de la traduction. Les 61 codons restants, les codons sens, codent les 20 acides aminés qu'on retrouve dans les protéines.

L'acide aminé correspondant au codon de l'ARNm est capté dans le milieu environnant, à l'aide de l'ARNt (de transfert) : lorsque le ribosome se déplace le long de la molécule d'ARNm, chaque codon d'ARNm est associé à une molécule d'ARNt portant ses trois nucléotides complémentaires⁹ (voir fig.5). Ces trois nucléotides sont donc les anticodons, toujours suivant la règle de complémentarité des bases nucléotidiques déjà mentionnée. Au fur et à mesure de la traduction, les acides aminés sont mis bout à bout.

⁹ Type d'ARN. Fiche explicative, [s.d.]. Nagwa [en ligne]. Disponible sur [nagwa.com/fr](https://www.nagwa.com/fr) (Consulté en mai 2023)

On obtient alors une chaîne polypeptidique (un peptide étant un enchaînement d'acides aminés) (voir fig.4), dont l'ordre suit l'enchaînement des codons de l'ARNm.

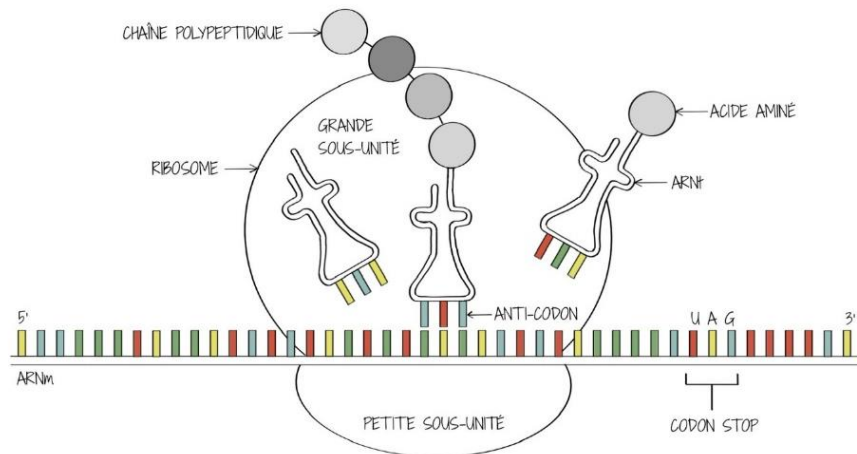


Figure 5 : processus de traduction d'un ARNm.
Illustration de Léonore Dor

Le repliement des protéines

Une fois la chaîne polypeptidique formée, elle doit se replier pour former la protéine finale. En effet, ce n'est pas tant la composition exacte d'une protéine (la séquence de ses acides aminés) qui importe, mais plutôt sa forme. Elle doit donc se replier pour adopter une forme bien précise, qui détermine sa fonction. Par exemple, la structure en Y des anticorps leur permettent de s'accrocher aux agents pathogènes pour déclencher une réaction immunitaire. De plus, un certain nombre de maladies sont dues au fait que certaines protéines n'adoptent pas la bonne configuration¹⁰.

¹⁰ Louapre, D. Le repliement des protéines : résolu par l'intelligence artificielle AlphaFold ? [VIDEO]. ScienceÉtonnante, 2021

Pour résumer simplement, le code génétique de chaque organisme est contenu dans ses cellules, sous formes d'ADN. Ce code correspond à des consignes, qui sont mises à l'abri dans le noyau. Pour être exécutées, les consignes sont rapportées à l'ARN messenger durant la **phase de transcription**. L'ARN messenger les transporte dans le cytoplasme de la cellule, jusqu'aux ribosomes. Ces derniers, assistés de l'ARN de transfert, lisent l'information de l'ARN messenger, durant la **phase de traduction**, et exécutent ainsi les consignes.

3. Le séquençage des génomes

Le séquençage* est ainsi défini par le Dictionnaire de l'Académie française :

« Séquençage, n. m. BIOCHIMIE. Opération permettant de déterminer la séquence d'une macromolécule. *Le séquençage d'une protéine, d'un polyside. Séquençage de l'A.D.N.* Par extension. *Entrepris en 1990, le séquençage du génome humain fut achevé en 2003.* »¹¹

Pourquoi séquencer

Le séquençage du génome fait référence à la détermination de la succession des bases nucléotidiques de l'ADN présent dans chaque cellule d'un organisme donné¹². Les premières techniques de séquençage datent de la fin des années soixante-dix et ont valu le prix Nobel de chimie aux chercheurs Walter Gilbert et Frederic Sanger en 1980. Aujourd'hui, et grâce aux nouvelles technologies et à l'intelligence artificielle, plusieurs dizaines de milliers de génomes sont déjà séquencés. Mais quelle en est l'utilité ?

Comme nous l'avons vu, c'est la séquence nucléotidique contenue dans l'ADN des cellules qui permet la fabrication des protéines et assigne leur rôle au sein du corps. Le code génétique est donc responsable de toutes les fonctions somatiques. Cela s'applique aussi bien à l'être humain qu'à un épi de maïs.

¹¹ Séquençage. Définition. Dictionnaire de l'Académie Française, 9^e édition (actuelle) [en ligne]. Disponible sur [dictionnaire-academie.fr](https://www.dictionnaire-academie.fr) (consulté en mai 2023)

¹² Furelaud, G., Esnault, Y. Le Séquençage des Génomes. École Normale Supérieure, Planet-Vie, 2014 /ca

Pouvoir connaître le code génétique du vivant, c'est avoir accès à ses instructions créatrices, ce qui a un impact considérable sur la recherche en biologie et ouvre de nouvelles voies d'exploration dans le domaine de la médecine (diagnostic, thérapeutique, prédiction, pronostic, prévention...) mais aussi d'autres disciplines comme l'anthropologie, l'agronomie, l'environnement, etc¹³.

La carte génétique

Bien souvent, un travail de cartographie est effectué en amont du séquençage génomique afin de faciliter cette tâche. Les techniques cartographiques utilisent des marqueurs* génétiques permettant de baliser les gènes dans le génome. Ces marqueurs correspondent à des séquences d'ADN repérables spécifiquement. La carte génétique permet de représenter l'ordonnement génomique ou chromosomique à une échelle plus large que le séquençage. On peut donc par exemple repérer les loci des gènes sur un chromosome.

Technique de séquençage de l'ADN

La difficulté à séquencer un génome est d'autant plus importante que celui-ci est grand et possède des séquences répétées. Pour exprimer la taille d'une séquence nucléotidique, la convention veut que l'on se serve du nombre de bases : les kilobases (kb) pour les milliers de bases, les mégabases (Mb) pour les millions, et les gigabases (Gb) pour les milliards. Ainsi, la taille du génome des organismes eucaryotes, par exemple celui des mammifères, qui mesure 2 à 3 Gb, requiert un travail cartographique préalable conséquent. Le premier génome d'un organisme cellulaire séquencé date de 1995. Pour sa part, le génome humain a vu sa première séquence séquencée en 2003, et il est aujourd'hui séquencé à 90 %. La difficulté à séquencer les 10 % restants tient aux techniques employées¹⁴.

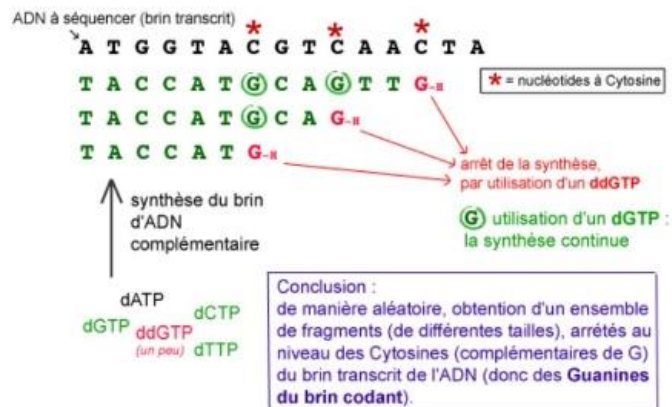
¹³ Lamoril, J., Ameziane, N., Deybach, JC. et al. Les techniques de séquençage de l'ADN : une révolution en marche, *Immuno-analyse & Biologie Spécialisée*, 2008, vol. 23, pp. 260-279

¹⁴ Le génome humain : de qui, pour qui, pourquoi ?. Communiqué, [s.d.]. Centre National de la Recherche Scientifique (CNRS) [en ligne]. Disponible sur [cnrs.fr](https://www.cnrs.fr) (consulté en avril 2023)

Une partie de l'ADN humain est constituée de fragments hautement répétés, elle est donc très difficile à séquencer et pratiquement dépourvue de gène.

Les techniques se basent sur les mécanismes de réplification de l'ADN. La méthode de Sanger utilise des ARN polymérases qui, comme nous l'avons vu (partie I.2 « Le code génétique, de l'ADN à la protéine »), peuvent synthétiser un brin complémentaire d'ADN, à partir d'un brin matrice en utilisant les désoxyribonucléotides (dNTP) présents dans le milieu. Pour le séquençage, les scientifiques ajoutent aussi à ce milieu des didésoxyribonucléotides (ddNTP) qui ont la particularité de stopper la synthèse lorsqu'ils sont ajoutés à la chaîne. Pour procéder, ils placent dans le milieu de réaction l'ADN à séquencer, l'ARN polymérase, des dNTP en grande quantité et des ddNTP en petite quantité (à Adénine, à Guanine, à Thymine ou à Cytosine). À certains moments aléatoire, l'ARN polymérase utilisera un ddNTP pour constituer la chaîne nucléotidique, et la synthèse s'arrêtera alors. L'opération est répétée dans un nouveau milieu réactionnel, avec à chaque fois des didésoxyribonucléotides différents (à A, G, T ou C)¹⁵ (voir fig.6). On obtient alors des fragments d'ADN complémentaires* de tailles variées, selon l'endroit où le ddNTP s'est inséré. Les fragments obtenus sont ensuite transférés sur un gel, et la séquence entière reconstituée.

Figure 6 : Utilisation du ddGTP dans le séquençage de l'ADN
Auteur : Gilles Furelaud



¹⁵ Delarue, M., Furelaud, G. Le Séquençage d'un ADN. École Normale Supérieure, Planet-Vie, 2014 /ca

Depuis 2006, les nanotechnologies employées permettent de séquencer simultanément des millions de fragments d'ADN sur une puce électronique de seulement quelques cm². Grâce à la bio-informatique, il est possible de reconstituer, à partir de ces millions de fragments, le génome d'un individu. Dès lors, nous sommes entrés dans l'ère du « séquençage à haut débit »¹⁶. Il est désormais possible de séquencer un génome humain en une journée, pour un coût inférieur à 1000 euros. Cependant, même s'il est possible de séquencer tout un génome, certaines parties du code génétiques ne sont pas encore connues. Il existe donc des gènes non identifiés, et quand bien-même ils le seraient, il n'est pas toujours possible de savoir quel gène est associé à quelle maladie et quels changements impliquent une pathologie.

Connaître notre génome fait partie intégrante de l'appréhension des maladies génétiques et des recherches biomédicales. Les principes génétiques explorés jusqu'à présent permettront dans un deuxième temps de comprendre l'origine et les facteurs du développement des maladies génétiques.

Nous pouvons d'ores et déjà noter que la terminologie liée à la génétique est dense et suppose plusieurs niveaux de connaissance, de plus en plus complexes. L'aperçu terminologique établi jusqu'ici donne une idée de la spécialisation linguistique de cette discipline. Grâce à cette mise en contexte de tous les termes, dont la plupart sont répertoriés dans le glossaire ou le lexique, nous possédons déjà les rudiments linguistiques pour une potentielle analyse textuelle ou traduction dans ce domaine, sans pour autant être des spécialistes.

¹⁶ Mathieu, M. (Tous chercheurs), Faurisson, F. (Inserm). Le séquençage du génome [VIDEO]. Inserm, 2018

II- Les maladies génétiques

Le Généthon est le tout premier laboratoire à s'être intéressé aux maladies génétiques. Il est issu d'une association fondée en 1958, qui identifie le gène responsable de la myopathie de Duchenne en 1986¹⁷. À partir de ce moment, la recherche française en génétique fait un bond en avant, et explore ainsi les gènes à l'origine de centaines de maladies génétiques, jusqu'alors inconnues ou mal identifiées. À l'échelle mondiale, environ 6 % des nouveau-nés sont atteints d'anomalies génétiques¹⁸. Ces anomalies sont dues à des mutations dans le génome. Cependant, les maladies génétiques peuvent aussi être causées par des mutations qui ne sont pas déjà présentes à la naissance.

1. La mutation génétique

Les mutations* correspondent à des changements dans la séquence de l'ADN. Elles sont donc à l'origine de l'existence des différents allèles* des gènes, c'est-à-dire des différentes variantes d'un même gène. Par exemple, pour le gène codant la couleur des yeux, on retrouve l'allèle yeux bleus, l'allèle yeux marrons, etc. En comparant le génome de deux individus pris au hasard, on dénombre 3,2 millions de nucléotides différents, ce qui ne représente qu'une variation de 0,1 %¹⁹. Cependant, même si les mutations permettent une grande variété génétique des génomes, elles sont aussi la source de certaines anomalies*, pouvant parfois entraîner des pathologies* plus ou moins graves chez l'être vivant, et de même héréditaires. La difficulté pour les généticiens réside dans la distinction entre une mutation responsable d'une maladie et une mutation agissant seulement comme facteur d'individualité.

¹⁷ Généthon, le laboratoire du Téléthon, a 30 ans [Ep1] [VIDEO]. AFM Téléthon, 2020

¹⁸ Statistique tirée de l'Organisation Mondiale de la Santé, 2023

¹⁹ Qu'est-ce qu'une maladie génétique ?. Institut Imagine [en ligne]. Disponible sur [institutimagine.org/fr](https://www.institutimagine.org/fr) (Consulté en janvier 2023)

Dans tous les cas, les mutations, en modifiant l'expression d'un gène, ont une incidence sur le phénotype*, soit « l'ensemble des caractères observables, apparents, d'un individu, d'un organisme dus aux facteurs héréditaires (génotype) et aux modifications apportées par le milieu environnant »²⁰. L'effet phénotypique d'une mutation se définit par comparaison avec le phénotype sauvage*, c'est-à-dire le phénotype sain ou standard.

Mutations spontanées ou induites

Une mutation génétique n'est pas nécessairement héritée. Elle peut apparaître naturellement après la fécondation, et avoir lieu de manière aléatoire. Dans ce cas, il s'agit d'une mutation spontanée*. Ces mutations affectent souvent même les cellules reproductrices, ce qui les rend transmissibles à la descendance.

Inversement, les mutations induites* peuvent être dues à certaines substances ou à des conditions dans l'environnement de l'organisme touché, substances ou conditions dites « mutagènes », car elles provoquent une modification des gènes. Généralement, ces mutations ne touchent pas toutes les cellules. C'est le cas notamment des cancers. Le cancer est toujours d'origine génétique, et les thérapies ciblent bel et bien la molécule d'ADN, mais la maladie n'est pas forcément héréditaire car elle peut être liée à des mutations induites. Toutefois, certains gènes sont promoteurs du cancer, c'est-à-dire qu'ils prédisposent l'individu à être plus facilement sujet à certaines mutations induites²¹.

L'épigénétique entre ici en jeu, puisque ce domaine de recherche concerne l'impact de l'environnement sur l'ADN. En effet, l'environnement peut agir en modifiant l'ADN, pas dans sa séquence primaire mais dans son expression²². Durant le développement de l'organisme, l'ADN des gènes accumule des marqueurs chimiques qui déterminent l'intensité de l'expression des gènes²³.

²⁰ Phénotype. Définition. Centre Nationale de Ressources Textuelles et Lexicales (CNRTL) [en ligne]. Disponible sur [cnrtl.fr](https://www.cnrtl.fr) (Consulté en janvier 2023)

²¹ Tiret, L. Adapté de propos recueillis en juin 2023

²² Tiret, L. Ibid

²³ Qu'est-ce que l'épigénétique ? Et en quoi est-elle liée au développement de l'enfant ?. Communiqué de 2023.

Ainsi, les changements induits par l'environnement correspondent à des modifications chimiques des paires de bases à des endroits clef, ce qui a ainsi un impact sur la manière dont le gène s'exprime.

Décalage du cadre de lecture

Nous avons vu précédemment que l'ARN messager était lu par les ribosomes, codon par codon, c'est-à-dire trois nucléotides par trois nucléotides. Ce cadre de lecture précis permet de traduire les codons en acides aminés, et d'ainsi former des protéines. Cependant, il arrive qu'une anomalie génétique provoque le décalage d'un ou deux nucléotides du cadre de lecture, rendant ainsi le message de l'ARN illisible²⁴ (voir fig.7).

Le décalage du cadre de lecture sur l'ARN messager aboutit à l'apparition d'un codon stop prématuré et donc à l'éventuelle présence d'une protéine tronquée. Toutes les mutations ne provoquent pas un décalage. Il en existe plusieurs sortes, chacune ayant des incidences variées sur le processus de transcription ou de traduction de l'ADN.



Figure 7 : La délétion de 2 nucléotides au niveau du 2ème triplet modifie complètement la phrase qui perd tout son sens. Illustration tirée et adaptée de l'AFM-Téléthon.

Center of the Developing Child. Harvard University

²⁴Cadre de lecture. Glossaire, 2016. AFM-Téléthon [en ligne]. Disponible sur afm-telethon.fr (consulté en mai 2023)

Substitutions, délétions, insertions et amplifications

Si la mutation consiste au remplacement d'un seul nucléotide par un autre, on parle de « substitution* ». Cette dernière représente près de 70 % des mutations. Il peut s'agir du remplacement d'une seule lettre (A, T, C ou G) dans la séquence de l'ADN, ou bien de plusieurs lettres. Ce changement modifie le cadre de lecture mais ne le décale pas. En revanche, une mutation peut aussi correspondre à l'ajout d'un nucléotide, ce qu'on appelle « insertion* ». Dans ce cas, le cadre de lecture est décalé. Par exemple, si on rajoute un nucléotide à A en début de séquence codante mais que celui-ci n'est normalement pas présent, tous les autres nucléotides seront décalés d'une position. Le décalage du cadre de lecture* peut aussi être provoqué par une mutation par « délétion* », lorsqu'un ou plusieurs nucléotides sont manquants. Les mutations par insertion ou délétion sont donc des mutations décalantes. En général, elles changent tous les acides aminés sur la chaîne polypeptidique formée en aval, de sorte que leurs effets sur le phénotype sont très prononcés. Les délétions et les insertions représentent près de 25 % des anomalies répertoriées.

Si elles se font par trois ou par un multiple de trois, le cadre de lecture demeure inchangé. Ce sont des insertions et délétions « en phase »²⁵. Toutefois, des mutations répétant plusieurs fois un même motif de trois nucléotides peuvent aussi aboutir à des conséquences fonctionnelles variables. En effet, dans le cas où les insertions ou délétions seraient tri-nucléotidiques, le cadre de lecture ne sera certes pas décalé, mais la protéine codée présentera une délétion ou une insertion d'un ou plusieurs acides aminés, puisqu'un codon code pour un acide aminé.

De plus, les anomalies caractérisées par la répétition d'un triplet sont souvent impliquées dans les mutations dynamiques, à l'origine de la plupart des maladies génétiques. Il est possible de distinguer les répétitions qui affectent la région codante (exon) des gènes et celles qui affectent les régions non codantes (introns), où la taille de la séquence répétée (amplifiée) peut atteindre 10 kb. Il existe aussi quelques cas où le motif nucléotidique répété est supérieur à trois nucléotides²⁶. On parle de mutations dynamiques car elles évoluent d'une génération à l'autre.

²⁵ Bechkri, S. Cours de génétique, L2 Sciences biologiques. Constantine : Université Frères Mentouri, Faculté des sciences de la nature et de la vie, 2020

²⁶ Hanna, N., Parfait, B., Vidaud, D., et al. Mécanismes et conséquences des mutations, *médecine/sciences*, 2005. Vol. 21, No 11, pp. 969-980

En effet, pour chaque nouvelle génération, l'âge d'apparition de la maladie est de plus en plus précoce et la gravité des symptômes s'accroît, en raison de l'amplification* du nombre de motifs répétés au cours des successions. On retrouve ce cas de figure la plupart du temps chez les personnes atteintes d'une maladie dégénérative*.

Au fil des ans, le séquençage a permis l'identification de certaines mutations à l'origine de maladies génétiques chez l'homme. Le diagnostic génotypique des maladies héréditaires fait ainsi partie intégrante du conseil en génétique et du diagnostic prénatal. De plus, identifier le type de mutation ainsi que sa position dans le génome permet d'adapter la méthode de correction qui pourra potentiellement être employée.

2. Les types de maladie génétique

À ce jour, plus de 7 000 maladies génétiques ont été recensées dans le monde²⁷. Selon la ou les mutations qui les provoquent, ces maladies sont catégorisées différemment.

Les maladies chromosomiques

La maladie chromosomique, caractérisée par une anomalie de structure d'un chromosome ou la présence d'un chromosome surnuméraire (ou bien encore le manque d'un chromosome), est liée à une mutation chromosomique. La plus connue de ces pathologies est la Trisomie 21, causée par la présence de trois chromosomes 21 au lieu de deux (en effet, comme nous l'avons mentionné dans la partie I.1. « Le génome », chaque cellule humaine est censée contenir 23 paires de chromosomes).

²⁷Qu'est-ce qu'une maladie génétique ?. Institut Imagine [en ligne]. Disponible sur institutimagine.org/fr (Consulté en janvier 2023)

La mutation chromosomique est bien plus importante en taille que les petites mutations génétiques puisqu'elle touche tout ou partie d'un chromosome, supposant ainsi de grands réarrangements génomiques par modification, suppression ou duplication de plusieurs gènes à la fois. Les quatre cas de figures des mutations chromosomiques sont la duplication, la délétion, l'inversion et la translocation de tout ou partie d'un ou plusieurs chromosomes (voir fig.8).

Dans le cas de la duplication (comme pour la Trisomie 21), l'organisme vivant concerné a dans ces cellules une ou plusieurs copies d'un ou plusieurs chromosomes ou d'une partie d'un chromosome. Cet ajout d'information génétique peut donc avoir des conséquences nuisibles sur le phénotype, mais constitue parfois un atout. Par exemple, il existe certaines plantes polyploïdes, ce qui signifie qu'elles ne possèdent pas seulement deux exemplaires de chaque chromosome (comme la plupart des animaux et des autres plantes), mais trois ou plus²⁸. Cela peut favoriser la capacité d'innovation génomique de la plante, la rendre plus résistante à certains pathogènes ou encore augmenter la taille de ses feuilles et de ses fleurs. Les techniques de duplication chromosomique sont ainsi employées lors de la sélection de plantes d'intérêt agronomique.

Lorsque tout ou partie d'un ou plusieurs chromosomes est manquante, on parle de délétion chromosomique. De même que pour les duplications, ces pertes de matériel génétique influent conséquemment sur le phénotype. Par exemple, le syndrome de Prader-Willi, caractérisé entre autres par une obésité précoce et des troubles comportementaux et psychiatriques, est dû à l'absence d'une partie du chromosome 15²⁹.

Les inversions chromosomiques, ou cassures chromosomiques, correspondent au détachement d'un segment de chromosome qui se rattache par la suite au reste du chromosome dans le sens inverse. Dans ce cas, il n'y a habituellement pas de perte ni de gain d'information génétique, mais simplement un réarrangement* de la séquence de l'ADN³⁰.

²⁸ Pontvianne, F. Dupliquer pour s'adapter ou comment accélérer l'évolution des plantes ?. Laboratoire génome et développement des plantes, 2020. Centre National de la Recherche Scientifique (CNRS) [en ligne]. Disponible sur [cnrs.fr](https://www.cnrs.fr) (Consulté en mai 2023)

²⁹ Syndrome de Prader-Willi. Protocole National de Diagnostic et de Soins (PNDS). Centre de référence du syndrome de Prader-Willi et autres Obésités Rares avec troubles du comportement alimentaire, 2021

³⁰ Inversion chromosomique. Lexique, mis à jour en 2022. Aquaportail [en ligne]. Disponible sur [aquaportail.com](https://www.aquaportail.com) (Consulté en mai 2023)

Chez l'être humain, les inversions chromosomiques n'ont généralement pas d'autres effets nuisibles que les problèmes de fertilité.

Enfin, lors d'une translocation, une partie d'un seul chromosome s'attache à un autre chromosome non homologue, c'est-à-dire qui n'appartient pas à la même paire, et donc qui ne porte pas les mêmes gènes. La translocation est dite « équilibrée » s'il n'y a ni perte ni duplication de matériel génétique. Dans ce cas, il peut n'y avoir aucun effet sur la personne porteuse. Cependant, en présence de cette mutation chromosomique, les chances de l'apparition d'une translocation dite « déséquilibrée » chez la descendance sont plus élevées. Dans ce cas, le descendant héritera du chromosome atteint avec plus ou moins de matériel génétique, ce qui sera probablement source de dysfonctionnements.

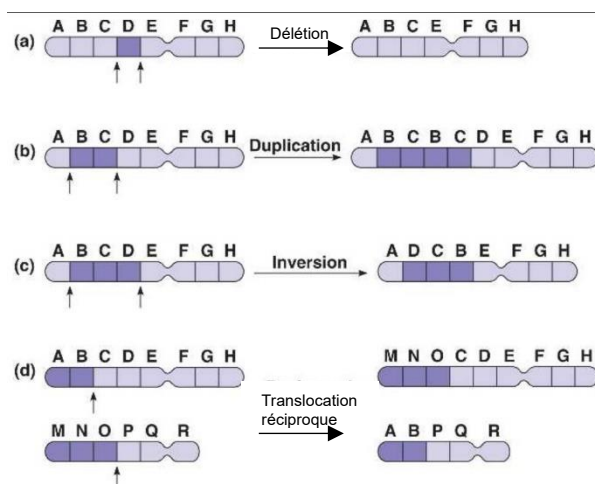


Figure 8 : les différents types de mutations chromosomiques. Tiré et adaptée de Pearson Education, Inc. 2008

Les maladies géniques

Les maladies géniques sont dues aux mutations génétiques (voir partie II.1. « La mutation génétique »). La maladie monogénique implique une mutation dans un seul gène, tandis que la maladie polygénique, de loin la plus fréquente, implique des mutations dans plusieurs gènes.

L'origine des maladies polygéniques, désormais plutôt appelées maladies multifactorielles, est souvent très difficile à déterminer car elles sont dues à plusieurs mutations dans plusieurs gènes, mais aussi à des facteurs de milieu³¹. Nous pouvons citer comme exemple de maladie polygénique le diabète, l'asthme, les maladies cardiovasculaires et neurologiques ou encore certains cancers.

Orphanet, le Portail des maladies rares et des médicaments orphelins, possède entre autres un inventaire, une classification et une encyclopédie des maladies rares ainsi que les gènes impliqués.

La transmission des maladies génétiques

Précisons tout d'abord que les versions chromosomiques de chaque paire de chromosomes présentes dans les gamètes (cellules reproductrices) correspondent aux versions qui seront transmises à la descendance. Par ce même principe, elles déterminent le sexe de l'enfant. En effet, la paire 23, qui correspond aux chromosomes sexuels, est soit composée d'un chromosome X et d'un chromosome Y (sexe masculin), soit de deux chromosomes X (sexe féminin). Ainsi, dans les gamètes femelles, ne se trouve qu'un chromosome X, tandis que pour les mâles, chaque gamète a soit un Y, soit un X. S'il y a un X, l'enfant aura une paire 23 XX (un chromosome X de la mère et un du père) et sera femelle. S'il y a un Y, l'enfant sera XY et donc mâle.

Le mode de transmission des maladies génétiques varie selon la mutation et sa présence ou non dans le génome des deux parents. En effet, l'enfant héritant, pour chaque paire de chromosome, d'un chromosome du père et d'un chromosome de la mère (Parie I.1. « Le génome »), il peut ou non hériter d'une mutation, selon les chromosomes sur lesquels il tombe.

³¹ Feingold, J. Maladies multifactorielles : un cauchemar pour le généticien, *médecine/sciences*, 2005. Vol. 21, No 11, pp. 927-933

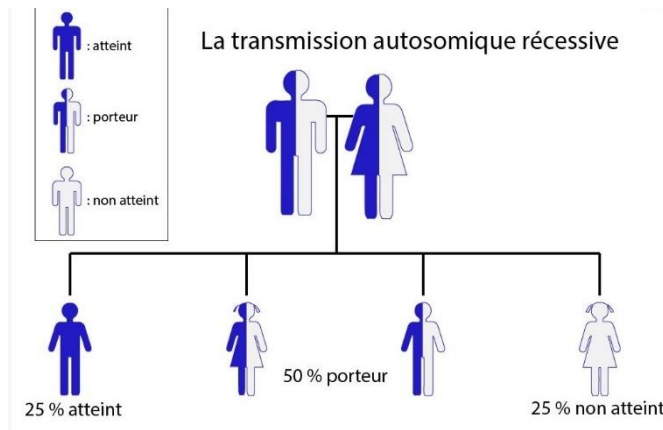
De plus, une mutation peut être héritée du génome des parents, mais elle peut aussi être nouvelle. Il s'agit alors d'une mutation *de novo** : aucun des génomes parentaux n'est porteur de la mutation et celle-ci apparaît par hasard. Il convient alors de préciser qu'une maladie génétique n'est pas nécessairement héréditaire. En effet, si les cellules reproductrices (gamètes) ne sont pas affectées par la mutation génétique, comme c'est le cas pour les maladies dues à des mutations induites (II.1. « La mutation génétique »), la maladie ne sera pas transmise.

Le terme de maladie héréditaire* est cependant réservé aux maladies génétiques. Le terme de maladie transmissible s'utilise plutôt lorsque la cause n'est pas génétique, comme pour les maladies sexuellement transmissibles ou celles liées à un facteur de milieu. Pour comprendre les mécanismes héréditaires des maladies génétiques, nous nous pencherons ici sur les différents modes de transmission.

La façon dont une maladie génétique se transmet diffère selon le type de chromosome sur lequel se trouve l'anomalie génétique en cause (autosome, chromosome sexuel (X, Y), chromosome de l'ADN mitochondrial*) mais aussi le caractère dominant ou récessif de l'anomalie génétique.

Si la mutation se trouve sur un chromosome autosome*, c'est-à-dire ne codant pas pour le caractère sexuel (contrairement aux chromosomes X et/ou Y de la 23^e paire), la transmission peut être autosomique dominante ou autosomique récessive. Dans le cas d'une maladie autosomique récessive, le caractère génétique de la maladie s'exprime chez les enfants ayant reçu le gène muté à la fois de leur père et de leur mère. Cela signifie que les deux parents doivent forcément être porteurs du gène muté sur l'un des chromosomes de leur paire. Ainsi s'explique le terme de récessif* : le caractère héréditaire doit être transmis par le père et la mère pour se manifester, car le gène muté doit se trouver sur les deux chromosomes de la paire chez l'enfant (voir fig.9).

Figure 9 : Le mode de transmission autosomique récessif. Tirée du CHU Sainte-Justine



Pour une maladie autosomique dominante, seule la transmission du caractère par le père ou la mère suffit pour qu'il se manifeste, donc il est possible que le gène muté ne soit présent que sur l'un des deux chromosomes de la paire dans le génome de l'enfant (voir fig.10). Nous pouvons aussi rappeler que le chromosome qui sera transmis à l'enfant dépend de la version du chromosome présente dans les gamètes du père et de la mère (puisqu'elles ne contiennent qu'un exemplaire de chaque paire, contrairement aux autres cellules).

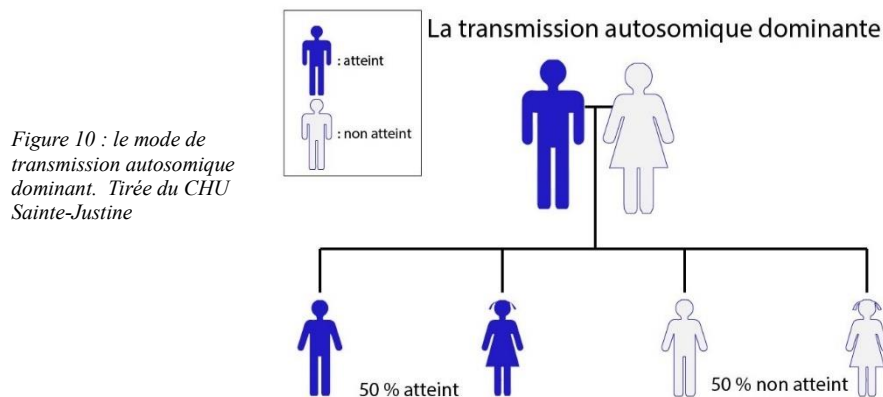
Lorsque le ou les gènes concernés par la mutation se trouvent sur le chromosome X, alors il s'agit d'une transmission « liée à l'X », qui peut de même être récessive ou dominante selon le type d'anomalie. Les maladies dont la transmission est liée à l'X atteint le plus souvent les garçons, puisque ces derniers n'ont qu'un seul chromosome X³². Les femmes porteuses n'ont quant à elle que 50 % de chance de transmettre la maladie (soit ses gamètes possèdent le chromosome X touché, soit elles possèdent le chromosome X non touché).

Dans de rares cas, la maladie s'hérite par transmission mitochondriale. Cette dernière est caractéristique des maladies dont la mutation génétique ne réside pas dans l'ADN du noyau des cellules, contrairement aux autres maladies génétiques, mais dans l'ADN des mitochondries*³³.

³² Comprendre les gènes. Article mis à jour en 2023. CHU Sainte-Justine [en ligne]. Disponible sur chusj.org (consulté en mai 2023)

³³ Maladies mitochondriales. Laboratoire Biologie mitochondriale, 2021. Institut Pasteur [en ligne]. Disponible sur pasteur.fr (consulté en janvier 2023)

Les mitochondries sont des organites présents dans le cytoplasme de toutes les cellules, essentielles au fonctionnement de ces dernières, et qui contiennent leur propre ADN. Or, elles se transmettent exclusivement par les ovules de la mère, c'est pourquoi on parle aussi de transmission maternelle.



Comprendre comment les maladies génétiques sont transmises fait partie de la manière dont les diagnostics prénataux sont interprétés et permet d'évaluer l'incidence d'une mutation chez une personne atteinte ou simplement porteuse. Le séquençage du génome constitue l'un des tests génétiques pouvant être effectués afin de déterminer la propension d'une personne à développer une maladie.

En nous penchant sur un cas concret de maladie génétique, il sera possible d'analyser tous les mécanismes en jeu dans les facteurs et le développement de dysfonctionnements liés à des mutations. Dans le texte-support que j'ai traduit, il m'a fallu m'intéresser tout particulièrement à la maladie de Duchenne. Appréhender cette maladie musculaire requiert certes de comprendre les principes génétiques que nous avons explorés jusqu'ici, mais aussi de connaître le rôle exact de certaines protéines dans le développement musculaire, et d'avoir quelques connaissances sommaires en anatomie générale et en médecine. De ce fait, il apparaît évident que la traduction d'un article spécialisé nécessite souvent des connaissances à la croisée d'un nombre varié de domaines.

Les interactions entre les différents champs de spécialisation, aussi bien dans le contenu que dans le langage emprunté, peuvent s'avérer complexes à identifier pour le traducteur. Il faut parfois approfondir ses compétences au regard de mentions seulement ponctuelles, car ces dernières peuvent être la clef d'une compréhension plus générale du texte ou de l'élucidation d'autres termes.

3. Le cas de la myopathie de Duchenne

Myopathie, n.f., du grec « Myo », qui signifie « muscle » et « pathos », soit « ce que l'on éprouve [de mal] »³⁴.

Les myopathies* sont des maladies génétiques liées à des mutations dans les gènes codant pour des protéines responsables de la fonction musculaire. Selon la mutation et la protéine touchée, les symptômes varient, mais dans tous les cas, les myopathies sont caractérisées par une forme de dysfonctionnement musculaire. La dystrophie* musculaire est une myopathie marquée par une dégénérescence progressive des cellules musculaires : au fur et à mesure du temps, les muscles s'atrophient, ce qui veut dire qu'ils perdent en volume, et donc s'affaiblissent.

Les cas les plus connus sont ceux de la dystrophie musculaire de Becker* (DMB) et de la dystrophie musculaire de Duchenne* (DMD), la dernière étant une forme aggravée de la première. La DMD touche dans le monde environ un garçon sur 3 500 à la naissance³⁵ et elle apparaît dès l'enfance. C'est une maladie monogénique récessive rare et liée à l'X (partie II.2. « Les types de maladie génétique »), ce qui explique qu'elle touche particulièrement les garçons (liée à l'X), mais aussi qu'elle peut se transmettre par des parents qui sont porteurs sains (récessive). Cependant, dans un tiers des cas de DMD, il s'agit d'une mutation *de novo*, qui n'est pas présente chez la mère.

³⁴Tiré des définitions de « myo » et « pathos ». Trésor de la langue française informatisé (TLFi) [en ligne]. Disponible sur atilf.atilf.fr (Consulté en mai 2023)

³⁵ La dystrophie musculaire : comment les cellules souches pourraient-elles aider ?. Article, mis à jour en 2016. EuroStemCell [en ligne]. Disponible sur eurostemcell.org/fr (consulté en mai 2023)

Dans tous les cas, la ou les mutations touchent le gène codant la dystrophine*, une protéine essentielle au maintien de l'intégrité musculaire*.

Afin de comprendre l'impact du manque de dystrophine sur la cohésion des muscles, nous pouvons nous pencher sur les mécanismes musculaires.

Le fonctionnement des muscles :

Nous nous intéresserons ici aux muscles striés squelettiques (par oppositions au muscle strié cardiaque* et aux muscles lisses), qui représentent 40 % de la masse du corps humain et sont liés aux fonctions motrices et à la production de force³⁶.

Le tissu des muscles squelettiques* est fait de faisceaux de fibres musculaires. Ces fibres sont de très grandes cellules, appelées myocytes*, entourées d'une membrane, le sarcolemme*. Les myocytes contiennent plusieurs noyaux, un cytosquelette* fait de filaments d'actine permettant de structurer la cellule, ainsi que des myofibrilles*, assemblées les unes aux autres dans le cytoplasme de la cellule. Chaque molécule de myofibrille se compose d'une chaîne de petites unités contractiles appelées sarcomères*³⁷. Enfin, dans un sarcomère se trouvent des filaments d'actine et de myosine*, les myofilaments (voir fig.11). Par interaction entre ces myofilaments, les sarcomères peuvent se rapprocher entre eux, ce qui raccourci la chaîne des sarcomères, contractant ainsi les myofibrilles, et par extension les cellules musculaires... et donc le muscle.

Les myofibrilles, qui sont composées de sarcomères, sont maintenues à la membrane de la cellule musculaire, le sarcolemme, par la protéine de la dystrophine. Cette très grande protéine, située juste en dessous du sarcolemme, assure la liaison entre, d'un côté, les filaments du cytosquelette reliés aux myofibrilles, et de l'autre, un complexe protéique transmembranaire, relié au sarcolemme³⁸.

³⁶ Bensalah, M. Fibrose musculaire : acteurs cellulaires et stratégies thérapeutiques. Thèse en Médecine humaine et pathologie. Paris : Sorbonne Université, 2019, 274 p

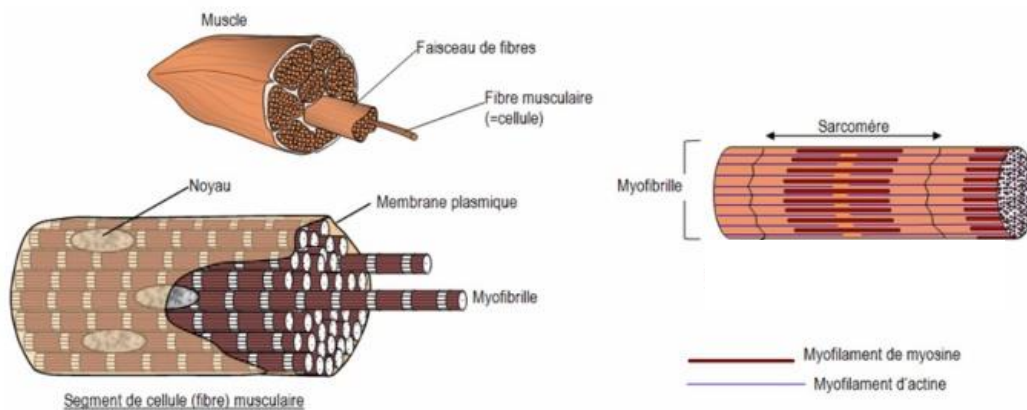
³⁷ Lemarchand, P. MOOC côté cours : La contraction musculaire [VIDEO]. Inserm, 2015

³⁸ Benchaouir, R., Meregalli, M., Farini, A., et al., Réhabilitation de la dystrophine humaine dans le modèle scid/mdx après greffe de cellules souches de patients DMD corrigées par saut d'exon. *Médecine/sciences*, 2008. Vol. 24, No 1, pp. 99-101

Le tout forme le complexe glycoprotéique associé à la dystrophine (CGD)* (voir fig.12). Le rôle de liaison de la dystrophine est donc essentiel au maintien des myofibrilles dans les cellules du muscle. En l'absence de dystrophine, les sarcolemmes sont déstabilisés à chaque contraction, ce qui abîme les muscles et finit par les détruire tout à fait.

En conditions normales, le muscle squelettique peut se régénérer intégralement après une lésion, et ce tout au long de la vie, grâce à un type de cellules souches adultes, les cellules satellites. Elles se situent le long des fibres musculaires et sont capables de générer de nouvelles fibres ainsi que de les fusionner aux fibres existantes. Cependant, lorsque la fréquence des lésions est trop élevée, comme c'est le cas en l'absence de dystrophine, les cellules satellites sont surmenées et la dégénérescence musculaire finit par prendre le dessus. L'une des voies explorées pour traiter la DMD se trouve justement dans l'implantation de ce type de cellules souches dans le sujet malade.

Figure 11 : Les différents niveaux d'organisation de la structure du muscle. Tirée et adaptée de la Banque de Schémas. SVT, Académie de Dijon.

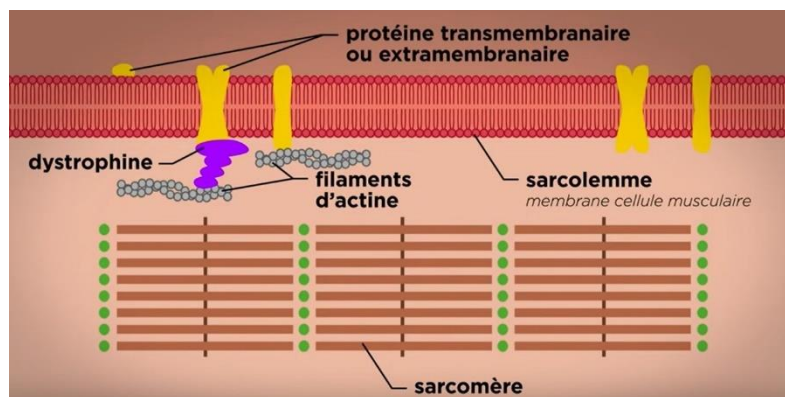


Causes

La DMD est due à une ou plusieurs mutations du gène codant la dystrophine (le gène DMD, long de 2,3 Gb) sur le chromosome X. Les mutations sont le plus souvent de larges délétions dans les exons 45 à 55. Dans le cas de la DMD, elles décalent le cadre de lecture de l'ADN, qui est alors mal retranscrit.

L'épissage de l'ARN pré-messager n'a pas lieu, ce qui aboutit à une non-production de dystrophine (parie II.1 « La mutation génétique »). Dans le cas de la dystrophie de Becker, un ou deux exons sont manquants mais le cadre de lecture n'est pas décalé, ce qui permet la production d'une dystrophine fonctionnelle mais tronquée³⁹.

Figure 12 : Le rôle de maintien de la dystrophine dans les myofibrilles. Extrait du MOOC de l'Inserm : La contraction musculaire



Symptômes

Dès l'âge de trois ans, les enfants atteints de la DMD montrent des signes de difficulté motrice, notamment une faiblesse musculaire au niveau des membres inférieurs. L'un des premiers signes cliniques est l'augmentation du taux de créatine-phosphokinase* comparé à un sujet sain, car il s'agit d'une enzyme libérée par les fibres musculaires lorsqu'elles sont endommagées⁴⁰. Au fil des ans, outre les troubles cognitifs, ce sont les jambes puis les bras et le tronc qui s'affaiblissent. Enfin, le muscle du cœur finit par être atteint, et le patient meurt généralement prématurément d'une insuffisance respiratoire*.

³⁹ Pichavant, C., Thérapie génique de la dystrophie musculaire de Duchenne - Utilisation de transgènes de la dystrophine chez le modèle canin. Thèse en biologie cellulaire et moléculaire. Québec : Faculté des études supérieures de l'Université Laval, 2010, 176 p.

⁴⁰ Pichavant, C. Ibid

À l'échelle somatique, plusieurs fonctions sont donc touchées, toutes dépendantes du fonctionnement musculaire : les muscles moteurs, le cœur, le cerveau, les systèmes respiratoire et digestif, etc.

À l'échelle cellulaire et musculaire, plusieurs symptômes sont identifiables, notamment la fibrose*, l'infiltration cellulaire* ou encore l'inflammation. L'inflammation survient lors de ruptures de la membrane des myocytes (les cellules musculaires). En effet, ces lésions provoquent une fuite de certains débris cellulaires (comme la créatine kinase), qui nécessitent l'intervention de cellules inflammatoires pour être nettoyés.

Traitements

Une partie de la recherche et des traitements en cours porte sur la limitation de l'impact de la maladie sur l'organisme, par exemple, la réduction de la fibrose ou de l'inflammation. L'autre partie porte sur la cause, c'est-à-dire sur le rétablissement de l'expression de la dystrophine dans le système. Il n'y a actuellement aucun traitement curatif de la DMD, mais la thérapie cellulaire consistant à transplanter des cellules productrices de dystrophine est déjà employée. Quant à elle, la voie de la thérapie génique actuellement appliquée profite à un certain type d'anomalie du gène de la DMD (environ 13 % des cas) et la mise sur le marché du médicament correspondant n'est que conditionnelle en Union européenne. Une autre méthode de thérapie génique vise à provoquer l'expression d'une dystrophine tronquée, comme elle l'est dans le cas de la myopathie de Becker ⁴¹ (voir partie III.1 « Les différentes approches »).

⁴¹ Dystrophie musculaire de Duchenne. Article mis à jour en 2023. AFM-Téléthon [en ligne]. Disponible sur afm-telethon.fr (consulté en juin 2023)

Bien que pour la plupart des maladies génétiques, des traitements contre les symptômes soient envisageables et couramment employés, les traitements curatifs sont encore loin d'être pratiqués en routine. La médecine de précision permettant l'atténuation de l'expression d'un gène défectueux ou bien même sa modification constitue des voies de recherche importantes pour soigner les maladies génétiques. Des outils de thérapie génique et de thérapie cellulaire sont déjà utilisés pour traiter certains patients et la recherche avance à grand pas. Dans cette troisième partie, nous étudierons les différentes méthodes de thérapie génique, toujours dans le domaine médical. De plus, en nous penchant sur le fonctionnement de CRISPR-Cas9, nous pourrions mesurer l'importance et l'impact de cet outil en médecine.

III- Les méthodes de la thérapie génique

La thérapie génique regroupe toutes les méthodes biotechnologiques visant à introduire du matériel génétique dans les cellules pour soigner une maladie génétique. Son utilisation est déjà répandue et certains médicaments de thérapie génique sont d'ailleurs sur le marché en Europe, aux États-Unis et en Chine.

1. Les différentes approches

Ajouter un gène

Première méthode de thérapie génique employée, la supplétion génique consiste à importer une copie d'un gène entier sain, pour qu'elle soit traduite en une protéine fonctionnelle dans les cellules. Cette approche ne modifie pas le gène défectueux mais ne fait que le compléter (voir fig.13). Ainsi, le génome de l'organisme concerné demeure inchangé, mais grâce à l'addition récurrente d'une information génétique saine, l'expression qui fait défaut peut être rétablie à des taux acceptables.

La stratégie de supplétion du gène avec un gène thérapeutique* peut être faite *in vivo* (directement dans l'organisme), ce qui suppose l'utilisation d'un vecteur (voir partie III.2 « Vecteurs et ciblage »), ou *in vitro*. Dans le deuxième cas, les cellules sont prélevées chez le patient, puis modifiées en laboratoire et réinjectées dans l'organisme⁴², ce qui limite les effets dans les organes non-ciblés. L'approche *ex vivo* est souvent utilisée pour les maladies sanguines car les cellules à modifier sont prélevables simplement par prise de sang. D'autres cellules sont en revanche impossibles à isoler ou à cultiver, comme les cellules du cerveau ; dans ce cas le gène thérapeutique est introduit *in vivo*, directement par injection dans les tissus malades, ou par voie sanguine.

Le gène thérapeutique peut être synthétisé en laboratoire, mais il peut aussi provenir de l'homme ou d'un autre organisme. Ainsi pouvons-nous appréhender le concept de la transgénèse*, une approche envisageable dans les applications biotechnologiques. Dans cette approche, le transgène*, défini par l'Institut national de la santé et de la recherche médicale comme « séquence génétique isolée, transférée d'un organisme à un autre »⁴³, devient thérapeutique.

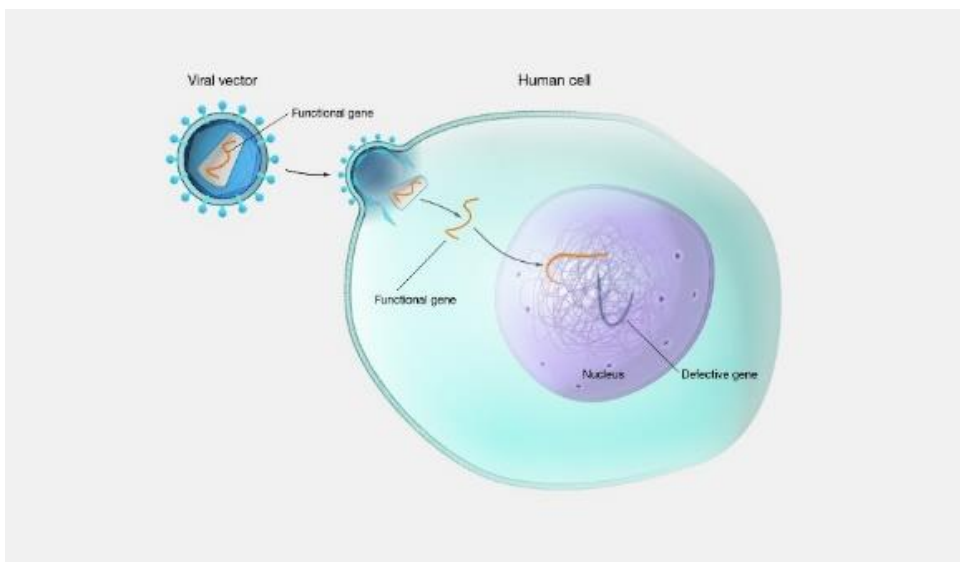


Figure 13 : Supplétion de gène *in vivo*. Tirée du National Institute of Health.

⁴² Thérapie génique - Une recherche de longue haleine qui porte ses fruits. Article mis à jour en 2018. Inserm [en ligne]. Disponible sur [inserm.fr](https://www.inserm.fr) (Consulté en juin 2023)

⁴³ Ibid

Altérer l'expression d'un gène

Certaines méthodes de thérapie génique consistent à influencer l'expression d'un gène en modifiant les ARN messagers ou pré-messagers, qui acheminent l'information génétique de l'ADN nucléaire aux ribosomes. Il ne s'agit pas ici d'introduire un nouveau matériel génétique, ou de modifier celui présent dans la cellule, mais plutôt d'intervenir au moment de la traduction de ce matériel génétique. Pour ce faire, il est possible d'utiliser des ARN interférents ou bien des oligonucléotides antisens* (OAN) qui agissent comme des sortes de messagers pour les messagers !

Les ARN interférents sont des petites séquences d'ADN ou d'ARN non codants qui, en se fixant à l'ARNm, vont empêcher le message d'être transmis et ainsi bloquer la production de la protéine ciblée⁴⁴. Ce sont donc des ARN inhibiteurs d'expression du gène qui agissent après l'étape de transcription (voir partie I.2 « Le code génétique, de l'ADN à la protéine »). Ils sont naturellement présents dans l'organisme et permettent de détruire des ARNm dont la transcription aurait été erronée. C'est ce qui est désigné par le processus de dégradation des ARNm médiée par les brins non-sens*, ces derniers étant des ARNm dont la création lors de la transcription a mené à l'apparition d'un codon stop prématuré. Lorsque l'organisme repère des ARNm non-sens, il se charge de les détruire pour empêcher qu'une protéine déficiente ne s'exprime. Certaines pathologies sont ainsi caractérisées par une dérégulation de l'expression des ARN interférents. La découverte des ARN interférents par les chercheurs Andrew Z. Fire et Craig C. Mello leur a valu le prix Nobel de médecine en 2006. En effet, ces ARN présentent un intérêt double : ils peuvent révéler la fonction d'un gène (en inhibant son expression, on peut observer les effets sur le phénotype), et bien sûr être utilisés comme outil thérapeutique en empêchant une mutation de s'exprimer⁴⁵.

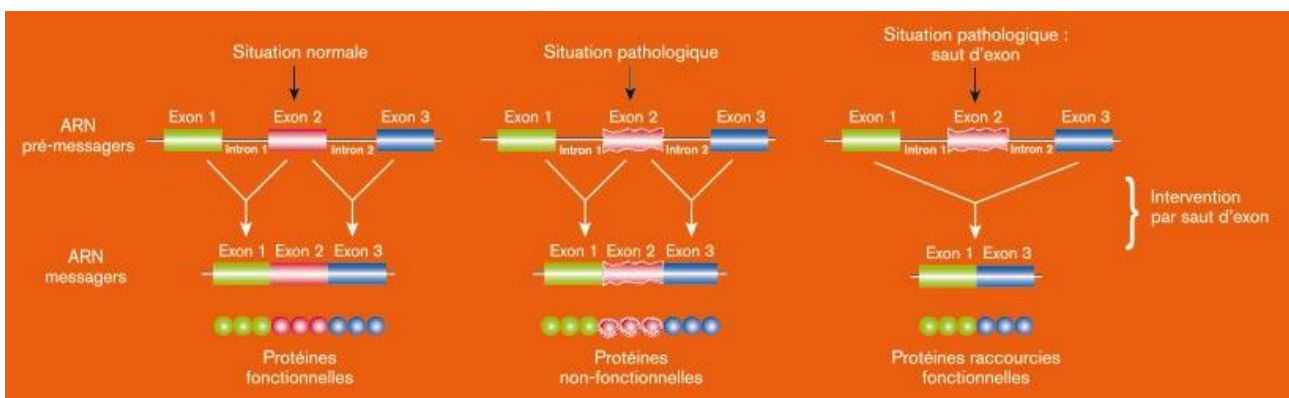
De leur côté, les oligonucléotides antisens sont des fragments d'ARN synthétiques dont la séquence est complémentaire à celle d'un ARNm ciblé.

⁴⁴ Thérapies à ARN. Un domaine thérapeutique en pleine expansion. Dossier mis à jour en 2022. Inserm [en ligne]. Disponible sur [inserm.fr](https://www.inserm.fr) (Consulté en juin 2023)

⁴⁵ F. Ibarondo, F., Camus, G. Une nouvelle classe d'ARN : les petits ARN interférents. École Normale Supérieure, Planet-Vie, 2005

Son intervention lors de l'étape de maturation de l'ARNpm vers l'ARNm, l'épissage, permet d'induire le saut ou au contraire l'incorporation d'exons, modifiant ainsi le message génétique et aboutissant à la production d'une protéine fonctionnelle. L'approche par « saut d'exon* », qui consiste à exclure un ou plusieurs exons mutés afin de rétablir le cadre de lecture de l'ARN, est notamment utilisée pour traiter la maladie de Duchenne⁴⁶ (voir fig.14). Les exons portant un message erroné sont supprimés, la protéine obtenue est certes amputée (moins d'exons signifie moins d'acides aminés) mais elle reste plus ou moins fonctionnelle comparée à une absence de production. Au contraire, d'autres approches visent à rétablir l'expression normale d'un gène de type sauvage en incluant un ou plusieurs exons dans l'ARN.

Figure 14 : Schéma de la technique du saut d'exon. Tiré et adapté de l'Inserm, F. Koulikoff



Comme pour toute méthode de thérapie, certains paramètres liés à la modification de l'ARNm, par inhibition de l'expression ou par modulation de l'épissage, ne sont pas parfaitement contrôlables. Par exemple, en venant volontairement supprimer certains exons sur l'ARNm, induisant ainsi, l'apparition d'un codon stop prématuré ou même encore l'élimination du codon stop, il est possible que l'organisme enclenche plus tôt que prévu le mécanisme de dégradation des ARN, ce qui annulerait les effets thérapeutiques recherchés.

⁴⁶ Le Guen, YT., Le Gall, T., Laurent, V., et al., Dystrophie musculaire de Duchenne : état actuel et perspectives thérapeutiques. *Bulletin de l'Académie Nationale de Médecine*, 2021. Vol. 205, No 5, pp. 509-518

« Éditer » le génome

Lorsque le gène muté est éliminé ou réparé directement dans l'ADN nucléaire de la cellule, on parle d'édition du génome. Cette stratégie, encore très expérimentale, suppose d'insérer dans la cellule deux composants : une enzyme spéciale capable de couper de l'ADN et un segment d'ADN qui correspond à une version saine d'une partie du gène afin de restaurer le génome. Modifier directement le génome suppose que la réparation soit transmise, en théorie à toutes les cellules filles issues de la division cellulaire de la cellule mère traitée (voir partie I.1. « Le génome »), ce qui veut dire que le traitement par édition génique serait définitif. Dans les faits, de nombreux facteurs et effets indésirables rendent cette approche imparfaite. Par exemple, le processus de dégradation des ARNm médiée par les brins non-sens peut annuler les effets réparateurs en empêchant les ARNm issus de la modification génomique de traduire à temps des protéines saines. Il faut aussi prendre en compte l'immunogénicité* des enzymes utilisées, c'est-à-dire leur capacité à provoquer une réaction immunitaire, ce qui pose un problème évident pour le corps hôte.

L'outil de CRISPR-Cas9, que nous étudierons en partie III.3, correspond à une méthode d'édition du génome. Ce cas particulier permet de comprendre le fonctionnement global de cette approche.

2. Vecteurs et ciblage

L'organisme humain possède un mécanisme de défense immunitaire solide et complexe. Insérer un corps étranger dans le noyau des cellules, type gène sain, brin d'ARN ou autre, n'est donc pas une tâche facile et suppose l'intervention d'un véhicule capable de franchir toutes les barrières du corps pour acheminer les produits à bon port.

Vecteurs

Une fois injectés de manière locale ou systémique, il faut rendre les composants de la thérapie génique assimilables, c'est-à-dire leur faire gagner les cellules cibles et leur faire franchir les membranes cellulaires. À cet égard, des vecteurs sont utilisés, le plus souvent viraux. Les vecteurs viraux*, qui sont employés dans plus de 75 % des essais de thérapie génique⁴⁷, profitent des propriétés pénétratives du virus pour livrer le matériel génétique désiré. Parmi ces vecteurs viraux, les virus adénoassociés* (soit AAV, notamment utilisés pour les maladies neuromusculaires comme la DMD) et les lentiviraux* sont largement utilisés. Ils constituent une nouvelle génération de vecteurs, plus sûre que l'ancienne dont les effets indésirables liés à la propagation du virus étaient parfois incontrôlables.

Le conditionnement du virus vecteur en un véhicule de livraison des composants de thérapie génique se fait en culture. La première étape consiste à le rendre inoffensif, c'est-à-dire à lui ôter les parties de son ADN responsables de ses propriétés pathogènes et de sa capacité à se reproduire dans une cellule hôte. Son immunogénicité est aussi réduite au maximum. Ne lui sont laissés que les gènes codants pour ses éléments constitutifs. Ensuite, le brin thérapeutique, tel un passager, est intégré au virus. Le virus obtenu pourra pénétrer la cellule et délivrer son matériel génétique dans le noyau. Selon son conditionnement, soit le virus ne s'intégrera pas aux chromosomes de la cellule (vecteur non intégratif), soit il s'intégrera au génome de la cellule ciblée (vecteur intégratif). Les vecteurs non intégratifs sont utilisés pour intégrer des cellules post-mitotiques, qui ne se diviseront plus. Dans ce cas, le matériel génétique délivré par le vecteur reste dans le noyau de la cellule pendant toute la durée de sa vie. Au contraire, les vecteurs intégratifs permettent de traiter les cellules qui se multiplient beaucoup.⁴⁸

Séquences de reconnaissance / Promoteurs

⁴⁷ Thérapie génique - Une recherche de longue haleine qui porte ses fruits. Article mis à jour en 2018. Inserm [en ligne]. Disponible sur [inserm.fr](https://www.inserm.fr) (Consulté en juin 2023)

⁴⁸ Thérapie génique. Glossaire, mis à jour en 2022. AFM-Téléthon [en ligne]. Disponible sur [afm-telethon.fr](https://www.afm-telethon.fr) (consulté en mai 2023)

Une fois dans le noyau de la cellule, le même problème se pose : comment faire en sorte que l'élément injecté dans la cellule soit lu et activé ? qu'il se fixe sur le bon locus de l'ADN, ou sur le bon brin d'ARN ? « Une des limites importantes de l'utilisation de la thérapie génique est l'absence de contrôle de l'expression du transgène »⁴⁹.

Pour induire l'expression d'un gène thérapeutique, complémentaire à la séquence ciblée dans l'organisme, il peut être combiné à un ADN régulateur, un promoteur*, qui contrôle la séquence principale du matériel génétique injecté⁵⁰, et qui se situe en amont de la séquence retranscrite⁵¹. Bien que présents naturellement dans l'organisme, les promoteurs peuvent aussi être synthétisés à des fins thérapeutiques. Il s'agit de séquences de reconnaissance silencieuses mais déterminantes dans l'expression du gène.

Pour reprendre de manière simplifiée, nous pourrions dire que la thérapie génique se fait en trois étapes. La première consiste à **repérer la mutation** responsable de la maladie et à trouver un élément génique qui puisse pallier le problème causé par cette mutation. La deuxième étape est de faire en sorte, grâce à un **vecteur**, que l'élément thérapeutique atteigne la cellule et y pénètre. La troisième étape fait intervenir un **promoteur**, assurant le bon fonctionnement de l'élément thérapeutique étranger au sein de la cellule malade.

⁴⁹ A. Bruhat, A., Chaveroux, C., Carraro, V., et. al. Réguler l'expression d'un transgène thérapeutique par un régime alimentaire carencé en un acide aminé indispensable. *Médecine/sciences*, 2017. Vol. 33, No 2, pp. 136-139

⁵⁰ Académie suisse des sciences naturelles, Dossier, Focus biologie synthétique, Thérapie génique

⁵¹ Ibarrondo, F., Camus, G. La transcription chez les eucaryotes. École Normale Supérieure, Planet-Vie, 2007 /ca

3. Le cas de CRISPR-Cas9

Outil développé en 2012 et depuis répandu chez les biologistes, la technique de CRISPR-Cas9 a valu en 2020 un prix Nobel à ses deux créatrices, les chercheuses Jennifer Doudna et Emmanuelle Charpentier. En venant inciser l'ADN de n'importe quel génome, il peut induire des mutations ciblées.

CRISPR-Cas9 est un outil d'édition du génome qui a rendu la modification d'un gène très accessible par sa simplicité. À l'origine, il s'agit d'une trouvaille fortuite : la découverte, dans le génome d'une bactérie, de petites séquences nucléotidiques (d'une trentaine de bases) répétées, qui sont des sortes de palindromes. Cela signifie qu'elles se lisent dans les deux sens (comme le mot « ressasser »). On les a plus tard appelées CRISPR* pour « Clustered Regularly Interspaced Short Palindromic Repeats » (courtes répétitions en palindrome regroupées et régulièrement espacées). Les séquences CRISPR des bactéries étaient entrecoupées d'autres séquences intermédiaires, d'à peu près la même taille.

Vingt ans après est découverte l'utilité des séquences entre les CRISPR, les « protospacers »⁵² : elles sont exactement identiques à de l'ADN de virus, et il s'agit du système immunitaire primitif de la bactérie. En effet, certains virus, les bactériophages, peuvent infecter des bactéries (comme ils infecteraient un être humain). Pour se défendre des bactériophages, les bactéries conservent un petit bout de leur ADN, leur permettant de les identifier et de les détruire s'ils reviennent. Le fait que cet ADN de virus soit stocké dans l'ADN de la bactérie le rend d'ailleurs héréditaire.

La bactérie possède des enzymes*, les protéines Cas (pour « CRISPR-associated »), capables d'identifier ces séquences d'ADN viral. Lorsque la bactérie subit une agression de virus, elle déclenche une transcription de toute la partie de son ADN dans lequel il y a les séquences palindromiques et les séquences intermédiaires (protospacers), pour obtenir un brin d'ARN complémentaire.

⁵² Tremblay, JP. CRISPR, un système qui permet de corriger ou de modifier l'expression de gènes responsables de maladies héréditaires. *Médecine/sciences*, 2015. Vol. 31, No 11, pp. 1014-1022

Après maturation, l'ARNpm obtenu se divise en plusieurs petits ARN guides* (ARNg). Les protéines Cas récupèrent ces derniers, puis partent à la recherche, dans la bactérie, de séquences complémentaires à l'ARNg. Si elles trouvent un bout d'ADN qui correspond, elles s'y accrochent et le découpent, ce qui neutralise le virus. Ces bactéries possèdent donc un système d'identification et de découpage de certaines séquences ciblées d'ADN (ici, celles des bactériophages). L'outil de CRISPR-Cas9 réutilise ce mécanisme pour l'appliquer à n'importe quelle séquence d'ADN ciblée.

Coupage de l'ADN avec l'outil CRISPR-Cas9

CRISPR-Cas9, tel des « ciseaux moléculaires », est utilisé comme système de chirurgie du génome. C'est un complexe qui combine une enzyme Cas, la Cas9, et un ARN guide complémentaire à une séquence ciblée, afin d'induire des cassures dites « double-brin »* dans l'ADN. À noter que certaines technologies CRISPR peuvent aussi induire une coupure simple brin.

La Cas9 (pour « CRISPR-associated protein 9 ») est souvent issue de la bactérie *S. Pyogène* et elle est programmable pour découper n'importe quel bout d'ADN de manière ciblée. L'ARN guide incorporé dans la Cas9 est constitué d'une séquence variable d'environ 20 nucléotides complémentaires à la séquence ciblée. C'est la séquence homologue⁵³. L'ARNg est aussi constitué d'une séquence constante (la séquence palindromique qui va former une épingle), et une séquence NGG ou NGA (appelée motif PAM*, pour « *protospacer adjacent motif* »), qui suit la séquence homologue et permet à la Cas9 de repérer l'endroit où elle doit se lier à l'ADN (voir fig. 15 et 16). Les motifs PAM sont très fréquents dans tous les génomes. Ainsi, un nombre très important de séquences peut être ciblé.

Ensemble, L'ARNg et la Cas9 forment un complexe capable de repérer une séquence d'ADN ciblée, et de la découper au niveau des deux brins. Ce complexe peut être intégré à une cellule humaine via un virus adénoassocié faisant office de vecteur.

⁵³Gilgenkrantz, H. La Révolution des CRISPR est en marche. *Médecine/sciences*, 2014, Vol. 31, No 12, pp. 1066-1069

Figure 15 : Représentation du système CRISPR-Cas9. Illustration tirée de Médecine/sciences, 2015

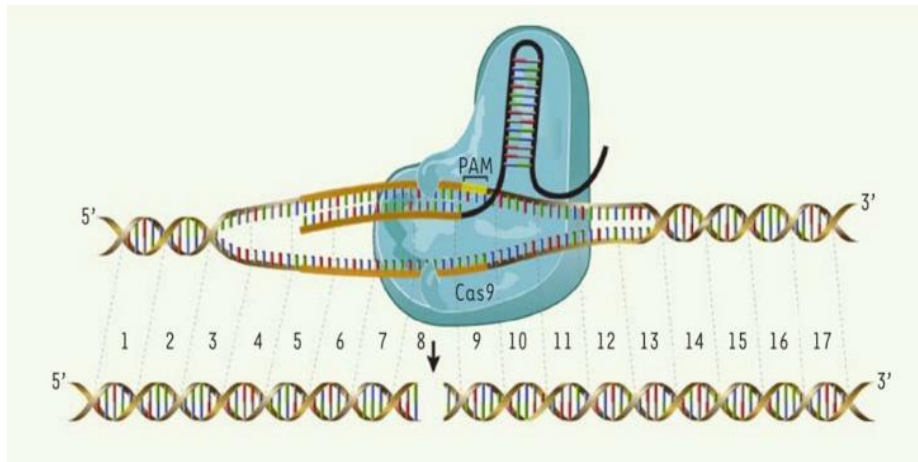
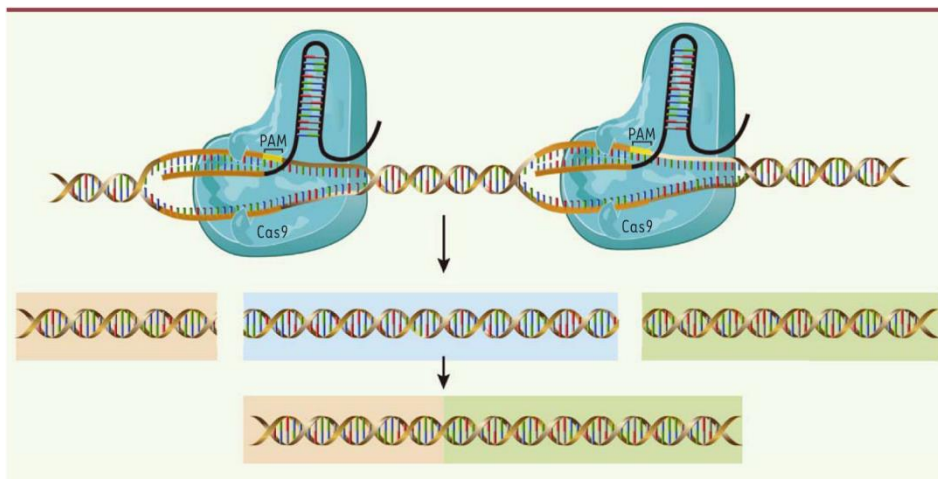


Figure 16 : Utilisation du système CRISPR/Cas pour éliminer une séquence d'ADN. Il est possible de couper, en utilisant deux ARNg, les deux brins d'ADN à deux endroits dans un même chromosome. Ceci permet d'enlever une séquence de quelques nucléotides ou même de centaines de milliers de nucléotides. Illustration tirée de Médecine/sciences, 2015



Recombinaison homologue ou non homologue

La Cas9 peut induire une coupure sur l'ADN, mais elle ne le répare pas. Afin de modifier l'ADN, il faut trouver un moyen de remplacer la séquence dont on a voulu se débarrasser, ou bien recoller l'ADN scindé. Il existe justement un mécanisme de réparation naturellement présent dans les cellules, dit de « recombinaison ».

Si un morceau d'ADN est endommagé, les cellules vont en permanence essayer de le réparer⁵⁴.

Dans le cadre d'une recombinaison homologue* (HDR, pour l'anglais *Homology-directed repair*), il suffit d'injecter une séquence d'ADN donneur, qui contient des séquences de nucléotides homologues à celles qui précèdent et suivent le site de coupure, pour que la cellule l'utilise lors du processus de réparation. En l'absence d'un ADN donneur, il y a recombinaison non homologue* (NHEJ, pour l'anglais *Non-homologous end-joining*, soit « jonction d'extrémités non-homologues »)⁵⁵, moins précise, souvent caractérisée par des micro-délétions ou micro-insertions d'un nombre variable de nucléotides. Si elles ne sont pas des multiples de trois nucléotides, ces petits indels (insertions et délétions) modifient le cadre de lecture du gène « résultant soit en un gène non fonctionnel soit dans le rétablissement d'un gène fonctionnel lorsque le gène ciblé ne l'était pas.⁵⁶ »

Pour citer une fois de plus Jacques P. Tremblay, professeur en médecine moléculaire à la Faculté de médecine de l'Université Laval, à Québec : « Les recombinaisons homologues conduisent à des modifications très précises, ce qui est très utile pour corriger des gènes impliqués dans des maladies héréditaires. Elles peuvent également permettre d'introduire au niveau d'un gène une séquence codant pour un rapporteur afin de détecter la protéine qui en résulte.⁵⁷ »

Risques

Malgré sa précision de ciblage, son efficacité et son accessibilité, l'outil de CRISPR-Cas9, de même que toute méthode d'édition du génome, n'est pas exempt de risques. Le premier concerne la précision du système.

⁵⁴ Louapre, D., Modifier le génome avec CRISPR [VIDEO]. ScienceEtonnante, 2015

⁵⁵ Tremblay, JP. CRISPR, un système qui permet de corriger ou de modifier l'expression de gènes responsables de maladies héréditaires. *Médecine/sciences*, 2015. Vol. 31, No 11, pp. 1014-1022

⁵⁶ Tremblay, JP. Ibid

⁵⁷ Tremblay, JP. Ibid

Comme l'explique Jennifer Doudna :

« As we transition into the next decade of CRISPR genome editing, several key challenges demand innovative solutions. Two of these are editing accuracy (i.e., specificity for the target site) and precision (i.e., producing the exact desired editing outcome). »⁵⁸

CRISPR-Cas9 est en effet capable de provoquer des effets hors cible, liés par exemple à des coupures qui n'ont pas été opérées au bon endroit sur la séquence d'ADN. Les effets hors cible dépendent aussi du vecteur et du guide utilisé pour permettre l'expression de la Cas9 dans la cellule. En outre, ces effets sont difficiles à évaluer à l'échelle d'un gène, et supposent l'utilisation de logiciels de traitement de données modernes.

De plus, il existe des risques liés à l'immunogénicité de Cas9. En effet, voici ce que nous expliquent Adita Mehta et Olivier M. Merkel, du Département de Pharmacie de l'Université Ludwig-Maximilian à Munich :

« Recent studies have demonstrated an innate and adaptive cellular immune response to Cas9 in mouse models and the presence of anti-Cas9 antibodies and T-cells in human plasma. Pre-existing immunity against therapeutic Cas9 delivery could decrease its efficacy in vivo and may pose significant safety issues. »⁵⁹

Le problème d'immunogénicité peut avoir un premier effet, moins grave : l'inefficacité du système (comme le processus de dégradation des ARNm médiée par les brins non-sens, qui annule les effets de l'édition génomique). Il peut aussi avoir un effet plus grave, celui de l'impact imprévu sur le corps lié à toute réaction immunitaire.

⁵⁸ Wang, JY., Doudna, AJ. CRISPR technology: A decade of genome editing is only the beginning. *Science*, 2023. Vol. 379, No. 6629

⁵⁹ Mehta, A., et Merkel, OM. Immunogenicity of Cas9 Protein. *Journal of Pharmaceutical Science*, 2019. Vol. 109, pp. 62-67

En conclusion, adapté du système immunitaire d'une bactérie, CRISPR-Cas9 est un outil très précis, ce qui limite, mais n'exclut pas, les risques d'effets secondaires. C'est une découverte bon marché et applicable à une grande variété de domaines biologiques, puisqu'elle s'adapte à n'importe quel génome de n'importe quelle espèce vivante. Ces ciseaux moléculaires ouvrent des perspectives très intéressantes en recherche et en médecine, étant employés à la fois pour permettre de comprendre le rôle d'un gène, mais aussi pour développer des thérapies contre les maladies génétiques héréditaires ou acquises. Ils constituent donc un potentiel développement thérapeutique pour soigner la DMD.

Cependant, c'est parce que cette technologie est démocratisée au sein de la communauté scientifique qu'elle est aussi potentiellement dangereuse. Pour lui attribuer une utilisation spécifique, il convient de mener des études, d'analyser les résultats, et d'avoir conscience de ses risques. Notre quatrième et dernière partie porte sur les champs de l'application et de la recherche scientifique, parts fondamentales de la compréhension d'un article scientifique pour un traducteur. Mener des expériences et interpréter des résultats impliquent un vocabulaire précis et rigoureux, des ordres et des unités de grandeurs spécifiques, ainsi qu'une structure interne déterminante dans la rédaction.

IV- Recherche, application et éthique

En entrant dans le domaine des sciences, il convient de ne pas confondre pratiques expérimentales avec scientifique fou. Justement, le mot « expérimental », caractérisé par l'emploi systématique de l'expérience⁶⁰, est sémantiquement opposé à celui de « abstrait ».

⁶⁰ Tiré de la définition de « expérimental ». *Le Grand Robert de la langue française*, 2022

Le monde de la recherche scientifique, et dans notre cas biomédicale et biotechnologique, est semé de conventions, de règles et de structurations précises. Grâce à elles, émanent de la recherche des résultats concrets et applicables. Nous explorerons dans cette dernière partie certaines des spécificités de la pratique des sciences, tant dans la précision de la langue que dans la rigueur des actes.

1. Comprendre et interpréter des résultats scientifiques

Le déroulement de la recherche :

La recherche scientifique peut être définie en reprenant les termes de l'article « Le protocole de recherche » dans la revue Hegel : « une activité entreprise par un chercheur en vue de la mise en évidence d'un fait ou d'un phénomène nouveau⁶¹ ». Elle implique plusieurs étapes, allant de l'idée nouvelle au résultat concret. Dans le cadre de notre exercice de compréhension de ce que constitue un article scientifique, nous pourrions considérer que trois étapes centrales y sont abordées : expérimentation, analyse et interprétation. Premièrement, après avoir établi un protocole de recherche agissant comme plan d'action, le chercheur doit procéder à la collecte de données. Il s'agit ici de la phase d'expérimentation ou d'essai. Dans le cadre de la recherche en thérapie génique, les essais précliniques sont par exemple effectués *in vivo* sur des animaux dont le modèle mutationnel ressemble à celui de l'homme, ou *in vitro* dans des cellules en culture. Certaines approches d'édition du génome pour soigner la dystrophie de Duchenne ont ainsi été observées chez des souris et des chiens. De plus, la durée de la phase d'expérimentation (études à court terme ou à long terme), la taille de la population étudiée, la variation des sujets étudiés (sujet malade et traité, sujet témoin non traité, sujets sauvages, etc.), sont autant de facteurs qui orienteront différemment les résultats obtenus et les conclusions qui en sont tirées.

⁶¹Bossali, F., Ndziessi, G., Paraiso Moussilao, N., et al. Le protocole de recherche : étape indispensable du processus de recherche garantissant la validité des résultats. *HEGEL*, 2015. No 1, pp. 23-28

Une seconde phase consisterait en « la compilation, l'analyse et l'interprétation des résultats⁶² », soit un processus qui implique certaines méthodes d'analyse permettant de donner un sens aux résultats obtenus lors de la phase d'expérimentation, et d'ainsi mesurer l'efficacité de ces expériences. En recherche biomédicale, les analyses sont de toute nature, qu'elle s'effectue à l'échelle somatique ou cellulaire. Dans le cadre de la recherche en édition du génome pour soigner la DMD, nous pouvons par exemple citer les analyses histologiques, c'est-à-dire qui ont trait aux tissus musculaires, mais aussi les tests PCR et le séquençage de l'ADN, qui ont pour but d'observer les effets des expériences, ou l'absence d'effet, sur l'ADN des cellules.

Analyser les tissus musculaires

Dans le cas de la DMD, analyser l'état des tissus musculaires revient à mesurer les effets potentiels des changements géniques sur le fonctionnement du muscle. Par exemple, pour observer le rétablissement ou non de la protéine de la dystrophine (voir partie II.3 « Le cas de la myopathie de Duchenne »), les chercheurs ont recours à des techniques d'immunomarquage* en biologie moléculaire, telles que l'analyse Western blot (un immunotransfert). La Western blot est une méthode qui permet de séparer dans un gel et selon leur taille les protéines que l'on souhaite observer, puis de les transférer vers une membrane où elles sont identifiées grâce à des anticorps spécifiques dirigés contre ces protéines.

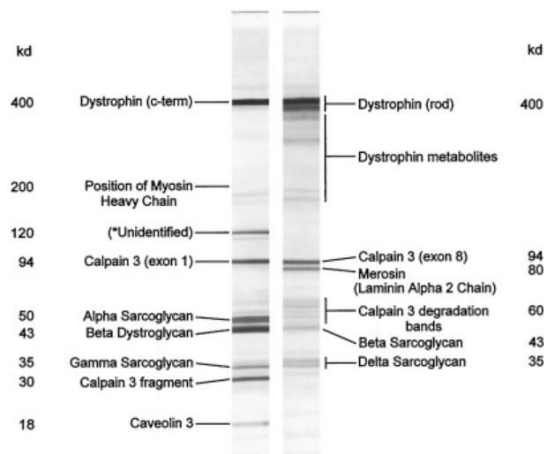
Cette analyse se déroule en plusieurs étapes. Tout d'abord, les échantillons contenant la protéine d'intérêt sont déposés sur un gel. Un courant électrique traversant le gel permet de séparer les protéines selon leur poids moléculaire, car plus la protéine est grande, moins elle migrera vite. On obtient ainsi des bandes homogènes de protéines⁶³ (voir Fig. 17). Les protéines sont ensuite transférées telles quelles sur une membrane, aux propriétés différentes du gel, et qui ressemble à du buvard. Cette migration sur la membrane permettra dans un troisième temps de révéler la position des protéines. En effet, la dernière étape consiste à utiliser des anticorps pour marquer la position des protéines.

⁶² Ibid

⁶³ Western Blot. Article [s.d.]. Ooreka [en ligne]. Disponible sur premiers-secours.ooreka.fr (Consulté en juin 2023)

Une première série d'anticorps spécifiques à la protéine vient se fixer à ces dernières. Une deuxième série d'anticorps est conjuguée à un marqueur fluorescent, radioactif ou enzymatique, et ils reconnaissent les anticorps de la première série⁶⁴. Grâce à l'ajout d'un réactif, une coloration, une image radiographique ou une émission de lumière (selon le marqueur utilisé) est provoquée, et une image de la position des protéines est ainsi créée (voir Fig. 17). La Western blot permet ainsi de déterminer l'abondance d'une protéine dans une cellule.

Figure 17 : bandes protéiques d'une Western blot d'un muscle squelettique humain étiquetée grâce à des anticorps spécifiques aux protéines indiquées. Illustration tirée de l'article « Multiplex Western Blotting System for the Analysis of Muscular Dystrophy Proteins », The American Journal of Pathology, 1999



D'autres méthodes de ce genre peuvent être employées pour cerner des protéines spécifiques dans les cellules d'une coupe tissulaire, telles que l'immunohistochimie et l'immunofluorescence (voir Fig. 18).

⁶⁴ David, L. MOOC côté labo : Technique d'étude des protéines cellulaires : western blot : les principes [VIDEO]. Inserm, 2017

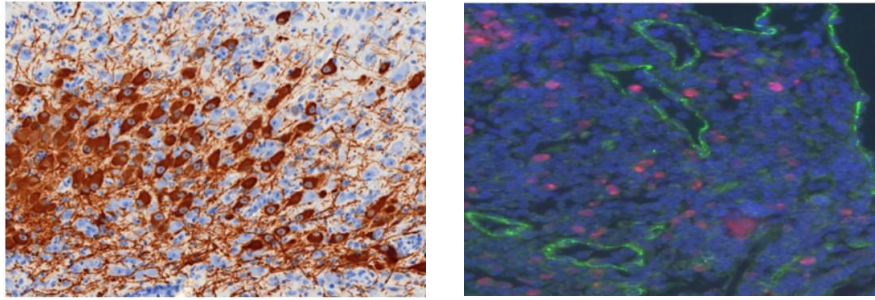


Figure 18 : colorations à l'immunohistochimie (à gauche) et à l'immunofluorescence (à droite) dans le cerveau d'un rat et d'une souris respectivement. Images tirées de l'Université d'Ottawa.

Analyser l'ADN

L'amplification génique par PCR* (pour « Polymerase Chain Reaction » en anglais, soit « amplification en chaîne par polymérase » en français) permet de détecter in vitro certaines régions spécifiques d'un ADN cible dans un produit. Ce processus mime celui, naturel, de synthèse de l'ADN⁶⁵. On obtient ainsi, en quelques heures et à partir d'un petit échantillon, des milliers voire des millions de copies d'un fragment d'ADN cible, sélectionné parmi des milliers d'autres fragments dans l'ADN. L'on comprend alors tout de suite l'intérêt de la PCR pour observer les résultats d'une opération d'édition génique. Si l'on veut observer si une séquence d'ADN a été corrigée in vivo, par exemple si elle a bien été coupée par l'enzyme Cas utilisée dans le cadre de l'utilisation de l'outil CRISPR-Cas9, il faut récupérer un échantillon de cellules et pouvoir sélectionner ce fragment d'ADN spécifique puis le séquencer (voir partie I.3 « Le séquençage du génome »). Les fragments spécifiques répliqués obtenus grâce à la PCR sont appelés « produits d'amplification* »⁶⁶, ou de manière anglicisée, « amplicons* ».

La PCR est une réaction en chaîne, ce qui signifie que ses cycles se répètent en boucle (en moyenne 20 à 40 cycles⁶⁷), et que la synthèse des copies d'ADN est donc exponentielle.

⁶⁵Analyses génétiques par PCR. Article mis à jour en 2021, Eurofins [en ligne]. Disponible sur [eurofins.fr](https://www.eurofins.fr) (Consulté en juin 2023)

⁶⁶Produit d'amplification. Définition mise à jour en 2011. France Terme [en ligne]. Disponible sur [culture.fr/franceterme](https://www.culture.fr/franceterme) (Consulté en juin 2023)

⁶⁷ La PCR (polymerase chain reaction). École Doctorale des Sciences de la Vie et de la Santé ULP, Strasbourg [en ligne]. Disponible sur [ed.vie-sante.unistra.fr](https://www.ed.vie-sante.unistra.fr) (Consulté en juin 2023)

Pour chaque cycle, trois paliers de températures correspondent respectivement à trois étapes de synthèse. Pour la PCR, l'ADN cible est extrait de l'échantillon et placé sur gel, et des amorces* sont utilisées pour reconnaître la séquence spécifique ciblée. Comme son nom l'indique, l'amorce, outil entièrement expérimental, amorce l'amplification d'une séquence. « Ce sont des fragments courts d'ADN, capables de s'hybrider de façon spécifique, grâce à la complémentarité des bases, sur l'un des deux brins d'ADN. Les amorces sont choisies de façon à encadrer la séquence d'ADN à amplifier. La taille de ces amorces est généralement d'une vingtaine de [nucléotides]. De plus, les amorces sont en très forte concentration par rapport à celle de l'ADN à amplifier.⁶⁸ ». Enfin, des nucléotides de synthèse et une enzyme polymérase interviennent pour synthétiser le nouveau brin d'ADN à partir du brin matrice, comme le fait naturellement l'ARN polymérase avec les désoxyribonucléotides présents dans le milieu (voir partie I.2 « Le code génétique, de l'ADN à la protéine »).

Lors du premier palier de température, qui correspond à la phase de dénaturation, l'ADN perd sa forme en double hélice et les deux brins complémentaires sont séparés. La deuxième phase, celle de l'hybridation*, se produit à une température moins élevée et correspond au moment où les amorces repèrent et se fixent à la séquence qui sera amplifiée⁶⁹. Enfin, lors de la phase d'élongation, l'enzyme polymérase utilise les nucléotides présents dans le milieu réactionnel pour synthétiser un fragment de brin complémentaire à celui encadré par les amorces. Lors du cycle suivant, les fragments synthétisés servent de brin matrice aux prochains brins répliqués, et ainsi de suite (voir Fig. 19).

Dans le cas des approches par édition du génome, il est possible, à partir des produits d'amplification, d'étudier la présence et l'abondance de certaines nouvelles insertions ou délétions (indels*) afin d'évaluer l'efficacité de l'outil utilisé.

⁶⁸ Ibid

⁶⁹ Ibid

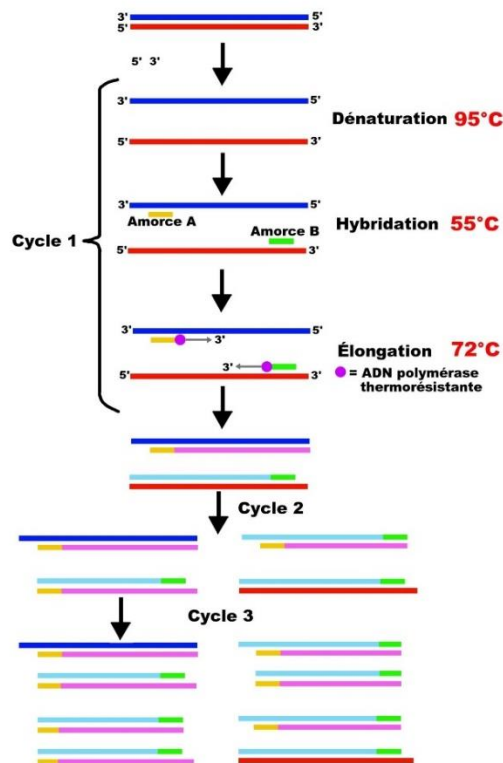


Figure 19 : Chronologie de la PCR.
Illustration de Cobas Amplicor.

Expliquer les résultats

Nous pourrions considérer que la dernière phase de toute recherche scientifique n'est pas seulement celle de l'interprétation des résultats, mais aussi et surtout de leur explication. En effet, une fois les expériences menées et les résultats analysés, il s'agit de rédiger un article sur l'étude menée, reprenant le protocole suivi, les méthodes employées, les essais menés et les résultats obtenus. Ce type d'article implique une rigueur structurelle permettant de comprendre l'aboutissement de la recherche ou d'attester une approche expérimentale. Les sujets étudiés sont mentionnés et dénommés spécifiquement, les durées des expériences, les quantités et les unités de mesures sont précisées, et tous les facteurs internes et externes sont pris en compte. On obtient alors, normalement, un document reprenant tous les codes en usage dans le domaine de recherche, codes qu'il convient de repérer et de comprendre si l'on souhaite procéder à la traduction d'un tel texte.

2. Recherche et applications concrètes

« La biologie moléculaire et les biotechnologies sont des domaines scientifiques relativement récents. Les premières grandes découvertes ne datent que de la fin du XIX^{ème} siècle, et de nombreuses avancées dans ces domaines n'ont été effectuées qu'au cours des dernières années. Entre la découverte de l'ADN au début du XX^{ème} siècle et son séquençage complet en 2004, il s'est écoulé presque un siècle, rempli de nouvelles recherches et inventions. »⁷⁰

Mis à part les enjeux éthiques que cela suppose (voir partie IV.3 « Questions éthiques »), ces avancées rentrent dans le cadre d'un effort scientifique de combiner plusieurs disciplines entre elles afin d'aboutir à des résultats concrets, et ce de manière rapide. Le séquençage du génome humain a ainsi été permis grâce aux avancées en informatique (intelligence artificielle, logiciels de traitement de données de pointe). De même que les perspectives futures de CRISPR-Cas9 se situent à la croisée des différents développements technologiques, tels que l'apprentissage automatique, les technologies de séquençage ou l'imagerie de cellules vivantes (voir fig. 20).

Les avancées majeures sont aussi rendues possibles par les partenariats entre la recherche académique et clinique, mais aussi par les efforts effectués du côté des associations de malades, des entreprises de biotechnologie et des laboratoires pharmaceutiques. À ce même titre, la recherche translationnelle* joue un rôle clef dans les progrès en médecine. La recherche translationnelle (ou recherche de transfert) se situe entre la recherche fondamentale, dont le travail consiste à comprendre les mécanismes à l'origine d'une maladie, et la recherche clinique qui vise à évaluer l'efficacité et la tolérance de nouveaux traitements sur les patients. Ainsi, en thérapie génique, « environ 2 000 essais cliniques ont été menés ou sont en cours depuis 1989, dont 65% dans le domaine du cancer, 11% contre les maladies monogéniques⁷¹ ».

⁷⁰Agay, L. Enjeux politiques, économiques et éthiques de CRISPR/Cas9 en Europe. Thèse en Sciences pharmaceutiques. Grenoble : Université Grenoble Alpes, Faculté de pharmacie, 2019. 112 p.

⁷¹ Thérapie génique - Une recherche de longue haleine qui porte ses fruits. Article mis à jour en 2018. Inserm [en ligne]. Disponible sur [inserm.fr](https://www.inserm.fr) (Consulté en juin 2023)

La plupart de ces essais sont de phase I ou II, c'est-à-dire qu'ils évaluent l'efficacité et la tolérance aux traitements chez quelques centaines de patients. Mais certains essais sont de phase III : ils visent à tester le traitement sur des milliers de patients pour voir si l'efficacité et la tolérance est identiques chez une grande variété de personnes⁷².

Une fois les études cliniques validées, les médicaments (ici les traitements de thérapie génique) peuvent être mis sur le marché. Cependant, le déséquilibre des approbations de mise sur le marché entre les pays témoigne des limites des progrès médicaux. Certains traitements ne sont pas approuvés car les recherches ne sont pas forcément transnationales, mais aussi parce que les perceptions éthiques, politiques, économiques et légales varient d'un État à l'autre.

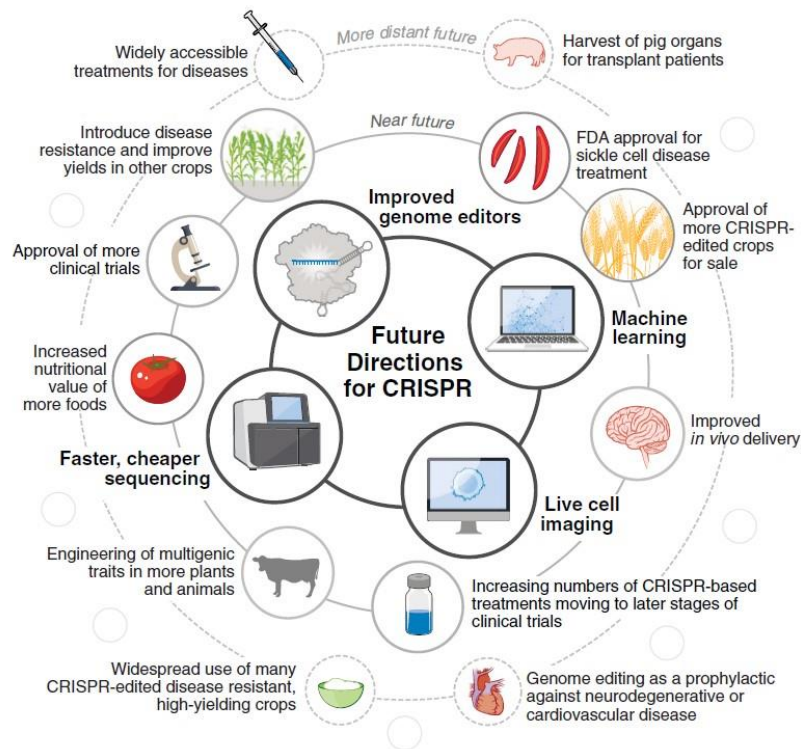


Figure 20 : perspectives futures de la technologie CRISPR dans les 10 années à venir. Tiré de l'article "CRISPR technology: A decade of genome editing is only the beginning", Science, 2023

⁷²=Phase clinique. Article [s.d.]. Interpharma, Bâle [en ligne]. Disponible sur interpharma.ch (Consulté en juin 2023)

3. Questions éthiques

Les questions éthiques que suppose l'intervention humaine sur le génome, et plus particulièrement l'édition du génome telle que définie plus haut, sont à profusion. L'analyse éthique et déontologique de cette discipline va de pair avec sa pratique et la manière dont elle est appréhendée. Parmi les questions globales qui reviennent, on retrouve les suivantes : « Quelle légitimité de modifier la nature pour l'homme ? », « Quelle égalité d'accès aux techniques d'édition du génome ? », « Est-ce que nous pouvons définir la valeur d'un individu par son code génétique ? », « Quelles conséquences pour les générations futures ? », « Comment articuler la notion de consentement avec les techniques d'édition du génome ? », et bien d'autres encore (voir la Cartographie des controverses en annexe 1).

Pour soigner les maladies génétiques, l'intérêt de la thérapie génique par rapport à des méthodes pharmacologique classiques est qu'une fois une mutation précocement identifiée, il est envisageable de faire des réparations au niveau cellulaire pour empêcher que la maladie ne se développe. Le problème est donc abordé à sa racine. Dans le cadre d'une réparation de mutations somatiques, soit sur un corps seulement, nous pourrions considérer que les enjeux éthiques sont faibles car on ne met en péril que l'organisme concerné. Toutefois, à partir du moment où il y a une modification du génome sur les gamètes, nous influençons des générations, ce qui s'avère plus problématique. A-t-on envie de devenir des modificateurs de la loi naturelle des choses ? C'est souvent la limite bioéthique actuelle, et arbitraire, que s'imposent les chercheurs en thérapie génique : guérir la personne mais s'interdire de toucher à l'hérédité.

Une loi bioéthique datant de 1994 interdit en France de pratiquer la thérapie génique germinale* chez l'homme : le gène thérapeutique introduit ou l'édition génomique induite ne peuvent s'opérer au niveau de cellules souches embryonnaires. Mais comme nous l'avons vu, la thérapie génique somatique* est quant à elle appliquée pour soigner certaines maladies génétiques. Ce type de thérapie ne touche pas aux cellules germinales et n'a pas d'influence héréditaire.

« En France, l'édition du génome est encadrée par l'article 16-4 du Code civil qui interdit la « *transformation (...)* des caractères génétiques dans le but de modifier la descendance de la personne » (11). Seules sont acceptées les interventions à des fins médicales, c'est-à-dire les recherches pour la prévention, le diagnostic ou le traitement d'une maladie génétique. »⁷³

Conclusion

Cette première partie nous aura permis de cerner les thématiques et les enjeux principaux du domaine de spécialisation choisi pour ce mémoire. Ces clefs en main, il s'avèrera plus aisé de procéder à une traduction spécialisée. En effet, même si la présentation du sujet reste de l'ordre de la vulgarisation, elle permet toutefois de se situer et d'appréhender le texte-support, que ce soit au niveau de la logique interne du document et des expériences réalisées, ou au niveau de la terminologie et des faits de langues employés.

La thérapie génique se trouve à la croisée de multiples domaines, de la génétique à la médecine, de la biologie à la technologie. La simplicité de son fonctionnement entraîne une complexité due à tous les facteurs qui entrent en jeu, complexité qu'il faut savoir démêler en tant que spectateur de cette discipline. Ce travail de décorticage a été effectué sans connaissances préalables ni expérience du sujet, ce qui a supposé la structure hiérarchique de l'exposé (disons, du plus simple au plus complexe). Le texte-support choisi a en lien avec cette partie qu'on y retrouve mentionnés, et non explicités, la plupart des notions étudiées jusqu'ici. Nous pourrions donc en conclure que la « Partie 1 : Exposé » agit comme un manuel à la compréhension du texte support.

⁷³ Consultation citoyenne sur l'édition du génome en Île-de-France, Faculté de médecine de l'Université Paris-Saclay, 2021, p. 5

Partie 2 : TEXTE-SUPPORT ET SA TRADUCTION

Dans la version cible :

- ❖ *En gras les termes faisant l'objet d'une fiche terminologique*
- ❖ *Apposés à un astérisque les termes vus dans la partie Exposé et/ou présents dans le glossaire ou les lexiques*
- ❖ *Soulignés, les passages justifiés dans la Partie 3 : stratégie de traduction*

Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy

Leonela Amoasii^{1,2}, John C.W. Hildyard³, Hui Li¹, Efrain Sanchez-Ortiz¹, Alex Mireault¹, Daniel Caballero¹, Rachel Harron³, Thaleia-Rengina Stathopoulou⁴, Claire Massey³, John M. Shelton⁵, Rhonda Bassel-Duby¹, Richard J. Piercy³, Eric N. Olson^{1*}

¹Department of Molecular Biology, Hamon Center for Regenerative Science and Medicine, Sen. Paul D. Wellstone Muscular Dystrophy Cooperative Research Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA. ²Exonics Therapeutics, 75 Kneeland Street, Boston, MA 02111, USA. ³Department of Clinical Science and Services, Comparative Neuromuscular Diseases Laboratory, Royal Veterinary College, London NW1 0TU, UK. ⁴Section of Anaesthesia and Analgesia, Royal Veterinary College, London NW1 0TU, UK. ⁵Department of Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA.

*Corresponding author. Email: eric.olson@utsouthwestern.edu

Mutations in the gene encoding dystrophin, a protein that maintains muscle integrity and function, cause Duchenne muscular dystrophy (DMD). The deltaE50-MD dog model of DMD harbors a mutation corresponding to a mutational “hot spot” in the human DMD gene. We used adeno-associated viruses to deliver CRISPR gene editing components to four dogs and examined dystrophin protein expression 6 weeks after intramuscular delivery (n=2) or 8 weeks after systemic delivery (n=2).

L'édition du génome rétablit l'expression de la dystrophine dans un modèle canin de la dystrophie musculaire de Duchenne

(...)

Des mutations* dans le gène* codant la dystrophine, une protéine* qui maintient l'intégrité et la fonction musculaire, causent la dystrophie* musculaire de Duchenne (DMD). Le model canin deltaE50-MD de la DMD est porteur d'une mutation dans la région « hotspot » des mutations du gène *DMD* chez l'humain. Nous avons utilisé des virus adénoassociés* pour délivrer les composants de l'édition génomique* CRISPR chez quatre chiens, et nous avons analysé l'expression de la protéine dystrophine 6 semaines après injection intramusculaire (n = 2) ou 8 semaines après injection systémique* (n = 2).

After systemic delivery in skeletal muscle, dystrophin was restored to levels ranging from 3 to 90% of normal, depending on muscle type. In cardiac muscle, dystrophin levels in the dog receiving the highest dose reached 92% of normal. The treated dogs also showed improved muscle histology. These large animal data support the concept that, with further development, gene editing approaches may prove clinically useful for the treatment of DMD.

The persistent contraction and relaxation of cardiac and skeletal muscles necessitates mechanisms that maintain the integrity of muscle membranes (1, 2). Dystrophin is a large scaffolding protein that supports muscle structure and function by linking the cytoskeleton with the sarcolemma of muscle tissue (1, 3, 4). Mutations in the dystrophin gene cause Duchenne muscular dystrophy (DMD), a disorder primarily affecting boys that is characterized by progressive muscle degeneration and atrophy, leading to premature death from cardiomyopathy and respiratory collapse (5). Thousands of mutations have been identified in the dystrophin gene, which spans ~2.5 megabases of DNA and contains 79 exons. Many of these mutations cluster into “hot spots”, most commonly in a region that spans exons 45 to 50, typically placing exon 51 out of frame with preceding exons and preventing expression of functional dystrophin protein (6, 7). Therapies that induce “skipping” of exon 51 restore the reading frame, and in principle could benefit approximately 13% of DMD patients (8). An oligonucleotide that allows skipping of exon 51 can restore dystrophin expression to 0.22 to 0.32% of normal levels after one year of treatment and has been approved for DMD patients (9–13).

Après distribution systémique au sein des muscles squelettiques, la dystrophine a été rétablie à un niveau allant de 3 à 90 % de l'expression normale, selon le type de muscle* examiné. Dans le muscle cardiaque* du chien ayant reçu la dose la plus importante, le niveau de dystrophine a atteint 92 % de l'expression normale. Les chiens traités ont aussi montré une amélioration de l'histologie* musculaire. Ce large ensemble de données obtenu chez l'animal de grande taille soutient le concept que l'édition du génome, pourvu que son développement soit poursuivi, serait cliniquement pertinente pour traiter la DMD.

La contraction* et le relâchement* continus du muscle cardiaque et des muscles squelettiques requièrent un mécanisme permettant de maintenir l'intégrité des membranes musculaires (1, 2). La dystrophine est une grosse protéine structurelle qui soutient la structure et la fonction musculaire en reliant le cytosquelette* au sarcolemme* du tissu musculaire (1, 3, 4). Des mutations dans le gène de la dystrophine sont responsables de la dystrophie musculaire de Duchenne, une maladie qui atteint principalement les garçons et qui se caractérise par une dégénérescence* progressive des muscles et leur atrophie*, qui conduit à une mort prématurée résultant d'une cardiomyopathie et d'une insuffisance respiratoire* (5).

Des milliers de mutations ont été identifiées dans le gène codant la dystrophine, qui s'étend sur ~2,5 mégabases* de l'ADN génomique* et contient 79 exons*. Une grande partie de ces mutations se regroupent en régions *hotspots*, dont la plus importante s'étend des exons 45 à 50, décalant ainsi l'exon 51 du cadre de lecture* des exons précédents et empêchant l'expression de la protéine dystrophine fonctionnelle (6, 7). Les thérapies qui induisent un « saut » de l'exon 51 permettent de rétablir le cadre de lecture et pourraient, en principe, bénéficier à environ 13 % des patients DMD (8). Un oligonucléotide* permettant le saut de l'exon 51 et capable de rétablir l'expression de la dystrophine à 0,22-0,32 % de son niveau normal après un an de traitement, a été approuvé pour traiter les patients DMD (9-13).

CRISPR/Cas9 gene editing can target DMD mutations and restore dystrophin expression in mice and muscle cells derived from human induced pluripotent stem cells (iPSCs) (14– 22). An essential step toward clinical translation of gene editing as a therapeutic strategy for DMD is the demonstration of efficacy and safety of this approach in large mammals.

The mutation carried by the deltaE50-MD canine model of DMD leads to loss of exon 50 and moreover can be corrected by skipping of exon 51, making this a valuable model for translational studies.

First identified as a naturally occurring, spontaneous mutation in Cavalier King Charles Spaniels (23) and now maintained on a beagle background, this model (in contrast to mice) exhibits many of the clinical and pathological features of the human disease, such as muscle weakness, atrophy and fibrosis (24).

To correct the dystrophin reading frame in the deltaE50-MD canine model (henceforth referred to as Δ Ex50) (Fig. 1A), we used *S. pyogenes* Cas9 coupled with a sgRNA to target a region adjacent to the exon 51 splice acceptor site (referred to as sgRNA-51) (Fig. 1B). The sgRNA-51 corresponded to a highly conserved sequence that differs by only one nucleotide between the human and dog genomes (fig. S1A). Cas9 coupled with each of these sgRNA-51 sequences introduced a genomic cut only in DNA of the respective species (fig. S1B).

For the *in vivo* delivery of Cas9 and sgRNA-51 to skeletal muscle and heart tissue in dogs, we used recombinant adeno- associated virus serotype 9 (referred to as AAV9), which displays preferential tropism for these tissues (25, 26). A muscle- specific creatine kinase (CK) regulatory cassette was used to drive expression of Cas9; three RNA polymerase III promoters (U6, H1 and 7SK) directed expression of the sgRNA, as described previously in mice (fig. S2) (18). AAV9-Cas9 and AAV9-sgRNA-51 were initially introduced into the cranial tibialis muscles of two 1 month-old dogs by intra-muscular (IM) injection with 1.2×10^{13} AAV9 viral genomes (vg) of each virus.

La technologie d'**édition du génome*** CRISPR/Cas9 peut cibler les mutations DMD et rétablir l'expression de la dystrophine chez les souris et dans les cellules musculaires dérivées de cellules souches* pluripotentes induites humaines (iPSC) (14-22). La démonstration de l'efficacité et de l'innocuité de cette approche chez les mammifères de grande taille constitue une étape essentielle vers une application clinique de l'édition du génome en tant que stratégie thérapeutique pour la DMD.

La mutation portée par le modèle canin deltaE50-MD de la DMD provoque la perte de l'exon 50 et peut être corrigée par saut de l'exon 51, ce qui rend ce modèle précieux pour les études translationnelles.

Ce modèle a d'abord été identifié chez l'épagneul Cavalier King Charles comme une mutation naturelle et spontanée et elle est désormais entretenue chez le Beagle. À l'inverse du modèle souris, le modèle canin partage beaucoup des aspects cliniques et pathologiques avec l'humain, telle que la faiblesse musculaire, l'atrophie et la fibrose* (24).

Afin de corriger le cadre de lecture de la dystrophine dans le modèle canin deltaE50-MD (ci-après Δ Ex50) (Fig. 1A), nous avons utilisé des Cas9* issues de *S. pyogenes* et combinées à un ARNsg* pour cibler une région adjacente au site accepteur d'épissage* de l'exon 51 (ci-après ARNsg-51) (Fig. 1B). L'ARNsg-51 correspondait à une séquence hautement conservée qui ne diffère que d'un nucléotide* entre le génome humain et le génome canin (fig. S1A). Associée à chacune de ces séquences d'ARNsg-51, Cas9 a induit une coupure génomique* uniquement dans l'ADN des espèces respectives (fig. S1B).

Pour effectuer la livraison in vivo de Cas9 et de l'ARNsg-51 dans les muscles squelettiques et cardiaque des chiens, nous avons utilisé un virus adénoassocié recombinant de sérotype 9 (ci-après AAV9), qui présente un tropisme préférentiel pour ces tissus (25, 26). Une cassette du promoteur muscle-spécifique du gène codant la créatine-kinase (CK)* a été utilisée pour permettre l'expression de Cas9 ; trois promoteurs* de l'ARN polymérase III (U6, H1 et 7SK) ont dirigé l'expression de l'ARNsg, comme précédemment décrit chez les souris (fig. S2) (18). Les vecteurs AAV9-Cas9 et AAV9-ARNsg-51 ont d'abord été administrés à deux chiots d'1 mois dans le muscle tibial crânial via une injection intramusculaire (IM) de $1,2 \times 10^{13}$ génomes viraux du AAV9 pour chaque vecteur viral.

Muscles were analyzed 6 weeks after injection. In vivo targeting efficiency was estimated within muscle biopsy samples by RT-PCR with primers for sequences in exons 48 and 53, and genomic PCR amplification products spanning the target site were subjected to amplicon deep-sequencing. The latter indicated that a mean of 9.96% of total reads contained changes at the targeted genomic site including insertions, deletions and substitutions (fig. S3). The most commonly identified mutations with a mean of 2.35% contained an adenosine (A) insertion immediately 3' to the Cas9 genomic cutting site (Fig. 1C). The deletions identified using this method encompassed a highly-predicted exonic splicing enhancer (ESE) site for exon 51 (18, 27, 28) (fig. S3A). However, this method does not identify larger deletions that might occur beyond the annealing sites of the primers used for PCR. Since these tissue samples contain a mixture of muscle and non-muscle cells, the method likely underestimates the efficiency of gene editing within muscle cells.

Sequencing of RT-PCR products of RNA from muscle of Δ Ex50 dogs injected intramuscularly with AAV9-Cas9 and AAV9-sgRNA-51 showed that deletion of exon 51 (Δ Ex50-51) allowed splicing from exon 49 to 52, which restores the dystrophin open reading frame (fig. S3B). On gels, the PCR product with the A insertion was indistinguishable in size from non-edited cDNA products, so we performed deep-sequencing analysis to quantify its abundance compared to other small insertions. Deep-sequencing of the upper band containing the non-edited cDNA product and reframed cDNA products indicated that a mean of 73.19% of total reads contained reframed cDNA products with an A insertion, a mean of 26.81% contained non-edited cDNA product, and the rest contained small deletions and insertions (fig. S3C).

Les muscles ont été analysés 6 semaines après injection. L'efficacité du ciblage in vivo a été estimée dans des échantillons* issus d'une biopsie musculaire par analyse RT-PCR* avec des **amorces*** spécifiques aux séquences dans les exons 48 et 53. Les produits d'amplification* de la PCR génomique recouvrant le site ciblé ont été soumis à un séquençage profond*. Ce dernier a indiqué qu'en moyenne 9,96 % des lectures contenaient des changements au niveau du site génomique ciblé, y compris des insertions*, des délétions* et des substitutions* (fig. S3). Parmi les mutations identifiées, les plus courantes, avec une moyenne de 2,35 %, comprenaient une insertion d'adénosine (A) en 3' du site de découpage* génomique de Cas9 (Fig. 1C). Les délétions identifiées grâce à cette méthode recouvraient une séquence prédite avec un haut niveau de confiance comme activatrice d'épissage exonique (AEE)* pour l'exon 51 (18, 27, 28) (fig. S3A). Cependant, cette méthode ne permet pas d'identifier des délétions plus importantes pouvant avoir lieu au-delà des sites d'**hybridation*** des amorces utilisées pour la PCR*. Étant donné que ces échantillons de tissu contiennent à la fois des cellules musculaires et non musculaires, il est probable que la méthode employée sous-estime l'efficacité de l'édition du génome au sein de cellules musculaires.

Le séquençage* des produits de RT-PCR de l'ARN provenant des muscles de chiens Δ Ex50 ayant reçu une injection intramusculaire des vecteurs AAV9-Cas9 et AAV9-ARNsg-51 a montré qu'une délétion de l'exon 51 (Δ Ex50-51) permettait un épissage de l'exon 49 à l'exon 52, rétablissant ainsi le cadre ouvert de lecture de la dystrophine (fig. S3B). Sur gel, la taille du produit PCR contenant une insertion d'A était indissociable des produits d'ADNc* non édités. Nous avons donc effectué un séquençage profond pour quantifier son abondance comparée aux autres micro-insertions. Le séquençage profond de la bande supérieure contenant le produit d'ADNc non édité et les produits d'ADNc avec un cadre de lecture rétabli a montré que parmi toutes les lectures, en moyenne 73,19 % contenaient des produits d'ADNc avec un cadre de lecture rétabli suite à l'insertion d'un A, 26,81 % contenaient des produits d'ADNc non édités, et le reste contenait des courtes délétions et insertions (fig. S3C).

However, nonsense mediated decay might impact the abundance of non-edited cDNA products. These data indicate that the two Δ Ex50 dogs injected with AAV9-Cas9 and AAV9-sgRNA-51 had a high frequency of reframing events (with cDNA products containing an A insertion in the sequence of exon 51) and exon 51 skipping events resulting from deletion of the highly conserved ESE region.

To evaluate the specificity of our gene editing approach, we analyzed predicted off-target genomic sites for possible promiscuous editing. A total of three potential genome-wide off-target sites (OT1 to OT3) (fig. S4) were predicted in coding exons and 4 in non-coding regions (fig. S4) by the CRISPR design tool (<http://crispr.mit.edu/>). We performed deep sequencing at the top predicted off-target sites within protein coding exons. None of these sites revealed significantly more sequence alterations than the background analysis performed with other regions of the amplicons (fig. S5).

To evaluate dystrophin correction at the protein level, we performed histological analysis of AAV9-injected cranial tibialis muscles 6 weeks after AAV9 injection. Dystrophin immunohistochemistry of muscle from Δ Ex50 dogs injected with AAV9-Cas9 and AAV-sgRNA-51 revealed widespread expression: the majority of fibers within the injected muscles expressed sarcolemmal dystrophin, albeit to varying levels (Fig. 1D). Western blot analysis confirmed the restoration of dystrophin expression in skeletal muscle (Fig. 2, A and B) to ~60% of wild-type levels. On average, 2% of wild-type levels of dystrophin were detected in the uninjected contralateral muscles, far more than could be attributed to rare revertant events, which typically represent fewer than 0.001% of fibers and are undetectable by Western blot in Δ Ex50 muscle. We attribute expression in uninjected contralateral muscles to leakage of AAV9 into circulation.

Cependant, la dégradation des ARNm médiée par les brins non-sens* pourrait réduire l'abondance de produits d'ADNc non édités. Ces données montrent que les deux chiens Δ Ex50 ayant reçu une injection des vecteurs AAV9-Cas9 et AAV9-ARNsg-51 présentaient une fréquence élevée de rétablissement du cadre de lecture* (pour les produits d'ADNc contenant une insertion d'A au niveau de la séquence de l'exon 51), et de sauts de l'exon 51 résultants de la délétion du site AEE hautement conservé.

Afin d'évaluer la spécificité de notre méthode d'édition du génome, nous avons analysé des sites génomiques hors cible* pour lesquels nous avons prédit des éditions possibles par promiscuité. Le logiciel d'identification des séquences cibles CRISPR* (<http://crispr.mit.edu/>) a prédit à l'échelle du génome un total de 3 sites potentiels hors cible (OT1 à OT3) dans les exons codants, et 4 dans les régions non codantes (fig. S4). Nous avons effectué un séquençage profond des sites hors cibles les plus prévisibles au niveau des exons codant des protéines. Aucun de ces sites ne présentait d'altérations de la séquence en nombre significativement plus élevé que dans d'autres régions des amplicons* sujettes à des analyses de fond (fig. S5)

Afin d'évaluer la correction de la quantité de la protéine dystrophine, nous avons effectué une analyse histologique du muscle tibial crânial 6 semaines après injection des AAV9. L'immunohistochimie de la dystrophine du muscle chez les chiens Δ Ex50 ayant reçu une injection des vecteurs AAV9-Cas9 et AAV-ARNsg-51 a révélé une expression répandue de la dystrophine : la majorité des fibres musculaires ayant reçu une injection exprimait une dystrophine au niveau du sarcolemme, bien qu'à des niveaux variables (fig. 1D). L'analyse en Western blot a permis de confirmer le rétablissement de l'expression de la dystrophine dans les muscles squelettiques (fig. 2, A et B) à ~60 % du niveau des individus sauvages*. En moyenne 2 % du niveau de dystrophine des individus sauvages ont été détectés dans les muscles controlatéraux non injectés, un chiffre bien trop élevé pour pouvoir l'attribuer à des effets révertants* rares, qui représentent habituellement moins de 0,001 % des fibres et ne sont pas détectables par Western blot dans le muscle Δ Ex50. Nous avons attribué les raisons d'une expression de dystrophine dans les muscles controlatéraux n'ayant pas reçu d'injection à une fuite de l'AAV9 dans la circulation sanguine.

As assessed by H&E staining, the injected muscles appeared to be normalized relative to muscles of untreated animals, with fewer hypercontracted or necrotic fibers, reduced edema and fibrosis, and fewer regions of inflammatory cellular infiltration (Fig. 2C). Immunohistochemistry for developmental myosin heavy chain (dMHC), a marker of regenerating fibers, revealed a marked reduction in developmental myosin (dMHC)-positive fibers within injected muscles (fig. S6).

Dystrophin nucleates a series of proteins into the dystrophin-associated glycoprotein complex (DGC) to link the cytoskeleton and extracellular matrix (3, 4). In Δ Ex50 mice, dogs, and DMD patients, these proteins are destabilized and do not localize to the sub-sarcolemmal region (4). Muscles injected with AAV9-Cas9 and AAV9-sgRNA-51 showed recovery of the DGC protein beta-dystroglycan compared to contralateral un-injected muscles (fig. S7). We conclude that – at least in a short time frame of 6 weeks—single-cut genomic editing using AAV9-Cas9 and AAV9-sgRNA-51 can efficiently restore dystrophin expression and assembly of the DGC in dystrophic muscles. Immunohistochemistry using canine-specific CD4 and CD8 T cell markers (fig. S8), showed no evidence of enhanced mononuclear cellular infiltration or relevant hematological abnormalities in the treated animals compared to untreated controls or reference ranges (fig. S9).

Based on the high dystrophin-correction efficiency observed following IM injection of AAV9-Cas9 and AAV9-sgRNA-51, we tested for rescue of dystrophin expression in two Δ Ex50 dogs after systemic delivery of gene editing components. The dogs at 1 month of age were injected intravenously with the viruses and analyzed 8 weeks later. We tested two doses (2×10^{13} vg/kg and 1×10^{14} vg/kg) of each of the two viruses (AAV9-Cas9 and AAV9-sgRNA-51). To avoid a possible immune reaction, we included a transient regimen of immune suppression with the high dose.

Évalués par la coloration à l'hématoxyline et l'éosine (H&E)*, les muscles injectés semblaient normalisés comparés aux muscles d'animaux non traités, avec moins de fibres hypercontractées ou nécrosées, une réduction de l'œdème* et de la fibrose, ainsi qu'une plus faible quantité de régions avec infiltration de cellules inflammatoires (fig. 2C). L'immunohistochimie révélant la chaîne lourde de myosine* développementale (dMHC), un marqueur des fibres en régénération, a montré une réduction marquée des fibres positives en dMHC dans les muscles ayant reçu une injection (fig. S6).

La dystrophine agrège une série de protéines dans le complexe glycoprotéique associé à la dystrophine (CGD)* afin de lier le cytosquelette à la matrice extracellulaire (3, 4). Chez les souris et les chiens ΔEx50 ainsi que les patients DMD, ces protéines sont déstabilisées et n'arrivent pas à se localiser à la région sous-sarcolemmique (4). Les muscles ayant reçu une injection des vecteurs AAV9-Cas9 et AAV9-ARNsg-51 affichaient un rétablissement de la bêta-dystroglycane* du CGD comparé aux muscles controlatéraux n'ayant pas reçu d'injection (fig. S7). Nous en concluons qu'une édition génomique à coupure unique* combinant l'usage des vecteurs AAV9-Cas9 et AAV9-ARNsg-51 peut rétablir efficacement l'expression de la dystrophine et la cohésion du CGD dans les muscles dystrophiques, au moins sur une période courte de 6 semaines. L'immunohistochimie utilisant les marqueurs* CD4 et CD8 spécifiques des lymphocytes T canins (fig. S8), n'a montré ni augmentation de l'infiltration cellulaire* mononucléaire ni anomalies hématologiques pertinentes chez les animaux traités comparés aux témoins non traités ou aux valeurs de référence (fig. S9).

Étant donné l'efficacité prononcée de la correction de la dystrophine observée après l'injection IM des vecteurs AAV9-Cas9 et AAV9-ARNsg-51, nous avons évalué le rétablissement de l'expression de la dystrophine chez deux chiens ΔEx50 après injection systémique des vecteurs d'édition génomique. Les chiots de 1 mois ont reçu une injection intraveineuse des vecteurs viraux, et ont été analysés 8 semaines plus tard. Nous avons testé deux doses (2×10^{13} vecteur/kg et 1×10^{14} vecteur/kg) pour chacun des deux vecteurs (AAV9-Cas9 et AAV9-ARNsg-51). Afin d'éviter une potentielle réponse immune, nous avons combiné la dose supérieure à un traitement immunosuppresseur transitoire.

Systemic delivery of 2×10^{13} vg/kg of each virus (total virus 4×10^{13} vg/kg) in Δ Ex50-Dog-#2A resulted in expression of virus in peripheral skeletal muscle samples, and to a lower extent in heart samples, as shown by qPCR analysis (fig. S10A). The delivery of 1×10^{14} vg/kg of each virus (total virus 2×10^{14} vg/kg) in Δ Ex50-Dog-#2B (via infusion) allowed more widespread expression of viral constructs in the peripheral skeletal muscle samples and in heart samples (fig. S10B). Systemic delivery of AAV9-Cas9 and AAV9-sgRNA-51 led to dystrophin expression in a broad range of muscles, including the heart, in gene-edited Δ Ex50 dogs at 8 weeks post-injection, and to a markedly greater extent than that achieved with the lower dose (Fig. 3).

To investigate the proportions of various indels generated by systemic delivery of AAV9-Cas9 and AAV9-sgRNA-51, we performed amplicon deep-sequencing analysis of the genomic DNA from heart, triceps and biceps muscles. The genomic deep-sequencing analysis revealed an increase of percentage of reads containing changes at the targeted genomic site, especially of the 1A insertion mutation in the samples from Dog-#2B compared to Dog-#2A (fig. S11).

Additionally, we performed tracking indels by decomposition (TIDE) (30) analysis at the genomic and cDNA levels, which showed an increase in numbers of indels in the samples from Dog-#2B compared to the samples from Dog-#2A (fig. S12). Testes analysis and Western blot analysis showed no activity of Cas9 and confirmed muscle specific expression of gene editing machinery (fig. S13). Western blot analysis confirmed the restoration of dystrophin expression in skeletal muscle (Fig. 4, A and B) to levels $\sim 50\%$, 20% and 3% of wild-type levels for the cranial tibialis, triceps, and biceps, respectively, after systemic delivery of 2×10^{13} vg/kg of each virus (total virus 4×10^{13} vg/kg). For Dog-#2B, which received 1×10^{14} vg/kg of each virus (total virus 2×10^{14} vg/kg), Western blot analysis showed restoration of dystrophin expression (Fig. 4, C, D and fig. S14) to levels $\sim 70\%$, 25% , 64% , 58% , 92% and 5% of wild-type levels for the cranial tibialis, triceps, biceps, diaphragm, heart and tongue muscles, respectively.

L'injection systémique de 2×10^{13} vecteur/kg de chaque virus (au total 4×10^{13} vecteur/kg de virus) chez le chien- Δ Ex50-#2A a conduit à l'expression des virus dans les échantillons issus de muscles squelettiques* périphériques et, dans une moindre mesure, dans les échantillons du cœur, comme le montre l'analyse qPCR* (fig. S10A). L'injection de 1×10^{14} vecteur/kg de chaque virus (2×10^{14} vecteur/kg au total) chez le chien- Δ Ex50-#2B (via administration intraveineuse) a permis une expression plus répandue des vecteurs viraux dans les échantillons issus de muscles squelettiques périphériques et les échantillons du cœur (fig. S10B). L'administration systémique des vecteurs AAV9-Cas9 et AAV9-ARNsg-51 a rétabli l'expression de la dystrophine dans une grande variété de muscles, dont le cœur, 8 semaines après injection chez les chiens Δ Ex50 génétiquement édités, et ce de manière bien plus prononcée avec la dose supérieure qu'avec la dose la plus faible administrée (fig. 3)

Pour évaluer la proportion des différents indels* générés par l'injection systémique des vecteurs AAV9-Cas9 et AAV9-ARNsg-51, nous avons analysé les amplicons de l'ADN génomique du cœur, des triceps et des biceps par séquençage profond. Cette analyse a démontré une augmentation du pourcentage de lectures comprenant des changements au niveau des sites génomiques ciblés, et surtout de mutations caractérisées par l'insertion 1A dans les échantillons provenant du chien-#2B comparé au chien-#2A (fig. S11).

De plus, nous avons recherché les indels par une analyse de décomposition (TIDE) (30) au niveau du génome et de l'ADNc, qui a démontré une augmentation du nombre d'indels dans les échantillons du chien-#2B comparés à ceux du chien-#2A (fig. S12). L'analyse au niveau des testicules et Western blot ont relevé l'absence d'activité de la Cas9, confirmant ainsi l'expression spécifique au muscle de cet outil d'édition génomique (fig. S13). Le Western blot a confirmé le rétablissement de l'expression de la dystrophine du muscle squelettique (fig. 4, A et B) à ~50 %, 20 % et 3 % des niveaux des individus sauvages dans le muscle tibial crânial, le triceps et le biceps respectivement, après une injection systémique de 2×10^{13} vecteur/kg de chaque vecteur viral (4×10^{13} vecteur/kg au total). Chez le chien-#2B ayant reçu 1×10^{14} vecteur/kg de chaque virus (2×10^{14} vecteur/kg au total), le Western blot a montré le rétablissement de l'expression de la dystrophine (fig. 4, C et D et fig. S14) à ~70 %, 25 %, 64 %, 58 %, 92 % et 5 % des niveaux des individus sauvages dans le muscle tibial crânial, le triceps, le biceps, le diaphragme, le cœur et les muscles de la langue respectivement.

Similar to what was seen after IM injection, muscles appeared normalized via H&E staining (Fig. 4E). Immunostaining of muscle sections from treated Δ Ex50 dogs also showed recovery of beta-dystroglycan expression (fig. S15) and widespread reduction in dMHC, a marker of muscle regeneration (fig. S16).

To determine hematological and biochemical parameters of the treated dogs compared to the controls, we collected blood samples the day before injection and then at 1, 2, 4, 6 and 8 weeks post-injection. The blood samples from all 4 dogs (healthy untreated, Δ Ex50 untreated, and Δ Ex50 dogs receiving 2×10^{13} vg/kg of each virus (total virus 4×10^{13} vg/kg) and 1×10^{14} vg/kg of each virus (total virus 2×10^{14} vg/kg) before and after injection were unremarkable (fig. S17). Hematology counts, serum electrolytes and kidney/liver function parameters remained within the normal ranges in all dogs. Additionally, blood samples were collected weekly for CK assessment: we observed a modest decline in serum CK activity in Dog-#2B treated with 1×10^{14} vg/kg of each virus (total virus 2×10^{14} vg/kg) (fig. S18) compared to Δ Ex50 untreated dog.

To evaluate the targeting efficiency of a human DMD mutation, we used a DMD iPSC line carrying a deletion from exon 48 to 50. Deletion of exons 48 to 50 leads to a frameshift mutation and appearance of a premature stop codon in exon 51. To correct the dystrophin reading frame, we introduced two concentrations of Cas9 and sgRNA-51 (26ng/ μ l, referred to as high, and 13ng/ μ l, referred to as low). Indel analysis showed 55.8% and 31.9% of indels for the high and low concentrations, respectively (fig. S19). Genomic deep-sequencing analysis revealed that 27.94% of mutations contained a single A insertion 3' to the PAM sequence for the high concentration condition and 19.03% for the low concentration condition (fig. S20A), as observed in mouse and dog cells with a similar sgRNA directed against same genomic locus.

De même qu'observé après injection IM, les muscles semblaient normalisés sur la coloration H&E (fig. 4E). L'immunomarquage* des sections musculaires chez les chiens ΔEx50 traités montrait également un rétablissement de l'expression de la bêta-dystroglycane (fig. S15) ainsi qu'une réduction généralisée de la dMHC (fig. S16).

Pour quantifier les valeurs des paramètres hématologiques et biochimiques des chiens traités comparés aux témoins, nous avons réalisé des prélèvements sanguins le jour précédant l'injection, puis 1, 2, 4, 6 et 8 semaines après injection. Les échantillons sanguins des 4 chiens (sain non traité, ΔEx50 non traité, ΔEx50 recevant 2×10^{13} vecteur/kg de chaque vecteur viral (4×10^{13} vecteur/kg au total) et ΔEx50 recevant 1×10^{14} vecteur/kg de chaque vecteur viral (2×10^{14} vecteur/kg au total)) avant et après injection étaient normaux (fig. S17). La numération formule*, les électrolytes sériques et les biomarqueurs des fonctions hépatiques et rénale sont restés normaux chez tous les chiens. Des échantillons sanguins ont aussi été prélevés chaque semaine pour une analyse des concentrations sanguines de CK. Nous avons ainsi pu observer une légère baisse de l'activité sérique des CK chez le chien-#2B recevant 1×10^{14} vecteur/kg de chaque vecteur viral (2×10^{14} vecteur/kg au total) (fig. S18) comparé au chien ΔEx50 non traité.

Pour évaluer l'efficacité du ciblage d'une mutation DMD chez l'humain, nous avons utilisé une lignée d'iPSC de DMD portant une délétion des exons 48 à 50. Une telle délétion entraîne une modification du cadre de lecture et l'apparition d'un codon stop* prématuré dans l'exon 51. Pour corriger le cadre de lecture de la dystrophine, nous avons injecté deux concentrations de Cas9 et d'ARNsg-51 (26 ng/μl, la concentration supérieure, et 13 ng/μl, la concentration inférieure). L'analyse des indels affichait 55,8 % et 31,9 % d'indels pour les concentrations supérieure et inférieure respectivement (fig. S19). Le séquençage profond du génome a montré que 27,94 % des mutations contenaient une insertion unique d'un A en 3' du motif PAM* dans le cas de la concentration supérieure, contre 19,03 % dans le cas de la concentration inférieure (fig. S20A), tel qu'observé dans les cellules de souris et de chien pour lesquelles un ARNsg similaire était dirigé vers le même locus* génomique.

DMD iPSCs treated with Cas9 and sgRNA-51 and induced to form cardiomyocytes (iCMs) showed restoration of dystrophin immunostaining (fig. S20B) and expression of dystrophin protein to levels that were 67 - 100% of the levels of WT cardiomyocytes, as measured by Western blot (fig. S20C and D).

It has been estimated that even 15% of normal levels of dystrophin would provide significant therapeutic benefits for DMD patients (30–32). Our results demonstrate the efficacy of single cut genome editing for restoration of dystrophin expression in a large animal model of DMD, reaching up to ~80% of wild-type levels in some muscles after 8 weeks. Longer-term studies are required to establish whether the expression of dystrophin and the maintenance of muscle integrity we observed are sustained.

This study, while encouraging, is preliminary and has several limitations, including the relatively small number of animals analyzed and the short duration of the analysis (6-8 weeks). The possibility of off-target effects of *in vivo* gene editing is a further potential safety concern. While our initial deep sequencing analysis of the top predicted off-target sites revealed no specific off-target gene editing above background levels in treated animals, it will be important to further assess possible off-target mutagenesis in longer term studies with greater numbers of animals. Recent studies reported large deletions and complex genomic rearrangements at target sites of CRISPR-Cas9 in mouse embryonic cells, hematopoietic progenitors and human immortalized epithelial cells (33). However, these cells are highly proliferative, more genomically unstable in culture and use different DNA repair pathways than somatic post-mitotic cells (such as muscle and heart cells) (34).

Les iPSC de DMD traitées par Cas9 et l'ARNsg-51, puis induites pour former des cardiomyocytes* (iCM), ont montré un rétablissement de l'immunomarquage de la dystrophine (fig. S20B) et de l'expression de la dystrophine allant de 67 % à 100 % du niveau dans les cardiomyocytes sauvages, tel que mesuré grâce au Western blot (fig. S20C et D).

Il a été estimé que même 15 % du niveau normal de dystrophine auraient des bienfaits thérapeutiques pour les patients DMD (30-32). Nos résultats démontrent l'efficacité d'une édition génomique à coupure unique afin de rétablir l'expression de la dystrophine dans un modèle DMD de grand animal, expression qui atteint les ~80 % du niveau des individus sauvages dans certains muscles après 8 semaines. Il faudrait procéder à des études à plus long terme afin de déterminer la durabilité de l'expression de la dystrophine et de l'intégrité musculaire* observées.

Bien qu'encourageante, cette étude n'est que préliminaire et comporte plusieurs limites, notamment la quantité relativement faible d'animaux analysés et la durée réduite (6-8 semaines) de l'analyse. La présence de potentiels effets hors cible liés à une édition génique in vivo peut aussi constituer un autre problème de sécurité. Bien que notre première analyse de séquençage profond des principaux sites hors cible prédits n'ait pas révélé d'édition génique spécifique hors cible qui dépasse un niveau normal chez les animaux traités, une évaluation plus détaillée de la possibilité d'une mutagenèse* hors cible restera nécessaire dans les études à plus long terme et avec une plus grande population étudiée. Des études récentes ont rapporté d'importantes délétions et des réarrangements génomiques* complexes au niveau des sites ciblés par CRISPR-Cas9 dans des cellules embryonnaires* de souris, des cellules hématopoïétiques et des cellules humaines épithéliales (33). Cependant, il s'agit de cellules très proliférantes, génomiquement plus instables en culture, et qui n'utilisent pas les mêmes voies de réparation de l'ADN que les cellules somatiques post-mitotiques* (comme les cellules musculaires et cardiaques) (34).

Future studies will be required to investigate the long-term genomic stability of gene-edited muscle tissues in vivo. Another potential concern with CRISPR/Cas9- mediated gene editing in vivo is immunogenicity of Cas9, particularly as expression persists post initial treatment. In this short-term study, we did not observe CD4- or CD8- positive cell infiltration of treated muscle, but longer-term studies with more sensitive assays must be performed. Additionally, although production of large quantities of AAV9 poses a challenge, doses of 2×10^{14} vg/kg have been successfully used in human gene therapy trials (35).

Finally, although gene editing and exon skipping oligos can both restore production of internally deleted dystrophin proteins, similar to those expressed in Becker muscular dystrophy, there are key distinctions between these approaches. Most notably, CRISPR gene editing may be permanent and not require re-delivery whereas oligos require continuous treatment. A corollary of this is that CRISPR treatment may be difficult to terminate if safety concerns arise.

Il faudra effectuer des études plus poussées pour examiner sur le long-terme la stabilité génomique des tissus musculaires édités in vivo. L'édition génique in vivo par CRISPR/Cas9 pose aussi le potentiel problème de l'immunogénicité* de Cas9, notamment lorsque son expression persiste au-delà du premier traitement. Dans cette étude à court terme, nous n'avons pas observé l'infiltration de cellules CD4- ou CD8-positives dans les muscles traités. Des études à plus long terme et des essais plus poussés sont tout de même nécessaires. De plus, même si une production importante de virus AAV9 constitue un défi, des doses de 2×10^{14} vecteur/kg ont été employées avec succès lors d'essais en thérapie génique* chez l'humain (35).

Enfin, même si l'édition génique et les oligonucléotides induisant un **saut d'exon*** peuvent toutes les deux rétablir in vivo la production de protéines de la dystrophine telle qu'elle s'exprime dans la dystrophie musculaire de Becker*, ces approches restent distinctes. La différence la plus notable est que l'édition génique via CRISPR pourrait être permanente, et ce sans qu'une nouvelle administration ne soit nécessaire, tandis que les oligonucléotides requièrent un traitement continu. Le contrepoint étant que l'arrêt du traitement par CRISPR serait difficile en cas d'effets indésirables sérieux.

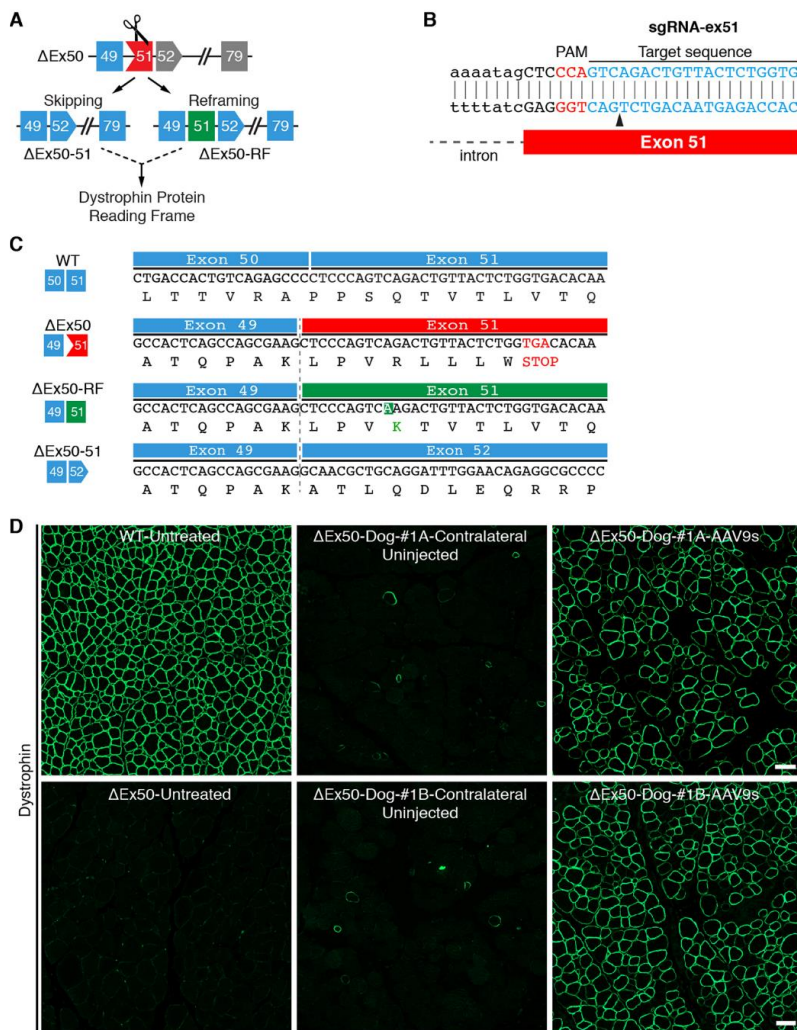


Fig. 1. Single cut CRISPR editing of canine exon 50 in vivo and in vitro. (A) Scheme showing the CRISPR/Cas9-mediated genome editing approach to correct the reading frame in Δ Ex50 dogs by reframing and skipping of exon 51. Gray exons are out of frame. (B) Illustration of sgRNA binding position and sequence for sgRNA-ex51. PAM sequence for sgRNA is indicated in red. Black arrow indicates the cleavage site. (C) Sequence of the RT-PCR products of the Δ Ex50-51 lower band confirmed that exon 49 spliced directly to exon 52, excluding exon 51. Sequence of RT-PCR products of Δ Ex50 reframed (Δ Ex50-RF). (D) Cranial tibialis muscles of Δ Ex50 dogs were injected with AAV9s encoding sgRNA-51 and Cas9 as schematized in Fig. 1 and analyzed 6 weeks later. Dystrophin immunohistochemistry staining of cranial tibialis muscle of wild type dog untreated, Δ Ex50 dog untreated, Δ Ex50 dogs contralateral (uninjected)

muscle and Δ Ex50 dogs injected with AAV9-Cas9 and AAV9-sgRNA-51 (referred as Δ Ex50-#1A-AAV9s and Δ Ex50-#1B-AAV9s). Scale bar: 50 μ m.

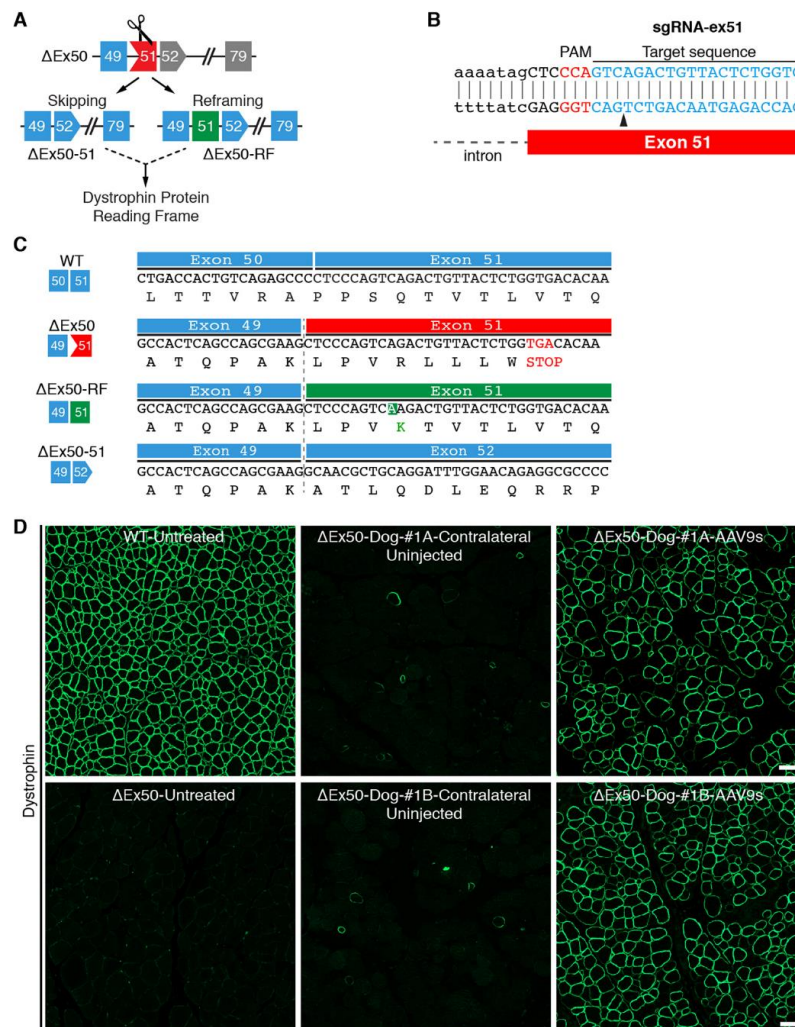


Fig. 1. Edition à cassure unique de l'exon 50 du chien in vivo et in vitro via CRISPR. (A) Schéma détaillant l'approche de l'édition génomique par CRISPR/Cas9 pour corriger le cadre de lecture chez les chiens Δ Ex50 en modifiant le cadre et en induisant un saut de l'exon 51. Les exons gris sont hors cadre. (B) Illustration de la position de liaison de l'ARNsg et de la séquence pour l'ARNsg-ex51. La séquence PAM pour l'ARNsg est signalée en rouge. La flèche noire indique le site de coupure*. (C) La séquence des produits de RT-PCR du brin inférieur Δ Ex50-51 a confirmé que l'exon 49 s'épissait directement avec l'exon 52, excluant l'exon 51. Séquence des

produits de RT-PCR du Δ Ex50 dont le cadre de lecture a été rétabli (Δ Ex50-RF). (D) Des chiens Δ Ex50 ont reçu une injection d'AAV9s permettant l'expression de l'ARNsg-51 et de la Cas9 dans le muscle tibial crânial, comme schématisé dans la Fig. 1, et une analyse a été effectuée 6 semaines plus tard. Coloration par immunohistochimie de la dystrophine du muscle tibial crânial de chien sauvage non-traité, de chien Δ Ex50 non traité, du muscle contralatéral de chiens Δ Ex50 (non-traités) et de chiens Δ Ex50 ayant reçu une injection des vecteurs AAV9-Cas9 et de AAV9-ARNsg-51 (AAV9- Δ Ex50-#1A et AAV9- Δ Ex50-#1B). Échelle : 50 μ m.

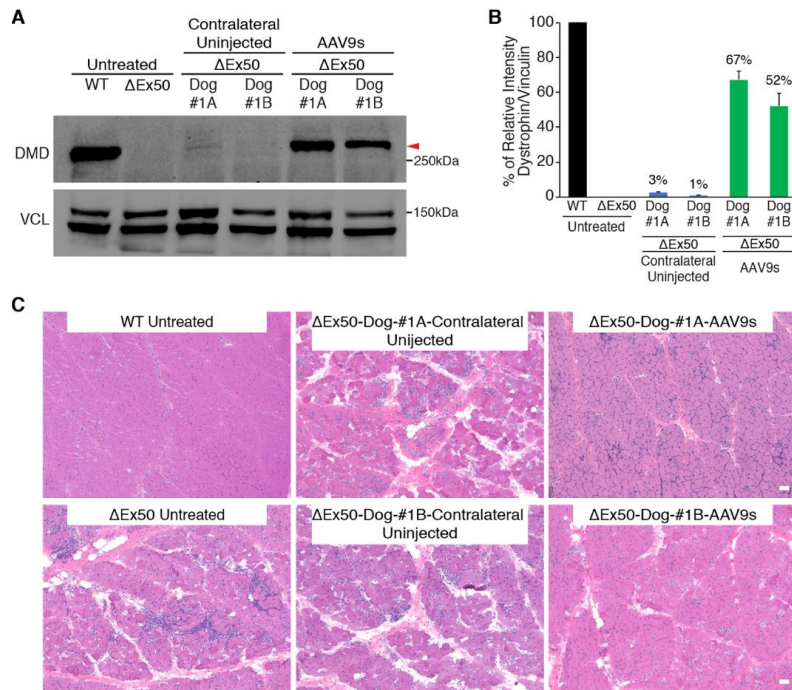


Fig. 2. Dystrophin correction following intramuscular delivery of AAV9-encoded gene editing components. (A) Western blot analysis of dystrophin (DMD) and vinculin (VCL) expression in cranial tibialis muscles 6 weeks after intramuscular injection in 2 dogs (#1A and #1B). (B) Quantification of dystrophin expression from blots after normalization to vinculin. (C) Histochemistry by hematoxylin and eosin (H&E) staining of cranial tibialis muscle from a wild type dog, Δ Ex50 dog untreated, Δ Ex50 contralateral uninjected and Δ Ex50 dogs injected intramuscularly with AAV9-Cas9 and AAV9-sgRNA-51 (referred as Δ Ex50-Dog-#1A- AAV9s and Δ Ex50-Dog-#1B- AAV9s). Scale bar: 50 μ m.

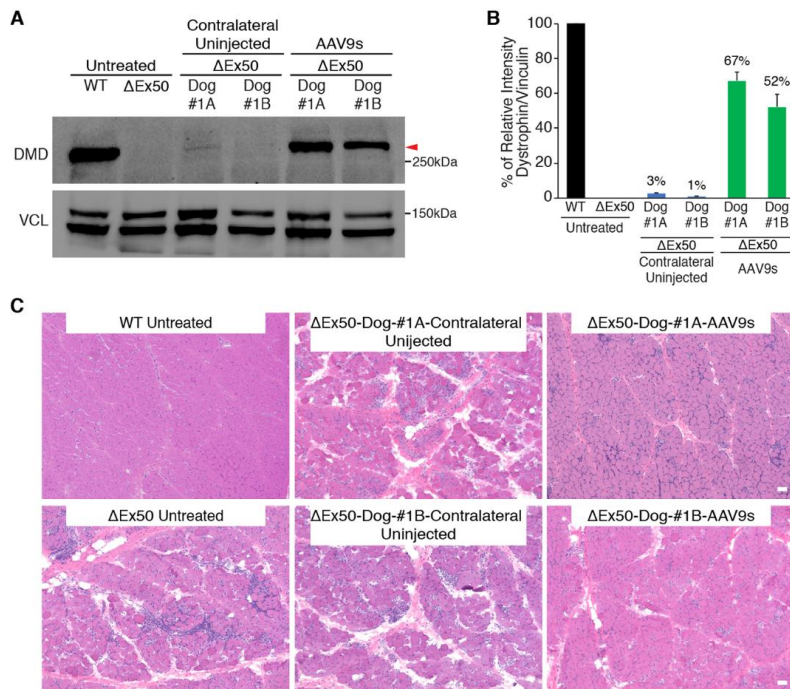


Fig. 2. Correction de la dystrophine après injection intramusculaire des vecteurs AAV9 délivrant le matériel génique. (A) Western blot de l'expression de la dystrophine (DMD) et de la vinculine* (VCL) dans le muscle tibial crânial 6 semaines après injection intramusculaire chez 2 chiens (#1A et #1B). (B) Quantification par Western blot de l'expression de la dystrophine après normalisation par la vinculine. (C) Histochimie par coloration à l'hématoxyline et éosine (H&E) du muscle tibial crânial d'un chien sauvage, d'un chien Δ Ex50 non-

traité, d'un muscle controlatéral Δ Ex50 non-injecté et des chiens Δ Ex50 ayant reçu une injection intramusculaire des vecteurs AAV9-Cas9 et AAV9-ARNsg-51 (AAV9 du chien- Δ Ex50-#1A-AAV9s et AAV9 du chien- Δ Ex50-#1B). Échelle : 50 μ m.

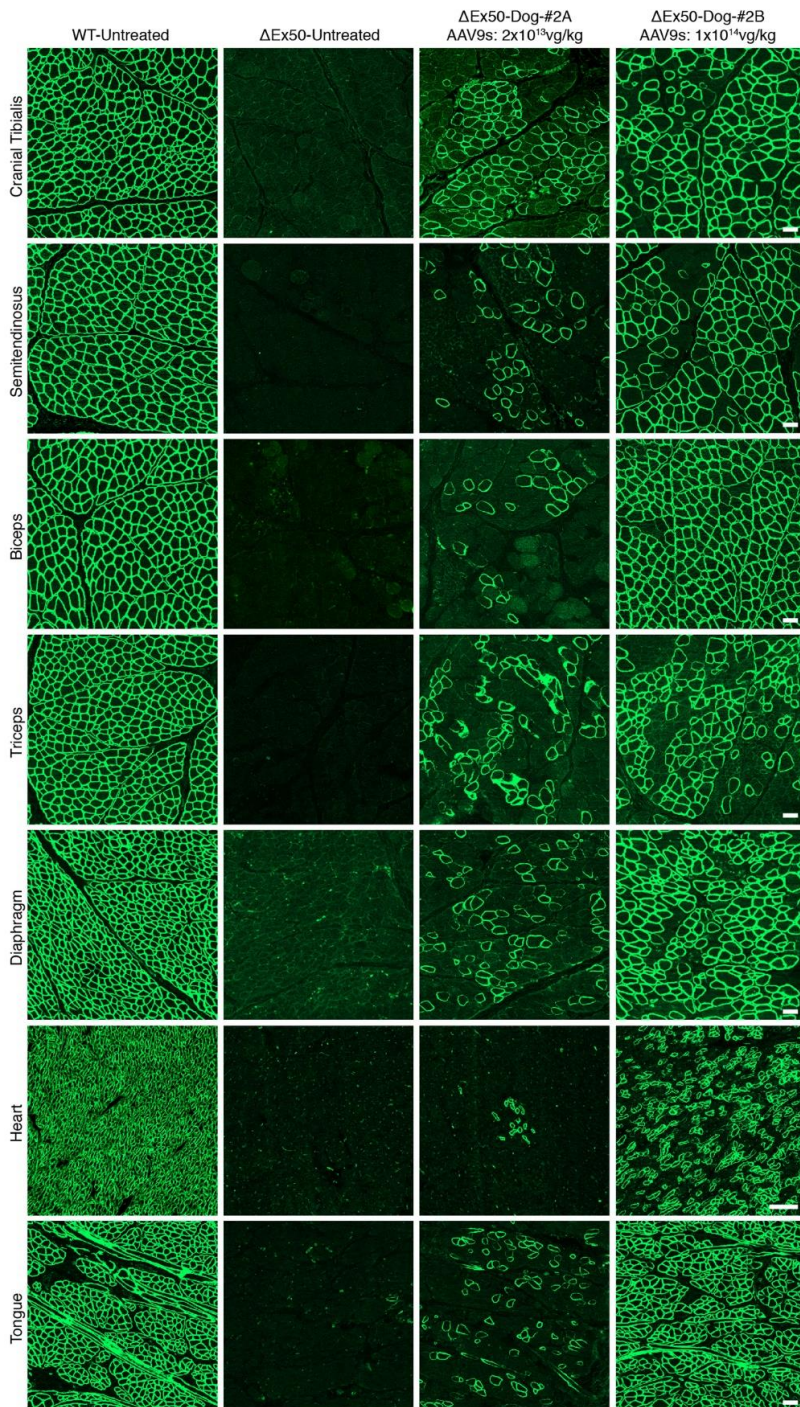


Fig. 3. Immunostaining of dystrophin following intravenous delivery of AAV9-encoded gene editing components. Dystrophin immunohistochemistry staining of cranial tibialis, semitendinosus, biceps, triceps, diaphragm, heart and tongue muscles of wild type dog, untreated Δ Ex50 dog, and Δ Ex50 dogs injected systemically with AAV9-Cas9 and AAV9-sgRNA at 2×10^{13} vg/kg (total virus 4×10^{13} vg/kg, referred as Δ Ex50-Dog #2A- AAV9s) and 1×10^{14} vg/kg (total virus 2×10^{14} vg/kg, referred as Δ Ex50- Dog #2B- AAV9s) for each virus.

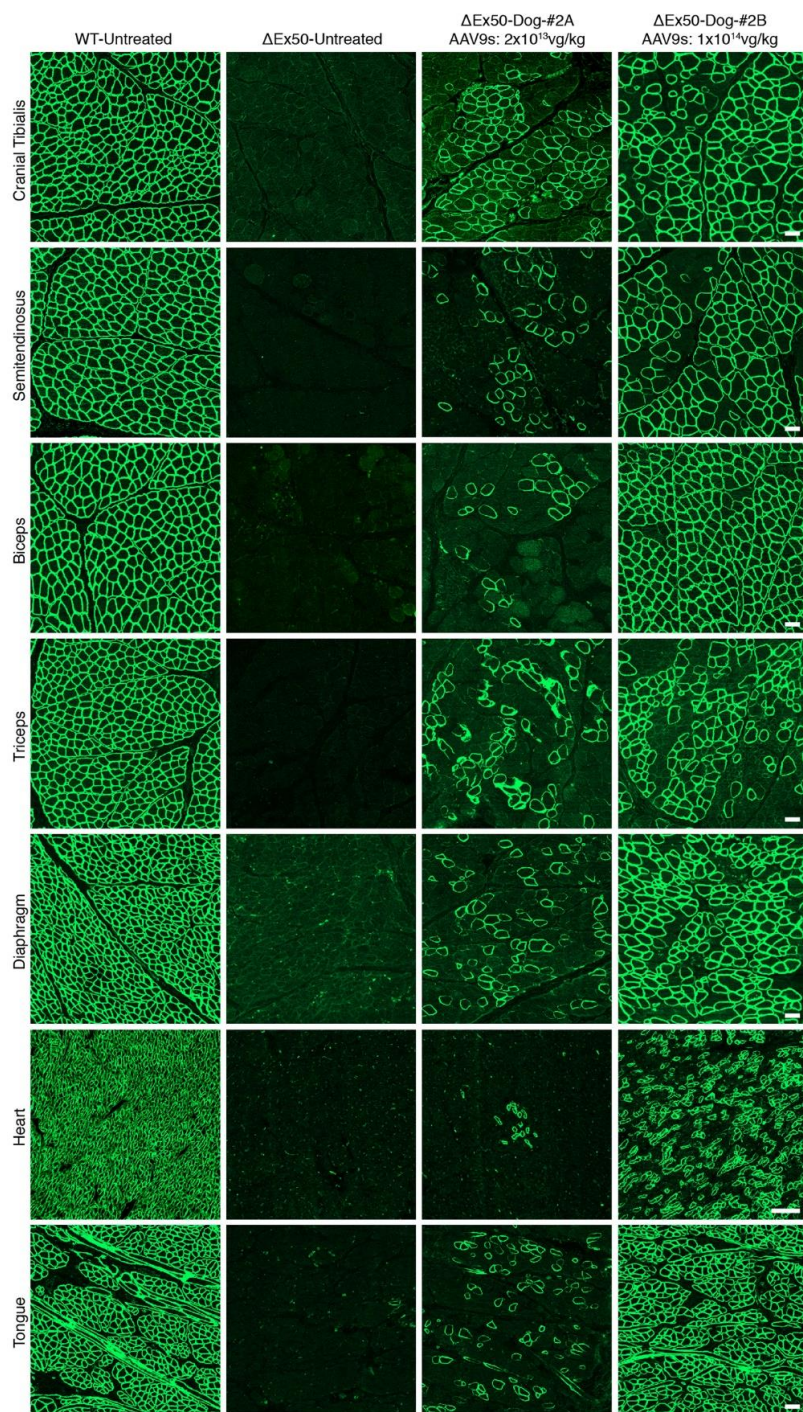


Fig. 3. Immunomarquage de la dystrophine après injection intraveineuse des vecteurs l'AAV9 délivrant le matériel d'édition génique. Coloration par immunohistochimie de la dystrophine dans le muscle tibial crânial, le muscle semi-tendineux, le biceps, le triceps, le diaphragme, le cœur et les muscles de la langue chez un chien sauvage, un chien ΔEx50 non traité et des chiens ΔEx50 ayant reçu une injection systémique des vecteurs AAV9-Cas9 et AAV9-sgARN de 2x10¹³ vecteur/kg (au total 4x10¹³ vecteur/kg de virus, correspondant aux AAV9 du chien-ΔEx50-#2A) et de 1x10¹⁴ vecteur/kg (au total 2x10¹⁴ vecteur/kg de virus correspondant aux AAV9 du chien-ΔEx50-#2B) de chaque vecteur viral.

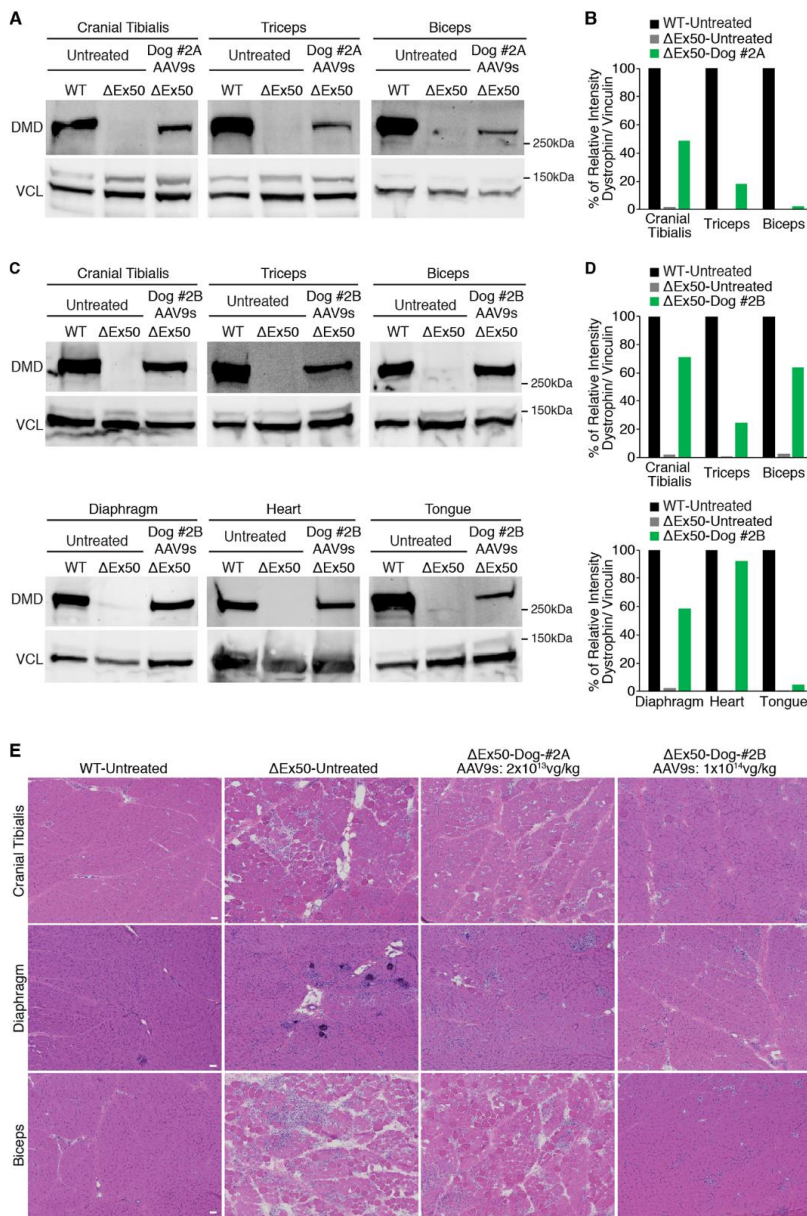


Fig. 4. Western blot of dystrophin and muscle histology following intravenous delivery of AAV9-encoded gene editing components. (A) Western blot analysis of dystrophin (DMD) and vinculin (VCL) of cranial tibialis, triceps, biceps muscles of wild type, untreated Δ Ex50, and Δ Ex50 injected with AAV9-Cas9 and AAV9-sgRNA at 2x10¹³ vg/kg for each virus (total virus 4x10¹³ vg/kg, referred as Δ Ex50-Dog #2A-AAV9s). (B) Quantification of dystrophin expression from blots after normalization to vinculin. (C) Western blot analysis of dystrophin (DMD) and vinculin (VCL) of cranial tibialis, triceps, biceps, diaphragm, heart, tongue muscles of wild type, untreated Δ Ex50, and Δ Ex50 injected with AAV9-Cas9 and AAV9-sgRNA at 1x10¹⁴ vg/kg (total virus 2x10¹⁴ vg/kg, referred as Δ Ex50-Dog #2B-AAV9s). (D) Quantification of dystrophin expression from blots after normalization to vinculin. (E) Hematoxylin and eosin (H&E) staining of cranial tibialis, diaphragm and biceps muscles of wild type, untreated Δ Ex50, and Δ Ex50 injected with AAV9-Cas9 and AAV9-sgRNA at 2x10¹³ vg/kg for each virus (total virus 4x10¹³ vg/kg) and 1x10¹⁴vg/kg for each virus (total virus 2x10¹⁴ vg/kg). Scale bar: 50 μ m.

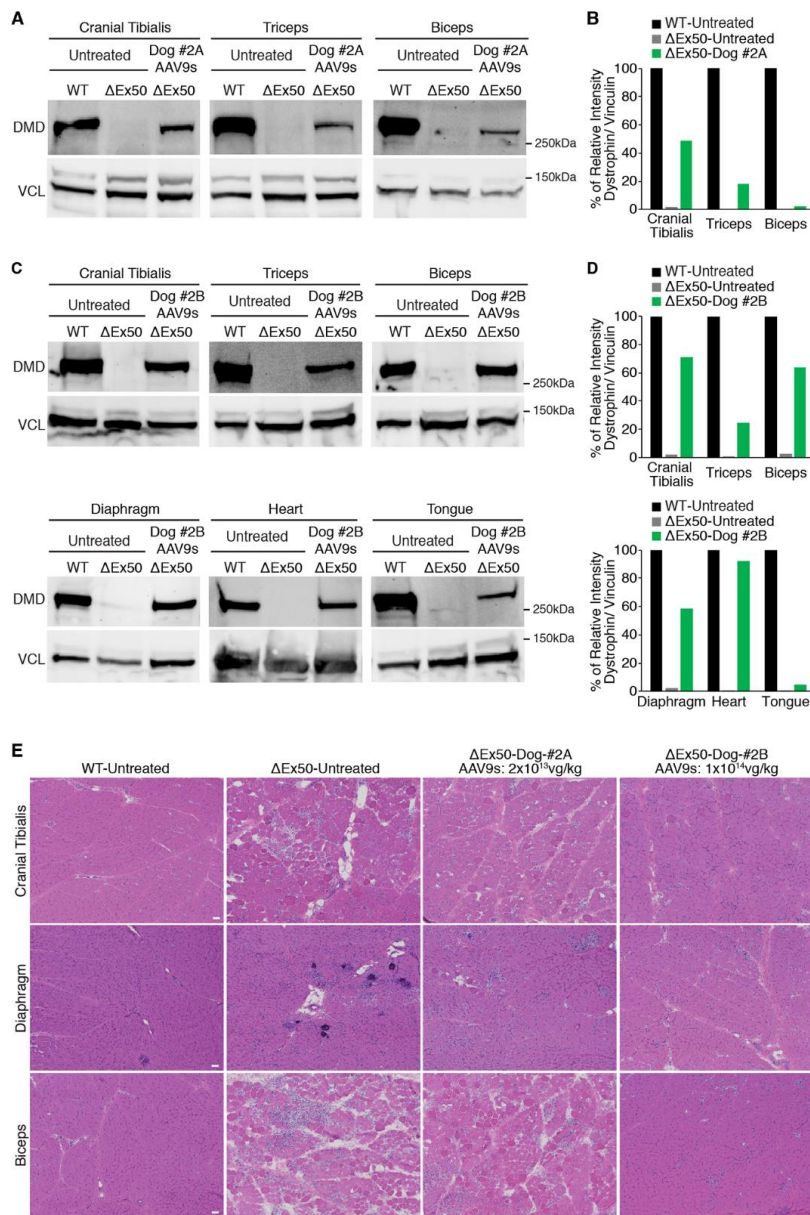


Fig. 4. Western blot de la dystrophine et histologie musculaire après injection intraveineuse des composants/du matériel d'édition génique intégré à l'AAV9. (A) Western blot de la dystrophine (DMD) et de la vinculine (VCL) dans le muscle tibial crânial, le biceps et le triceps chez un chien sauvage, un chien ΔEx50 non traité et un chien ΔEx50 ayant reçu une injection des vecteurs AAV9-Cas9 et AAV9-sgARN, chacun à 2×10^{13} vecteur/kg (au total 4×10^{13} vecteur/kg de virus, correspondant aux AAV9 du chien-ΔEx50-#2A). (B) Quantification par Western blot de l'expression de la dystrophine après normalisation par la vinculine. (C) Western blot de la dystrophine (DMD) et de la vinculine (VCL) dans le muscle tibial crânial, le biceps, le triceps, le diaphragme, le cœur et les muscles de la langue chez un chien sauvage, un chien ΔEx50 non traité et un chien ΔEx50 ayant reçu une

injection des vecteurs AAV9-Cas9 et AAV9-sgARN, chacun à 1×10^{14} vecteur/kg (au total 2×10^{14} vecteur/kg de virus correspondant aux AAV9 du chien-ΔEx50-#2B). (D) Quantification par Western blot de l'expression de la dystrophine après normalisation par la vinculine. (E) Coloration H&E du muscle tibial crânial, du diaphragme et du biceps chez un chien sauvage, un chien ΔEx50 non traité et des chiens ΔEx50 ayant reçu une injection des vecteurs AAV9-Cas9 et AAV9-ARNsg, chacun à 2×10^{13} vecteur/kg (au total 4×10^{13} vecteur/kg) et à 1×10^{14} vecteur/kg (au total 2×10^{14} vecteur/kg). Échelle : 50 μm

Partie 3 : STRATÉGIE DE TRADUCTION

I- Présentation du texte-support

Titre : Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy

Amoasii, L., Hildyard, JCW., Li, H., et al., *Science*, 2018. Vol.362, N°6410, pp. 86-91

Cet article présente une stratégie thérapeutique utilisant une technologie innovante de correction du génome (Crispr-Cas9) et montre qu'il a été possible de corriger une maladie génétique. Il s'agit essentiellement d'un article de méthodologie, d'une démonstration de rigueur scientifique. Il traite plus précisément d'expériences menées sur des chiots atteints de la dystrophie musculaire de Duchenne (DMD) et des résultats obtenus à court terme (6 à 8 semaines) à la suite d'une injection de deux composants de Crispr-Cas9 (un ARN guide qui permet de cibler la séquence à réparer dans l'ADN, et la protéine Cas9 qui permet de découper l'ADN à cet endroit-là). Le modèle canin est particulièrement intéressant à étudier car les mutations, les symptômes et les réactions immunitaires sont très proches du modèle humain, ce qui permet d'évaluer l'efficacité thérapeutique de la DMD.

Choix de l'article

L'étude scientifique faisant ici l'objet d'une traduction a tout d'abord été sélectionnée pour la pertinence du sujet qu'il traite. J'ai considéré que l'avantage de ce domaine très actuel était qu'il semblait très médiatisé et que les ressources de vulgarisation ne manquaient pas (vidéos pédagogiques ou de divertissement, documentaires, travaux scientifiques, publications de centres de recherche, etc. – voir la bibliographie commentée). De plus, l'intérêt pour moi de la question de la traduction scientifique dans un domaine de recherche très pointu, et surtout actuel (comme ici CRISPR-Cas9), est que l'essentiel se fait en anglais, et beaucoup d'aspects terminologiques et phraséologiques ne sont pas forcément (bien) institutionnalisés.

La traduction pouvait donc être un réel défi.

En effet, la thérapie génique est une discipline qui évolue rapidement et dont le langage se meut à une cadence semblable. Il est donc intéressant de se demander dans quelle mesure sa traduction doit rester traditionnelle ou pas. D'un côté, les sciences de la vie ont une histoire linguistique longue, et ses spécificités se sont ancrées dans l'usage au fil du temps. De l'autre, l'accélération du développement des nouvelles technologies et la mondialisation du langage des sciences supposent une déformation de ce parler pour s'adapter à notre époque.

Cet article m'a été fourni par mon référent, qui travaille, entres autres, sur cette question, parmi d'autres articles similaires. J'ai considéré qu'il avait un niveau de spécialisation suffisamment complexe pour susciter un travail de recherche et de préparation conséquent, et que le travail de vulgarisation préalable me permettait de présenter le sujet de manière claire et exhaustive, dans la mesure du possible.

II- Méthode de travail et démarches de recherche

La méthodologie employée pour la traduction de l'article reprend les différentes méthodes de traduction qui m'ont été enseignées tout au long de mes deux années de master professionnel, aussi bien en cours de traductions technique et générale, qu'en cours de terminologie. Du fait de la complexité du sujet et de ma connaissance très restreinte du domaine, le processus a été particulièrement long car il s'est accompagné de recherches abouties et de discussions avec mon référent, au cours desquelles nous avons notamment abordé des questions techniques mais aussi d'usage. Il est donc concevable que l'exercice ne reflète pas dans l'absolu la réalité du métier, qui devrait plutôt s'orienter, dans ce cas-là, vers un apprentissage progressif (et des études en biologie !) menant à la spécialisation. Cependant, les méthodes à long terme restent les mêmes que celles qui seraient appliquées dans un cadre professionnel.

J'ai jugé que l'utilisation des outils de traduction assistée par ordinateur n'était pas pertinente dans mon cas, étant démunie de ressources préalables conséquentes, et cette traduction faisant l'objet d'un exercice ponctuel. C'est pourquoi la traduction s'est déroulée sur un tableau en deux colonnes dans une page de traitement de texte, avec apposition de nombreux commentaires pour chaque doute, questionnement, alternative ou justification (voir annexe 2).

Dans cette partie, je commencerai par décrire le déroulé complet de l'exécution de la traduction et de toutes les tâches satellites qui l'accompagnent. Je procéderai ensuite à la justification de certains de mes choix terminologiques, qui pourront éclairer mes choix dans leur ensemble mais aussi les fiches terminologiques qui suivent. Je me focaliserai enfin sur certaines difficultés de traduction, propres au domaine d'étude et aux spécificités du type de texte choisi.

1. Déroulé

Avant toute chose, il convient de noter que le processus de traduction du texte support s'est évidemment fait conjointement avec celui de la recherche et de la rédaction de la première partie (« Exposé »). En effet, même si des recherches préalables sont nécessaires pour aborder le texte de spécialisation, celui-là même fourni des pistes concrètes et des indices permettant d'orienter la recherche. Ainsi, chronologiquement, l'exercice est caractérisé par des va-et-vient entre traduction de l'article et investigation.

Avant même de me plonger concrètement dans la lecture du texte-support, et considérant qu'il fallait d'abord avoir un aperçu général du domaine scientifique dans lequel je m'apprêtais à m'immerger, j'ai organisé une discussion avec un spécialiste en thérapie génique, mon référent pour ce mémoire, qui a bien voulu me faire une introduction du concept d'édition du génome. Une fois combinée à l'analyse contextuelle de l'article, cette première mise en jambe permet une mise en contexte générale en vue d'être capable de situer précisément mon texte dans le domaine.

Les éléments extratextuels tels que la date de publication, les auteurs ou la source sont autant d'indices permettant d'identifier le niveau de spécialisation du texte et le public visé. L'on peut considérer le domaine de la thérapie génique pour soigner les maladies génétiques comme une sphère. Voir la sphère dans son ensemble, bien que de manière encore très abstraite, d'une part, et considérer les enjeux du texte d'autre part, permet de localiser le texte plus ou moins précisément à l'intérieur de cette sphère. Grâce à cette première approche, j'ai pu très facilement identifier la finalité du texte et donc décider du but de ma traduction. Cet article étant écrit par et pour des spécialistes, il se veut aussi être pour des spécialistes dans la langue cible, donc il ne s'agira en aucun cas de procéder à de la vulgarisation, ne serait-ce qu'imperceptiblement.

J'ai ensuite effectué des recherches, en français d'abord, puis en anglais, pour prendre une certaine habitude avant de me plonger dans la lecture de l'article que j'allais traduire. J'aurai fait l'erreur d'orienter ces premières recherches sur des sujets trop pointus, alors que je manquais déjà des éléments plus élémentaires, comme les concepts généraux de la génétique ou les caractéristiques définitives d'une maladie génétique. En effet, j'ai surtout essayé dans un premier temps de trouver des articles semblables à celui que j'allais traduire, pour avoir de quoi comparer mon texte. Ma première (vraie) lecture du texte s'est donc effectuée en essayant d'y appliquer mes connaissances sommaires. La deuxième lecture a été plus pointilleuse, surlignant tous les termes nouveaux et/ou inconnus et les termes dont j'avais le pressentiment qu'ils seraient clefs, puis en surlignant tous les passages qui m'étaient incompréhensibles pour l'instant.

Le but de ces deux lectures était de dégager la structure du texte, sa logique interne et sa structure chronologique. Au travers d'une schématisation très visuelle du texte, il m'apparaît plus facile d'en appréhender tous les aspects. Dans un premier temps, tout ce travail est bilingue. En effet, à ce stade, la langue ou le souci de traduction n'a pas d'importance pour moi car il s'agit de dégager le sens, plus ou moins dépourvu de tout signifiant. La compréhension est conceptuelle et non pas textuelle. Cette schématisation constitue un travail de décorticage qui agit comme une déverbalisation du langage pour me libérer de toute terminologie que je ne saisis pas. Par exemple, j'ai ainsi pu comprendre les fonctionnements internes du texte, sans pour autant avoir parfaitement accès aux techniques biotechnologiques employées ou à quoi se référaient les termes médicaux.

Enfin, la dernière étape avant de commencer la traduction a consisté en un retour sur chaque passage obscur, accompagné d'une recherche documentaire plus approfondie. La recherche s'effectuait essentiellement par mot-clef sur Internet, ou par repérage de passages similaires dans la documentation anglophone que j'avais amassée dans la première phase, et par comparaison avec des explications similaires dans la documentation française. C'est ainsi que j'ai créé, au fur et à mesure et de manière officieuse, une sorte de corpus comparable (voir bibliographie « ressources académiques »), qui a précédé la compilation plus concrète de corpus anglais et français survenue plus en aval de la traduction (voir partie II.3 « Corpus comparables »). Cette manière de se documenter et d'éclairer certains passages justifie l'importance du décorticage en amont des concepts a priori. En effet, si je veux pouvoir comparer deux passages similaires dans une source anglaise et une source française, il faut que je saisisse les concepts et la logique abordés dans chacune de ces sources afin de pouvoir les mettre en parallèle de manière sémantique. Progressivement, ce premier corpus comparable m'a donc aidé à la fois à élucider les passages compliqués, et à repérer des équivalences terminologiques et phraséologiques.

Une fois le texte appréhendé, sa structure intégrée et les termes repérés, j'ai effectué un premier jet de traduction. Pour ce faire je ne me suis pas arrêtée à chaque fois que je butais, palliant le manque de compréhension de certains passages par des tournures pansement, pour y revenir plus tard. Cette méthode a donné un premier jet plus semblable à un squelette de traduction qu'à un texte complet. Mais cette ébauche a permis une meilleure connaissance du texte et de sa géographie. Tout au long de ce processus, j'ai pris des notes à part, de manière très personnelle et de sorte à former une leçon simplifiée et schématique qui, en l'appliquant au texte à traduire, m'aidait à me situer dans son déroulement et son contenu. En parallèle, j'ai effectué une recherche terminologique continue et fragmentaire (qui s'est ensuite accompagnée d'une recherche terminologique exhaustive). Si je ne trouvais pas d'équivalence pour un terme, ou bien que j'hésitais, je continuais tout de même ma traduction, considérant que c'était en continuant à me documenter ou en laissant le texte décanter que je trouverai sûrement une solution. Progressivement j'ai complété les vides et rectifié des passages.

L'utilisation des corpus comparables compilés sur *Sketch Engine* et des différentes fonctionnalités de *Sketch Engine* pour multiples faits de langues sont intervenus lors du deuxième jet de traduction, une fois que le premier avait été mis en chair. La transformation du brouillon de traduction en une traduction presque finale s'est effectuée en lissant tous les derniers passages où j'avais buté et en complétant la terminologie et la phraséologie manquante, puis en éliminant toutes les hésitations. Cette étape s'est accompagnée de quelques discussions avec mon référent sur des questions d'usage mais aussi, dans une moindre mesure, de compréhension. Ces deux aspects ont contribué à clarifier tout à fait des segments pour lesquels subsistaient des doutes. J'ai ensuite fait reposer le texte une première fois avant de procéder à une relecture complète, toujours en gardant en tête le texte source, et en effectuant des allers-retours progressifs. L'agencement sous forme de tableau (texte source à gauche, texte source à droite) a facilité le repérage de potentielles omissions, d'écarts de sens ou de faux sens. Lors de cette relecture, j'ai à nouveau opéré un lissage final et quelques reformulations. La première relecture a donc pu valider mon texte, par la correction des erreurs de contenu et la correction des fautes d'orthographe et de typographie.

Après avoir réadapté la mise en page dans sa forme finale, j'ai à nouveau laissé reposer le texte plusieurs jours, et fait une dernière lecture, sans texte source, pour m'assurer de la cohérence rédactionnelle et sémantique de mon texte cible. La phase finale a confirmé que le niveau de spécialisation était identique à celui du texte source, et par conséquent que le lectorat visé était le bon. Par exemple, j'ai constaté que je n'étais pas tombée dans les pièges que peuvent parfois impliquer les étoffements : même s'il m'a fallu, en tant que traductrice, faire une heure de recherche pour comprendre un bout de phrase et le traduire, cela ne veut pas dire que le lecteur visé, lui, ne comprend pas directement (il n'est pas traducteur, il est spécialisé dans l'édition du génome, il n'a pas besoin que je lui prouve que j'ai compris de quoi il s'agissait).

Pour conclure la phase de traduction, le texte français a été relu par mon référent, qui lui-même avait aidé à valider certains usages. Sa perception était déterminante pour savoir si j'avais rempli ou non les objectifs que je m'étais donnés.

2. Corpus comparables

Lors de la traduction de mon texte support, j'ai été confrontée à un nombre considérable de problèmes terminologiques et de phraséologie. En effet, le jargon scientifique employé est très loin de mon usage linguistique habituel. Pour répondre à ces problèmes, j'ai effectué beaucoup de recherches à partir d'articles similaires, de vidéos, de documentaires, de plateformes de traduction, et de glossaires. L'utilisation de corpus comparables s'est avérée très utile pour m'assister dans certains de mes choix terminologiques et phraséologiques en les confirmant ou en les infirmant, au regard des usages observés.

Dans ce cadre, la compilation de corpus comparables sur Sketch Engine, ainsi que l'utilisation des différentes fonctionnalités de cet outil ont pu m'aider à me faire une idée générale du vocabulaire employé dans les articles scientifiques du même style que le mien. Par exemple, grâce à la fonctionnalité *Wordlist* et en faisant une recherche sur la fréquence de l'emploi de tous les verbes, j'ai pu avoir des exemples concrets de verbes qui s'utilisent dans la démonstration scientifique, la stratégie thérapeutique ou la description d'une expérience et de ses résultats. Sketch Engine m'a aussi permis d'établir le profil de certains mots ou termes, pour comprendre dans quel cadre ils étaient employés, avec quels autres mots et à quelle fréquence, me familiarisant ainsi avec ce parler particulier. Consciente des imperfections de l'outil, j'ai tout de même extrait de ma traduction ou du texte-support plus d'une cinquantaine de faits de langue que j'ai analysés sous l'angle des corpus compilés.

Je détaillerai ci-dessous succinctement la méthode de compilation de corpus utilisée par le biais de *One Click Term* et Sketch Engine, et les difficultés de cet exercice.

Compilation des corpus EN et FR

Pour procéder à la compilation d'un corpus en anglais (corpus EN), la langue source de mon texte-support, j'ai commencé par passer l'article dans *One Click Term* et j'en ai extrait une liste des candidats-termes (les 150 premiers, le reste n'étant pas pertinent).

J'ai nettoyé l'Excel des résultats obtenus en ôtant le bruit potentiel et en décidant pour chaque candidat s'il s'agit d'un terme ou d'un mot-clef pertinent ou non. Cette première base terminologique (TB) se constitue de 112 termes contre 129 que j'avais préalablement relevés manuellement. Le nettoyage de l'Excel n'a peut-être pas été parfaitement méticuleux, mais cela n'a pas d'importance étant donné que j'avais déjà procédé à une extraction manuelle auparavant, donc je savais quels termes je voulais retenir. L'extraction par *One Click Term* m'a permis de comparer le résultat avec mon propre relevé, et de confirmer la pertinence de mes choix, par exemple au regard de la fréquence de l'emploi de certains termes.

Pour compiler un corpus de documents en anglais, j'ai fait une recherche sur le web (à partir de Sketch Engine) par mot-clef. À partir de la TB obtenue précédemment, j'ai décidé de choisir des mots-clefs dont la pertinence me paraissait justifiée afin d'obtenir des textes similaires à mon texte-support, à la fois concernant les domaines abordés et le niveau de spécialisation. Ma première recherche comportait donc les mots suivants : « Cas9 », « sgRNA », « exon » et « dystrophin ». Ainsi, chaque combinaison donnerait lieu à des textes en génétique et maladie génétique (en combinant les termes « sgRNA », « exon » et « dystrophin »), ou en thérapie génique et génétique (en combinant les termes « Cas9 » ; « sgRNA » et « exon »), ou en maladie génétique et thérapie génique (en combinant tous les termes, de la manière la plus idéale, puisque ce sont les deux domaines de mon texte-support). Avant de valider la compilation, j'ai pris garde de ne conserver que les documents dont la source me paraissait fiable.

J'ai ainsi obtenu un corpus en anglais (corpus EN) d'un total de 74 textes et 338.213 mots. La majorité des documents étaient des travaux scientifiques et/ou universitaires. À partir de ce corpus EN, j'ai effectué une extraction des termes grâce à la fonctionnalité *Keywords* de Sketch Engine, en ne conservant ici aussi que les 150 premiers termes, puis en procédant au nettoyage.

Grâce au relevé terminologique en langue source, j'ai pu créer un premier corpus comparable en langue cible (corpus FR). Pour cette première compilation, j'ai fait le choix d'utiliser les mêmes mots-clefs qu'en anglais, à partir d'hypothèses de traduction en langue cible. Nous verrons par la suite que ce n'était pas forcément la bonne décision.

J'ai donc utilisé les termes « Cas9 », « exon », « dystrophine » et « ARNsg ». Avec cette recherche, j'ai obtenu un total de 154.505 mots, ce qui est faible comparé au corpus EN. La majorité des textes étaient de vulgarisation, ou bien didactiques, avec beaucoup de doublons (plusieurs sites reprennent les mêmes textes, ce qui fausse les statistiques et les analyses que je peux en tirer par la suite sur Sketch Engine). Enfin, une grande partie des documents étaient des traductions de l'anglais, dont je ne pouvais juger avec certitude de la qualité.

En comparant les corpus EN et FR, il apparaissait que le corpus FR avait 15 % de textes en moins, mais 54 % de mots en moins. On peut en conclure que les textes dans le corpus FR étaient en moyenne beaucoup plus courts, ce qui n'est pas étonnant étant donné que la majorité d'entre eux étaient des textes de presse. De plus, dans leur qualité, les corpus EN et FR étaient très déséquilibrés. En effet, étant donné le domaine de recherche et la disponibilité des ressources sur Internet, on obtient facilement un corpus français très peu spécialisé. La plupart des articles sont de vulgarisation, journalistiques, plutôt courts, et le corpus est en grande partie parasité par des articles scientifiques au titre ou à l'abstract français, mais au corps de texte anglais. Ce déséquilibre s'explique par le fait que la recherche en thérapie génique s'effectue majoritairement en anglais, même au sein des équipes françaises de chercheurs qui ont tendance à publier en anglais. L'autre raison est que ce sujet a été très médiatisé lorsque le prix Nobel de chimie a été remis en 2018 aux chercheuses ayant inventé la technologie de Crispr-Cas9. La plupart des articles s'adressent donc au grand public.

La difficulté de mon texte-support est qu'il mélange deux, voire trois domaines : la thérapie génique (plus précisément la technologie de ciseaux moléculaires) et les maladies génétiques, mais aussi dans une moindre mesure l'histologie (étude et analyse des tissus musculaires). Il fallait donc choisir des mots-clefs appartenant à ces deux (trois) domaines afin d'obtenir un corpus complet touchant à tous les potentiels problèmes de phraséologie auxquels j'étais confrontée.

Les problèmes des mots-clefs choisis dans la première compilation du corpus FR sont multiples. Premièrement, ce sont des termes identiques en anglais et en français (à part « dystrophine »), donc les résultats obtenus sont très souvent des textes anglophones. Ensuite, ils ne sont peut-être pas suffisamment spécialisés.

En examinant les textes obtenus, je me suis rendu compte que beaucoup de textes didactiques employaient ces termes spécialisés, ce qui n'était pas aussi pertinent pour l'exercice, et il manquait toute une partie portant sur les stratégies thérapeutiques et les essais cliniques ou en laboratoire. Pour viser ce type de texte, il fallait ajouter un terme dans le domaine de l'imagerie médicale, de l'histologie, ou de la stratégie thérapeutique.

Au regard de ces défauts, j'ai donc procédé à une seconde compilation de corpus FR avec de nouveaux mots-clefs. Il s'agissait de mélanger les termes anglais et français, car certains termes ne s'emploient qu'en anglais, mais que je voulais tout de même tomber sur des textes français. Il fallait aussi mélanger des termes plus généraux avec des termes plus spécialisés, pour tomber sur des articles avec le même niveau de spécialisation que mon texte-support, mais quand même obtenir suffisamment de ressources afin d'obtenir un corpus conséquent. Les mots-clefs utilisés pour la deuxième compilation sont les suivants : « Cas9 », « dystrophine », « immunohistochimie », « séquençage » et « épissage ». Après avoir effectué un tri, j'ai obtenu 51 documents pour 131.106 mots. Les textes semblaient plus spécialisés, l'ensemble homogène, mais beaucoup de textes ne portaient pas spécifiquement sur la thérapie génique.

Même une fois des mots-clefs plus pertinents choisis, je suis restée confrontée à l'insuffisance des ressources en français sur Internet. Afin que les corpus FR et EN soient bien équilibrés et que le corpus FR devienne pertinent pour m'aider dans ma traduction, il a fallu procéder à un ajout manuel de textes, préalablement sélectionnés. Pour ce faire, plusieurs plateformes étaient à ma disposition : *Google scholar*, *Jstor*, la bibliothèque en ligne de la BSN, etc. De plus, il a été pertinent de jeter un œil aux bibliographies des articles que je possédais déjà. J'ai aussi contacté mon référent à l'occasion d'une discussion pour mon mémoire et il m'a fourni quelques ressources utiles que je pouvais intégrer aussi bien au corpus FR qu'au corpus EN. Après une recherche bibliographique plus approfondie et avec quelques documents que j'avais déjà employés dans le cadre de l'écriture de mon mémoire, j'ai ainsi enrichi manuellement le corpus FR, mais aussi dans une moindre mesure le corpus EN. J'ai surtout ajouté des textes sur la dystrophie musculaire de Duchenne, car c'est cet aspect-là qui manquait à cause des mots-clefs sélectionnés. Finalement, mon nouveau corpus FR se composait de 72 documents pour 339.199 mots. Les corpus obtenus n'étaient pas particulièrement fournis, mais il faut garder en tête que le domaine reste pointu et relativement récent.

Enfin, j'ai effectué une extraction des termes à partir du corpus FR, toujours avec la fonctionnalité *Keyword*, me permettant ainsi d'établir des correspondances potentielles entre termes sources et cibles.

Travail terminologique et traduction à l'aide du corpus

Vérification de certains usages terminologiques

J'ai utilisé les corpus comparables par le biais de Sketch Engine tout d'abord pour effectuer des recherches terminologiques : emploi et sens d'un terme en contexte, statistiques sur la fréquence d'utilisation, collocations pertinentes, contextes définitoires, etc. J'ai ainsi pu déterminer si certains faits de langue étaient des termes ou non, j'ai amélioré ma compréhension de certains termes grâce aux différents contextes d'utilisation, et j'ai pu leur trouver des équivalences en passant du corpus EN au corpus FR. Enfin, j'ai résolu des hésitations de choix terminologiques. Les parties suivantes constituent certains exemples illustrant la manière dont j'ai employé les différentes fonctionnalités de Sketch Engine.

Rechercher une équivalence FR du terme “immunostaining”

Pour ce terme, je pars avec plusieurs hypothèses, sans toutefois savoir la quelle est la meilleure : « immunomarquage », « immunocoloration » ou « immunofluorescence ».

J'utilise la fonctionnalité *Wordlist* et je recherche [*noun*] + [*starting with “immuno”*]. Les deux premiers termes les plus fréquemment utilisés sont tous les deux des termes désignant un type d'analyse (des techniques d'immunomarquage), mais je sais que le premier terme « immunohistochimie » correspond à « immunohistochemistry » en anglais. Donc je passe le deuxième terme « immunofluorescence » dans le concordancier pour avoir des pistes sur une potentielle correspondance avec « immunostaining ». Je comprends en contexte que l'immunofluorescence est un type d'immunomarquage. Or, compte tenu du contexte définitoire de « immunostaining » dans le corpus EN, ce terme correspond plutôt à un hyperonyme. Dans la liste des mots commençant par « immuno », « immunomarquage » n'est employé que 3 fois, mais son contexte définitoire semble mieux correspondre à « immunostaining ».

Avec des recherches plus approfondies sur la définition de ce terme, je confirme ce choix. Cet exercice montre les limites de Sketch Engine, ou plus spécifiquement de mes corpus, qui restent déséquilibrés à plusieurs niveaux.

Rechercher une équivalence FR du terme « primer »

Dans le corpus EN, fonctionnalité *Concordance*, je fais une recherche simple de « primer » pour observer avec quels autres mots ce terme est souvent utilisé (en alignant le *Left context* puis le *Right context*). Je peux faire de même à partir de la fonctionnalité *Word Sketch*, surtout en regardant la liste dans la catégorie [*modifiers* of « primer »] car ce sont ces mots qui m'aideront le mieux à faire une recherche dans le corpus FR. Je vois que « primer » est souvent utilisé avec le nom « PCR », les adjectifs « antisense » et « reverse », et avec les verbes « use » et « amplify ». À partir de cette observation, je peux étudier les termes en français qui correspondraient à « primer » dans sa définition, et qui apparaissent souvent avec les noms, adjectifs et verbes énumérés. Dans le corpus FR, je passe ces derniers (en français : « PCR », « antisens », « utiliser », « amplifier ») dans *Concordance* et *Word Sketch* pour émettre une hypothèse sur l'équivalence FR de « primer ». Le terme français qui ressort de cette recherche est « amorce ». Pour vérifier mon hypothèse, j'utilise le *CQL builder* dans *Concordance* et fais les requêtes suivantes :

- [lemma="amorce"][]{}{0,4}[lemma="PCR"]
- [lemma="PCR"][]{}{0,4}[lemma="ammorce"]
- [lemma="amorce"][]{}{0,4}[lemma="antisens"]
- [lemma="utiliser"][]{}{0,4}[lemma="ammorce"]

Ces concordances en contexte me permettent de valider que l'environnement lexical de « amorce » en français est le même que celui de « primer » en anglais.

Savoir si « human induced pluripotent stem cells » est un terme

Je souhaite savoir si cette expression est un terme figé ou bien s'il s'agit seulement d'une apposition d'adjectifs qualifiant le terme « stem cells » (en français « cellules souches »).

Pour ce faire, j'utilise la fonctionnalité *Concordance* dans le corpus EN et je cherche « pluripotent stem cells ». En alignant le *Left context*, j'observe beaucoup d'occurrences pour « induced pluripotent stem cells » et « human pluripotent stem cells ». J'en conclus que « human » et « induced » sont deux adjectifs qualifiant « pluripotent stem cells », donc ils sont interchangeables. Pour trouver l'équivalence en français, j'utilise dans le corpus FR la fonctionnalité concordance et je recherche [cellules souches * *] afin de trouver une potentielle traduction de « pluripotent » et de « human » ou « induced ». Je trouve une grande quantité d'occurrences pour « cellules souches pluripotentes humaines » et « cellules souches pluripotentes induites ». Je trouve même « cellules souches pluripotentes induites humaines » et « cellules souches pluripotentes humaines induites ». De même qu'on retrouve l'abréviation « iPSC humaines » ou « hPSC » ou même « hiPSC ». Je confirme que « cellules souches pluripotentes (PSC) » s'accompagne de manière interchangeable avec « humaines » et « induites ».

Co-occurrences et phraséologie

Sketch Engine m'a aussi permis de résoudre des hésitations contextuelles sur certains faits de langues, que ce soit à propos de l'ordre syntaxique, des prépositions correctes, ou des emplois figés ou non.

« Target a region » / « target site » / « targeting efficiency » / « targeted genomic site »

Je veux confirmer que le lemme « cibler » s'emploie aussi bien dans tous les faits de langue cités ci-dessus, et de manière habituelle, ou bien si certains phraséologismes se composent avec un autre mot. Pour commencer, je fais une recherche simple dans *Concordance* (corpus FR) de [cibl*], puis grâce à l'option *Collocations*, je recherche les co-occurrences pertinentes éloignées à plus ou moins 3 mots (*positions*) (*Left ou Right context*) (voir l'annexe 3). Cela me permet d'observer la fréquence des occurrences avec le lemme « cibler ». Ainsi, je vois que l'occurrence la plus fréquente se fait avec « séquence », mais il y a aussi beaucoup d'occurrences avec « gène » et avec « région » ou son pluriel « régions ».

Cela me permet aussi de voir ces occurrences en contexte (je sélectionne les occurrences qui m'intéressent et elles s'affichent dans le concordancier).

Je confirme ainsi que « région cible », « région ciblée » et « cibler une région » s'emploient abondamment, de même que « séquence cible », « séquence d'ADN ciblée », « site cible » ou encore « fréquence de ciblage » et « ciblant le gène ».

Rétablir ou restaurer ?

Lors de ma traduction, j'hésite à traduire « restore the reading frame » par « rétablir le cadre de lecture » ou par « restaurer le cadre de lecture ». Je vais donc utiliser la fonctionnalité *Word Sketch Difference* pour comparer l'usage de « rétablir » avec celui de « restaurer ». Dans la liste [*objects of* rétablir/restaurer] je trouve le mot « cadre » et regarde le nombre d'occurrences qu'il y a pour chaque verbe. Dans *Concordance*, je remarque que les deux verbes sont utilisés à quasiment la même fréquence (voir l'annexe 3). J'en conclus qu'il n'y a pas de phraséologie particulière associée à ce fait de langue et que je pourrai choisir l'un ou l'autre verbe.

Déterminer l'article correct

Un problème récurrent auquel je suis confrontée lors de ma traduction est celui de l'absence de déterminant en anglais pour des termes qui ne sont pas des mots officiels, ni de la langue française, ni de la langue anglaise. Par exemple, les termes « sgARN » (pour « simple guide RNA ») et « Cas9 » désignent respectivement un type d'acide ribonucléique et une protéine. Seulement, il n'est pas dit que ces deux abréviations prennent un déterminant, et si c'est le cas, rien n'indique en grammaire les quels il faudrait utiliser. Par exemple, dans le cas de « Cas9 », sachant qu'il s'agit d'une protéine (ou enzyme), dirait-on « l'injection in vivo de Cas9 », « l'injection in vivo de la Cas9 » ou « l'injection in vivo de la protéine Cas9 » ? Grâce à la fonctionnalité *Concordance*, je recherche la fréquence d'utilisation de « la Cas9 » et le résultat est au nombre de 154.

En faisant de même avec « la protéine Cas9 » j'obtiens 34 résultats, et en tapant seulement « Cas9 », j'observe une quantité conséquente d'occurrences sans article. Je peux déjà éliminer l'emploi de « la protéine Cas9 », surtout dans un contexte aussi spécialisé que celui de mon texte-support.

Il en va de même pour les formulations raccourcies en anglais comme « AAV9-Cas9 » et « AAV9-sgRNA-51 » par exemple dans « AAV9-Cas9 and AAV9-sgRNA-51 were initially introduced into the cranial tibialis muscles » (p.62, paragr.5). La question se pose d'autant plus que « AAV9 » désigne un virus (adénoassocié de sérotype 9), donc s'il est associé à « Cas9 », une protéine, il faudra trouver le genre correct. Un bon moyen de contourner le problème serait de désigner « AAV9-Cas9 » par un autre nom ou en utilisant une périphrase. Or, l'on parle souvent des « composants de Crispr-Cas9 » ou des « complexes viraux ». On pourrait donc faire une hypothèse sur l'emploi de « les complexes/composants AAV9-Cas9 et AAV9-ARNsg-51 ». En cherchant « complexe » dans *Concordance* et en alignant le *Right context* on tombe sur 26 occurrences de « complexe Cas9/ARN » (qui proviennent d'au moins deux sources différentes). En faisant de même avec « composant », on ne tombe sur aucun résultat concluant. De plus, en cherchant « AAV* » on tombe sur plusieurs termes composés (AAV-Cas9 ; AAV9-SaCas9 ; etc.) précédés de la préposition « d' » ou du déterminant « l' ». Ces complexes sont donc plutôt genrés au masculin, sans doute car ils sont considérés comme des virus (un nom masculin). Par la suite, je comprends que le complexe désigne l'ensemble des composants CRISPR, mais que les virus AAV-Cas9 et AAV9-ARNsg-51 sont les « vecteurs » viraux qui forme le complexe CRISPR-Cas9. Or, le mot « vecteur » est aussi masculin.

« Drive expression »

Le verbe « drive » dans son appréciation générale a plusieurs traductions en français, mais dans un contexte spécialisé (« A muscle-specific creatine kinase (CK) regulatory cassette was used to drive expression of Cas9 »), il a une signification bien spécifique. Je cherche à trouver un verbe qui corresponde, ou si possible un fait de langue figé qui soit équivalent à « drive expression », que ce soit une tournure verbale ou nominale.

Dans *Word Sketch*, je vais d'abord afficher le concordancier à partir de la liste de la catégorie [*verbs with "expression" as object*], et je fais défiler jusqu'à trouver les tournures qui m'intéressent. Je retiens « permet l'expression de Cas9 et de l'ARN-guide » et « contrôler l'expression d'une endonucléase ». Dans les listes des catégories [*verbs with "expression" as subject*] et [*modifiers of "expression"*] je ne trouve rien de pertinent. Dans *Word Sketch Difference*, en comparant les usages de « contrôler » et « permettre », je me rends compte que « permettre » est bien plus adéquat pour parler de « l'expression de la Cas9 ». En faisant une recherche d'occurrence à partir de mes hypothèses personnelles (« mener l'expression » et « influencer l'expression »), je n'obtiens aucun résultat. Bien sûr, cela ne veut pas dire que mes hypothèses n'ont jamais été employées, mon corpus FR n'étant ni parfait ni exhaustif. Cependant, il s'agit d'un bon indicateur d'usage. J'opte donc pour « permettre l'expression ».

L'utilisation de corpus comparables (langue source et langue cible) s'est avérée très utile dans le cadre d'une traduction de type spécialisée. Les usages courants spécialisés sont souvent connus du traducteur aguerri, mais peu de celui qui découvre un domaine en même temps qu'il le traduit. L'analyse des faits de langues dans les corpus grâce aux différentes fonctionnalités de Sketch Engine a permis à la fois de se familiariser avec l'environnement linguistique du domaine technique, mais aussi de valider ou d'invalidier des choix terminologiques et phraséologiques spécifiques.

Cependant, même si les statistiques et les profils lexicaux et syntaxiques présents dans les corpus constituent des pistes solides pour la traduction du texte-support, l'outil comporte son lot de défauts. Ces imperfections viennent notamment de la compilation des corpus, essentiellement basée sur des textes puisés sur Internet, dont la fiabilité n'est pas toujours avérée, que ce soit en français ou en anglais. De plus, de nombreuses sources reprennent par effet d'écho les mêmes textes ou passages, ce qui peut considérablement fausser les statistiques. Du côté du corpus français, bien des documents sont sûrement des traductions de l'anglais. Leur qualité et leur provenance ne permettent pas d'authentifier les usages observés.

Ces problèmes sont d'autant plus présents dans le domaine de spécialisation que j'ai choisi pour mon mémoire, la recherche se faisant essentiellement en anglais, et les sujets étant relativement récents et en constante évolution. Mais toutes les hésitations auxquelles j'ai été confrontée ont aussi constitué une aubaine : à ce stade, si un usage n'était pas encore bien instauré dans la langue, je pouvais prendre une décision autonome et appuyée, et mon choix de traduction dépendait entre autres de ce que je jugeais être le plus judicieux. La partie III.2 « Justification des choix de traduction » éclaire justement ces choix et leur origine.

3. Corpus alignés

En sus des corpus comparables, l'utilisation de corpus alignés s'est avérée clef dans la résolution de certains problèmes de traduction. En effet, leur disponibilité offre un accès direct à des faits de langues pertinents, déjà traduits. L'utilisation de ressources bilingues a l'avantage d'être simple et de nécessiter peu de recherches, convenu que l'on garde un regard critique sur la qualité de ces documents. Le problème déjà mentionné dans la partie « corpus comparables » subsiste : des textes en français qui sont des traductions de l'anglais comportent le risque de ne pas être idiomatiques, et leur qualité ne peut parfaitement attestée. Toutefois, certaines ressources bilingues sont parfaitement exploitables, surtout lorsqu'elles sont spécialisées.

Notons par exemple la profusion d'articles scientifiques dont l'*abstract* (résumé) est bilingue. Le rôle de l'*abstract* a cette double qualité qu'il reprend toutes les notions clefs de l'article auquel il appartient, et ce à un niveau de langue propre au domaine d'expertise concerné, mais aussi qu'il assure un repérage rapide des éléments qu'on en voudrait extraire dans les deux langues. Dans le cas d'un exercice de traduction de l'anglais vers le français, l'idéal est d'avoir accès à des articles initialement rédigés en français (donc très probablement par des chercheurs francophones), et dont l'abstract aurait été ensuite traduit en anglais, par exemple pour permettre de toucher la communauté scientifique plus largement. L'intérêt de ce cas est qu'il assure l'idiomaticité de la version française, tandis que la perfection de l'anglais n'a pas besoin d'être avérée, étant donné que nous sommes déjà en possession des faits de langues anglais que nous voulons vérifier et traduire.

D'autant plus que, comme déjà expliqué, la plupart des publications dans le domaine de la thérapie génique s'effectuant en anglais, il y a de fortes chances que les chercheurs ayant rédigé l'abstract en connaissent les bons usages. À cet égard, certains articles disponibles dans la bibliothèque en ligne de *médecine/sciences* (partie « Résumé » en français, partie « Abstract » en anglais), ou sur *ScienceDirect* (« Résumé »/« Summary »), jouent un rôle clef dans la constitution de corpus alignés.

Nous pourrions aussi mentionner les articles académiques écrits en français mais exclusivement inspirés de ressources anglophones, type thèse ou mémoire et qui, en l'occurrence ici, sont pour la plupart d'origine canadienne. Les ressources canadiennes sont bien évidemment à prendre avec des pincettes, mais dans le domaine des sciences, et surtout d'une discipline aussi jeune que celle de la thérapie génique et l'édition du génome, nous pouvons considérer que les équivalences sont identiques entre le français français et le français canadien. Les ressources utilisées seront explorées plus en détail dans la partie « Bibliographie critique ».

Enfin, il convient de citer les sites en recherche biologique officiels européens, d'associations ou d'organisations internationales, qui regorgent de communiqués et de documents didactiques ou spécialisés, bilingues ou trilingues. Nous pouvons par exemple citer *EuroStemCell*, un « réseau de scientifiques et académiques » qui fournit « des informations validées par des experts indépendants et des outils pédagogiques ayant fait leurs preuves concernant les cellules souches et leur impact sur la société » et financé par « le programme de recherche et d'innovation de l'Union européenne « Horizon 2020 »⁷⁴ ». Notons aussi la pertinence des ressources suisses, souvent multilingues de manière officielle, comme le site de l'Académie suisse des sciences naturelles (SCNAT) qui propose ses ressources en allemand, français, italien et anglais. La SCNAT est une institution de longue date qui promeut le savoir scientifique et les échanges interdisciplinaires⁷⁵.

⁷⁴ Site d'EuroStemCell, disponible en ligne sur www.eurostemcell.org

⁷⁵ Communiqué de presse de la SCNAT, « L'Académie suisse des sciences naturelles façonne la science et la Suisse depuis 200 ans », disponible en ligne sur <https://scnat.ch/fr>

Sa crédibilité à l'échelle nationale et internationale fait de son site une source bilingue fiable.

En conclusion, le déroulé de la traduction a impliqué une grande variété de ressources, beaucoup d'allers-retours, et beaucoup de questionnements, aussi triviaux soient-ils. L'utilisation de corpus comparables et alignés nous a montré qu'aucune ressource n'est mauvaise, si tant est qu'elle soit employée de manière adéquate. Une illustration du processus de traduction par la description détaillée de certains problèmes de traductions et par la justification de certains choix de traduction nous permettra dans un deuxième temps de concrétiser la résolution des problèmes soulevés par le texte-support.

III- Repérage et analyse des principaux problèmes et difficultés de traduction

1. Justification des choix de traduction

Comme nous avons pu le pressentir à la lumière des choix méthodologiques de travail, la traduction a été effectuée majoritairement sous l'angle des usages actés. Les formulations consacrées auront toujours primé sur le souci de justesses syntaxique et grammaticale, même si ces deux aspects fonctionnent ensemble la plupart du temps. Cette considération des usages, écrits ou parlés, va de pair avec une approche de la traduction selon la théorie fonctionnelle du *skopos*, théorie selon laquelle tout texte doit être envisagé selon sa finalité et le public auquel il se destine.

Ainsi, dans le cadre d'une étude scientifique de rigueur, écrite par et pour des spécialistes, chercheurs dans le domaine de la thérapie génique, et qui plus est dans le milieu vétérinaire, la structure a été respectée en tout point, le lexique du domaine expérimental repris avec précision et les formulations abrégées, évidemment, non explicitées et bien souvent calquées. Les exemples donnés dans cette partie III viennent illustrer cette approche où l'utilité a prévalu sur l'esthétisme. Ils reprennent et justifient de manière précise et détaillée des choix terminologiques et phraséologiques.

J'éclairerai seulement ici dans un premier temps deux propositions de traduction pour des termes sans correspondance exacte. Ces deux choix de traduction émanent d'une discussion avec mon référent spécialiste. Le premier concerne la traduction du concept de « reframing », tel qu'employé dans les faits de langues suivants :

« These data indicate that the two Δ Ex50 dogs injected with AAV9-Cas9 and AAV9-sgRNA-51 had a high frequency of reframing events (with cDNA products containing an A insertion in the sequence of exon 51) and exon 51 skipping events resulting from deletion of the highly conserved ESE region. » (p.66, paragr.1)

« Scheme showing the CRISPR/Cas9-mediated genome editing approach to correct the reading frame in Δ Ex50 dogs by reframing and skipping of exon 51 » (p.78, paragr.1)

« Deep-sequencing of the upper band containing the non-edited cDNA product and reframed cDNA products indicated that a mean of 73.19% of total reads contained reframed cDNA products » (p.64, paragr.2)

Il s'agit d'un néologisme qui n'existe pas en français et qui désigne un événement biotechnologique propre à l'édition du génome. Son emploi dans un texte spécialisé montre à quel point le langage de la thérapie génique a dû évoluer concomitamment à l'évolution des avancées biotechnologiques. Il est intéressant de voir que, là où énormément d'anglicismes ou de mots anglais sont repris tels quels dans ce domaine, soit par manque de temps (les équivalences n'ont pas encore parfaitement été mises en place), soit par facilité, il semble que le concept de « reframing » ait échappé à ce phénomène. Il est aussi pertinent de rappeler que le mot a été concrétisé en anglais, alors que la recherche en français est tout aussi innovante dans le domaine de l'édition génique (notons par exemple le travail non négligeable d'Emmanuelle Charpentier, généticienne française).

Pour traduire convenablement ce terme, il convient aussi de noter une différence de sens entre le « reframing » et le « frameshifting », le premier correspondant à un rétablissement du cadre de lecture (qui pourra alors être lu, 3 nucléotides par 3 nucléotides, puis traduit en protéine), le deuxième à une modification ou un décalage du cadre de lecture (qui ne pourra pas forcément être lu si le changement n'aboutit pas en des multiples de 3 nucléotides) (voir dans la partie Exposé II.1 « La mutation génétique).

Dans le cadre de cette traduction, il ne s'agit pas d'inventer de nouveaux mots, puisque la version cible doit être cohérente et adaptée au public. Après vérification auprès de mon référent, je confirme que « reframed cDNA » ne peut se traduire par « ADNc recadré », car « recadré » n'existe pas et ne s'emploie pas ⁷⁶. L'autre possibilité serait de dire « rephasé ». Le terme existe en effet, mais il ne colle pas. L'utiliser constituerait un contresens. Ici, il n'y a donc pas d'autre solution que d'opter pour la périphrase : « les produits d'ADNc avec un cadre de lecture rétabli » (p.66, paragr.2). Bien que plus lourd qu'en anglais, ce fait de langue permet de rester fidèle au sens du texte et juste dans le langage et évite des tournures trop aventureuses qui n'auraient pas leur place ici.

Une périphrase a de même été nécessaire pour traduire « wild-type levels » comme dans les phrases suivantes :

« Western blot analysis confirmed the restoration of dystrophin expression in skeletal muscle (Fig. 2, A and B) to ~60% of wild-type levels » (p.67, paragr.3)

« On average, 2% of wild-type levels of dystrophin were detected in the uninjected contralateral muscles, far more than could be attributed to rare revertant events » (p.66, paragr.3).

En l'état, une telle contraction sémantique n'est pas envisageable en français, et un contournement est nécessaire pour rétablir le sens des phrases en français. En effet, « wild-type » en anglais fait référence à l'état dit « sauvage » d'un individu type, que ce soit au niveau de son phénotype ou de son génotype. On emploie en français le mot « sauvage », plus par fidélité à l'origine historique que par soucis de véracité.

⁷⁶ Tiret, L. Adapté de propos recueillis en juin 2023.

De ce fait, un individu « sauvage » ne correspond pas à un individu qu'on trouverait dans la nature, par opposition à un individu « domestique ». Il ne correspond pas non plus à un individu dont le génotype serait originel, entièrement dépourvu de mutations. D'ailleurs ce dernier n'existe pas. En fait, un individu sauvage correspondrait à un individu dont le phénotype est la médiane de la variabilité d'un phénotype sauvage, mais lui-même a subi des variations mutationnelles ⁷⁷. Il n'est donc pas envisageable de traduire « wild-type levels » par « niveaux sauvages », car ce ne sont pas les « niveaux » qui sont sauvages, mais plutôt, les individus chez lesquels on aurait observé des niveaux (ici en l'occurrence, de dystrophine) moyens ou normaux.

C'est pourquoi j'ai décidé de développer la formulation en français, tout en faisant en sorte qu'elle reste la plus succincte possible : « L'analyse en Western blot a permis de confirmer le rétablissement de l'expression de la dystrophine dans les muscles squelettiques (fig. 2, A et B) à ~60 % du niveau des individus sauvages* » (p.68, parag.3). Cette tournure se comprend mieux que « niveau sauvage » et évite de trop rallonger en disant « niveau présent chez les individus sauvages ».

Ces deux traductions constituent des cas à part dans la globalité de la traduction, puisqu'il s'agit des deux seuls moments où le manque d'équivalence supposait un peu d'inventivité. Toutefois, les choix ont toujours été appuyés par un œil expert et par l'observation des usages.

2. Difficultés de traduction

Les difficultés de traduction auxquelles j'ai été confrontée éclairent mes choix de traduction. Ils sont de nature variée, mais toutefois scindables en deux catégories, qui sont relativement universelles à toute traduction : les difficultés liées à la compréhension du sujet du texte, et celles liées à la reformulation. Ces difficultés sont aussi intrinsèquement liées au domaine de spécialisation et au type de document choisi.

⁷⁷ Ibid

Difficultés de compréhension

Dans un premier temps, même en ayant compris les généralités d'un sujet, il est commun d'être confronté à de nombreux passages techniques obscures. Même avec toute la volonté du monde, je ne suis pas experte en thérapie génique.

Points techniques

Une erreur que j'ai commise et dont je me suis vite rendu compte a été de négliger toute la partie de mon sujet concernant la Dystrophie musculaire de Duchenne, ses aspects médicaux, et physiologiques. Pourtant, cette partie est indispensable pour comprendre de manière exhaustive les résultats de l'étude de mon texte-support, notamment tout ceux ayant rapport à l'histologie. Comprendre les grands principes des différentes méthodes de thérapie génique, et même leurs spécificités, ne suffit pas à appréhender une application concrète, en contexte, qui implique un pan médical crucial.

Certains passages du texte m'ont donc posé un problème de compréhension, or sans compréhension, le risque de faire des erreurs de traduction est considérable. Par exemple, un passage qui a été pour moi difficile à comprendre est le suivant :

« Immunohistochemistry for developmental myosin heavy chain (dMHC), a marker of regenerating fibers, revealed a marked reduction in developmental myosin (dMHC)-positive fibers within injected muscles ». (p.68, paragr.1)

Le sens de cette phrase n'a pas l'air compliqué à comprendre, mais il suppose de connaître un peu plus finement le rôle de la dMHC dans les cellules musculaires. En effet, en toute logique, et si l'on veut prouver que l'expérience effectuée fonctionne, il faudrait que les résultats soient positifs. Tout dans cet article indique que les résultats de l'utilisations de l'outil CRISPR-Cas9 pour soigner la DMD sont plutôt concluants. Mais cette phrase porte à confusion : il y est affirmé qu'en effectuant des analyses, une présence réduite de dMHC a été détectée. Or il nous est aussi expliqué que la dMHC est un marqueur de fibres en voie de régénération. Pourquoi alors une réduction de la dMHC constituerait un résultat positif si l'objectif est d'éviter que les fibres musculaires ne dégèrent ? Voici la question que je pose à mon référent.

« Lors du processus de réparation du muscle, celui-ci fait appel à la cellule initiale chargée de fabriquer le muscle en premier lieu : le myoblaste. Le myoblaste est la cellule souche responsable de la formation des muscles dans l’embryon. La récapitulation de la production de cette cellule souche implique que le muscle réemprunte le même problème que pendant son développement dans l’embryon. Transitoirement, le myoblaste exprime de la dMHC (une myosine développementale), qui n’est normalement pas présente dans les fibres adultes. La dMHC agit donc comme un marqueur de la réparation musculaire en cours, et elle signifie que les cellules souches ont bien fonctionné face à la sénescence précoce des cellules musculaires. Donc s’il y a moins de myosine développementale (dMHC), cela veut dire qu’il y a moins de fibres musculaires en voie de réparation, donc que de manière générale les muscles s’abîment moins. Un muscle qui ne casserait pas, donc qui ne serait pas atteint de dystrophie musculaire, ne présenterait que de la myosine adulte, et non pas développementale. »⁷⁸

Grâce à cette explication, j’ai pu rétablir la logique de cette phrase dans le contexte de l’article, et confirmer que je ne faisais pas de contresens dans ma traduction.

Disponibilité des ressources

Les difficultés liées aux ressources ont pu poser problème pour la rédaction de l’exposé, ou encore comme nous l’avons vu, pour la compilation de corpus (voir partie II.3 « Corpus comparables »), mais aussi pour certains passages à traduire. En effet, c’est en me familiarisant avec la littérature en thérapie génique que je peux intégrer les usages à ma traduction. L’Académie française ou France Terme pourront me donner des pistes, mais ces sources ne valideront jamais un emploi exact dans le domaine étudié. Cependant, la première difficulté de traduction liée aux ressources est que les ressources francophones sont très lacunaires. En génétique et en médecine, la documentation spécialisée et/ou académique ne manque pas. Les sujets fondamentaux sur lesquels la couche de spécialité vient s’appuyer sont faciles à intégrer et à comprendre. Mais quand il en vient aux méthodes de thérapie génique et d’édition du génome, il est difficile de trouver des expériences similaires à celles du texte-support qui aient été décrites en français. La récence relative de cette discipline, surtout appliquée dans le milieu médical, suppose deux mouvements contraires : la profusion de littérature de vulgarisation, et la carence en littérature spécialisée.

⁷⁸ Tiret, L. Propos retranscrits, 2023

En interrogeant certains spécialistes du domaine (un étudiant en médecine à Descartes, un doctorant en thérapie génique à l'ENS et mon référent, enseignant et chercheur en thérapie génique à l'École Vétérinaire de Maison Alfort), j'ai pu confirmer mon a priori qui était que, même si la communauté scientifique francophone travaille en français au quotidien, ou en français et en anglais, la majorité des publications se fait en anglais. Un usage linguistique est peut-être donc bien installé ou consacré au sein de la communauté, mais il est rendu très inaccessible au traducteur qui ne ferait pas partie de cette communauté. Très souvent, mes seuls moyens de correctement vérifier les usages étaient tout simplement de poser la question à un spécialiste !

Le manque de ressources francophone s'est aussi accompagné du problème des ressources obsolètes. La thérapie génique est un domaine qui évolue à une rapidité fulgurante. Pour chaque détail de mon exposé ou de ma traduction, il est difficile de systématiquement trouver des articles qui soient à jour, surtout en français. Par exemple, les statistiques sur le séquençage des génomes n'ont rien à voir entre 2003 et 2023. Entre temps, les progrès informatiques ont augmenté de manière exponentielle la taille des génothèques et l'ampleur des recensements génomiques. La plupart de mes chiffres datent de 2008. Cette obsolescence ne pose pas de problème pour comprendre les méthodes biotechnologiques et génétiques dans leur ensemble, mais elle rend la connaissance des évolutions du domaine, des médicaments de thérapie génique sur le marché, des réglementations, du nombre de maladies génétiques recensées, ou encore de la part du génome humain réellement séquencée très floue. Il est possible de se perdre dans le flot de l'information au sujet des avancées et des découvertes, et il faudra toujours rester vigilant sur les dates de publication. À titre d'exemple, entre le moment de la publication de mon texte-support (2018) et la rédaction de mon mémoire (2023), mon référent m'a informé que certains aspects liés à la sénescence des cellules souches musculaires avaient été fraîchement élucidés, faisant ainsi la lumière sur des résultats donnés dans mon texte-support.

Difficultés de reformulation

Dans un deuxième temps, les problèmes de transfert de langue furent nombreux. Les travaux scientifiques impliquent d'une part une rédaction particulière qui peut paraître inhabituelle à certains endroits, et d'autre part beaucoup de raccourcis d'usage, qui sont propres à tout texte spécialisé. Les difficultés de reformulation sont souvent liées à celles relevant de la compréhension, car c'est parfois les quelques subtilités de sens qui m'échappaient, qui étaient la clef d'une bonne traduction. De plus, j'ai été confrontée en grande partie à la question de l'étoffement, parfois pour des faits de langue restreint. Comme nous le verrons dans cette partie, la question prédominante a été celle de savoir si la concision devait à tout prix primer sur les efforts de justesse de la langue.

Ordres de grandeur et unités

Certaines unités de mesure étant peu communes dans le langage courant, elles nécessitent une réflexion plus poussée. Je fais notamment référence à l'unité utilisée pour quantifier les doses de vecteurs viraux qui sont injectés lors des expériences effectuées.

Ces mentions sont récurrentes : « We tested two doses (2×10^{13} vg/kg and 1×10^{14} vg/kg) of each of the two viruses (AAV9-Cas9 and AAV9-sgRNA-51). » (p.68, paragr.3), ou bien « Systemic delivery of 2×10^{13} vg/kg of each virus (total virus 4×10^{13} vg/kg) in Δ Ex50-Dog-#2A » (p.70, paragr.1), ou encore « For Dog-#2B, which received 1×10^{14} vg/kg of each virus (total virus 2×10^{14} vg/kg), Western blot analysis showed restoration of dystrophin expression » (p.70, paragr.3).

D'après sa première occurrence dans l'article (« intra-muscular (IM) injection with 1.2×10^{13} AAV9 viral genomes (vg) of each virus »), on peut facilement en déduire que l'unité « vg/kg » correspond ici à la quantité de génome viral (*viral genome* – vg) par kilogramme. En sondant nombre d'autres articles spécialisés (recherche par mot clef et utilisation de la fonction contrôle + F), je ne trouve pas de mention faite de « génome viral par kilogramme ». J'en conclus d'abord que cette quantification est exclusivement employée en anglais, ou alors quasi inexistante en français.

Cependant, en cherchant « viral genome per kilogram », je trouve par association beaucoup d'occurrences pour « vector genomes (vg) » et « vector genomes per kilogram », à la fois dans mon corpus EN sur *Sketch Engine*, et sur des bibliothèques en ligne spécialisées (*American Society for Biology*, *Oxford Academic*, *ScienceDirect*, etc.). Je peux supposer qu'il y a une légère erreur dans mon texte-support, ou du moins un usage qui diffère. Les doses injectées sont quantifiées en utilisant une quantité de génomes de vecteur, quel que soit le vecteur (voir partie Exposé III.2 « Vecteurs et ciblage »), par kilogramme. Subsiste la question de savoir si cette unité de grandeur se traduit.

Ma première théorie est que « vg/kg » ne devrait pas s'employer en français puisqu'en parlant de « génome de vecteur par kilogramme », on l'écrirait « gv/kg ». Seulement, cette unité, comme la plupart des faits de langues de mon article, ont pour origine l'anglais, et il est plus que probable qu'elle soit laissée en anglais dans les textes français. Mais mes recherches de ce côté-là ne sont pas concluantes. Je trouve cependant des mentions de « vecteurs génomiques/kg⁷⁹ » et de « vecteur/kg⁸⁰ » dans des notices de posologie médicamenteuse et des études scientifiques, qui plus est, avec des ordres de grandeurs similaires aux chiffres de mon texte-support (1×10^{14} , 2×10^{14} , etc.). Ces mentions étant rares et leurs sources non francophones à l'origine, je n'exclue pas de m'écarter de cette traduction pour mon texte. En effet, ayant considéré tous ces paramètres, je demeure résolue dans la volonté de rendre ma traduction la plus concise possible, quitte à ce qu'elle ne soit pas parfaitement francisée. Je préférerai donc les options « vecteur/kg », pour la simplicité et la clarté, ou bien une non-traduction (« vg/kg »), pour la fidélité et la concision.

Spécificités et précision de la terminologie scientifique.

La notion de précision est consubstantielle au langage scientifique. Chaque mot compte : un « contenu » n'est pas une « quantité », qui n'est pas une « concentration ».

⁷⁹ *The Medical Letter on Drugs and Therapeutics*. Édition française (traduction : Dr. Pierre-Alain Plan), 2019. Vol. 41, No 17

⁸⁰ Rossi, A., Salvetti, A. Intégration des vecteurs AAV et mutagenèse insertionnelle. *Médecine/sciences*, 2016. Vol. 32, No 2, pp. 167-174

Ainsi, j'ai été confrontée à plusieurs cas où plusieurs équivalences françaises pour un seul terme anglais donnaient l'impression de se valoir, mais un examen plus approfondi prouvait que ce n'était pas le cas. Par exemple, quelle serait la différence entre « étudier », « examiner », « mesurer » et « analyser » ? Ou la différence entre « envoyer », « transmettre », « administrer » et « injecter » ? Au fur et à mesure du travail de traduction, nous observerons ainsi que la protéine de la dystrophine n'est pas « grande », mais « grosse ». En effet, ce dernier terme permet de sous entendre que cette protéine est « lourde, en kDalton, ce qui est le cas...L'une des plus lourdes, 427 kDa ! »⁸¹. De même qu'on ne parle pas de « sûreté », mais « d'innocuité », pas de « comptage », mais de « numération », pas de « livraison », mais « d'administration ». Le sens du détail est donc réellement important. Pour finir, nous pourrions noter que la phraséologie scientifique interdit de dire « les échantillons de muscles squelettiques » et il faut préférer « les échantillons issus de muscles squelettiques » (p.72, paragr.1) car c'est la biopsie qui correspond à un prélèvement. Dans la définition biologique et juridique, toute molécule issue de ce prélèvement devient un échantillon. Le sang est un prélèvement ; les cellules sanguines ou le plasma (le sang moins les globules rouge), un échantillon. Ici, une PCR est faite à partir de l'ADN extrait des cellules du muscle prélevé (biopsie), c'est donc bien un échantillon issu de muscles⁸².

Anglicismes et absence de traductions consacrées

Dans la majorité des cas, la recherche terminologique a été relativement simple du fait de la transparence de la plus grande partie des termes employés dans l'article, comme en attestent d'ailleurs lexiques et glossaires (voir Partie 4 « Analyse terminologique»). La similitude morphologique des termes anglais et français est due à trois cas de figure : le lexique de la biologie est majoritairement fondé sur une étymologie latine, reprise par l'anglais et par le français ; les mots les plus récents ont souvent l'anglais pour origine, et donc ils sont adaptés en français depuis l'anglais, ou bien adoptés tels quels en français ; enfin, une partie des faits de langue correspond à des abréviations qui n'appartiennent officiellement à aucune langue.

⁸¹ Tiret, L. Commentaires de traduction. Juillet 2023.

⁸² Tiret, L. Commentaires de traduction. Juillet 2023.

Cependant, cette transparence comporte aussi des dangers. Certains termes faisant l'objet d'une fiche terminologique montrent d'ailleurs à quel point la transparence peut être trompeuse, car elle implique de potentielles polysémies et entraîne parfois un manque de spécificité ou tout simplement un glissement ou une erreur de sens. Ainsi, plusieurs passages étaient pour moi trompeurs car ils paraissaient simples à comprendre.

C'est le cas du passage suivant : « In ΔEx50 mice, dogs, and DMD patients, these proteins are destabilized and do not localize to the sub-sarcolemmal region. » (p.68, paragr.2)

L'anglais ici comporte plusieurs problèmes pour un francophone. Le premier est l'emploi de la préposition « to » dans « do not localize to the sub-sarcolemmal region. ». En effet, le verbe « localize » est habituellement accompagné d'une préposition lorsqu'il est sous sa forme participiale (« localized within » ; « localized at », etc) mais moins lorsqu'il est sous sa forme infinitive, comme ici. Cependant, il est intéressant de voir grâce à *Sketch Engine* que dans le corpus *English Web 2021*, « localize » dans les phrases prépositionnelles est le plus souvent accompagné de la préposition « to » (20 344 occurrences, contre 11 013 avec « to » en deuxième position). De plus, quelle que soit la préposition avec laquelle ce verbe est employé, les co-occurrences sont surtout employées dans les domaines de la science et de la médecine. Nous pourrions en comprendre qu'il s'agit d'un fait de langue spécialisé.

La deuxième difficulté de ce passage se trouve dans l'épithète « sub-sarcolemmal » qui qualifie « region ». Sans confondre « sarcolemme » avec « sarcomère » (le premier étant la membrane cellulaire du myocyte, le deuxième l'unité de base des myofibrilles, toujours dans le myocyte), j'ai pu comprendre qu'il s'agissait de la région située en-dessous du sarcolemme. Mais il a fallu bien distinguer la région sarcolemme de la région sarcomère, car le sens aurait été différent, même si à l'échelle cellulaire, les deux sont très proches d'un point de vue spatial. Ensuite, il a fallu me décider entre « sous-sarcolemmique » et « subsarcolemmale », les deux n'existant pas officiellement mais étant plus ou moins employés, à la fois dans des thèses francophones et dans des articles scientifiques⁸³.

⁸³ Dias Wilson, V. Adaptations du métabolisme musculaire en réponse à l'exercice et à une supplémentation en antioxydants chez des patients atteints de Dystrophie Facioscapulohumérale. Thèse en Biologie santé. Montpellier : Université de Montpellier, École doctorale Sciences chimiques et biologiques pour la santé, 2015. 256 p. ; Myopathie avec agrégats tubulaires. Encyclopédie des maladies rares, avec les gènes associés. Portail des maladies

De façon similaire, j'ai dû faire face à certains termes intraduisibles, dans le sens où il n'y avait pas d'équivalent acceptables, ou bien où l'anglais avait été communément adopté en français. C'est le cas de l'emploi de « Western blot » et des « indels ». Par exemple dans :

« Western blot analysis confirmed the restoration of dystrophin expression in skeletal muscle » (p.71, paragr.3) ou dans « To investigate the proportions of various indels generated by systemic delivery of AAV9-Cas9 and AAV9-sgRNA-51, we performed amplicon deep-sequencing analysis of the genomic DNA from heart, triceps and biceps muscles. » (p.71, paragr.2)

D'un côté, il est facile de trouver des explications sur ce qu'est une analyse Western blot, et par conséquent d'observer l'utilisation de ce terme en contexte, en anglais ou en français. De l'autre, il n'est pas facile de savoir s'il existe un équivalent français qui existe. En outre, il faut noter que « blot », qui signifie « buvard » en anglais, permet de comprendre que le terme « Western Blot » désigne la deuxième étape de la méthode de transfert (voir dans la partie Exposé IV.1 « Comprendre et interpréter des résultats scientifiques »), mais que par abus de langage, il est employé pour parler de l'intégralité du procédé⁸⁴.

Ici, la difficulté se situe dans le fait que bien d'autres types d'analyses du même type ont un terme consacré, en anglais aussi bien qu'en français. Cependant, la « Western blot » ne semblait pas être traduite (elle avait bien l'équivalent « immunotransfert », mais dont l'usage n'était pas attesté dans les textes spécialisés non pédagogiques), ce qui rendait son utilisation en français plus complexe : devrait-on dire « l'analyse Western blot », « la Western blot », « le test de Western blot », « l'analyse par Western blot », « le Western blot » ? Les variantes sont sans fin, et elles sont toutes plus ou moins présentes dans les travaux scientifiques rédigés en français, que ce soit par des francophones ou des noms francophones. Une telle souplesse dans la formulation constitue une difficulté majeure pour un traducteur.

rare et des médicaments orphelins [en ligne]. Disponible sur orpha.net (consulté en juin 2023)

⁸⁴ Western Blot. Article [s.d.]. Ooreka [en ligne]. Disponible sur premiers-secours.ooreka.fr (Consulté en juin 2023)

En effet, n'étant pas spécialiste lui-même, et ne pouvant vérifier des usages plus communs que d'autres, ou plus logiques, ou plus grammaticalement corrects, il ne peut faire un choix de traduction qui soit plus judicieux qu'un autre. Cependant, cet exemple prouve aussi que le langage scientifique, bien que rigoureux et stricte dans certains emplois, peut aussi être moins rigide que la langue courante, et donc que la traduction de tels faits de langue n'implique pas de trop grands risques. Par conséquent, ma manière de procéder a été de préciser à la première occurrence qu'il s'agissait d'un type d'analyse histologique « L'analyse en Western Blot » p.68, paragr.3), puis de choisir par la suite l'expression la plus simple et la plus fluide, ce qui permettait ainsi au syntagme de bien s'intégrer au reste de la phrase dans laquelle il se situait. Par exemple : « Le Western blot a confirmé le rétablissement de l'expression de la dystrophine du muscle squelettique ». (p.72, paragr.3)

Le mot « indel » est encore à part puisque ses occurrences dans les corpus sont déjà plutôt rares en anglais, ce qui permettait mal, dans un premier temps, d'identifier sa définition exacte. De fait, ce mot n'existe pas dans les dictionnaires et, sans pour autant en avoir l'apparence, il constitue un mot-valise, construite à partir des mots anglais « insertion » et « délétion » (INsertionDELetion). Ces deux noms caractérisent des cas de figure observables lorsque l'enzyme Cas9 induit une coupure de l'ADN, et qu'il y a recombinaison du brin, par ajout ou retrait de nucléotides. Le fait d'amalgamer ces deux processus, pourtant inverses, permet de désigner en un seul terme les effets ou l'absence d'effet d'une coupure de l'ADN par l'outil CRISPR. L'utilisation de ce terme est donc spécifique à l'analyse des méthodes de thérapie génique, bien que les termes « insertion » et « deletion » se retrouvent aussi de manière plus générale dans le domaine de la biologie cellulaire, en anglais comme en français (« insertion » et « délétion »). Nous pourrions considérer qu'un mot-valise du même genre en français se fasse selon un autre mode de découpe, ou bien avec un accent sur le [e] de « indel » pour reprendre l'accent dans « délétion », mais il faut une fois de plus se rappeler que toutes ces dénominations sont relativement récentes et tirées, la plupart du temps, de l'anglais. C'est pourquoi le terme en anglais est repris tel quel en français. Mais alors, quel genre lui attribuer ?

Ce qui est surprenant ici est que « indel » reprend deux mots féminins, pourtant, il est employé au masculin, peut-être car il s'agit d'un marqueur ⁸⁵.

En conclusion, laisser un mot en anglais dans la traduction suppose une réflexion sur la structure syntaxique et grammaticale française construite autour de ce mot, aspects qui sont parfois plus difficiles à gérer que de trouver une simple équivalence française.

Formes abrégées

Les faits de langues qui m'ont de même posé problème étaient relatifs aux abréviations, majoritairement utilisées pour étiqueter des sujets de l'expérience, un type de dose employé, un échantillon analysé, etc. : tout ce qui était présent en plusieurs exemplaires et qui devait être différenciable. Ma première question concernait le fait de savoir si je devais garder certaines abréviations anglaises si le terme abrégé avait un équivalent en français. Par exemple, pour « A total of three potential genome-wide off-target sites (OT1 to OT3) » (p.66, pragr.2), sachant que « OT » correspond aux initiales de « off-target », que j'ai traduit par « hors-cible », dois-je remplacer les acronymes « OT1 » et « OT3 » par « HC1 » « HC3 » ? Dans ce cas-là, il m'apparaissait trop confus de traduire un acronyme si court et compréhensible, surtout sachant qu'il n'était employé qu'une fois dans l'article. En revanche, mon raisonnement a été inverse pour les occurrences telles que « sgRNA-51 » dans « For the in vivo delivery of Cas9 and sgRNA-51 to skeletal muscle and heart tissue in dogs » (p.62, paragr.5). Ici, j'ai pensé que la désignation « ARNsg » était couramment employée telle quelle pour parler d'un ARN guide simple, et que, sa version anglaise n'étant pas sous sa forme abrégée dans les appellations du type « AAV9-sgRNA-51 », je pouvais me permettre d'utiliser une version française, comme « AAV9-ARNsg-51 ».

Ces réflexions ne sont pas si évidentes et ne vont pas forcément de soi.

⁸⁵ Ben Ayed, R., Mohamed Ali Triki, MA., Rebai, A. Identification de marqueur 'InDel' au niveau du gène Cu-Zn-superoxydodismutase (SOD) par des approches moléculaires et bioinformatiques. *Journal of Arid Arboriculture and Olive Growing*, 2022. Vol. 1, No 2, pp 1-8

Ainsi, mon dernier exemple est d'autant plus intéressant qu'il s'agit d'une désignation, donc peut-être qu'il ne faut pas traduire le contenu afin de rester fidèle à la structure, de même que nous ne traduirions pas un prénom.

En effet, deux chiens atteints de la DMD et traités avec les composants de CRISPR-Cas9, mais avec des doses différentes, sont désignés ainsi : « ΔEx50-Dog-#2A » et « ΔEx50-Dog-#2B » (p.70, paragr.1). Admettons que ces deux désignations constituent des références. Si je les traduits respectivement par « Chien-ΔEx50-#2A » et « Chien-ΔEx50-#2B », est-ce que je change l'étiquetage de l'étude, et donc d'une certaine manière, fausse les résultats ? En effet, si l'on considère que la structure de la rédaction doit absolument être conservée à l'identique dans la traduction pour des raisons de fidélité à la rigueur du protocole expérimentale, il faudrait de même être fidèle à toutes les références de l'expérience.

Seulement, et dans le contexte bien précis de cet article, il est souvent fait mention des sujets de l'expériences de manière détournée ou bien plus souple que l'étiquetage stricte cité ci-dessus. Par exemple, les mêmes sujets traités sont appelés « dogs », « ΔEx50 dogs », « ΔEx50 dogs injected with AAV9-Cas9 and AAV-sgRNA-51 », « treated dogs », et même, par extension, « injected muscles ». D'une certaine manière, ces nombreux exemples prouvent que la rigidité de l'étiquetage des outils de l'expérience (sujets, doses, échantillons), n'est pas absolument primordiale si tant est que l'on comprenne lequel est concerné exactement, par opposition à ses semblables. Dans le cas des sujets, il faudra par exemple bien faire la différence entre un chien malade non traité (chien témoin), un chien malade traité (avec une dose plus ou moins élevée) et un chien au phénotype sauvage (chien sain). Il m'a donc semblé qu'à partir du moment où cette classification était strictement respectée, afin éviter toute confusion ou erreur dans les résultats, il m'était possible de traduire « ΔEx50-Dog-#2A » par « Chien-ΔEx50-#2A ».

Raccourcis d'usage

L'omission de prépositions est propre au langage scientifique. Est-ce par négligence ? Par souci d'allègement des phrases déjà lourdes ? Des tournures techniquement incorrectes doivent-elles alors être conservées à certains endroits afin de rester fidèle au langage scientifique ?

Et peut-on considérer qu'elles ne sont pas incorrectes s'il s'agit de tournures de type « raccourcis d'usage » ? Ces raccourcis d'usage ou de dénominations reprennent d'ailleurs les difficultés évoquées dans le cadre des abréviations : il faut savoir établir la limite entre un étiquetage figé et rigoureux, et une simple dénomination raccourcie permettant de simplifier le texte. Ainsi, j'ai dû me demander si je pouvais me permettre de parler de « mutations DMD » (p.63, paragr.1), en reprenant le modèle anglais, ou bien s'il fallait rétablir l'intégralité du propos se cachant derrière ce syntagme court en disant par exemple « mutations responsables de la DMD ». Il en va de même pour « DMD patients » (p.74, paragr.2), qui peut être traduit simplement par « patients DMD », ou bien plus correctement par « patients atteints de la DMD ». Dans un article de presse, on ne dirait pas « un patient DMD », mais dans une étude scientifique, nous pouvons considérer qu'une fois de plus il s'agit d'une désignation figée pour parler d'un sujet.

De plus, la question de l'absence de préposition vaut aussi pour des faits de langues qui ne sont pas nécessairement figés. Ainsi, « post injection », comme dans « we collected blood samples the day before injection and then at 1, 2, 4, 6 and 8 weeks post-injection. » (p.72, paragr.2), ne désigne aucun « outil » de l'étude, que ce soit un sujet de l'expérience, un échantillon ou une dose, puisqu'il s'agit d'un groupe adjectival et non pas d'un sujet. Il faudrait donc le traduire par « après l'injection » ou « après une injection » et non pas par « après injection ». Pourtant, en observant les co-occurrences dans *Sketch Engine*, et dans mon corpus FR spécialisé, je trouve une seule occurrence pour les deux premières propositions, et sept pour la dernière (provenant de sources différentes à chaque fois).

Les difficultés posées par les raccourcis d'usage montrent l'importance de s'adapter au jargon qui imprègne un texte. La nécessité de respecter certains de ces raccourcis émane d'usages officiels mais rendus corrects par leur emploi prononcé. De plus, ils permettent de simplifier les phrases, voire de les rendre plus claires car plus essentielles.

Étoffement des syntagmes anglais

J'ai aussi été amenée à faire face à un problème inverse à celui du raccourci. De manière plus générale, une traduction de l'anglais vers le français confronte souvent le traducteur au problème des syntagmes nominaux anglais parfois interminables, ceux-ci étant composés de nombreuses appositions de mots et dont l'absence de préposition fait douter des hiérarchies internes. Dans le cadre d'une traduction technique, le problème est davantage important qu'il suppose de devoir délimiter les termes présents à l'intérieur des syntagmes, si terme il y a (voir la partie II.2 « Corpus comparables : savoir si « human induced pluripotent stem cells » est un terme »). Il est pourtant parfois nécessaire d'étoffer ces syntagmes, déjà longs, non pas pour qu'ils deviennent plus explicites en français, mais simplement grammaticalement corrects. Ce cas de figure est moins courant dans ce type de texte, qui suppose d'être concis, mais rallonger les syntagmes s'avère inévitable en français, si l'on veut que les phrases soient compréhensibles et que les hiérarchies grammaticales soient conservées.

J'ai ainsi déjà mentionné le cas des « vecteurs AAV9-Cas9 et AAV9-sgRNA-51 » dans la partie II.2 « Corpus comparables : déterminer l'article correct ». L'ajout du mot « vecteur » permet d'éviter de réfléchir au genre des abréviations utilisées, mais aussi de regrouper les deux composants de CRISPR, tels qu'ils le sont dans les doses injectées.

Mais j'ai aussi eu du mal à éviter certains ajouts dans d'autres cas. Une fois de plus, où dois-je placer le curseur entre le grammaticalement correct et le grammaticalement concis ? Ainsi, devais-je simplement traduire « injected muscles », comme dans « the majority of fibers within the injected muscles expressed sarcolemmal dystrophin, albeit to varying levels. » (p.66, paragr.3), par « muscles injectés » ? Là où certains raccourcis me semblaient justifiés car ils n'empiétaient pas sur le sens (« après injection » au lieu de « après l'injection », « patient DMD » au lieu de « patient atteint de la DMD »), ici l'absence d'étoffement aurait constitué une erreur. Injecter est un verbe transitif, et son participe, même adjectivé comme ici, ne peut être employé sans complément.

Ces dernières difficultés de traduction mentionnées (étoffements nécessaires des syntagmes nominaux) montrent qu'il est difficile d'éviter le foisonnement en français, dans la plupart des cas.

Pour éviter le foisonnement, peut-on se permettre d'adapter la langue pour faire primer des phrases succinctes et ne conservant que l'essentiel sémantique ? Étant donné le public de cet article, l'on pourrait opiner que oui, car l'information se doit d'être brute. La beauté et la perfection de la langue n'est pas primordiale dans ce contexte. En revanche, le texte doit rester compréhensible, extrêmement clair, et garder sa structure, pour éviter le risque d'introduire des erreurs liées au protocole expérimental. Il faut donc, quoi qu'il arrive, conserver la logique interne des phrases, les conjonctions utilisées et les hiérarchies entre chaque groupe grammatical, tout en se permettant certains raccourcis terminologiques et phraséologiques, quoi qu'ils puissent ne pas être parfaitement grammaticalement corrects.

Cependant, l'avantage d'un texte rigoureux dans sa structure est que la fidélité de sa traduction implique très peu de remaniements à grande échelle (par exemple, sur un paragraphe ou une phrase entière), voire pas du tout. C'est pourquoi la plupart des difficultés de reformulation étaient focalisées sur des faits de langue relativement courts, et dont l'incidence était limitée à des portions de texte retréintes.

En conclusion, la stratégie de traduction adoptée dans le cadre de cet exercice a tenté de suivre une seule et même ligne directrice, celle de la concision et du respect de la structure du texte, qui ne se fasse pas au détriment d'une déformation de la langue française. Bien sûr, dans les faits, les exceptions sont nombreuses et elles montrent que les intuitions du traducteur, son niveau d'expérimentation et sa capacité à appréhender les enjeux du texte, font partie intégrante du processus de traduction.

Partie 4 : ANALYSE TERMINOLOGIQUE

Fiches terminologiques

Le choix des fiches terminologiques s'explique par les termes dont certaines ambiguïtés doivent être radiées, au sein des domaines de la thérapie génique et/ou de la biologie moléculaire contemporaine. Certains termes, dont la transparence de l'anglais au français ne semble pas justifier leur étude, trouvent leur intérêt dans les subtilités de leur usage, dans leur polysémie et dans les risques de confusion qu'ils entraînent parfois. La constitution de ces fiches se fonde sur l'analyse des usages à travers des corpus de textes spécialisés et généraux, mais aussi la réutilisation de glossaires et de dictionnaires reconnus, et quelques discussions avec mon référent.

Cinq paires de fiches terminologiques sont ici présentées, chacune composée d'une fiche française et d'une fiche anglaise, pour les équivalents d'un même terme. Les champs suivants composent tout ou partie des fiches :

VE : vedette (terme et ses synonymes). EN pour les termes anglais, FR pour les termes français.

EN : English

FR : Français

DF : définition

DOM : domaine

CTX : emploi du terme en contexte

COL : collocations

ID : Identification de l'auteur

NT : notes

EXP : compléments d'informations utiles

USG : marques d'usage

GRM : indications grammaticales

ETY : étymologie

DER : mots dérivés

HOM : Homonymes

ANT : antonymes

SPE : termes spécifiques subordonnés au terme vedette

GEN : termes génériques superordonnés au terme vedette

REL : renvoi associatif à d'autres termes

RF : sources. SEC RF pour une source modifiée

Fiche terminologique n°1

VE EN genome editing [1] gene editing [2]
DF: Method that lets scientists change the DNA of organisms in vivo, by adding, removing or altering genetic material at particular locations in the genome.
DOM: cell biology, genetics, medicine
CTX: An alternative strategy to systemic delivery is direct injection to a specific tissue, which largely avoids editing other unintended tissues and organs. However, direct injection results in genome editing of relatively few cells within a localized space and is practical only for organs directly accessible to such injection.
COL: <i>n.</i> * technology; * tool; CRISPR *; <i>adj.</i> multiplex *; therapeutic *
ID: ESIT MEM23 EDO
Notes: EXP: By extension, “genome editing” corresponds to the biotechnologies used to edit the genome, such as TALENs or Zinc finger nucleases. USG: although “gene editing” is widely used, “genome editing” is more proper as it allows for a more general approach of the method. A gene specific editing would imply that the edition can only occur in gene regions of the DNA. However, the technologies employed are not restricted to genes and can act in any region of the genome, although gene regions are most commonly targeted. DER: genome editor GEN: gene therapy REL: end joining; recombination; indel; genomic cut
RF: What is genome editing?. Article last updated in 2019. National Human Genome Research Institute [online]. Available at genome.gov [SEC DEF]; What are genome editing and CRISPR-Cas9?. Article last updated in 2022. National Library of Medicine. NIH [online]. Available at medlineplus.gov [SEC EXP]; Wang, JY., Doudna, AJ. CRISPR technology: A decade of genome editing is only the beginning. Science, 2023. Vol. 379, No. 6629, p.7 [CTX]

VE FR édition du génome [1] modification du génome [2] édition génomique [3] édition génique [4]

DF : Modification ciblée du génome qui consiste à ajouter, éliminer ou modifier une ou plusieurs bases dans une séquence d'ADN, procédant ainsi à une réécriture du matériel génétique afin de corriger une ou plusieurs mutations *in vivo*.

DOM : biologie cellulaire, génétique, médecine

CTX : Nous avons donc développé des outils innovants, sûrs et efficaces pour faire de la thérapie génique non additive. Ils consistent en une expression transitoire par un plasmide de la Cas9 et de ses ARNg qui permettent, en présence d'une forte concentration d'oligonucléotide simple brin homologue à la séquence cible, d'obtenir une édition du génome définitive, et sûre.

COL : *v.* permettre l'* ; montrer une * ; révéler une * ; obtenir une * ; *gn.* technique de l'*

ID : ESIT MEM23 EDO

Notes :

EXP : Par extension, le terme « édition du génome » désigne l'ensemble des techniques biotechnologiques de modifications du génome, telles que les méganucléases, les nucléases à doigts de zinc, les TALENs ou CRISPR.

USG : selon la délégation générale à la langue française et aux langues de France (DGLFLF), le terme « édition du génome » est une désignation en français insatisfaisante car l'emploi de « édition » est impropre dans ce contexte. Il devrait être officiellement remplacé par « réécriture du génome ». L'usage contemporain est pourtant catégorique : l'emploi du nom « édition », pourtant un anglicisme, est rentré dans le langage courant ainsi que celui des spécialistes.

GEN : thérapie génique

REL : insertion ; délétion ; substitution ; cassure ; recombinaison

RF : Consultation citoyenne sur l'édition du génome en île-de-France. Faculté de médecine de l'Université Paris-Saclay, 2021, p. 3 [SEC DF] [SEC EXP]; Quillot, E. Propos tirés de « La DGLFLF et le dispositif d'enrichissement de la langue française ». Conférence du 20 janvier 2023. Ministère de la culture et DGLFLF [USG] ; Boutin, J. Alternative à la thérapie génique additive par utilisation du système CRISPR-Cas9 pour la correction génique des maladies héréditaires du globule rouge. Thèse en biologie médicale. Bordeaux : Université de Bordeaux, U.F.R. des Sciences médicales, 2017 Médecine humaine et pathologie, 2017, 95 p. [CTX]

Fiche terminologique n°2

VE EN annealing [1] hybridization [2] binding [3]
DF: A process in which two complementary or nearly complementary single-stranded DNA and/or RNA molecules bond together to form a double-stranded molecule.
DOM: cell biology, genetics, biochemistry
CTX: (...) in the presence of HIV-1 NC, annealing of a single-stranded DNA molecule to a complementary DNA strand that would yield a more stable double-stranded product was favored over annealing to alternative complementary DNA strands that would form less stable duplex products.
COL: <i>n.</i> * site; strand *; single-strand *; DNA *
ID: ESIT MEM23 EDO
Notes: EXP 1: The annealing process can either occur naturally within the cells, or can be experimental, as it is the case for polymerase chain reactions, where the annealing of the primers to the DNA is due to a change in temperature. USG: [2] and [3] also commonly used but do not constitute terms per se. DER: <i>v.</i> to anneal HOM: By extension, annealing corresponds to the step following denaturation during the PCR, in which the primers hybridize with the DNA strand that is to be amplified. ANT: denaturation GEN: pairing
RF: Hybridization. Glossary of Genomic and Genetic Terms, last updated in 2023. National Human Genome Research Institute [online]. Available at genome.gov/genetics-glossary [SEC DF]; Morrical, SW. DNA-Pairing and Annealing Processes in Homologous Recombination and Homology-Directed Repair. <i>Cold Spring Harbor Perspectives in Biology</i> , 2015. Vol.7, N°2 [SEC DF]; Tsuchihashi, Z., Brown, PO. DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. <i>ASM Journals. Journal of Virology</i> , 1994. Vol.68, N°9 [CTX]

VE FR hybridation [1] fusion [2]

DF : processus de jonction de deux séquences complémentaires d'acides nucléiques simple brin et qui conduit à la formation d'un double brin ou duplex.

DOM : biologie cellulaire, biologie moléculaire, génétique, biochimie

CTX : En principe, 20 paires de bases (pb) sont plus que suffisantes pour être spécifiques d'un site cible unique dans le génome humain. Une hybridation parfaite de l'ARN et de l'ADN n'est cependant pas nécessaire pour qu'il y ait coupure par la Cas9.

COL : site d'* ; * de l'ADN et de l'ARN ; * des nucléotides ; étape d'* ; température d'*

ID : ESIT MEM23 EDO

Notes :

EXP 1 : L'hybridation est basée sur le principe de complémentarité des bases nucléiques entre deux brins d'ADN (A avec T et C avec G) ou entre un brin d'ARN et un brin d'ADN (A avec U et C avec G).

EXP 3 : Par extension, l'hybridation dite « moléculaire » correspond à une technique biotechnologique permettant de mettre en évidence une séquence nucléotidique.

EXP 4 : par extension, l'hybridation désigne la combinaison de tout ou partie de deux séquences d'ADN complémentaires.

DER : v. hybrider

HOM : Par extension, l'hybridation, ou renaturation, correspond à l'étape du processus de PCR suivant l'étape de dénaturation et lors de laquelle les amorces reconnaissent et se fixent à leurs séquences complémentaires en reformant des liaisons hydrogène.

ANT : dénaturation

REL : amorce ; séquence ; nucléotide ; complémentarité

RF : Les outils de génétique moléculaire. Réseau d'Enseignement en Génétique, Université de Tours, 2008 [SEC DF] ; Hybridation. Lexique mis à jour en 2021. Aquaportail [en ligne], disponible aquaportail.com [SEC DF] ; Tremblay, JP. CRISPR, un système qui permet de corriger ou de modifier l'expression de gènes responsables de maladies héréditaires. *Médecine/sciences*, 2015. Vol. 31, No 11, pp. 1014-1022 [CTX] ; La PCR (polymerase chain reaction). École Doctorale des Sciences de la Vie et de la Santé ULP, Strasbourg [en ligne]. Disponible sur ed.vie-sante.unistra.fr [SEC EXP 2]

Fiche terminologique n°3

VE EN cut [1] break [2] cleavage [3] nick [4]
DF: In vivo induced cutting of a DNA sequence at a targeted location by disruption of the phosphodiester bond between two consecutive nucleotides in a DNA strand.
DOM: cell biology, molecular biology, genetics
CTX: Cas9 coupled with each of these sgRNA-51 sequences introduced a genomic cut only in DNA of the respective species.
COL: <i>ng.</i> single *; *ting site; * site; * region; double-strand *; genomic *; DNA* <i>v.</i> induce a *
ID: ESIT MEM23 EDO
Notes: EXP: The first step of any genome editing technic is to induce a cut at a targeted site in the genome. The following step is either a non-homologous end-joining repair (gene knock-out), or a homology-directed repair (gene knock-in). DER: cutting ANT: join; end-joining; ligation GEN: editing REL: Non-homologous end-joining (NHEJ); homology-directed repair (HDR); rearrangement; indel; frameshift; reframing
RF: What are genome editing and CRISPR-Cas9?. Article last updated in 2022. National Library of Medicine. NIH [online]. Available at medlineplus.gov [SEC DF]; Komor, AC., Badran, AH., Liu, DR. CRISPR-based technologies for the manipulation of eukaryotic genomes. <i>Cell</i> , 2017. Vol.168, pp.20-36 [SEC DF] [EXP]. Amoasii, L., Hildyard, JCW., Li, H., et al. Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. <i>Science</i> , 2018. Vol.362, N°6410, pp. 86-91 [CTX]

VE FR coupure [1] cassure [2] césure [3]

DF : coupure induite d'une liaison phosphodiester entre deux nucléotides adjacents sur un des deux brins d'acides nucléiques.

DOM : biologie cellulaire, biologie moléculaire, génétique

CTX : **1.** L'inactivation génique a été réalisée par création d'une coupure double brin dans un exon du gène du CCR5. **2.** Une fois linéarisée, la matrice donneuse peut s'intégrer au niveau du site génomique de coupure.

COL : *gn.* site de * ; site de la * ; * de l'ADN ; * du brin ; * unique ; * monobrin ; * double brin ; * génomique ; * hors cible. *v.* générer une * ; induire une *

ID : ESIT MEM23 EDO

Notes :

USG : La désignation principale fournie par France terme est « cassure ». Toutefois, l'usage le dément.

DER : *n.* découpage ; *adj.* coupé

GEN : édition

REL : recombinaison homologue ; recombinaison non homologue

RF : Cassure d'un brin. Définition mise à jour en 2000. France Terme, Ministère de la Culture [en ligne]. Disponible sur culture.fr/franceterme/terme, [SEC DF] [SEC USG] ; Boutin, J. Alternative à la thérapie génique additive par utilisation du système CRISPR-Cas9 pour la correction génique des maladies héréditaires du globule rouge. Thèse en biologie médicale. Bordeaux : Université de Bordeaux, U.F.R. des Sciences médicales, 2017 Médecine humaine et pathologie, 2017, 95 p. [CTX 1] ; Cullot, G. Génotoxicité des systèmes CRISPR-Cas9. Thèse en biologie cellulaire et physiopathologie. Bordeaux : Université de Bordeaux, École doctorale Sciences de la vie et de la santé, 2019, p. 97 [CTX 2]

Fiche terminologique n°4

VE EN exon skipping [1]
DF: Technic aiming to restore the production of a rescued protein by removing the mutated exons during the splicing of the mRNA.
DOM: cell biology, molecular biology, biotechnology
CTX: 1. In patients DL 363.2 and DL 589.2, the targeted exon skipping enlarged the deletions to span exons 44–54 and 50–55, respectively. 2. As illustrated in Fig. 2, additional skipping of exon 51, induced by an exon-51-targeted oligonucleotide, restores the reading frame.
COL: <i>ng.</i> multi-*; * therapy; single *; * approach; antisense-mediated *; <i>v.</i> induce *; cause *
ID: ESIT MEM23 EDO
Notes: USG: The noun phrase “exon skipping” can take other grammatical forms depending on the linguistical context. For instance, “Skipping of the exons spanning the targeted region” is favored over “exon skipping of multiple exons spanning the targeted region” (see CTX 2). HOM: By extension, “exon skipping” corresponds to a therapy using an exon-skipping approach to eliminate a mutation. GEN: gene therapy REL: splicing, oligonucleotide
RF: Goyenvalle, A., Vulin, A., Fougousse, F., et al. Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. <i>Science</i> , 2004. Vol.306, p.1796-9 [SEC DF]; Aartsma-Rus, A., Janson, AAM., Kaman, WE., et al. Therapeutic antisense-induced exon skipping in cultured muscle cells from six different DMD patients, <i>Human Molecular Genetics</i> , 2003. Vol.12, N°8, p. 907-914 [CTX 1]; Kole, R., Krief, AM. Exon skipping therapy for Duchenne muscular dystrophy. <i>Advanced Drug Delivery Reviews</i> , 2015. Vol.87, p.104-107 [CTX 2]

VE FR saut d'exon [1]

DF : Technique qui intervient au stade de l'épissage et qui induit la suppression d'un ou plusieurs exons sur le brin nucléaire in vivo pour rétablir la fabrication d'une protéine à la fois tronquée mais fonctionnelle.

DOM : biologie cellulaire, biologie moléculaire, biotechnologie

CTX : 1. De là est née l'idée de molécules permettant le saut d'exons afin de rétablir l'expression de dystrophine, sous une forme tronquée mais fonctionnelle.

2. Cette suppression entraîne un saut d'exon de l'exon 51 lors de l'épissage et permet de rétablir le cadre de lecture.

COL : v. entraîner un * ; induire un * ; permettre un * ;

ID : ESIT MEM23 EDO

Notes :

USG : Contrairement à l'anglais, l'emploi de « saut d'exon » a tendance à rester figé, même s'il ne correspond pas grammaticalement au contexte linguistique dans lequel il est employé. Par exemple, on privilégiera « saut d'exon de l'exon 51 » à « saut d'exon 51 », même si les deux sont en usage (voir CTX 2).

GEN : thérapie génique

REL : épissage

RF : Nau, JY. Saut d'exon contre myopathie. *Revue médicale suisse*, 2004 [SEC DF].
Pichavant, C., Thérapie génique de la dystrophie musculaire de Duchenne - Utilisation de transgènes de la dystrophine chez le modèle canin. Thèse en biologie cellulaire et moléculaire. Québec : Faculté des études supérieures de l'Université Laval, 2010, 176 p. [CTX 1] ; Cullot, G. Génotoxicité des systèmes CRISPR-Cas9. Thèse en biologie cellulaire et physiopathologie. Bordeaux : Université de Bordeaux, École doctorale Sciences de la vie et de la santé, 2019, p. 63 [CTX 2]

Fiche terminologique n°5

VE EN primer [1]
DF: A short single-stranded DNA fragment synthesized in a lab and used in certain laboratory techniques to initiate the amplification process of a specific DNA sequence in the genome.
DOM: cell biology; biotechnology; medicine; genomics
CTX: Two peanut cultivars, 25 unadapted germplasm lines of <i>A. hypogaea</i> , [...] and 29 diploid wild species of <i>Arachis</i> were evaluated for variability using primers of arbitrary sequence to amplify segments of genomic DNA.
COL: v. using *; * flanking; labelling of *; * binding <i>ng.</i> PCR *; reverse *; forward *; * set; * sequence; set of *; antisense *; random *; specific *; pair of *; labelled *
ID: ESIT MEM23 EDO
Notes: EXP 1: In the context of a PCR, the primer hybridizes with the sample DNA and defines the region that will be amplified, resulting in millions of copies in a very short timeframe. Primers are also used in DNA sequencing and other experimental processes. EXP 1: Once hybridized to the strand, the primer acts as a substrate for a DNA polymerase, which creates a complementary strand via sequential addition of deoxynucleotides. EXP 2: Primers are usually 18 to 25 nucleotides long. SYN: synthetic oligonucleotide REL: PCR; annealing; hybridize; binding site; amplification
RF: Shchelochkov, OA. Primer. Glossary updated in 2023. National Human Genome Research Institute [online]. Available at genome.gov ; Halward, T., Stalker, T., LaRue, E., et al. Use of single-primer DNA amplifications in genetic studies of peanut (<i>Arachis hypogaea</i> L.). <i>Plant Molecular Biology</i> , 1992. Vol.18, pp.315-325 [CTX]; Mackay, IM., Arden, KE., Nitsche, A. Real-time PCR in virology. <i>Nucleic Acids Research</i> , 2002. Vol.30, N°6, pp.1292-1305 [EXP 2]

VE FR amorce [1]

DF : Courte séquence d'ADN ou d'ARN expérimentale nécessaire pour amorcer la réaction de synthèse d'un brin complémentaire à partir d'un brin matrice.

DOM : biologie cellulaire ; biotechnologie ; médecine

CTX : L'exon 4 a été ciblé avec les amorces suivantes qui encadrent la zone d'édition ciblée par l'oligonucléotide : 5' TAGTTCCAGGCACATAGTAAGCAC 3' (amorce sens) 5' AGGAGGTGAACAACGAATAGACAG 3' (amorce anti-sens).

COL : * d'ARN ; * d'ADN ; *de PCR ; * encadrant ; * spécifique ; * universelle ; * aléatoire ; * de séquence ; * sens ; * antisens

ID : ESIT MEM23 EDO

Notes :

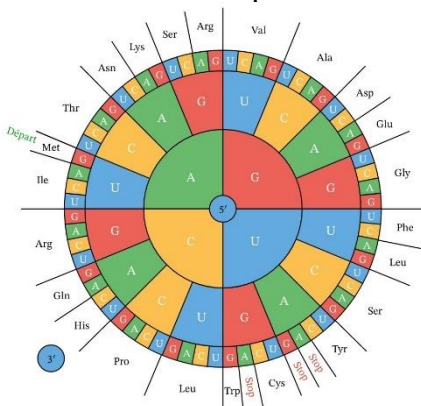
EXP : Quoique similaires dans leur définition, « amorce » et « promoteur » ne sont pas synonymes. L'amorce est utilisée dans le cadre d'une réaction en chaîne par polymérase (PCR).

DER : amorcer

REL : PCR ; hybridation ; amplification

RF : Tiret, L. Propos recueillis en juin 2023 [SEC DF] ; Dallaire, L. Amorce. Glossaire de génétique médicale et moléculaire, mis à jour en 2006 [en ligne]. Disponible sur atlasgeneticsoncology.org [SEC DF] ; Boutin, J. Alternative à la thérapie génique additive par utilisation du système CRISPR-Cas9 pour la correction génique des maladies héréditaires du globule rouge. Thèse en biologie médicale. Bordeaux : Université de Bordeaux, U.F.R. des Sciences médicales, 2017 Médecine humaine et pathologie, 2017, p.57 [CTX]

Glossaire FR-EN

ADN, acide désoxyribonucléique, A. D. N.	DNA, desoxyribonucleic acid
<p>DF : Molécule qui porte les instructions génétiques de la vie et composée de deux brins s'enroulant l'un autour de l'autre pour former une double hélice.</p> <p>EXP : Le contenu de cette information est le « code » de synthèse de toutes les protéines de l'organisme.</p> <p>EXP : L'ADN dit nucléaire est présent dans le noyau des cellules eucaryotes sous la forme de chromosomes.</p> <p>USG : Le sigle anglais DNA ou D. N. A., s'est employé en français jusque dans les années 1990. Cet anglicisme, déconseillé, est en général remplacé par le sigle français ADN devenu usuel.</p> <p>SPE : ADN nucléaire ; ADN mitochondrial</p> <p>RF : ENS Lyon, Glossaire de la génétique, à « ADN », [en ligne], disponible acces.ens-lyon.fr [DF] [EXP] ; <i>Le Grand Robert de la langue française</i>, à « ADN » [USG].</p>	
acide aminé	amino acid
<p>DF : Molécule à la base de la constitution des protéines et autres peptides.</p> <p>EXP : Chez l'Homme, comme chez de nombreuses espèces, seuls vingt acides aminés différents sont incorporés dans les protéines lors de la traduction.</p>	
<div style="display: flex; align-items: center;">  <div style="margin-left: 20px;"> <p>Schéma des différentes combinaisons de nucléotides formant des codons et les acides aminés pour lesquels ils codent. Tiré de Nagwa.com</p> </div> </div>	
<p>RF : Camus, G. Les acides aminés. Article mis à jour en 2006. Planet-Vie. ENS Lyon [en ligne]. Disponible sur planet-vie.ens.fr (consulté en mai 2023) [DF] [EXP]</p>	
ARN messenger, ARNm	Messenger RNA, mRNA
<p>DF : Acide nucléique chargé de transmettre l'information codée dans les gènes pour permettre la synthèse des protéines nécessaires au fonctionnement des cellules.</p> <p>EXP : L'ARN messenger est synthétisé durant la phase de transcription du gène puis traduit par les ribosomes dans le cytoplasme de la cellule.</p>	

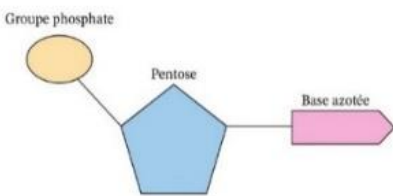
RF : Secret de fabrication : C'est quoi un ARN messenger ?. Article mis à jour en 2020. Inserm [en ligne]. Disponible sur inserm.fr (Consulté en mai 2023) [SEC DF]	
ARN polymérase	polymerase RNA
DF : Enzyme qui catalyse la transcription en synthétisant l'ARN pré-messager à partir d'une matrice d'ADN. RF : ARN polymérase : qu'est-ce que c'est ?. Article [s.d.]. FUTURA [en ligne]. Disponible sur futura-sciences.com (Consulté en juin 2023) [SEC DF]	
ARNsg, ARN guide simple	sgRNA, simple guide RNA
DF : ARN s'associant à une enzyme ou à un complexe protéique, qui, lorsqu'il s'apparie à une séquence d'ARN ou d'ADN complémentaire, permet à l'enzyme ou au complexe protéique de se positionner sur cet ARN ou ADN complémentaire. EXP : Dans la technologie de CRISPR-Cas9, un ARNsg est combiné à une enzyme Cas. Il est complémentaire à la séquence d'ADN ciblée et permet à la Cas de repérer où elle doit découper le brin d'ADN. RF : ARN guide. Définition mise à jour en 2018. France Terme [en ligne]. Disponible sur culture.fr/franceterme/terme (Consulté en mai 2023) [DF]	
chromosome	chromosome
DF : Élément constitutif du génome dans le noyau cellulaire, composé d'une longue molécule d'ADN organisée en gènes et porteuse des facteurs déterminants de l'hérédité. EXP : Le génome humain est constitué de 46 chromosomes (23 paires). RF : Le génome humain : de qui, pour qui, pourquoi ?. Communiqué, [s.d.]. Centre National de la Recherche Scientifique (CNRS) [en ligne]. Disponible sur cnrs.fr (consulté en avril 2023) [SEC DF] ; <i>Le Grand Robert de la langue française</i> , à « chromosome » [SEC DF].	
Cas	Cas
DF : Enzyme présente dans certaines bactéries, capable d'identifier des séquences d'ADN viral afin de les neutraliser en les découpant. EXP : En édition du génome, les protéines Cas sont associées à un ARN guide (complémentaire à la séquence cible) et permettent de découper un brin d'ADN de manière ciblée. ETY : de l'anglais <i>CRISPR-associated</i> (associée à CRISPR) SPE : Cas9 RF : Louapre, D., Modifier le génome avec CRISPR [VIDEO]. ScienceEtonnante, 2015 [SEC DF]	
cellule souche	stem cell
DF : Cellule présente naturellement dans l'organisme, capable de s'autorenouveler et de se différencier en d'autres types cellulaires afin de reconstituer le stock de cellules spécialisées qui ont été épuisées ou endommagées. EXP 1 : Les cellules souche ne sont pas physiologiquement spécialisées, comme le serait un myocyte (cellule musculaire) ou un globule rouge (cellule sanguine). EXP 2 : En thérapie cellulaire, les cellules souches sont élevées en culture et administrées au patient afin de l'aider à régénérer ou recréer des tissus détruits.	

<p>SPE : cellule souche embryonnaire; cellule souche pluripotente induite ; cellule souche tissulaire</p> <p>RF : Cox, C., Smith, A. Types de cellules souches et leurs applications. Fiche-info [s.d.]. EuroStemCell [en ligne]. Disponible sur eurostemcell.org (Consulté en juin 2023) [SEC DF] [SEC EXP 1] ; Cellule souche : qu'est-ce que c'est ?. Article [s.d.]. FUTURA [en ligne]. Disponible sur futura-sciences.com (Consulté en juin 2023) [SEC DF] [SEC EXP 2]</p>	
codon	codon
<p>DF : Dans un acide nucléique, triplet de nucléotides voisins, désigné par les initiales des noms des trois bases respectives, et dont l'ordre séquentiel constitue l'information qui commande et spécifie la synthèse d'un acide aminé spécifique.</p> <p>EXP : Sur l'ARNm, les codons sont reconnus par des ARNt et traduits en acides aminés par les ribosomes, pour donner lieu à une chaîne polypeptidique, qui par repliement formera une protéine. Le code génétique est ainsi lu trois nucléotides par trois nucléotides.</p> <p>GEN : exon</p> <p>RF : <i>Le Grand Robert de la langue française</i>, à « codon » [SEC DF].</p>	
codon stop, codon non-sens, codon de terminaison	terminator, stop codon, nonsense codon, chain-terminating codon
<p>DF : Codon qui désigne la fin du message génétique et qui détermine, par conséquent, la fin de la synthèse de la protéine.</p> <p>EXP : Il existe trois codons stop : UAA, UAG et UGA.</p> <p>RF : Codon STOP. Glossaire, mis à jour en 2015. AFM-Téléthon [en ligne]. Disponible sur afm-telethon.fr (consulté en mai 2023) [SEC DF] [SEC EXP]</p>	
complexe glycoprotéique associé à la dystrophine ; complexe dystrophine-glycoprotéines ; CGD	dystrophin glycoprotein complex; DGC
<p>DF : Ensemble formé par la dystrophine en position subsarcolemmale et les glycoprotéines en position transmembranaire, et responsables du maintien de l'intégrité de la fibre musculaire.</p>	
<p>Schéma du CGD, à cheval sur le sarcolemme du myocyte. Tiré de assistancescolaire.com (La cellule musculaire : une structure spécialisée dans son propre raccourcissement).</p>	
<p>RF : Cullot, G. Génotoxicité des systèmes CRISPR-Cas9. Thèse en biologie cellulaire et physiopathologie. Bordeaux : Université de Bordeaux, École doctorale Sciences de la vie et de la santé, 2019, 176 p. [SEC DF]</p>	

CRISPR, Crispr	CRISPR, Crispr
<p>DF : Biotechnologie constituée d'un complexe formé par un ARN guide et une enzyme Cas, permettant d'identifier un gène ciblé et de découper une séquence de ce gène in vivo.</p> <p>ETY : abréviation, de l'anglais <i>Clustered Regularly Interspaced Short Palindromic Repeats</i> (courtes répétitions en palindrome regroupées et régulièrement espacées).</p> <p>SYN : ciseaux moléculaires ; ciseaux génétiques</p> <p>GEN : biotechnologie</p> <p>SPE: CRISPR-Cas9</p> <p>RF : Wang, JY., Doudna, AJ. CRISPR technology: A decade of genome editing is only the beginning. <i>Science</i>, 2023. Vol. 379, No. 6629 [SEC DF]</p>	
décalage du cadre de lecture	frame shift
<p>DF : Phénomène provoqué par une mutation liée à des délétions ou des insertions qui ne sont pas des multiples de trois paires de bases, ce qui modifie la lecture du code génétique.</p> <p>SYN : déphasage du cadre de lecture ; mutation déphasante ; mutation du cadre de lecture</p> <p>ANT : rétablissement du cadre de lecture</p> <p>RF : Décalage du cadre de lecture. Définition [s.d.]. Glossaire de la PCR. ENS Lyon [en ligne]. Disponible sur ens-lyon.fr (Consulté en juin 2023) [SEC DF]</p>	
dystrophie musculaire de Duchenne	Duchenne muscular dystrophy
<p>DF : Maladie rare d'origine génétique, caractérisée par une atrophie et une faiblesse musculaires progressives dues à une dégénérescence des muscles squelettiques, lisses et cardiaques.</p> <p>EXP : La dystrophie musculaire de Duchenne étant une maladie autosomique récessive liée à l'X, elle atteint principalement les garçons (entre un sur 3 500 et un sur 9 300 à la naissance)</p> <p>SYN : DMD ; dystrophinopathie de Duchenne ; myopathie de Duchenne</p> <p>GEN : myopathie</p> <p>REL : Dystrophie musculaire de Becker</p> <p>RF : Dystrophie musculaire de Duchenne. Inventaire, classification et encyclopédie des maladies rares, avec les gènes associés. Portail des maladies rares et des médicaments orphelins [en ligne]. Disponible sur orpha.net (consulté en avril 2023) [DF] [SEC EXP]</p>	
dystrophine	dystrophin
<p>DF : Protéine géante dont la principale fonction est d'assurer une cohésion mécanique entre le cytosquelette et la matrice des cellules musculaires.</p> <p>RF : Mirouse, V. Des fonctions inattendues pour la dystrophine, impliquée dans la myopathie de Duchenne. Article de l'Institut de Génétique Reproduction et Développement (iGRéD), 2020. CNRS [DF]</p>	
enzyme	enzyme
<p>DF : Substance protéinique qui facilite ou accroît une réaction biochimique.</p> <p>DER : enzymatique ; enzymologie</p> <p>RF : <i>Le Grand Robert de la langue française</i>, à « enzyme » [DF].</p>	

exon	exon
<p>DF : séquence codante d'un gène, par opposition à l'intron. ANT : intron SPE : codon</p>	
gamète, cellule sexuelle	gamete
<p>DF : Cellule sexuée mâle ou femelle, destinée à la fécondation et dont le noyau comporte un nombre N de chromosomes égal à la moitié de celui des cellules originelles. DER : gamétique ; gaméticide SPE : spermatozoïde ; ovule RF : Gamète. Définition mise à jour en 2021 /ca. Centre National de Ressources Textuelles et Lexicales (CNRTL) [en ligne]. Disponible sur cnrtl.fr/definition (Consulté en juin 2023) [SEC DF]</p>	
gène	gene
<p>DF : Fragment d'ADN contenant toutes les informations nécessaires pour produire un ARN ou, le plus souvent, une protéine. EXP : Les gènes contiennent des séquences codantes (exons) et non codantes (introns). RF : Le génome humain : de qui, pour qui, pourquoi ?. Communiqué, [s.d.]. Centre National de la Recherche Scientifique (CNRS) [en ligne]. Disponible sur cnrs.fr (consulté en avril 2023) [DF]</p>	
génomome	genome
<p>DF : Ensemble de l'information génétique d'un organisme contenu dans chacune de ses cellules sous la forme de chromosomes. EXP : Le support matériel du génome est l'ADN, sauf chez certains virus où il s'agit d'ARN. RF : Le génome humain : de qui, pour qui, pourquoi ?. Communiqué, [s.d.]. Centre National de la Recherche Scientifique (CNRS) [en ligne]. Disponible sur cnrs.fr (consulté en avril 2023) [DF] [EXP]</p>	
histologie	histology
<p>DF : Science qui traite de la structure des tissus et des cellules qui constituent les êtres vivants. RF : <i>Le Grand Robert de la langue française</i>, à « histologie » [DF].</p>	
immunomarquage	immunostaining
<p>DF : Étude par imagerie de tissus biologiques par des techniques de marquage immunologique utilisant des anticorps spécifiques. REL : immunohistochimie ; immunofluorescence ; Western blot ; immunoélectrophorèse RF : <i>Le Grand Robert de la langue française</i>, à « immunohistochimie » [SEC DF] ; Immunomarquage. Définition mise à jour en 2015. Dictionnaire médical. HOPITAL.fr [en ligne]. Disponible sur hopital.fr/Le-dico-medical (Consulté en juin 2023)</p>	
injection systémique	systemic delivery
<p>DF : Injection par laquelle le produit injecté emprunte la circulation sanguine pour atteindre</p>	

<p>son site d'action SYN : injection par voie générale ANT : injection locale SPE : injection parentérale ; voie entérale ; injection intramusculaire RF : Boulanger, T. Les Formes pharmaceutiques et les voies d'administration. Pharmacie galénique. Institut de formation en soins infirmiers, CH Carcassonne, 2014 [SEC DF]</p>	
intron	intron
<p>DF : Séquence nucléotidique non codante d'un gène, par opposition à l'exon. EXP 1 : Les introns sont présents dans les ARN prématures (ou pré-messagers), transcrits à partir du brin matrice d'ADN comportant le gène d'intérêt. Pendant la phase d'épissage, les introns sont éliminés et l'ARN messager (ou mature) n'est constitué que d'exons. EXP 2 : Le rôle des introns est encore peu connu, mais il est possible que ces séquences nucléotidiques jouent un rôle dans la manière dont le gène concerné s'exprime et est transcrit. ANT : exon RF : Liu, PP. Intron. Article mis à jour en 2023. National Human Genome Research Institute [en ligne]. Disponible sur genome.gov (Consulté en juin 2023)</p>	
indel, InDel, INDEL	indel
<p>DF : Ensemble des événements d'insertion et/ou de délétion d'un ou plusieurs nucléotides au niveau du site de découpage d'une séquence nucléotidique génomique. EXP : Mot-valise composé de la première syllabe du mot « insertion » et de celle du mot « délétion ». RF : Sehn, JK. Chapter 9 - Insertions and Deletions (Indels). <i>Clinical Genomics</i>. Academic Press, 2015, pp. 126-150 [SEC DF]</p>	
locus	locus
<p>DF : Localisation d'un gène ou d'un de ses allèles sur un chromosome. GRM : <i>pl. loci</i> RF : <i>Le Grand Robert de la langue française</i>, à « locus » [DF].</p>	
mutation	mutation
<p>DF : Altération du matériel génétique entraînant une modification durable de certains caractères du fait de la transmission héréditaire de ce matériel de génération en génération. RF : ENS Lyon, Glossaire de la génétique, à « mutation », [en ligne], disponible sur acces.ens-lyon.fr [SEC DF]</p>	
motif PAM	PAM sequence, protospacer adjacent motif
<p>DF : Courte séquence nucléotidique permettant au complexe CRISPR de s'hybrider au brin d'ADN ciblé dans la cellule. EXP : Le motif PAM se situe à côté de la séquence complémentaire à celle ciblée, à un locus déterminé. Dans le complexe CRISPR-Cas9, le motif PAM correspond à la suite des trois nucléotides NGG ou NGA. ETY : de l'anglais <i>protospacer adjacent motif</i>.</p>	

<p>ETY : de l'anglais <i>protospacer adjacent motif</i>. RF : Tired, L. [SEC DF] [SEC EXP] ; Tremblay, JP. CRISPR, un système qui permet de corriger ou de modifier l'expression de gènes responsables de maladies héréditaires. <i>Médecine/sciences</i>, 2015. Vol. 31, No 11, pp. 1014-1022 [SEC EXP].</p>	
mutagénèse	mutagenesis
<p>DF : Production de mutations due à l'action d'agents physiques ou chimiques. USG : peut aussi s'écrire mutagenèse. RF : <i>Le Grand Robert de la langue française</i>, à « mutagenèse » [DF].</p>	
myocyte, cellule musculaire, fibre musculaire	myocytes, muscle cell, muscle fiber
<p>DF : Cellule spécialisée dans la contraction musculaire. SPE : cellule musculaire striée ; cellule musculaire lisse ; cardiomyocyte RF : Myocyte : qu'est-ce que c'est ?. Article [s.d.]. FUTURA [en ligne]. Disponible sur futura-sciences.com (Consulté en juin 2023) [SEC DF]</p>	
nucléotide	nucleotide
<p>DF : Constituant élémentaire de l'ADN et de l'ARN, composé d'un sucre (désoxyribose ou ribose), d'un phosphate et d'une base azotée.</p> <div style="text-align: center;">  <p>Composition d'un nucléotide. Tiré du site Nagwa.com</p> </div>	
<p>RF : ENS Lyon, Glossaire de la génétique, à « nucléotide », [en ligne], disponible sur acces.ens-lyon.fr [DF]</p>	
PCR, amplification génique, réaction en chaîne par polymérase	PCR, polymerase chain reaction
<p>DF : Technique d'amplification in vitro permettant de recopier de manière exponentielle un fragment d'ADN grâce à l'utilisation d'une enzyme polymérase afin d'obtenir une quantité suffisante de ce fragment pour l'étudier ou l'utiliser dans des expériences. ETY : 1989 ; sigle angl. de <i>polymerase chain reaction</i> « réaction en chaîne par polymérase ». RF : ENS Lyon, Glossaire de la génétique, à « Amplification génique », [en ligne], disponible sur acces.ens-lyon.fr [SEC DF] ; <i>Le Grand Robert de la langue française</i>, à « PCR » [SEC DF] [ETY].</p>	
phénotype sauvage	wild-type
<p>DF : Forme naturelle ou de référence d'un phénotype. EXP : Le phénotype d'un individu correspond à l'ensemble de ses caractères observables, résultant de l'interaction entre son génotype et les effets de son environnement. Dans l'absolu, le phénotype sauvage serait celui d'un individu original sans mutation. Dans les faits, n'importe quel individu est le fruit, dans l'histoire, de plusieurs mutations. C'est</p>	

<p>pourquoi le phénotype sauvage désigne dans l'usage la médiane de la variabilité d'un phénotype sauvage. En comparaison avec un phénotype malade, il désigne le phénotype sain.</p> <p>RF : Tiret, L. [SEC DF] [SEC EXP] ; ENS Lyon, Glossaire de la génétique, à « Amplification génique », [en ligne], disponible sur acces.ens-lyon.fr [SEC EXP]</p>	
produit d'amplification, amplicon	amplicon, amplification product
<p>DF : Séquence spécifique d'ADN double brin obtenue en plusieurs exemplaires lors d'une réaction d'amplification en chaîne par polymérase (PCR).</p> <p>EXP : Les produits d'amplification sont analysés ou séquencés pour identifier l'expression d'un gène.</p> <p>RF : Produit d'amplification. Définition mise à jour en 2011. France Terme [en ligne]. Disponible sur culture.fr/franceterme/terme (Consulté en mai 2023) [SEC DF]</p>	
promoteur	promotor
<p>DF : Séquence d'ADN située en amont de la séquence codant la protéine, indispensable à l'expression du gène.</p> <p>EXP : Le promoteur n'est pas universel : il doit être adapté à la cellule receveuse en cas de transfert du gène.</p> <p>RF : ENS Lyon, Glossaire de la génétique, à « promoteur », [en ligne], disponible sur acces.ens-lyon.fr [DF] [EXP].</p>	
protéine	protein
<p>DF : Molécule présente dans les organismes vivants et composée d'un enchaînement d'acides aminés.</p> <p>EXP : Les protéines remplissent différentes fonctions dans la cellule, notamment des fonctions de structure et des fonctions enzymatiques.</p> <p>RF : ENS Lyon, Glossaire de la génétique, à « protéine », [en ligne], disponible sur acces.ens-lyon.fr [SEC DF] [EXP].</p>	
révertant	revertant
<p>DF : Dérivé par réversion d'une cellule ou d'un organisme mutant.</p> <p>ETY : emprunté à l'anglais <i>revertant</i> (1955), de <i>to revert</i> « revenir, retourner ».</p> <p>RF : <i>Le Grand Robert de la langue française</i>, à « révertant » [DF] [SEC ETY].</p>	
thérapie génique	gene therapy
<p>DF : Technique biologique permettant l'introduction de gènes sains dans les cellules pour qu'ils assurent les fonctions que les gènes défectueux ne parviennent pas à assurer.</p> <p>SPE : édition du génome ; thérapie cellulaire</p> <p>RF : Consultation citoyenne sur l'édition du génome en île-de-France. Faculté de médecine de l'Université Paris-Saclay, 2021, p. 3 [SEC DF].</p>	
transgène	transgene
<p>DF : Gène introduit dans le génome d'un organisme par génie génétique.</p> <p>RF : ENS Lyon, Glossaire de la génétique, à « transgène », [en ligne], disponible acces.ens-lyon.fr [DF].</p>	

Lexique FR-EN

FR	en
acide aminé	amino acid
activatrice d'épissage exonique (AEE)	exonic splicing enhancer (ESE)
ADN, acide désoxyribonucléique	DNA, desoxyribonucleic acid
ADN complémentaire, ADNc	complementary DNA, cDNA
ADN mitochondrial	mitochondrial DNA
ADN nucléaire	genomic DNA
allèle	allele
autosome, chromosome somatique, euchromosome	autosome, somatic chromosome, euchromosome
amorçe	primer
anomalie	anomaly
ARN messenger, ARNm	messenger RNA, mRNA
ARN polymérase	RNA polymerase
ARN pré-messenger, ARNm précurseur, ARNpm	pre-messenger RNA, precursor mRNA, pmRNA
ARN ribosomique, ARNr, ARN ribosomal	ribosomal RNA, rRNA, ribosomic rRNA
ARN guide simple, ARNsg	simple guide RNA, sgRNA
ARN de transfert, ARNt, ARN soluble	transfer RNA, tRNA, soluble RNA
atrophie	atrophy
base azotée, nucléobase	nitrogenous base, nucleobase
cadre de lecture	reading frame
cardiomyocytes, myocytes cardiaques	cardiomyocytes, cardiac muscle cell
Cas9	Cas9
cassure double-brin	double-stranded break
cellule	cell
cellule embryonnaire	embryonic cell
cellule eucaryote	eukaryotic cell

cellule fille	daughter cell
cellule post-mitotique	post-mitotic cell
cellule souche	stem cell
cellule souche pluripotente induite	induced pluripotent stem cell (iPSC)
chromosome	chromosome
codon	codon
codon stop, codon non-sens, codon de terminaison	terminator, stop codon, nonsense codon, chain-terminating codon
coloration à l'hématoxyline et l'éosine (H&E)	hematoxylin and eosin (H&E) staining
complexe glycoprotéique associé à la dystrophine (CGD), complexe dystrophine-glycoprotéine	dystrophin-associated glycoprotein complex (DGC)
contraction (musculaire)	contraction
coupure génomique	genomic cut
coupure unique	single cut
créatine kinase (CK), créatine phosphokinase (CPK)	creatine kinase (CK), creatine phosphokinase (CPK)
CRISPR, Crispr	CRISPR, Crispr
cytoplasme	cytoplasm
cytosquelette	cytoskeleton
décalage du cadre de lecture	frameshift
dégénérescence musculaire	muscle degeneration
dégénérescence cellulaire	cell degeneration
dégradation des ARNm médiée par les brins non-sens	nonsense mediated decay, NMD
délétion	deletion
désoxyribonucléotide	deoxyribonucleotide
dystroglicane	dystroglican
dystrophie	dystrophy
dystrophie musculaire de Becker	Becker muscular dystrophy
dystrophie musculaire de Duchenne	Duchenne muscular dystrophy
dystrophine	dystrophin

échantillon	sample
édition du génome, modification du génome, Édition génomique, édition génique	genome editing, gene editing
enzyme	enzyme
exon	exon
fibrose	fibrosis
gamète, cellule sexuelle	gamete
gène	gene
génom	genome
hématologie	hematology
héréditaire	hereditary
histologie	histology
hors cible	off-target
hybridation, fusion	annealing, hybridization, binding
immunogénicité	immunogenicity
immunomarquage	immunostaining
immunosuppression	immune suppression
indel, InDel, INDEL	indel, InDel, INDEL
infiltration cellulaire	cell infiltration
injection systémique	systemic injection
insertion	insertion
insuffisance respiratoire	respiratory collapse
intégrité musculaire	muscle integrity
intron	intron
locus	locus
maladie chromosomique	chromosome disorder, chromosomal disorder
maladie dégénérative	degenerative disease
maladie héréditaire	inherited disease, hereditary disease
maladie monogénique	single-gene disease, monogenic disease

maladie polygénique, maladie multigénique	polygenic disease
marqueur	tag
mégabase	megabase
mitose	mitosis
mitochondrie, chondriosome	mitochondria, mitochondrion, chondriosome
motif PAM	PAM sequence
muscle	muscle
muscle cardiaque, myocarde	cardiac muscle, heart muscle, myocardium
muscle squelettique	skeletal muscle
mutagénèse	mutagenesis
mutation	mutation
mutation <i>de novo</i>	de novo mutation
mutation induite	induced mutation
mutation spontanée	spontaneous mutation
myocyte, cellule musculaire, fibre musculaire	myocyte, muscle cell, muscle fiber
myofibrille, fibrille musculaire	myofibril, myofibrilla, muscular fibril
myopathie	myopathy
myosine	myosin
nucléotide	nucleotide
œdème	edema
oligonucléotide	oligonucleotide, oligo
paire de bases	base pair
pathologie	pathology
PCR	PCR, polymerase chain reaction
PCR quantitative, qPCR	quantitative PCR, qPCR,
phénotype	phenotype
phénotype sauvage	wild-type phenotype
polypeptide	polypeptide
produit d'amplification, amplicon, produit	amplicon, PCR product

de PCR	
promoteur	promotor
protéine	protein
réarrangement génomique	genome rearrangement
recherche translationnelle	translational studies
recombinaison homologue	homologous recombination, homology directed repair (HDR)
recombinaison non homologue, jonction d'extrémités non homologues	non-homologous end joining (NHEJ)
relâchement	relaxation
rétablissement du cadre de lecture	reframing
révertant	revertant
ribosome	ribosome
RT-PCR	RT-PCR, reverse transcriptase PCR
sarcolemme	sarcolemma
sarcomère	sarcomere
saut d'exon	exon skipping
séquençage	sequencing
site accepteur d'épissage	splice acceptor site
site de coupure, site de cassure	cutting site, cleavage site
substitution	substitution
thérapie génique	gene therapy
transgène	transgene
vecteur viral	viral vector
virus	virus
virus adénoassocié, virus adéno-associé, AAV	adeno-associated virus (AAV)
vinculine	vinculin

Lexique EN-FR

En	Fr
adeno-associated virus (AAV)	virus adénoassocié, virus adéno-associé, AAV
allele	allèle
amino acid	acide aminé
amplicon, PCR product	produit d'amplification, amplicon, produit de PCR
anomaly	anomalie
annealing, hybridization, binding	hybridation, fusion
atrophy	atrophie
autosome, somatic chromosome, euchromosome	autosome, chromosome somatique, euchromosome
base pair	paire de bases
Becker muscular Dystrophy	dystrophie musculaire de Becker
cardiac muscle, heart muscle, myocardium	muscle cardiaque, myocarde
cardiomyocytes, cardiac muscle cell	cardiomyocytes, myocytes cardiaques
Cas9	Cas9
cell	cellule
cell degeneration	dégénérescence cellulaire
cell infiltration	infiltration cellulaire
chromosome	chromosome
chromosome disorder, chromosomal disorder	maladie chromosomique
codon	codon
complementary DNA, cdna	ADN complémentaire, ADNc
contraction	contraction (musculaire)
creatine kinase (CK), creatine phosphokinase (CPK)	créatine kinase (CK), créatine phosphokinase (CPK)
CRISPR, Crispr	CRISPR, Crispr

cutting site, cleavage site	site de coupure
cytoplasm	cytoplasme
cytoskeleton	cytosquelette
daughter cell	cellule fille
de novo mutation	mutation <i>de novo</i>
degenerative disease	maladie dégénérative
deletion	délétion
deoxyribonucleotide	désoxyribonucléotide
DNA, desoxyribonucleic acid	ADN, acide désoxyribonucléique
double-stranded break	cassure double-brin
Duchenne muscular dystrophy	dystrophie musculaire de duchenne
dystroglican	dystroglicane
dystrophin	dystrophine
dystrophin-associated glycoprotein complex (DGC)	complexe glycoprotéique associé à la dystrophine (CGD), complexe dystrophine-glycoprotéine
dystrophy	dystrophie
edema	œdème
embryonic cell	cellule embryonnaire
enzyme	enzyme
eukaryotic cell	cellule eucaryote
exon	exon
exon skipping	saut d'exon
exonic splicing enhancer (ESE)	activatrice d'épissage exonique (AEE)
fibrosis	fibrose
frameshift	décalage du cadre de lecture
gamete	gamète, cellule sexuelle
gene	gène
gene therapy	thérapie génique
genome	génom
genome editing, gene editing	édition du génome, modification du génome, Édition génomique, édition génique
genome rearrangement	réarrangement génomique
genomic cut	coupure génomique

hematology	hématologie
hematoxylin and eosin (H&E) staining	coloration à l'hématoxyline et l'éosine (H&E)
hereditary	héréditaire
histology	histologie
homologous recombination, homology directed repair (HDR)	recombinaison homologue
immune suppression	immunosuppression
immunogenicity	immunogénicité
immunostaining	immunomarquage
indel, InDel, INDEL	indel, InDel, INDEL
induced mutation	mutation induite
induced pluripotent stem cell (iPSC)	cellule souche pluripotente induite
inherited disease, hereditary disease	maladie héréditaire
insertion	insertion
intron	intron
locus	locus
megabase	mégabase
messenger RNA, mrna	ARN messenger, ARNm
mitochondria, mitochondrion, chondriosome	mitochondrie, chondriosome
mitosis	mitose
muscle	muscle
muscle degeneration	dégénérescence musculaire
muscle integrity	intégrité musculaire
mutagenesis	mutagénèse
mutation	mutation
myocyte, muscle cell, muscle fiber	myocyte, cellule musculaire, fibre musculaire
myofibril, myofibrilla, muscular fibril	myofibrille, fibrille musculaire
myopathy	myopathie
myosin	myosine
nitrogenous base, nucleobase	base azotée, nucléobase

non-homologous end joining (NHEJ)	recombinaison non homologue, jonction d'extrémités non homologues
nonsense mediated decay, NMD	dégradation des ARNm médiée par les brins non-sens
nucleotide	nucléotide
off-target	hors cible
oligonucleotide, oligo	oligonucléotide
PAM sequence	motif PAM
pathology	pathologie
PCR, polymerase chain reaction	PCR
phenotype	phénotype
polygenic disease	maladie polygénique, maladie multigénique
polypeptide	polypeptide
post-mitotic cell	cellule post-mitotique
pre-messenger RNA, precursor mRNA, pmRNA	ARN pré-messager, ARNm précurseur, ARN _{pm}
primer	amorçe
promotor	promoteur
protein	protéine
quantitative PCR, qPCR,	PCR quantitative, qPCR
reading frame	cadre de lecture
reframing	rétablissement du cadre de lecture
relaxation	relâchement
respiratory collapse	insuffisance respiratoire
revertant	révertant
ribosomal RNA, rRNA, ribosomic RNA	ARN ribosomique, ARN _r , ARN ribosomal
ribosome	ribosome
RNA polymerase	ARN polymérase
RT-PCR, reverse transcriptase PCR	RT-PCR
sample	échantillon
sarcolemma	sarcolemme
sarcomere	sarcomère
sequencing	séquençage
simple guide RNA, sgRNA	ARN guide simple, ARN _{sg}

single cut	coupure unique
single-gene disease, monogenic disease	maladie monogénique
skeletal muscle	muscle squelettique
splice acceptor site	site accepteur d'épissage
spontaneous mutation	mutation spontanée
stem cell	cellule souche
substitution	substitution
systemic injection	injection systémique
tag	marqueur
terminator, stop codon, nonsense codon, chain-terminating codon	codon stop, codon non-sens, codon de terminaison
transfer RNA, tRNA, soluble RNA	ARN de transfert, ARNt, ARN soluble
transgene	transgène
translational studies	recherche translationnelle
vinculin	vinculine
viral vector	vecteur viral
virus	virus
wild-type phenotype	phénotype sauvage

BIBLIOGRAPHIE CRITIQUE

Bibliographie anglaise

Ressources pédagogiques

Instituts de recherche et académies scientifiques

- National Institutes of Health [en ligne]. Disponible sur www.nih.gov/

Le NIH est l'agence nationale de recherche médicale des États-Unis et fait partie du Department of Health and Human Services. Le NIH met à disposition plusieurs ressources en ligne de qualité et spécialisées en médecine, telles que :

- La National Library of Medicine [en ligne]. Disponible sur www.nlm.nih.gov.

Cette bibliothèque en ligne donne accès à de nombreux articles et travaux scientifiques tirés de revues scientifiques pointues, telles que le *European Journal of Paediatric Neurology*, ou *Life Sciences*.

- MedlinePlus [en ligne]. Disponible sur <https://medlineplus.gov>.

MedlinePlus est une plateforme qui répertorie et met fréquemment à jour une quantité d'informations sur toutes sortes de maladies. L'organisation de ses fiches rend la lecture très simple et les informations facilement repérables. Certains articles sont sous forme de questions-réponses (par ex : « Muscular Dystrophy » medlineplus.gov/musculardystrophy), et des ressources plus poussées sont toujours fournies afin d'en apprendre plus (« Scientific journal articles for further reading »).

- Le National Human Genome Research Institute [en ligne]. Disponible sur www.genome.gov/genetics-glossary

Le NHGRI est une organisation fondée en 1989 et qui collabore avec les communautés scientifiques et médicales en vue de faire avancer la recherche

génomique.

Cette ressource m'a été utile pour son glossaire en génétique fourni et détaillé. Chaque fiche explicative d'un terme constitue un article rédigé par un scientifique ou un professionnel spécialisé en génétique au sein de l'Institut. Les articles s'accompagnent de schémas explicatifs clairs, dont certains ont été repris dans ce mémoire. Cependant, les explications restent de l'ordre de la vulgarisation et ne permettent donc pas d'identifier des faits de langue très spécialisés.

- EuroStemCell [en ligne]. Disponible sur <https://www.eurostemcell.org/fr>

EuroStemCell est un réseau de scientifiques et d'académiques européen qui fournit des informations pédagogiques sur les cellules souches. Le site est disponible en plusieurs langues, dont l'anglais et le français, ce qui permet d'avoir accès aux mêmes informations de manière bilingue. La qualité textuelle et de contenu aident à accéder à des informations plus abouties que celles disponibles sur le NHGRI. Toutefois, je n'ai pas pu trouver d'information sur les éventuelles traductions de tous les articles.

Ressources académiques

Articles scientifiques (scientific papers)

- Wang, JY., Doudna, AJ. CRISPR technology: A decade of genome editing is only the beginning. *Science*, 2023. Vol. 379, No. 6629

Cet article, coécrit par l'une des chercheuses ayant mis au point l'outil CRISPR-Cas9, fait un état des lieux de cette technologie. La fiabilité de la source (autrices et magazine dans lequel il est publié) permet d'apaiser certaines confusions liées au domaine de spécialisation. De plus, cet article est à la fois très explicatif et contextualisant, mais aussi très spécialisé. Je considère qu'il est accessible seulement

en ayant étudié un peu les caractéristiques de l'édition du génome au préalable.

Cet article m'a été très utile pour la création des fiches terminologiques et la recherche des faits de langue, mais aussi pour l'étude de l'état de la recherche actuelle dans ce domaine. En effet, le texte date de 2023 et il est difficile de faire plus récent. Il reprend tous les points positifs majeurs de CRISPR, mais aussi les failles à améliorer. Il envisage les perspectives futures des applications de l'édition du génome et les questions techniques, économiques et politiques que suscite la technologie CRISPR.

- Lim, KR., Yoon, C., Yokota, T. Applications of CRISPR/Cas9 for the Treatment of Duchenne Muscular Dystrophy. *Journal of Personalized Medicine*, 2018. Vol. 8, No 38

La similarité du sujet traité dans ce *paper* avec celui de mon texte-support a fait de cette ressource une merveilleuse façon de comparer les faits de langue anglophones. L'avantage de l'article est aussi qu'il se focalise sur la contextualisation, il est donc très explicatif et très concret. Il permet de se familiariser avec les usages.

Enfin, de nombreux articles et travaux scientifiques ont été incorporés aux corpus comparables afin d'étudier la fréquence des faits de langue et les emplois en contexte.

Bibliographie française

Ressources pédagogiques

Cours du secondaire et universitaires, MOOC

La disponibilité en ligne de cours et de travaux pratiques du secondaire ou de l'université a constitué pour moi un réel avantage pour me plonger dans le monde de la génétique. Ces ressources m'ont été très utiles pour comprendre de manière claire, parfois un peu succinctes, des aspects déjà très institutionnalisés de la biologie : les grands principes de la génétique, la division cellulaire, le principe de la PCR, etc. Souvent, ces cours ont été pensés pour introduire

un sujet et développer son fonctionnement de manière structurée.

Cependant, les sujets plus récents ou encore peu institutionnalisés, comme les méthodes de thérapie génique, ne sont pas abordés dans ces ressources. Pour ces sujets, j'ai donc dû créer des structures de mon côté, et faire sens d'une multitude de ressources pointues et autonomes, pour les regrouper et aboutir à un fil logique simple et exhaustif. Quant aux cours, ateliers, MOOC et TP utilisés, nous pouvons citer les suivants à titre d'exemple :

- Bechkri, S. Cours de génétique, L2 Sciences biologiques. Constantine : Université Frères Mentouri, Faculté des sciences de la nature et de la vie, 2020. 101 p.

Ce cours, destiné à des élèves de licence en Sciences biologiques, Biotechnologies et Écologie et Environnement, a l'avantage d'être exhaustif et de suivre un ordre de complexité croissant. La table des matières permet également de rapidement recueillir une information. Cependant, certains aspects explicatifs sont restés flous pour moi, sûrement car il s'agit d'un cours de L2. J'ai donc souvent dû compléter ce que j'y trouvais avec d'autres ressources, ne serait-ce que pour m'assurer de la qualité du propos. Sur ce point, les références bibliographiques Dr. Bechkri S. reprennent des classiques de la génétique (Génétique générale, 1985 ; Introduction à l'analyse génétique 4^e édition, 2006 ; etc.).

- Plateforme pédagogique e-enseignement de l'académie de Nice [en ligne]. Disponible sur e-enseignement.ac-nice.fr

Cette plateforme met à disposition de ses élèves mais aussi du public extérieur des cours ou des activités créées par les enseignants. Ainsi sont accessibles des cours, des ressources informatiques, des extraits sonores et des vidéos, travaux collaboratifs, et autres.

- ENS Lyon, Glossaire de la génétique [en ligne], disponible acces.ens-lyon.fr ; Planet-Vie. ENS Lyon [en ligne]. Disponible sur planet-vie.ens.fr

L'ENS Lyon possède plusieurs en ressources en ligne, qui constituent des valeurs sûres. Les glossaires et les fiches pédagogiques sont nombreuses (explication de la PCR, de la qPCR, etc.), les articles disponibles sur Planet-vie sont rédigés par des enseignants chercheurs spécialistes du domaine.

Vidéos et documentaires

L'avantage des ressources vidéo est qu'elles peuvent être utilisées dans une approche plus kinesthésique de la traduction. En effet, quitte à ne pas pouvoir visiter un laboratoire de génétique, je peux observer quelqu'un menant une expérience et qui a été filmé. Cette approche visuelle me permettra de mieux comprendre ce que je traduits concrètement, et de m'assurer que le sens et les tournures sont donc adéquates. J'ai par exemple ainsi procédé avec la vidéo suivante :

- David, L. MOOC côté labo : Technique d'étude des protéines cellulaires : western blot : les méthodes [VIDEO]. Inserm, 2017

Cette vidéo est avant tout pédagogique (puisque'il s'agit d'un MOOC), mais donne accès à une expérience autrement difficilement vérifiable. Grâce à cette vidéo, j'ai pu mieux comprendre la chronologie du Western Blot, et ôter certains doutes face à mon texte-support (notamment la question de savoir ce à quoi faisait référence la « upper-band » (p.64, paragr.2)).

Je me suis aussi servie d'autres vidéos, destinées au grand public, telles que :

- Lemarchand, P. MOOC côté cours : La contraction musculaire [VIDEO]. Inserm, 2015.
- Louapre, D., Modifier le génome avec CRISPR [VIDEO]. ScienceEtonnante, 2015

Ces vidéos peuvent être de bons supports pédagogiques et elles donnent accès au parler plus oral du domaine de spécialisation.

Instituts de recherche et académies scientifiques

De nombreuses institutions scientifiques offrent un contenu francophone de qualité et attesté par la communauté scientifique. À ce titre, nous pouvons citer les sites suivants :

- L'AFM-Téléthon [en ligne]. Disponible sur www.afm-telethon.fr

L'AFM-Téléthon est une association française « de militants, malades et parents de malades, concernés par des maladies génétiques, rares, évolutives et lourdement invalidantes, les maladies neuromusculaires. » (description sur le site). Le site de l'AFM-Téléthon possède un glossaire à visée pédagogique fréquemment mis à jour. Il m'a notamment servi à appréhender les notions de cadre de lecture, de rétablissement

et de modification du cadre de lecture.

L'association met aussi en ligne des articles et des vidéos sur les gènes, les maladies génétiques et la dystrophie musculaire de Duchenne.

- L'Institut Imaginaire [en ligne]. Disponible sur www.institutimagine.org/fr

L'Institut Imaginaire est un pôle européen de recherche, de soins et d'enseignement sur les maladies génétiques qui rassemble 1000 médecins, chercheurs et personnels de santé. Cette référence bibliographique a été validée par mon référent spécialiste, enseignant chercheur à l'École vétérinaire de Maison Alfort. Le site de l'institut présente l'avantage d'être bilingue français – anglais. Il permet d'avoir accès à des statistiques solides et des articles rédigés par des spécialistes.

- L'Inserm [en ligne]. Disponible sur www.inserm.fr

L'Institut nationale de la santé et de la recherche médicale est un établissement dédié à la recherche biologique et médical. Il est public et placé sous la tutelle du ministère de la Santé et du ministère de la Recherche. L'Inserm propose des dossiers sur les avancées biotechnologiques très complets. Nous pouvons par exemple citer le dossier « Thérapie génique - Une recherche de longue haleine qui porte ses fruits » datant de 2018. Ce dernier fait le profil détaillé des applications géniques en médecine, des différentes méthodes de thérapie génique et des médicaments sur le marché. Nous pouvons de même mentionner le dossier « Thérapies à ARN – Un domaine thérapeutique en pleine extension », datant de 2022.

Nous pouvons citer de nombreux autres instituts spécialisés et leurs Laboratoires de recherche associés, et dont le niveau de langue permet une analyse francophone satisfaisante des emplois actés dans le domaine de la thérapie génique : le CNRS [en ligne]. Disponible sur www.cnrs.fr ; l'Institut Pasteur [en ligne]. Disponible sur www.pasteur.fr ; le Portail des maladies rares et des médicaments orphelins, Orphanet [en ligne]. Disponible sur orpha.net ; l'Espace éthique de la région Île-de-France [en ligne]. Disponible sur www.espace-ethique.org ;

Centres hospitaliers

Les centres hospitaliers (CH) et centre hospitaliers universitaires (CHU) recèlent de documentation académique et de documentation dédiée au grand public dans le domaine de la médecine et des maladies rares. J'ai notamment utilisé les articles disponibles sur le site du CHU Sainte-Justine (CHU mère-enfant) [en ligne] , disponibles sur chusj.org et fréquemment mis à jour.

Ressources académiques

Travaux de thèse

Les travaux de thèse français ou francophones en médecine et en biologie ont soutenu ma recherche et ma traduction du début à la fin. D'une part, les introductions constituent un bon moyen de se familiariser à un sujet de manière simple. D'autre part, les spécificités techniques du domaine de recherche m'ont permis d'observer les faits de langue dans un contexte académique. Cependant, il faut noter que la plupart des références citées dans ces travaux de thèses sont anglophones. Il y a en effet plus de thèses françaises en thérapie génique et en édition du génome, que de publications scientifiques françaises sur le même sujet. Nous pouvons donc nous demander dans quelle mesure les tournures employées dans les travaux sont officiellement correctes. À supposer que bien des doctorants travaillent en quotidien au français, nous pouvons considérer que leurs emplois constituent des arguments d'autorité.

- Cullot, G. Génotoxicité des systèmes CRISPR-Cas9. Thèse en biologie cellulaire et physiopathologie. Bordeaux : Université de Bordeaux, École doctorale Sciences de la vie et de la santé, 2019, 176 p.
- Boutin, J. Alternative à la thérapie génique additive par utilisation du système CRISPR-Cas9 pour la correction génique des maladies héréditaires du globule rouge. Thèse en biologie médicale. Bordeaux : Université de Bordeaux, U.F.R. des Sciences médicales, 2017 Médecine humaine et pathologie, 2017, 95 p.

Comme en attestent les notes de bas de page et les références citées dans le corps du mémoire, les deux thèses mentionnées ci-dessus ont amplement été exploitées pour la création des fiches terminologiques, et pour la compréhension de certains faits de langues très spécifiques dans mon texte-support (comme la dégradation des ARNm médiée par les brins non-sens, l'utilisation de vecteurs viraux adénoassociés, etc.).

- Pichavant, C., *Thérapie génique de la dystrophie musculaire de Duchenne - Utilisation de transgènes de la dystrophine chez le modèle canin*. Thèse en biologie cellulaire et moléculaire. Québec : Faculté des études supérieures de l'Université Laval, 2010, 176 p.

Cette thèse est québécoise, mais dans le cadre du domaine de la thérapie génique et de la médecine, les différences de langue n'ont pas d'incidence. D'ailleurs, je n'en ai pas vraiment trouvé. De plus, ce travail m'a été particulièrement utile pour comprendre les enjeux des méthodes thérapeutiques liées au cas particulier de la myopathie de Duchenne. De ce fait, beaucoup de passages étaient très similaires aux cas de figures présents dans mon texte-support, ce qui a permis d'établir des comparaisons linguistiques et sémantiques.

Articles scientifiques

Beaucoup de travaux scientifiques sont disponibles dans les revues scientifiques comme *Médecine/sciences* ou *Science*. Il s'agit de comptes rendus d'expériences et d'essais de phase I, II ou III, d'études scientifiques approfondies et autres travaux spécialisés. En outre, j'ai manuellement incorporé ces articles (ainsi que des travaux de thèse) aux corpus comparables dans *Sketch Engine*.

- Tremblay, JP. CRISPR, un système qui permet de corriger ou de modifier l'expression de gènes responsables de maladies héréditaires. *Médecine/sciences*, 2015. Vol. 31, No 11, pp. 1014-1022

J'ai fait usage de cet article à bien des reprises. Il a beaucoup appuyé ma recherche terminologique. Il m'a permis de bien appréhender le sujet. Sa rédaction est pointue et claire.

- Cherel, Y., Rouger, K., Larcher, T., et al. Thérapie cellulaire de la Dystrophie musculaire de Duchenne : Essai préclinique chez le chien myopathe. *Bulletin de l'Académie Vétérinaire de France*, Académie Vétérinaire de France, 2011. Vol. 164, No 3, pp. 211-216

L'avantage du thème étudié ici est qu'il concerne des données vétérinaires canines, là où bien des expériences sont faites avec des souris. Linguistiquement, nous pouvons considérer que les différences sont inexistantes. Mais sans connaissance du sujet, plus le domaine d'étude se rapproche de celui de mon texte support, plus la compréhension m'en est facilitée.

- Fardeau, M., Braun, S., Romero, NB. À propos d'un essai de Phase I de thérapie génique effectué avec un plasmide contenant l'intégralité du gène dystrophine dans la Myopathie de Duchenne/Becker. *Journal de la Société en biologie*, 2005. Vol. 199, pp. 29-32.

Cet article datant de 2005, il permet de voir les évolutions dans le domaine du traitement de la DMD par la thérapie génique, et ce de manière très spécifique. Il permet de même d'observer l'évolution des usages entre les années 2000 et les années 2020 dans la communauté scientifique française

Consultation citoyenne sur l'édition du génome en Île-de-France

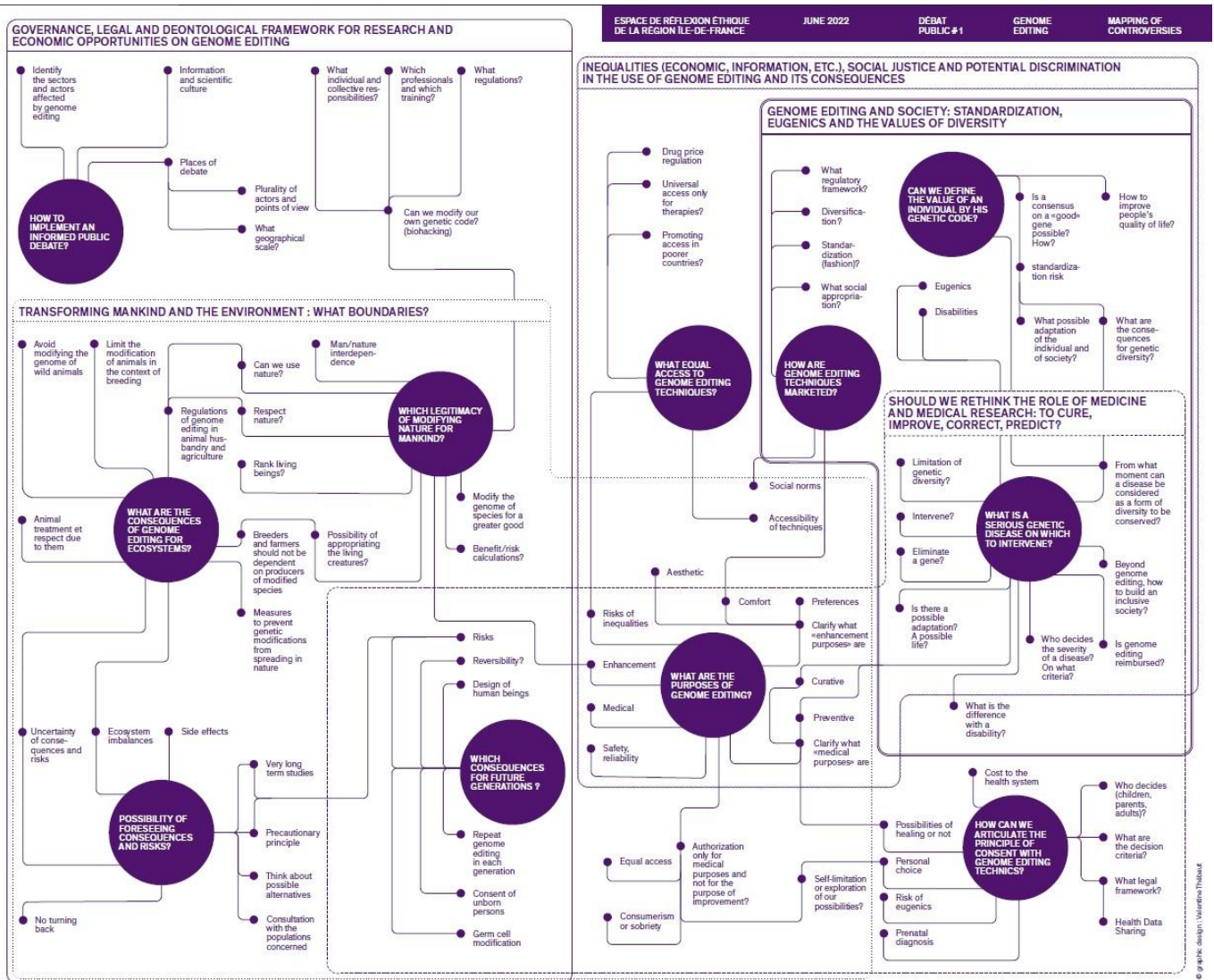
Dossier documentaire pour les participants au débat. Espace éthique de la région Île-de-France, en partenariat avec la Faculté de médecine de l'Université Paris-Saclay, l'Inserm, la Global Citizens' Assembly on Genome editing, et CNERER, 2021. 16 p.

Ce document m'a été conseillé par mon référent spécialiste. Il est composé d'une première partie sur la mise en contexte et l'histoire de l'édition du génome, d'une deuxième partie sur ses applications scientifiques et techniques, et d'une dernière partie sur son cadre réglementaire. En cela, il constitue une approche quasiment holistique de l'édition du génome, ce qui le rend inestimable. La qualité de la langue y est avérée et en adéquation avec son temps, l'approche est explicative soutenue. J'ai utilisé cette ressource notamment pour la partie Exposé (et plus particulièrement pour comprendre les enjeux éthiques de l'édition du génome), mais aussi pour étayer mon glossaire.

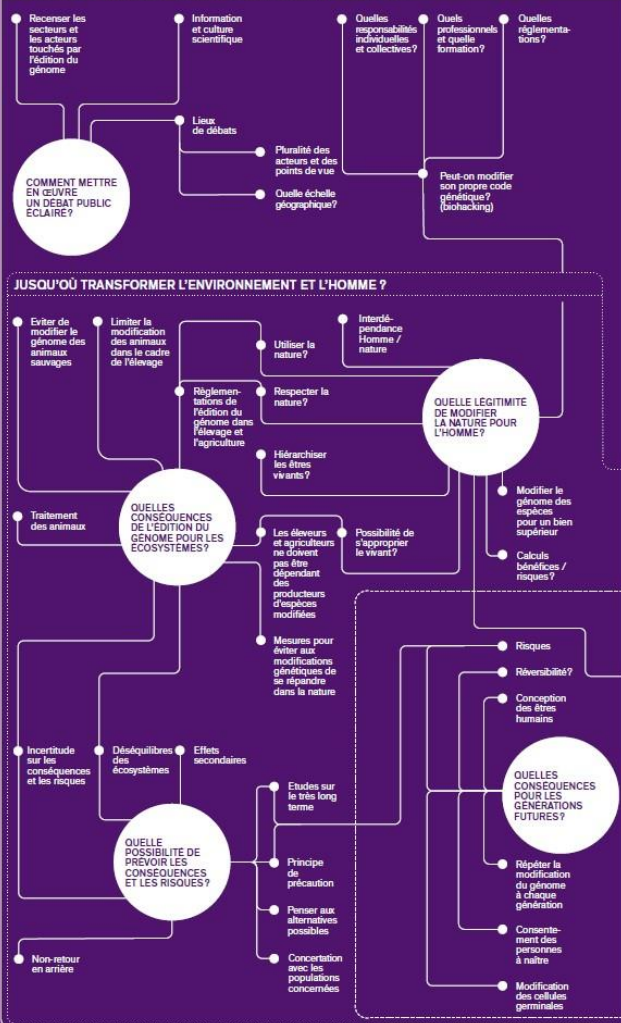
ANNEXE

Annexe 1

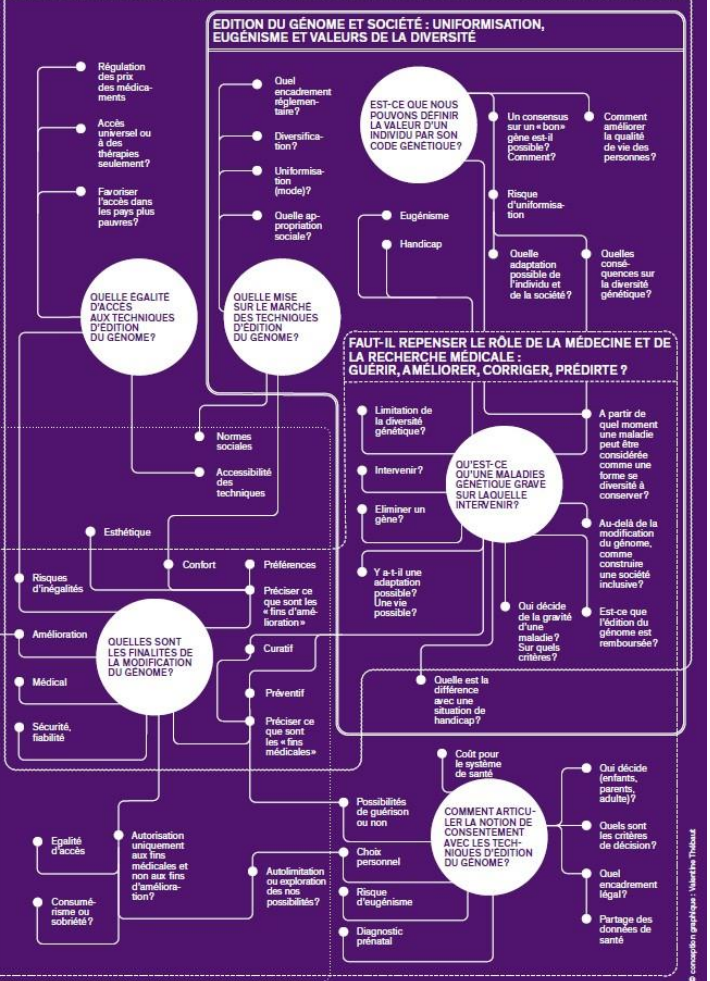
Cartographie des controverses, document émis en anglais et en français par l'Espace de réflexion éthique de la région Île-de-France, en partenariat avec la Faculté de médecine de l'Université Paris-Saclay. Juin 2022



GOUVERNANCE ET ENCADREMENT LÉGAL ET DÉONTOLOGIQUE DE LA RECHERCHE ET DES DÉBOUCHÉS ÉCONOMIQUES SUR L'ÉDITION DU GÉNOME



INÉGALITÉS, JUSTICE SOCIALE ET DISCRIMINATIONS POTENTIELLES DANS L'USAGE DE L'ÉDITION DU GÉNOME ET SES CONSÉQUENCES



© conception graphique : Wantha tribaux

Annexe 2

Traduction à deux colonnes

DMD is the demonstration of efficacy and safety of this approach in large mammals.

The mutation carried by the deltaE50-MD canine model of DMD leads to loss of exon 50 and moreover can be corrected by skipping of exon 51, making this a valuable model for translational studies. First identified as a naturally occurring spontaneous mutation in Cavalier King Charles Spaniels (23) and now maintained on a beagle background, this model (in contrast to mice) exhibits many of the clinical and pathological features of the human disease, such as muscle weakness, atrophy and fibrosis (24).

To correct the dystrophin reading frame in the deltaE50-MD canine model (henceforth referred to as ΔEx50) (Fig. 1A), we used *S. pyogenes* Cas9 coupled with a sgRNA to target a region adjacent to the exon 51 splice acceptor site (referred to as sgRNA-51) (Fig. 1B). The sgRNA-51 corresponded to a highly conserved sequence that differs by only one nucleotide between the human and dog genomes (Fig. S1A).

For the in vivo delivery of Cas9 and sgRNA-51 to skeletal muscle and heart tissue in dogs, we used recombinant adeno-associated virus serotype 9 (referred to as AAV9), which displays preferential tropism for these tissues (25, 26). A muscle-specific creatine kinase (CK) regulatory cassette was used to drive expression of Cas9; three RNA polymerase III promoters (U6, H1 and

de l'efficacité et de la sûreté de cette approche chez les mammifères de grande taille constitue une avancée majeure vers une application clinique de l'édition du génome en tant que stratégie thérapeutique pour la DMD.

La mutation portée par le modèle canin deltaE50-MD de la DMD se caractérise par la perte de l'exon 50 et peut être corrigée par saut de l'exon 51, ce qui rend ce modèle précieux pour les études d'application. Ce modèle a d'abord été identifié chez l'épagneul Cavalier King Charles comme une mutation naturelle et spontanée et elle est désormais maintenue chez le Beagle. À l'inverse du modèle des souris, le modèle canin partage beaucoup des aspects cliniques et pathologiques du modèle humain, tels que la faiblesse musculaire, l'atrophie et la fibrose (24).

Afin de corriger le cadre de lecture de la dystrophine dans le modèle canin deltaE50-MD (ci-après ΔEx50) (Fig. 1A), nous avons utilisé les protéines Cas9 de *S. pyogenes* couplées à un ARN guide permettant de viser une région adjacente au site d'épissage accepteur de l'exon 51 (ci-après ARNsg-51) (Fig. 1B). L'ARNsg-51 correspondait à une séquence hautement conservée, qui diffère d'un nucléotide seulement entre le génome humain et le génome canin (Fig. S1A). Associé à l'édition de ces séquences d'ARNsg-51, Cas 9 génère/induit une coupure génomique seulement dans l'ADN des espèces respectives (Fig. S1B).

Pour procéder à la livraison in vivo de Cas9 et de l'ARNsg-51 dans les muscles squelettiques et le muscle cardiaque des chiens, nous avons utilisé un virus adeno-associé recombinant de sérotype 9 (ci-après AAV9), qui manifeste une tropisme pour ces tissus (25, 26). L'utilisation d'une cassette régulatoire régulée de créatine-kinase (CK) a permis l'expression

TSK) directed expression of the sgRNA, as described previously in mice (Fig. S2) (15). AAV9-Cas9 and AAV9-sgRNA-51 were initially introduced into the cranial tibialis muscles of two 1-month-old dogs by intramuscular (IM) injection.

Muscles were analyzed 6 weeks after injection. In vivo targeting efficiency was estimated with muscle biopsy samples by RT-PCR and genomic PCR amplification products spanning the target site were subjected to amplicon deep-sequencing. The latter indicated that a mean of 9.96% of total reads contained changes at the targeted genomic site including insertions, deletions and substitutions (Fig. S3).

The most commonly identified mutations with a mean of 2.35% contained an adenosine (A) insertion immediately 3' to the Cas9 genomic cutting site (Fig. 1C). The deletions identified using this method encompassed a highly-predicted splicing enhancer (ESE) site for exon 51 (15, 27, 28) (Fig. S3A).

However, this method does not identify larger deletions that might occur beyond the annealing sites of the primers used for PCR. Since these tissue samples contain a mixture of muscle and non-muscle cells, the method likely underestimates the efficiency of gene editing within muscle cells.

de Cas9; trois promoteurs d'ARN polymérase III (U6, H1 et TSK) ont dirigé l'expression de l'ARNsg, comme précédemment décrit chez les souris (Fig. S2) (15). Les complexes AAV9-Cas9 et AAV9-ARNsg-51 ont d'abord été administrés à deux chiots d'1 mois via injection intramusculaire avec l'ARNsg-51 dans le muscle tibiaux crâniens.

Les muscles ont été analysés 6 semaines après injection. L'efficacité de ciblage in vivo a été évaluée au sein d'échantillons d'une biopsie musculaire par analyse RT-PCR avec des amorces spécifiques aux séquences des / dans les exons 48 et 53.

Des produits d'amplification du PCR génomique couvrant le site cible ont subi un séquençage profond des amplicons. Ce dernier a indiqué qu'en moyenne 9,96% des lectures contenaient des changements au niveau du site génomique ciblé, y compris des insertions, des délétions et des substitutions (Fig. S3). Parmi les mutations identifiées, les plus courantes, avec une moyenne de 2,35%, comprenaient une insertion d'adénosine (A), tout juste à l'extrémité 3' du site de découpage / épilage génomique de Cas9 (Fig. 1C). Les délétions identifiées grâce à cette méthode concernent en partie une séquence hautement prévisible pour l'exon 51 (15, 27, 28) (Fig. S3A).

Cependant, cette méthode ne permet pas d'identifier des délétions plus importantes pouvant avoir lieu au-delà des sites d'hybridation des amorces utilisées pour la PCR. Étant donné que ces échantillons de tissu contiennent à la fois des cellules musculaires et non musculaires, il est possible que la méthode employée sous-estime l'efficacité de l'édition du génome au sein de cellules musculaires.

Le séquençage des produits de

- **Eva Bar** Où est un promoteur de sgRNA assurant la transcription des 3 gènes AAV9 ?
- **Eva Bar** Pourquoi séquençage ? Pas possible car le séquençage de gènes par Cas9 dans ce séquençage par exon ?
- **Eva Bar** Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?

Sequencing of RT-PCR products of RNA from muscle of ΔEx50 dogs injected intramuscularly with AAV9-Cas9 and AAV9-sgRNA-51 showed that deletion of exon 51 (ΔEx50-51) allowed splicing from exon 48 to 52, which restores the dystrophin open reading frame (Fig. S3B). Deep-sequencing analysis so performed deep-sequencing analysis to quantify its abundance compared to other small insertions. Deep-sequencing of the upper band containing the non-edited cDNA product and reformed cDNA products indicated that a mean of 73.19% of total reads contained reframed cDNA products with an A insertion; a mean of 26.81% contained non-edited cDNA product, and the rest contained small deletions and insertions (Fig. S3C).

However, nonsense mediated decay might impact the abundance of non-edited cDNA products. These data indicate that the two ΔEx50 dogs injected with AAV9-Cas9 and AAV9-sgRNA-51 had a high frequency of

RT-PCR de l'ARN provenant des chiens ΔEx50 ayant reçu une injection intramusculaire des complexes AAV9-Cas9 et AAV9-ARNsg-51 a montré qu'une délétion de l'exon 51 (ΔEx50-51) permettait l'épissage de l'exon 48 à l'exon 52, rétablissant ainsi le cadre ouvert de lecture de la dystrophine (Fig. S3B). Sur gel, la taille du produit PCR contenant une insertion d'A était indissociable des produits d'ADN non édité. Nous avons donc effectué une analyse de séquençage profond pour quantifier son abondance comparée aux autres petites insertions. Le séquençage profond de la ligne supérieure contenant le produit d'ADN non édité et les produits d'ADN réassemblés montrés que parmi toutes les lectures, en moyenne 73,19% contenaient des produits d'ADN recadrés avec une insertion d'A, 26,81% contenaient des produits d'ADN non édité, et le reste contenait des délétions et des insertions mineures / micro-délétions et micro-insertions (Fig. S3C).

Cependant, la dégradation des ARN non-sensés pourrait avoir un effet sur la quantité élevée/abondance de produits d'ADN non édité. Ces données montrent que les deux chiens ΔEx50 ayant reçu une injection des complexes AAV9-Cas9 et AAV9-

(Fig. S4) by the CRISPR design tool (<http://crispr.mit.edu>). We performed deep sequencing at the top predicted off-target sites within protein coding exons. None of these sites revealed significantly more sequence alterations than the background analysis performed with other regions of the amplicons (Fig. S5).

To evaluate dystrophin correction at the protein level, we performed histological analysis of AAV9-injected cranial tibialis muscles 6 weeks after AAV9 injection. Dystrophin immunohistochemistry of muscle from ΔEx50 dogs injected with AAV9-Cas9 and AAV9-sgRNA-51 revealed widespread expression; the majority of fibers within the injected muscles expressed dystrophin, albeit to varying levels (Fig. 1D).

Western blot analysis confirmed the restoration of dystrophin expression in

(<http://crispr.mit.edu>) a relevé un total de 3 potentiels sites hors cible à l'échelle du génome (OT1 à OT3) sur les exons codants, et 4 dans les régions non-codantes (Fig. S4). Nous avons effectué un séquençage profond au niveau des premiers sites hors-cibles relevés pour les exons codants des protéines. Aucun de ces sites ne présentait significativement plus d'altérations de la séquence que dans d'autres régions des amplicons sujettes à des analyses de fond (Fig. S5).

Afin d'évaluer le rétablissement/correction de la dystrophine au niveau de la protéine, nous avons effectué une analyse histologique du muscle tibiaux crâniens injectés de AAV9 6 mois post injection. L'immunohistochimie du muscle chez les chiens ΔEx50 ayant reçu une injection des complexes AAV9-Cas9 et AAV9-ARNsg-51 a révélé une expression étendue/dépendante de la dystrophine : la majorité des fibres des muscles ayant reçu une injection affichait une expression de dystrophine étendue/dépendante au niveau du sarcolemme, même si les niveaux étaient variables (Fig. 1D).

L'analyse Western blot a permis de confirmer le rétablissement de l'expression

- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?
- **Eva Bar** Pourquoi séquençage ? Pourquoi séquençage ?

Annexe 3

Exemples des différentes utilisations des corpus comparables sur *Sketch Engine*.

Recherche simple [cibl*] dans le corpus FR (Concordance) et option Collocations

CONCORDANCE Get more space 🔍 📄 🔗 🔄 👤

simple cibl* • 486
1,143.63 per million tokens • 0.11%

COLLOCATIONS

BASIC ADVANCED ABOUT

Attribute? Range? KWIC GO

Collocations CHANGE CRITERIA BACK TO CONCORDANCE

	Word	Cooccurrences?	Candidates?	T-score	MI	LogDice ↓		Word	Cooccurrences?	Candidates?	T-score	MI	LogDice ↓	
1	<input type="checkbox"/> séquence	51	463	7.07	6.59	10.78 ...		11	<input type="checkbox"/> r	96	5,614	9.14	3.90	9.01 ...
2	<input type="checkbox"/> ADN	69	813	8.19	6.21	10.77 ...		12	<input type="checkbox"/> brin	10	151	3.11	5.86	9.01 ...
3	<input type="checkbox"/> hors	19	27	4.35	9.27	10.25 ...		13	<input type="checkbox"/> leur	12	345	3.35	4.93	8.89 ...
4	<input type="checkbox"/> ARNg	16	133	3.96	6.72	9.73 ...		14	<input type="checkbox"/> sur	24	1,211	4.62	4.12	8.86 ...
5	<input type="checkbox"/> gène	25	556	4.87	5.30	9.62 ...		15	<input type="checkbox"/> régions	8	33	2.79	6.23	8.82 ...
6	<input type="checkbox"/> exon	17	256	4.05	5.86	9.55 ...		16	<input type="checkbox"/> paire	7	35	2.63	7.45	8.78 ...
7	<input type="checkbox"/> région	12	136	3.42	6.27	9.30 ...		17	<input type="checkbox"/> qui	22	1,179	4.40	4.03	8.76 ...
8	<input type="checkbox"/> 54	10	74	3.14	6.88	9.19 ...		18	<input type="checkbox"/> Le	16	761	3.78	4.20	8.72 ...
9	<input type="checkbox"/> thérapie	12	219	3.39	5.58	9.12 ...		19	<input type="checkbox"/> la	114	8,415	9.78	3.57	8.71 ...

Différence d'usage entre « rétablir » et « restaurer » grâce à Word Sketch Difference puis au concordancier

WORD SKETCH DIFFERENCE CRISPR_FR_2

rétablir 26x | restaurer 36x

Show relation
objects of "rétablir/restaurer"

← mostly with rétablir | equally frequently with rétablir and restaurer | mostly with restaurer →

CONCORDANCE CRISPR_FR_2

CQL cadre + restaurer • 7
16.47 per million tokens • 0.0016%

Left context | KWIC | Right context

1	ulaval.ca	Ngs ciblant des introns permettait de générer de larges délétions et de restaurer un cadre de lecture propice à l'expression d'une dystrophine tronquée
2	ulaval.ca	ique. Cette délétion permet la formation d'un exon hybride qui restaure non seulement le cadre de lecture du gène de la dystrophine[1], mais
3	theses.fr	jène de la dystrophine. La technique du saut d'exon permet de restaurer le cadre de lecture dans le cas de nombreuses mutations et d'obtenir
4	uliege.be	sociés AAV9. Des ARNGs ciblant l'exon 51 ont été utilisés pour restaurer le cadre de lecture du gène de la dystrophine chez des chiens dystrop
5	theses.fr	e avec le système CRISPR/Cas9. Cette approche a permis de restaurer le cadre de lecture du gène, l'expression d'une dystrophine entière. C
6	===NONE===	rigée par la cellule par recombinaison non homologue, ce qui pourrait restaurer le bon cadre de lecture par les INDELS qui se forment lors de la répar
7	===NONE===	à la dystrophine, elles sont dues à des événements de saut d'exon qui restaurent le cadre ouvert de lecture du gène [6]. Aspects cliniques D'un

CONCORDANCE CRISPR_FR_2

CQL cadre + rétablir • 6
14.12 per million tokens • 0.0014%

Left context | KWIC | Right context

1	canada.ca	recombinaison non homologue (NHEJ) pour produire un exon hybride, rétablir le cadre de lecture et permettre la synthèse d'un edystrophine tronquée
2	===NONE===	lure, lors de l'épissage de l'ARNm, un ou plusieurs exons mutés afin de rétablir le cadre ouvert de lecture [57]. Pour cela, un ou plusieurs oligor
3	===NONE===	ion entraine un saut d'exon de l'exon 51 lors de l'épissage et permet de rétablir le cadre de lecture. (B) Illustration du système à deux AAVs sér
4	===NONE===	ture, soit l'épissage d'un ou plusieurs exons additionnels permettant de rétablir le 21 cadre de lecture de l'ARNm chez des patients ayant une délétion
5	===NONE===	ertions ou des micro-délétions au niveau du site de coupure pourraient rétablir le cadre de lecture de l'ARNm. Cette approche fut testée dans l
6	===NONE===	ilisé sur 7 patients dystrophiques dans le but d'épisser l'exon 51 afin de rétablir le cadre de lecture de leur dystrophine. Le morpholino fut inject

INDEX

A

acide aminé, 14, 15, 22, 23, 39, 135, 137, 142
Acide aminé, 14, 23, 42, 135, 137, 143, 148
activatrice d'épissage exonique (AEE), 66, 143, 149
ADN, 5, 7, 8, 10, 11, 12, 13, 16, 17, 18, 19, 20, 21, 22, 23, 25, 28, 29, 33, 38, 40, 41, 42, 43, 44, 45, 46, 47, 50, 52, 53, 55, 62, 64, 72, 76, 87, 100, 117, 126, 128, 130, 134, 135, 136, 139, 140, 141, 142, 143, 148, 149
ADN complémentaire, 18, 66, 68, 72, 107, 128, 136, 143, 148
ADN mitochondrial, 10, 28, 135, 143
ADN nucléaire, 10, 38, 40, 135, 143
Allèle, 20, 140, 143, 148
Amorce, 53, 66, 98, 128, 134, 143, 151
Amplicon, 52, 53, 65, 66, 67, 68, 71, 72, 116, 142, 146, 148
Anomalie, 20, 22, 23, 24, 28, 29, 35, 143, 148
ARN de transfert, 13, 14, 16, 137, 143, 152
ARN guide simple, 64, 66, 68, 70, 72, 74, 76, 80, 82, 86, 95, 101, 118, 136, 143, 151
ARN messenger, 12, 13, 14, 16, 22, 38, 39, 40, 135, 136, 137, 140, 143, 150, 151
ARN polymérase, 11, 13, 18, 53, 64, 136, 143, 151
ARN pré-messenger, 12, 34, 38, 44, 136, 143, 151
ARN ribosomique, 13, 14, 143, 151
Atrophie, 62, 64, 138, 143, 148

B

Base azotée, 7, 8, 141, 143, 150

C

Cadre de lecture, 12, 22, 23, 33, 39, 46, 62, 64, 66, 68, 74, 80, 100, 107, 132, 138, 143, 144, 147, 149, 151, 157
Cardiomyocyte, 75, 76, 141
Cardiomyocytes, 75, 76, 143, 148
Cas9, 5, 36, 40, 43, 44, 45, 46, 47, 48, 52, 55, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 94, 95, 96, 100, 101, 106, 109, 112, 116, 117, 118, 119, 121, 125, 126, 128, 129, 130, 132, 134, 136, 137, 138, 140, 143, 148, 154, 155, 159
Cassure double-brin, 143, 149
Cellule, 5, 7, 9, 10, 11, 13, 14, 16, 21, 24, 25, 27, 28, 29, 31, 32, 33, 35, 36, 38, 40, 41, 42, 44, 45, 46, 47, 49, 50, 51, 52, 55, 57, 64, 66, 70, 73, 74, 76, 78, 98, 104, 109, 110, 111, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 148, 149, 150, 151, 152, 154
Cellule embryonnaire, 143, 149
Cellule eucaryote, 143, 149
Cellule fille, 144, 149
Cellule post-mitotique, 144, 151
Cellule souche, 136, 137, 144, 150, 152
Cellule souche pluripotente induite, 144, 150
Chromosome, 7, 11, 17, 24, 25, 26, 27, 28, 29, 33, 136, 140, 143, 144, 145, 148
chromosome somatique, 143, 148
Codon, 12, 14, 22, 23, 38, 39, 73, 74, 137, 139, 144, 148, 152
Codon stop, 13, 22, 38, 39, 74, 137, 144, 152
Collapsus respiratoire, 34, 62, 145, 151

Coloration à l'hématoxyline et l'éosine, 70, 144, 150
complexe dystrophine-glycoprotéine, 33, 70, 137, 144, 149
Contraction, 32, 33, 61, 62, 107, 141, 144, 148, 157
Coupure, 44, 45, 47, 64, 70, 76, 117, 128, 130, 144, 147, 149, 152
Coupure génomique, 64, 144, 149
Coupure unique, 144, 152
Créatine kinase, 35, 63, 64, 73, 74, 101, 144, 148
CRISPR, 5, 36, 40, 43, 44, 46, 47, 48, 52, 55, 59, 60, 63, 64, 67, 68, 75, 76, 77, 78, 79, 80, 87, 106, 109, 117, 119, 121, 125, 126, 128, 129, 130, 132, 134, 136, 137, 138, 140, 141, 144, 148, 154, 155, 157, 159, 160
Cytoplasme, 9, 10, 14, 16, 30, 32, 135, 144, 149
Cytosquelette, 32, 62, 70, 138, 144, 149

D

Décalage du cadre de lecture, 22, 138, 144, 149
Dégénérescence cellulaire, 144, 148
Dégénérescence musculaire, 144, 150
Dégradation des ARNm médiée par les brins non-sens, 38, 40, 47, 68, 144, 151, 160
Dégradation des ARNm non-sens, 38, 40, 47, 68, 144, 151, 160
Délétion, 23, 25, 33, 46, 53, 66, 68, 74, 76, 117, 126, 138, 140, 144, 149
Dystroglicane, 144, 149
Dystrophie, 31, 34, 35, 39, 49, 60, 62, 78, 87, 96, 109, 110, 115, 132, 138, 144, 148, 149, 158, 160, 161
Dystrophie musculaire de Becker, 31, 78, 138, 144, 148
Dystrophie musculaire de Duchenne, 31, 32, 33, 34, 35, 39, 41, 48, 50, 59, 60, 61, 62, 63, 64, 69, 70, 73, 74, 75, 76, 81, 82, 85, 86, 87, 96, 109, 115, 119, 120, 121, 131, 132, 138, 144, 149, 158, 160, 161
Dystrophine, 32, 33, 34, 35, 50, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 95, 96, 108, 117, 132, 137, 138, 144, 149, 160, 161

E

Échantillon, 50, 52, 53, 66, 72, 74, 118, 119, 120, 145, 151
Édition du génome, 5, 40, 43, 46, 47, 49, 50, 52, 53, 57, 58, 60, 62, 64, 66, 68, 70, 72, 76, 78, 80, 84, 86, 89, 92, 104, 106, 110, 126, 136, 142, 145, 149, 154, 159, 161
Enzyme, 13, 14, 34, 40, 43, 44, 52, 53, 100, 117, 136, 138, 141, 145, 149
Exon, 11, 12, 23, 33, 39, 61, 62, 63, 64, 65, 66, 67, 68, 73, 74, 77, 79, 80, 94, 95, 106, 130, 131, 132, 134, 137, 139, 140, 145, 147, 149

F

Fibrose, 32, 35, 64, 70, 145, 149

G

Gamète, 7, 27, 28, 29, 57, 139, 145, 149
Gène, 5, 7, 8, 10, 11, 13, 17, 18, 19, 20, 21, 23, 25, 26, 27, 28, 29, 31, 32, 33, 35, 36, 37, 38, 39, 40, 41, 42, 43, 46, 47, 48, 57, 60, 62, 64, 99, 100, 115, 118, 128, 130, 135, 136, 138, 139, 140, 141, 142, 145, 149, 158, 160, 161
Génome, 5, 7, 16, 17, 19, 20, 24, 25, 27, 29, 30, 36, 40, 43, 44, 46, 48, 52, 55, 57, 64, 68, 72, 74, 87, 111, 112, 113, 126, 128, 136, 139, 142, 145, 149, 155, 157, 161

H

Hématologie, 145, 150
Héréditaire, 145
Histologie, 5, 62, 86, 95, 96, 109, 139, 145, 150
Hors cible, 47, 68, 76, 130, 145, 151
Hybridation, 53, 66, 128, 134, 145, 148

I

Immunogénicité, 40, 41, 47, 78, 145, 150
Immunomarquage, 50, 74, 76, 84, 97, 139, 145, 150
Immunosuppression, 145, 150
Indel, 46, 53, 71, 72, 73, 74, 116, 117, 118, 125, 129, 140, 145, 150
Infiltration cellulaire, 35, 70, 145, 148
Injection systémique, 60, 70, 72, 84, 139, 145, 152
Insertion, 23, 46, 53, 65, 66, 67, 68, 71, 72, 73, 74, 106, 117, 126, 138, 140, 145, 150
Insuffisance respiratoire, 34, 62, 145, 151
Intégrité musculaire, 32, 76, 145, 150
Intron, 140, 145

K

Kilobase, 17

L

Locus, 11, 42, 73, 74, 140, 145, 150

M

Maladie chromosomique, 24, 145, 148
Maladie dégénérative, 145, 149
Maladie héréditaire, 28, 145, 150
Maladie monogénique, 31, 145, 152
Maladie polygénique, 26, 146, 151
Marqueur, 17, 21, 51, 70, 109, 110, 118, 146, 152
Mégabase, 17, 62, 146, 150
Mitochondrie, 146, 150
Mitose, 9, 10, 146, 150
Motif PAM, 44, 74, 140, 146, 151
Muscle, 14, 31, 32, 33, 34, 50, 59, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 101, 107, 109, 110, 116, 117, 118, 119, 121, 131, 138, 141, 143, 144, 145, 146, 148,

150, 152

Muscle squelettique, 146, 152
Mutagénèse, 76, 141, 146, 150
Mutation, 5, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 31, 33, 38, 42, 43, 57, 59, 60, 61, 62, 63, 64, 65, 66, 71, 72, 73, 74, 87, 107, 108, 120, 126, 131, 138, 140, 141, 146, 149, 150, 152
Mutation *de novo*, 146, 149
Mutation induite, 146, 150
Mutation spontanée, 146, 152
myocarde, 146, 148
Myocyte, 10, 31, 32, 35, 64, 66, 76, 109, 110, 115, 136, 137, 138, 141, 143, 146, 148, 150
Myofibrille, 32, 115, 146, 150
Myopathie, 20, 31, 35, 50, 115, 132, 138, 146, 150, 160, 161
Myosine, 32, 70, 110, 146, 150

N

Nucléotide, 7, 8, 11, 12, 14, 20, 22, 23, 44, 46, 53, 64, 107, 117, 128, 130, 137, 140, 141, 146, 151

O

Œdème, 70, 146, 149
Oligonucléotide, 38, 62, 78, 126, 134, 146, 151

P

Paire de bases, 146, 148
Pathologie, 146
PCR, 50, 52, 65, 66, 71, 72, 79, 80, 98, 127, 128, 133, 134, 138, 141, 142, 146, 147, 148, 151, 155, 156
PCR quantitative, 71, 72, 146, 151, 156
Phénotype, 21, 23, 25, 38, 107, 119, 141, 146, 151, 152
Phénotype sauvage, 141, 146, 152
Polypeptide, 14, 15, 23, 137, 146, 151
Promoteur, 12, 21, 42, 64, 134, 142, 147, 151

Protéine, 10, 11, 12, 13, 14, 15, 16, 18, 22, 23, 30, 31, 32,
36, 38, 39, 40, 43, 46, 50, 51, 53, 60, 62, 68, 70, 78,
87, 100, 101, 107, 132, 135, 136, 137, 138, 139, 142,
147, 151, 157

R

Réarrangement génomique, 25, 76, 147, 149
Recherche translationnelle, 55, 147, 152
Recombinaison homologue, 45, 46, 130, 147, 150
Recombinaison non homologue, 46, 130, 147, 151
Relâchement, 62, 147, 151
Rétablissement du cadre de lecture, 147, 151
Révertant, 142, 147, 151
Ribosome, 14, 16, 22, 38, 135, 137, 147, 151
RT-PCR, 65, 66, 79, 80, 147, 151

S

Sarcolemme, 32, 62, 68, 115, 147, 151
Sarcomère, 32, 115, 147, 151
Saut d'exon, 32, 39, 78, 132, 147, 149

Séquençage, 16, 17, 18, 19, 24, 30, 50, 52, 55, 66, 68, 72,
74, 76, 96, 111, 147, 151

Site accepteur d'épissage, 64, 147, 152

Site de coupure, 46, 80, 147, 149

Substitution, 23, 65, 66, 126, 147, 152

T

Thérapie génique, 5, 6, 34, 35, 36, 37, 38, 41, 42, 49, 55,
56, 57, 58, 78, 88, 89, 94, 95, 96, 104, 106, 109, 110,
111, 117, 123, 126, 130, 132, 134, 142, 147, 149, 156,
158, 159, 160, 161

Transgène, 34, 37, 42, 132, 142, 147, 152, 160

V

Vecteur viral, 41, 70, 72, 74, 147, 152, 160

Vinculine, 82, 86, 147, 152

Virus, 9, 41, 43, 44, 60, 63, 64, 65, 66, 67, 68, 69, 70, 71,
72, 73, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 101, 106,
112, 113, 116, 118, 119, 121, 127, 139, 147, 148, 152

Virus adénoassocié, 41, 67, 68, 101, 113, 119, 147, 148