



HAL
open science

Nouvelle technique de détection simultanée des variant ponctuels et des copy number variants dans l'obésité monogénique

Mehdi Derhourhi

► **To cite this version:**

Mehdi Derhourhi. Nouvelle technique de détection simultanée des variant ponctuels et des copy number variants dans l'obésité monogénique. Médecine humaine et pathologie. Université de Lille, 2018. Français. NNT : 2018LILUS029 . tel-02077477

HAL Id: tel-02077477

<https://theses.hal.science/tel-02077477>

Submitted on 22 Mar 2019

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

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

UNIVERSITÉ DU DROIT ET DE LA SANTÉ - LILLE

ÉCOLE DOCTORALE BIOLOGIE - SANTÉ

Année universitaire 2018-2019

Thèse Soutenue en Vue de l'Obtention du Titre de Docteur de l'Université de Lille

Discipline : *Génétique, génomique, bio-informatique*

Spécialité : *Génétique humaine*

Nouvelle technique de détection simultanée des variant ponctuels et des copy number variants dans l'obésité monogénique

Par Mehdi DERHOURHI

Sous la direction du **Docteur Amélie BONNEFOND**

Soutenue le 19 décembre 2018 devant le jury composé de :

Docteur Richard REDON	Rapporteur
Professeur Karine CLEMENT	Rapporteur
Docteur Philippe LEFEBVRE	Examineur
Professeur Jacques BECKMANN	Examineur
Docteur Amélie BONNEFOND	Directrice de thèse

Remerciements :

Je tiens à remercier les membres de mon jury, Karine Clement, Jacques Beckmann, Richard Redon et Philippe Lefebvre pour avoir accepté d'évaluer mon travail de thèse. Merci également au docteur Jean Muller pour sa participation à mes comités de suivi de thèse.

Merci à Philippe Froguel de m'avoir permis d'effectuer cette thèse et d'avoir mis à ma disposition les moyens nécessaires à sa réalisation.

Merci à Amélie Bonnefond, ma directrice de thèse, de m'avoir encadré durant ces 3 ans, de m'avoir formé et aidé, et sans qui ce projet n'aurait jamais vu le jour.

Merci à Louise Montagne avec qui j'ai fait équipe durant ce projet pour son aide précieuse durant toute sa durée.

Merci à Dorothée Thuillier pour m'avoir appris à coder, et sa participation à la réalisation de ce projet.

Merci à Bénédicte Toussaint pour toute son aide et sa contribution lors de toute les étapes de cette thèse.

Merci à toute l'équipe du laboratoire avec qui j'ai eu grand plaisir à travailler, avec qui j'ai beaucoup appris, mais aussi avec qui je me suis bien amusé durant ces années.

Merci à mes parents et mon frère pour leur indéfectible soutien et confiance dans tout ce que j'ai fait. Je vous dois ce que je suis devenu et ce que je deviendrai, et ne pourrai jamais assez vous en remercier.

Et enfin merci à l'ensemble de ma famille et de mes amis pour m'avoir accompagné durant ces années et qui ont tous contribué d'une façon ou d'une autre à l'aboutissement de ce projet.

Résumé :

La génétique, et par extension le séquençage de l'ADN, sont des outils qui ont transformé la compréhension des mécanismes impliqués dans la survenue de nombreuses pathologies, dont l'obésité. Les technologies aujourd'hui à notre disposition nous permettent de déterminer rapidement si un patient est ou non porteur d'un évènement génétique pouvant expliquer sa pathologie. L'une des techniques les plus utilisées en diagnostic aujourd'hui est le séquençage d'exome, ou WES, qui permet une excellente détection des mutations ponctuelles dans les régions codantes du génome. Mais d'autres évènements comme les *copy number variants*, ou CNV, peuvent également expliquer certaines pathologies, dont l'obésité, via entre autres les CNV de la région 16p11.2. Actuellement, la technique de référence pour la détection de ces *copy number variants* est l'analyse de puces CGH (*Comparative Genomic Hybridization*), mais celles-ci ne permettent pas de détecter des mutations non répertoriées au préalable lors de la création de la puce. Sur le principe, le séquençage d'exome peut lui aussi être utilisé pour détecter les CNV, mais son absence de couverture des régions non codantes du génome ne permet pas une détection efficace de ces CNV, car ceux-ci peuvent survenir sur l'ensemble du génome, en englobant des régions codantes et non codantes au sein d'un seul évènement. Le séquençage génome complet peut détecter ces deux types d'évènement, mais son cout est encore élevé ce qui freine sa démocratisation, et l'analyse de données associées nécessite d'importantes ressources informatiques, et le rend difficilement utilisable en diagnostic de routine en l'état actuel des choses.

Il est donc pour l'instant nécessaire d'avoir recours à deux techniques différentes pour couvrir ces deux types d'évènements génétiques. Cela implique d'utiliser des échantillons parfois très précieux à deux reprises, de supporter les couts liés à deux techniques diagnostiques (d'environ 450 euros pour le séquençage d'exome au laboratoire et un cout un peu plus élevé pour une puce à ADN dans un laboratoire clinique), et d'allonger les temps de rendu de résultats et donc la durée d'établissement du diagnostic du patient.

Cet état de fait nous a conduit à développer une technique de séquençage, que nous avons nommé CoDE-seq (*Copy number variation Detection and Exome sequencing*), et qui permettra la

détection simultanée de ces deux types d'évènements, pour diminuer les temps d'établissement de diagnostics, leurs couts, et la quantité d'échantillon nécessaire.

Ce travail a nécessité deux aspects : la mise au point technique et la mise au point analytique.

La mise au point technique est passée par la création d'une nouvelle « capture », permettant une détection correcte des mutations ponctuelles de l'exome et des CNV de tout le génome.

La mise au point analytique a consisté à définir la méthode à employer, et à permettre d'arriver à une détection fiable, à la fois sensible et spécifique, des CNV sur l'ensemble du génome. Une fois ces CNV identifiés, la question de leur signification fonctionnelle se pose également, et une seconde partie de ma thèse porte sur l'étude de cette signification fonctionnelle, via l'étude de la conformation spatiale de la chromatine et de l'influence des CNV sur celle-ci.

Abstract:

Genetics, and by extension DNA sequencing, are tools that have modified the understanding of the mechanisms involved in genetic diseases, like obesity. Today's technology has allowed us to rapidly find if a patient carries a genetic event that may explain his/her pathology. One of the most used technology for diagnostic is exome sequencing, or WES, which enables an excellent detection of point mutations in coding regions of the genome. However other events, such as copy number variations, or CNV, can also explain some pathologies, like a severe form of obesity due to CNV in the chr16p11.2 region. Actually, the gold standard method for an accurate detection of CNV is array CGH, but this technology cannot detect new point mutations. Exome sequencing can be used to detect CNV, but the lack of coverage in non-coding regions limits CNV detection sensitivity. Of note, whole genome sequencing can detect both CNVs and point mutations, but it is still very expensive and needs huge informatics capacities, which is an obvious limitation for a routine diagnostic use.

For now, we have had to use two different methods in order to accurately detect both CNVs and point mutations. In other words, we have had to use precious samples two times, to assume the cost of two different methods (which is nearly 450 euros in the laboratory for exome sequencing, and a bit more for array CGH in a clinical laboratory), and to consider the time of the realization of two different methods in order to achieve a complete diagnostic.

In this context, we aimed to develop an innovative sequencing method, named CoDE-seq (Copy number variation Detection and Exome sequencing), which would allow us to simultaneously detect both CNVs and point mutations, in order to reduce the time of diagnostic, the cost, and the needed quantity of sample.

This work included the method conception, and the data analysis steps. The method conception has been done through the creation of a new capture enabling the detection of point mutations in the exome, and CNVs all along the genome. Furthermore, the data analysis step included the choice of the bioinformatics methods to be used, in order to get a specific and sensitive CNV detection, all along the genome.

We were also interested in the functional significance of identified CNV, and tried to decipher it by the study of chromatin spatial conformation and the influence of these CNV.

Table des matières :

Remerciements :	2
Résumé :	3
Abstract:	5
Liste des abréviations :	9
Introduction :	10
L'obésité :	10
<i>Epidémiologie :</i>	15
Prévalence et tendances :	15
Facteurs de risques connus :	17
Le rôle de l'empreinte métabolique :	20
Comorbidités et mortalité :	21
<i>Mécanismes / Physiopathologie :</i>	23
Le tissu adipeux :	23
Types :	23
<i>Changements pathologiques dans le tissu adipeux :</i>	25
Régulation de la balance énergétique :	27
<i>Le poids de la génétique et de l'épigénétique :</i>	28
Epigénétique :	31
Etudes gènes-candidats :	32
Mutation des gènes de la leptine et du récepteur de la leptine :	33
<i>Régulation de la prise alimentaire :</i>	33
<i>Régulation de la dépense énergétique et de l'oxydation des graisses :</i>	34
<i>La leptine en tant que déclencheur métabolique de la puberté :</i>	35
<i>La leptine comme médiateur de la régulation nutritionnelle de la fonction immune :</i>	36
Mutations perturbant la signalisation via la mélanocortine :	36
<i>Régulation de la prise énergétique, de la dépense énergétique et de l'oxydation des graisses :</i>	38
<i>Régulation de la croissance :</i>	39
Mutations des molécules en aval de MC4R :	39
Syndromes d'obésité développementale incluant un dysfonctionnement ciliaire :	40
Les désordres génétiques sujets à empreinte :	41
Les Copy Number Variants :	42
Les variants génétiques communs associés à un risque plus élevé de survenue de l'obésité :	43
Trouver de nouvelles étiologies dans l'obésité humaine :	44
<i>Diagnostic, recherche et prévention :</i>	45
Phénotypes obèses :	45
Evaluation :	46
Prévention :	47
Interventions sur le style de vie :	50
L'effet des Régimes :	50
Activité physique :	51
Gestion médicalisée :	52
Chirurgie bariatrique :	53
Qualité de vie :	55
Nutrition de précision individualisée :	57
<i>Perspectives :</i>	58
Physiopathologie :	58

Prévention :.....	58
Médecine de précision :.....	59
Séquençage de l'ADN :.....	60
Les débuts du séquençage :.....	60
L'invention du séquençage de l'ADN :.....	61
Le passage au génome humain :.....	63
Le séquençage d'ADN massivement parallèle :.....	65
Le séquençage en temps réel sans amplification :.....	68
Les puces à ADN ou puces CGH :.....	72
Les applications du séquençage de l'ADN :.....	73
L'assemblage de génomes <i>de novo</i> :.....	73
Le reséquençage de génomes :.....	75
Les applications cliniques du séquençage :.....	77
Méthodes de détection des copy number variants :.....	78
Méthode de l'alignement des paires de <i>reads</i> :.....	78
Méthode des <i>reads</i> tronqués :.....	79
Méthode de la profondeur des <i>reads</i> :.....	79
Méthode de l'assemblage <i>de-novo</i> :.....	80
La détection des CNV par le séquençage de longs fragments :.....	81
Le futur du séquençage de l'ADN :.....	82
Déterminer la diversité du génome :.....	82
Séquençage de populations :.....	82
Biologie développementale :.....	82
Séquenceurs temps-réel portables :.....	83
Autres usages :.....	83
Mon travail de thèse :.....	84
Matériel et méthodes :.....	86
Echantillons et cohortes utilisées :.....	86
Conception de la méthode :.....	87
<i>Capture</i> :.....	87
<i>Design de la capture</i> :.....	88
Librairie :.....	90
Enrichissement :.....	91
Séquençage :.....	93
Analyse bioinformatique et analyse des variants :.....	93
Résultats :.....	98
Discussion et Perspectives :.....	109
Conformation spatiale de la chromatine :.....	111
NG-CaptureC :.....	113
Matériel et méthode :.....	117
<i>Analyse Bioinformatique</i> :.....	119
Résultats préliminaires :.....	119

Discussion :..... 123
Bibliographie : 125
Supplementary data :..... 145

Liste des abréviations :

ACTH : Adréno Cortico Tropic Hormone	ARN : Acide Ribonucléique
ADCY3 : adenylyl cyclase 3	ARNm : Acide Ribonucléique messenger
ADN : Acide Désoxyribonucléique	BAC : Bacterial Artificial Chromosome
AHO : dystrophie héréditaire d'Albright	BBS : Bardet-Biedl Syndrome
AMPC: Adénosine monophosphate cyclique	
BDNF : Brain-Derived Neurotrophic Factor	DXA : absorptiométrie biphotonique à rayons X
BED : Browser Extensible Data	FDA : Food and Drug Administration
bp : paire de bases	FISH : Fluorescent In Situ Hybridization
CGH : Comparative Genomic Hybridization	FTO: Fat mass and obesity
CNV : Copy Number Variant	Gb : giga base
db/db: Diabetic mouse	GH : Growth Hormone
dNTP : deoxyribo Nucléotide Tri Phosphate	
GHRH : GH-releasing hormone	kb: Kilobase
GPR54 : G Proteine Receptor 54	Kcal : kilo calorie
GWAS: Genome Wide Association Study	LEP : leptine
HGP : Human Genome Project	LEPR : Leptine Receptor
Hi-C : genome wide conformation capture	MAF : Minimum Allele Frequency
HIF3A : facteur de transcription inductible par hypoxie 3A	Mb : mega base
IMC : Indice de Masse corporelle	MC4R : récepteur 4 de la mélanocortine
IRM: Imagerie par Résonance Magnétique	mmHg : millimètre de mercure
IRX3 : Iroquois-class homeodomain protein 3	MSH : Melanocyte Stimulating Hormone
IRX5 : Iroquois-class homeodomain protein 5	NGS : Next Generation Sequencing
ob/ob: Obese mouse	NIPT : dépistage prénatal non invasif
OGTT: Oral Glucose Tolerance Test	RNA-seq: Ribonucleic Acid sequencing
OMS : Organisation Mondiale de la santé	SBS : Sequencing By Synthesis
OSV : Obésité Saint-Vincent	SH2B1 : Src homology 2 B adaptor protein 1
pb: paire de base	SIM1 : single minded 1
PCR: Polymerase Chain Reaction	SNP: Single Nucleotide Polymorphism
PCSK1 : pro-hormone convertase 1	TrkB : tyrosine kinase receptor tropomyocin-related kinase B
POMC : pro-opiomélanocortine	µg : micro gramme
PWS : Prader-Willi Syndrome	WAGR : Wilms tumour-aniridia syndrome
QMPSF : quantitative multiplex polymerase chain reaction of short fluorescent fragments	WES : Whole Exome Sequencing
qPCR: quantitative Polymerase Chain Reaction	WGS : Whole Genome Sequencing

Introduction :

L'obésité :

L'obésité est considérée comme une épidémie du 21^{ème} siècle. Elle est définie par une augmentation de la masse grasse ayant des conséquences pour la santé, avec une accumulation excessive de tissu adipeux accompagnée par une inflammation légère, chronique et systémique. L'obésité est associée au diabète de type 2, aux maladies cardio-vasculaires, à certains types de cancers ainsi que d'autres pathologies (1). Certains des signes cliniques associés à l'obésité, et en particulier une distribution abdominale des graisses, définissent le syndrome métabolique (2), qui est considéré comme facteur de risque du diabète de type 2 et des maladies cardio-vasculaires. De façon simplifiée, on peut considérer qu'une augmentation de la masse grasseuse résulte d'un déséquilibre entre les calories absorbées et les dépenses énergétiques. Selon ce point de vue, l'obésité est la résultante d'une faible activité physique (un style de vie sédentaire), et d'une surconsommation de nourriture à haute teneur énergétique dépassant les besoins de l'individu. Mais en réalité, l'étiologie de l'obésité est bien plus complexe (3). D'autres facteurs comme le statut socio-économique, l'environnement et les habitudes de vie, ainsi que l'interaction génotype-phénotype doivent être pris en considération pour comprendre la survenue de l'obésité, étant donné que tous ces facteurs influencent la prise alimentaire, la thermogénèse, la mobilisation des acides gras depuis leur stockage vers leur oxydation, et la répartition des dépôts de graisse dans des régions de stockage ou non.

Le contrôle du poids corporel ainsi que sa composition doivent prendre en compte l'absorption énergétique et la dépense énergétique, ainsi que le dépôt de masse grasseuse, qui sont interconnectés et sous le contrôle des systèmes nerveux et endocrine, auxquels participent différentes hormones et neurotransmetteurs (4). De nombreux éléments modifient ces processus de régulation : les facteurs environnementaux (manque de sommeil, changement d'horaires de travail, ou température ambiante), le régime alimentaire, le niveau d'activité physique, le microbiote intestinal, les perturbateurs endocriniens (représentant les éléments chimiques intervenant sur le système endocrinien), les facteurs liés à la reproduction (comme une fertilité

plus importante associée avec une augmentation de la masse adipeuse), les médicaments, ou les facteurs épigénétiques avec effets intergénérationnels (5).

Une autre difficulté rencontrée lors de l'étude de l'obésité est l'hétérogénéité des individus obèses. Aux deux extrêmes, on trouve l'obésité gynoïde, lors de laquelle un excès de graisse sous-cutanée est retrouvé autour des hanches et de la taille (corps en forme de poire ou obésité gynoïde, commune chez les femmes), et l'obésité viscérale, lors de laquelle la graisse (principalement du tissu adipeux omental et mésentérique) est concentrée dans la région abdominale (corps en forme de pomme ou obésité androïde). L'obésité viscérale est commune chez les hommes et a tendance à avoir un impact plus important sur la santé, en particulier concernant les risques de maladies cardiovasculaires.

Pour déterminer si quelqu'un est obèse et à quel point, plusieurs méthodes ont été mises au point, comme la détermination basée sur l'anthropométrie, l'analyse de l'impédance bioélectrique, la sensitométrie et les méthodes basées sur l'imagerie (6). Parmi ces méthodes, et bien qu'elle soit imprécise, celle de l'indice de masse corporelle (IMC) est la plus utilisée (7). L'IMC estime grossièrement l'adiposité et identifie le surpoids et l'obésité en se basant sur le poids d'un individu exprimé en kilogramme (kg) et divisé par sa taille au carré exprimée en mètre (m²). La classification de l'organisation mondiale de la santé (OMS ou WHO) utilisant l'IMC définit le déficit pondéral lorsque l'IMC < 18,5 kg/m², le poids normal lorsque 18,5 < IMC < 24,9 kg/m², le surpoids lorsque 25 < IMC < 30 kg/m², l'obésité lorsque l'IMC ≥ 30 kg/m², l'obésité sévère lorsque l'IMC ≥ 35 kg/m², et l'obésité morbide lorsque l'IMC ≥ 40 kg/m². L'IMC peut être complété en mesurant également le tour de taille pour déterminer la présence d'une obésité viscérale ou sous cutanée. De la même façon qu'il a été montré que la faible présence de graisse au niveau des hanches (ajustée par rapport à la taille ou au tour de taille) serait un facteur protecteur contre certaines pathologies, le ratio tour de taille sur tour de hanche ou tour de taille sur la taille (hauteur) ont également été proposés pour affiner les risques de l'obésité pour la santé. Néanmoins pour des raisons pratiques liées à la facilité de mesurer la taille et le poids, l'IMC reste l'indicateur le plus utilisé à travers le monde (8).

Différentes approches et traitements ont été développés et prescrits : éducation et contrôle diététique, programmes d'activité physique, pharmacothérapie et chirurgie bariatrique. La

chirurgie bariatrique a été associée à une amélioration de l'équilibre métabolique (incluant l'arrêt de la progression ou la récession du diabète de type 2, bien que ce phénomène soit parfois transitoire) indépendant de la perte de poids, même si des effets indésirables ont également été rapportés (9).

Aujourd'hui, 39% de la population mondiale est obèse ou en surpoids, malgré des décennies d'efforts pour enrayer la progression de l'épidémie (**Figures 1 et 2**). Cette prévalence se traduit par un coût en frais de santé équivalent à 2,8% du produit intérieur brut mondial, ou approximativement 2 trillions de dollars (10). Pour stopper l'épidémie, des stratégies de traitement individuel comme la médecine de précision devraient être mises en place en association avec des approches plus générales, comme la prévention. Récemment, plusieurs pays comme le Royaume Uni ont débuté des recherches dans ce domaine. Ces recherches incluent le génotypage de plus d'un million de personnes et la recherche d'associations avec le phénotype. En complément, des stratégies de santé public ont été mises en place, comme la taxation sur les produits trop gras ou trop sucrés, ainsi que des approches nutritionnelles personnalisées (11).

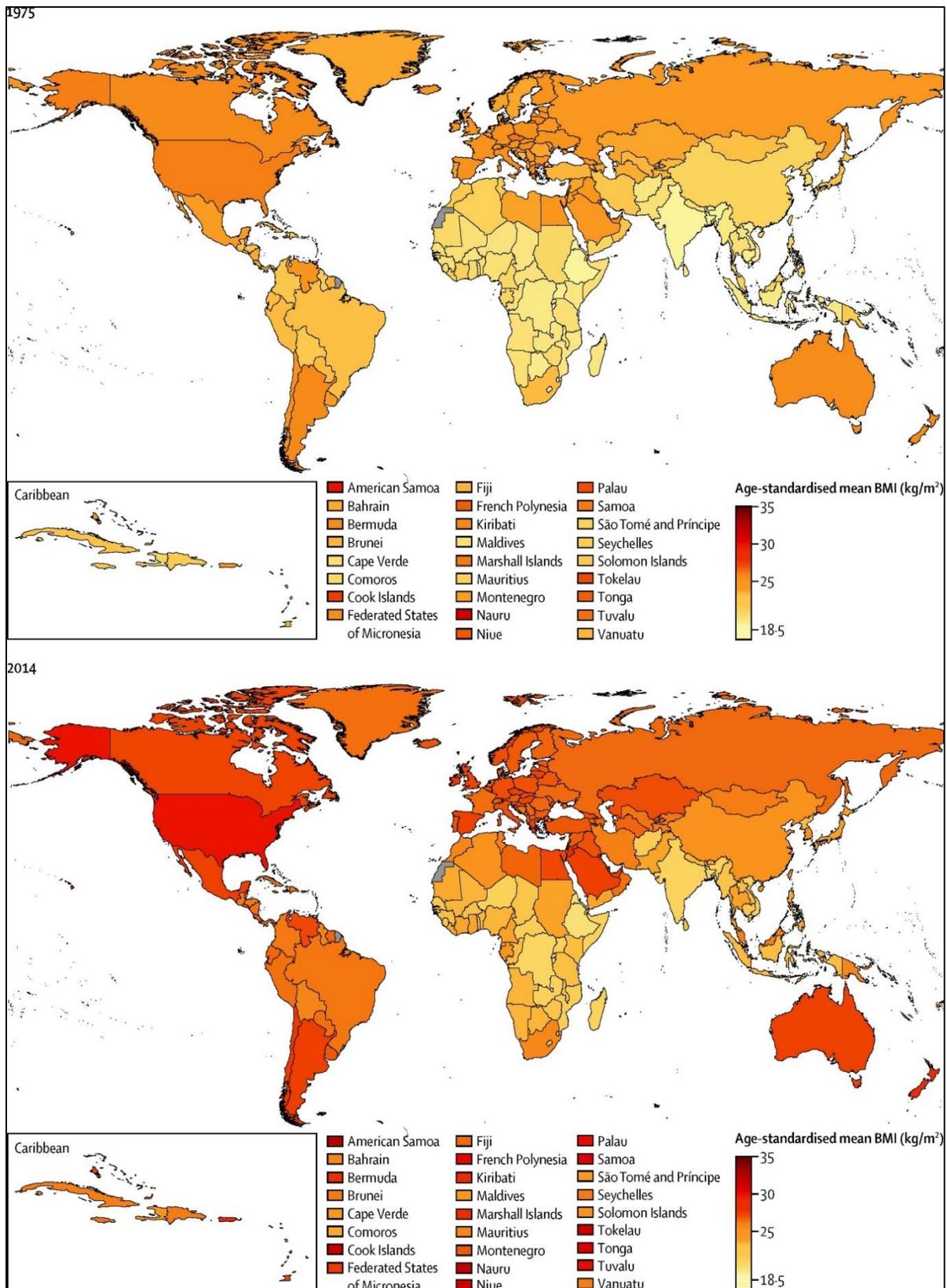


Figure 1 : Variations de l'IMC chez les hommes entre 1975 et 2014 par pays (12).

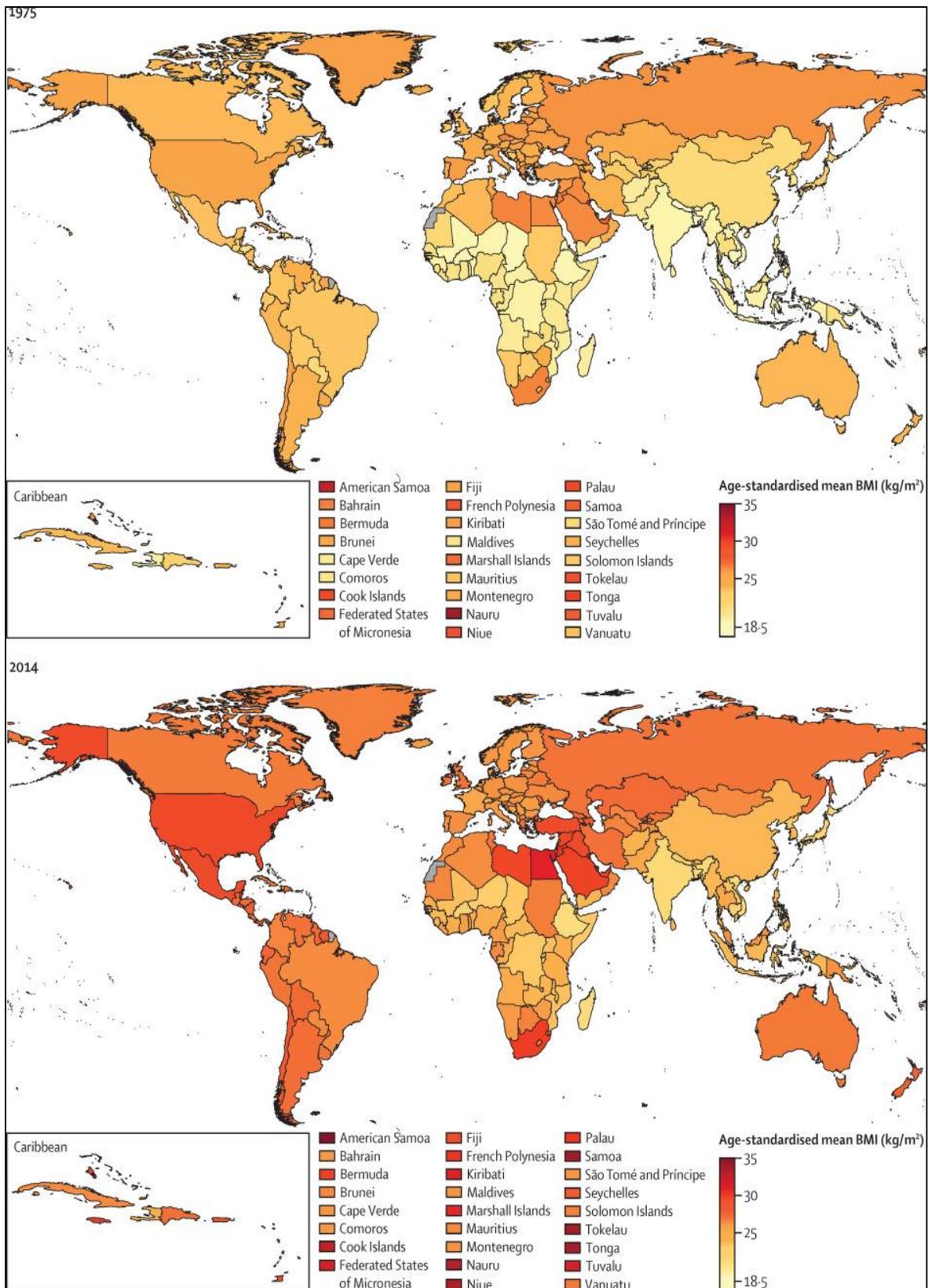


Figure 2 : Variations de l'IMC chez les femmes entre 1975 et 2014 par pays (12).

Epidémiologie :

Prévalence et tendances :

Selon l’OMS, plus de 2,1 milliards d’adultes étaient obèses ou en surpoids en 2014, dont 1,5 milliards de personnes en surpoids et 640 millions de personnes obèses. La prévalence estimée de l’obésité en 2014 était de 10,8% chez les hommes et 14,9% chez les femmes (13). Les données indiqueraient que le sexe féminin est associé à un risque accru de survenue de l’obésité, bien que le surpoids soit davantage présent chez les hommes (12). Néanmoins, *the Global Burden of Disease Study 2013* a rapporté une prévalence similaire pour le surpoids et l’obésité entre hommes et femmes, à plus de 36% (14).

Dans plusieurs études conduites aux Etats-Unis, les afro-américains ont montré une prévalence plus élevée pour l’obésité extrême que les autres ethnies. Les populations asiatiques ont un IMC plus bas que les populations caucasiennes, mais elles sont plus enclines aux dépôts de graisse viscérale, ce qui les rend plus susceptibles au développement du diabète de type 2 avec un IMC faible que les populations caucasiennes (15).

Entre 1980 et 2008, la moyenne globale de l’IMC a augmenté de 0,4 kg/m² chez les hommes et 0,5 kg/m² chez les femmes par décennie. Le pourcentage d’adultes avec un IMC de 25 kg/m² ou plus a augmenté entre 1980 et 2008 de 28,8% à 36,9%, et de 29,8% à 38% chez les hommes et les femmes, respectivement ((12) et **Figure 3**).

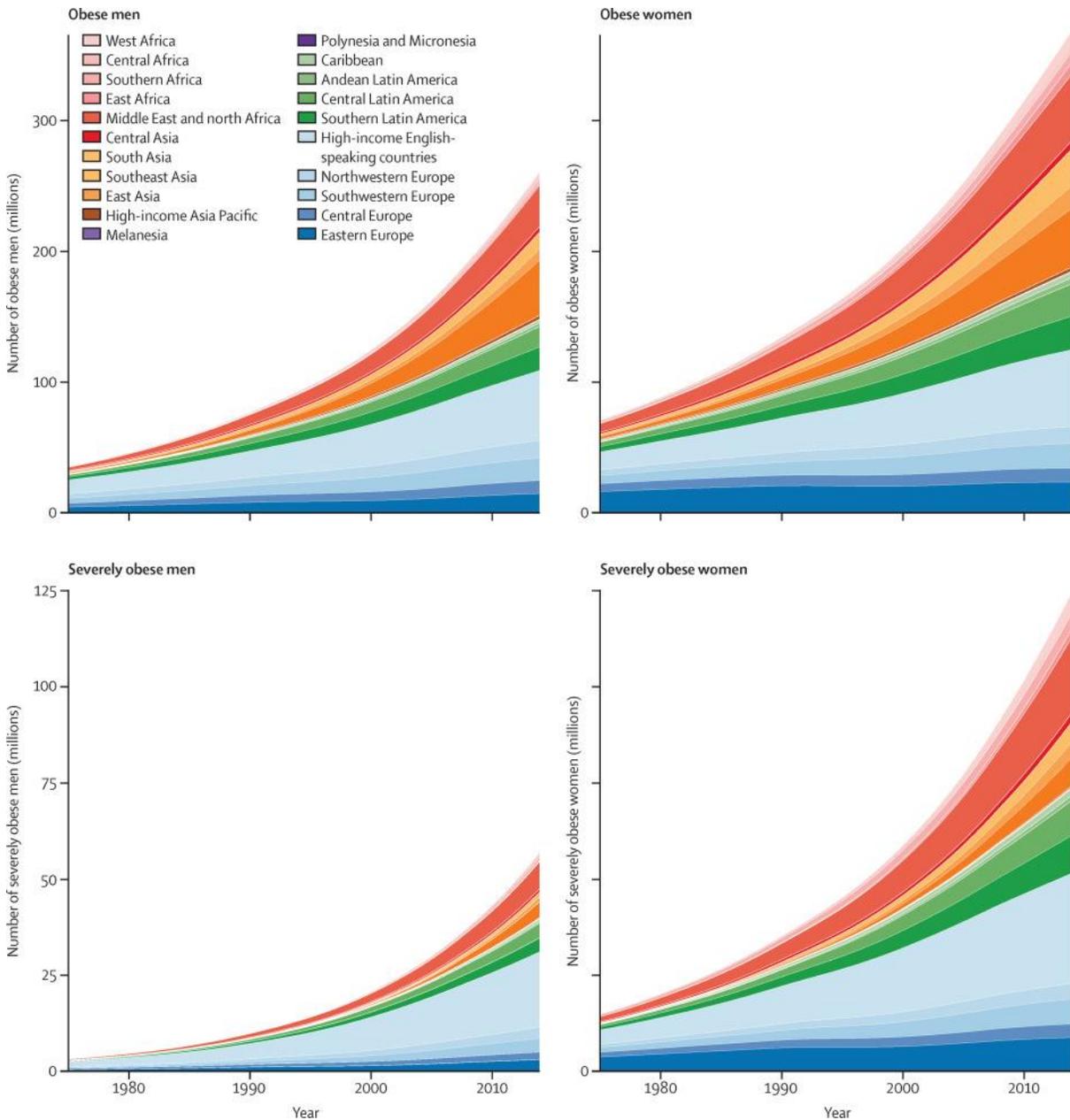


Figure 3 : Progression de l'obésité dans différentes populations entre 1975 et 2014. Les individus obèses ont un IMC > 30 kg/m², et les individus sévèrement obèses ont un IMC > 35 kg/m². (12)

D'ici 2030, les prévisions font état que 57,8 % (soit 3,3 milliards de personnes) de la population adulte aura un IMC de 25 kg/m² ou plus (16). De ce fait, l'impact de l'ensemble des maladies associées à l'obésité augmentera dans les années qui viennent. Dans de nombreux pays et régions, et pas uniquement en Europe ou aux Etats-Unis, le nombre d'adultes obèses ou en surpoids est plus élevé que celui d'adultes de poids normal (12). Les conséquences indésirables

de l'obésité pour la santé publique représentent aujourd'hui une plus grande menace que la faim ou la malnutrition (17).

En 2013-2014, le nombre d'enfants et d'adolescents (2-19 ans) souffrants d'obésité était estimé à 110 millions, soit le double de 1980, bien que la population mondiale a elle aussi augmenté. La prévalence de l'obésité chez ces mêmes enfants était estimée à 5% en 2014 (14). Et les tendances concernant le surpoids et l'obésité pour les enfants entre 11 et 15 ans en Amérique de Nord et certains pays d'Europe entre 2002 et 2010 font état d'une stabilisation de leur prévalence, même si celle-ci reste élevée. L'obésité infantile est associée avec l'apparition de complications métaboliques et de maladies chroniques à l'âge adulte (18).

Facteurs de risques connus :

A première vue, les principaux facteurs responsables de l'épidémie d'obésité sont facilement identifiables : la suralimentation et la sédentarité. Une légère surconsommation alimentaire journalière semble être un contributeur important pour l'apparition de l'obésité, bien que ce soit un phénomène plus complexe (19). Des approches épidémiologiques, comme des analyses bio-informatiques, doivent être utilisées pour élucider les causes de l'apparition de l'obésité, dont l'interconnexion entre le style de vie, l'environnement, la physiologie, la génétique, et les facteurs sociaux et économiques.

Un des aspects les plus complexes dans l'étude de l'obésité concerne les différences fortement marquées entre individus dans l'accumulation de graisse. Cette hétérogénéité se retrouve dans la façon dont le corps stocke les graisses, et toutes les variations dans la façon dont il peut stocker, indépendamment de la quantité à stocker. Une telle hétérogénéité des différents phénotypes obèses rend l'utilisation des précédentes études toutes rassemblées sous le terme unique « obésité » très complexe. Il a d'ailleurs été suggéré d'utiliser le terme « obésités », plutôt que « obésité », étant donné qu'on ne peut considérer tous les cas d'obésité de la même façon (20). Face à ce constat, le développement de l'imagerie et des techniques de physiologie intégrative, dont des éléments permettant de « phénotyper » des individus obèses et de déterminer la composition du corps et l'emplacement des dépôts de graisse (comme un isotope permettant de

marquer la présence de dépôts de graisse brune), a permis de grandes avancées. De ce fait, il pourrait devenir possible d'identifier de larges sous-groupes d'individus obèses possédant des caractéristiques communes et qui pourraient également partager des facteurs étiologiques.

La recherche de gènes à effet majeur pour l'obésité a eu pour résultat l'identification de formes monogéniques de l'obésité. Bien que relativement rare, leurs prévalences sont plus élevées dans les populations consanguines. Les mutations du gène *MC4R* (qui code pour le récepteur 4 de la mélanocortine) comptent pour jusqu'à 5% des cas d'obésité extrêmement précoce (21). Néanmoins, la génétique seule ne peut expliquer la relativement récente et rapide augmentation de la prévalence de l'obésité. Il est devenu clair qu'elle résulte de l'interaction de nombreux gènes de susceptibilité (dont certains ont été identifiés) avec de nombreux facteurs liés au style de vie (22). De plus, le rôle entre autre des approches génétiques et nutritionnelles, comme l'étude des désordres métaboliques, du patrimoine épigénétique, ou les mesures métabolomiques, sont des champs d'études suscitant de plus en plus d'intérêt pour découvrir les origines, les causes et les conséquences de l'obésité, et pour permettre une prise en charge individualisée des patients.

Plusieurs facteurs environnementaux et biologiques ont un effet sur l'absorption et la dépense énergétique. Par exemple, les comportements alimentaires ne sont pas uniquement influencés par des facteurs génétiques ou biologiques, mais également par l'abondance ou non des opportunités de manger, la quantité et qualité de la nourriture disponible et son prix ainsi que l'environnement (qui affecte par exemple l'accès à la nourriture et les dépenses physiques). Dans ce contexte, les firmes agroalimentaires ont adopté de puissantes stratégies de marketing pour encourager la population à manger davantage. Le vaste choix de nourriture à haute teneur énergétique est également vecteur de l'obésité. De plus, le fait de manger hors de chez soi, qui est associé à l'obésité via l'augmentation de la taille des portions consommées ainsi qu'à d'autres facteurs, a progressé lors des dernières décennies (23). L'activité physique, un autre facteur comportemental, est également conditionnée par des facteurs socioéconomiques et culturels (par exemple la façon de se déplacer, ou l'importance de l'utilisation des ordinateurs ou des jeux-vidéos, ainsi que les hobbies). Une association conjointe entre la progression du style de vie sédentaire et l'augmentation de l'apport énergétique peut en grande partie expliquer l'augmentation de l'IMC moyen dans les pays à haut niveau de vie, mais également dans les pays

moyennement et peu développés (24). Le nombre d'heures de sommeil a diminué, augmentant la fatigue durant les heures d'activité et la surconsommation de nourriture, prédisposant de ce fait à la sédentarité et à l'apparition de l'obésité (25). De façon intéressante, la « sieste méditerranéenne » a été inversement associée avec le risque d'apparition de l'obésité, si la durée de celle-ci est inférieure à 30 minutes (25). De plus, le travail de nuit est associé avec le manque de sommeil, dont une des conséquences est l'augmentation de la surcharge pondérale (5). Malgré cela, la question visant à savoir si l'obésité est majoritairement liée à une consommation excessive de calories ou à une dépense énergétique insuffisante fait toujours débat, les différents éléments de la balance énergétique étant vraisemblablement importants.

En plus de la quantité de calories, le type et la qualité de celles-ci (par exemple, des graisses saturées comparées aux graisses insaturées, tout comme les sources de graisses, de carbohydrates et de protéines) influencent également la balance énergétique et le poids à long terme. Néanmoins, cette aire de recherche est toujours sujette à controverse. L'influence des différents nutriments sur le poids, et l'impact de nutriments en particulier (par exemple l'effet thermogénique des protéines) est important sur l'obésité en favorisant la surconsommation d'énergie ou par effet métabolique direct encore mal connu (26). Un régime alimentaire sain (défini par une consommation équilibrée de nutriments, avec entre 10 et 20% de l'énergie provenant des protéines, moins de 30% des lipides, et entre 50 et 55% des carbohydrates) est associé de façon inverse avec la prise de poids et le risque de survenue de l'obésité. Cette découverte est également vraie pour une consommation élevée de fruits et noix, de légumes, de grains complets, de yaourts, et pour le passage au régime alimentaire méditerranéen (27,28). A l'inverse, une surconsommation de boissons sucrées, de chips, de frites, de viandes rouges et transformées, d'aliments préparés de façon industrielle, de graisses trans, de grains transformés et d'aliments avec sucres ajoutés a été associée avec une prise de poids (29). Plusieurs études longitudinales supportent le fait de consommer des repas composés de nourriture saine et de carbohydrates complexes avec un indice glycémique bas (une valeur fixée par la vitesse à laquelle un aliment donné fait augmenter la glycémie) et insistent sur l'importance d'un comportement basé sur la consommation raisonnée de nourriture de bonne qualité pour prévenir l'apparition de l'obésité (30).

Le rôle de l’empreinte métabolique :

Plusieurs périodes critiques concernant l’apparition de l’obésité existent, avant la naissance, lors de l’enfance et de l’adolescence. L’empreinte métabolique, définie comme la programmation du métabolisme durant les périodes prénatales et néonatales aux niveaux génétique et épigénétique (31), pourrait affecter de façon permanente le risque de développer une pathologie. Des données provenant d’études longitudinales sur des cohortes dès la naissance fournissent des estimations sur le développement de l’obésité lors de l’enfance et l’adolescence, ainsi que concernant les effets du génotype et de l’environnement sur l’IMC de l’enfance à l’âge adulte.

Lors de la période prénatale, un gain excessif de poids par la mère, en particulier lors du début de la grossesse (les première 20 semaines), est un facteur de risque pour le développement du surpoids chez les enfants. Le poids à la naissance est un bon marqueur de la nutrition fœtale. Une surnutrition ou une sous nutrition durant la vie fœtale peuvent déclencher des voies métaboliques qui seront plus tard responsables de l’apparition de l’obésité (32). Un poids à la naissance élevé a été associé à un plus haut risque de survenue de l’obésité, alors qu’un faible poids à la naissance a été associé à une forte proportion de masse grasseuse, indépendante de l’IMC, et à une obésité abdominale chez les adolescents (33).

Chez les enfants, une augmentation rapide du poids a également été associée à un haut risque de surpoids plus tard dans la vie. Le type et la quantité de nourriture calorique consommée par la mère durant la grossesse et l’allaitement, et même par les deux parents durant la grossesse, ont été associés à l’apparition de complications métaboliques à l’âge adulte (34). Un allaitement prolongé a été associé à une diminution du risque d’apparition du surpoids au cours de la vie (35). Une réduction du contenu protéique dans les préparations pour enfants en bas âge est associée, dans un essai clinique randomisé, à un IMC moins élevé et une diminution du risque de survenue de l’obésité lors de l’enfance (36).

Durant l’enfance, un rebond adipocytaire (**Figure 4**) précoce est associé à un IMC élevé, à une grande quantité de tissu adipeux sous-cutané ainsi qu’à un tour de taille élevé à l’âge adulte (37). L’adolescence est une étape du développement sensible au niveau nutritionnel car le rythme de croissance s’accélère, avec pour conséquence une augmentation des besoins énergétiques et

nutritionnels. L'étape de survenue de la puberté influence l'IMC et est associée avec l'IMC de l'enfance à l'âge adulte, où une puberté précoce augmente le risque de survenue de l'obésité plus tard dans la vie (38).

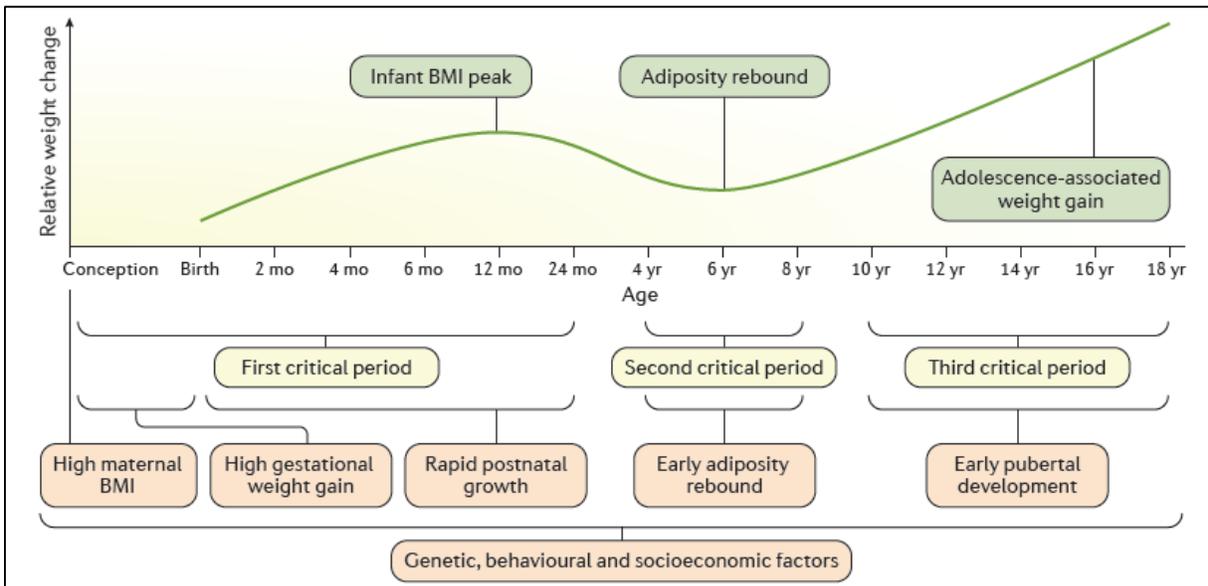


Figure 4 : Périodes critiques de la vie pour le développement de l'obésité (8). La première période dure de la conception à la fin de la deuxième année de vie, où l'enfant atteint un pic de son IMC aux alentours du 7ème mois de vie. Puis l'IMC atteint un minimum aux alentours de 5 à 7 ans, avant de réaugmenter, ce qui correspond au rebond adipocytaire (39).

Comorbidités et mortalité :

Plusieurs études portant sur des populations ont montré que des individus obèses sont plus à risque de développer un grand nombre de complications de santé qui contribuent à leur mort prématurée par rapport à des individus de poids normal (13).

Le syndrome métabolique est une des associations de facteurs de risque liée à l'obésité les plus fréquentes. Selon la définition et les critères harmonisés fournis par la Fédération Internationale du Diabète, L'Association Américaine du Cœur et le *US National Heart, Lung and Blood Institute*, le syndrome métabolique est diagnostiqué lorsque 3 des 5 critères suivants sont retrouvés (2):

-Obésité viscérale : Un tour de taille de plus de 94cm chez les hommes et 80cm chez les femmes pour les pays occidentaux.

-Hypertriglycéridémie : plus de 150mg par dl ou patient sous traitement visant à diminuer les triglycérides.

-Faible taux de lipoprotéine de haute densité : moins de 40 mg / dL chez les hommes et moins de 50 mg / dL chez les femmes.

-Pression sanguine élevée : pression sanguine systolique supérieure à 130mmHg, pression diastolique supérieure à 85mmHg, ou patient sous traitement anti-hypertensif avec historique d'hypertension.

-Glycémie élevée : glycémie à jeun supérieure à 100mg / dL ou patient sous traitement visant à diminuer la glycémie.

D'autres comorbidités communes associées au syndrome métabolique sont les désordres endocriniens (comme le diabète de type 2), des problèmes respiratoires (comme l'apnée du sommeil), des maladies cardiovasculaires (comme l'athérosclérose ou les crises cardiaques) et des cancers (comme les cancers de l'endomètre ou du foie) (40). De plus, l'obésité a des effets négatifs sur l'humeur ou peut engendrer des soucis psychologiques, mais également agir négativement sur les fonctions cognitives (41).

Les études épidémiologiques sur les comorbidités et les complications associées à l'obésité sont limitées pour plusieurs raisons. Une des difficultés rencontrées lors de larges études de populations est de définir le groupe de « poids normal » de référence. Bien que des bonnes pratiques aient suggéré qu'un IMC compris entre 18,5 et 25 kg/m² était compatible avec la bonne santé, une étude récente ayant rassemblé les 4 millions d'individus provenant de 189 études se déroulant sur en moyenne 13,7 ans a montré que l'IMC associé à la plus faible mortalité était compris entre 20 et 25 kg/m², et a également montré que les valeurs d'IMC plus faible étaient associées à une mortalité accrue (42). Ce rapport a mis en évidence la controverse visant les études précédentes voulant que des individus en surpoids ou faiblement obèses (entre 30 et 34,9 kg/m²) ont une mortalité plus faible que des individus non obèses. Néanmoins, ces études ont inclus des individus minces (IMC entre 18,5 et 20 kg/m²) dans la catégorie des individus de « poids normal ». Ces individus de faible poids sont probablement responsables de la mortalité plus élevée du groupe de « poids normal », étant donné que certains étaient plus susceptibles d'être des fumeurs ou d'être atteints d'une maladie chronique responsable de leur poids plus faible (43).

Quand le *Global BMI Mortality Collaboration international group* a exclus les fumeurs et les individus atteints d'une maladie chronique, il est devenu clair que le surpoids et la faible obésité sont associés à une mortalité accrue par rapport aux individus de « poids normal ».

Certaines études ont suggéré que l'IMC maximum atteint durant la vie est plus important que l'IMC lors de l'étude de son influence sur la mortalité. Lors d'une méta-analyse utilisant les données de 239 cohortes, le surpoids et l'obésité étaient significativement associés à la mortalité totale, ainsi qu'à celle liée aux maladies cardiovasculaires ou aux cancers, par rapport au groupe des individus de « poids normal », après avoir tenu compte de tous les paramètres cités précédemment, comme le fait de fumer (42). Ces associations sont valables dans les populations asiatiques, australiennes, néo-zélandaises, européennes et d'Amérique du nord. En plus de la mortalité, l'obésité a également été associée à l'insulino-résistance, l'hypertension, et l'hypercholestérolémie (44).

Mécanismes / Physiopathologie :

Le poids du corps, et par extension l'obésité, est déterminé par l'interaction de la génétique, de l'environnement et de facteurs psychosociaux agissant via plusieurs médiateurs physiologiques de la prise alimentaire et de la dépense énergétique qui affectent l'accumulation de graisse.

Le tissu adipeux :

Types :

Il existe deux grands types de tissus adipeux : les tissus adipeux blanc et brun. Ils ont tous deux un rôle crucial dans la sensibilité et la réponse aux changements concernant la balance énergétique du corps. Les adipocytes du tissu adipeux brun contribuent à la dépense énergétique via le phénomène de thermogénèse et de maintien de la température du corps. Le tissu adipeux brun est abondant chez les nouveau-nés et est principalement localisé au niveau des régions inter-scapulaires et supra-claviculaires, mais également autour des reins, du cœur, de l'aorte, du pancréas et de la trachée. Ces dépôts adipeux bruns diminuent avec l'âge, mais

peuvent toujours être retrouvés chez l'adulte. L'abondance de tissu adipeux brun a été négativement associée à l'IMC, bien que son rôle précis dans la survenue de l'obésité soit toujours assez obscur (45). Les adipocytes blancs retrouvés dans le tissu adipeux blanc sont les plus abondants chez l'humain et ont longtemps été considérés comme jouant uniquement un rôle de stockage énergétique. Néanmoins, des études ont montré que ces cellules sécrètent plusieurs substances bioactives (appelés adipocytokines ou adipokines) (46). En plus, un type mineur mais distinct d'adipocytes, appelés adipocytes « beiges /brite » est retrouvé dans des dépôts d'adipocytes blanc en tant que nouveau type d'adipocytes dissipant de l'énergie.

Les différents tissus adipeux sont classés en tissus adipeux sous-cutané ou viscéral, en fonction de leur localisation dans le corps (**Figure 5**). Le tissu adipeux viscéral comprend le tissu adipeux abdominal (principalement du tissu adipeux omental), le tissu adipeux périrénal et péricardial. Le type le plus étudié est le tissu intra abdominal, étant donné que son accumulation est associée avec le développement des symptômes correspondant au syndrome métabolique.

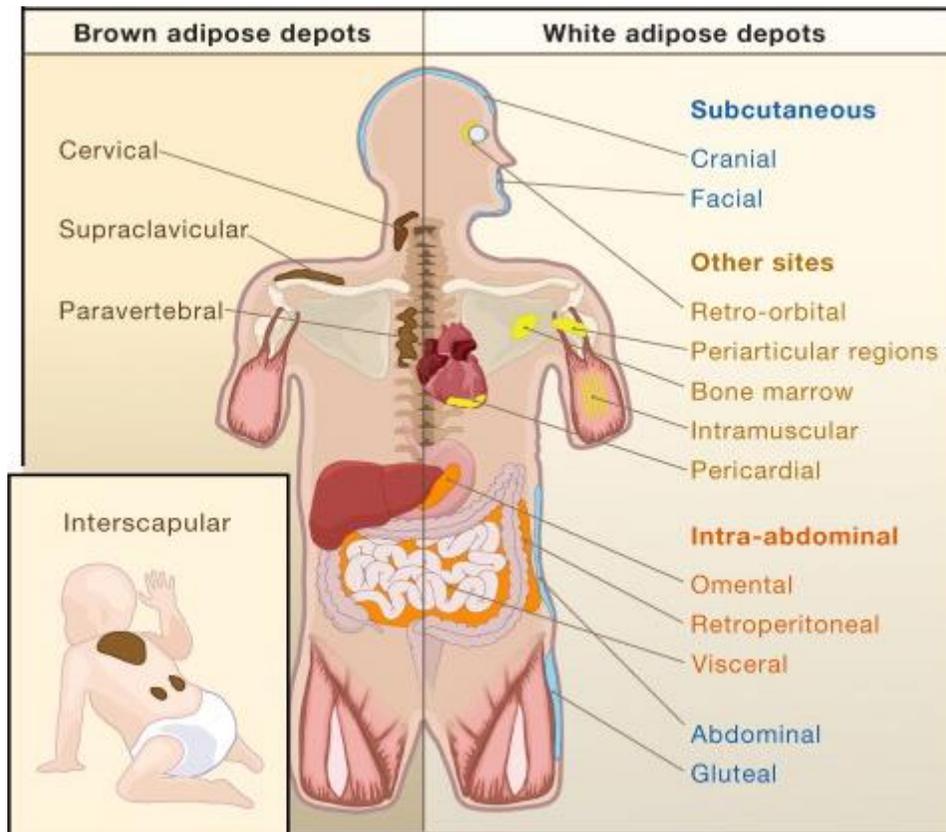


Figure 5 : Distribution des types de tissu adipeux (47).

Changements pathologiques dans le tissu adipeux :

Bien que la liste des conséquences cliniques associées avec l'obésité soit plutôt longue, le nombre et l'importance de ces complications de santé semblent être dépendants de la présence additionnelle d'excès de graisse au niveau interne (comme les dépôts de graisse intra-abdominale) ainsi qu'autour de tissus normalement exempts de graisse comme les reins, le cœur ou le foie. Ce phénomène de dépôts au niveau des organes se nomme dépôt de graisse ectopique (48). Il semble que la façon dont le tissu adipeux gère le surplus énergétique et les tensions dues à une croissance pathologique de ces adipocytes ait un lien très fort avec le risque cardiométabolique d'un individu, sa sensibilité à l'insuline et sa dyslipidémie (qui correspond à un haut niveau de triglycérides ou un faible niveau de lipoprotéine de cholestérol à haute densité). Par exemple, si le tissu adipeux sous cutané gère cet excès énergétique en provoquant une hyperplasie des adipocytes le composant (c'est-à-dire une croissance du tissu par la multiplication des cellules le composant), ce tissu adipeux grandissant normalement agira comme un « puit métabolique » protégeant les tissus non gras (comme le cœur, le foie, les reins ou le pancréas) contre les effets néfastes des dépôts de graisse ectopique. Par contre, si ce même tissu adipeux sous cutané ne parvient pas à se développer via le phénomène d'hyperplasie (et ce pour des raisons encore mal connues), les triglycérides stockées contribueront à l'hypertrophie des adipocytes (c'est-à-dire une croissance du tissu via le grossissement des cellules le composant), jusqu'à ce que ces adipocytes saturent et ne soient plus capable de grossir, menant soit à leur rupture et à un recrutement massif de macrophages, et / ou à une augmentation de la sécrétion par ces adipocytes d'adipokines pro-inflammatoires et à une diminution de la sécrétion d'adipokines anti-inflammatoires, comme l'adiponectine, par ces adipocytes hypertrophiés (49). Ces phénomènes contribuent à l'inflammation et à l'insulinorésistance du tissu concerné. De plus, étant donné que la capacité à grossir de ces adipocytes atteint ses limites, l'excès de triglycérides n'aura pas de place pour être stocké dans le tissu adipeux sous cutané et finira par être stocké dans des lieux où ils auront un effet néfaste, comme le foie, le cœur, les reins ou le pancréas (**Figure 6**). Ces dépôts de graisse ectopiques sont également générateurs d'athérosclérose, favorisant le diabète et l'environnement pro-inflammatoire des cellules (50).

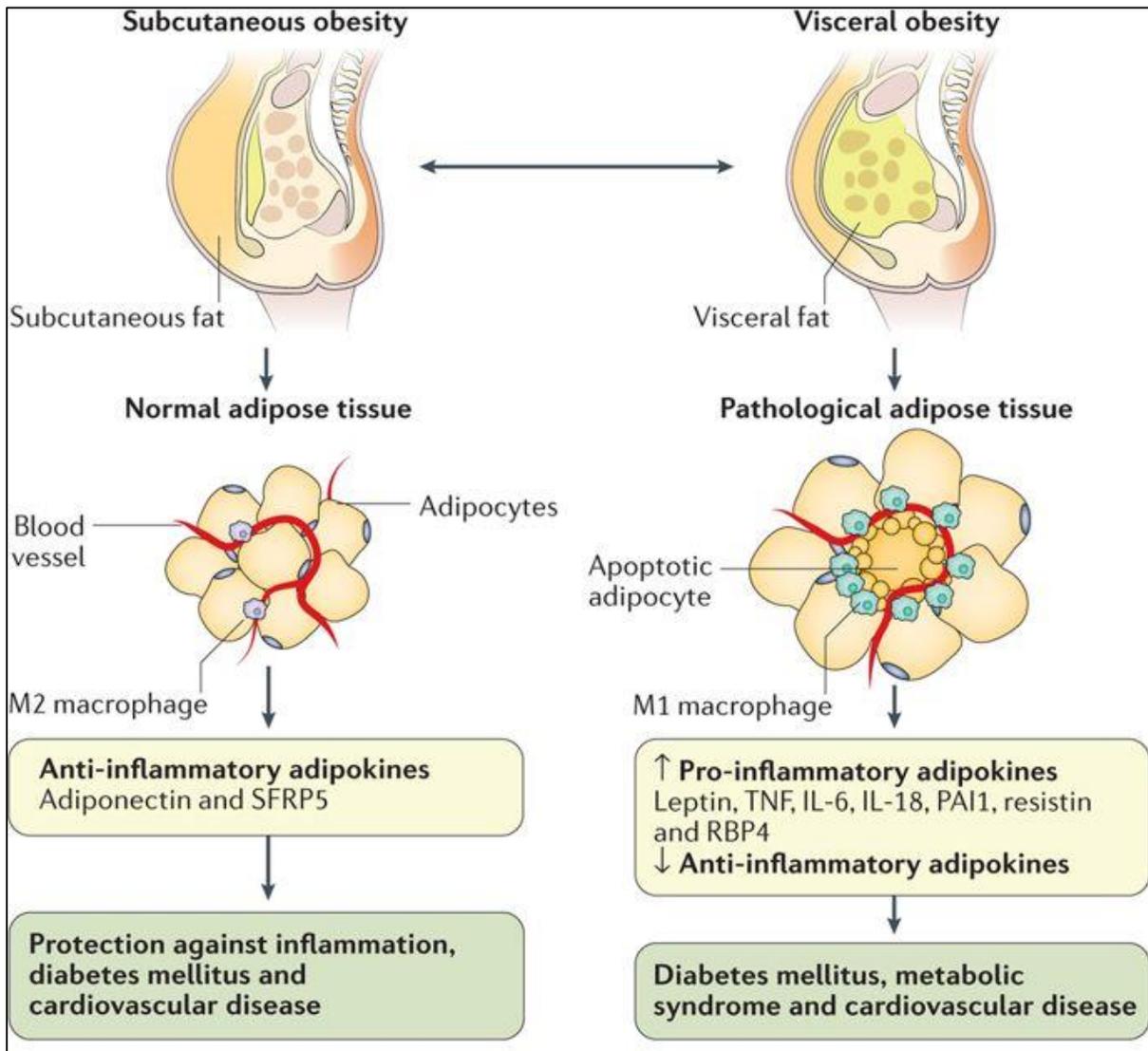


Figure 6 : Différents effets induits par l'obésité viscérale et l'obésité sous-cutanée (8).

L'obésité viscérale est souvent accompagnée d'un excès de graisse dans le foie, le cœur et les reins, alors que cette association est moins évidente dans le cas de l'obésité sous cutané (51). Bien que l'excès de tissu adipeux viscéral et l'excès de graisse au niveau du foie sont souvent associés, plusieurs individus présentant un excès de tissu adipeux viscéral n'ont pas de dépôts de graisse au niveau du foie, et inversement. Il est évident que l'identification des raisons (génétique, hormonale, environnementale) conduisant aux différents dépôts de graisses ectopiques régionaux est une problématique qui se doit d'être approfondie.

De plus, tous les dépôts de graisse sous-cutanés ne sont pas égaux et les données actuellement disponibles indiquent que le « puits métabolique » le plus sain pour la santé serait le dépôt de graisse glutéo-femorale (49). Néanmoins, les dépôts de graisse sous cutanée au niveau du bas du corps pourraient ne pas être neutres concernant leurs effets sur la santé. Il se pourrait même qu'ils soient protecteur contre plusieurs risques comme les risques de développer un diabète de type 2 ou une maladie cardio-vasculaire (52). Au vue de ces données, l'état sévère de perturbation du système métabolique dont souffrent les patients atteints de lipodystrophie (c-à-d manquant de tissu adipeux sous-cutané) montre l'importance du tissu adipeux sous cutané sain (53).

Régulation de la balance énergétique :

La consommation de nourriture stimule le déclenchement de signaux gastro-intestinaux véhiculés par extension mécanique ou par des hormones paracrines, et par les nutriments eux-mêmes qui modulent l'appétit, via plusieurs neurotransmetteurs, des peptides provenant de l'axe intestins-cerveau, des acides aminés et des neuropeptides. Le système nerveux autonome et plusieurs hormones circulantes ont également été impliqués dans la réponse métabolique à l'absorption de nourriture et au métabolisme des nutriments, qui affecte l'appétit, la thermogénèse et le dépôt de graisses, entre autres processus (54) (**Figure 7**).

Le rôle du microbiote est également très étudié, et il a été découvert que le microbiote d'individus obèses est moins diversifié et a une proportion différente de *firmicutes* et de *bactéroidetes* par rapport à celui d'une personne de poids normal (55). Et bien que davantage d'études soient nécessaires, il semble que ces perturbations dans les proportions du microbiote favorisent l'apparition d'un environnement inflammatoire, perturbent le métabolisme des nutriments, influencent la gestion de la balance énergétique et agissent sur la façon dont l'énergie est stockée et extraite. D'autres facteurs endogènes associés à l'obésité pourraient induire des perturbations dans le contrôle de l'appétit, des soucis lors de l'adipogénèse, lors des nombreux processus inflammatoires et lors des perturbations dans le métabolisme des lipides (5).

Le poids de la génétique et de l'épigénétique :

Jusqu'à 70% des variations entre individus concernant le poids seraient dues à des variations génétiques (56). L'identification de gènes influençant la susceptibilité à l'obésité peut fournir des pistes sérieuses concernant les mécanismes physiopathologiques qui déterminent la régulation du poids et le stockage des graisses, ce qui pourrait à terme permettre de déterminer de nouvelles approches de prévention et de traitement.

Les recherches sur l'obésité monogénique ont mis en évidence la biologie de l'obésité dans la population générale. La détermination de l'emplacement des mutations causant l'obésité monogénique chez la souris a été une des premières stratégies dans la recherche de gènes influençant le poids. Les résultats les plus connus concernant cette approche incluent la découverte des gènes codant la leptine (*LEP*), et son récepteur (*LEPR*), le récepteur 4 à la mélanocortine (*MC4R*), ou encore le gène de la pro-opiomélanocortine (*POMC*) (**Figure 7**). Ces gènes jouent un rôle sur le poids via plusieurs voies affectant le système nerveux central (57). Des mutations dans les orthologues humains de ces gènes sont à l'origine de l'obésité monogénique (**Figure 8 et Table 1**) (58).

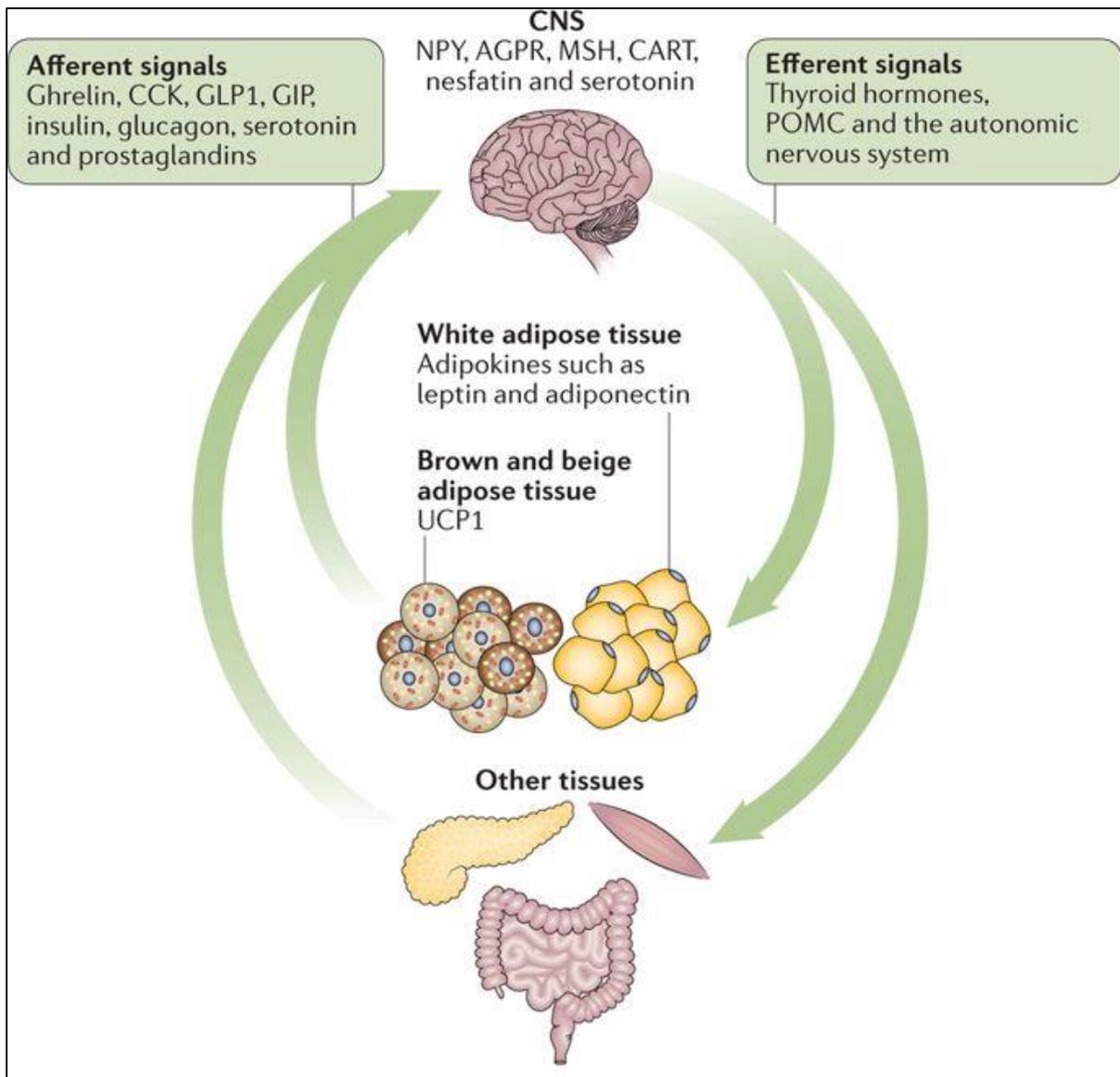


Figure 7 : Interactions intervenant dans le contrôle de de la faim et de la satiété (8).

De larges études d'association pangénomique (GWAS) qui testent l'association entre des millions de variants génétiques communs et la survenue de l'obésité, sans hypothèse préalable, ont identifié plus de 300 loci génétiques influençant le risque d'obésité. Le succès majeur de ces GWAS a été la découverte du locus du gène *FTO* (*Fat mass and obesity-associated protein*). En effet un cluster de variants non codants au niveau de *FTO* a montré une très forte association avec le risque de survenue de l'obésité et la variation de l'IMC (59). Les mécanismes biologiques entourant cette association commencent seulement à être mis en évidence. *FTO* est exprimé dans

l'hypothalamus, et son expression est régulée par des facteurs liés à la nutrition (60). Les souris *Fto* KO sont plus petites avec une augmentation de la dépense énergétique (61) alors que les souris surexprimant *Fto* sont obèses (62). FTO a été montré comme étant une dioxygénase ayant la capacité de déméthyliser la 3-méthyl-thymine dans l'ADN *in vitro*, mais des preuves de cette fonction *in vivo* restent à établir (60).

Le locus *FTO* pourrait réguler l'expression du gène *RPGRIP1L* à proximité, ou des gènes *IRX3* et *IRX5* à plus longue distance pour influencer le poids, en régulant l'appétit, la thermogénèse, le phénomène de brunissement adipocytaire et des mécanismes épigénétiques associés à l'obésité ou à l'IMC (63). D'autres GWAS ont identifié des loci associés aux mécanismes liés aux lipides ou au rapport tour de taille sur tour de hanches. Des analyses concernant les voies métaboliques ou les tissus touchés par le stockage des graisses ont été effectuées concernant les loci associés à l'IMC, et ont montré un rôle fondamental du système nerveux central dans la régulation du poids. Les tissus possédant une expression enrichie pour ces gènes comprennent l'hypothalamus, la glande pituitaire et l'hippocampe. Cette expression enrichie concerne des voies liées à la fonction synaptique et aux neurotransmetteurs (64). Les mêmes analyses appliquées aux loci associés avec le rapport tour de taille sur tour de hanches ont mis en évidence des voies impliquées dans l'adipogénèse, l'angiogénèse, l'insulino-résistance et les mécanismes affectant la prise en charge des lipides absorbés (65).

Malgré ces associations significatives, les effets prouvés de ces loci sur l'obésité sont peu nombreux. De tous les loci à effet établi sur l'obésité, *FTO* à l'effet le plus important sur l'obésité, bien que celui-ci soit relativement faible : chaque allèle à risque augmente le risque de survenue de l'obésité de 1,2 à 1,32 fois, et l'IMC de 0,37kg/m² (ce qui correspond à 1,06 kg pour une personne de 1,70m). L'effet des autres variants associés à l'IMC est compris entre 0,08 et 0,3 kg/m² (65). L'ensemble des variants associés à l'IMC combinés explique moins de 5% des variations de l'IMC. Les effets de ces mêmes variants sur le rapport tour de hanche sur tour de taille sont de même amplitude. Néanmoins, des interactions entre des facteurs génétiques et environnementaux pourraient expliquer d'importantes différences individuelles de poids pour un même environnement

Epigénétique :

Les processus épigénétiques, incluant la méthylation de l'ADN, les modifications d'histones et les ARNs non codants qui allument ou éteignent l'expression des gènes sans changement au niveau de leur séquence d'ADN, sont sensibles à des facteurs extérieurs (par exemple, le régime alimentaire et l'activité physique), et sont réversibles et peuvent être transmis à la génération suivante. Les processus épigénétiques sont cellulaires, temporels et tissulaires spécifiques, ce qui rend difficile l'étude de leur rôle chez les humains. Ce constat est particulièrement vrai pour l'obésité, étant donné le rôle central du cerveau dans la régulation du poids. Les preuves du rôle de l'épigénétique concernant l'obésité proviennent majoritairement d'études animales, et plus rarement d'études sur des humains, même si plus de 600 articles ont été publiés à ce sujet. Par exemple, les garçons nés de femmes ayant jeûné durant la première moitié de leur grossesse lors de la famine Hollandaise avaient un risque de survenue de l'obésité significativement plus élevé que les sujets auxquels ils ont été comparés, ce qui a plus tard été relié à des modifications épigénétiques (66). Cette étude intergénérationnelle suggère qu'une privation alimentaire de mères enceintes peut avoir des effets non génétiques persistants sur le poids de la génération suivante. Par exemple, bien que la méthylation du gène codant le facteur de transcription inductible par hypoxie 3A (*HIF3A*) soit considérée comme associée à un IMC plus élevé, cette association a été déterminée comme une conséquence d'un IMC élevé, et non une cause de celui-ci (67). Néanmoins, le système HIF a été montré comme jouant un rôle clé dans la dépense énergétique et l'apparition de l'obésité (68). Deux autres études épigénétiques de grande ampleur, comprenant 10 000 et 7 800 individus, ont identifié un grand nombre de sites de méthylation de l'ADN associés à l'IMC : 187 sites méthylés dans la première étude, et 83 sites méthylés dans la seconde (69,70). Ces sites de méthylation ont majoritairement été identifiés en tant que conséquences, et non comme causes de la variation de l'IMC. De manière générale, davantage de preuves sont nécessaires concernant un rôle (causal ou non) de l'épigénétique sur l'obésité, mais les difficultés rencontrées pour étudier l'épigénome au bon moment et dans les bons tissus sont une barrière importante pour le succès des études chez l'humain.

Etudes gènes-candidats :

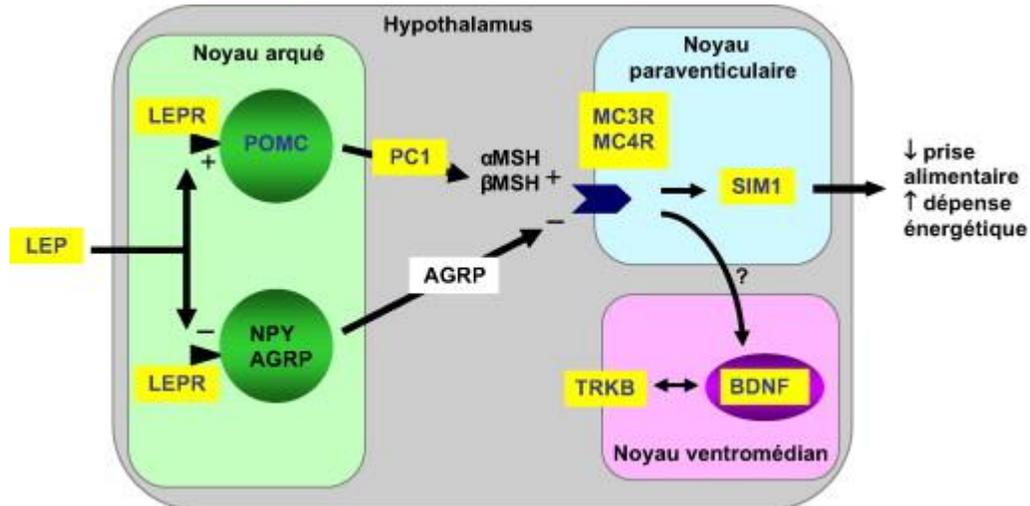


Figure 8 : Schema des gènes impliqués dans la voie mélanocortine (71).

Les études gènes-candidats basées sur le séquençage de gènes connus pour causer une obésité sévère dans des modèles animaux ont montré que ces mêmes gènes contribuaient également à l'obésité infantile monogénique chez l'Homme. Des expériences de parabiose entre lignées de souris souffrant sévèrement d'obésité, comme les souris *ob/ob* ou *db/db* ont suggéré l'existence d'un facteur circulant intervenant dans la régulation du poids (72). L'identification de cette hormone, la leptine, a ouvert la voie à l'étude des circuits moléculaires et physiologiques contrôlant la régulation énergétique. La leptine est une hormone circulante de 16 kDa dont les niveaux corrèlent fortement avec la masse grasse (73). Un jeûne prolongé provoque une chute des niveaux de leptine, ce qui provoque de nombreux changements dans l'absorption et la dépense énergétique, ou encore la fonction neuroendocrine de façon à maintenir l'homéostasie énergétique (74). La plupart des effets physiologiques de la leptine passent par le système nerveux central, et en particulier par l'hypothalamus, qui est le lieu de la plus forte expression de l'isoforme longue du récepteur de la leptine (75). Plusieurs études ont montré que la leptine stimule l'expression de la pro-opiomélanocortine (POMC) dans les neurones primaires localisés dans le noyau arqué de l'hypothalamus. POMC est largement modifié d'un point de vue post-transcriptionnel pour générer les peptides de mélanocortine, qui activent les récepteurs de la mélanocortine (notamment MC3R et MC4R dans l'hypothalamus) pour moduler plusieurs fonctions du système nerveux central, de la glande adrénaline et de la peau (76). *L'alpha*

mélanocyte stimulating hormone (α -MSH) lors de sa fixation au récepteur de la mélanocortine contribue à la suppression de la prise alimentaire. De plus, la leptine inhibe les voies orexigéniques, qui sont médiées par des neurones exprimant l'antagoniste de la mélanocortine, l'Agouti-related protein, ainsi que le neuropeptide Y (NPY). NPY supprime l'expression de *POMC*. Ces deux groupes de neurones récepteurs primaires de la leptine ont des projections vers un second ordre de neurones exprimant le récepteur de la mélanocortine 4 (MC4R). Une extinction ciblée de *MC4R* chez la souris provoque un accroissement de la prise alimentaire et augmente la quantité de masse maigre (77). Ces voies hypothalamiques interagissent avec d'autres centres nerveux pour réguler l'appétit et moduler les signaux vers la périphérie ce qui régule le métabolisme et la dépense énergétique. L'obésité humaine peut donc résulter d'une série de défauts dans la voie impliquant la leptine et la mélanocortine.

Mutation des gènes de la leptine et du récepteur de la leptine :

Des frameshifts homozygotes, mutations nonsense ou faux sens du gène de la leptine (*LEP*) ont été identifiées et résultent en une incapacité à produire de la leptine. Les patients avec ces mutations sont sévèrement obèses dès leur jeune âge, et leurs niveaux de leptine sont sous les niveaux de détection (78). Ce défaut congénital en leptine peut être traité par des injections quotidiennes de leptine humaine recombinante, corrigeant toutes les anomalies métaboliques chez ces patients. Jusqu'à 3% des patients sévèrement obèses sont porteurs d'une mutation perte de fonction touchant le récepteur de la leptine (*LEPR*) (79). Les phénotypes associés avec une déficience en leptine ou en son récepteur sont très semblables et montrent que chez les humains, la leptine est un régulateur central de la balance énergétique et du comportement alimentaire, de la fonction neuroendocrine et de l'immunité. Ces patients peuvent maintenant bénéficier d'un traitement par setmelanotide par voie injectable, un agoniste de MC4R (80).

Régulation de la prise alimentaire :

Les patients déficients en leptine ou en son récepteur naissent avec un poids normal mais prennent rapidement du poids lors des premiers mois de vie, ce qui entraîne une obésité sévère.

Les éléments marquants de ces déficiences sont l'hyperphagie, un comportement agressif en cas de privation alimentaire, ainsi qu'une satiété non fonctionnelle, qui entraîne la recherche de nourriture très rapidement après la fin d'un repas (81). L'effet principal de la prise de leptine se retrouve sur la prise alimentaire, avec une normalisation de l'hyperphagie et une satiété fonctionnelle. La leptine intervient également dans le mécanisme de récompense alimentaire. L'imagerie IRM a montré des changements impactant la circulation sanguine, qui traduisent des changements dans l'activation de neurones, et ce en réponse à des images représentant de la nourriture comparé à des images n'en représentant pas (82). Lors de l'état de déficience en leptine, la présentation de ces images représentant de la nourriture était associée avec une augmentation marquée de l'activation neuronale au niveau du striatum ventral, une aire associée au plaisir et à la récompense. Ces données semblent indiquer que le striatum ventral n'est pas directement impliqué dans le fait d'aimer la nourriture ingérée, mais dans le fait de vouloir ingérer cette nourriture (83). Chez les rongeurs, les récepteurs à la leptine sont exprimés au niveau du système mésolimbique, connu pour gérer la récompense associée aux différentes substances. L'administration de leptine diminue le seuil d'activation des neurones dopaminergiques associés à ce système, et le KO des récepteurs à la leptine au niveau de ces neurones augmente la prise alimentaire (84).

Régulation de la dépense énergétique et de l'oxydation des graisses :

Chez les individus déficients en leptine, les niveaux du métabolisme basal et de la dépense énergétique sont adaptés à la composition du corps de ces patients. Il a été montré que l'administration de leptine ne diminue pas le niveau du métabolisme basal suite à la perte de poids. Et étant donné que la perte de poids par d'autres moyens est associée avec une diminution de ce métabolisme basal, cette absence d'effet sur le métabolisme basal est notable (81). Le phénomène d'adaptation métabolique à la perte de poids est bien connu, et correspond à une diminution du métabolisme basal plus importante par rapport à celle que la perte de poids seule aurait normalement dû provoquer. D'autres études ont par la suite confirmé le fait que la leptine agit sur la dépense énergétique chez les humains, étant donné que la diminution de la dépense

énergétique basale observée après une perte de poids de 10% est atténuée par l'administration de leptine (85).

Une déficience en leptine est caractérisée par une accumulation préférentielle de masse grasse sous cutanée, et la perte de poids résulte préférentiellement en une perte de masse grasse (81). Chez les souris, la leptine stimule l'oxydation des acides gras dans les muscles via la stimulation de l'AMP kinase (86). Et chez les humains déficients en leptine, une mauvaise oxydation des graisses a été retrouvée (87). Des anomalies de la fonction nerveuse sympathique ont été retrouvées chez des adultes déficients en production de leptine avec en particulier des anomalies dans les fonctions efférentes liées à la thermogénèse (88). Des modifications de la fonction thyroïdienne, liée elle aussi à la dépense énergétique, ont également été observées chez des patients déficients en leptine ou pour son récepteur (89). Des données obtenues sur des rongeurs montrent que la leptine est nécessaire pour la biosynthèse et la sécrétion de la thyrotropine-releasing hormone. Une déficience complète en leptine est associée avec un niveau modéré d'hypothyroïdisme hypothalamique caractérisé par un faible niveau de thyroxine circulante et un haut niveau de thyroïde-stimulating hormone, qui elle est inactive. Des effets similaires ont été observés chez des humains ayant perdu du poids et ont été corrigés par administration de leptine. Cela est également valable dans certains cas chez des patients déficients en récepteur à la leptine (90,91).

La leptine en tant que déclencheur métabolique de la puberté :

Une déficience en leptine est associée à l'hypogonadisme hypogonadotrope et au développement pubertaire non fonctionnel (92). L'administration de leptine permet la survenue d'un développement pubertaire avec un *timing* relativement normal, ce qui suggère que la leptine est un facteur permissif de la survenue de la puberté chez l'humain (81). Néanmoins la survenue de menstruations retardées mais spontanées chez des patientes déficientes en leptine est possible. Il semblerait que l'excès de masse grasse chez ces patientes provoquerait la sécrétion de suffisamment d'œstrogènes par réaction d'aromatase pour provoquer la survenue de menstruations irrégulières même en l'absence de caractères sexuels secondaires pleinement

développés. La leptine semble agir sur le système reproducteur via plusieurs autres molécules incluant la kisspeptine, qui via le récepteur GPR54 provoque la libération de GH-releasing hormone, ainsi que via des cellules nerveuses incluant le *premmamillary nucleus* (93).

La leptine comme médiateur de la régulation nutritionnelle de la fonction immunitaire :

La leptine stimule la réponse inflammatoire, la prolifération de lymphocyte T et la production de cytokines Th1 durant le jeûne chez des souris normales et chez des souris *ob/ob*, ce qui montre son rôle sur la réponse immunitaire. Il semblerait également que la leptine soit impliquée dans la médiation de la réponse systémique à l'état septique et aux maladies parasitaires (94). Les patients déficients en leptine ont plus fréquemment des infections et des anomalies marquées du nombre et du bon fonctionnement des lymphocytes T, fréquence qui se normalise lors d'un traitement en leptine (81). Ces effets multiples de la leptine sur la réponse immunitaire innée et adaptative suggèrent que la modulation immunitaire par la leptine pourrait avoir un rôle thérapeutique chez les humains.

Mutations perturbant la signalisation via la mélanocortine :

Gène	Symbole	Phénotypes associés
Leptine	<i>LEP</i>	Obésité sévère précoce, hyperphagie, hypogonadisme hypogonadotrope
Récepteur de la Leptine	<i>LEPR</i>	Obésité sévère précoce, hyperphagie, hypogonadisme hypogonadotrope
Pro-opiomélanocortine	<i>POMC</i>	Obésité précoce, hyperphagie, insuffisance surrénale, parfois cheveux roux et peau pale
Proconvertase 1	<i>PCSK1</i>	Hyperphagie, obésité précoce, proinsuline élevée, hypoglycémie post-prandiale, hypogonadisme
récepteur 4 à la mélanocortine	<i>MC4R</i>	Obésité précoce, hyperphagie, augmentation de la masse grasse et maigre
facteur neurotrophique dérivé du cerveau	<i>BDNF</i>	Obésité sévère, hyperphagie
tyrosine kinase receptor tropomyosin-related kinase B	<i>TRKB</i>	Obésité précoce sévère, hyperphagie
single minded 1	<i>SIM1</i>	Obésité, hyperphagie, parfois associée à une déficience intellectuelle (avec des troubles ressemblant au syndrome de Prader-Willi)
Adénylate Cyclase 3	<i>ADCY3</i>	Obésité, hyperphagie, anosmie

Table 1 : Tableau récapitulatif des phénotypes associés aux gènes impliqués dans la voie mélanocortine (95,96).

La leptine module la balance énergétique via une combinaison de voies dépendant ou non de la mélanocortine. Chez les humains, des mutations perte de fonction de *POMC* mènent à l'hyperphagie, une obésité précoce, une déficience en adrenocorticotrophine (ACTH) associée à une défaillance cardiorespiratoire à la naissance par insuffisance surrénale aigue, et une hypo pigmentation de la peau et parfois des cheveux (97). Les porteurs de mutations « perte de fonction » hétérozygotes ont un risque plus élevé que la moyenne de devenir en surpoids ou obèse. Il a été décrit que plusieurs mutations hétérozygotes de *POMC* et pertes de fonction des alpha et beta melanocyte-stimulating hormones (alpha et bêta-MSH) augmentent significativement le risque d'obésité mais ne sont pas toujours associées à l'obésité (98). Cela met en évidence un rôle de la beta MSH dans le contrôle de l'homéostasie énergétique. Via la caractérisation fonctionnelle de ces mutations associées à l'obésité, il a été prouvé que le N terminal de *POMC* joue un rôle dans le tri et la sécrétion des peptides dérivés de *POMC* (99). Ces patients peuvent dorénavant également bénéficier d'un traitement par setmelanotide (100).

La pro-hormone convertase 1 (*PCSK1*) est une enzyme impliquée dans la coupure de *POMC* en ACTH, elle-même ensuite clivée en alpha-MSH par la carboxypeptidase E. Les humains déficients pour *PCSK1* sont sévèrement obèses et sont déficients en glucocorticoïdes, souffrent d'un hypogonadisme hypogonadotrope et d'une hypoglycémie postprandiale, qui survient à la suite d'une transformation non fonctionnelle de la pro-insuline en insuline (101). Des niveaux élevés de pro insuline plasmatique dans le contexte de faibles niveaux en insuline mature sont la base du diagnostic de ce trouble. Un signe clinique majeur touchant ces patients est la très faible absorption de leur intestin grêle lors de la période néonatale, attribué à une maturation non fonctionnelle des propeptides des cellules enteroendocrines et nerveuses qui expriment *PCSK1* dans l'intestin (102).

Des mutations de *MC4R* sont retrouvées chez 5 à 6% des patients souffrant d'obésité sévère précoce, et environ chez une personne sur 1000 dans la population générale du Royaume-Uni, ce qui en fait l'une des maladies monogéniques les plus fréquentes (103). Les mutations de *MC4R* sont héritées de manière codominante, avec une pénétrance et une expression variable chez les porteurs hétérozygotes. La plupart de ces mutations de *MC4R* causant une maladie agissent sur l'expression et l'adressage du récepteur vers la surface cellulaire (103). *MC4R* est maintenant

séquencé de plus en plus souvent lors du diagnostic d'enfants sévèrement obèses. Dans le cas de sujets porteurs d'une mutation homozygote, on retrouve une obésité sévère associée à une hyperphagie dès les premiers mois. Dans le cas d'une mutation hétérozygote, le phénotype est moins sévère, avec une obésité généralement plus tardive voir parfois absente (104,105).

Régulation de la prise énergétique, de la dépense énergétique et de l'oxydation des graisses :

La déficience de MC4R a été caractérisée comme provoquant de l'hyperphagie, une augmentation de la masse maigre, via une relation génotype-phénotype proportionnelle au degré de dysfonction du récepteur (106). Ces caractéristiques sont également retrouvées pour les patients porteurs d'une mutation hétérozygote affectant la production d'alpha et bêta-MSH (107). D'autres études ont également mis en évidence que MC4R joue un rôle dans l'oxydation des graisses la répartition des nutriments. Ces effets sont également observés chez les rongeurs et semblent être régulés par le système nerveux sympathique. Certaines des origines du signal sympathique vers le tissu adipeux blanc ont été identifiées au niveau de zones hypothalamiques exprimant MC4R, en particulier le noyau para-ventriculaire. De plus, l'administration intracérébroventriculaire d'agonistes de la mélanocortine a montré une augmentation de l'activité nerveuse sympathique dans le tissu adipeux blanc et une modulation du métabolisme du triacylglycérol favorisant la mobilisation des lipides (108). Un blocage central du récepteur à la mélanocortine par voie pharmacologique mène à une prise de poids moins importante chez des souris triplement déficientes pour la voie adrénérique (récepteurs beta-adrénériques bêta 1, 2 et 3) que chez des souris contrôles, ce qui suggère une diminution du tonus sympathique vers le tissu adipeux blanc, qui favorise l'accumulation de graisses, et pourrait donc être un mécanisme important par lequel la déficience en MC4R mène à l'obésité (108). Cette action sur le tonus sympathique se retrouve également par une modification de la pression artérielle chez les patients porteurs d'une mutation de *MC4R*. Les effets chez l'Homme du setmelanotide, un agoniste de MC4R, sur la diminution de la prise de poids et de l'hyperphagie, renforcent cette hypothèse (80,100).

Régulation de la croissance :

Des mutations de *MC4R* sont parfois associées à une accélération de la croissance disproportionnée par rapport au degré d'obésité des patients atteints. Une hyperinsulinémie est retrouvée chez de jeunes enfants déficient en *MC4R*, qui persiste à la suite de l'enfance et est un facteur prédominant de l'augmentation de la croissance (106). Alors que la sécrétion pulsatile d'hormone de croissance (GH) est supprimée chez les contrôles obèses, les niveaux de GH n'étaient parfois pas complètement supprimés chez les adultes déficients en *MC4R*, ce qui suggère que la suppression de la sécrétion pulsatile de GH lors de l'obésité peut être médiée par *MC4R* (109). La sécrétion de GH est sous le contrôle réciproque de la GH-releasing hormone (GHRH), qui détermine l'amplitude du pic de GH, et de la somatostatine, qui détermine les niveaux de GH en inhibant le relargage de la GHRH et la sécrétion de GH. Bien qu'il n'y ait pas d'étude concernant le relargage de GH chez les souris KO pour *MC4R*, l'axe somatotrope a été étudié chez la souris agouti, qui surexprime la protéine Agouti, un antagoniste naturel de *MC4R*. Il a été montré que bien que l'expression de l'ARN messager de *GHRH* dans le noyau arqué de la souris obèse agouti était inchangée, le contenu en peptide de la somatostatine et les niveaux de ARNm de la somatostatine dans le noyau périventriculaire étaient diminués de 50% chez les modèles wild-type (110).

Mutations des molécules en aval de *MC4R* :

Le facteur neurotrophique dérivé du cerveau (BDNF) est un des nombreux facteurs de croissance nerveux qui intervient via le tyrosine kinase receptor tropomyosin-related kinase B (TrkB). Des injections de BDNF réduisent la prise de nourriture chez les modèles de souris déficients pour *MC4R*, ce qui suggère que les intervenants BDNF et TrkB se trouveraient en aval de *MC4R* sur la voie de signalisation (111). Les souris haploinsuffisantes et celles chez qui *BDNF* a été supprimé après la naissance sont obèses, hyperphagiques et hyperactives. Il a été rapporté un enfant sévèrement obèse, avec une mémoire à court terme déficiente, et retard de développement, qui était porteur d'une mutation faux sens de-novo impactant la fonction de TrkB (112). Egalement un patient possédant une inversion de-novo touchant le locus de *BDNF* et impactant son

expression. Il a également été montré chez des patients souffrant du syndrome de WAGR une série de délétions couvrant le locus de *BDNF* associées à l'obésité précoce (113).

Une autre molécule que l'on pense agir en aval des neurones exprimant MC4R est *single minded 1* (*SIM1*), un facteur de transcription impliqué dans le développement des noyaux paraventriculaires et supraoptiques de l'hypothalamus. Une translocation équilibrée *de novo* qui touche *SIM1*, et les mutations faux sens de *SIM1* causent une obésité sévère et un retard de développement plus ou moins poussé (114). Les souris hétérozygotes déficientes pour *SIM1* sont hyperphagiques et obèses et ont une croissance augmentée (115). Par ailleurs une étude récente parue dans *JCI* montre que 9 nouveaux variants hétérozygotes de *SIM1* réduisent son activité et conduisent à des phénotypes hyperphagiques et obèses (116,117). Les cibles transcriptionnelles de *SIM1* sont inconnues, mais une de ses cibles potentielles est le neuropeptide ocytocine. Les niveaux d'ARN messagers d'ocytocine sont diminués de façon proportionnelle avec la diminution de l'expression du gène *SIM1* dans les modèles de souris déficients pour *SIM1*. L'hyperphagie observée chez les animaux haploinsuffisants pour *SIM1* est diminuée par administration d'ocytocine, et accentuée par administration d'antagonistes aux récepteurs de l'ocytocine (118). Plusieurs mutations homozygotes du gène de l'Adénylate Cyclase 3 (*ADCY3*) ont également été montré comme étant la cause de cas d'obésités accompagnés d'hyperphagie dans une population consanguine pakistanaise dans laquelle 30% des cas d'obésité sévères sont monogéniques (95).

Syndromes d'obésité développementale incluant un dysfonctionnement ciliaire :

Le syndrome de Bardet-Biedl (BBS) est une maladie autosomique récessive caractérisée par des difficultés d'apprentissage, une polydactylie, une dystrophie rétinienne, un hypogonadisme, des anomalies rénales, et une obésité. Le BBS est génétiquement hétérogène, avec des mutations dans plus de 21 gènes identifiés (119). Beaucoup de ces gènes semblent impacter des protéines localisées dans le corps basal, un élément clef du cil que l'on pense important pour la détection intracellulaire chez les cellules mammifères dont les neurones (120). D'autres anomalies de la fonction ciliaire (comme les syndromes d'Alström ou de Carpenter) peuvent également mener à l'obésité, mais avec des manifestations cliniques différentes (121). Des études récentes ont

suggéré une connexion entre la fonction ciliaire et la voie de la leptine. Le KO conditionné de protéines impliquées dans le transport intraflagellaire chez la souris conduit à l'hyperphagie et l'obésité. Ce phénotype est retrouvé quand la perte de la fonction ciliaire est limitée aux neurones et quand celle-ci touche spécifiquement les neurones exprimant POMC (122). Le rôle des cils dans des populations clés de neurones impliquées dans l'homéostasie énergétique est amené à devenir un sujet d'étude majeur dans l'avenir.

Les désordres génétiques sujets à empreinte :

Le syndrome de Prader-Willi (PWS) est un syndrome associant hypotonie, retard mental, petite taille, hypogonadisme, hyperphagie et obésité, et qui est causé par l'absence d'expression de gènes du chromosome 15q11-13, du fait de leur empreinte (le phénomène épigénétique déterminant conduisant à l'expression unique de la copie maternelle ou paternelle) à une expression uniquement de la copie paternelle du segment de chromosome. De hautes concentrations sériques de ghréline et la perte des neurones producteurs d'ocytocine dans le cerveau des patients atteints suggèrent un rôle majeur de ces éléments (123,124).

La dystrophie héréditaire d'Albright (AHO) est une maladie autosomique dominante provoquée par une mutation de *GNAS*, qui diminue l'expression de la protéine Gs-alpha. La transcription maternelle de ces mutations mène à l'AHO, qui se traduit par une petite taille, une obésité, des défauts squelettiques, et un odorat non fonctionnel, et conduit également à une résistance à différentes hormones (comme l'hormone parathyroïde) qui active la protéine Gs au niveau de ses tissus cibles. La transmission paternelle de ces mutations conduit uniquement au phénotype de l'AHO (125). Les souris déficientes en Gs-alpha d'origine maternelle sont obèses, et comme les souris déficientes pour *MC4R*, possèdent un tonus sympathique réduit, mais ne sont pas hyperphagiques (126). Les signaux transitant par Gs-alpha semblent jouer sur les effets de la mélanocortine sur la prise alimentaire, alors que les signaux transitant par par Gs-alpha semblent médier les effets de la mélanocortine sur le tonus sympathique et la dépense énergétique.

Les Copy Number Variants :

Les *copy number variant* (CNV), sont des variations génétiques incluant des gains (duplications ou insertions), des pertes (délétions) ou des réarrangements, et sont estimés comme participants pour 18% de la variabilité héritée au niveau de l'expression génique ((127) et **Figure 9**). Les mécanismes provoquant leur apparitions sont nombreux, complexes, et encore mals connus (128). Leur taille minimum est définie à 1kb, mais les CNV peuvent atteindre la taille d'un chromosome entier. Plusieurs *Copy Number Variants* (CNV) rares ont été identifiés de façon récurrente chez des patients sévèrement obèses, par rapport aux contrôles de poids normal (129,130). Les auteurs d'une étude ont analysé par puces CGH 300 individus sévèrement obèses, et pour certains également atteints de déficiences intellectuelles, issus de la cohorte *Genetics of Obesity Study*. Ils se sont intéressés aux patients non mutés pour un gène connu de l'obésité monogénique, et ont recherché les délétions rares fortement enrichies chez ces individus par rapport à plus de 7000 contrôles du *Welcome Trust Case Control Consortium 2*. Ils ont identifié une délétion d'environ 220 kb au niveau de la région chr 16p11.2, et touchant quelques gènes, dont *SH2B1*, connu pour son implication dans les voies de la leptine et de l'insuline (130). Les patients touchés par ce CNV prennent du poids lors des premières années de vie, avec une hyperphagie et des niveaux d'insuline plasmatique très élevés comparés à des individus d'âge et d'IMC comparables. Ce phénotype se rapproche de celui de souris ayant subi une délétion ciblée de *SH2B1* (131). Ils ont également identifié une délétion d'un autre segment d'en moyenne 593kb, toujours au niveau du chromosome 16p11.2 et couvrant près de 30 gènes dont *SH2B1*. Cette délétion a été également trouvée très fortement associée à l'obésité ainsi qu'à la déficience intellectuelle dans une autre étude parue en 2010, et lors de laquelle les auteurs se sont intéressés aux données de puces CGH, de qPCR (*quantitative Polymerase Chain Reaction*), de QMPSF (*quantitative multiplex polymerase chain reaction of short fluorescent fragments*) et de FISH (*Fluorescent In Situ Hybridization*) issues de plusieurs milliers de patients de différentes cohortes utilisées pour des GWAS ou adressés à des centres de diagnostic pour une obésité sévère associée ou non à une déficience intellectuelle (129). Par ailleurs, cette délétion est également associée à une macrocephalie (132). Il est à noter que la duplication en miroir de cette délétion

du segment de chromosome 16p11.2 de 593kb est associée avec une diminution de l'IMC, des déficiences intellectuelles accompagnées de troubles alimentaires, ainsi qu'avec une microcéphalie. Cette duplication provoque donc un phénotype en miroir concernant l'obésité, ainsi que pour la circonférence crânienne, par rapport à la délétion (133).

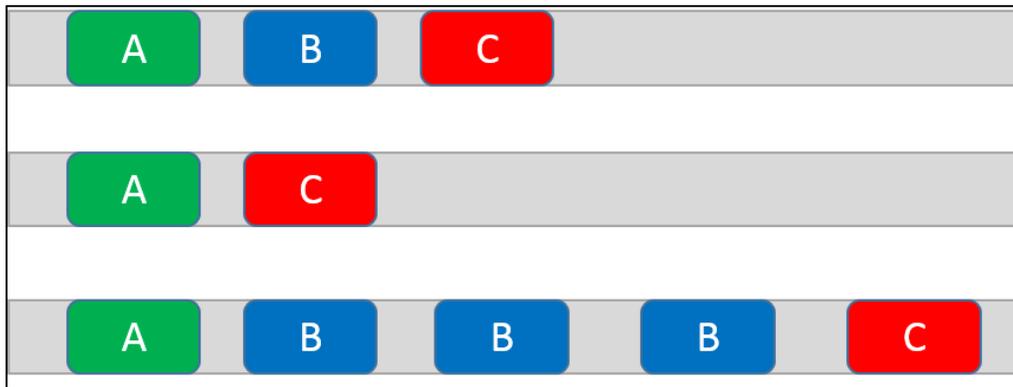


Figure 9 : Exemple de CNV. En haut, le génome de référence. Au milieu, le segment B a disparu, présence d'un CNV de type délétion. En bas, le segment B est présent en plusieurs exemplaires, présence d'un CNV de type duplication.

Les variants génétiques communs associés à un risque plus élevé de survenue de l'obésité :

En génotypant plusieurs centaines de milliers de SNPs sur plus de 75% du génome, les genome wide association studies (GWAS) effectués sur des cohortes et s'intéressant à l'IMC ou à des cas-contrôles conçus pour l'étude l'obésité ont permis l'identification de nombreux loci d'intérêt. Ces études sont basées sur la supposition que l'héritabilité des maladies communes est constituées d'un grand nombre de variants pathogènes communs, ayant une fréquence de l'allèle mineur (MAF) supérieur à 1% (8). Les premiers loci détectés contenaient les variants du gène *FTO* (59) et des variants environ 200 kb en aval de *MC4R* (134).

Il est important de noter que les GWAS identifient des associations entre des traits phénotypiques ou des maladies et des régions du génome, et non des gènes. Beaucoup des éléments identifiés se trouvent dans des régions non codantes du génome, et l'interprétation du rôle de ces éléments se base en grande partie sur les annotations connues au niveau de ces régions. Les gènes se trouvant au niveau de régions d'intérêt issues de GWAS sont classés comme candidats potentiels

en tant que cause de l'association, mais l'association pourrait également provenir d'interactions génétiques à longue distance, ou d'autres mécanismes encore moins connus. La compréhension des mécanismes biologiques à l'œuvre lors de ces associations est encore balbutiante.

Trouver de nouvelles étiologies dans l'obésité humaine :

Les variants fréquents découverts par les GWAS ont tous en commun d'être à effets modérés sur le risque d'obésité (*odds ratio* par allèle compris entre 1,2 et 1,5), et la variabilité expliquée par les loci issus de GWAS reste limitée, autour de 10%. Il est possible que l'héritabilité des phénotypes liés à l'obésité ait été surestimée, étant donné que les effets de l'environnement partagé entre individus, et dans le cas de jumeaux, de l'environnement in utero, sont difficiles à séparer de ceux hérités (8).

Concernant l'obésité, les variants communs influençant la susceptibilité à l'obésité, et les variants rares plus pénétrants, dont des CNV, peuvent se trouver aux mêmes loci (par exemple au niveau de *MC4R* ou *SH2B1*) (134). Actuellement, la recherche de variants génétiques rares se base sur le séquençage de nouvelle génération, ou séquençage à haut débit, via le séquençage d'exome entier (WES), grâce auquel des dizaines de milliers de variants codants sont identifiés par individu (135). Tous les variants ne seront pas identifiés, à cause de variabilités dans les profondeurs de couverture, de mauvais alignements des séquences (*reads*), et de la complexité des régions répétées du génome. Les variants détectés doivent être minutieusement examinés pour identifier les variants pathogènes potentiels. Des données fonctionnelles, ainsi que des prédictions des conséquences des variants sur les protéines codées peuvent également être incluses comme critères. Alors que les méthodes d'analyses informatiques et statistiques progressent en permanence, la puissance du séquençage sera au final limitée par le matériel génétique séquencé. Comme toutes les autres approches sans hypothèse, les approches méthodologiques initiales ont généralement eu tendance à séquencer des individus appariés (136), ou opposés au niveau du spectre phénotypique (par exemple comparer des individus normo-glycémiques avec des individus extrêmement obèses) (137). Quasiment toutes les études visant à trouver des gènes associés à l'obésité se sont intéressées aux phénotypes extrêmes d'obésité. Mais les modèles

animaux, ainsi que d'autres modèles ou des analogies avec d'autres maladies suggèrent qu'il serait utile d'étudier des individus minces et en bonne santé. La minceur semble être un trait stable et héritable tout comme l'obésité (138). L'application de l'approche du séquençage à des groupes de patients particuliers, aussi bien minces qu'obèses, a le potentiel d'apporter de grands progrès dans la découverte de gènes liés au métabolisme humain.

La technologie du séquençage a déjà été utilisée pour le diagnostic génétique et l'identification des nouvelles protéines impactées dans plusieurs maladies mendéliennes, mais son implication concernant des maladies complexes comme l'obésité est plus compliquée (139). Comprendre le potentiel pathogénique de mutations individuelles dans le contexte de variation au niveau d'autres loci pouvant également intervenir complexifie le développement d'approches physiologiquement appropriées au niveau de modèles cellulaires ou animaux pour décortiquer la causalité de ces variants.

Bien qu'il soit difficile de faire la distinction entre des effets purement environnementaux et des interactions entre l'environnement et des facteurs épigénétiques, la mesure des variations épigénétiques à l'échelle du génome a été récemment rendue possible par des techniques comme les puces (*microarray*) spécifiques ciblant des sites méthylés. Chez les primates, l'exposition de la mère à une nourriture riche en graisse durant la grossesse a d'importantes répercussions sur le développement de maladies métaboliques chez sa progéniture (140). Chez l'humain, une étude a montré l'existence de plusieurs régions méthylées différentes entre individus, mais stables dans le temps, et qui sont associées avec l'IMC (141).

De nombreux autres éléments influençant ou régulant la survenue et l'évolution de l'obésité seront encore découverts, mais leur transposition en éléments utilisables dans une pratique clinique reste un gros défi à venir.

Diagnostic, recherche et prévention :

Phénotypes obèses :

Comme mentionné précédemment, les patients obèses souffrant de désordres métaboliques, comme l'insulino-résistance, ou la dyslipidémie, ont régulièrement un excès de graisse

abdominale. Les scanners ou l'imagerie par résonance magnétique (IRM) peuvent fournir une estimation de la quantité de graisse stockée dans les différents tissus, dépôts qui varient grandement entre individus, mais ils ne sont pas utilisés en routine pour réaliser un diagnostic de l'obésité. La présence simultanée d'une forte adiposité combinée avec une faible masse ainsi qu'une altération de la fonction musculaire (la sarcopénie) pourrait être associée avec une prévalence augmentée d'anomalies cardio-métaboliques par rapport à des individus de même IMC mais avec une masse musculaire plus importante (142). Ce diagnostic est connu sous le nom « d'obésité sarcopénique ».

Le phénotype obèse métaboliquement sain est une catégorie utilisée pour décrire les individus obèses qui ne présentent pas les critères cliniques du syndrome métabolique (143). A l'âge adulte, le principal élément prédicteur du passage d'un phénotype métaboliquement sain à celui d'un phénotype présentant les symptômes du syndrome métabolique est l'accumulation de graisse viscérale (144). Néanmoins, les discussions sont nombreuses concernant le phénotype métaboliquement sain, car celui-ci est mal défini et a également tendance à se présenter comme un phénomène transitoire, voire à ne pas être totalement bénin étant donné que ce n'est souvent qu'une question de temps avant que les complications habituelles pour un individu obèse ne se développent (143).

Enfin, une obésité à poids normal a également été décrite comme une pathologie métabolique. Plusieurs études ont mis en évidence les risques de santé pour les individus atteints par ce phénomène, en particulier des populations asiatiques, dont les valeurs limites d'IMC normal sont généralement entre 23 et 24 kg/m², pour les hommes et les femmes, respectivement (145). Ce qui met également en évidence que les valeurs d'IMC doivent être adaptés aux différentes ethnies.

Evaluation :

L'IMC doit être mesuré pour tous les individus dans des conditions d'évaluations cliniques. La chose étrange dans l'utilisation de l'IMC dans la caractérisation de l'obésité est qu'il n'est pas un excellent indicateur de l'adiposité, mais un très bon estimateur du risque de développer des

maladies cardiovasculaires. Et l'utilisation du rapport tour de taille sur tour de hanche améliore encore la précision de cet indice pour la prédiction du risque de développer des maladies cardiovasculaires (146). Les mesures de l'IMC devraient être accompagnées d'indices représentatifs de la forme que prend cette obésité (comme le tour de taille) et d'autres indicateurs (comme la pression sanguine, la glycémie ou les niveaux de cholestérol) de façon à identifier précisément le risque encouru par les individus obèses.

Le surpoids et l'obésité sont définis chez les adultes par l'utilisation de l'IMC. Et chez les enfants, l'IMC Z-score est souvent préféré (un score basé sur la déviation standard de l'IMC), de façon à analyser plus facilement différentes ethnies (147). Pour surmonter les limites de l'IMC dans la pratique clinique, en plus de l'addition du tour de taille, plusieurs indicateurs concernant l'obésité ont été développés. Ces indices ont été développés de façon à prendre en compte d'autres facteurs, comme la dyslipidémie et l'homéostasie glucidique, et à faciliter la pratique des professionnels de santé pour la catégorisation des différents types d'obésité rencontrés et la bonne adaptation des traitements aux risques encourus par les différents patients. On peut citer par exemple l'*Edmonton Obesity Staging System*, le *Cardiometabolic Staging System*, l'ATPIII panel et le *Framingham risk score*, entre autres (148).

Des mesures de l'épaisseur des plis cutanés et des analyses de l'impédance bioélectrique peuvent également être utilisées en clinique pour mesurer la quantité totale de graisse. Néanmoins la mesure de ces plis et de la conductivité du corps ne sont que des estimateurs de la masse grasse et de la masse maigre du corps.

Il existe néanmoins des méthodes de référence pour mesurer la graisse totale du corps, mais leur usage est limité aux laboratoires, étant donné la nécessité d'avoir recours à des techniques lourdes incluant la densitométrie ou des méthodes basées sur l'imagerie, comme l'IRM ou les scanners (149).

Prévention :

La prévention de l'obésité doit se focaliser sur le maintien de la perte de poids ou sur le contrôle de la prise de poids. Des stratégies de prévention potentielles sont les programmes publics de

promotion d'un mode de vie sain, ainsi que leurs équivalents marketing provenant d'entreprises privées, visant les habitudes de vie et l'environnement des patients.

De façon générale, ces interventions doivent commencer le plus tôt possible, et également chez les mères. Néanmoins le recul concernant ces interventions prénatales est limité. En effet une revue systématique des différentes interventions ayant eu lieu durant les premières étapes de la vie (**Figure 4**) met en évidence que seules deux ont eu lieu lors de la période prénatale, sans effet sur l'IMC des enfants l'ayant suivie (150). Ces interventions incluent des pratiques visant le mode de vie durant la grossesse, comme du conseil diététique, du coaching et de l'activité physique, le suivi et le contrôle du diabète gestationnel via des conseils diététiques, un contrôle de la glycémie, ou l'administration d'insuline si nécessaire (151). Comme exemple d'interventions efficaces, on peut citer que 8 visites à domicile par des professionnels de santé se concentrant sur le régime alimentaire de l'enfant et son activité physique ont été effectuées durant les 24 premiers mois de la vie de l'enfant (152). Ou encore des professionnels de santé ayant apporté des conseils pour le régime alimentaire de la mère et de l'enfant, en combinant des visites à domicile et des séances en groupes (153). Selon une autre revue, quelques études effectuées chez des enfants de moins de deux ans et visant à encourager de bonnes pratiques alimentaires en combinant de l'éducation nutritionnelle et de l'activité physique ont eu lieu, mais sans effet apparent sur le poids des enfants les ayant suivies (154).

Même si l'efficacité de ces programmes de prévention de l'obésité chez les enfants ont montré des résultats divers et parfois conflictuels, plusieurs stratégies ont prouvé leur efficacité, comme la mise en place d'une éducation nutritionnelle, une augmentation de la quantité d'heures d'activité physiques dans les programmes scolaires, et la mise à disposition de nourriture compatible avec un mode de vie sain à l'école (155). Peu d'informations sont actuellement disponibles concernant l'efficacité des stratégies visant l'environnement de vie des enfants, ou les technologies de monitoring santé concernant les enfants et les parents. Pratiquer ces différentes stratégies d'information en tant que soins de première ligne pourrait améliorer les pratiques parentales et encourager les habitudes alimentaires plus saines, mais les études actuelles ne font encore état que d'un effet limité.

La complexité de l'obésité nécessite des interventions à plusieurs niveaux qui requièrent une approche systémique. Les approches à plusieurs niveaux et à plusieurs éléments se concentrent sur le fait de changer les habitudes de vie, en partant de l'individu, souvent des enfants ou des adolescents, en passant par les écoles, jusqu'aux différentes communautés dont les enfants font partie. Pour développer efficacement ces stratégies, celles-ci doivent prendre en compte les populations auxquelles elles s'adressent, ainsi que les différentes opportunités qu'elles ont pour les toucher. Selon le rapport de l'OMS intitulé *Ending Childhood Obesity* (156), les différents facteurs à prendre en compte incluent le lien entre la santé maternelle et infantile, l'éducation nutritionnelle et l'éducation de santé en général, la perception des parents de ce qu'est une croissance et un développement sain pour leur enfant, la pression des pairs, le régime alimentaire familial et les habitudes de vie sportive, ainsi que le rôle de l'environnement construit et de la place de la nourriture dans celui-ci. Un point de vue très large est nécessaire, étant donné que se focaliser sur un nombre limité d'éléments ou uniquement sur l'aspect santé a peu de chances d'être efficace. Enfin malgré le grand nombre de limites identifiées concernant les programmes de prévention de l'obésité infantile, ceux-ci doivent continuer, étant donné qu'aucune approche n'est à l'heure actuelle plus efficace. L'auto-suivi est également un élément très important dans la perte de poids (157).

Les objectifs actuels du management de l'obésité sont centrés sur la perte de poids avec une approche progressive commençant par des traitements à bas risque, comme des interventions visant le style de vie, des changements d'alimentation et de l'activité physique en tant que première approche, suivi ensuite par des traitements médicamenteux ou la chirurgie dans certains cas. Des bonnes pratiques ont été mises en place aux Etats-Unis, au Royaume-Uni et en Europe pour aider les professionnels de santé dans le traitement des patients souffrant d'obésité (158,159). La réponse à ces différents traitements varie selon les individus, et les stratégies de perte de poids peuvent être différentes des stratégies de maintien du poids.

Interventions sur le style de vie :

Les modifications du comportement dans des programmes visant le style de vie ont été une partie importante des programmes de réduction du poids depuis plus d'un quart de siècle. Une méta-analyse a montré une perte de poids moyenne de 3,01 kg lors de changements de style de vie (160). Ces changements aident les patients à comprendre et surveiller leur comportement alimentaire, des éléments qui déclenchent leurs excès, en passant par les lieux, la nourriture consommée et la vitesse à laquelle elle est consommée, jusqu'aux conséquences de ces excès et comment les éviter. Ces interventions sur le style de vie consistent également en des stratégies pour aider les patients à mettre en place un comportement sain et volontaire, à apprendre à dominer leurs pulsions et à gérer le stress. Les innovations les plus récentes dans l'univers de ces interventions visant le style de vie sous l'apparition des outils de suivi en ligne, incluant des mails de suivi automatique, du conseil et des thérapies en ligne (161).

L'effet des Régimes :

Les régimes en tant que méthodes pour contrôler le poids ont seulement été rigoureusement testés par essais randomisés depuis 20 ans. Une balance énergétique négative via restriction calorique est habituellement utilisée pour réaliser une perte de poids lors d'un régime, et tous les régimes produisent en moyenne un effet. Une restriction de 500 kcal (kilocalories) par jour ou une restriction calorique individualisée correspondant à -30% des apports habituels, ou encore un régime avec 1200 kcal absorbée par jour pour les femmes et 1500 kcal pour les hommes est généralement ce qui est recommandé. De façon intéressante, améliorer le régime alimentaire dans sa qualité sans une réduction de l'apport calorique (par exemple le passage au régime alimentaire méditerranéen) peut également permettre de perdre du poids (162). Dans une méta-analyse de plusieurs régimes populaires, les régimes pauvres en carbohydrates permettent en moyenne une perte de poids de 7,25 kg après 12 mois, en comparaison des 7,27 kg de perte moyenne de 12 mois pour les régimes pauvres en graisses (163). Les différences de perte de poids entre individus lors de régimes sont très élevées, certaines personnes perdant beaucoup de poids

et d'autres gagnants même du poids. Des variations génétiques pourraient expliquer au moins une partie des différences dans la perte de poids entre individus lors de régimes (164).

Des différences dans la composition en macronutriments de ces différents régimes n'en favorisent pas toujours un par rapport aux autres, ce qui inclut un stress pour le patient dans le choix du régime à effectuer. Il est préférable de favoriser un régime alimentaire sain à un régime uniquement de restriction alimentaire, comme le passage au régime alimentaire méditerranéen, qui peut de plus être utilisé avec un faible apport calorique (165). Les régimes pauvres en graisses sont associés à des niveaux plus faibles de lipoprotéines de faible densité, et les régimes riches en graisses à des niveaux plus faibles de triglycérides et de lipoprotéines de haute densité. Mais ces effets sont valables à court terme et disparaissent sur le long terme quand la perte de poids atteint un plateau et que les poids ne sont plus différents entre les différents groupes de régimes (166).

Pour un maintien à long terme de la perte de poids, l'adhésion à un régime alimentaire sain est importante, mais la composition en macronutriments pourrait également jouer un rôle. Une consommation plus élevée de protéines, de nourriture avec un indice glycémique faible et une diminution de la consommation de graisse pourraient aider à maintenir la perte de poids, alors que les régimes riches en carbohydrates pourraient favoriser la reprise de poids (166). Les recommandations pour le contrôle du poids devraient prendre en considération la qualité générale du régime utilisé et surtout les effets à long terme de ce régime, ainsi que le rôle d'acides aminés spécifiques, des acides gras consommés, des substances bioactives consommées, le nombre et les horaires des repas, ainsi que la répartition en différents macronutriments (167).

Activité physique :

Comme pour tous les individus sédentaires, les recommandations actuelles sont d'augmenter progressivement l'exercice physique aérobie chez les patients obèses, via des exercices peu intenses comme la marche rapide, pour atteindre un objectif de plus de 150 minutes d'activité par semaine (168). Cette stratégie a des bénéfices pour la santé indépendamment de la perte de poids, étant donné qu'une activité physique modérée diminue le risque de développer un diabète

ou une maladie cardiovasculaire, ce qui est potentiellement lié au fait que cette activité physique modérée peut diminuer le dépôt de tissu adipeux viscéral et les dépôts de graisses ectopiques (169). Une méta-analyse a montré que l'activité physique permet une réduction de poids de 1 à 1,5 kg plus importante que via l'utilisation d'un régime seul (170). Pour le maintien du poids à long terme, 60 à 90 min d'exercice par jour pourraient être requis (171). Lors de l'essai clinique *Look AHEAD*, l'augmentation de l'activité physique durant la première année n'a pas été maintenue à 4 ans, et a résulté en une augmentation du poids par rapport à la fin de la première année, même si le poids des patients est resté inférieur à leur poids initial 8 ans après le début de l'étude (172). Le type d'activité physique (par exemple, de l'aérobic contre des exercices de résistance, ou des exercices de forte ou faible intensité) ne semble pas jouer sur la perte de poids. L'efficacité de l'activité physique sur la réduction du poids a également été montrée chez des patients porteurs de variants génétiques associés à l'obésité, comme ceux touchant le gène *FTO* (173).

Gestion médicalisée :

Les médicaments utilisables chez les patients obèses sont approuvés dans le cadre de leur utilisation conjointe avec un régime adapté et de l'exercice physique, et aucun n'a été approuvé pour une utilisation lors de la grossesse, lors de l'allaitement ou chez les enfants (159). L'utilisation de ces médicaments doit être réservée à des patients avec une obésité à risque modéré ou élevé (soit un IMC supérieur à 30 kg/m², ou à 27 kg/m² si des comorbidités sont présentes). Etant donné que ces traitements sont prescrits à des patients ayant des difficultés pour perdre ou maintenir leur poids, un historique des raisons des échecs dans la gestion du poids doit être pris en compte. Tous ces traitements fonctionnent en aidant les patients à mieux suivre leur régime, à part pour l'orlistat ou le cetilistat (uniquement disponibles au Japon), qui aident à mettre en place un régime pauvre en graisses. De ce fait, ces médicaments ne doivent être utilisés que dans le cadre d'un effort réel de suivi d'un régime. Il n'y a jusqu'ici aucun médicament idéal. Tous ceux disponibles peuvent être efficaces chez un patient. Si un patient n'a pas perdu 4 à 5% de son poids après 3 mois, le traitement médicamenteux doit être arrêté et un nouveau

traitement basé ou non sur un autre médicament doit être mis en place (159). La disponibilité de ces différents médicaments varie selon les pays. L'utilisation d'amphétamines, de méthamphétamines ou de phénmetrazine est connue et approuvée par la FDA (*Food and Drugs Administration*) pour des traitements à court terme (soit moins de 12 semaines), mais pas pour un traitement à long terme d'un patient obèse.

Chirurgie bariatrique :

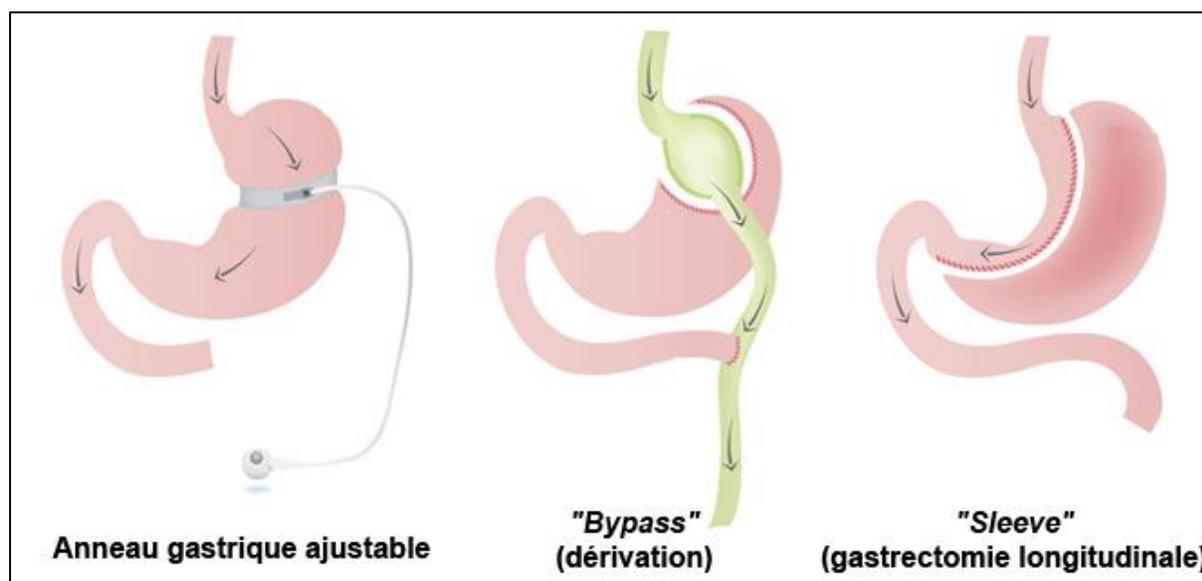


Figure 10 : Différentes opérations de chirurgie bariatrique (174).

L'utilisation de la chirurgie bariatrique (également connu sous le nom de chirurgie métabolique) a été rapidement adoptée comme un traitement possible des obésités sévères, et ce en partie grâce aux procédures laparoscopiques à faibles risques. Le critère pour envisager une de ces chirurgies chez un patient est que celui-ci ait un IMC supérieur à 40 kg/m² ou à 35 kg/m² avec comorbidités, comme l'hypertension ou la dyslipidémie. Les patients souffrants d'un pré-diabète ou d'une apparition récente du diabète peuvent être candidats à la chirurgie bariatrique avec un IMC compris entre 30 et 35 kg/m² (175). La chirurgie bariatrique peut être réalisée au cas par cas chez des adolescents sévèrement obèses. Près d'un demi-million de chirurgies bariatriques ont été réalisées dans le monde en 2013 (176). Il existe toute un éventail de procédures et d'interventions qui résultent en des niveaux variables de perte de poids et ont chacune leurs

propres risques et bénéfiques, qui doivent être pris en compte dans le choix de l'une ou l'autre de ces procédures pour un patient donné (**Figure 10**) (177).

L'évolution de la santé des patients ayant subi une chirurgie bariatrique montre généralement de bons résultats. L'étude *Swedish Obese Subject* a suivi 2 000 patients durant 20 ans et ayant été opérés selon 3 protocoles différents : la gastroplastie verticale calibrée, l'anneau gastrique ou le bypass gastrique Roux-en-Y (**Figure 10**). La mortalité chez ces patients a diminué de 24%, principalement grâce un risque réduit d'infarctus du myocarde, et chez les femmes un risque réduit de cancer par rapport au groupe ayant reçu les soins non chirurgicaux (178). D'autres comorbidités, comme le diabète de type 2 ou l'apnée du sommeil sont également réduits, ce qui permet une amélioration de la qualité de vie du patient. La perte de poids moyenne suite à ces interventions est de 23% à 1 an et de 18% à 20 ans.

Le contrôle de la glycémie s'améliore rapidement chez les patients obèses diabétiques de type 2 suite à la chirurgie bariatrique, ce qui suggère qu'une partie de l'amélioration de l'état métabolique est indépendante de la perte de poids (178). Une étude randomisée de la chirurgie bariatrique contre un traitement médical intensif contre le diabète de type 2 a montré un meilleur contrôle de la glycémie suite à l'intervention chirurgicale (179). Une étude indique également que le risque futur de développer des complications micro ou macrovasculaires est également réduit suite à la chirurgie bariatrique (180). La rémission du diabète de type 2 est fortement induite par la perte de poids, bien que certaines hormones comme le glucagon-peptide 1, le polypeptide YY ou la ghreline aient également été impliqués (181). Mais ces données doivent être mises en considération avec les risques encourus lors de la chirurgie, qui peut entraîner des complications chez un nombre non négligeable de patients. Parmi les 6 études incluses dans une revue, un seul décès a été rapporté lors de l'utilisation de la chirurgie laparoscopique, mais le taux de réopération était de 6,7 à 24% dans le cas du bypass gastrique Roux-en-Y, et de 3,3 à 34% dans le cas de l'anneau gastrique (182).

Certains patients trouvent difficile de s'adapter aux nombreux changements liés à la quantité et le type de nourriture qu'ils peuvent consommer suite à l'intervention. Une surveillance et complémentation à vie sont nécessaires pour les carences en vitamines et minéraux dus aux difficultés d'absorption suivant l'opération. Le *dumping syndrome*, le reflux gastro-intestinal ou

l'hypoglycémie peuvent être particulièrement handicapants et difficiles à traiter. Le regain de poids peut également être un problème, et les chirurgies de corrections sont porteuses de nombreux risques sans garanties de succès (183).

Par ailleurs, il a été montré que les patients porteurs de mutation du gène *MC4R* subissaient davantage de complications suite à une opération de chirurgie bariatrique, comme des troubles de l'alimentation (184). Ces données sont encore préliminaires mais attirent l'attention sur une problématique à étudier, qui pourrait conduire à la réalisation systématique d'une recherche de variants avant la réalisation d'une chirurgie bariatrique.

Qualité de vie :

Le poids de l'obésité et du surpoids est considérable et représente une priorité pour le système de santé. D'ici 2030, les coûts de santé liés à l'obésité atteindront 48 à 66 milliards de dollars par an aux Etats-Unis (40).

L'augmentation du risque pour la population de développer les comorbidités liées à l'obésité, comme le diabète de type 2, les maladies cardiovasculaires et certains types de cancers représente une grande partie de ce poids financier (40). Néanmoins, le surpoids et l'obésité sont également connus pour avoir un effet conséquent sur le statut émotionnel d'un individu, son estime de soi et sa santé psychologique (185). Le surpoids et l'obésité sont également associés à une qualité de vie liée à la santé moins élevée que pour les individus de poids normal. Des études conduites aux États-Unis, au Canada, en Europe ou en Australie ont relevé une tendance montrant que quand l'IMC augmente, la qualité de vie liée à la santé diminue (186). Par exemple, dans la cohorte *Rancho Bernardo*, en Californie, le groupe de poids normal avait le score le plus élevé, suivi par le groupe d'individus maigres, puis le groupe d'individus en surpoids, et enfin le groupe d'individus obèses (187).

L'estimation de la qualité de vie liée à la santé pour l'*US Medical Expenditure Survey* a montré qu'à l'échelle des Etats-Unis, les hommes et femmes en surpoids perdaient respectivement 270 000 et 1,8 millions d'années de vie en bonne santé ajustées par an par rapport au groupe d'individus de poids normal. Les hommes et femmes obèses perdaient eux respectivement 1,9 et

3,4 millions d'années de vie en bonne santé par an par rapport au groupe d'individus de poids normal. Les auteurs de l'étude suggèrent que ces différences entre sexe peuvent s'expliquer par 3 raisons principales : les hommes très musclés ont plus de chances d'être inclus par erreur dans le groupe des individus en surpoids, voire même des individus obèses. L'obésité chez les femmes a tendance à plutôt impacter la morbidité que la mortalité, et celles-ci ont en plus une espérance de vie plus longue que celle des hommes. Et enfin l'impact psychologique du surpoids ou de l'obésité est plus important chez les femmes que chez les hommes (188).

Un cercle vicieux peut apparaître dans lequel une moins bonne qualité de vie liée à la santé et une diminution de la santé psychologique peuvent agir comme déclencheur d'une prise de poids importante (189). Une quantification de la perte de poids nécessaire pour obtenir une amélioration cliniquement parlant de la qualité de vie est nécessaire. Par exemple, deux larges études, *Nurses-I* et *Nurses-II Health Studies*, ont montré qu'une diminution des scores de bien-être physique était associée à une faible activité physique, et qu'une augmentation des scores de bien-être physique était associée à une activité physique plus élevée, mais le rôle de cette activité physique est limité sur le bien-être car elle génère également des soucis physiques, comme des douleurs (190). Pour la majorité des gens souffrant d'obésité sévère, des changements bénéfiques minimes concernant la qualité de vie liée à la santé ont seulement été atteints après une perte de poids de plus de 20% sur 2 ans (191).

Bien qu'il semble évident que la perte de poids mènerait à une amélioration de la qualité de vie liée à la santé, l'ampleur de la perte de poids nécessaire pour une amélioration significative de la qualité de vie liée à la santé chez des individus obèses a été peu étudiée, et très peu de publications se sont intéressées à la corrélation entre le changement de poids et le changement de qualité de vie liée à la santé. Récemment, une revue de la littérature des Etats-Unis, incluant 20 publications qui ont analysé des publications pour mesurer la qualité de vie a montré qu'une perte de poids cliniquement significative (généralement plus de 5%) après chirurgie bariatrique serait associée à une amélioration de la qualité de vie liée à la santé. Dans les études sans chirurgie bariatrique avec perte de poids supérieur à 5%, des améliorations de certains aspects de la qualité de vie liée à la santé ont pu être notées, bien que la causalité soit encore à démontrer. A la fois dans les études concernant la chirurgie bariatrique ou celles ne l'ayant pas utilisée, plusieurs

aspects de la qualité de vie liée à la santé ont progressé, en particulier ceux liés à l'aspect physique de la qualité de vie, mais rarement dans l'aspect mental de la qualité de vie (192).

Chez les enfants et adolescents, le surpoids et l'obésité ne mènent généralement pas immédiatement à des comorbidités, mais sont associés avec des styles de vie sédentaires, une estime de soi faible, une exclusion sociale, une réussite scolaire diminuée et une faible qualité de vie (193).

Nutrition de précision individualisée :

Le rôle de la génétique, de l'épigénétique, de la nutriginomique (l'étude des interactions entre la nutrition et les gènes fondée sur les technologies « omiques ») ou la nutrition personnalisée doivent être mieux compris pour améliorer la prise en charge de l'obésité. Néanmoins, il est important de comparer les valeurs individuelles et épidémiologiques des approches personnalisées avec les valeurs à l'échelle d'une population. La nutrition de précision basée sur des données phénotypiques et génotypiques semble prometteuse mais ne permettra vraisemblablement pas de régler l'épidémie d'obésité dans son ensemble, étant donné que le problème est surtout sociétal. Les éléments suivants doivent être pris en compte pour la médecine et la nutrition de précision dans le but d'individualiser les traitements : les antécédents familiaux, le style de vie et les antécédents cliniques du patient, les préférences alimentaires, les allergies et intolérances, l'épigénétique, la culture et la religion, le statut socio-économique, le nombre d'heures de sommeil et les périodes de travail, les éventuelles heures supplémentaires à effectuer ou encore la chronobiologie (l'étude des rythmes biologiques). La nutrition de précision dépend des données génétiques et épigénétiques ainsi que des informations phénotypiques sur l'individu (194). Les nouvelles approches « omiques », incluant la nutriginétique, la transcriptomique, la métabolomique, la métagénomique (par l'étude du microbiote) ou encore l'épigénétique, permises grâce aux nouvelles technologies, et notamment l'ensemble des évolutions du séquençage humain, sont importantes pour l'implémentation future de la nutrition individualisée, associées à une mesure plus précise de la composition du corps, et l'arrivée des

analyses *big data* et bio-informatiques pour l'intégration et l'interprétation de toute ces données (195).

Perspectives :

Physiopathologie :

Une meilleure compréhension de l'appétit ou des préférences alimentaires est nécessaire, et doit inclure la satiété, les récepteurs du goût, les voies nerveuses et les neurotransmetteurs. La capacité de thermogénèse du tissu adipeux brun et les récentes découvertes du tissu adipeux beige sont également des domaines suscitant de plus en plus d'intérêt (196). De ce point de vue, une meilleure compréhension de la transformation d'adipocytes blancs en adipocytes beiges puis bruns pour augmenter la thermogénèse est nécessaire.

Prévention :

Il est certain qu'une attention plus importante ainsi qu'un soutien financier plus important doivent être mis en place pour la prévention de l'obésité. Au niveau d'une population, des politiques publiques et stratégies économiques sont nécessaires pour améliorer les environnements physiques et alimentaires, l'industrie agroalimentaire et le système de santé de façon à inverser la tendance de la progression de l'obésité. De plus, de nombreuses communautés travaillent à faciliter l'accès à l'activité physique et l'augmentation du nombre d'heures de sommeil en modifiant les horaires télévisés et en augmentant la quantité d'aires dites « vertes » adaptées à l'activité physique, de trottoirs, ou de pistes cyclables (197). De plus, habiter à haute altitude semble prévenir la survenue de l'obésité et l'apparition du syndrome métabolique (198). Une hypothèse est que la plus faible disponibilité d'oxygène à haute altitude augmente la dépense calorique à activité comparable par rapport aux altitudes plus faibles, et que cet effort augmente la santé cardio-respiratoire sur le long terme. Plusieurs approches comportementales ont été conçues pour les adultes et les enfants, suggérant que les politiques de santé publique alimentaire doivent coexister avec une approche de la nutrition individualisée (199).

Médecine de précision :

La médecine de précision commence à se développer dans la prise en charge des cas d'obésité sévère. Le séquençage de l'ADN permet maintenant de déterminer d'éventuelles causes génétiques de l'obésité, et dans certains cas d'obésités monogéniques, de mettre en place un traitement.

Des administrations de leptine peuvent permettre aux patients mutés pour le gène de la leptine *LEP* de réguler leur prise alimentaire, de perdre du poids, et de retrouver une puberté fonctionnelle (200). Et les patients mutés pour le gène *LEPR* ou pour *POMC* pourront bientôt bénéficier d'un traitement à base de setmelanotide (80,100).

Séquençage de l'ADN :

La recherche génétique s'appuie aujourd'hui principalement sur le séquençage de l'ADN, comme dans le cas de l'obésité.

Le développement des technologies de séquençage de l'ADN possède une histoire riche, avec plusieurs changements de paradigmes ayant eu lieu en quelques décennies.

Les débuts du séquençage :

Fred Sanger a voué toute sa carrière scientifique à la détermination de la séquence primaire de ce qu'on connaît aujourd'hui comme l'ADN (Acide Déoxyribo Nucléique), car il pensait que le fait de connaître la structure chimique spécifique des molécules biologiques était nécessaire pour comprendre leur fonctionnement (201). Mais malgré l'état actuel des capacités de séquençage et ses performances sur l'ADN, ce sont les protéines et l'ARN dont les structures ont été étudiées en premier.

La première séquence protéique, celle de l'insuline, a été élucidée au début des années 1950 par Fred Sanger, qui a fragmenté ses deux chaînes, étudié séparément chacun des fragments puis rassemblé ses résultats pour obtenir une séquence complète. Son travail a montré que les protéines ont un schéma défini d'acides aminés successifs (202). Par la suite le développement de la dégradation d'Edman, une élimination répétée du résidu N-terminal de la chaîne peptidique, a simplifié le séquençage des protéines (203). Bien que ces méthodes soient lourdes, de nombreuses protéines ont été séquencées jusqu'à la fin des années 1960, et il est devenu évident que chaque séquence protéique variait entre les espèces et entre les individus.

Au cours des années 1960, le séquençage de l'ARN a suivi le même schéma général : l'ARN (Acide Ribo Nucléique) d'une espèce était d'abord fragmenté par des RNAses, puis les morceaux étaient séparés par chromatographie et électrophorèse, puis chaque fragment était étudié par digestion séquentielle via exonucléase, et la séquence finale était déduite en chevauchant les différents fragments. La première séquence d'ARN élucidée, celle du tRNA (ARN de transfert) de l'alanine, a nécessité le travail de 5 personnes durant 3 ans avec un gramme de matériel pur isolé de 140

kg de levure pour déterminer ses 76 nucléotides (204). Ce processus a été grandement simplifié par l'arrivée des techniques « d'empreinte » qui se basent sur la séparation de fragments d'ARN marqués radioactivement et leur visualisation en 2 dimensions, et où leur position traduisait leur taille et leur séquence (205).

L'invention du séquençage de l'ADN :

Les premières tentatives de séquençage de l'ADN étaient complexes à mettre en place. En 1968, Wu a décrit une méthode basée sur l'utilisation d'amorces pour l'extension qui a permis de déterminer 12 bases au niveau des extrémités cohésives du bactériophage lambda (206). En 1973, Gilbert et Maxam ont élucidé 24 bases du site de liaison au répresseur du lactose, en le copiant sous forme d'ARN puis en séquençant ses fragments. Cela a pris 2 ans : une base par mois (207). En 1976, le développement de 2 méthodes permettant de décoder des centaines de bases par jour a révolutionné le domaine. Ces deux méthodes (la méthode de la terminaison de chaîne par Sanger et Coulson (**Figure 11**), et la méthode de dégradation chimique par Maxam et Gilbert (**Figure 12**)) utilisaient la distance entre un marqueur radioactif et les positions occupées par chacune des bases pour déterminer l'ordre de ces bases. La méthode Sanger implique 4 extensions d'une amorce marquée par une ADN polymérase, chacune avec des traces d'un nucléotide terminateur d'extension, de façon à produire des fragments de différentes longueurs (208). La méthode Gilbert utilise des fragments d'ADN marqués aux extrémités, puis crée des clivages spécifiques de chacune des bases par 4 réactions chimiques différentes (209). Dans les deux cas, les différents fragments obtenus sont séparés par longueur à la base près via électrophorèse sur gel de polyacrylamide, pour déterminer la séquence du fragment d'ADN. Ces gels de 4 lignes, une par base, sont ensuite placés sous rayons X pour produire une image sous forme d'échelle à partir de laquelle la séquence du fragment d'ADN étudié pouvait être directement lue.

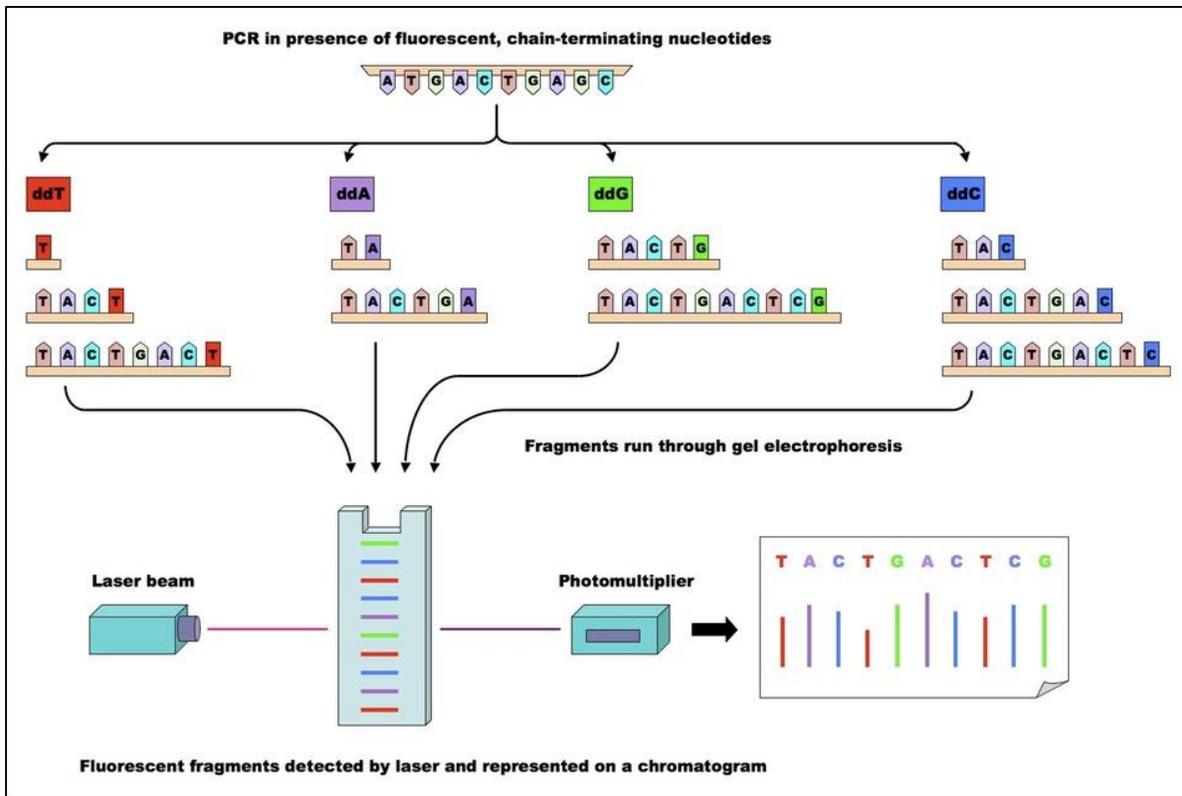


Figure 11 : Méthode de séquençage Sanger (210).

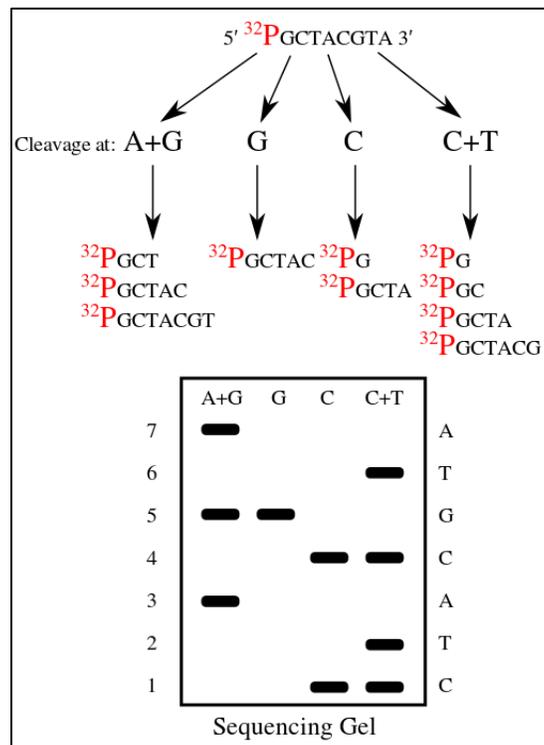


Figure 12 : Méthode de séquençage Maxam-Gilbert (211).

Ces deux méthodes ont été très rapidement appliquées. Le séquençage shotgun, lors duquel le séquençage de clones est suivi par l'assemblage des séquences en s'appuyant sur le chevauchement entre elles, a été décrit par Staden en 1979 (212), et utilisé pour assembler des génomes *de novo* comme le bactériophage lambda dès 1982 (213). En 1987, des machines automatiques basées sur la méthode de fluorescence Sanger ont été mises au point par Smith, Hood et Applied Biosystem, et pouvaient générer 1 000 bases par jour (214). La quantité de données séquencées a augmenté exponentiellement, suivant plus ou moins la loi de Moore et conduisant à la création de bases de données, comme GenBank, et d'outils de recherche, comme BLAST (215), ce qui a encore augmenté la valeur de chaque séquence et mis en place un esprit général de partage des données. En 1982, un demi-million de bases étaient sur GenBank, et en 1986, près de 10 millions.

Le passage au génome humain :

Le projet génome humain (HGP) a vu l'émergence d'une stratégie dite de « clonage hiérarchique », où de larges fragments de génome humain ont été clonés dans des chromosomes bactériens artificiels. L'ADN de chacun de ces chromosomes bactériens artificiels a ensuite été fragmenté, trié par taille et cloné de nouveau. Ces clones individuels ont ensuite été cultivés, et leur ADN a été isolé. L'ADN purifié issu de ces clones a ensuite été séquencé par technique Sanger, le signal issu du scan des gels par laser analysé, et la séquence produite et alignée par chevauchement. Étant donné le nombre d'étapes différentes, et la nécessité du bon fonctionnement de chacune d'elles, la perspective d'un séquençage humain à coût raisonnable était à l'époque inenvisageable.

De ce fait, il est devenu clair que chacune des étapes du processus devait être grandement améliorée pour espérer parvenir à une méthode massivement utilisable, ce qui a pu être obtenu dans les années 1990. Les principales améliorations ont consisté en : un passage des amorces marquées aux terminateurs marqués, ce qui a permis de ne réaliser qu'une seule réaction au lieu des 4 réactions parallèles précédemment nécessaires (216). Puis l'arrivée d'une T7 ADN polymérase mutante a permis d'incorporer plus facilement les terminateurs marqués (217).

Ensuite, les réactions d'amplifications linéaires ont largement réduits la quantité de matériel nécessaire a permis la miniaturisation (218). La purification de l'ADN par billes magnétiques a simplifié la préparation de l'ADN avant séquençage (219). Le séquençage d'ADN double brin a permis d'utiliser des plasmides pour le clonage et l'avènement du séquençage en double-sens (*paired-end*). L'électrophorèse capillaire a supprimé l'usage des gels et simplifié l'interprétation des signaux fluorescents. Enfin le passage à une méthodologie industrielle avec des contrôles qualité, a permis une automatisation, et des procédures standardisées.

La mise en place de protocoles efficaces et standardisés représente seulement la moitié du travail. Il a également été nécessaire de développer toute la partie logicielle, par exemple pour isoler les clones, interpréter les fluorescences, ou assembler les séquences obtenues. L'édition manuelle des séquences lues a été remplacée par la mise en place des *phred* scores, qui ont permis d'utiliser des indicateurs de qualité pour les séquences lues (220). Puis ces séquences pouvaient être assemblées informatiquement en se basant sur leur score de qualité pour obtenir de longs fragments d'ADN. Avec l'arrivée de génomes complexes, les séquences répétées sont devenues un souci très fréquent, et même après un séquençage par *shotgun* très profond, certaines séquences n'étaient toujours pas représentées, ce qui provoquait des discontinuités dans la séquence finale obtenue. Le séquençage *paired-end*, c'est-à-dire le séquençage de chaque fragment à partir de ses deux extrémités, a aidé à combler certaines de ces discontinuités, et celles restantes pouvaient ensuite être comblées par séquençage direct en se servant des séquences alentour comme amorces. Certains problèmes restants ne pouvaient être résolus que par observation et validation direct des séquences.

Le procédé général est ensuite resté relativement stable, mais de nombreuses améliorations dans l'efficacité et la rapidité de chacune des étapes a permis de fortement diminuer le cout au cours des années 1990, en parallèle des avancées informatiques permettant une automatisation de plus en plus poussée du procédé. En 2001, quelques groupes académiques possédaient des lignes de séquençage automatisées produisant près de 10 millions de bases par jour. Les logiciels d'assemblage de génomes progressaient en parallèle, avec par exemple les assembleurs TIGR ou Celera capables de gérer des génomes de plus en plus complexes (221). La progression dans la complétion de génomes de plus en plus larges et complexes surpassait la loi de Moore, avec un

doublent de la taille des génomes complétés tous les ans, avec *Haemophilus influenza*, 2 megabases (Mb) en 1995, puis *Saccharomyces cerevisiae*, 12 Mb en 1996, et *Caenorhabditis elegans*, 100 Mb en 1998 (222). Le projet génome humain (HGP) de 30 fois la taille de celui de *C. elegans* et contenant bien plus de séquences répétées a vu sa première version complétée en 2001, et sa version finale en 2004 (223). Un projet parallèle au projet HGP a été réalisé par Craig Venter et Celera, qui a commencé par une stratégie pilote sur *Drosophila melanogaster*, 175 Mb en 2000 (224).

En 2004, les instruments de séquençage produisaient des séquences de 600 à 700 bases à environ 1 dollars la séquence, et les progrès concernant la technique Sanger sont devenus moins fréquents. La fin du projet génome humain a également rendu flou l'avenir du séquençage.

Le séquençage d'ADN massivement parallèle :

Durant les années 1980 et 1990, plusieurs groupes ont cherché des alternatives au séquençage par électrophorèse. Bien que ces efforts n'aient pas payé à cette époque, à la fin des années 2000, le séquençage « massivement parallèle » ou « séquençage de nouvelle génération », ou encore « NGS », avait pratiquement dépassé le Séquençage Sanger. Les technologies de séquençage NGS se démarquent fortement du séquençage par électrophorèse de plusieurs manières, mais le changement majeur est la capacité de multiplexage. A la place d'un tube par réaction, une librairie complexe de plusieurs échantillons d'ADN est immobilisée sur une surface en deux dimensions, avec tous les échantillons accessibles pour un seul volume réactionnel. A la place d'un clonage bactérien, les échantillons sont amplifiés par une amplification *in vitro* pour générer des copies de chaque échantillon à séquencer. A la place de mesurer la longueur des fragments, le séquençage comprend des cycles biochimiques, comme l'incorporation via une polymérase de nucléotides fluorescents, et de l'imagerie. On désigne également cette méthode sous le nom de séquençage par synthèse (*sequencing by synthesis* ou SBS) (225).

Bien que l'amplification ne soit pas forcément nécessaire (226), le multiplexage extrêmement poussé du NGS, avec des millions à milliards de séquences immobilisées, a été rendu possible par l'amplification *in vitro* de ces fragments. L'approche la plus simple, appelée « amplification en

ponts », implique l'amplification d'une librairie complexe avec des amorces immobilisées sur une surface, de façon à ce que toutes les copies de chaque séquence restent sous la forme d'un cluster séparés des autres (227). C'est l'approche utilisée actuellement par l'ensemble des séquenceurs de la société Illumina (San Diego, USA). Cette amplification peut également être réalisée par émulsion, où chaque fragment est immobilisé sur une bille incluse dans une goutte d'eau entourée d'huile, de façon à ce que chaque fragment soit amplifié séparément de tous les autres, et que chaque bille ne comporte les copies que d'un seul fragment. Ces billes sont ensuite immobilisées sur une surface pour le séquençage (228). Une troisième méthode consiste en une amplification en cercle pour générer des « *nanoballs* » qui sont ensuite fixées sur support et séquencées (229).

Il existe ensuite 3 grandes stratégies de séquençage par synthèse. L'approche du pyroséquençage de Ronaghi et Nyrèn emploie l'ajout étape par étape de chaque deoxyribonucléotide (dNTP). Cet ajout de dNTP relargue un pyrophosphate, qui alimente la génération de lumière par la luciférase de luciole (230). Une seconde stratégie utilise la spécificité de la DNA-ligase pour lier des oligonucléotides fluorescents aux fragments disposés sur un support en fonction de leur séquence (229). Une troisième approche, la plus utilisée à l'heure actuelle, utilise l'incorporation par étape de deoxyribonucléotides fluorescents par une polymérase (231). L'élément nécessaire au SBS passant par une polymérase a été le développement de la terminaison réversible, l'utilisation de dNTP fluorescents compatibles avec la terminaison réversible, et une polymérase permettant l'incorporation par chaque fragment d'une seule base lors de chaque cycle. Puis l'étape d'imagerie permet de déterminer laquelle des 4 couleurs a été incluse par chaque fragment, à la suite de laquelle a lieu la fin du cycle où l'élément bloquant et les éléments fluorescents sont éliminés avant de débiter le cycle suivant (232). Cette approche a été initiée par l'entreprise Solexa en 1998.

Les premières plateformes de séquençage NGS sont arrivées en 2005, avec le reséquençage de *Escherichia coli*, le séquençage *de novo* de *Mycoplasma genitalium*, et le reséquençage du phi174 et d'un BAC (Bacterial Artificial Chromosome) humain. Ces études ont démontré l'utilité des fragments très courts, ou *reads* très courts, utilisés par les technologies de séquençage à haut débit de l'époque, si l'on possède un génome de référence sur lequel les aligner. La plateforme

Solexa a permis en 3 ans le séquençage du génome humain avec des fragments *paired-end* de 35 bases de longueur (233).

En 2005, le premier instrument de NGS a été commercialisé par la société 454. Il a ouvert le séquençage à des laboratoires de taille bien moindre que celle des centres ayant participé au projet génome humain. Cette démocratisation du séquençage a largement modifié le domaine du séquençage, en permettant l'émergence de nouvelles méthodes, nouveaux résultats, nouveaux génomes séquencés, et ce en provenance de nombreux laboratoires (225).

A la différence de l'époque de monopole d'Applied Biosystems sur le séquençage électrophorétique, plusieurs entreprises dont 454 (plus tard acquise par Roche), Solexa (plus tard acquise par Illumina), Agencourt (plus tard acquise par Applied Biosystems), Helicos, Complete Genomics ou Ion Torrent, se sont livrés une importante compétition sur le NGS, qui a permise des changements rapides du domaine avec des nouveaux instruments fréquemment lancés (**Figure 13**). Entre 2007 et 2012, le coût brut par base du séquençage d'ADN a été réduit de 4 ordres de magnitude (234).

Depuis 2012, le rythme des innovations a largement diminué, tout comme la compétition. Plusieurs acteurs n'existent plus et Illumina est maintenant la plateforme dominante. Néanmoins, les progrès ont été considérables depuis les débuts du NGS en 2005. La longueur des *reads* séquencés est maintenant de plusieurs centaines de bases, avec 99,9 % de précision. Plus d'un milliard de fragments indépendants, soit plus d'un terabase de séquences, peut être généré en 2 jours sur un seul instrument (le NovaSeq d'Illumina), le tout pour quelques milliers de dollars. Cela dépasse les 23 milliards de bases générées au cours du projet génome humain d'un facteur 40 (225).

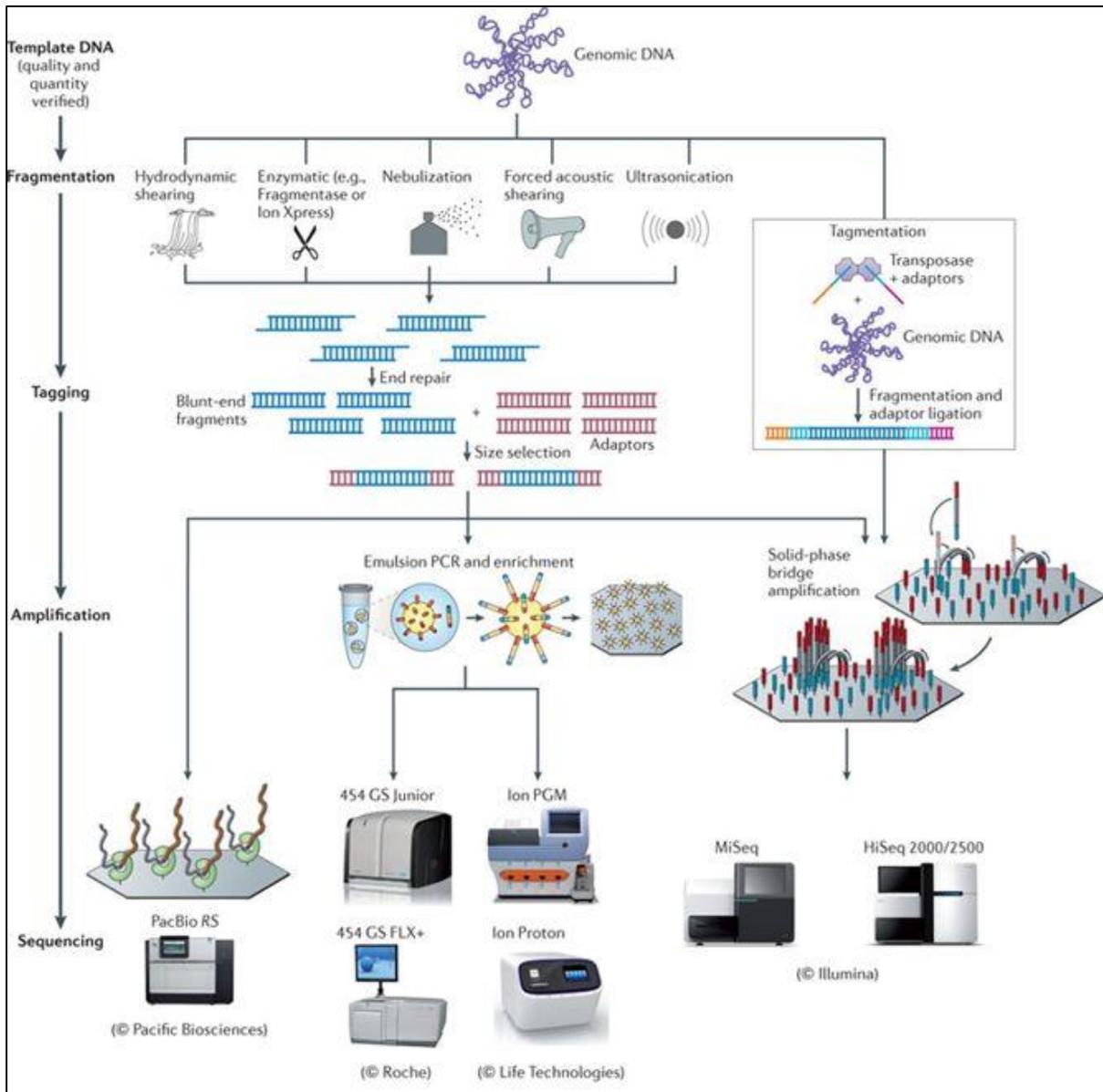


Figure 13 : Panels non exhaustif des technologies de séquençage à haut débit.

Le séquençage en temps réel sans amplification :

Quasiment toutes les plateformes ou instruments mentionnés précédemment nécessitent une amplification des échantillons à séquencer. Et cette amplification conduit à des biais, comme les erreurs de copie, les biais lors de la lecture des séquences, l'information perdue comme la méthylation, en plus du temps et de la complexité engendrée. Idéalement, le séquençage se ferait sur l'échantillon sans traitement préalable, de façon précise, et sans limite de longueur. Pour en

arriver là, plusieurs groupes ont depuis les années 1980 explorés d'autres approches du NGS. Beaucoup se sont avérées être des voies sans issue, mais deux d'entre elles ont conduit à l'arrivée de plateformes de séquençage en temps réel et sans amplification qui bousculent une nouvelle fois le domaine.

La première approche, menée par l'entreprise Pacific Biosciences (Menlo Park, USA) (**Figure 14**), est d'observer de façon optique et en temps réel l'action d'une polymérase sur l'ADN à séquençer (235). Un « *zero mode waveguide* », qui est un trou de diamètre inférieur à la moitié de la longueur d'onde de la lumière observée, permet de limiter la fluorescence à un faible volume d'espace au niveau duquel une seule polymérase se trouve avec l'ADN qu'elle réplique. Cela permet de ne détecter que la fluorescence émise par les nucléotides incorporés par cette polymérase en particulier. La polymérase utilisée permet la lecture de *reads* de 10kb, voire parfois 100kb. Le débit des instruments Pacific Biosciences est toujours plus d'un ordre de magnitude inférieur à celui des meilleurs instruments de NGS, mais proche de ce qu'était ce débit maximum il y a quelques années. Le taux d'erreur est toujours élevé, près de 10%, mais aléatoire dans ses positions. La combinaison des avantages de la technique Pacific Biosciences, c'est-à-dire une grande tolérance des séquences extrêmes comme celles très riches en bases GC, des erreurs distribuées aléatoirement, des *reads* longs et une couverture redondante des régions séquencées permet des assemblages de génome *de novo* de très grande qualité, particulièrement précis et continus.

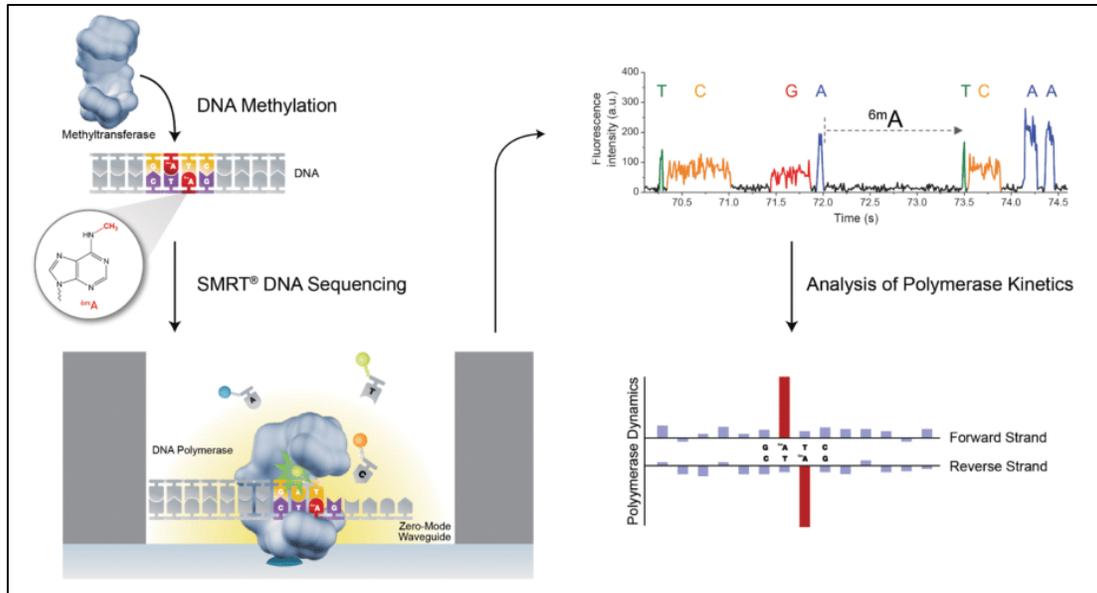


Figure 14 : Technique de séquençage Pacific Biosciences (236).

La deuxième approche est le séquençage Nanopore (**Figure 15**) (237). Ce concept a été théorisé dans les années 1980, et est basé sur l'idée que les schémas des variations dans les flux d'ions qui se produisent lorsqu'un brin unique d'ADN traverse un canal de très faible diamètre permettent de déduire la séquence du brin d'ADN. Plusieurs décennies ont été nécessaires pour passer du concept à la réalité. L'un des premiers problèmes concerne le passage d'un brin d'ADN à travers un pore de dimensions nanométriques, qui est si rapide que la quantité d'ion traversant par le pore en même temps que chaque nucléotide est trop faible pour en déduire un signal. Au cours du temps des solutions ont été trouvées, comme le ralentissement du passage de l'ADN, à l'aide de protéines servants de nucléopores davantage adaptées, et l'analyse plus fine des résultats obtenus (238). Ces avancées ont mené au succès à la fois académique et industriel d'Oxford Nanopore Technologies (Oxford, UK). La longueur de leurs *reads* est du même ordre que celle des *reads* de Pacific Bioscience, voire supérieure, avec un record de 900kb (239). Une des grandes différences entre cette technologie et les autres est la portabilité des systèmes Nanopore qui peuvent être de la taille d'une clé usb, car ils se basent sur une détection de signaux électroniques et non optiques. Des soucis persistent, les erreurs sont plus importantes que pour un séquenceur à *reads* courts classiques et non randomisées, mais la technique progresse encore rapidement.

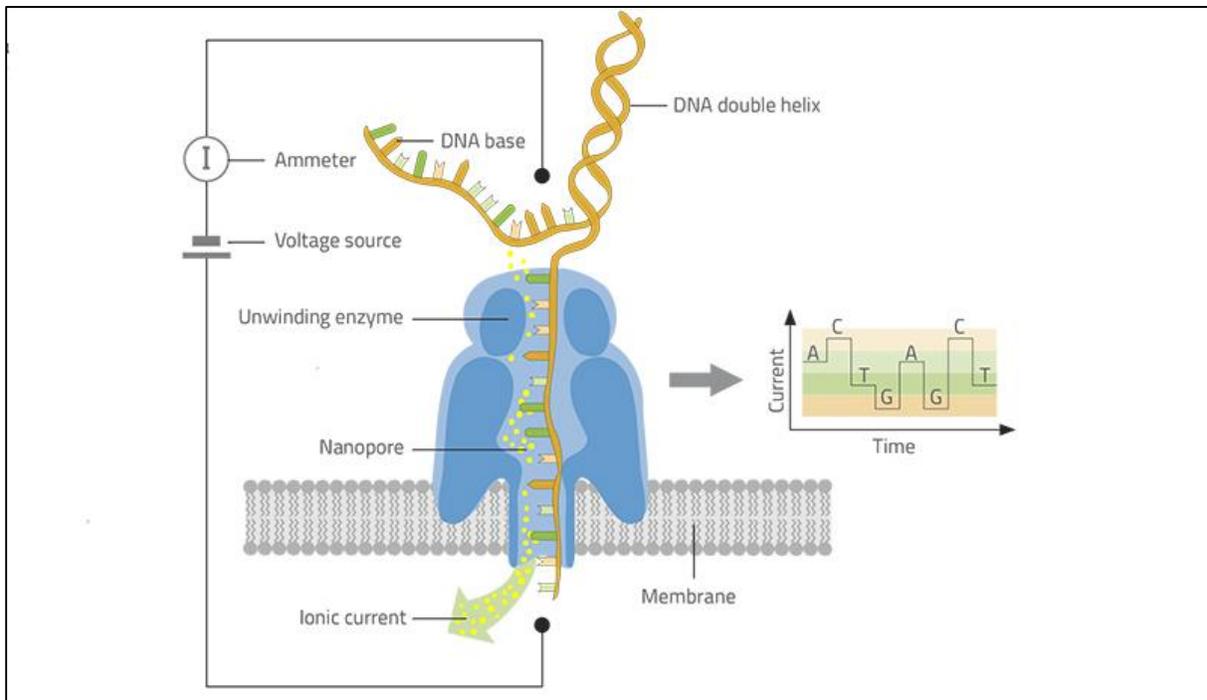


Figure 15 : Technique de séquençage Nanopore (240)

Le séquençage devrait idéalement détecter les modifications de l'ADN, et sur ce point les méthodes de Pacific Biosciences et Oxford Nanopore Technologie sont sans concurrence de la part du séquençage à *reads* courts. Elles permettent en effet la détection des modifications nucléotidiques, comme la méthylation ou l'acétylation des bases séquencées (241). Elles pourraient même permettre de séquencer directement l'ARN, ou même les protéines (242).

Depuis 1977, les technologies de séquençage de l'ADN ont rapidement évolué, et continuent encore de le faire. Bien qu'illumina soit actuellement l'acteur dominant du marché, ce marché s'est agrandi avec de nouvelles technologies actuellement de niches, mais amenées dans le futur à se tailler la part du lion. Toutes ces technologies continueront encore à diminuer en coût et augmenter en débit, et de nouvelles technologies actuellement en développement arriveront pour certaines un jour sur le marché, ce qui modifiera une nouvelle fois le domaine tout entier (225).

Les puces à ADN ou puces CGH :

Les puces à ADN, ou puces CGH (Comparative Genomic Hybridization), permettent de déterminer la présence ou l'absence de fragments d'ADN préalablement connus le long du génome. Leur principe de fonctionnement repose sur l'utilisation de lames de verre au niveau desquelles des fragments d'ADN connus sont accrochés. Ces fragments d'ADN peuvent être en nombre variables (de quelques centaines pour les premiers modèles, à plusieurs millions aujourd'hui) et la position des différents fragments est connue lors de la fabrication de la puce. Lors de leur utilisation, on met en contact la puce avec les fragments d'ADN à étudier, après les avoir préalablement marqués avec un fluorophore à leur extrémité. Ces fragments d'ADN iront ensuite s'hybrider ou non avec l'un des fragments présents sur la puce, suivant la correspondance de leurs séquences. La puce sera ensuite lue en appliquant un laser à sa surface, pour provoquer l'émission de lumière par les fluorophores des différents fragments s'étant hybridés sur la puce. La détermination de l'emplacement et de l'intensité de la lumière émise permettra ensuite de déterminer quelles sont les séquences présentes dans l'ADN étudié, et leurs quantités respectives ((243) et **Figure 16**). Ces puces sont majoritairement utilisées en génétique humaine pour la recherche de SNP connus, ou la recherche de CNV par détermination de la quantité d'ADN de différents fragments du génome.

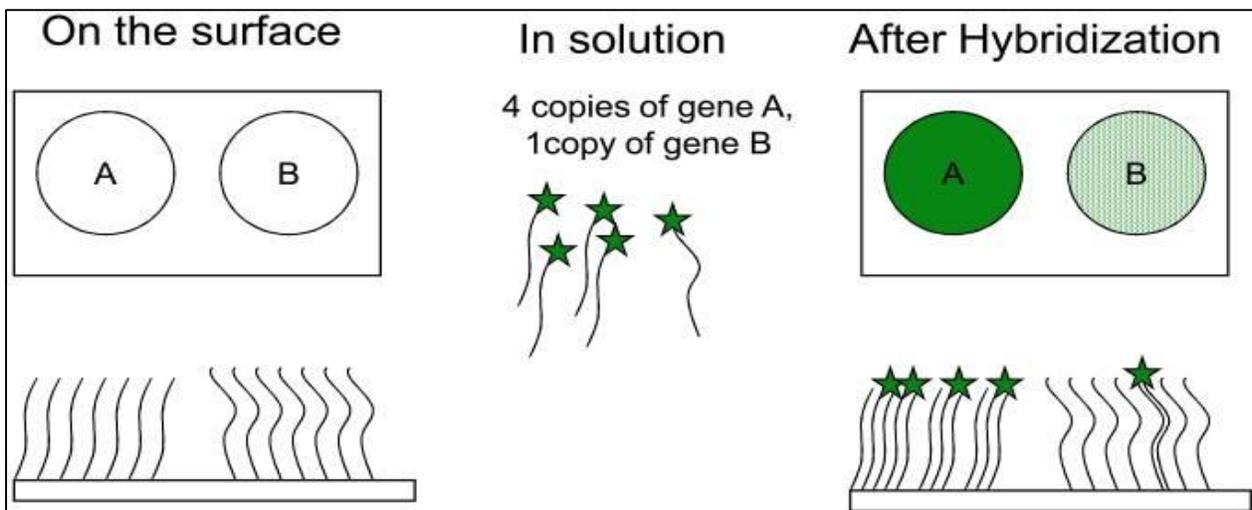


Figure 16 : Principe de fonctionnement des puces à ADN.

Les applications du séquençage de l'ADN :

L'étendue des applications du séquençage de l'ADN s'est également élargie ces dernières années, en suivant les possibilités offertes par les nouvelles techniques de séquençage.

L'assemblage de génomes *de novo* :

Durant ses 25 premières années, le principal objectif du séquençage de l'ADN a été l'élucidation complète ou partielle de génomes. Les vrais débuts du séquençage Sanger ont eu lieu lors du séquençage du premier génome (le phiX174, un génome viral de 5,4kb), qui fut assemblé à la main (244). Mais le séquençage est seulement une des techniques ayant permis l'assemblage de génomes complexes. Si la séquence d'ADN était aléatoire, son assemblage par seul chevauchement de fragments séquencés aurait été possible. Mais l'ADN comportant de larges séquences répétées, et le séquençage étant lui aussi biaisé, il est donc impossible d'assembler des génomes complexes avec le seul chevauchement de *reads* de moins d'un kilobase. Il manque l'information de continuité du génome étudié.

Pour le Projet Génome Humain (245), cette information de continuité provenait de plusieurs sources. Les cartes génétiques établies à partir de la segmentation des polymorphismes génétiques, qui fournit des informations sur l'ordre des séquences de l'échelle du *read* à celle du chromosome. Les cartes physiques ont également été utilisées et établies à partir de BAC clonés, puis analysés par digestion avec des enzymes de restriction pour identifier les chevauchements et assembler une « carte » du génome. Les différents clones ont ensuite été séquencés et assemblés par *shotgun*. Le séquençage *paired-end*, qui séquence chaque fragment d'ADN à partir des deux extrémités a également été très utile en permettant d'obtenir la séquence de fragments plus longs dans leur intégralité (246). Selon la méthode de clonage utilisée, la longueur couverte pouvait varier de quelques dizaines à plusieurs centaines de kb. Une couverture redondante de 8 à 10X de profondeur, associée aux différentes sources d'information de continuité a permis l'assemblage du génome, et ce, avec en moyenne une erreur toutes les 100000 bases. Et enfin des analyses de régions plus précises ont permis de combler les trous restants et d'éliminer les éventuelles ambiguïtés.

En parallèle, l'effort mené par Craig Venter et son équipe, associés à l'équipe de l'assembleur Celera, se sont directement basés sur le séquençage *paired-end*, et en utilisant les cartes physiques du génome comme intermédiaires (224). L'assembleur Celera a également fortement facilité l'alignement des séquences obtenues, en étant plus tolérant et flexible que les aligneurs précédents. Mais à cause de sa tolérance plus élevée, il n'a pas permis d'obtenir un génome d'aussi bonne qualité. Le génome de référence actuel est un descendant de celui du Projet Génome Humain établi en 2004, amélioré avec le temps grâce aux efforts du Genome Reference Consortium, avec de nouvelles versions régulièrement publiées (247).

Avec les avancées du Séquençage de Nouvelle Génération vers l'année 2005, le nombre de génomes assemblés *de novo* a largement augmenté. La combinaison de *reads* courts et de génomes contenant des séquences répétées a pu être surmontée par l'arrivée de nouveaux algorithmes d'assemblage, comme EULER ou Velvet (248). Néanmoins, et surtout concernant les génomes de grande taille, leur qualité était généralement moindre que celle du Projet Génome Humain. Bien que les *reads* de faible longueur soient souvent les coupables désignés, ils ne sont pas la principale raison du manque de qualité. Celle-ci provient essentiellement du manque de techniques complémentaires au NGS pour apporter l'information de continuité de ces génomes. Les bibliothèques d'ADN établies *in vitro* sont moins « riches » que celles réalisées par clonage, et le NGS manque de techniques permettant d'obtenir des équivalents aux cartes génétiques et physiques utilisées lors du Projet Génome Humain.

L'avenir du séquençage *de novo* est néanmoins prometteur. On peut citer l'arrivée des méthodes permettant de séparer des fragments génomiques de grandes tailles pour les séquencer, tout en gardant l'information de liaison entre les fragments séparés, en marquant les fragments à différents endroits avant de les réduire en plus petits fragments séquençables (249). Également les méthodes comme l'Hi-C (*genome wide conformation capture*) qui permettent d'assembler physiquement des séquences jusqu'à l'échelle du chromosome (250). Et enfin les nouvelles méthodes de Pacific Biosciences et Oxford Nanopore Technologies, où la préparation des très long *reads* en vue du séquençage devient davantage problématique que le séquençage lui-même. L'absence d'amplification des fragments séquencés permet une grande qualité d'assemblages *de novo*, même chez des génomes contenant énormément de GC. Ces longs *reads* ont conduit à la

réapparition d'aligneurs spécialisés, comme à l'époque du séquençage électrophorétique, mais amélioré pour gérer les hauts taux d'erreurs et les différentes plateformes disponibles (247). En combinant ces longs *reads*, et une information de continuité à échelle encore plus importante, des assemblages *de novo* de qualité équivalente à celle du Projet Génome Humain de l'ère Sanger deviennent possibles (247).

Le reséquençage de génomes :

A la suite du Projet Génome Humain, l'étape suivante était clairement de référencer toutes les variations génétiques humaines, et donc de reséquençer d'autres personnes. Etant donné le coût élevé du séquençage Sanger, il a principalement été utilisé pour référencer l'emplacement des variants, et ceux-ci étaient ensuite génotypés par puce à ADN, technique bien moins coûteuse. En 2001, alors que le NGS n'existait pour ainsi dire pas encore, a commencé à émerger dans les esprits le but du « génome à 1000\$ », qui permettrait l'ouverture du séquençage massif à la population générale, et qui a ensuite été officialisé en tant qu'objectif par le programme *Revolutionary DNA Sequencing Technologies* lancé par le *National Human Genome Research Institute* (NHGRI). 220 million de dollars ont été mis à la disposition de 40 entités académiques et 27 entités commerciales qui ont permis la plupart des développements technologiques cités précédemment, et également pour supporter directement ou non toutes les plateformes au succès commercial (225).

Le reséquençage, qui consiste donc à aligner des *reads* sur un génome de référence pour identifier les variants génétiques qui le composent, est une tâche très différente de la réalisation d'un assemblage *de novo*. Une couverture redondante de chaque position (c'est-à-dire de nouveaux *reads* alignés sur les mêmes emplacements) est nécessaire pour détecter certains variants hétérozygotes ainsi que pour distinguer les erreurs des variants véritables. Certains logiciels comme SAMtools ou GATK ont adaptés la méthode basée sur le *phred* score pour juger de la confiance en une base aux méthodes de NGS (251). Généralement, les *reads* courts du NGS peuvent être placés à un emplacement unique au niveau du génome, en particulier lors des séquençages *paired-end*. Néanmoins les *reads* courts ne permettent pas de détecter les variants

dans les zones répétées, ainsi que les variants structuraux. Ce phénomène a pu être quantifié grâce aux reséquences effectués par les nouvelles méthodes sans amplification, comme par Pacific Biosciences (252). De même, la conservation des haplotypes, c'est-à-dire de la provenance de chaque copie séquencée, ne peut être réalisée que par des techniques complémentaires aux séquences cités précédemment, dont l'usage se répand de plus en plus (253).

Le génome issu du Projet Génome Humain a été construit à partir de l'ADN de plusieurs individus différents, mais principalement d'un individu de la ville de Buffalo, dans l'état de New York, issu en proportions environ égales d'ancêtres Européens et Américains. Le premier individu à avoir son propre génome séquencé entièrement était Craig Venter, en 2007, qui participait au projet de génome Celera (224). Puis l'année suivante Jim Watson par 454 Technologies, puis 2 individus anonymes et des cellules germinales et tumorales sur technologie Solexa / Illumina (233,254), et enfin 5 individus par Complete Genomics (229). Cette époque a aussi été marquée par le développement des séquences ciblées par une capture, pour trouver une alternative au séquençage de génome trop cher pour la plupart des projets, avec les premiers séquences d'exome ne ciblant que les 1- 2% du génome codant des protéines (135).

Avec la diminution des coûts de séquençage et le rapprochement de l'objectif des 1000\$ par génome, de plus en plus d'individus ont été séquences, par génome ou exome. Le projet 1000 Génomes, lancé en 2008, a permis de séquencer le génome de quelques centaines d'individus avec une faible couverture en 2010, puis de quelques milliers en 2015 (255). L'*Exome Sequencing Project* a lui permis de séquencer 6 500 exomes en 2013 (256). La *Genomes Agregation Database* regroupe aujourd'hui près de 120 000 exomes et 15 000 génomes. Et les projets *Genomics England*, *GenomeAsia 100K* et *NHLBI TOPMed* visent chacun le séquençage de 100 000 personnes sur les 2 ans à venir. Etant donné que ces grands projets ne représentent qu'une fraction de l'ensemble des séquences effectués, il est probable que plus d'un million d'individus aient déjà vu leur exome ou leur génome séquencé (225).

Les applications cliniques du séquençage :

Notre capacité à séquencer des génomes humains a largement dépassé notre capacité à en interpréter les variations, ce qui explique en partie pourquoi la médecine ne s'est pas encore pleinement appropriée le séquençage de génome. Mais il existe néanmoins des domaines dans lesquels le séquençage a prouvé son utilité clinique.

Une des aires de cette utilité clinique du séquençage concerne les tests prénataux non-invasifs (NIPT). Il a été montré en 2008 que le comptage de fragments d'ADN libérés par le fœtus dans le sang maternel permet de détecter des aneuploïdies chromosomiques (257). Des tests diagnostiques basés sur cette stratégie ont été très rapidement adoptés, et près d'un demi-million de femmes enceintes ont bénéficié d'un séquençage de génome à faible profondeur pour les NIPT.

Une des premières applications cliniques du séquençage d'exome a été de découvrir de nouveaux gènes pour des maladies Mendéliennes, et de réaliser le diagnostic des patients étant touchés par celles-ci (135). Le séquençage d'exome a également permis de découvrir qu'une bonne partie des maladies neuro-développementales pouvaient être attribuées à des mutations *de novo* dans des séquences codantes (258). Le séquençage d'exome est de plus en plus utilisé en tant qu'outil diagnostique, en particulier chez les enfants, avec un taux de diagnostic par séquençage d'exome dans le cadre d'une suspicion de maladie mendélienne atteignant 25%, et en augmentation (259). La compréhension du cancer a également été transformée par le séquençage de l'ADN. Le reséquençage massif du génome humain a mis en évidence la diversité génétique des cancers (260). Il a permis de changer la prise en charge clinique des cancers en rendant possible les thérapies ciblées selon le patrimoine génétique du cancer, en permettant les diagnostics non invasifs grâce au séquençage des cellules cancéreuses ou à l'ADN tumoral relâché dans la circulation sanguine, ou encore en permettant l'identification de mutations affectant certaines protéines des cellules cancéreuses, ce qui permet de mettre au point des « vaccins personnalisés ». Beaucoup de progrès sont encore nécessaires, mais la prise en charge du cancer est en pleine transformation avec les évolutions du séquençage.

Méthodes de détection des copy number variants :

La première méthode employée pour réaliser la détection des CNVs est la méthode FISH (*Fluorescent In Situ Hybridization*) (261). Par la suite, au début des années 2000, les puces CGH (*Comparative Genomic Hybridization*) ont commencé à être employées pour cet usage, et sont encore aujourd'hui largement utilisées en tant qu'outils diagnostic (**Figure 16** et (262), malgré des limitations liées à la résolution ainsi qu'à la couverture du génome obtenue, leur impossibilité de détecter les mutations non connues, ainsi que leur incapacité à détecter les translocations.

Mais le développement du séquençage de nouvelle génération, ou NGS, a largement modifié le paysage du séquençage de l'ADN. L'arrivée de séquenceurs à débit toujours plus élevé a permis de diminuer les coûts de ces analyses, et d'augmenter la résolution disponible, jusqu'à l'arrivée du séquençage d'un génome humain complet en quelques jours pour un coût relativement raisonnable. Ces avancées ont considérablement modifié les approches disponibles pour l'analyse des mutations ponctuelles, mais également des CNV, avec l'apparition de nombreuses méthodes et outils bioinformatiques dédiés à leur détection à partir de donnée de séquençage à haut débit. La détection des CNV par séquençage nécessite des approches particulières, différentes de la détection des mutations qui s'intéresse uniquement à la séquence des *reads*.

Il existe à l'heure actuelle plusieurs stratégies pour procéder à la détection de CNV à partir de données NGS : celle basée sur l'alignement des paires de *reads* (*paired-end mapping*, PEM), celle utilisant les *reads* tronqués (*split reads*, SR), celle utilisant la profondeur des *reads* (*reads depth*, RD), celle de l'assemblage *de-novo* (AS), ainsi que les méthodes combinant plusieurs de ces approches (263).

Méthode de l'alignement des paires de *reads* :

Cette méthode est la première à avoir été développée pour l'analyse des CNVs à partir de données de séquençage (264). Elle ne fonctionne qu'avec des *reads paired-end*, et se base sur la distance entre les deux *reads* d'une même paire lors de l'alignement (**Figure 17**). Lors du séquençage *paired-end* de fragment d'ADN d'une taille donnée, la distance attendue entre le début de chacun des *reads* d'une paire doit correspondre à la taille de ce fragment. Si cette distance entre les deux

reads d'une paire est plus ou moins importante que celle attendue, on en déduit la présence d'un CNV de type duplication ou délétion. Cette méthode permet également d'identifier des évènements structuraux comme les insertions, ou inversions, mais cela ne fonctionne que si les points de cassures sont présents au niveau d'un des *reads*. De la même façon, la détection des duplication ou délétions ne fonctionne que si le début et la fin de l'évènement sont présent dans les fragments séquencés (265).

Méthode des *reads* tronqués :

Cette méthode utilise des paires de *reads* dans lesquelles un des reads s'aligne correctement sur le génome, et l'autre ne s'aligne pas correctement ou seulement partiellement. Ces reads partiellement alignés peuvent apporter le point précis de cassure d'un CNV ou d'un variant structural. Ils sont ensuite découpés et réalignés plus finement pour détecter avec précision le point de cassure (**Figure 17** et (263)). Cette méthode comporte néanmoins le même inconvénient que la méthode précédente, à savoir que le point de cassure doit se situer au niveau des fragments séquencés, et nécessite également des *reads* de longueur importante pour maximiser les chances d'y inclure le point de cassure.

Méthode de la profondeur des *reads* :

L'utilisation de la profondeur des *reads* pour déterminer la présence de CNV se base sur le principe voulant que la profondeur de couverture d'une région du génome est liée à la quantité de matériel génétique correspondant à cette région. Cette méthode permet théoriquement de déterminer le nombre exact de copies, qui sera lié à la profondeur de couverture, alors que les autres méthodes ne le peuvent pas car elles ne prennent en compte que les informations de position des reads (**Figure 17**). Cette approche peut être utilisée en combinaison avec du séquençage de génome complet (WGS) ou avec une approche séquençage d'exome (WES). Le WGS permettra une détection des CNV sur l'ensemble du génome avec une excellente précision concernant les points de cassure des évènements détectés, mais à un prix plus élevé que le WES,

qui ne séquence que les régions codantes et est moins précis sur l'emplacement des points de cassure.

La détermination des CNV par la méthode de profondeur des *reads* peut être appliquée avec plusieurs configurations. Lors de l'application à un seul échantillon, la stratégie appliquée consiste à calculer une profondeur moyenne de couverture, et identifier les régions s'en écartant significativement. Dans le cas d'une étude avec deux échantillons appariés, comme la comparaison d'un tissu sain à un tissu tumoral, l'échantillon sain peut servir de référence concernant la profondeur de séquençage pour la détection des CNVs. Et dans le cas d'une étude comportant de nombreux échantillons, ceux-ci peuvent être utilisés pour former une ligne de base des profondeurs de couverture qui sera comparée aux échantillons analysés (263).

Méthode de l'assemblage *de-novo* :

A la différence des autres méthodes s'appuyant sur une référence préexistante pour procéder à l'alignement des *reads* séquencés et ensuite en déduire la présence des CNVs, la méthode de l'assemblage *de novo* procède à l'assemblage des *reads* séquencés sans référence préexistante. L'assemblage réalisé est ensuite comparé à une référence pour réaliser la détection des CNVs. Cette méthode fonctionne avec une faible profondeur de couverture pour l'assemblage et donc la détection des CNVs, mais nécessite une couverture des points de cassures de ces CNVs (**Figure 17**).

Il est également possible d'utiliser une combinaison de plusieurs méthodes, de façons à combler les faiblesses de celles-ci (263).

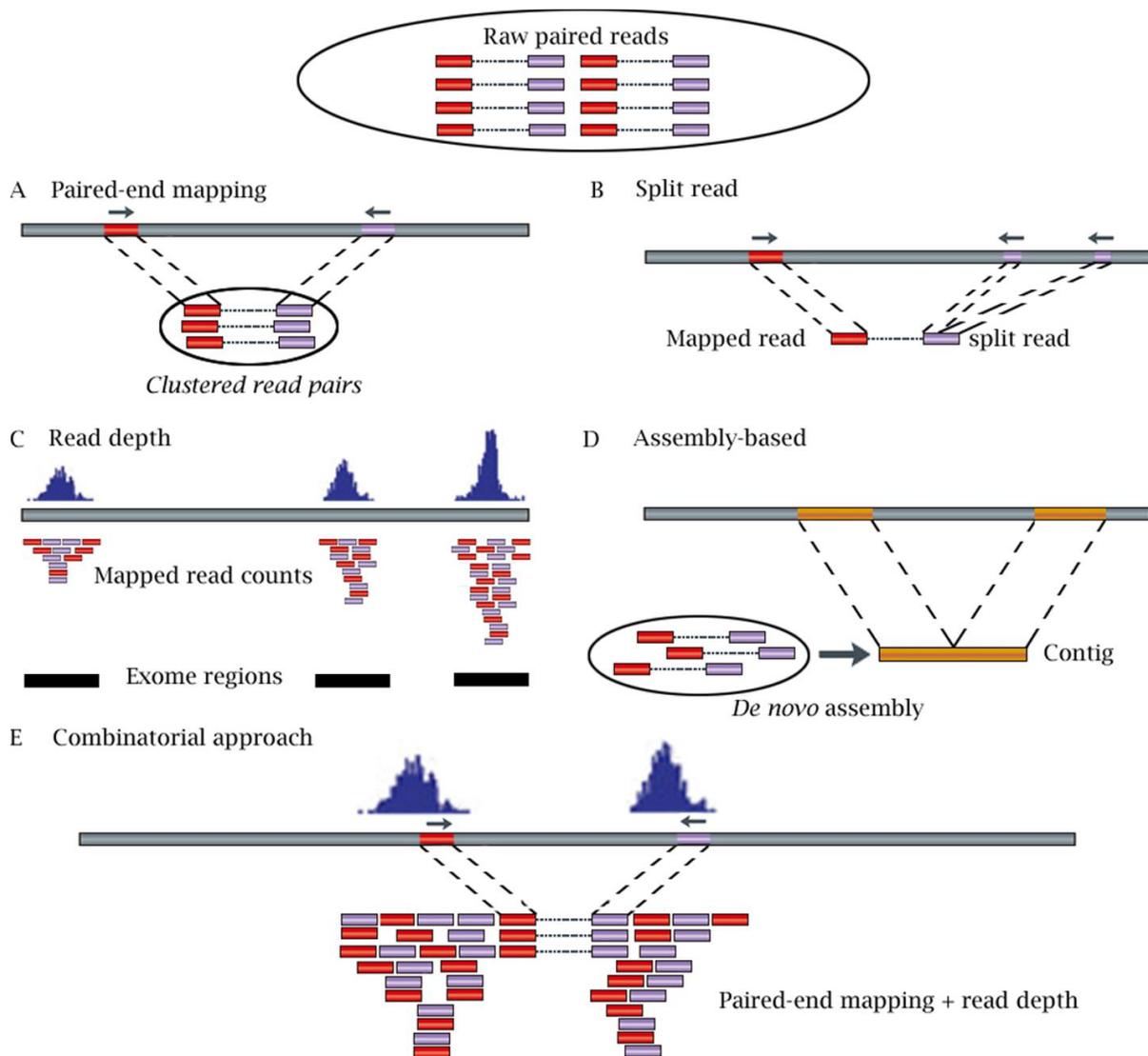


Figure 17 : Méthode de détections de CNV par séquençage de reads courts (263).

La détection des CNVs par le séquençage de longs fragments :

Les nouvelles technologies de séquençage permettant de séquencer de longs fragments sans amplification préalable sont de plus en plus utilisées, en particulier pour la détections de variant structuraux comme les CNVs grâce aux nombreux avantages qu'elles procurent. Ces technologies, représentées par Oxford Nanopore et Pacific Bioscience, permettent de séquencer de longs fragments de plusieurs dizaines voire même centaines de kb. Ces très longs fragments permettent ensuite par alignement direct de détecter la présence de CNVs ou de variants structuraux comme les inversions par alignement sur une référence, de façon similaire à la méthode des *reads*

tronqués. Mais ces méthodes permettent de détecter des évènements de taille bien plus importante, et sans limitation concernant les zones du génome où ces évènements ont lieu car elles ne nécessitent pas d'amplification (266). Néanmoins le taux d'erreur de séquençage encore élevé (près de 10%) reste un frein pour leur adoption massive et pour la détermination précise des points de cassure des évènements détectés (267).

Le futur du séquençage de l'ADN :

Déterminer la diversité du génome :

Un génome complet, incluant les télomères, centromères, tous les chromosomes sans discontinuité ou ambiguïté n'a jusqu'ici été accompli que pour un seul organisme eucaryote (268). Avec les évolutions technologiques, ces barrières seront certainement franchies dans le futur. De la même façon les millions d'espèces découvertes ou encore à découvrir sur Terre seront vraisemblablement un jour séquencées, de façon à obtenir une vue globale de la génétique du monde du vivant, et apporter des informations sur les mécanismes de son fonctionnement, comme par exemple l'élaboration des protéines (269).

Séquençage de populations :

Dans peu de temps, 0,1% de la population terrestre aura bénéficié du séquençage de son ADN, à un degré plus ou moins poussé. Le nombre de mutations *de novo* référencées dépasse le nombre de nucléotides du génome humain. L'agrégation de millions de génomes différents pourrait permettre d'établir une carte de tous les génotypes compatibles avec la vie, et donc du génome humain dans son ensemble (270).

Biologie développementale :

Bien que nous sachions que nous nous développons à partir d'une seule cellule, la compréhension des mécanismes impliqués dans ce développement reste très limitée. Le séquençage de cellules

uniques existe déjà, mais l'idéal serait en plus d'en conserver le contexte, de pouvoir réaliser le séquençage de l'ADN, de l'ARN, ou des protéines d'une cellule *in situ* (271). La mise en place de « codes-barres » cellulaires transportables lors des échanges intercellulaires permettrait de cartographier ces échanges (272).

Séquenceurs temps-réel portables :

Les séquenceurs Minlon de Oxford Nanopore Technologies pèsent 70g et permettent de produire des données en 30min. La prochaine étape serait la dissémination de cette technologie, pour mesurer l'environnement et la vie humaine en général, comme l'air, la nourriture ou nos propres corps, et en partageant ces données à travers un immense réseau.

Autres usages :

Le séquençage de nouvelle génération peut également servir à d'autres applications, comme le stockage de données (273), ou la nano-ingénierie, en procurant de nouvelles manières de stocker l'information, ou de concevoir des machines.

Le séquençage en est encore à ses débuts, et le rythme des évolutions qu'il subit et procure permet d'entrevoir la place qu'il sera amené à prendre dans le monde et dans nos vies, comme celle que peuvent avoir une technologie comme la microscopie.

Mon travail de thèse :

La génétique, et par extension le séquençage de l'ADN, sont des outils qui ont transformé la compréhension des mécanismes impliqués dans la survenue de nombreuses pathologies, dont l'obésité. Les technologies aujourd'hui à notre disposition nous permettent de déterminer rapidement si un patient est ou non porteur d'un évènement génétique pouvant expliquer sa pathologie. L'une des techniques les plus utilisées en diagnostic aujourd'hui est le séquençage d'exome, ou WES, qui permet une excellente détection des mutations ponctuelles dans les régions codantes du génome. Mais d'autres évènements comme les *copy number variants*, ou CNV, peuvent également expliquer certaines pathologies, dont l'obésité, via entre autres les CNV de la région 16p11.2. Actuellement, la technique de référence pour la détection de ces *copy number variants* est l'analyse de puces CGH (*Comparative Genomic Hybridization*), mais celles-ci ne permettent pas de détecter des mutations non répertoriées au préalable lors de la création de la puce. Sur le principe, le séquençage d'exome peut lui aussi être utilisé pour détecter les CNV, mais son absence de couverture des régions non codantes du génome ne permet pas une détection efficace de ces CNV, car ceux-ci peuvent survenir sur l'ensemble du génome, en englobant des régions codantes et non codantes au sein d'un seul évènement. Le séquençage génome complet peut détecter ces deux types d'évènement, mais son cout est encore élevé ce qui freine sa démocratisation, et l'analyse de données associées nécessite d'importantes ressources informatiques, et le rend difficilement utilisable en diagnostic de routine en l'état actuel des choses.

Il est donc pour l'instant nécessaire d'avoir recours à deux techniques différentes pour couvrir ces deux types d'évènements génétiques. Cela implique d'utiliser des échantillons parfois très précieux à deux reprises, de supporter les couts liés à deux techniques diagnostiques (d'environ 450 euros pour le séquençage d'exome au laboratoire et un cout un peu plus élevé pour une puce à ADN dans un laboratoire clinique), et d'allonger les temps de rendu de résultats et donc la durée d'établissement du diagnostic du patient.

Cet état de fait nous a conduit à développer une technique de séquençage, que nous avons nommé CoDE-seq (*Copy number variation Detection and Exome sequencing*), et qui permettra la

détection simultanée de ces deux types d'évènements, pour diminuer les temps d'établissement de diagnostics, leurs couts, et la quantité d'échantillon nécessaire.

Ce travail a nécessité deux aspects : la mise au point technique et la mise au point analytique.

La mise au point technique est passée par la création d'une nouvelle « capture », permettant une détection correcte des mutations ponctuelles de l'exome et des CNV de tout le génome.

La mise au point analytique a consisté à définir la méthode à employer, et à permettre d'arriver à une détection fiable, à la fois sensible et spécifique, des CNV sur l'ensemble du génome. Une fois ces CNV identifiés, la question de leur signification fonctionnelle se pose également, et une seconde partie de ma thèse porte sur l'étude de cette signification fonctionnelle, via l'étude de la conformation spatiale de la chromatine et de l'influence des CNV sur celle-ci

Matériel et méthodes :

Echantillons et cohortes utilisées :

Dans le but de mettre au point cette technique, nous avons choisi de réaliser en premier lieu une étude preuve de concept, pour démontrer que la technique est fiable et utilisable en lieu et place des puces CGH associées au séquençage d'exome. Pour ce faire, nous avons eu accès à l'ADN d'une cohorte de 40 enfants et jeunes adultes déficients intellectuels sans diagnostic préalable en provenance du Centre de Génétique Chromosomique de l'hôpital Saint Vincent de Paul, à Lille, dont la particularité est qu'elle a déjà subi une recherche de CNV par puces CGH, mais aucune analyse concernant leurs éventuelles mutations. L'accès à cette cohorte a été rendu possible grâce à notre collaboration avec le Professeur Bruno Delobel, ainsi qu'avec le docteur Louise Montagne. Cette cohorte préalablement analysée nous a permis d'étalonner notre technique, en comparant nos résultats avec ceux obtenus par les puces CGH. Elle se compose de 24 hommes et 16 femmes, ayant eu accès à la puce CGH Agilent 60K v2 en raison de leurs troubles neurologiques : 13 d'entre eux souffrent de déficience intellectuelle, 12 ont des capacités d'apprentissage limitées, 5 souffrent à la fois de déficience intellectuelle et de troubles du comportement, et 3 à la fois de troubles du comportement et de capacités d'apprentissage limitées.

Une fois la technique en place, nous l'avons appliquée à une cohorte de 30 enfants obèses sans diagnostic préalable (15 filles et 15 garçons) provenant de la cohorte *Obesity Saint Vincent (OSV)* de l'hôpital Saint Vincent de Paul, à Lille. L'accès à cette cohorte a été rendu possible grâce à notre collaboration avec le Professeur Bruno Delobel, ainsi qu'avec le docteur Louise Montagne. Cette cohorte est composée d'enfants obèses avec un IMC supérieur à 30, recrutés entre 2015 et 2016, et qui continuera à être alimentée jusqu'en 2020. Tous les patients la composant ont été inclus en raison de leur obésité syndromique (c'est-à-dire associée par exemple à une déficience intellectuelle, des troubles alimentaires, des troubles du comportement, un handicap physique ou visuel, une ou plusieurs malformations, des troubles du sommeil, ou un hypogonadisme), de

leur obésité isolée (si par exemple leurs deux parents ne sont pas obèses), ou de leur obésité consanguine. Des données cliniques ont été obtenues pour tous ces enfants, incluant les éventuels évènements *in utero*, les paramètres à la naissance, leur développement psychomoteur, leurs éventuels désordres neurologiques, l'âge d'apparition de l'obésité, leurs paramètres métaboliques, leurs habitudes alimentaires, les éventuels traitements, leur activité physique, leur niveau scolaire et leur historique familial. Trente-trois des parents de ces enfants ont également été analysés dans le but d'établir si les évènements détectés chez leurs enfants sont *de novo* ou non (**Supplementary Table B**).

Egalement 12 enfants ou jeunes adultes, 7 garçons et 5 filles sans diagnostic préalable, ont été adressés au laboratoire dans le but d'obtenir un diagnostic via cette nouvelle méthode. Ces individus souffrent d'une obésité sévère et/ou d'une déficience intellectuelle. Les parents d'un de ces enfants ont également été analysés dans le but d'établir si les évènements détectés chez leur enfant sont *de novo* ou non (**Supplementary Table B**).

Conception de la méthode :

Notre problématique principale était de parvenir à détecter les éventuels CNV d'un patient avec une qualité au moins équivalente à une détection réalisée dans un laboratoire clinique via l'utilisation de puces CGH, tout en obtenant une détection des mutations ponctuelles dans les régions codantes aussi performante qu'avec le séquençage d'exome entier. Pour ce faire, nous avons choisi de partir d'une solution de séquençage d'exome entier classique, et de l'améliorer pour combler ses lacunes dans la détection de CNV.

Capture :

Le séquençage d'exome (WES) se différencie du séquençage de génome (WGS) car il ne cible que 2 à 3% du génome, en l'occurrence les régions codantes. Cette approche ciblée est permise par l'ajout d'une étape dans le protocole de WES par rapport à celui du WGS : la capture.

La capture consiste en la sélection des fragments d'ADN de la librairie d'échantillons qui sont ensuite séquencés. Le principe général repose sur l'utilisation de sondes complémentaires des régions du génome que l'on souhaite conserver, et donc séquencer. Ces sondes sont couplées à un élément permettant de les conserver lors des étapes de lavages pour éliminer les fragments d'ADN non désirés (par exemple un couplage avec de la biotine, qui se liera ensuite à un élément fixe de streptavidine). Lors de leur ajout à la librairie, ces sondes s'hybrident aux fragments qui leur correspondent au sein de la librairie. Plusieurs lavages permettent ensuite de ne conserver que les sondes et les fragments qui y sont associés. Les fragments sont ensuite détachés des sondes et les sondes sont à leur tour éliminées. On obtient alors un pool des fragments sélectionnés, ce qui permet un séquençage ciblé.

Le principal défaut du WES pour la détection de CNV concerne sa capture, qui ne permet pas une détection uniforme des CNV sur l'ensemble du génome. En effet les zones codantes sont parfois séparées de grands espaces non couverts par le WES. Or la détection de CNV nécessite une couverture régulière du génome, pour pouvoir les détecter dans leur ensemble et être le plus précis possible sur leurs dimensions.

Nous avons donc cherché à améliorer cette capture WES, en y ajoutant de nouvelles sondes. Pour ce faire, nous sommes partis de la capture WES existante *SeqCap EZ MedExome Enrichment*, commercialisée par NimbleGen (Madison, USA), qui comme son nom l'indique est spécifique de l'exome. Nous avons ensuite utilisé la possibilité offerte par le fabricant d'ajouter des sondes à une capture existante, via leur capture *NimbleGen SeqCap EZ ChoiceXL*. Il nous a donc fallu définir comment ajouter de nouvelles sondes à cette capture d'exome classique.

Design de la capture :

Pour améliorer cette capture de WES, l'objectif était de combler les espaces entre les sondes existantes. Il a donc fallu définir une stratégie d'ajout de ces sondes, en tenant compte de plusieurs paramètres : l'espace sans sonde maximum recherché pour obtenir une détection correcte, le nombre maximum de sondes pouvant être ajouté, et le surcout (de séquençage) engendré par l'ajout de ces sondes.

L'objectif général étant de développer une technique pouvant concurrencer le WES ou les puces CGH, il est nécessaire que le cout par échantillon soit d'ordre comparable à ces deux techniques. L'ajout d'un nombre trop important de sondes risquait de faire grimper ces couts (actuellement d'environ 450 euros pour un WES au laboratoire), en particulier au niveau du séquençage, et de tendre vers le WGS, ce qui rendrait la technique bien plus difficile à diffuser. La capture WES classique séquence un total d'environ 65Mb (mega bases) sur les 3Gb (giga bases) du génome. Nous avons choisi de ne pas augmenter les coûts de plus de moitié, de façon à conserver un tarif compatible avec une utilisation diagnostique, et donc de nous limiter à environ 90Mb pour le total de la nouvelle capture exome « enrichie ». Les sondes à ajouter recommandées par le fabricant couvrent en moyenne 120 bases par sonde. Ces paramètres nous ont permis de déterminer qu'il était possible de placer suffisamment de sondes de façon à ne pas obtenir d'espace entre deux sondes supérieures à 20kb. Nous avons ensuite affiné cette stratégie en divisant le génome en deux catégories de priorités différentes : une partie du génome consistant en des « zones d'intérêt », au niveau desquels des CNV ont été répertoriés comme en rapport avec différentes pathologies, et au niveau desquelles la densité des sondes recherchés doit être importante ((274) et **Supplementary Table 1**). Le total de ces zones d'intérêt couvre entre 1/4 et 1/3 du génome. Et une partie du génome consistant en des zones de priorité plus faible au niveau desquelles la densité de sondes peut être moins importante, c'est-à-dire tout ce qui ne rentre pas dans la première catégorie. Cette stratégie affinée inclue des sondes espacées au maximum de 10kb dans les zones d'intérêt, et espacées de 25kb dans le reste du génome.

Pour répartir ces nouvelles sondes, nous nous sommes appuyés sur le fichier BED (*browser extensible data*) de la capture WES classique, qui reprend les positions génomiques de toutes les sondes de cette capture. Nous avons relevé tous les espaces entre ces sondes, à la fois dans les zones d'intérêt que nous priorisons, et dans le reste du génome. Tous ceux de plus de 10kb dans les régions d'intérêt se sont vu attribuer des sondes à intervalle de 10kb, et tous ceux de plus de 25kb dans le reste du génome se sont vu attribuer des sondes à intervalle de 25kb (**Figure 18**).

Une fois la capture théorique conçue, celle-ci a dû faire l'objet d'une validation de la part de l'équipe de designers de NimbleGen, en tenant compte de leurs contraintes, comme la

compatibilité avec la capture WES existante, et la possibilité de placer ou non des sondes dans les différentes régions du génome (zones répétées, trop riches en GC, ...).

Après de nombreux échanges, une solution satisfaisante a été trouvée, qui consiste en l'addition de la capture WES classique (soit 226 072 sondes) et d'une capture contenant toutes les sondes supplémentaires nécessaires à la détection correcte de CNV (soit 155 563 sondes), pour un total d'environ 91,5 Mb (soit 381 635 sondes).

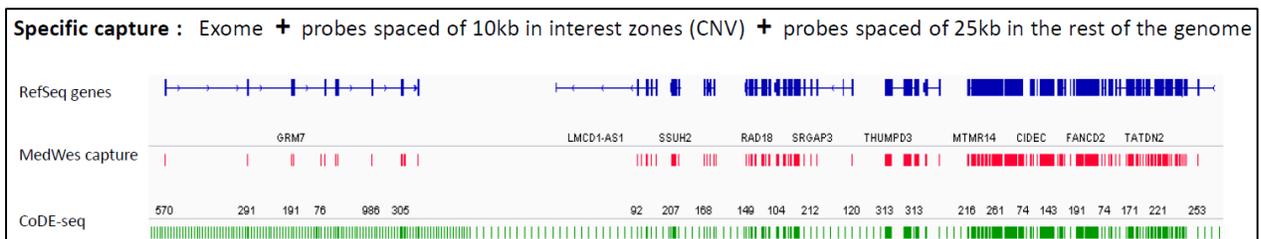


Figure 18 : Représentation schématique des sondes de notre nouvelle capture CoDE-seq par rapport à celles de la capture WES classique *NimbleGen SeqCap EZ MedExome Enrichment*, au niveau d'une portion du chromosome 3. On peut voir que la capture CoDE-seq reprend l'ensemble des sondes de la capture exonique classique, et y ajoute des sondes à intervalle moyen de 10kb dans les zones d'intérêt pour la survenue de CNV (partie gauche de l'image) et tous les 25kb en moyenne au niveau du reste du génome (centre de l'image).

Librairie :

Nous avons utilisé 1µg d'ADN de chacun des patients du projet, qui nous a été soit fourni directement sous forme d'ADN extrait, soit sous forme de sang dont nous avons extrait l'ADN. Ces ADN ont été fragmentés à l'aide d'un appareil de sonication Covaris E220 de la société Covaris (Woburn, USA) présent au laboratoire, pour obtenir des fragments de taille compatible avec le séquençage d'environ 200 bases. Les fragments obtenus ont ensuite subi une réparation les extrémités soniquées, pour permettre ensuite leur ligation aux adaptateurs de séquençage Illumina grâce au kit HTP *library preparation kit* de la société KAPA (Bâle, Suisse). Cette étape est nécessaire à la fixation des fragments à séquençer sur la *flowcell*, ou puce de séquençage, et a été effectuée sur un automate Hamilton STARlet de la société Hamilton (Reno, USA). Les échantillons ont ensuite été amplifiés par *polymerase chain reaction*, ou PCR, en utilisant des

amorces complémentaires des adaptateurs de séquençage. Les échantillons amplifiés ont ensuite été purifiés lors d'une étape de sélection de taille pour ne garder que ceux compatibles avec le séquençage et éliminer les éléments parasites, puis quantifiés par l'appareil LabChip GX de Perkin Elmer (Wellesley, USA) (**Figure 19**).

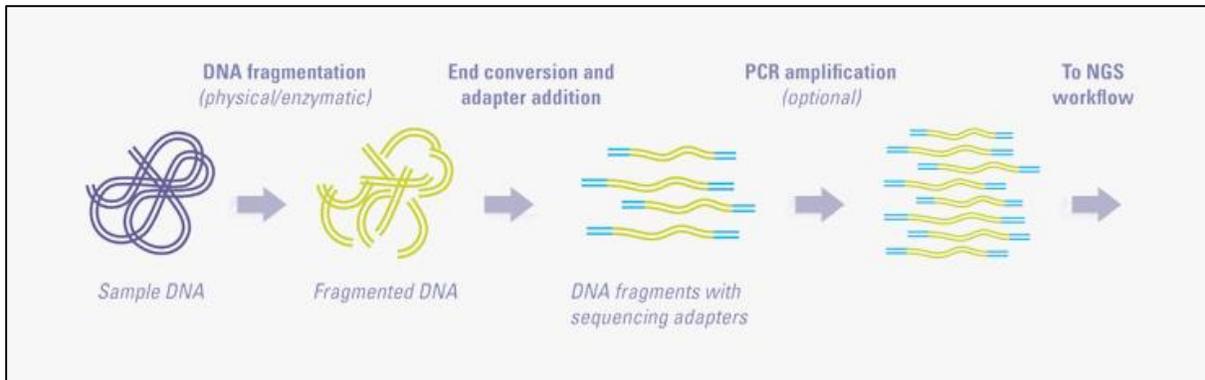


Figure 19 : Schéma des étapes de la préparation d'une bibliothèque d'ADN pour séquençage de nouvelle génération (275).

Enrichissement :

A la suite de ces étapes, les échantillons ont été répartis en pools de 5 échantillons constitués de 1µg d'ADN par pool pour procéder à l'étape d'enrichissement. Lors de cette étape, les pools d'échantillons ont été mis en présence des sondes biotinylées correspondant à la fusion de la capture WES classique et des sondes supplémentaires de détection de CNV pendant 72h à 47°C. Les pools d'échantillons capturés ont ensuite été purifiés grâce aux kits de lavage *SeqCap Hybridization* et *Wash Kit* de la compagnie Roche (Bâle, Suisse) sur l'automate *Bravo Automated Liquid Handling Platform* de la société Agilent (Santa Clara, USA), de façon à éliminer tous les fragments non capturés. Ces captures ont ensuite été amplifiées grâce au kit *Hifi HotStart ReadyMix* de la société Nimblegen et quantifiées sur les appareils Caliper LifeScience Labchip GX de la société Perkin Elmer (Waltham, USA) et Qubit fluorometric quantitation assays de la société Thermo Fisher Scientific (Waltham, USA) (**Figure 20**).

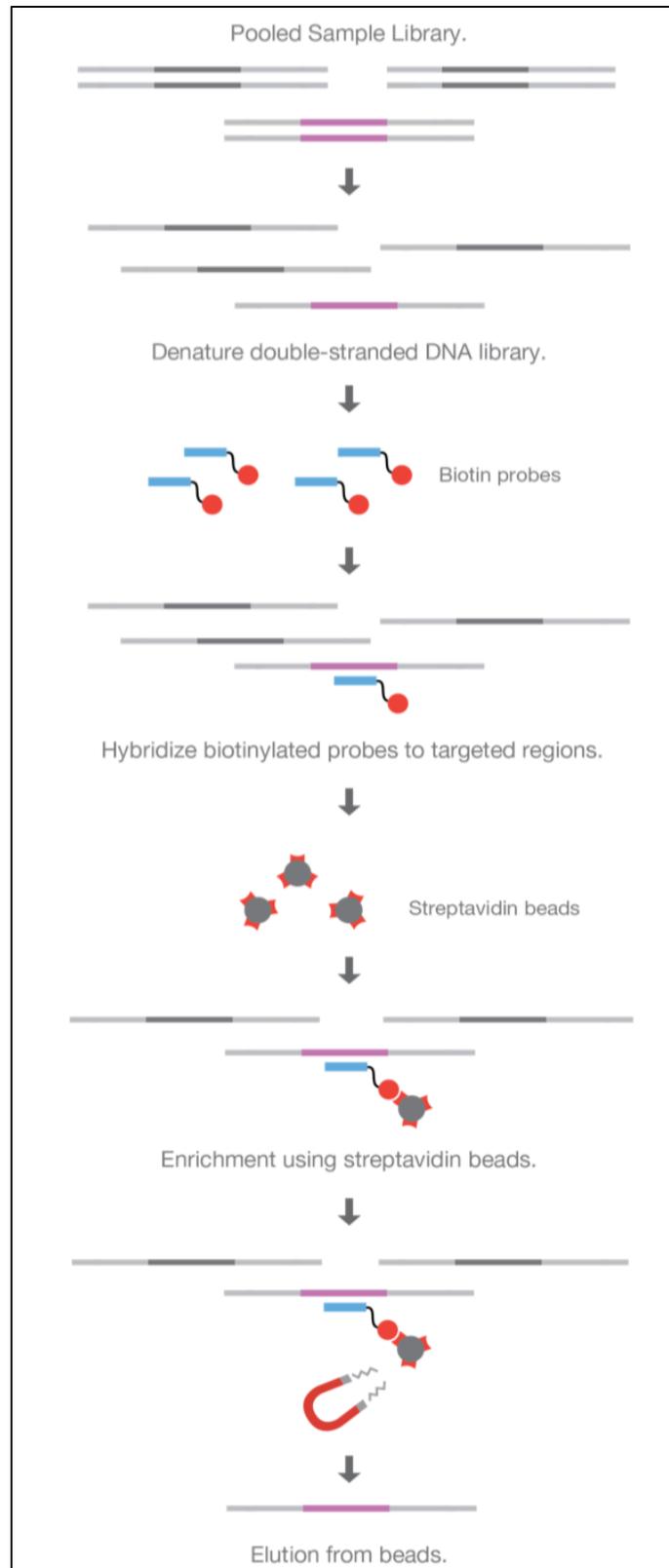


Figure 20 : Schéma des différentes étapes du procédé de capture (276).

Séquençage :

Les échantillons ont été séquençés sur le séquenceur HiSeq 4000 d'Illumina, en poolant 5 échantillons par ligne sur les flowcells, et en utilisant le protocole de 2x150 bp reads. 4 runs ont été nécessaires pour séquençer la totalité des échantillons du projet (**Figure 21**).

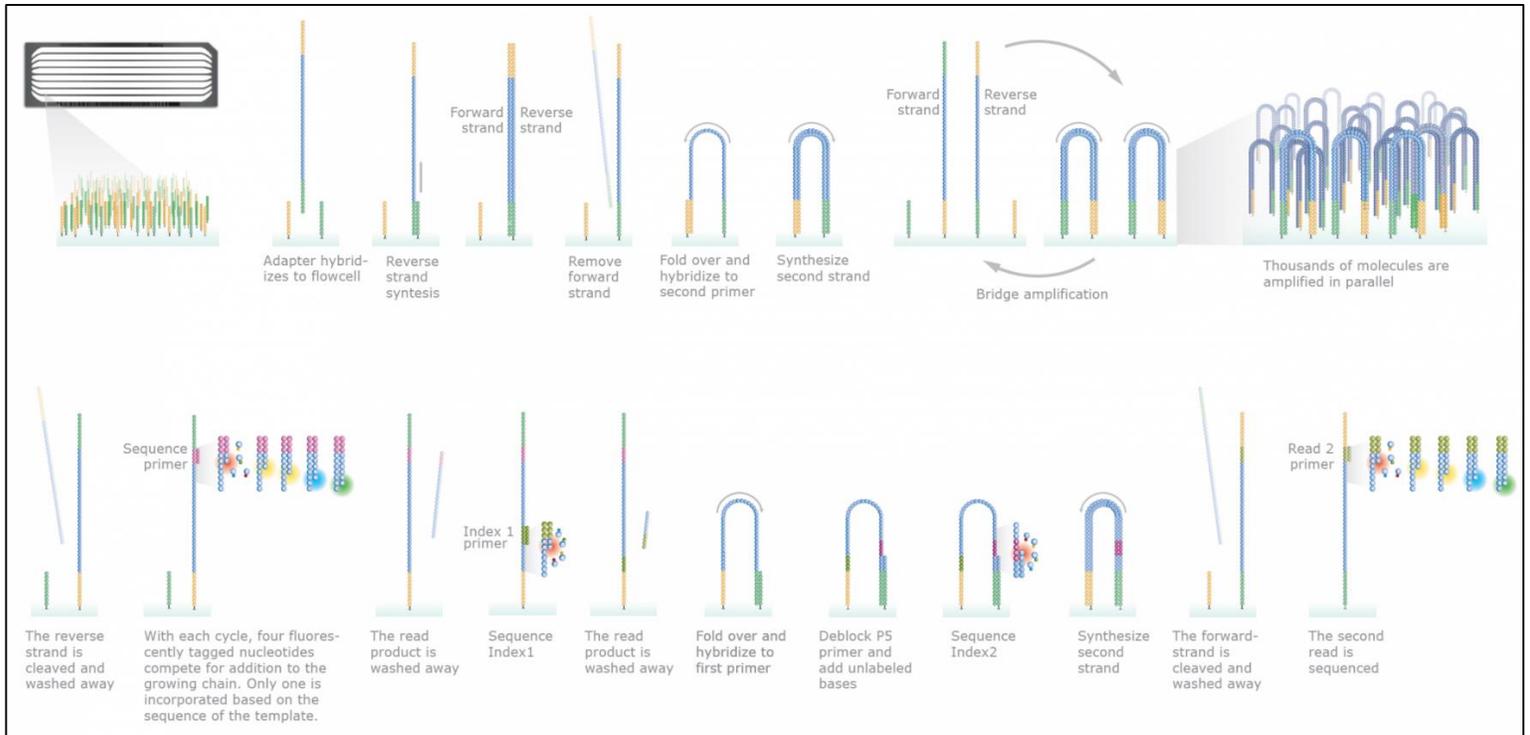


Figure 21 : Schéma de la technologie de séquençage SBS (sequencing by synthesis) de Illumina (277).

Analyse bioinformatique et analyse des variants :

Les reads séquençés ont ensuite été démultiplexés, c'est-à-dire attribués à un des échantillons, grâce au logiciel bcl2fastq (Illumina, version 2.19.1), pour générer des fastq normaux. Ces fastq ont ensuite été alignés sur le génome, en version hg38, en utilisant le logiciel BWA (Burrow-Wheeler Aligner, version 0.7.15) (278). En moyenne 145 millions de reads ont été alignés pour chacun des échantillons, dont 95,7% \pm 0,8% ont été alignés correctement. L'étape suivante, celle

de la recherche des variants, a été réalisée à l'aide du logiciel GATK (Genome Analysis ToolKit, version 3.7) (262). Tous les variants ponctuels ainsi que les petites insertions et délétions référencés ont été annotés, en utilisant l'outil Ensembl Perl Application Program Interfaces (version 89) ainsi que des scripts en langage Perl développés au laboratoire pour inclure les données des bases dbSNP (version 135) et dbNSFP (version 3.4).

Une fois l'ensemble des variants annotés, nous nous sommes focalisés sur les variants rares touchants des gènes impliqués dans les formes d'obésité mendélienne ou de déficience intellectuelle, ainsi que les pathologies en lien avec l'obésité, comme l'hypercholestérolémie, le diabète et la lipodystrophie. L'expertise nécessaire à l'interprétation des variants concernant la déficience intellectuelle nous a été apportée par les docteurs Jean Muller et Amélie Piton du laboratoire de diagnostic moléculaire de l'hôpital universitaire de Strasbourg. Pour l'interprétation des variants ponctuels, nous nous sommes basés sur les critères de l'*American College of Medical Genetics and Genomics (ACMG)* (279), qui fournissent un tableau de critères validés ou non par le variant, ce qui donne ensuite une classification pour celui-ci, pouvant être « *pathogenic* », « *likely pathogenic* », « *of uncertain significance* » jusque « *benign* » (**Figure 22** et **Table 2**). L'ensemble des variants considérés lors de cette classification étaient couverts par au minimum 20 *reads* en profondeur, et possédaient un phred score de qualité supérieur à 150.

	Benign		Pathogenic			
	Strong	Supporting	Supporting	Moderate	Strong	Very Strong
Population Data	MAF is too high for disorder <i>BA1/BS1</i> OR observation in controls inconsistent with disease penetrance <i>BS2</i>			Absent in population databases <i>PM2</i>	Prevalence in affecteds statistically increased over controls <i>PS4</i>	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product <i>BP4</i> Missense in gene where only truncating cause disease <i>BP1</i> Silent variant with non predicted splice impact <i>BP7</i>	Multiple lines of computational evidence support a deleterious effect on the gene /gene product <i>PP3</i>	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before <i>PM5</i> Protein length changing variant <i>PM4</i>	Same amino acid change as an established pathogenic variant <i>PS1</i>	Predicted null variant in a gene where LOF is a known mechanism of disease <i>PVS1</i>
Functional Data	Well-established functional studies show no deleterious effect <i>BS3</i>		Missense in gene with low rate of benign missense variants and path. missenses common <i>PP2</i>	Mutational hot spot or well-studied functional domain without benign variation <i>PM1</i>	Well-established functional studies show a deleterious effect <i>PS3</i>	
Segregation Data	Non-segregation with disease <i>BS4</i>		Co-segregation with disease in multiple affected family members <i>PP1</i>	Increased segregation data →		
De novo Data				<i>De novo</i> (without paternity & maternity confirmed) <i>PM6</i>	<i>De novo</i> (paternity & maternity confirmed) <i>PS2</i>	
Allelic Data		Observed in <i>trans</i> with a dominant variant <i>BP2</i> Observed in <i>cis</i> with a pathogenic variant <i>BP2</i>		For recessive disorders, detected in <i>trans</i> with a pathogenic variant <i>PM3</i>		
Other Database		Reputable source w/out shared data = benign <i>BP6</i>	Reputable source = pathogenic <i>PP5</i>			
Other Data		Found in case with an alternate cause <i>BP5</i>	Patient's phenotype or FH highly specific for gene <i>PP4</i>			

Figure 22 : Critères de pathogénicité de l’American College of Medical Genetics and Genomics (279).

<p>Pathogenic</p> <ol style="list-style-type: none"> 1 Very Strong (PVS1) AND ≥ 1 Strong (PS1–PS4) OR ≥ 2 Moderate (PM1–PM6) OR 1 Moderate (PM1–PM6) and 1 Supporting (PP1–PP5) OR ≥ 2 Supporting (PP1–PP5) ≥ 2 Strong (PS1–PS4) OR 1 Strong (PS1–PS4) AND ≥ 3 Moderate (PM1–PM6) OR 2 Moderate (PM1–PM6) AND ≥ 2 Supporting (PP1–PP5) OR 1 Moderate (PM1–PM6) AND ≥ 4 Supporting (PP1–PP5) <p>Likely Pathogenic</p> <ol style="list-style-type: none"> 1 Very Strong (PVS1) AND 1 Moderate (PM1–PM6) OR 1 Strong (PS1–PS4) AND 1–2 Moderate (PM1–PM6) OR 1 Strong (PS1–PS4) AND ≥ 2 Supporting (PP1–PP5) OR ≥ 3 Moderate (PM1–PM6) OR 2 Moderate (PM1–PM6) AND ≥ 2 Supporting (PP1–PP5) OR 1 Moderate (PM1–PM6) AND ≥ 4 Supporting (PP1–PP5) <p>Benign</p> <ol style="list-style-type: none"> 1 Stand-Alone (BA1) OR ≥ 2 Strong (BS1–BS4) <p>Likely Benign</p> <ol style="list-style-type: none"> 1 Strong (BS1–BS4) and 1 Supporting (BP1–BP7) OR ≥ 2 Supporting (BP1–BP7)
--

Table 2 : Règles de combinaison des critères de l’*American College of Medical Genetics and Genomics* pour définir la pathogénicité d’un variant (279).

Concernant l’analyse des CNV, nous avons utilisé le logiciel XHMM (pour eXome Hidden Markov Model, version 1.0) (280). Ce logiciel se base sur les comparaisons de profondeurs de couverture entre cas et témoins au niveau des sondes de la capture pour déterminer la présence ou l’absence d’un CNV sur une région donnée (**Figure 23**). Cette profondeur de couverture a été calculée par le logiciel GATK (version 3.7) (251). Le logiciel commence par créer une ligne de base en moyennant les profondeurs des témoins, en l’occurrence l’ensemble des patients de l’étude, qui servira de base de comparaison pour la détection de CNV chez un individu donné. Cette de ligne de base normalise les profondeurs de chaque sonde via une analyse de composante principale, et le logiciel réalise ensuite la détection de CNV pas l’utilisation de la méthode de la chaîne de Markov cachée.

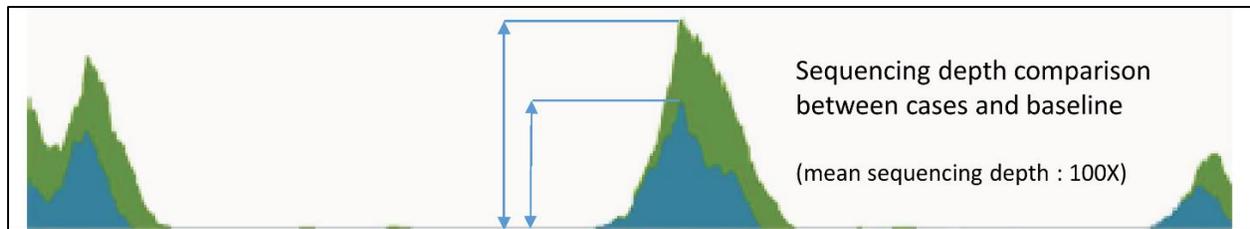


Figure 23 : Principe de la détection de CNV par comparaison des profondeurs de séquençage.

Un second logiciel de détection de CNV a également été utilisé, CANOES (CNVs with an Arbitrary Number Of Exome Samples) (281), de façon à comparer les résultats issus du logiciel XHMM et à les infirmer ou non. Ce logiciel fonctionne de manière similaire à XHMM, en utilisant les variations de profondeurs de séquençage entre cas et témoins, et procède ensuite en créant une ligne de base qui normalise les profondeurs de chaque sonde via une analyse de composante principale, puis en réalisant la détection de CNV pas l'utilisation de la méthode de la chaîne de Markov cachée.

Lors de notre étude témoin, nous avons comparé les CNV détectés par CoDE-seq avec ceux détectés par la puce CGH précédemment réalisée. Ces CNV précédemment connus nous ont par ailleurs permis d'étalonner correctement notre méthode de détection, de façon à maximiser la spécificité, tout en gardant une bonne sensibilité. Nous avons également procédé à la détection des CNV en utilisant uniquement des données issues de la capture WES classique, pour comparer l'efficacité du WES avec celle de CoDE-seq, en ne réalisant la détection que sur les sondes incluses dans le BED de la capture WES classique.

Résultats :

Au total, 145 échantillons ont été séquencés : les 40 patients déficients intellectuels de l'étude témoin, les 30 patients de la cohorte OSV, 12 patients adressés individuellement au laboratoire, ainsi que les parents de certains de ces patients. Pour l'ensemble de ces échantillons, les résultats de séquençage sont tout à fait corrects, avec une profondeur moyenne de $120 \pm 31,5$ reads. Cette profondeur élevée est nécessaire à la bonne détection des CNV, étant donné que celle-ci se base uniquement sur la profondeur de séquençage. Concernant les détails de cette couverture, en moyenne $99,1 \pm 0,1$ % des 91,5Mb de la capture sont couverts par au moins 8 reads, $97,0 \pm 2,7$ % par au moins 20 reads, et $83,2 \pm 11,2$ % par au moins 50 reads.

Les précédentes analyses par puces CGH ont permis de détecter 97 CNV chez ces 40 patients (59 duplications et 38 délétions). La taille moyenne de ces événements détectés via les puces CGH est de 1,1Mb, avec un minimum à 57kb et un maximum à 31Mb.

Avec notre système CoDE-seq, les analyses de détection de CNV sur les patients de l'étude pilote ont permis de détecter au total 181 CNV chez ces 40 patients (102 duplications et 79 délétions), localisés sur l'ensemble des chromosomes. La taille moyenne de ces événements détectés par CoDE-seq est de 720kb, avec un minimum à 1,8kb et un maximum à 31Mb.

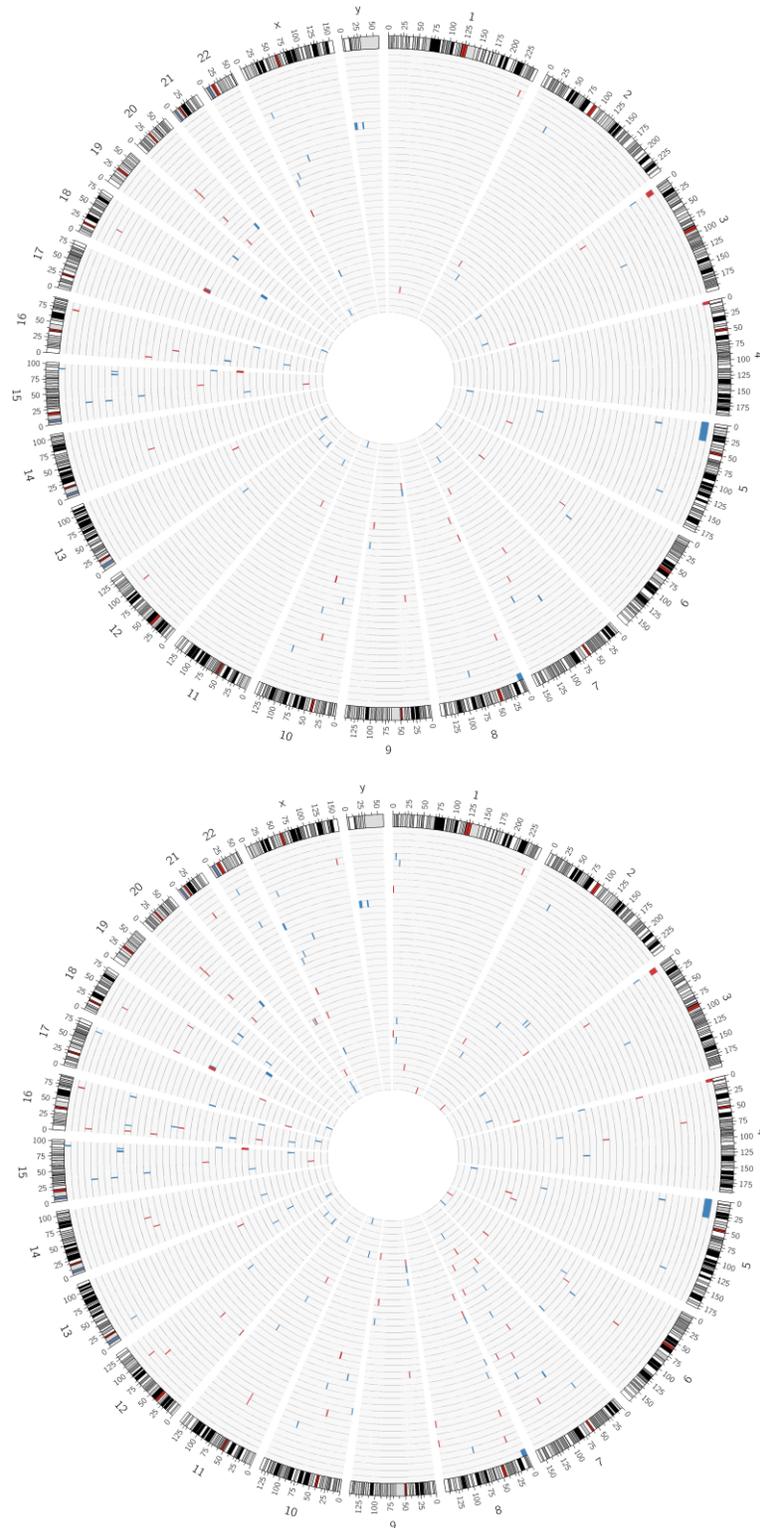


Figure 24 : CNV détectés chez les 40 patients de la cohorte de l'étude témoin par puces CGH (en haut) et par méthode CoDE-seq (en bas). Chacun des 40 cercles concentriques représente un patient. Les CNV de type duplication sont représentés en bleu, et les délétions en rouge. Représentation générée par le logiciel Circos (282).

L'élément le plus important de cette étude pilote est que l'ensemble des 97 CNV détectés (100%) par les puces CGH ont été retrouvés par CoDE-seq. Mais 67% des CNV retrouvés par CoDE-seq ont une taille plus importante, de 5% minimum, que lors de leur détection par puce à ADN (**Figure 24** et **Table 3**). Ces différences de tailles s'expliquent par la densité de sondes plus importante de notre méthode comparée aux puces CGH utilisées, qui permet de détecter plus finement l'emplacement des extrémités de ces CNV (**Supplementary Table D**).

84 CNV (46%) détectés par CoDE-seq n'ont pas du tout été détectés par les puces CGH, principalement à cause du manque de sondes des puces CGH dans les régions où ces CNV ont été détectés. La taille moyenne de ces nouveaux CNV est de 162kb, avec un minimum à 1,8kb, et un maximum à 2,5Mb.

Nous avons également réalisé la détection des CNV uniquement avec les sondes incluses dans la capture WES classique, de façon à vérifier si CoDE-seq est plus performante ou non que le WES pour la détection de CNV. Le WES a permis la détection de 114 CNV. La taille moyenne de ces CNV est de 855kb, avec un minimum à 1,8kb, et un maximum à 27Mb. L'ensemble de ces 114 CNV ont également été détectés par CoDE-seq, mais 55% d'entre eux ont une taille plus importante d'au moins 5% via CoDE-seq comparée au WES.

Soixante-sept CNV (37% des 181 initiaux) détectés par CoDE-seq n'ont pas été détectés par WES. La taille moyenne de ces CNV est de 197kb, avec un minimum à 11kb, et un maximum à 667kb. Parmi les 114 CNV détectés par WES, 42 d'entre eux (37%) n'ont pas été détectés par puces CGH, et parmi les 97 CNV détectés par puces CGH, 25 d'entre eux (26%) n'ont pas été détectés par WES.

Enfin parmi les 181 CNV détectés par CoDE-seq, 42 d'entre eux (23%) n'ont été détectés ni par WES ni par puces CGH (**Supplementary Figure A**). Pourtant, 24 de ces 42 CNV ont lieu dans des régions codantes du génome. La taille moyenne de ces 42 nouveaux CNV est de 144kb, avec un minimum à 11kb, et un maximum à 456kb (**Figure 25**).

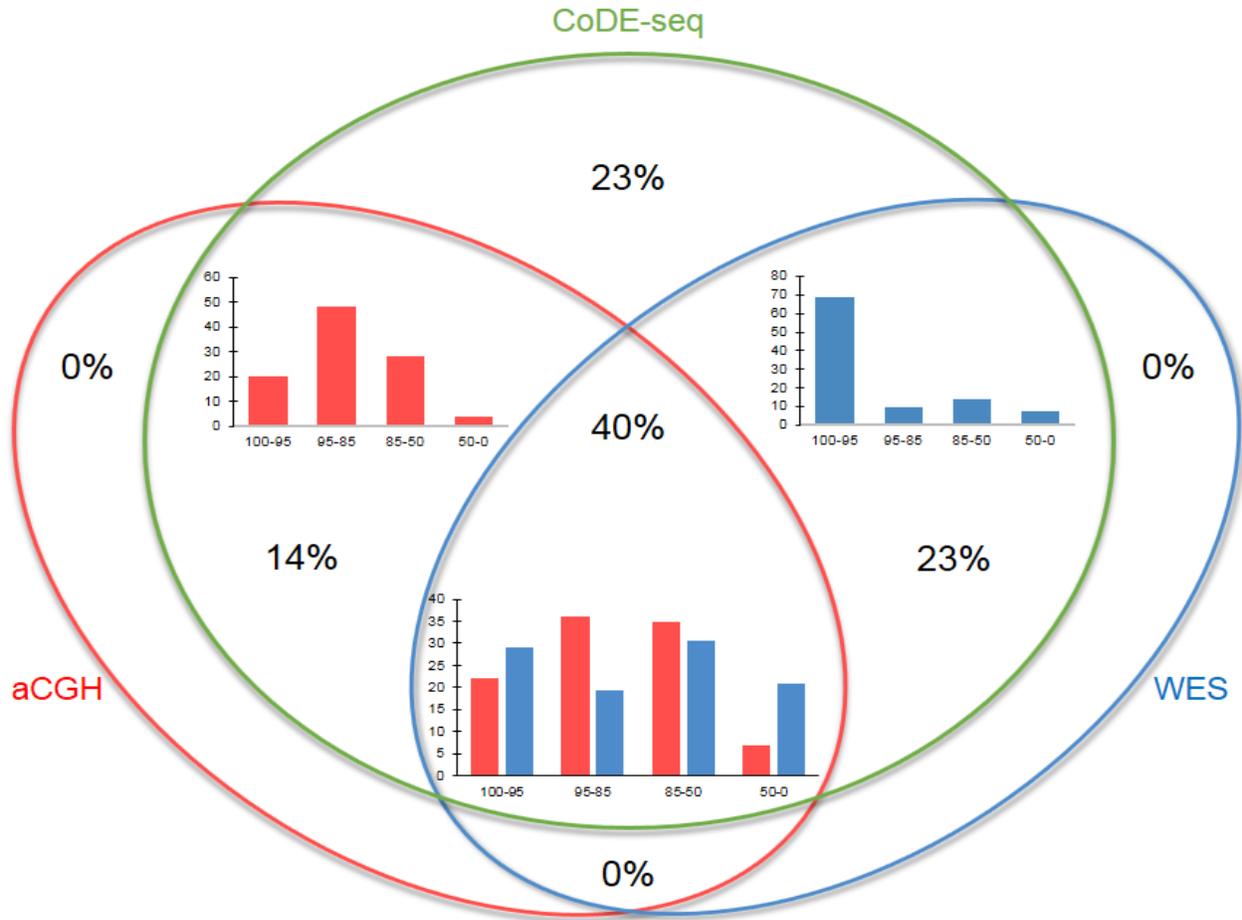


Figure 25 : Détection via les méthodes CoDE-seq, séquençage d'exome et puce CGH des CNV chez les 40 patients de la cohorte de l'étude pilote. Les histogrammes représentent la fréquence en % de CNV détectés par puce à ADN (en rouge) ou par séquençage d'exome (en bleu) selon leur taille en % comparée à celle détectée par CoDE-seq.

Ces différences dans la qualité de détection sont principalement dues à la différence du nombre de sondes incluses par chacune de ces techniques, ainsi qu'à leur répartition. En effet, les puces CGH utilisées lors du diagnostic des 40 patients de l'étude pilote comportaient environ 60 000 sondes. La capture NimbleGen SeqCap EZ MedExome Enrichment 226 072 sondes, et notre capture CoDE-seq 381 635 sondes. Ces sondes additionnelles permettent de détecter de nouveaux événements, et d'être plus précis sur la détection de ces événements (**figures 26 et 27**). La détection de CNV est donc plus sensible via CoDE-seq que par le même procédé appliqué à une capture WES ou par des puces CGH utilisées en routine en services hospitaliers.

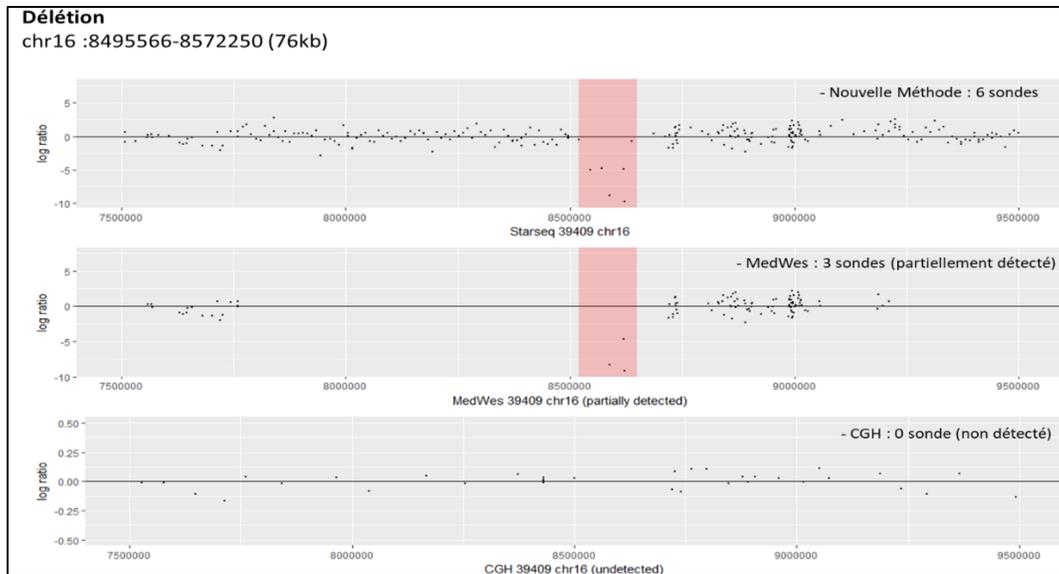


Figure 26 : Exemple d'un CNV de type délétion sur le chromosome 16 par les 3 techniques : CoDE-seq en haut, capture exome au centre, et puce CGH en bas. Les points représentent les sondes utilisées, avec en ordonnée le log-ratio de profondeur du patient analysé par rapport à la ligne de base de l'ensemble des autres patients de l'étude, en en abscisse les positions génomiques en bases. On remarque que la puce CGH ne l'a pas détecté car aucune sonde n'est présente dans cette zone.

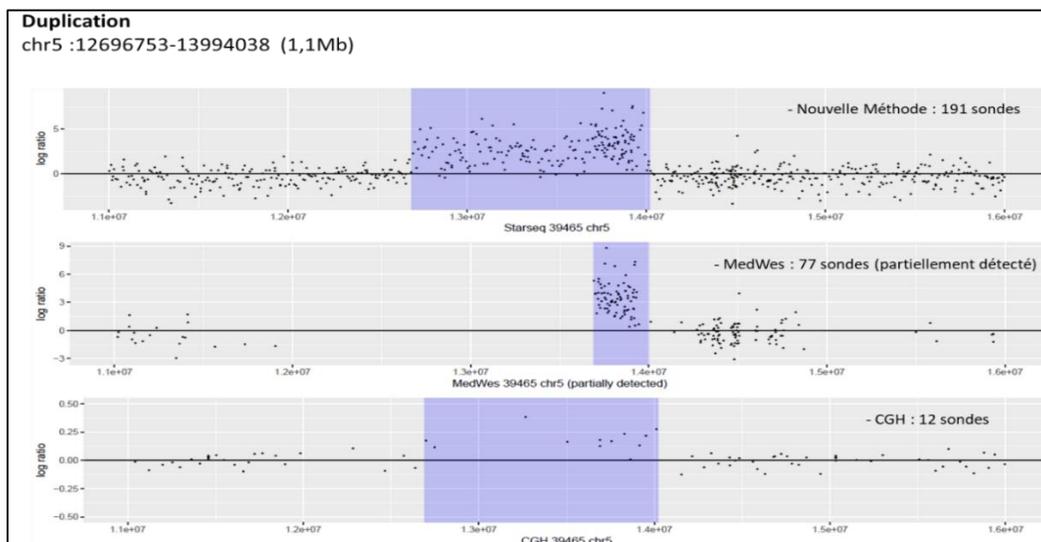


Figure 27 : Exemple d'un CNV de type duplication sur le chromosome 5 investigué par les 3 techniques : CoDE-seq en haut, capture exome au centre, et puce CGH en bas. Les points représentent les sondes utilisées, avec en ordonnée le log-ratio de profondeur du patient analysée par rapport à la ligne de base de l'ensemble des autres patients de l'étude, en en abscisse les positions génomiques en bases. On remarque que la capture exome ne l'a détecté que partiellement car elle ne possède que peu de sondes dans cette zone.

Nous avons par ailleurs procédé à une comparaison des résultats entre nos deux logiciels de détection de CNV : XHMM et CANOES. Cette comparaison avait deux objectifs : vérifier si les résultats obtenus par XHMM lors de l'utilisation de CoDE-seq sont retrouvés et donc confirmés par CANOES. Et regarder si les résultats communs à ces deux logiciels sont davantage robustes lors de l'utilisation de CoDE-seq que lors de l'utilisation du séquençage d'exome classique.

Concernant les 89 CNV détectés par XHMM au niveau de l'ensemble des chromosomes hors X et Y chez les 40 patients de la cohorte de l'étude pilote (CANOES ne réalise pas nativement la détection des CNV sur ces deux chromosomes) lors de l'utilisation de CoDE-seq, 88 (98%) ont été détectés de nouveau par CANOES, avec des tailles relativement similaires (**Table 3**).

Concernant maintenant la détection de CNV via XHMM lors de l'utilisation d'une capture WES classique au niveau de l'ensemble des chromosomes hors X et Y chez ces mêmes 40 patients de la cohorte de l'étude pilote, seuls 64 des 89 CNV avaient été détectés. CANOES est parvenu à détecter 70 de ces 89 CNV. Et 61 (68%) ont été retrouvés par à la fois CANOES et XHMM lors de l'utilisation d'une capture exome classique (**Table 3**).

CHR	START	END	SAMPLE	TYPE	XHMM_CoDE-seq	XHMM_MedWes	CANOES_CoDE-seq	CANOES_MedWes
chr18	32160092	32220277	28189	DEL	60	30	60.185	30.236
chr8	85828677	86153000	28189	DUP	324	104	324.323	104.562
chr10	80042528	80202743	29352	DUP	160	80	179.501	94.214
chr17	73856209	74671695	29352	DUP	815	467	815.486	467.783
chr6	145645449	145856364	29352	DUP	211	NA	210.915	NA
chr13	22595208	23079332	30028	DUP	484	NA	484.124	NA
chr13	20489343	20822317	30028	DUP	333	332	332.974	332.974
chr22	22646288	23364631	30028	DUP	718	669	719.948	670.878
chr15	36667786	36950430	30226	DEL	283	NA	254.012	258.779
chr5	8704478	9305400	30226	DUP	601	222	600.922	222.695
chr5	11266288	11521607	30226	DUP	255	NA	265.065	NA
chr1	71579368	71962893	31425	DEL	384	NA	383.525	342.553
chr12	79522765	79872962	31425	DUP	350	287	359.153	349.909
chr12	15594901	15813060	31425	DUP	218	137	218.159	137.891
chr11	32931712	33068868	32776	DUP	137	137	137.156	137.156
chr7	101210302	101616335	32776	DUP	406	108	414.258	414.258
chr16	29568896	30206106	32996	DUP	637	542	707.529	581.916
chr3	157391680	158223677	32996	DUP	832	793	831.997	793.409
chr3	15309522	15568614	35144	DUP	259	238	259.092	238.138

chr3	14089868	14198145	35144	DUP	108	85	133.312	85.28
chr9	4354549	4526950	35144	DEL	172	NA	172.401	NA
chr1	237142228	237388211	35361	DEL	246	117	245.983	117.738
chr12	99648143	99744292	35361	DEL	96	NA	96.149	96.149
chr13	69513474	70397597	35871	DUP	884	406	884.123	406.066
chr9	4718499	5535070	35871	DUP	817	804	804.137	804.137
chr9	3648286	3958614	35871	DUP	310	108	310.328	NA
chr2	66390896	66817732	36499	DUP	427	135	426.836	135.761
chr6	113900368	113970939	36499	DEL	71	30	70.571	29.909
chr16	75880696	76581426	36504	DEL	701	281	725.64	281.068
chr3	43153	7993841	36504	DEL	7951	7414	7950.688	7420.68
chr5	41442	31162895	36504	DUP	31121	26875	31141.15	26875.139
chr3	195175598	195451820	36578	DEL	276	267	276.222	267.289
chr5	116172508	116242834	36578	DEL	70	NA	70.326	NA
chr7	133131697	133317413	36578	DEL	186	NA	185.716	64.336
chr16	84529537	84768392	36616	DUP	239	201	238.855	201.569
chr2	44320188	44570056	36616	DEL	250	229	249.868	229.507
chr11	31520033	31707194	36711	DEL	187	162	187.161	162.071
chr15	22781870	23129536	36711	DUP	348	342	257.713	252.891
chr15	98507942	101861073	36981	DEL	3353	3169	3380.937	3169.611
chr18	71138517	80257123	36981	DUP	9119	5489	9110.834	5489.391
chr9	123557538	123960973	36981	DEL	403	371	403.435	372.392
chr5	42633203	42810867	37146	DUP	178	122	NA	121.999
chr8	8847017	9224537	37146	DEL	378	265	377.52	265.051
chr16	16492861	18175294	37403	DUP	1682	353	1682.433	362.172
chr13	91409277	91650157	37465	DEL	241	NA	240.88	NA
chr4	72955190	73177559	37465	DUP	222	119	222.369	119.479
chr9	131009179	131164184	37465	DUP	155	155	155.005	155.005
chr15	100048822	100341928	37588	DUP	293	293	293.106	293.106
chr8	2949276	3108782	37588	DEL	160	160	159.506	159.506
chr12	111744618	111871215	37872	DUP	127	127	126.597	126.597
chr15	92191681	92670133	38100	DUP	478	261	478.452	261.649
chr4	85715	3992441	38100	DEL	3907	3682	3971.477	3682.597
chr8	205215	7121027	38100	DUP	6916	6815	6973.779	6815.798
chr15	55223860	55324856	38114	DEL	101	101	100.996	100.996
chr19	29798945	31173810	38206	DUP	1375	898	1374.865	898.489
chr20	14719244	14867459	38206	DEL	148	NA	158.133	NA
chr18	47535	7791187	38213	DEL	7744	7726	7743.822	7726.927
chr18	7810771	8293586	38213	DUP	483	365	482.815	365.377
chr20	59032190	64329015	38213	DUP	5297	5255	5255.728	5255.728
chr10	53619410	54881439	38506	DEL	1262	721	1262.029	857.745

chr7	80405243	80581524	38506	DEL	176	NA	176.281	53.318
chr16	28811431	29052826	38645	DEL	241	227	294.924	294.924
chr6	72045966	72284159	38645	DEL	238	187	238.193	187.243
chr9	28621496	28701352	38785	DEL	80	NA	89.756	NA
chr10	24275739	24513481	38818	DUP	238	238	237.742	NA
chr20	14719244	14887403	38818	DEL	168	NA	168.159	NA
chr6	89029179	89699753	38818	DUP	671	618	669.863	618.182
chr3	17214183	17470105	38839	DEL	256	NA	255.922	NA
chr7	108564156	108887149	38839	DEL	323	NA	347.897	320.405
chr10	66990109	67167854	38991	DUP	178	NA	193.04	NA
chr16	6982497	7333124	38991	DEL	351	NA	325.682	NA
chr14	39000163	39096148	39147	DEL	96	62	120.81	62.931
chr15	36516961	36895301	39147	DUP	378	NA	379.755	316.89
chr7	124143633	124810634	39268	DUP	667	NA	667.001	NA
chr20	14768964	14897332	39402	DEL	128	NA	128.368	NA
chr3	97610768	97899303	39402	DUP	289	289	289.727	289.727
chr10	51344735	51887971	39409	DEL	543	NA	543.236	337.028
chr20	14176304	14564346	39409	DEL	388	NA	388.042	2.011
chr7	73316540	74798284	39409	DUP	1482	1482	1522.302	1494.909
chr15	83460498	84111657	39465	DUP	651	547	670.83	547.044
chr15	89573720	89652971	39465	DUP	79	77	79.251	77.406
chr5	12696753	13994038	39465	DUP	1297	239	1316.603	252.515
chr15	31739875	32243158	39509	DUP	503	NA	503.283	NA
chr8	15693720	15951771	39509	DEL	258	33	258.051	33.624
chr10	101383666	101628102	39612	DUP	244	197	244.436	197.786
chr2	44530376	45069527	39622	DUP	539	459	539.151	459.533
chr5	140482549	140702153	39622	DUP	220	220	219.604	219.604
chr15	31729880	32225059	39638	DUP	495	NA	522.629	NA
chr3	1107666	1308155	39638	DUP	200	150	200.489	NA

Table 3 : Comparaison des résultats obtenus sur les 40 patients de l'étude pilote par les logiciels XHMM et CANOES, avec la nouvelle capture CoDE-seq et avec la capture WES NimbleGen SeqCap EZ MedExome Enrichment. Les tailles indiquées sont en kb et les coordonnées en hg38.

Une fois cette étude preuve de concept validée, nous avons pu utiliser CoDE-seq pour procéder à la recherche de mutations ponctuelles et de CNV potentiellement pathogéniques pouvant expliquer les phénotypes de nos patients, que ce soit concernant les pathologies concernant le domaine de la déficience intellectuelle ou celui de l'obésité. Cette recherche de diagnostic a été effectuée sur l'ensemble des 82 patients de notre étude.

Nous nous sommes pour cela intéressé aux évènements touchant une liste de gènes connus pour être associés avec l'une ou l'autre de ces pathologies (**Supplementary Table C**).

Nous avons trouvé 13 CNV potentiellement pathogéniques (7 duplications et 6 délétions) d'en moyenne 3,5Mb, avec un minimum de 241kb et un maximum de 30,7Mb (**Table 4**). Ces CNV ont été trouvés chez 13 patients différents (15,9% d'entre eux). Ils sont localisés au niveau des chromosomes 5, 7, 15, 16 et X. Ils ont auparavant été associés à l'obésité ou aux pathologies associées à la déficience intellectuelle pour lesquels nous les avons considérés comme pathogéniques (**Supplementary Table A**).

Patient	Sex	Chr	Start	End	Size	Probes	Type	Phenotype
32996	F	16	29568896	30206106	637.21	215	Gain	Intellectual disability [psychomotor retardation, language delay], short stature, thinness, dysmorphic feature
36504	F	5	41442	31162895	30743.13	3989	Gain	Intellectual disability
36981	F	15	98507942	101861073	3353.13	513	Loss	Intellectual disability [psychomotor retardation], failure to thrive, short stature, thinness, microcephaly, dysmorphic feature
38213	F	20	59032190	64329015	5296.83	1180	Gain	Intellectual disability [psychomotor retardation, spatial disorientation], aggressive behavior, short stature
38645	M	16	28811431	29052826	241.4	136	Loss	Intellectual disability [psychomotor retardation, language delay]
38839	M	X	53231708	53878306	646.6	185	Gain	Intellectual disability [psychomotor retardation], neonatal hypotonia, thinness
39409	F	7	73316540	74798284	1481.74	339	Gain	Intellectual disability [psychomotor retardation, language delay], behavioral disorder
39509	M	15	31739875	32243158	503.28	58	Gain	Intellectual disability [language delay], seizures
39638	F	15	31729880	32225059	495.18	60	Gain	Intellectual disability
OSV20	M	16	29572106	30188122	616.02	211	Loss	Intellectual disability [language delay], obesity
OSV21	F	16	29642461	30194811	552.35	203	Loss	Intellectual disability [language delay], obesity
OSV46	F	16	28811431	29052826	241.4	136	Loss	Seizures, behavioral disorder, hypermetropia, obesity, dysmorphic feature
U2	F	16	29590876	30188122	597.25	209	Loss	Intellectual disability [psychomotor retardation], severe obesity

Table 4 : CNV « *likely pathogenic* » chez l'ensemble des 82 patients du projet.

Six des 13 patients chez lesquels un CNV potentiellement pathogéniques a été détecté sont touchés au niveau de la région du chromosome 16p11.2. Trois d'entre eux sont porteurs d'une délétion d'environ 600kb au niveau proximal de la région chr16p11.2, et présentent une déficience intellectuelle ainsi qu'une obésité marquée. Un des patients est porteur de la duplication miroir des délétions précédentes, au niveau proximal de la région 16p11.2, toujours d'environ 600kb, et souffre de déficience intellectuelle et de malformations, tout en étant de petite taille et maigre. C'est un cas intéressant de CNV où la délétion et duplication provoquent

des phénotypes miroir concernant la masse corporelle (133). Enfin 2 patients présentent une délétion d'environ 200kb au niveau distal de la région 16p11.2, et souffrent de déficience intellectuelle pour l'un d'entre eux, et de crises convulsives associées avec des troubles de comportement, une hypermétropie, une obésité et des malformations pour le second.

Nous avons également détecté 45 CNV chez un total de 33 individus, classés comme étant de signification incertaine (27 duplications et 18 délétions) d'en moyenne 195kb, avec un minimum de 9,8kb et un maximum de 718kb (**Supplementary Table E**). Ces CNV ont lieu au niveau de gènes impliqués soit dans une pathologie associée à la déficience intellectuelle, soit à l'obésité, ou alors au niveau de régions connues pour être touchées par des CNV associées avec ces pathologies, mais sont classés comme de signification incertaine car leur taille est inférieure à celle des CNV connus pour être associés à ces pathologies (**Supplementary Table F**).

Les parents de certains des patients ont également été analysés concernant leurs CNV, de façon à déterminer si les éventuels CNV retrouvés chez leurs enfants sont *de novo* ou non. Nous n'avons trouvé aucun CNV *de novo* potentiellement impliqué dans les pathologies touchant les patients pour lesquels nous avons pu avoir accès à l'ADN des parents.

Les mutations ponctuelles ont également été analysées chez l'ensemble des 82 patients de l'étude (dont les mutations n'avaient pas été analysées auparavant). Nous sommes parvenus à établir un diagnostic basé sur la présence d'une mutation ponctuelle touchant un gène associé à une maladie Mendélienne correspondant au phénotype du patient chez 14 de nos patients (soit 17,1 % d'entre eux).

Parmi les mutations pathogéniques ou probablement pathogéniques, deux d'entre elles touchent le gène *MC4R* et ont été retrouvées chez deux patients atteints d'obésité morbide, de déficience intellectuelle sévère, de troubles du comportement et de malformations. La première de ces deux mutations a été détectée chez le patient OSV37, et est une mutation hétérozygote NM_005912.2:c.52C > T/p.Arg18Cys, considérée comme pathogénique selon les critères ACMG car déjà rapportée comme délétère (103) et préalablement étudiée de façon fonctionnelle (283). La seconde de ces mutations a été détectée chez le patient OSV37, et est une mutation hétérozygote NM_005912.2:c.202del/p.Ala68Glnfs*2, considérée comme pathogénique selon les critères ACMG car provoquant l'apparition d'un codon stop, et n'ayant jamais été rapportée

dans la base gnomAD. Une mutation non-sens du gène *PTEN* a également été trouvée chez un patient souffrant de lipomes abdominaux, dont la mère, atteinte de cancers de la thyroïde et de l'utérus, est elle aussi porteuse.

Un des patients souffrant d'une forme de déficience intellectuelle a pu être diagnostiqué via la détection de deux mutations distinctes dans le gène *MTHFR* expliquant son phénotype.

Dix autres patients ont pu être diagnostiqués pour leur déficience intellectuelle par la présence d'une mutation ponctuelle au niveau des gènes suivants : *ATRX*, *MECP2*, *RAI1*, *EEF1A2*, *KDM5B*, *NF1*, *SL6AT1*, *TSC1*, et *NEXMIF / KIAA2022*.

Nous avons également détecté 8 autres mutations chez 7 patients différents, concernant des formes monogéniques de diabète, d'hyperinsulinémie, de convulsions, d'hypercholestérolémie et de déficience intellectuelle, bien que ces traits n'étaient pas ceux pour lesquelles ces patients nous ont été adressés (**Table 5**).

Patient	Sex*	Gene	Mutation	Status	Parental origin**	Pathogenicity	ACMG criteria
39465	F	<i>WFS1</i>	c.409_424dup / p.Val142Glyfs*110	het	M	Pathogenic	PVS1, PS1, PP2, PP4
OSV11	F	<i>NF1</i>	c.3989A>C / p.Glu1330Ala	het	Unk	Likely pathogenic	PM1, PM2, PP2, PP4
OSV17	F	<i>KCNJ11</i>	c.151G>T / p.Glu51*	het	M	Pathogenic	PVS1, PS1, PM2, PP2, PP4
OSV24	M	<i>CPA6</i>	c.544C>T / p.Arg182*	het	M	Pathogenic	PVS1, PS1, PP2, PP4
OSV24	M	<i>STAP1</i>	c.827-1del / -	het	F	Pathogenic	PVS1, PM2, PP2, PP4
OSV48	M	<i>PPARG</i>	c.419G>A / p.Arg140His	het	M	Pathogenic	PS1, PS3, PP2, PP4
OSV53	M	<i>TSC1</i>	c.2665_2666delinsAT / p.Glu889Ile	hom	Unk	Likely pathogenic	PM1, PM2, PP2, PP3, PP4
OSV58	M	<i>POGZ</i>	c.3026G>A / p.Arg1009His	het	Unk	Likely pathogenic	PM1, PM2, PP2, PP4
OSV71	M	<i>HRAS</i>	c.367C>T / p.Arg123Cys	het	Unk	Likely pathogenic	PS1, PP2, PP3, PP4
OSV77	M	<i>TSC1</i>	c.827C>A / p.Ser276Tyr	het	F	Likely pathogenic	PM1, PM2, PP2, PP3, PP4
OSV86	M	<i>HCN2</i>	c.1269C>G / p.Ser423Arg	het	D	Likely pathogenic	PS2, PM2, PP2, PP4

Table 5 : Mutations « pathogéniques » ou « probablement pathogéniques » chez l'ensemble des 82 patients du projet.

Enfin la quasi-totalité des patients étaient porteurs de mutations ponctuelles de signification incertaine (**Supplementary Table G**).

Discussion et Perspectives :

A travers ce projet comprenant une étude pilote et une étude plus étendue, nous sommes parvenus à mettre au point une méthode permettant de réaliser la détection simultanée de deux types d'évènements génétiques auparavant recherchés de façon séparée : les mutations ponctuelles et les CNV.

Les mutations ponctuelles étaient classiquement recherchées par séquençage d'exome, et la méthode CoDE-seq offre des résultats équivalents étant donné qu'elle procède en réalisant le séquençage de l'entièreté des régions codantes du génome, comme le fait le séquençage d'exome classique.

Les CNV étaient eux recherchés en pratique clinique par l'utilisation de puces CGH. Notre méthode a montré qu'elle fournissait des résultats plus complets que ceux obtenus par les puces utilisées dans la pratique clinique hospitalière, en redétektant l'ensemble des CNV trouvés par ces puces, en permettant d'affiner les limites de ces CNV, et en en trouvant de nouveaux. Ces nouveaux CNV ont pu être détectés étant donné que notre méthode inclus davantage de sondes que les puces CGH utilisées en clinique. Nous avons également pu confirmer que le WES n'était pas une méthode adaptée à la détection de CNV étant donné les différences retrouvées à la fois entre les résultats issus de CoDE-seq et ceux issus du WES classique, et entre les résultats issus de deux logiciels de détection de CNV sur les mêmes données issues de WES, qui sont bien plus robustes lors de l'utilisation de CoDE-seq que lors de l'utilisation du séquençage d'exome classique.

En plus de son efficacité supérieure, cette nouvelle méthode est également avantageuse d'un point de vue cout. Elle représente un surcoût de l'ordre de 25% de celui d'une analyse WES classique, soit environ 600 euros, ce qui la rend bien plus abordable que la réalisation conjointe des deux techniques de WES (450 euros) et de puces CGH (> 600 euros). Par ailleurs, le fait de ne réaliser qu'une seule technique permet de réaliser des économies de temps, à la fois concernant la préparation des échantillons, la technique elle-même, et l'interprétation des données. A Lille, lors d'une demande de diagnostic via puces CGH, le délai dépasse souvent 6 mois, auquel il faut ajouter en cas de résultats négatifs plusieurs semaines pour le diagnostic via séquençage

d'exome. Dans le cas de l'utilisation de CoDE-seq, 6 semaines permettent d'obtenir des résultats à la fois pour les CNV et les mutations ponctuelles.

L'objectif est maintenant de diffuser cette technologie, dans le but de modifier les pratiques hospitalières lors de la suspicion de la présence d'une anomalie génétique nécessitant la réalisation d'un séquençage d'exome ou d'une puce CGH.

A l'avenir le domaine du séquençage de l'ADN changera encore radicalement. La diminution du coût du séquençage de l'exome ou du génome humain diminuera encore largement, jusqu'à rendre le séquençage du génome compatible avec une utilisation massive. Mais les méthodes au coût actuel inférieur à celui du séquençage du génome deviendront elles aussi encore plus abordables. Une méthode comme CoDE-seq continuera donc à être compétitive par rapport au séquençage du génome, en restant plus accessible pour les applications ou les lieux où le séquençage du génome sera encore trop cher.

Les nouvelles méthodes de séquençage sans amplification, comme celles commercialisées par Oxford Nanopore ou Pacific Bioscience, vont également transformer les pratiques actuelles. Elles finiront vraisemblablement par remplacer les technologies et méthode de séquençage basées sur une amplification des échantillons puis sur la génération de *reads* courts, dont le principal représentant est aujourd'hui Illumina. Mais avant cela ces méthodes devront résoudre plusieurs soucis, comme le haut taux d'erreurs actuellement incompatible avec une utilisation diagnostique lors de la recherche de mutations rares, ou le faible débit de données générées en comparaison des meilleurs séquenceurs de *reads* courts.

On peut maintenant se poser la question de la conséquence clinique et fonctionnelle de ces CNV. La littérature actuelle est très peu développée concernant ce sujet, et la plupart des CNV connus pour être pathogéniques le sont par analyse d'association, sans mécanisme connu.

Nous supposons que les CNV pourraient agir sur les phénotypes des patients en étant porteur en modifiant la conformation spatiale de la chromatine chez ces patients (284).

Conformation spatiale de la chromatine :

La chromatine est stockée dans les cellules sous une forme extrêmement compactée, permettant aux près de 2mètres de la chaîne d'ADN humaine de tenir dans les noyaux cellulaires de quelques μm de diamètre. Cette forme compactée possède une structure très organisée, au niveau de la laquelle certains segments linéairement séparés sont très proches spatialement ((285) et **Figure 28**).

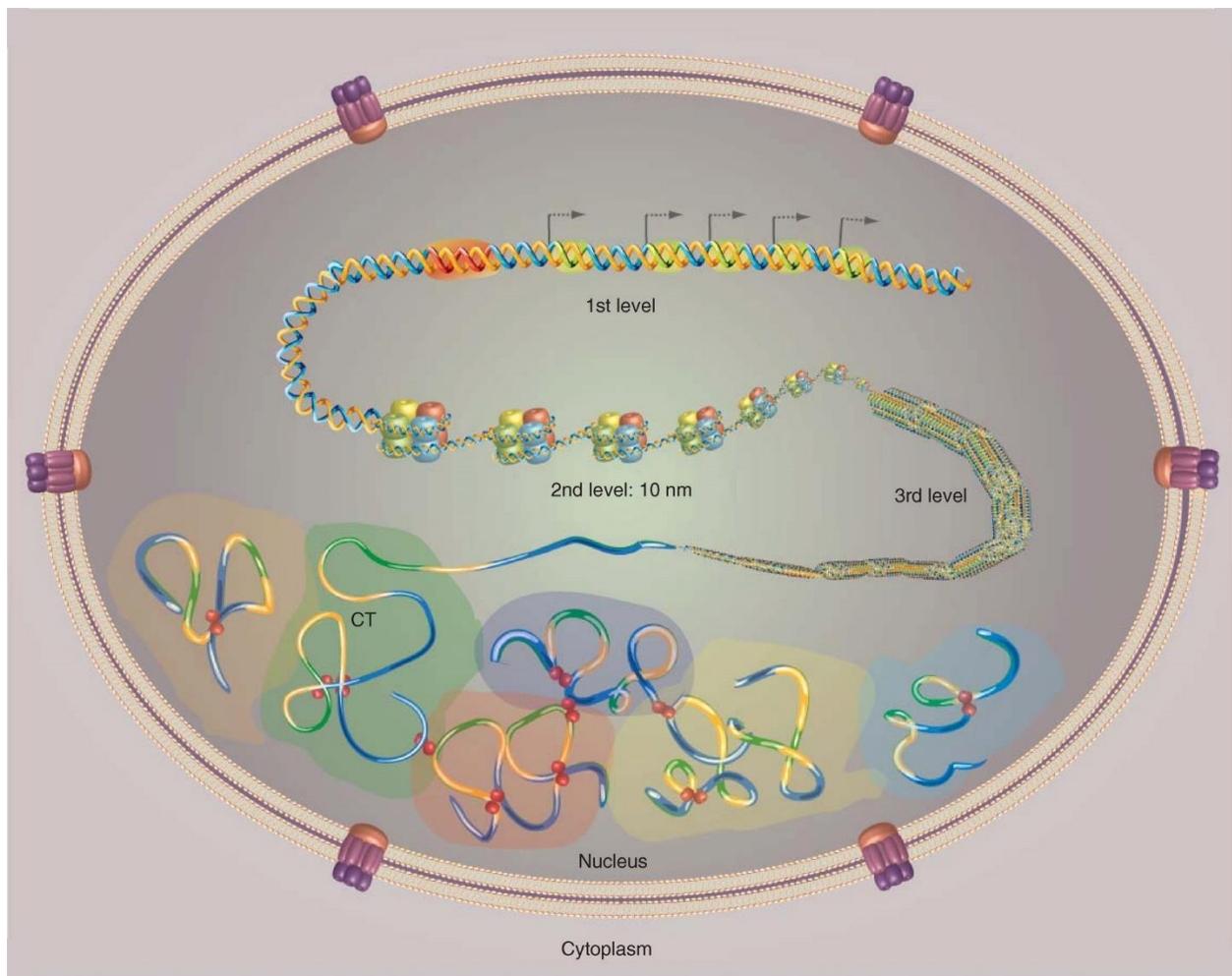


Figure 28 : Organisation de la chromatine dans la cellule (284).

Cette organisation pourrait être modifiée lors de la survenue d'un évènement génétique, comme une mutation ponctuelle, ou un CNV (**Figure 29**).

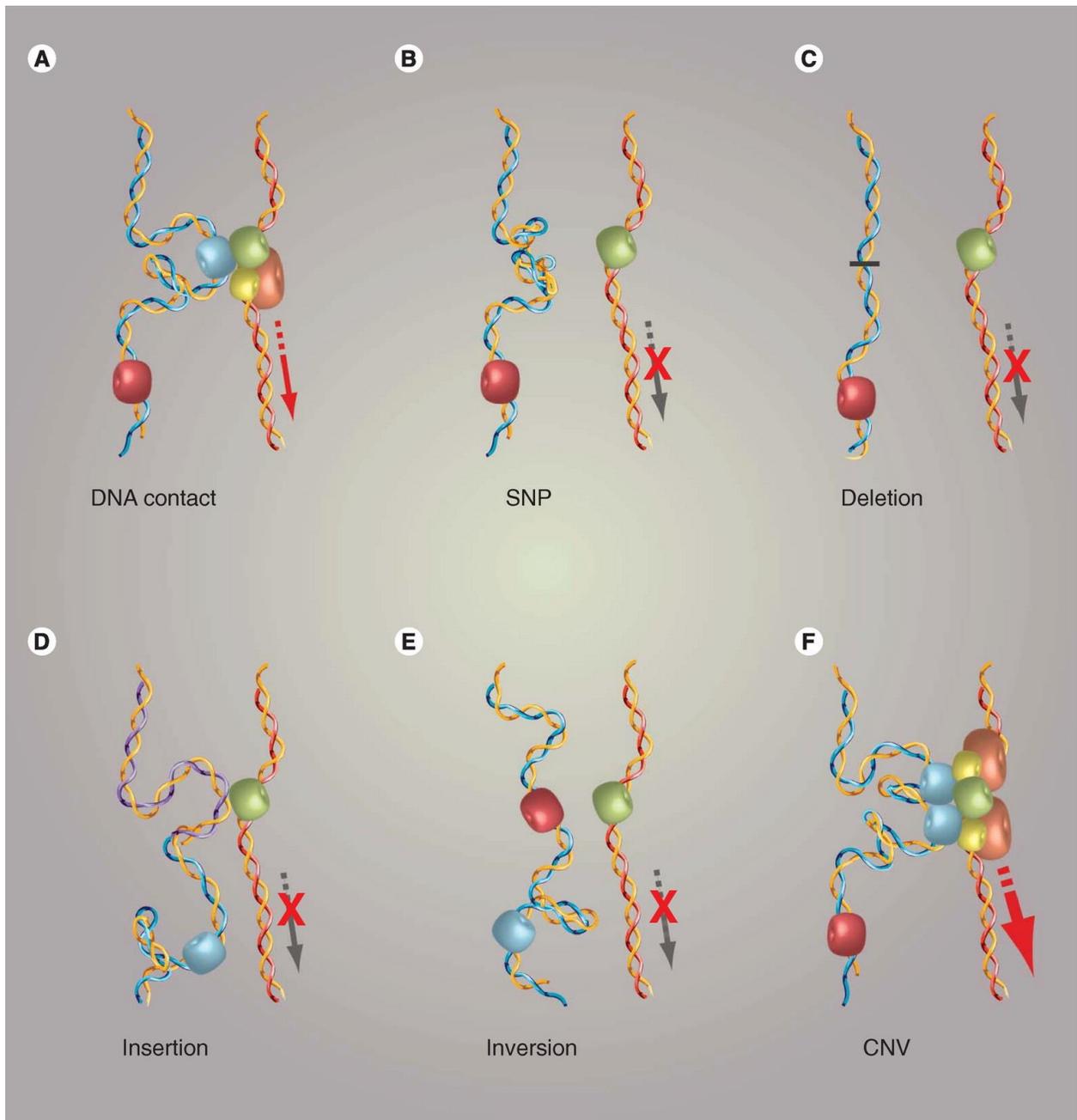


Figure 29 : Impact d'éventuelles variations du génome sur la conformation de la chromatine (284).

Ces variations génétiques pourraient, si elles surviennent au niveau d'une région régulatrice ou d'un gène, provoquer des perturbations dans l'expression de différents gènes, même si ceux-ci

sont linéairement très éloignés. Ce mécanisme pourrait expliquer l'impact de certains CNV ayant lieu apparemment loin de tout gène impliqué dans la pathologie à laquelle il est associé.

NG-CaptureC :

Il existe de nombreuses techniques d'analyse de la conformation spatiale de la chromatine, comme le 3C, le 4C ((286) et **Figure 30**), ou l'Hi-C cité précédemment (250). Elles se basent toutes sur le fait de « *cross-linker* » l'ADN, c'est-à-dire de bloquer les liaisons entre l'ADN et les protéines qui y sont fixées, pour en maintenir la conformation spatiale (comme pour la technique du ChIP), de l'extraire des cellules, de le fragmenter puis de procéder à une étape de ligation pour créer des fragments « hybrides » de faible taille reliant deux zones proches dans l'espace, et enfin d'éliminer les protéines pour ne conserver que les fragments d'ADN hybrides.

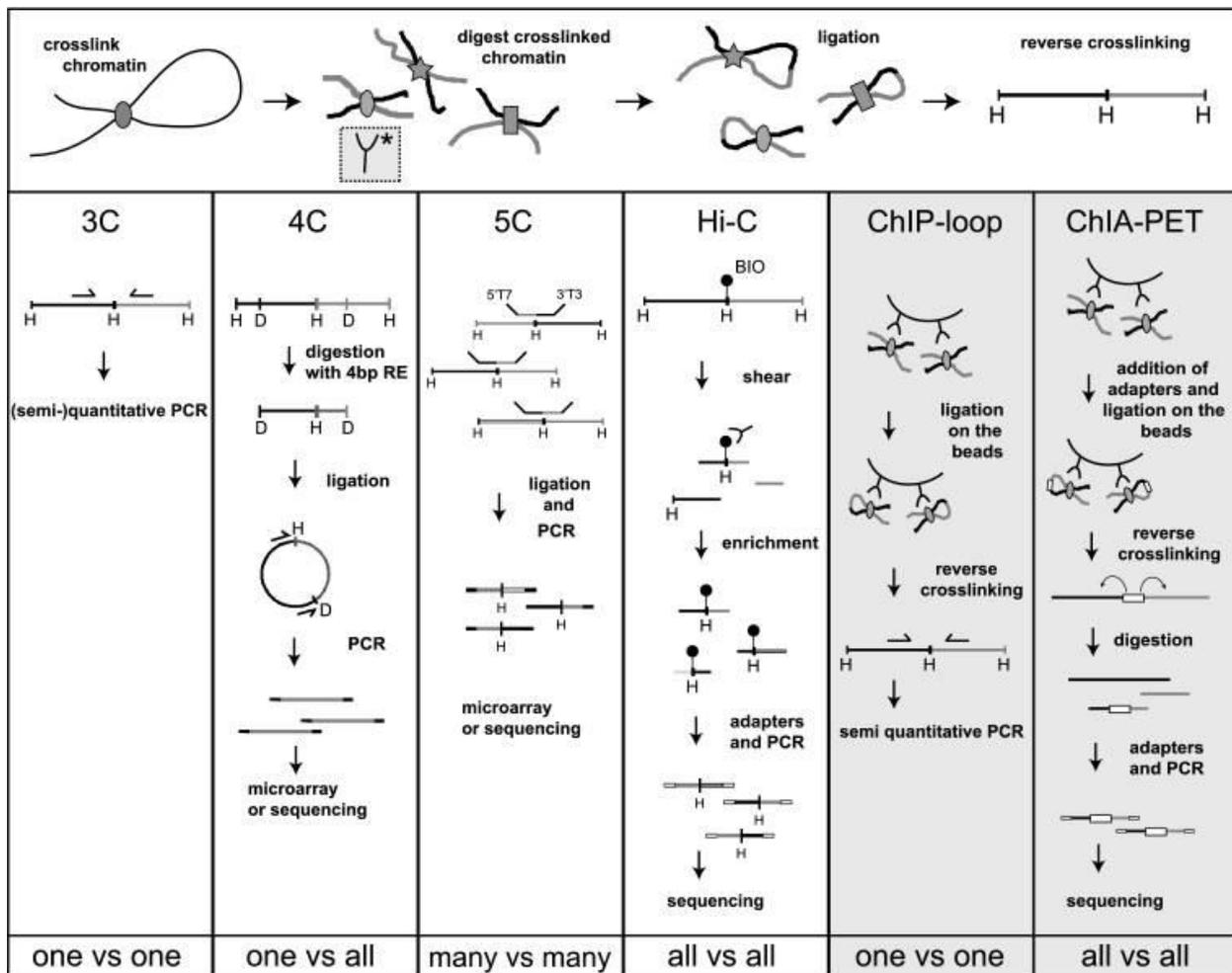


Figure 30 : Panel de plusieurs techniques d'analyse de la conformation spatiale de l'ADN (286).

Pour tenter d'étudier l'impact de certains CNV sur la conformation spatiale de la chromatine, nous avons choisi d'utiliser une technique appelée NG-CaptureC (287) que nous avons modifié, en remplaçant la première étape de digestion par enzyme de restriction par une étape de sonication (**Figure 31**). Cette technique a pour avantage de permettre une analyse de type « 1 contre tous », en étudiant l'ensemble des interactions entre une région du génome et le reste du génome. Nous l'avons modifiée pour ne pas être dépendant des sites de coupures des enzymes de restriction, et pouvoir analyser les interactions avec toutes les zones de notre choix. Par ailleurs, cette technique permet de quantifier les interactions, en s'appuyant sur la profondeur de séquençage, en reads, pour chacune des interactions répertoriées. En effet, plus une interaction sera représentée, plus celle-ci sera séquencée.

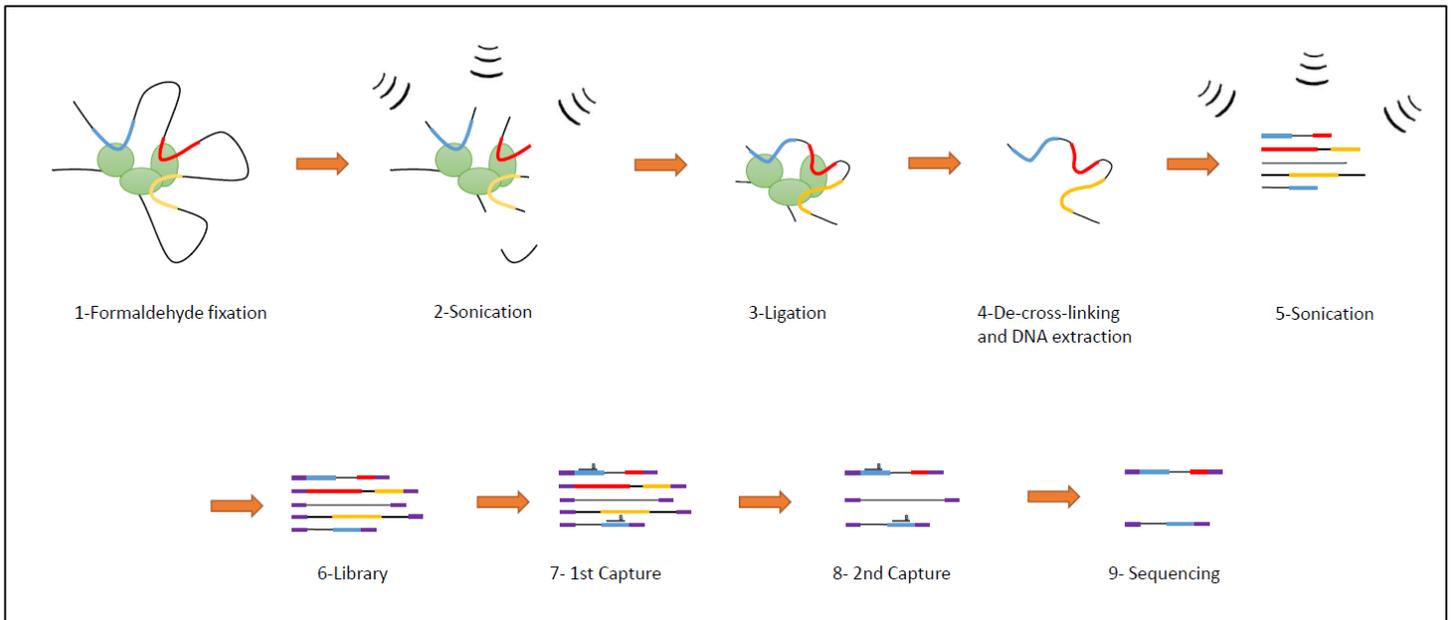


Figure 31 : Les différentes étapes de notre méthode se basant sur celle du NG-CaptureC.

Pour tester notre technique avant de l'appliquer à nos CNV d'intérêt, nous avons choisi de nous intéresser au gène *FTO*, et plus particulièrement à une zone intronique comprise entre le premier et deuxième exon (**Figure 32**). Cette zone possède plusieurs variants nucléotidiques (SNP) associés avec un IMC plus élevé que chez les sujets non mutés qui avaient été identifiés par les premières études GWAS. Les sujets homozygotes mutés pèsent en moyenne 3kg de plus que des sujets non mutés (288).

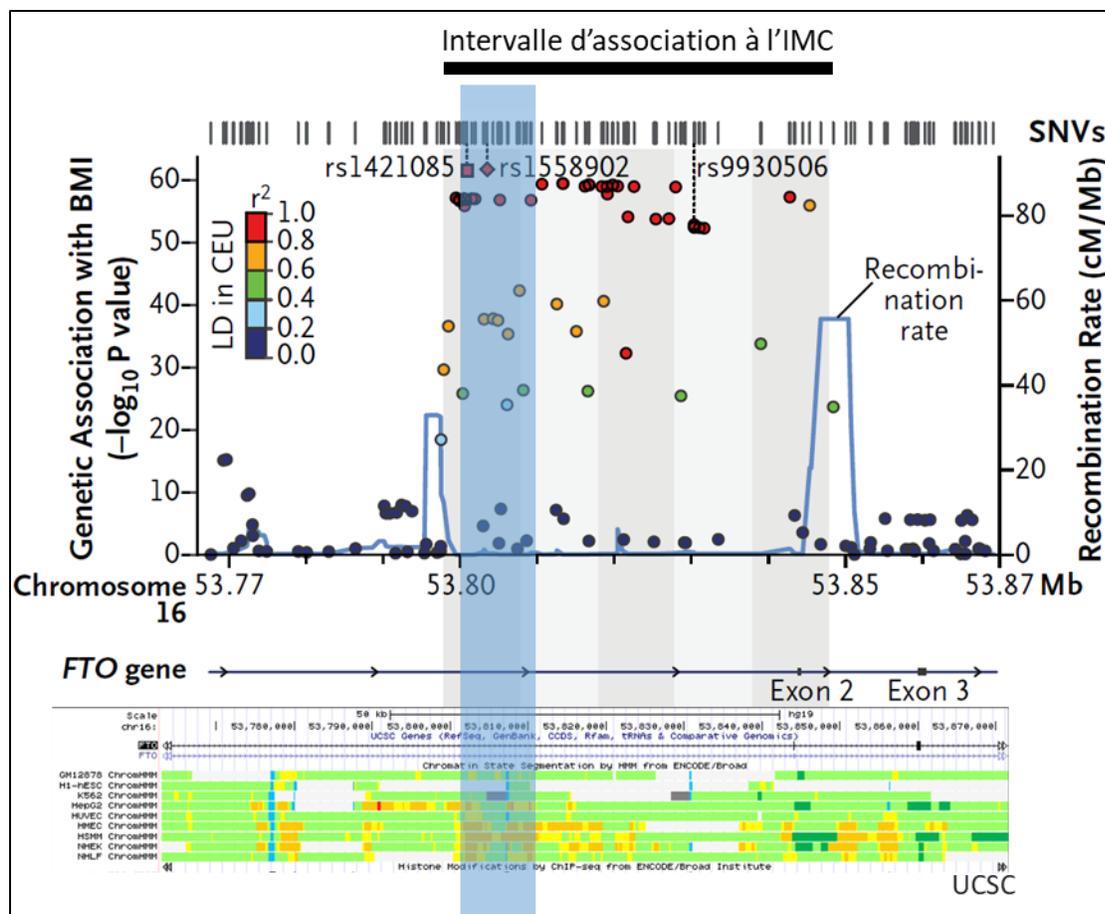


Figure 32 : Intervalle d'association à l'IMC au niveau du gène *FTO*. Surligné en bleu, notre zone d'intérêt pour l'étude preuve de concept (288).

Nous avons choisi, pour notre premier test, de nous concentrer sur le SNP rs1421085, l'un des plus étudiés et associés à l'augmentation de l'IMC, ainsi que sur le promoteur de *FTO*, de façon à comparer nos résultats avec ceux obtenus chez la souris ((289) et **Figure 33**). En effet les interactions entre les SNPs associés à l'IMC et les gènes entourant *FTO* font l'objet de controverses. Des études montrent que les gènes *IRX3* et *IRX5* sont les cibles des interactions avec les SNPs de *FTO* permettant d'expliquer cette association (288), et d'autres études s'orientent vers d'autres gènes, comme *RPGRIP1L* (290).

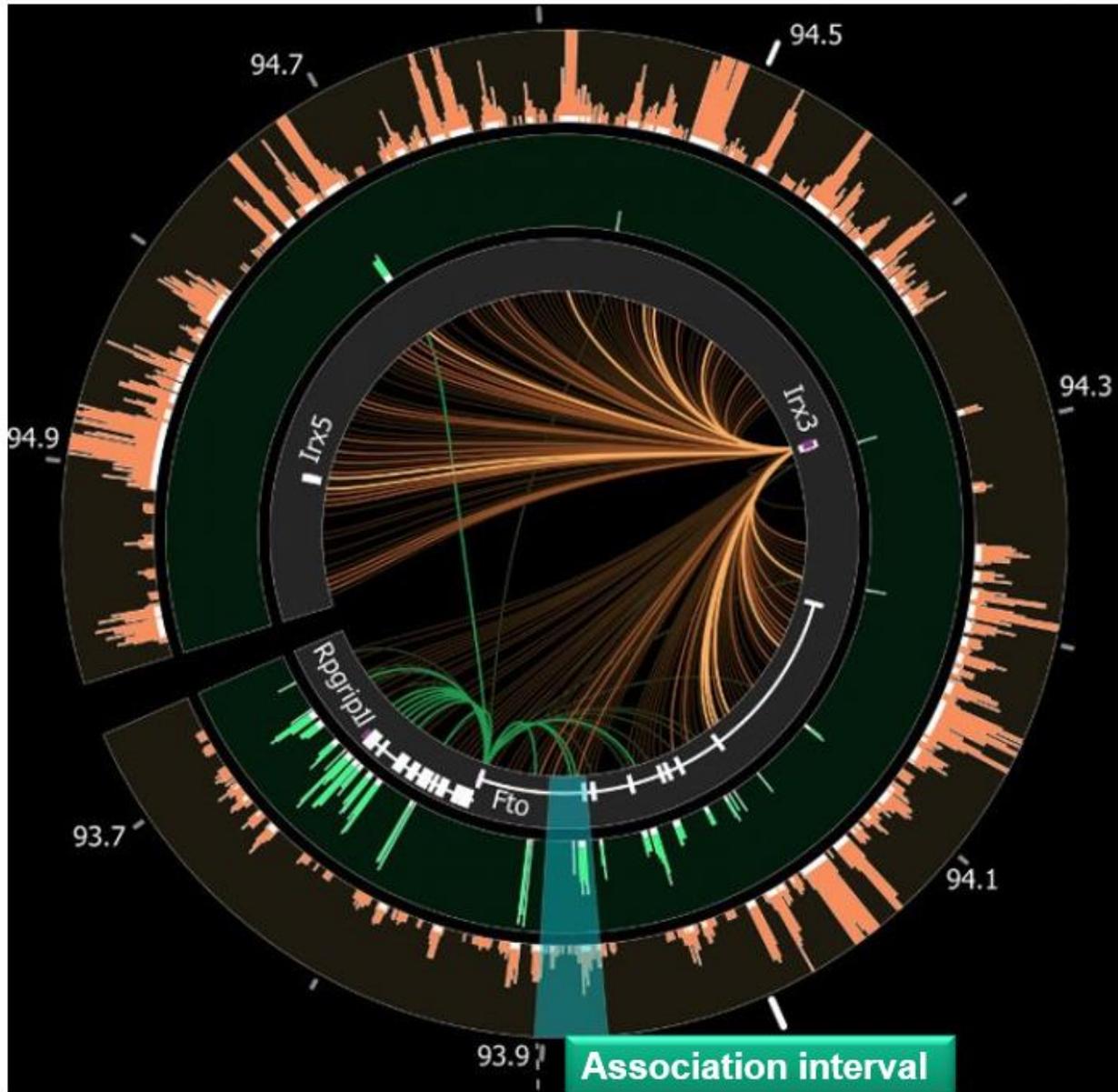


Figure 33 : Associations avec le promoteur du gène *fto* chez la souris (289). Chacun des liens présents dans le rond central représente une interaction entre deux points dans la zone du génome comprenant le gène *FTO*, et les barres verticales au niveau des cercles extérieurs représentent l'intensité de ces interactions.

Matériel et méthode :

Pour tester notre technique, il est nécessaire d'utiliser des cellules fraîches, de façon à ce que l'ADN soit toujours dans sa conformation fonctionnelle. Nous avons donc commencé à travailler

sur une lignée humaine de cellules de foie, les IHH, dans lesquelles *FTO* est connu pour être exprimé (**Figure 34**).

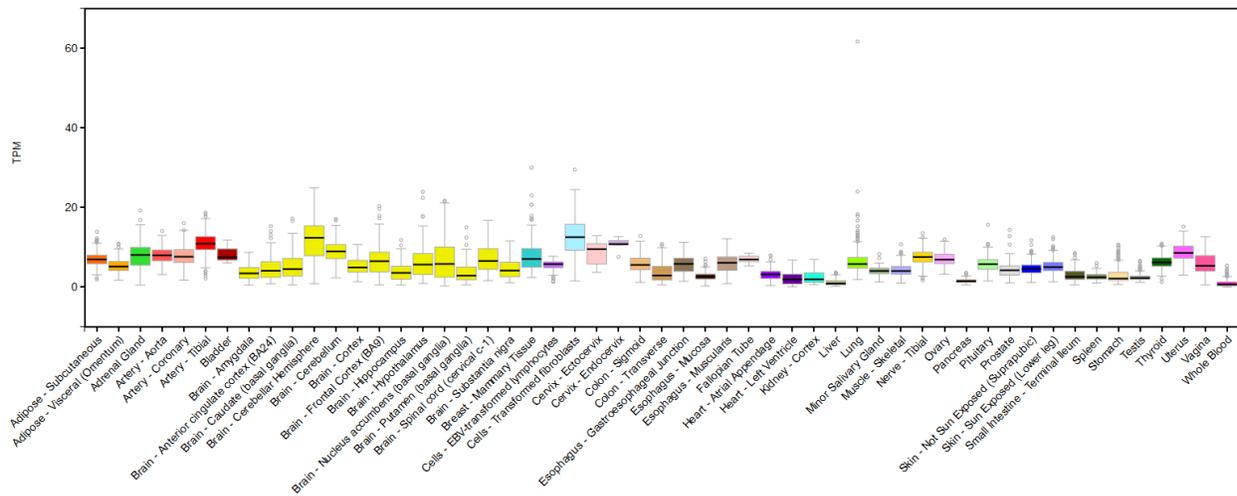


Figure 34 : Expression du gène *FTO* dans différents tissus, représenté en nombre de transcrits par million.

La conformation de l'ADN de ces cellules est ensuite fixée, ou cross-linkée par ajout de formaldéhyde, en suivant le protocole du kit *truChIP Chromatin Shearing Kit with Formaldehyde* de la société Covaris. Les cellules et l'ADN qu'elles contiennent ont ensuite été fragmentés sur l'appareil de sonication E220 de la société Covaris. Une fois l'ADN fragmenté, nous avons procédé à l'étape de ligation par ajout de DNA ligase T4, puis de suppression du cross-link par ajout de protéinase K, et enfin purifié cet ADN sur colonne grâce au kit NucleoSpin Tissue de la société Macherey-Nagel.

Nous avons ensuite effectué une librairie séquençable à partir de l'ADN obtenu. Nous avons utilisé 1µg d'ADN que nous avons fragmenté à l'aide d'un appareil de sonication Covaris E220 de la société Covaris, pour obtenir des fragments de taille compatible avec le séquençage d'environ 200 bases. Les fragments obtenus ont ensuite subi une réparation des extrémités soniquées, pour permettre ensuite leur ligation aux adaptateurs de séquençage Illumina grâce au kit *HTP library preparation kit* de la société KAPA. Les échantillons ont ensuite été amplifiés par *polymerase chain reaction*, ou PCR, en utilisant des amorces complémentaires des adaptateurs de séquençage. Les échantillons amplifiés ont ensuite été filtrés lors d'une étape de sélection de taille pour ne garder

que ceux compatibles avec le séquençage et éliminer les éléments parasites, puis quantifiés par l'appareil LabChip GX de Perkin Elmer (**Figure 19**).

Nous avons ensuite procédé aux 2 captures successives à l'aide de sondes d'ADN biotynilées de 120 bases ciblant le promoteur de *FTO* et le SNP rs1421085. Enfin à l'issue de ces captures, l'ADN a été purifié grâce au kit de lavage *SeqCap Hybridization and Wash Kit* sur l'automate Agilent *Bravo Automated Liquid Handling Platform*, de façon à éliminer tous les fragments non capturés. Ces captures ont ensuite été amplifiées grâce au kit Nimblegen *Hifi HotStart ReadyMix* et quantifiées sur les appareils Caliper *LifeScience Labchip GX* (Perkin Elmer) et Qubit *fluorometric quantitation assays* (Thermo Fisher Scientific) (**Figure 20**).

L'étape de séquençage a été effectuée sur le séquenceur MiSeq de la société Illumina.

Analyse Bioinformatique :

Une fois le séquençage effectué, il a été nécessaire de traiter des *reads* hybrides sans point de liaison connu, c'est-à-dire des *reads* constitués de deux séquences d'ADN correspondant à deux zones distinctes du génome, sans point de liaison défini entre eux. Il est difficile pour un aligneur de réussir à aligner ces *reads* sur le génome, étant donné qu'ils correspondent à deux zones distinctes, sans point de séparation défini. Nous avons donc choisi d'utiliser un logiciel nommé w4Cseq (291), spécifiquement développé pour gérer cette problématique.

Résultats préliminaires :

Nos premiers tests se sont intéressés au promoteur du gène *FTO*, de façon à étalonner la technique. Nos premiers résultats ont montré un problème d'enrichissement en régions interagissant avec le promoteur de *FTO*, ce qui montrait un souci lors de l'étape de capture. Nous avons par la suite amélioré nos paramètres (passage de 24 à 72h de capture, et de 47 à 65°C) et obtenu une nette amélioration (**Figure 35**). Nous avons pu observer plusieurs centaines d'interactions avec l'ensemble du génome, dont l'analyse est toujours en cours.

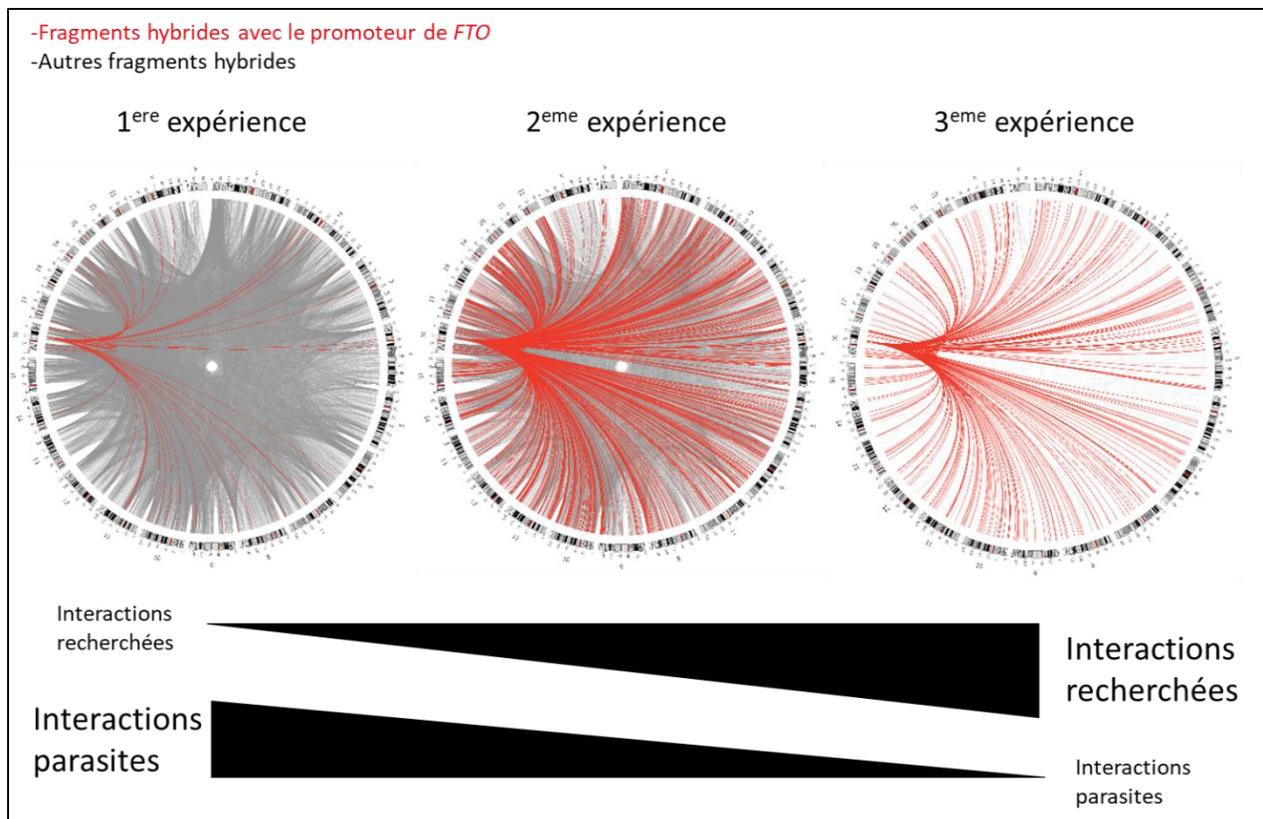


Figure 35 : Premiers résultats de notre protocole de NG-CaptureC sur le promoteur de *FTO*. Représentation via le logiciel Circos. Les liens rouges représentent les interactions entre le promoteur du gène *FTO* et le reste du génome, et les liens gris ceux ne comprenant pas le promoteur du gène *FTO*.

Nous nous sommes par la suite intéressés au SNP rs1421085 et à ses éventuelles interactions avec le reste du génome. D'après un article du *New England Journal of Medicine*, ce SNP est connu pour interagir avec les gènes *IRX3* et *IRX5* (288). Nous n'avons pas retrouvé ces interactions parmi nos résultats. Néanmoins, nous avons trouvé plusieurs interactions avec des gènes, comme le gène codant pour la Pleckstrine (*PLEK*), pour la 5-Oxoprolinase (*OPLAH*) ou pour la Zinc Finger Protein 251 (*ZNF251*). Nous avons également pu observer plusieurs centaines d'interaction entre cette zone et le reste du génome, mais en des points différents de celles concernant le promoteur *FTO*, avec beaucoup plus d'interactions inter-chromosomiques que dans la plupart des études (**Figure 36**). Ces résultats ont ensuite été reproduits avec succès dans les deux cas, avec conservation de la plupart des interactions de forte intensité, ce qui montre la robustesse de la technique.

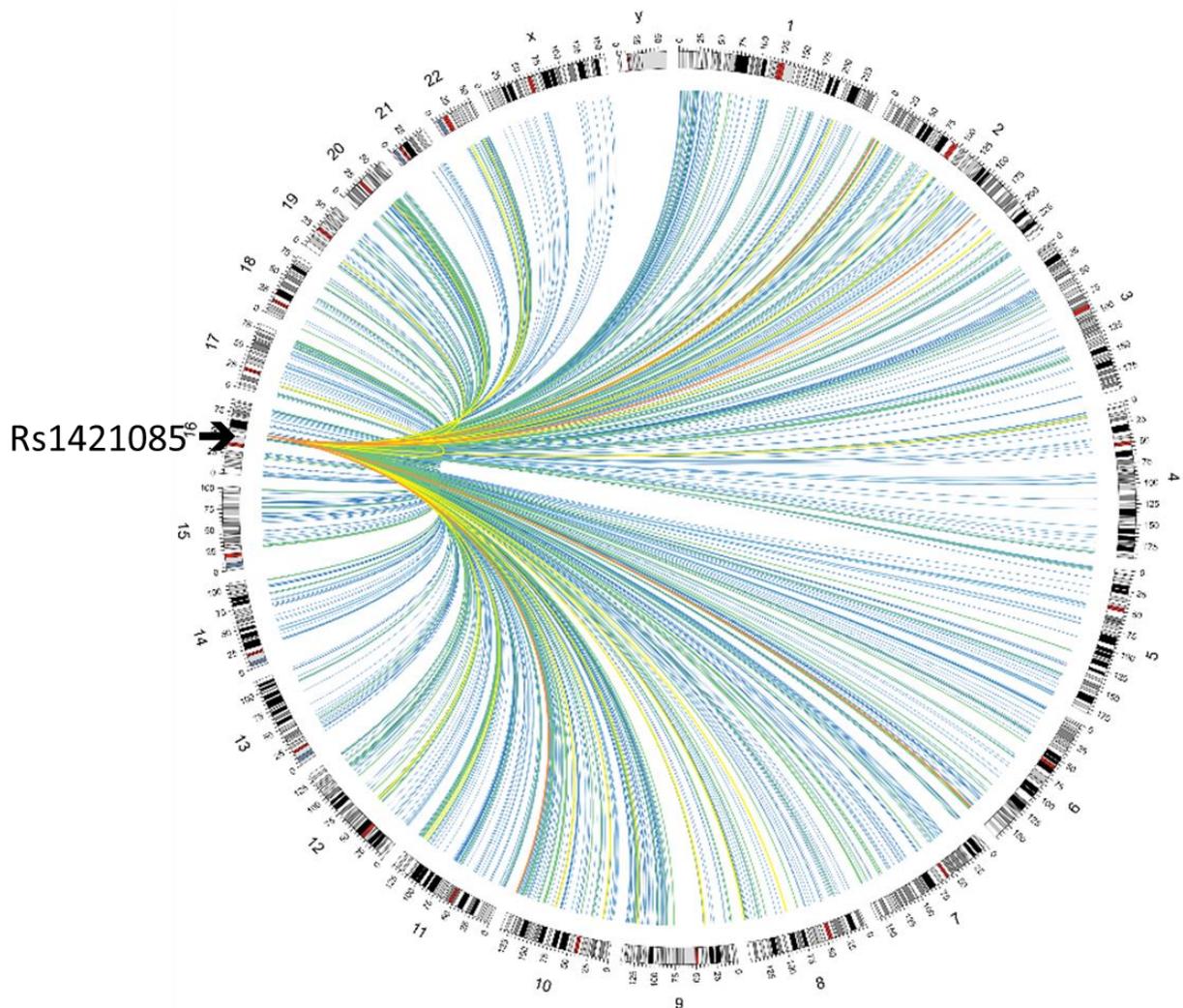


Figure 36 : Représentation des interactions entre le SNP rs1421085 et le reste du génome. Les couleurs représentent le nombre de *reads* par interaction, avec moins de 5 *reads* en bleu, de 5 à 10 en vert, de 10 à 25 en jaune, de 25 à 50 en orange, et plus de 50 en rouge.

Nous avons ensuite cherché à confirmer nos interactions avec le SNP rs1421085, pour confirmer les résultats, en recherchant les interactions à partir des 4 cibles de rs1421085 à la p-value d'interaction avec la plus basse. Trois d'entre-elles ont lieu au niveau de gènes (*PLEK*, *OPLAH* et *ZNF251*) du chromosome 8, et la dernière dans une zone non codante du chromosome 10 (**Figure 37**).

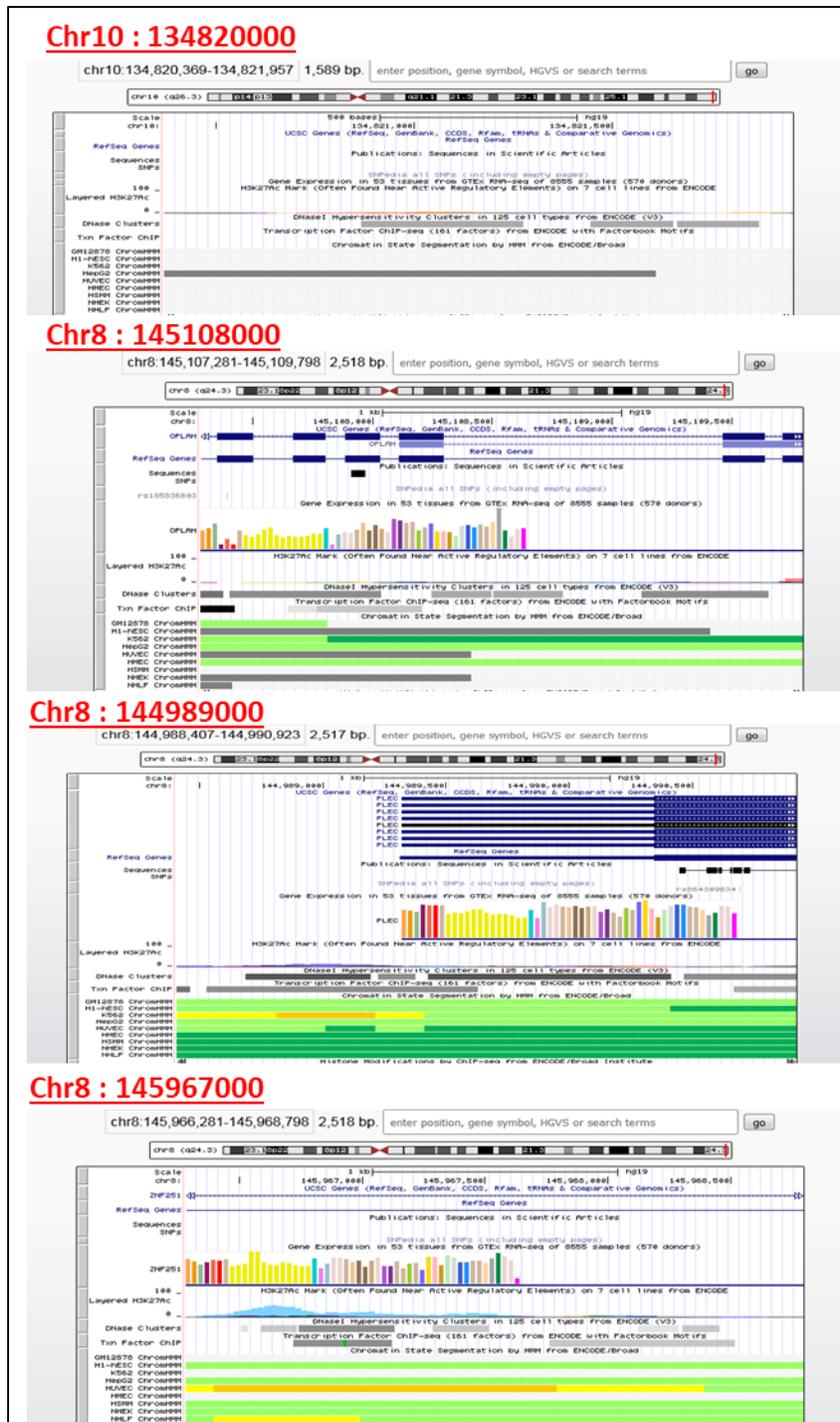


Figure 37 : Zones d'interaction avec rs1421085 à la p-value la plus basse. Les coordonnées sont basées sur le génome en version hg19.

Nous sommes parvenus à retrouver une interaction avec rs1421085 dans à partir de 2 des 4 cibles de rs1421085 : celle avec le gène *ZNF251* et la zone non codante du chromosome 10 (**Figure 38**).

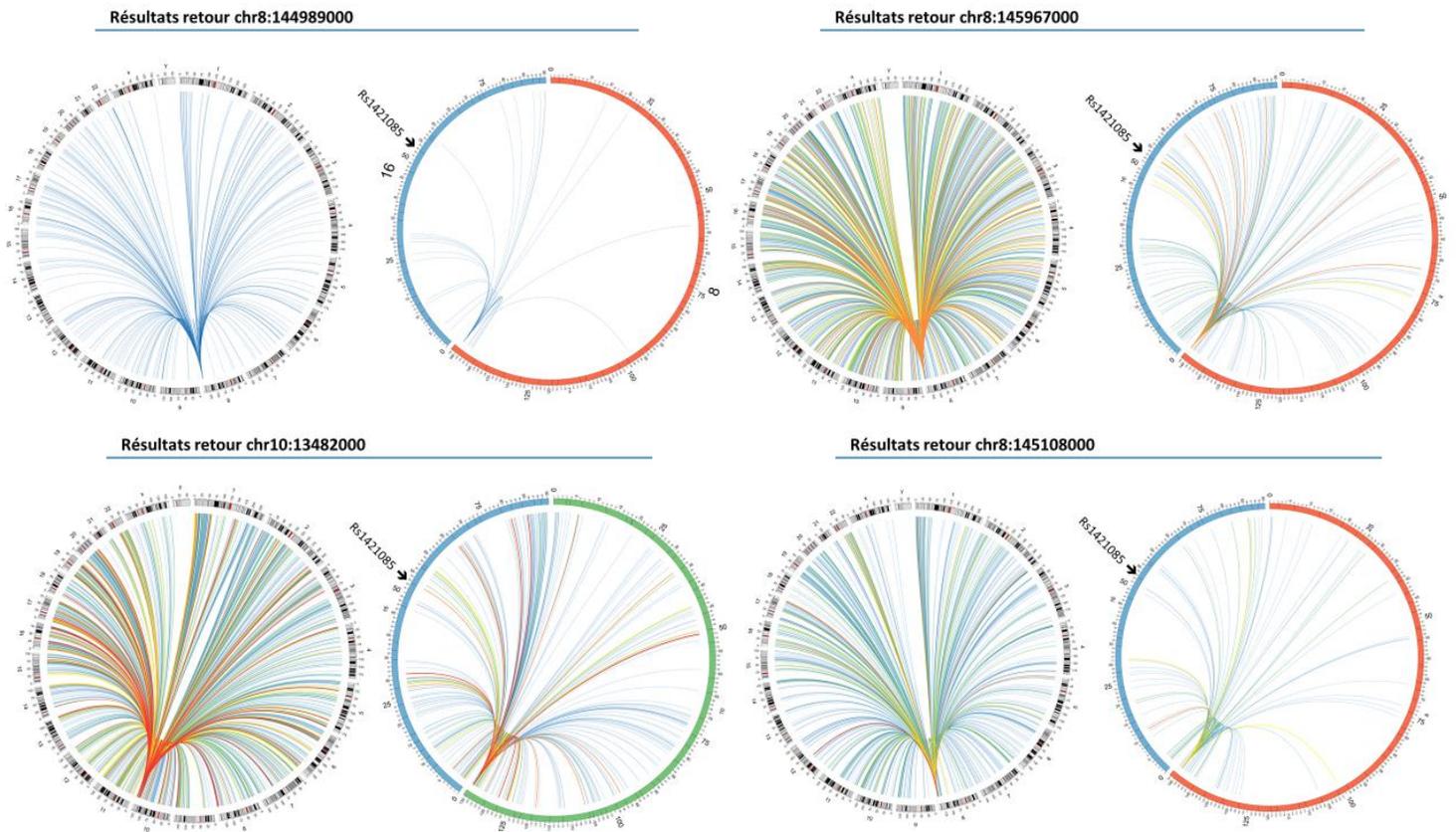


Figure 38 : Représentation des interactions entre les 4 zones d'interactions avec le SNP rs1421085 à la p-value la plus basse et le reste du génome. Les couleurs représentent le nombre de *reads* par interaction, avec moins de 5 *reads* en bleu, de 5 à 10 en vert, de 10 à 25 en jaune, de 25 à 50 en orange, et plus de 50 en rouge. On peut voir qu'une interaction avec le SNP rs1421085 n'a été trouvée que pour deux de ces 4 zones.

Discussion :

Nos résultats montrent que les interactions spatiales de la chromatine sont complexes et nombreuses. Nous parvenons à détecter davantage d'interactions que lors des études précédentes ayant concerné *FTO*, en particulier concernant les interactions inter-chromosomiques. Cela souligne les difficultés d'interprétation des GWAS et de l'attribution des effets d'un SNP, qui ont tendance à se focaliser sur les gènes les plus proches, alors que des interactions longues distances peuvent avoir lieu. Néanmoins, nous ne retrouvons pas les interactions entre le SNP rs1421085 et les gènes *IRX3* et *IRX5*. De plus, certaines de ces interactions ont été retrouvées dans les deux sens, ce qui souligne la robustesse de cette étude

pilote. Concernant les deux interactions non retrouvées dans les deux sens, il est possible qu'un séquençage insuffisant soit la cause de leur absence. Les prochaines étapes de ce travail seront d'effectuer ces analyses sur d'autres tissus, comme des préadipocytes primaires humains, des lignées de neurones dopaminergiques grâce à une collaboration avec le Docteur David Devos (Lille), ou des neurones hypothalamiques grâce à une collaboration avec le Docteur Giles Yeo (Cambridge). Cela nous permettra de valider notre méthode et ses résultats, avant de démarrer l'études des possibles variations de structures de la chromatine engendrées par les CNV, ou les mutations ponctuelles, sur les tissus de nos patients.

Bibliographie :

1. Williams EP, Mesidor M, Winters K, Dubbert PM, W, B S. Overweight and obesity: prevalence, consequences, and causes of a growing public health problem. *Curr Obes Rep.* 2015;4:363–370.
2. Alberti KG. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity Circulation. 2009;120:1640–1645.
3. Sellayah D, Cagampang FR, Cox RD. On the evolutionary origins of obesity: a new hypothesis. *Endocrinology* 155. 2014.
4. Martinez JA. Body-weight regulation: causes of obesity. *Proc Nutr Soc.* 2000;59:337–345.
5. McAllister EJ. Ten putative contributors to the obesity epidemic. *Crit. Rev. Food Sci. Nutr.* 2009;49(868–913).
6. Fosbol MO, Zerahn B. Contemporary methods of body composition measurement. *Clin. Physiol Funct Imaging.* 2015;35:81–97.
7. Javed A, Jumean M, Murad MH, Okorodudu D, Kumar S, Somers VK, et al. Diagnostic performance of body mass index to identify obesity as defined by body adiposity in children and adolescents: a systematic review and meta-analysis. *Pediatr Obes.* juin 2015;10(3):234-44.
8. González-Muniesa P, Martínez-González M-A, Hu FB, Després J-P, Matsuzawa Y, Loos RJF, et al. Obesity. *Nature Reviews Disease Primers.* 15 juin 2017;3:17034.
9. Urdampilleta A, González-Muniesa P, Portillo MP, Martínez JA. Usefulness of combining intermittent hypoxia and physical exercise in the treatment of obesity. *J Physiol Biochem.* juin 2012;68(2):289-304.
10. Dobbs R. How the world could better fight obesity. McKinsey: Global Institute; 2014.
11. Malik VS, Willett WC, Hu FB. Global obesity: trends, risk factors and policy implications. *Nat Rev Endocrinol.* 2013;9:13–27.
12. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *The Lancet.* 2 avr 2016;387(10026):1377-96.
13. Obesity and overweight [Internet]. World Health Organization. [cité 17 juill 2018]. Disponible sur: <http://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>
14. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet.* 30 août 2014;384(9945):766-81.

15. Nazare J-A, Smith JD, Borel A-L, Haffner SM, Balkau B, Ross R, et al. Ethnic influences on the relations between abdominal subcutaneous and visceral adiposity, liver fat, and cardiometabolic risk profile: the International Study of Prediction of Intra-Abdominal Adiposity and Its Relationship With Cardiometabolic Risk/Intra-Abdominal Adiposity. *Am J Clin Nutr.* oct 2012;96(4):714-26.
16. Finkelstein EA, Khavjou OA, Thompson H, Trogdon JG, Pan L, Sherry B, et al. Obesity and severe obesity forecasts through 2030. *Am J Prev Med.* juin 2012;42(6):563-70.
17. Bhupathiraju SN, Hu FB. Epidemiology of obesity and diabetes and their cardiovascular complications. *Circ Res.* 2016;118:1723–1735.
18. Ahluwalia N, Dalmaso P, Rasmussen M, Lipsky L, Currie C, Haug E, et al. Trends in overweight prevalence among 11-, 13- and 15-year-olds in 25 countries in Europe, Canada and USA from 2002 to 2010. *Eur J Public Health.* avr 2015;25 Suppl 2:28-32.
19. Hall KD, Sacks G, Chandramohan D, Chow CC, Wang YC, Gortmaker SL, et al. Quantification of the effect of energy imbalance on bodyweight. *Lancet.* 27 août 2011;378(9793):826-37.
20. Vague J. The degree of masculine differentiation of obesities: a factor determining predisposition to diabetes, atherosclerosis, gout, and uric calculous disease. *Am J Clin Nutr.* févr 1956;4(1):20-34.
21. Saeed S, Bonnefond A, Manzoor J, Shabbir F, Ayesha H, Philippe J, et al. Genetic variants in LEP, LEPR, and MC4R explain 30% of severe obesity in children from a consanguineous population. *Obesity (Silver Spring).* 2015;23(8):1687-95.
22. Flier JS. Obesity wars: molecular progress confronts an expanding epidemic. *Cell.* 23 janv 2004;116(2):337-50.
23. Bes-Rastrollo M. A prospective study of eatingaway-from-home meals and weight gain in aMediterranean population: the SUN (SeguimientoUniversidad de Navarra) cohort. *Public Health.* 2010;13:1356–1363.
24. Vandevijvere S, Chow CC, Hall KD, Umali E. S, A B. Increased food energy supply as a major driver of the obesity epidemic: a global analysis. *Bull. World Health Organ.* 2015;93:446–456.
25. Sayon-Orea C. Association between sleeping hours and siesta and the risk of obesity: the SUNMediterranean cohort. *Obes. Facts.* 6:337–347.
26. Ludwig DS. Lifespan weighed down by diet. *Vol. 315.* 2016. 2269–2270 p.
27. Razquin C. A 3 years follow-up of aMediterranean diet rich in virgin olive oil is associated with high plasma antioxidant capacity and reduced body weight gain. *Eur J Clin Nutr.* 2009;63(1387–1393).
28. Estruch R, Ros E, Salas-Salvadó J, Covas M-I, Corella D, Arós F, et al. Retraction and Republication: Primary Prevention of Cardiovascular Disease with a Mediterranean Diet. *N Engl J Med* 2013;368:1279-90. *N Engl J Med.* 21 2018;378(25):2441-2.
29. Smith JD. Changes in intake of protein foods, carbohydrate amount and quality, and long-term weight change: results from 3 prospective cohorts. *Am. J Clin Nutr.* 2015;101:1216–1224.

30. Tobias DK. Effect of low-fat diet interventions versus other diet interventions on long-term weight change in adults: a systematic review and meta-analysis. *Lancet Diabetes Endocrinol.* Lancet Diabetes Endocrinol. 3.
31. Hanley B, Dijane J, Fewtrell M, Grynberg A, Hummel S, Junien C, et al. Metabolic imprinting, programming and epigenetics - a review of present priorities and future opportunities. *Br J Nutr.* juill 2010;104 Suppl 1:S1-25.
32. Lin X, Lim IY, Wu Y, Teh AL, Chen L, Aris IM, et al. Developmental pathways to adiposity begin before birth and are influenced by genotype, prenatal environment and epigenome. *BMC Med.* 07 2017;15(1):50.
33. Yu ZB, Han SP, Zhu GZ, Zhu C, Wang XJ, Cao XG, et al. Birth weight and subsequent risk of obesity: a systematic review and meta-analysis. *Obes Rev.* juill 2011;12(7):525-42.
34. Druet C, Stettler N, Sharp S, Simmons RK, Cooper C, Smith GD, et al. Prediction of childhood obesity by infancy weight gain: an individual-level meta-analysis. *Paediatr Perinat Epidemiol.* janv 2012;26(1):19-26.
35. Arenz S, Ruckerl R, Koletzko B, von Kries R. Breast-feeding and childhood obesity--a systematic review. *Int J Obes Relat Metab Disord.* oct 2004;28(10):1247-56.
36. Weber M, Grote V, Closa-Monasterolo R, Escribano J, Langhendries J-P, Dain E, et al. Lower protein content in infant formula reduces BMI and obesity risk at school age: follow-up of a randomized trial. *Am J Clin Nutr.* mai 2014;99(5):1041-51.
37. Ohlsson C, Lorentzon M, Norjavaara E, Kindblom JM. Age at adiposity rebound is associated with fat mass in young adult males--the GOOD study. *PLoS ONE.* 2012;7(11):e49404.
38. Prentice P, Viner RM. Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. *Int J Obes (Lond).* août 2013;37(8):1036-43.
39. Rolland-Cachera MF, Deheeger M, Maillot M. B, F. Early adiposity rebound: causes and consequences for obesity in children and adults. *Int. J Obes (Lond).* 2006;30:11-17.
40. Wang YC, McPherson K, Marsh T, Gortmaker SL, Brown M. Health and economic burden of the projected obesity trends in the USA and the UK. *Lancet.* 2011;815-825.
41. Jauch-Chara K, Oltmanns KM. Obesity--a neuropsychological disease? Systematic review and neuropsychological model. *Prog Neurobiol.* mars 2014;114:84-101.
42. Global BMI Mortality Collaboration null, Di Angelantonio E, Bhupathiraju S, Wormser D, Gao P, Kaptoge S, et al. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet.* 20 2016;388(10046):776-86.
43. Flegal KM, Kit BK, Orpana H, Graubard BI. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *JAMA.* 2 janv 2013;309(1):71-82.

44. Després J-P, Lemieux I. Abdominal obesity and metabolic syndrome. *Nature*. 14 déc 2006;444(7121):881-7.
45. Sacks H, Symonds ME. Anatomical locations of human brown adipose tissue: functional relevance and implications in obesity and type 2 diabetes. *Diabetes*. juin 2013;62(6):1783-90.
46. Matsuzawa Y. The metabolic syndrome and adipocytokines. *FEBS Lett*. 2006;580:2917–2921.
47. Gesta S, Tseng Y-H, Kahn CR. Developmental origin of fat: tracking obesity to its source. *Cell*. 19 oct 2007;131(2):242-56.
48. Shulman GI. Ectopic fat in insulin resistance, dyslipidemia, and cardiometabolic disease. *N Engl J Med*. 18 sept 2014;371(12):1131-41.
49. Karpe F, Pinnick KE. Biology of upper-body and lower-body adipose tissue--link to whole-body phenotypes. *Nat Rev Endocrinol*. févr 2015;11(2):90-100.
50. Rosen ED, Spiegelman BM. What we talk about when we talk about fat. *Cell*. 16 janv 2014;156(1-2):20-44.
51. Ross R, Berentzen T, Bradshaw AJ, Janssen I, Kahn HS, Katzmarzyk PT, et al. Does the relationship between waist circumference, morbidity and mortality depend on measurement protocol for waist circumference? *Obes Rev*. juill 2008;9(4):312-25.
52. Neeland IJ, Turer AT, Ayers CR, Berry JD, Rohatgi A, Das SR, et al. Body fat distribution and incident cardiovascular disease in obese adults. *J Am Coll Cardiol*. 19 mai 2015;65(19):2150-1.
53. Patni N, Garg A. Congenital generalized lipodystrophies--new insights into metabolic dysfunction. *Nat Rev Endocrinol*. sept 2015;11(9):522-34.
54. Camilleri M. Peripheral mechanisms in appetite regulation. *Gastroenterology*. mai 2015;148(6):1219-33.
55. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, et al. Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science*. 6 sept 2013;341(6150):1241214.
56. Elks CE, den Hoed M, Zhao JH, Sharp SJ, Wareham NJ, Loos RJF, et al. Variability in the heritability of body mass index: a systematic review and meta-regression. *Front Endocrinol (Lausanne)*. 2012;3:29.
57. Myers MG, Leibel RL. Lessons From Rodent Models of Obesity. In: De Groot LJ, Chrousos G, Dungan K, Feingold KR, Grossman A, Hershman JM, et al., éditeurs. *Endotext* [Internet]. South Dartmouth (MA): MDText.com, Inc.; 2000 [cité 4 sept 2018]. Disponible sur: <http://www.ncbi.nlm.nih.gov/books/NBK279123/>
58. van der Klaauw AA, Farooqi IS. The hunger genes: pathways to obesity. *Cell*. 26 mars 2015;161(1):119-32.
59. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*. 11 mai 2007;316(5826):889-94.

60. Gerken T, Girard CA, Tung Y-CL, Webby CJ, Saudek V, Hewitson KS, et al. The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science*. 30 nov 2007;318(5855):1469-72.
61. Fischer J, Koch L, Emmerling C, Vierkotten J, Peters T, Brüning JC, et al. Inactivation of the Fto gene protects from obesity. *Nature*. 16 avr 2009;458(7240):894-8.
62. Church C, Moir L, McMurray F, Girard C, Banks GT, Teboul L, et al. Overexpression of Fto leads to increased food intake and results in obesity. *Nat Genet*. déc 2010;42(12):1086-92.
63. Claussnitzer M, Hui C-C, Kellis M. FTO Obesity Variant and Adipocyte Browning in Humans. *N Engl J Med*. 14 2016;374(2):192-3.
64. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 12 févr 2015;518(7538):197-206.
65. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Mägi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature*. 12 févr 2015;518(7538):187-96.
66. Ravelli GP, Stein ZA, Susser MW. Obesity in young men after famine exposure in utero and early infancy. *N Engl J Med*. 12 août 1976;295(7):349-53.
67. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aïssi D, Wahl S, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet*. 7 juin 2014;383(9933):1990-8.
68. Jiang X, Xue M, Fu Z, Ji C, Guo X, Zhu L, et al. Insight into the Effects of Adipose Tissue Inflammation Factors on miR-378 Expression and the Underlying Mechanism. *Cellular Physiology and Biochemistry*. 2014;33(6):1778-88.
69. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature*. 05 2017;541(7635):81-6.
70. Mendelson MM, Marioni RE, Joehanes R, Liu C, Hedman ÅK, Aslibekyan S, et al. Association of Body Mass Index with DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach. *PLoS Med*. janv 2017;14(1):e1002215.
71. Pigeyre M, Romon M. [Genetic obesities]. *Ann Endocrinol (Paris)*. déc 2007;68(6):430-7.
72. Coleman DL, Hummel KP. Effects of parabiosis of normal with genetically diabetic mice. *Am J Physiol*. nov 1969;217(5):1298-304.
73. Considine RV, Sinha MK, Heiman ML, Kriauciunas A, Stephens TW, Nyce MR, et al. Serum immunoreactive-leptin concentrations in normal-weight and obese humans. *N Engl J Med*. 1 févr 1996;334(5):292-5.
74. Ahima RS, Prabakaran D, Mantzoros C, Qu D, Lowell B, Maratos-Flier E, et al. Role of leptin in the neuroendocrine response to fasting. *Nature*. 18 juill 1996;382(6588):250-2.
75. Elmquist JK, Bjørbaek C, Ahima RS, Flier JS, Saper CB. Distributions of leptin receptor mRNA isoforms in the rat brain. *J Comp Neurol*. 15 juin 1998;395(4):535-47.

76. Cone RD. Anatomy and regulation of the central melanocortin system. *Nat Neurosci.* mai 2005;8(5):571-8.
77. Huszar D, Lynch CA, Fairchild-Huntress V, Dunmore JH, Fang Q, Berkemeier LR, et al. Targeted disruption of the melanocortin-4 receptor results in obesity in mice. *Cell.* 10 janv 1997;88(1):131-41.
78. Montague CT, Farooqi IS, Whitehead JP, Soos MA, Rau H, Wareham NJ, et al. Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature.* 26 juin 1997;387(6636):903-8.
79. Farooqi IS, Wangensteen T, Collins S, Kimber W, Matarese G, Keogh JM, et al. Clinical and Molecular Genetic Spectrum of Congenital Deficiency of the Leptin Receptor. *N Engl J Med.* 18 janv 2007;356(3):237-47.
80. Clément K, Biebermann H, Farooqi IS, Van der Ploeg L, Wolters B, Poitou C, et al. MC4R agonism promotes durable weight loss in patients with leptin receptor deficiency. *Nat Med.* mai 2018;24(5):551-5.
81. Farooqi IS, Matarese G, Lord GM, Keogh JM, Lawrence E, Agwu C, et al. Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency. *J Clin Invest.* oct 2002;110(8):1093-103.
82. Farooqi IS, Bullmore E, Keogh J, Gillard J, O’Rahilly S, Fletcher PC. Leptin regulates striatal regions and human eating behavior. *Science.* 7 sept 2007;317(5843):1355.
83. Rosenbaum M, Sy M, Pavlovich K, Leibel RL, Hirsch J. Leptin reverses weight loss–induced changes in regional neural activity responses to visual food stimuli. *J Clin Invest.* 1 juill 2008;118(7):2583-91.
84. Hommel JD, Trinko R, Sears RM, Georgescu D, Liu Z-W, Gao X-B, et al. Leptin receptor signaling in midbrain dopamine neurons regulates feeding. *Neuron.* 21 sept 2006;51(6):801-10.
85. Rosenbaum M, Goldsmith R, Bloomfield D, Magnano A, Weimer L, Heymsfield S, et al. Low-dose leptin reverses skeletal muscle, autonomic, and neuroendocrine adaptations to maintenance of reduced weight. *J Clin Invest.* déc 2005;115(12):3579-86.
86. Minokoshi Y, Kim Y-B, Peroni OD, Fryer LGD, Müller C, Carling D, et al. Leptin stimulates fatty-acid oxidation by activating AMP-activated protein kinase. *Nature.* 17 janv 2002;415(6869):339-43.
87. Galgani JE, Greenway FL, Caglayan S, Wong M-L, Licinio J, Ravussin E. Leptin replacement prevents weight loss-induced metabolic adaptation in congenital leptin-deficient patients. *J Clin Endocrinol Metab.* févr 2010;95(2):851-5.
88. Ozata M, Ozdemir IC, Licinio J. Human leptin deficiency caused by a missense mutation: multiple endocrine defects, decreased sympathetic tone, and immune system dysfunction indicate new targets for leptin action, greater central than peripheral resistance to the effects of leptin, and spontaneous correction of leptin-mediated defects. *J Clin Endocrinol Metab.* oct 1999;84(10):3686-95.

89. Gibson WT, Farooqi IS, Moreau M, DePaoli AM, Lawrence E, O'Rahilly S, et al. Congenital leptin deficiency due to homozygosity for the Delta133G mutation: report of another case and evaluation of response to four years of leptin therapy. *J Clin Endocrinol Metab.* oct 2004;89(10):4821-6.
90. Rosenbaum M, Murphy EM, Heymsfield SB, Matthews DE, Leibel RL. Low dose leptin administration reverses effects of sustained weight-reduction on energy expenditure and circulating concentrations of thyroid hormones. *J Clin Endocrinol Metab.* mai 2002;87(5):2391-4.
91. Huvenne H, Le Beyec J, Pépin D, Alili R, Kherchiche PP, Jeannic E, et al. Seven novel deleterious LEPR mutations found in early-onset obesity: a Δ Exon6-8 shared by subjects from Reunion Island, France, suggests a founder effect. *J Clin Endocrinol Metab.* mai 2015;100(5):E757-766.
92. Strobel A, Issad T, Camoin L, Ozata M, Strosberg AD. A leptin missense mutation associated with hypogonadism and morbid obesity. *Nat Genet.* mars 1998;18(3):213-5.
93. Chan Y-M, Broder-Fingert S, Seminara SB. Reproductive functions of kisspeptin and Gpr54 across the life cycle of mice and men. *Peptides.* janv 2009;30(1):42-8.
94. Tschöp J, Nogueiras R, Haas-Lockie S, Kasten KR, Castañeda TR, Huber N, et al. CNS leptin action modulates immune response and survival in sepsis. *J Neurosci.* 28 avr 2010;30(17):6036-47.
95. Saeed S, Bonnefond A, Tamanini F, Mirza MU, Manzoor J, Janjua QM, et al. Loss-of-function mutations in ADCY3 cause monogenic severe obesity. *Nat Genet.* févr 2018;50(2):175-9.
96. Albuquerque D, Stice E, Rodríguez-López R, Manco L, Nóbrega C. Current review of genetics of human obesity: from molecular mechanisms to an evolutionary perspective. *Molecular Genetics and Genomics.* août 2015;290(4):1191-221.
97. Krude H, Biebermann H, Luck W, Horn R, Brabant G, Grüters A. Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. *Nat Genet.* juin 1998;19(2):155-7.
98. Lee YS, Challis BG, Thompson DA, Yeo GSH, Keogh JM, Madonna ME, et al. A POMC variant implicates beta-melanocyte-stimulating hormone in the control of human energy balance. *Cell Metab.* févr 2006;3(2):135-40.
99. Creemers JWM, Lee YS, Oliver RL, Bahceci M, Tuzcu A, Gokalp D, et al. Mutations in the amino-terminal region of proopiomelanocortin (POMC) in patients with early-onset obesity impair POMC sorting to the regulated secretory pathway. *J Clin Endocrinol Metab.* nov 2008;93(11):4494-9.
100. Kühnen P, Clément K, Wiegand S, Blankenstein O, Gottesdiener K, Martini LL, et al. Proopiomelanocortin Deficiency Treated with a Melanocortin-4 Receptor Agonist. *N Engl J Med.* 21 juill 2016;375(3):240-6.
101. Jackson RS, Creemers JW, Ohagi S, Raffin-Sanson ML, Sanders L, Montague CT, et al. Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. *Nat Genet.* juill 1997;16(3):303-6.

102. Jackson RS, Creemers JWM, Farooqi IS, Raffin-Sanson M-L, Varro A, Dockray GJ, et al. Small-intestinal dysfunction accompanies the complex endocrinopathy of human proprotein convertase 1 deficiency. *J Clin Invest.* nov 2003;112(10):1550-60.
103. Vaisse C, Clement K, Durand E, Hercberg S, Guy-Grand B, Froguel P. Melanocortin-4 receptor mutations are a frequent and heterogeneous cause of morbid obesity. *J Clin Invest.* juill 2000;106(2):253-62.
104. Dubern B, Bisbis S, Talbaoui H, Le Beyec J, Tounian P, Lacorte J-M, et al. Homozygous null mutation of the melanocortin-4 receptor and severe early-onset obesity. *J Pediatr.* juin 2007;150(6):613-7, 617.e1.
105. Santini F, Maffei M, Pelosini C, Salvetti G, Scartabelli G, Pinchera A. Melanocortin-4 receptor mutations in obesity. *Adv Clin Chem.* 2009;48:95-109.
106. Farooqi IS, Keogh JM, Yeo GSH, Lank EJ, Cheetham T, O'Rahilly S. Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N Engl J Med.* 20 mars 2003;348(12):1085-95.
107. Biebermann H, Castañeda TR, van Landeghem F, von Deimling A, Escher F, Brabant G, et al. A role for beta-melanocyte-stimulating hormone in human body-weight regulation. *Cell Metab.* févr 2006;3(2):141-6.
108. Nogueiras R, Wiedmer P, Perez-Tilve D, Veyrat-Durebex C, Keogh JM, Sutton GM, et al. The central melanocortin system directly controls peripheral lipid metabolism. *J Clin Invest.* nov 2007;117(11):3475-88.
109. Martinelli CE, Keogh JM, Greenfield JR, Henning E, van der Klaauw AA, Blackwood A, et al. Obesity due to melanocortin 4 receptor (MC4R) deficiency is associated with increased linear growth and final height, fasting hyperinsulinemia, and incompletely suppressed growth hormone secretion. *J Clin Endocrinol Metab.* janv 2011;96(1):E181-188.
110. Martin NM, Houston PA, Patterson M, Sajedi A, Carmignac DF, Ghatei MA, et al. Abnormalities of the somatotrophic axis in the obese agouti mouse. *Int J Obes (Lond).* mars 2006;30(3):430-8.
111. Xu B, Goulding EH, Zang K, Cepoi D, Cone RD, Jones KR, et al. Brain-derived neurotrophic factor regulates energy balance downstream of melanocortin-4 receptor. *Nat Neurosci.* juill 2003;6(7):736-42.
112. Yeo GSH, Connie Hung C-C, Rochford J, Keogh J, Gray J, Sivaramakrishnan S, et al. A de novo mutation affecting human TrkB associated with severe obesity and developmental delay. *Nat Neurosci.* nov 2004;7(11):1187-9.
113. Han JC, Liu Q-R, Jones M, Levinn RL, Menzie CM, Jefferson-George KS, et al. Brain-derived neurotrophic factor and obesity in the WAGR syndrome. *N Engl J Med.* 28 août 2008;359(9):918-27.
114. Hung C-CC, Luan J, Sims M, Keogh JM, Hall C, Wareham NJ, et al. Studies of the SIM1 gene in relation to human obesity and obesity-related traits. *Int J Obes (Lond).* mars 2007;31(3):429-34.
115. Michaud JL, Boucher F, Melnyk A, Gauthier F, Goshu E, Lévy E, et al. Sim1 haploinsufficiency causes hyperphagia, obesity and reduction of the paraventricular nucleus of the hypothalamus. *Hum Mol Genet.* 1 juill 2001;10(14):1465-73.

116. Bonnefond A, Raimondo A, Stutzmann F, Ghossaini M, Ramachandrapa S, Bersten DC, et al. Loss-of-function mutations in SIM1 contribute to obesity and Prader-Willi-like features. *J Clin Invest.* juill 2013;123(7):3037-41.
117. Ramachandrapa S, Raimondo A, Cali AMG, Keogh JM, Henning E, Saeed S, et al. Rare variants in single-minded 1 (SIM1) are associated with severe obesity. *J Clin Invest.* 1 juill 2013;123(7):3042-50.
118. Kublaoui BM, Gemelli T, Tolson KP, Wang Y, Zinn AR. Oxytocin deficiency mediates hyperphagic obesity of Sim1 haploinsufficient mice. *Mol Endocrinol.* juill 2008;22(7):1723-34.
119. Ullah A, Umair M, Yousaf M, Khan SA, Nazim-Ud-Din M, Shah K, et al. Sequence variants in four genes underlying Bardet-Biedl syndrome in consanguineous families. *Mol Vis.* 2017;23:482-94.
120. Ansley SJ, Badano JL, Blacque OE, Hill J, Hoskins BE, Leitch CC, et al. Basal body dysfunction is a likely cause of pleiotropic Bardet-Biedl syndrome. *Nature.* 9 oct 2003;425(6958):628-33.
121. Ainsworth C. Cilia: tails of the unexpected. *Nature.* 9 août 2007;448(7154):638-41.
122. Davenport JR, Watts AJ, Roper VC, Croyle MJ, van Groen T, Wyss JM, et al. Disruption of intraflagellar transport in adult mice leads to obesity and slow-onset cystic kidney disease. *Curr Biol.* 18 sept 2007;17(18):1586-94.
123. Cummings DE, Clement K, Purnell JQ, Vaisse C, Foster KE, Frayo RS, et al. Elevated plasma ghrelin levels in Prader Willi syndrome. *Nat Med.* juill 2002;8(7):643-4.
124. Swaab DF, Purba JS, Hofman MA. Alterations in the hypothalamic paraventricular nucleus and its oxytocin neurons (putative satiety cells) in Prader-Willi syndrome: a study of five cases. *J Clin Endocrinol Metab.* févr 1995;80(2):573-9.
125. Weinstein LS, Chen M, Liu J. Gs(alpha) mutations and imprinting defects in human disease. *Ann N Y Acad Sci.* juin 2002;968:173-97.
126. Chen M, Wang J, Dickerson KE, Kelleher J, Xie T, Gupta D, et al. Central nervous system imprinting of the G protein G(s)alpha and its role in metabolic regulation. *Cell Metab.* juin 2009;9(6):548-55.
127. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. *Nature.* 23 nov 2006;444(7118):444-54.
128. Hastings P, Lupski JR, Rosenberg SM, Ira G. Mechanisms of change in gene copy number. *Nat Rev Genet.* août 2009;10(8):551-64.
129. Walters RG, Jacquemont S, Valsesia A, de Smith AJ, Martinet D, Andersson J, et al. A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature.* 4 févr 2010;463(7281):671-5.
130. Bochukova EG, Huang N, Keogh J, Henning E, Purmann C, Blaszczyk K, et al. Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature.* févr 2010;463(7281):666-70.
131. Ren D, Li M, Duan C, Rui L. Identification of SH2-B as a key regulator of leptin sensitivity, energy balance, and body weight in mice. *Cell Metab.* août 2005;2(2):95-104.

132. Shinawi M, Liu P, Kang S-HL, Shen J, Belmont JW, Scott DA, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *J Med Genet.* mai 2010;47(5):332-41.
133. Jacquemont S, Reymond A, Zufferey F, Harewood L, Walters RG, Kutalik Z, et al. Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature.* 31 août 2011;478(7367):97-102.
134. Loos RJF, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, et al. Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat Genet.* juin 2008;40(6):768-75.
135. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature.* 10 sept 2009;461(7261):272-6.
136. Roach JC, Glusman G, Smit AFA, Huff CD, Hubley R, Shannon PT, et al. Analysis of Genetic Inheritance in a Family Quartet by Whole Genome Sequencing. *Science.* 30 avr 2010;328(5978):636-9.
137. Price AL, Kryukov GV, de Bakker PIW, Purcell SM, Staples J, Wei L-J, et al. Pooled association tests for rare variants in exon-resequencing studies. *Am J Hum Genet.* 11 juin 2010;86(6):832-8.
138. Bulik CM, Allison DB. The genetic epidemiology of thinness. *Obes Rev.* mai 2001;2(2):107-15.
139. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, et al. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet.* janv 2010;42(1):30-5.
140. McCurdy CE, Bishop JM, Williams SM, Grayson BE, Smith MS, Friedman JE, et al. Maternal high-fat diet triggers lipotoxicity in the fetal livers of nonhuman primates. *J Clin Invest.* févr 2009;119(2):323-35.
141. Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T, et al. Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci Transl Med.* 15 sept 2010;2(49):49ra67.
142. Stenholm S, Harris TB, Rantanen T, Visser M, Kritchevsky SB, Ferrucci L. Sarcopenic obesity: definition, cause and consequences. *Curr Opin Clin Nutr Metab Care.* nov 2008;11(6):693-700.
143. Stefan N, Häring H-U, Hu FB, Schulze MB. Metabolically healthy obesity: epidemiology, mechanisms, and clinical implications. *Lancet Diabetes Endocrinol.* oct 2013;1(2):152-62.
144. Hwang Y-C, Hayashi T, Fujimoto WY, Kahn SE, Leonetti DL, McNeely MJ, et al. Visceral abdominal fat accumulation predicts the conversion of metabolically healthy obese subjects to an unhealthy phenotype. *Int J Obes (Lond).* sept 2015;39(9):1365-70.
145. Sun H, Ren X, Chen Z, Li C, Chen S, Wu S, et al. Association between body mass index and mortality in a prospective cohort of Chinese adults. *Medicine (Baltimore) [Internet].* 12 août 2016 [cité 5 sept 2018];95(32). Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4985299/>
146. Savva SC, Lamnisos D, Kafatos AG. Predicting cardiometabolic risk: waist-to-height ratio or BMI. A meta-analysis. *Diabetes Metab Syndr Obes.* 24 oct 2013;6:403-19.

147. Dong B, Wang Z, Wang H-J, Ma J. Associations between adiposity indicators and elevated blood pressure among Chinese children and adolescents. *Journal of Human Hypertension*. avr 2015;29(4):236-40.
148. Guo F, Moellering DR, Garvey WT. The progression of cardiometabolic disease: validation of a new cardiometabolic disease staging system applicable to obesity. *Obesity (Silver Spring)*. janv 2014;22(1):110-8.
149. Seabolt LA, Welch EB, Silver HJ. Imaging methods for analyzing body composition in human obesity and cardiometabolic disease. *Ann N Y Acad Sci*. sept 2015;1353:41-59.
150. Blake-Lamb TL, Locks LM, Perkins ME, Woo Baidal JA, Cheng ER, Taveras EM. Interventions for Childhood Obesity in the First 1,000 Days A Systematic Review. *Am J Prev Med*. 2016;50(6):780-9.
151. Gillman MW, Oakey H, Baghurst PA, Volkmer RE, Robinson JS, Crowther CA. Effect of Treatment of Gestational Diabetes Mellitus on Obesity in the Next Generation. *Diabetes Care*. mai 2010;33(5):964-8.
152. Wen LM, Baur LA, Simpson JM, Rissel C, Wardle K, Flood VM. Effectiveness of home based early intervention on children's BMI at age 2: randomised controlled trial. *BMJ*. 26 juin 2012;344:e3732.
153. Navarro JI, Sigulem DM, Ferraro AA, Polanco JJ, Barros AJD. The double task of preventing malnutrition and overweight: a quasi-experimental community-based trial. *BMC Public Health*. 9 mars 2013;13:212.
154. Ciampa PJ, Kumar D, Barkin SL, Sanders LM, Yin HS, Perrin EM, et al. Interventions aimed at decreasing obesity in children younger than 2 years: a systematic review. *Arch Pediatr Adolesc Med*. déc 2010;164(12):1098-104.
155. Summerbell CD. Evidence-based recommendations for the development of obesity prevention programs targeted at preschool children. *Obes. Rev* 13. 2012;1:129-132.
156. Gluckman P, Nishtar S, Armstrong T. Ending childhood obesity: a multidimensional challenge. *Lancet*. 21 mars 2015;385(9973):1048-50.
157. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. *J Am Diet Assoc*. janv 2011;111(1):92-102.
158. American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Expert Panel, 2013. Executive summary: Guidelines (2013) for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Obesity Society published by the Obesity Society and American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Based on a systematic review from the The Obesity Expert Panel, 2013. *Obesity (Silver Spring)*. juill 2014;22 Suppl 2:S5-39.
159. Apovian CM, Aronne LJ, Bessesen DH, McDonnell ME, Murad MH, Pagotto U, et al. Pharmacological management of obesity: an endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. févr 2015;100(2):342-62.

160. Leblanc ES, O'Connor E, Whitlock EP, Patnode CD, Kapka T. Effectiveness of primary care-relevant treatments for obesity in adults: a systematic evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med.* 4 oct 2011;155(7):434-47.
161. Tate DF. A series of studies examining Internet treatment of obesity to inform Internet interventions for substance use and misuse. *Subst Use Misuse.* 2011;46(1):57-65.
162. Mozaffarian D. Food and weight gain: time to end our fear of fat. *Lancet Diabetes Endocrinol.* 2016;4(8):633-5.
163. Johnston BC, Kanters S, Bandayrel K, Wu P, Naji F, Siemieniuk RA, et al. Comparison of weight loss among named diet programs in overweight and obese adults: a meta-analysis. *JAMA.* 3 sept 2014;312(9):923-33.
164. Bray GA, Siri-Tarino PW. The Role of Macronutrient Content in the Diet for Weight Management. *Endocrinol Metab Clin North Am.* 2016;45(3):581-604.
165. Martinez JA, Navas-Carretero S, Saris WHM, Astrup A. Personalized weight loss strategies-the role of macronutrient distribution. *Nat Rev Endocrinol.* déc 2014;10(12):749-60.
166. Bray GA, Siri-Tarino PW. The Role of Macronutrient Content in the Diet for Weight Management. *Endocrinol Metab Clin North Am.* 2016;45(3):581-604.
167. de la Iglesia R, Loria-Kohen V, Zulet MA, Martinez JA, Reglero G, Ramirez de Molina A. Dietary Strategies Implicated in the Prevention and Treatment of Metabolic Syndrome. *Int J Mol Sci.* 10 nov 2016;17(11).
168. Global Recommendations on Physical Activity for Health [Internet]. Geneva: World Health Organization; 2010 [cité 7 sept 2018]. (WHO Guidelines Approved by the Guidelines Review Committee). Disponible sur: <http://www.ncbi.nlm.nih.gov/books/NBK305057/>
169. Després J-P. Body fat distribution and risk of cardiovascular disease: an update. *Circulation.* 4 sept 2012;126(10):1301-13.
170. Wu T, Gao X, Chen M, van Dam RM. Long-term effectiveness of diet-plus-exercise interventions vs. diet-only interventions for weight loss: a meta-analysis. *Obes Rev.* mai 2009;10(3):313-23.
171. Jakicic JM, Marcus BH, Lang W, Janney C. Effect of exercise on 24-month weight loss maintenance in overweight women. *Arch Intern Med.* 28 juill 2008;168(14):1550-9; discussion 1559-1560.
172. Look AHEAD Research Group. Eight-year weight losses with an intensive lifestyle intervention: the look AHEAD study. *Obesity (Silver Spring).* janv 2014;22(1):5-13.
173. Rampersaud E, Mitchell BD, Pollin TI, Fu M, Shen H, O'Connell JR, et al. Physical Activity and the Association of Common FTO Gene Variants With Body Mass Index and Obesity. *Arch Intern Med.* 8 sept 2008;168(16):1791-7.

174. Quelle place pour la chirurgie de l'estomac ? - EurekaSanté par VIDAL [Internet]. EurekaSanté. [cité 7 sept 2018]. Disponible sur: <https://eurekasante.vidal.fr/maladies/metabolisme-diabete/obesite.html>
175. Schauer PR, Bhatt DL, Kirwan JP, Wolski K, Aminian A, Brethauer SA, et al. Bariatric Surgery versus Intensive Medical Therapy for Diabetes - 5-Year Outcomes. *N Engl J Med*. 16 2017;376(7):641-51.
176. Bray GA, Frühbeck G, Ryan DH, Wilding JPH. Management of obesity. *Lancet*. 7 mai 2016;387(10031):1947-56.
177. Inge TH, Courcoulas AP, Jenkins TM, Michalsky MP, Helmrath MA, Brandt ML, et al. Weight Loss and Health Status 3 Years after Bariatric Surgery in Adolescents. *N Engl J Med*. 14 janv 2016;374(2):113-23.
178. Sjöström L. Review of the key results from the Swedish Obese Subjects (SOS) trial - a prospective controlled intervention study of bariatric surgery. *J Intern Med*. mars 2013;273(3):219-34.
179. Ikramuddin S, Korner J, Lee W-J, Connett JE, Inabnet WB, Billington CJ, et al. Roux-en-Y gastric bypass vs intensive medical management for the control of type 2 diabetes, hypertension, and hyperlipidemia: the Diabetes Surgery Study randomized clinical trial. *JAMA*. 5 juin 2013;309(21):2240-9.
180. Sjöström L, Peltonen M, Jacobson P, Ahlin S, Andersson-Assarsson J, Anveden Å, et al. Association of bariatric surgery with long-term remission of type 2 diabetes and with microvascular and macrovascular complications. *JAMA*. 11 juin 2014;311(22):2297-304.
181. Nguyen KT, Korner J. The sum of many parts: potential mechanisms for improvement in glucose homeostasis after bariatric surgery. *Curr Diab Rep*. 2014;14(5):481.
182. Colquitt JL, Pickett K, Loveman E, Frampton GK. Surgery for weight loss in adults. *Cochrane Database Syst Rev*. 8 août 2014;(8):CD003641.
183. Chang S-H, Stoll CRT, Song J, Varela JE, Eagon CJ, Colditz GA. The effectiveness and risks of bariatric surgery: an updated systematic review and meta-analysis, 2003-2012. *JAMA Surg*. mars 2014;149(3):275-87.
184. Bonnefond A, Keller R, Meyre D, Stutzmann F, Thuillier D, Stefanov DG, et al. Eating Behavior, Low-Frequency Functional Mutations in the Melanocortin-4 Receptor (MC4R) Gene, and Outcomes of Bariatric Operations: A 6-Year Prospective Study. *Diabetes Care*. 2016;39(8):1384-92.
185. Perez-Cornago A, de la Iglesia R, Lopez-Legarrea P, Abete I, Navas-Carretero S, Lacunza CI, et al. A decline in inflammation is associated with less depressive symptoms after a dietary intervention in metabolic syndrome patients: a longitudinal study. *Nutr J*. 24 avr 2014;13:36.
186. Vallis M. Quality of life and psychological well-being in obesity management: improving the odds of success by managing distress. *Int J Clin Pract*. mars 2016;70(3):196-205.
187. Groessl EJ, Kaplan RM, Barrett-Connor E, Ganiats TG. Body mass index and quality of well-being in a community of older adults. *Am J Prev Med*. févr 2004;26(2):126-9.

188. Muennig P, Lubetkin E, Jia H, Franks P. Gender and the burden of disease attributable to obesity. *Am J Public Health*. sept 2006;96(9):1662-8.
189. Kolotkin RL, Crosby RD, Williams GR, Hartley GG, Nicol S. The relationship between health-related quality of life and weight loss. *Obes Res*. sept 2001;9(9):564-71.
190. Pan A, Kawachi I, Luo N, Manson JE, Willett WC, Hu FB, et al. Changes in body weight and health-related quality of life: 2 cohorts of US women. *Am J Epidemiol*. 1 août 2014;180(3):254-62.
191. Warkentin LM, Majumdar SR, Johnson JA, Agborsangaya CB, Rueda-Clausen CF, Sharma AM, et al. Weight loss required by the severely obese to achieve clinically important differences in health-related quality of life: two-year prospective cohort study. *BMC Med*. 15 oct 2014;12:175.
192. Kroes M, Osei-Assibey G, Baker-Searle R, Huang J. Impact of weight change on quality of life in adults with overweight/obesity in the United States: a systematic review. *Curr Med Res Opin*. 2016;32(3):485-508.
193. de Beer M, Hofsteenge GH, Koot HM, Hirasing RA, Delemarre-van de Waal HA, Gemke RBJ. Health-related-quality-of-life in obese adolescents is decreased and inversely related to BMI. *Acta Paediatr*. mai 2007;96(5):710-4.
194. Ferguson LR, De Caterina R, Görman U, Allayee H, Kohlmeier M, Prasad C, et al. Guide and Position of the International Society of Nutrigenetics/Nutrigenomics on Personalised Nutrition: Part 1 - Fields of Precision Nutrition. *J Nutrigenet Nutrigenomics*. 2016;9(1):12-27.
195. Goni L, Cuervo M, Milagro FI, Martínez JA. Future Perspectives of Personalized Weight Loss Interventions Based on Nutrigenetic, Epigenetic, and Metagenomic Data. *J Nutr*. 9 mars 2016;
196. Wu J, Cohen P, Spiegelman BM. Adaptive thermogenesis in adipocytes: is beige the new brown? *Genes Dev*. 1 févr 2013;27(3):234-50.
197. Giles-Corti B, Vernez-Moudon A, Reis R, Turrell G, Dannenberg AL, Badland H, et al. City planning and population health: a global challenge. *Lancet*. 10 2016;388(10062):2912-24.
198. Díaz-Gutiérrez J, Martínez-González MÁ, Pons Izquierdo JJ, González-Muniesa P, Martínez JA, Bes-Rastrollo M. Living at Higher Altitude and Incidence of Overweight/Obesity: Prospective Analysis of the SUN Cohort. *PLoS ONE*. 2016;11(11):e0164483.
199. Kelley CP, Sbrocco G, Sbrocco T. Behavioral Modification for the Management of Obesity. *Prim Care*. mars 2016;43(1):159-75, x.
200. Huvenne H, Dubern B, Clément K, Poitou C. Rare Genetic Forms of Obesity: Clinical Approach and Current Treatments in 2016. *Obes Facts*. juin 2016;9(3):158-73.
201. Sanger F. Sequences, sequences, and sequences. *Annu Rev Biochem*. 1988;57:1-28.
202. Frederick Sanger - Nobel Lecture: The Chemistry of Insulin [Internet]. NobelPrize.org. [cité 8 sept 2018]. Disponible sur: <https://www.nobelprize.org/prizes/chemistry/1958/sanger/lecture/>
203. Edman P. A method for the determination of amino acid sequence in peptides. *Arch Biochem*. juill 1949;22(3):475.

204. Holley RW, Apgar J, Everett GA, Madison JT, Marquisee M, Merrill SH, et al. STRUCTURE OF A RIBONUCLEIC ACID. *Science*. 19 mars 1965;147(3664):1462-5.
205. Sanger F, Brownlee GG, Barrell BG. A two-dimensional fractionation procedure for radioactive nucleotides. *J Mol Biol*. sept 1965;13(2):373-98.
206. Wu R, Kaiser AD. Structure and base sequence in the cohesive ends of bacteriophage lambda DNA. *J Mol Biol*. 14 août 1968;35(3):523-37.
207. Gilbert W, Maxam A. The Nucleotide Sequence of the lac Operator. *Proc Natl Acad Sci U S A*. déc 1973;70(12 Pt 1-2):3581-4.
208. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*. déc 1977;74(12):5463-7.
209. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA*. févr 1977;74(2):560-4.
210. Sequencing | BioNinja [Internet]. [cité 8 sept 2018]. Disponible sur: <http://www.vce.bioninja.com.au/aos-3-heredity/molecular-biology-technique/sequencing.html>
211. Maxam–Gilbert sequencing. In: Wikipedia [Internet]. 2018 [cité 8 sept 2018]. Disponible sur: https://en.wikipedia.org/w/index.php?title=Maxam%E2%80%93Gilbert_sequencing&oldid=846228101
212. Staden R. A strategy of DNA sequencing employing computer programs. *Nucleic Acids Res*. 11 juin 1979;6(7):2601-10.
213. Sanger F, Coulson AR, Hong GF, Hill DF, Petersen GB. Nucleotide sequence of bacteriophage lambda DNA. *J Mol Biol*. 25 déc 1982;162(4):729-73.
214. Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, et al. Fluorescence detection in automated DNA sequence analysis. *Nature*. 12 juin 1986;321(6071):674-9.
215. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 5 oct 1990;215(3):403-10.
216. Prober JM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, Zagursky RJ, et al. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*. 16 oct 1987;238(4825):336-41.
217. Tabor S, Richardson CC. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc Natl Acad Sci USA*. juill 1987;84(14):4767-71.
218. Craxton M. Linear amplification sequencing, a powerful method for sequencing DNA. *Methods*. 1 août 1991;3(1):20-6.
219. DeAngelis MM, Wang DG, Hawkins TL. Solid-phase reversible immobilization for the isolation of PCR products. *Nucleic Acids Res*. 25 nov 1995;23(22):4742-3.
220. Phred, Phrap, and Consed [Internet]. [cité 8 sept 2018]. Disponible sur: <http://www.phrap.org/phredphrapconsed.html>

221. Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, et al. A whole-genome assembly of *Drosophila*. *Science*. 24 mars 2000;287(5461):2196-204.
222. *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*. 11 déc 1998;282(5396):2012-8.
223. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*. 21 oct 2004;431(7011):931-45.
224. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science*. 16 févr 2001;291(5507):1304-51.
225. Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, Schloss JA, et al. DNA sequencing at 40: past, present and future. *Nature*. 11 oct 2017;550(7676):345-53.
226. Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, et al. Single-molecule DNA sequencing of a viral genome. *Science*. 4 avr 2008;320(5872):106-9.
227. Mitra RD, Church GM. In situ localized amplification and contact replication of many individual DNA molecules. *Nucleic Acids Res*. 15 déc 1999;27(24):e34.
228. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 15 sept 2005;437(7057):376-80.
229. Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*. 1 janv 2010;327(5961):78-81.
230. Ronaghi M, Karamohamed S, Pettersson B, Uhlén M, Nyrén P. Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem*. 1 nov 1996;242(1):84-9.
231. Mitra RD, Shendure J, Olejnik J, Edyta-Krzyszowska-Olejnik null, Church GM. Fluorescent in situ sequencing on polymerase colonies. *Anal Biochem*. 1 sept 2003;320(1):55-65.
232. Seo TS, Bai X, Kim DH, Meng Q, Shi S, Ruparel H, et al. Four-color DNA sequencing by synthesis on a chip using photocleavable fluorescent nucleotides. *Proc Natl Acad Sci USA*. 26 avr 2005;102(17):5926-31.
233. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 6 nov 2008;456(7218):53-9.
234. DNA Sequencing Costs: Data [Internet]. National Human Genome Research Institute (NHGRI). [cité 9 sept 2018]. Disponible sur: <https://www.genome.gov/27541954/dna-sequencing-costs-data/>
235. Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science*. 31 janv 2003;299(5607):682-6.
236. SMRT Sequencing: Epigenetics [Internet]. PacBio. [cité 9 sept 2018]. Disponible sur: <https://www.pacb.com/smrt-science/smrt-sequencing/epigenetics/>

237. Deamer D, Akeson M, Branton D. Three decades of nanopore sequencing. *Nat Biotechnol.* 06 2016;34(5):518-24.
238. Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, et al. The potential and challenges of nanopore sequencing. *Nat Biotechnol.* oct 2008;26(10):1146-53.
239. Jain M, Koren S, Miga KH, Quick J, Rand AC, Sasani TA, et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat Biotechnol.* avr 2018;36(4):338-45.
240. Decoding DNA with a pocket-sized sequencer | www.scienceinschool.org [Internet]. [cité 9 sept 2018]. Disponible sur: <https://www.scienceinschool.org/content/decoding-dna-pocket-sized-sequencer>
241. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature Methods.* juin 2010;7(6):461-5.
242. Zhao Y, Ashcroft B, Zhang P, Liu H, Sen S, Song W, et al. Single-molecule spectroscopy of amino acids and peptides by recognition tunnelling. *Nat Nanotechnol.* juin 2014;9(6):466-73.
243. Bumgarner R. DNA microarrays: Types, Applications and their future. *Curr Protoc Mol Biol.* janv 2013;0 22:Unit-22.1.
244. Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes CA, et al. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature.* 24 févr 1977;265(5596):687-95.
245. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature.* 15 févr 2001;409(6822):860-921.
246. Edwards A, Voss H, Rice P, Civitello A, Stegemann J, Schwager C, et al. Automated DNA sequencing of the human HPRT locus. *Genomics.* avr 1990;6(4):593-608.
247. Schneider VA, Graves-Lindsay T, Howe K, Bouk N, Chen H-C, Kitts PA, et al. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Res.* 2017;27(5):849-64.
248. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* mai 2008;18(5):821-9.
249. Adey A, Kitzman JO, Burton JN, Daza R, Kumar A, Christiansen L, et al. In vitro, long-range sequence information for de novo genome assembly via transposase contiguity. *Genome Res.* déc 2014;24(12):2041-9.
250. Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat Biotechnol.* déc 2013;31(12):1119-25.
251. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* sept 2010;20(9):1297-303.

252. Chaisson MJ, Huddleston J, Dennis MY, Sudmant PH, Malig M, Hormozdiari F, et al. Resolving the complexity of the human genome using single-molecule sequencing. *Nature*. 29 janv 2015;517(7536):608-11.
253. Snyder MW, Adey A, Kitzman JO, Shendure J. Haplotype-resolved genome sequencing: experimental methods and applications. *Nat Rev Genet*. juin 2015;16(6):344-58.
254. Wang J, Wang W, Li R, Li Y, Tian G, Goodman L, et al. The diploid genome sequence of an Asian individual. *Nature*. 6 nov 2008;456(7218):60-5.
255. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature*. 1 oct 2015;526(7571):68-74.
256. Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, et al. Analysis of 6,515 exomes reveals a recent origin of most human protein-coding variants. *Nature*. 10 janv 2013;493(7431):216-20.
257. Chiu RWK, Chan KCA, Gao Y, Lau VYM, Zheng W, Leung TY, et al. Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc Natl Acad Sci USA*. 23 déc 2008;105(51):20458-63.
258. Vissers LELM, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. *Nat Rev Genet*. janv 2016;17(1):9-18.
259. Yang Y, Muzny DM, Xia F, Niu Z, Person R, Ding Y, et al. Molecular Findings Among Patients Referred for Clinical Whole-Exome Sequencing. *JAMA*. 12 nov 2014;312(18):1870-9.
260. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. *Science*. 16 nov 2007;318(5853):1108-13.
261. Buysse K, Delle Chiaie B, Van Coster R, Loeys B, De Paepe A, Mortier G, et al. Challenges for CNV interpretation in clinical molecular karyotyping: lessons learned from a 1001 sample experience. *Eur J Med Genet*. déc 2009;52(6):398-403.
262. Carter NP. Methods and strategies for analyzing copy number variation using DNA microarrays. *Nat Genet*. juill 2007;39(7 Suppl):S16-21.
263. Zhao M, Wang Q, Wang Q, Jia P, Zhao Z. Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. *BMC Bioinformatics*. 2013;14(Suppl 11):S1.
264. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF, et al. Paired-end mapping reveals extensive structural variation in the human genome. *Science*. 19 oct 2007;318(5849):420-6.
265. Medvedev P, Stanciu M, Brudno M. Computational methods for discovering structural variation with next-generation sequencing. *Nat Methods*. nov 2009;6(11 Suppl):S13-20.
266. Shin SC, Ahn DH, Kim SJ, Lee H, Oh T-J, Lee JE, et al. Advantages of Single-Molecule Real-Time Sequencing in High-GC Content Genomes. *PLoS ONE*. 2013;8(7):e68824.
267. Mason-Suares H, Landry L, S. Lebo M. Detecting Copy Number Variation via Next Generation Technology. *Current Genetic Medicine Reports*. sept 2016;4(3):74-85.

268. Nozaki H, Takano H, Misumi O, Terasawa K, Matsuzaki M, Maruyama S, et al. A 100%-complete sequence reveals unusually simple genomic features in the hot-spring red alga *Cyanidioschyzon merolae*. *BMC Biol.* 10 juill 2007;5:28.
269. Ovchinnikov S, Park H, Varghese N, Huang P-S, Pavlopoulos GA, Kim DE, et al. Protein Structure Determination using Metagenome sequence data. *Science.* 20 janv 2017;355(6322):294-8.
270. Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y. Identifying personal genomes by surname inference. *Science.* 18 janv 2013;339(6117):321-4.
271. Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Yang JL, Ferrante TC, et al. Highly multiplexed subcellular RNA sequencing in situ. *Science.* 21 mars 2014;343(6177):1360-3.
272. Peikon ID, Kebschull JM, Vagin VV, Ravens DI, Sun Y-C, Brouzes E, et al. Using high-throughput barcode sequencing to efficiently map connectomes. *Nucleic Acids Res.* 7 juill 2017;45(12):e115.
273. Scaling up DNA data storage and random access retrieval | bioRxiv [Internet]. [cité 9 sept 2018]. Disponible sur: <https://www.biorxiv.org/content/early/2017/03/07/114553>
274. Montagne L, Derhourhi M, Piton A, Toussaint B, Durand E, Vaillant E, et al. CoDE-seq, an augmented whole-exome sequencing, enables the accurate detection of CNVs and mutations in Mendelian obesity and intellectual disability. *Molecular Metabolism.* juill 2018;13:1-9.
275. Next-Generation Sequencing Library Preparation—Thermo Scientific - HK [Internet]. [cité 9 sept 2018]. Disponible sur: <https://www.thermofisher.com/hk/zt/home/brands/thermo-scientific/molecular-biology/thermo-scientific-specialized-molecular-biology-applications/sequencing-thermo-scientific/next-generation-sequencing-library-preparation-thermo-scientific.html>
276. Sequencing Technology | Sequencing by synthesis [Internet]. [cité 9 sept 2018]. Disponible sur: <https://emea.illumina.com/science/technology/next-generation-sequencing/sequencing-technology.html>
277. MiSeq Illumina sequencing system at Lyon [Internet]. Biofidal. [cité 9 sept 2018]. Disponible sur: <http://www.biofidal-lab.com/details-miseq+illumina+sequencing+system+at+lyon-50.html>
278. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 15 juill 2009;25(14):1754-60.
279. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* mai 2015;17(5):405-24.
280. Fromer M, Purcell SM. Using XHMM Software to Detect Copy Number Variation in Whole-Exome Sequencing Data: Using XHMM Software to Detect Copy Number Variation in Whole-Exome Sequencing Data. In: Haines JL, Korf BR, Morton CC, Seidman CE, Seidman JG, Smith DR, éditeurs. *Current Protocols in Human Genetics* [Internet]. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2014 [cité 9 juill 2018]. p. 7.23.1-7.23.21. Disponible sur: <http://doi.wiley.com/10.1002/0471142905.hg0723s81>

281. Backenroth D, Homsy J, Murillo LR, Glessner J, Lin E, Brueckner M, et al. CANOES: detecting rare copy number variants from whole exome sequencing data. *Nucleic Acids Res.* 1 août 2014;42(12):e97.
282. Krzywinski M, Schein J, Birol Í, Connors J, Gascoyne R, Horsman D, et al. Circos: An information aesthetic for comparative genomics. *Genome Res.* sept 2009;19(9):1639-45.
283. Srinivasan S, Lubrano-Berthelier C, Govaerts C, Picard F, Santiago P, Conklin BR, et al. Constitutive activity of the melanocortin-4 receptor is maintained by its N-terminal domain and plays a role in energy homeostasis in humans. *J Clin Invest.* oct 2004;114(8):1158-64.
284. Crutchley JL, Wang XQD, Ferraiuolo MA, Dostie J. Chromatin conformation signatures: ideal human disease biomarkers? *Biomark Med.* août 2010;4(4):611-29.
285. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science.* 15 févr 2002;295(5558):1306-11.
286. de Wit E, de Laat W. A decade of 3C technologies: insights into nuclear organization. *Genes Dev.* 1 janv 2012;26(1):11-24.
287. Davies JOJ, Telenius JM, McGowan S, Roberts NA, Taylor S, Higgs DR, et al. Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat Methods.* janv 2016;13(1):74-80.
288. Claussnitzer M, Dankel SN, Kim K-H, Quon G, Meuleman W, Haugen C, et al. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N Engl J Med.* 3 sept 2015;373(10):895-907.
289. Smemo S, Tena JJ, Kim K-H, Gamazon ER, Sakabe NJ, Gómez-Marín C, et al. Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature.* 12 mars 2014;507(7492):371-5.
290. Rask-Andersen M, Almén MS, Schiöth HB. Scrutinizing the FTO locus: compelling evidence for a complex, long-range regulatory context. *Hum Genet.* nov 2015;134(11-12):1183-93.
291. Cai M, Gao F, Lu W, Wang K. w4CSeq: software and web application to analyze 4C-seq data. *Bioinformatics.* 01 2016;32(21):3333-5.

Supplementary data :

Table A. Régions du génome connues pour être touchées par des CNV.

<i>Chr</i>	<i>Start</i>	<i>End</i>	<i>Phenotype</i>	<i>Name</i>	<i>Locus</i>	<i>Ref</i>
1	10001	12780116	Deeply set eye / Intellectual disability / Delayed cranial suture closure / Muscular hypotonia / Pointed chin / Seizures	1p36 microdeletion syndrome	1p36	(292–298)
1	145686996	146048496	Schizoaffective disorder / Thrombocytopenia-Absent Radius syndrome	1q21.1 deletion	1q21.1	(299,300)
1	147061832	148411223	Intellectual disability / Microcephaly	1q21.1 recurrent microdeletion	1q21.1	(127,301–307)
1	147061832	148411223	Intellectual disability / Autism	1q21.1 recurrent microduplication	1q21.1	(127,305–308)
2	44074	1193224	Autism	2p25.3 deletion	2p25.3	(309)
2	17088783	18060650	Autism	2p24.2 duplication	2p24.2	(309)
2	44183312	44362445	Hypotonia-cystinuria	2p21 microdeletion syndrome	2p21	(310–314)
2	59058561	61592680	Abnormality of the face / Hydronephrosis / Camptodactyly of finger / Intellectual disability / Downslanted palpebral fissures / Microcephaly / Feeding difficulties in infancy / Telecanthus / Ptosis / Optic disc hypoplasia	2p15-16.1 microdeletion syndrome	2p15-16.1	(315,316)
2	196060397	204342216	Behavioural-Psychiatric Abnormality / Cleft palate / High palate / Inguinal hernia / Feeding difficulties in infancy / Seizures / Intellectual disability / Short stature	2q33.1 deletion syndrome	2q33.1	(317,318)
2	239048167	239400949	Brachydactyly syndrome / Deeply set eye / Round face / Intellectual disability / Thin upper lip vermilion	2q37 monosomy	2q37	(319,320)
3	2173349	2869900	Autism	3p26 deletion	3p26	(321)
3	2173349	2869900	Autism	3p26 duplication	3p26	(321)
3	195999964	197617792	Intellectual disability / Long face / Prominent nasal bridge / Short philtrum	3q29 microdeletion syndrome	3q29	(322,323)
3	195999964	197617792	Intellectual disability / Microcephaly / Obesity	3q29 microduplication syndrome	3q29	(323,324)
4	1567470	2108509	Intellectual disability / Microcephaly / Muscular hypotonia / Seizures / Small for gestational age	Wolf-Hirschhorn Syndrome [deletion]	4p15.2	(325–328)
5	10001	12533192	Intellectual disability / Microcephaly / Weak, high-pitched voice	Cri du Chat Syndrome [deletion]	5p	(329–331)
5	8205499	9068862	Autism / Severe speech delay	861-kb deletion 5p15.3	5p15.3	(332)

5	112707504	112846239	Neoplasm of the colon	Familial Adenomatous Polyposis [deletion]	5q22.2	(333,334)
5	176297633	177625115	Accelerated skeletal maturation / Intellectual disability / Macrocephaly / Mandibular prognathia / Tall stature	Sotos [deletion] syndrome	5q35	(335–340)
7	73330452	74728334	Delayed speech and language development / Intellectual disability / Short attention span	7q11.23 syndrome	duplication 7q11.23	(341,342)
7	73330452	74728334	Aortic valve stenosis / Microdontia / Everted lower lip vermilion / Intellectual disability / Short stature / Malar attering / Short attention span	Williams-Beuren Syndrome [deletion]	7q11.23	(343–352)
7	96688766	96709891	2-3 toe syndactyly / Cleft palate / Abnormality of the genital-urinary system / Cutaneous finger syndactyly / Median cleft lip / Ectrodactyly (hands-feet) / Lacrimal duct aplasia / Sensorineural hearing impairment	Split hand/foot malformation 1 (SHFM1) [deletion]	7q21.3	(353–360)
8	8242533	11907120	Abnormality of the heart / Cryptorchidism / Atrioventricular canal defect / Congenital diaphragmatic hernia / Defect in the atrial septum / Hyperactivity / Intellectual disability	8p23.1 syndrome	deletion 8p23.1	(361–366)
8	8242533	11907120	Abnormal facial shape / Behavioural-Psychiatric Abnormality / Delayed speech and language development / Intellectual disability / Malformation of the heart and great vessels	8p23.1 syndrome	duplication 8p23.1	(367–382)
8	76314229	76854003	Downturned corners of mouth / Full cheeks / Exaggerated cupid's bow / High forehead / Intellectual disability / Low-set ears / Micrognathia / Prominent ears / Ptosis / Opacification of the corneal stroma	8p21.11 syndrome	microdeletion 8p21.11	(383)
9	137618991	137836126	Abnormality of the heart / Anteverted nares / Intellectual disability / Midface retrusion / Muscular hypotonia / Protruding tongue / Short nose / Synophrys / Microcephaly	Kleefstra syndrome [9q subtelomeric deletion]	9q subtelomeric	(384–389)
10	79864244	87451243	Autism	10q22.3-23.3 deletion	10q22.3-23.3	(390)
11	31784791	32435541	Abnormality of the male genitalia / Aniridia / Renal neoplasm / Intellectual disability	WAGR [deletion] syndrome	11p13	(391–394)
11	43973250	46030899	Osteochondromas / Enlarged parietal foramina / Intellectual disability and delayed development of speech, motor skills and social skills / Brachycephaly / Short philtrum	Potocki-Shaer [deletion] syndrome	11p11.2-p13	(395)
12	970834	1237305	Childhood apraxia of speech	12p13.33 syndrome	microdeletion 12p13.33	(396)
12	64678139	68251745	Intellectual disability / Osteopoikilosis / Proportionate short stature	12q14 Syndrome	microdeletion 12q14	(397)
15	23374765	28193120	EEG abnormality / Microcephaly / Intellectual disability / Schizophrenia / Seizures / Truncal ataxia	Angelman (Type 1) [deletion] syndrome	15q11-q13	(299,398–404)

15	23374765	28193120	Feeding difficulties in infancy / Schizophrenia / Hypogonadism / Intellectual disability / Muscular hypotonia / Truncal obesity	Prader-Willi syndrome (Type 1) [deletion]	15q11-q13	(299,405–412)
15	22749354	28193120	EEG abnormality / Microcephaly / Intellectual disability / Schizophrenia / Seizures / Truncal ataxia	Angelman syndrome (Type 2) [deletion]	15q11-q13	(299,398–404)
15	22749354	28193120	Feeding difficulties in infancy / Hypogonadism / Intellectual disability / Muscular hypotonia / Truncal obesity / Schizophrenia	Prader-Willi syndrome (Type 2) [deletion]	15q11-q13	(299,405–412)
15	30618103	32153206	Abnormality of the face / Schizophrenia / Intellectual disability / Seizures	15q13.3 microdeletion syndrome	15q13.3	(299,413)
15	74120302	75680570	Abnormality of the face / Intellectual disability / Joint laxity / Microcephaly / Short stature / Hypospadias / Small for gestational age	15q24 recurrent microdeletion syndrome	15q24	(413–415)
15	98814741	101981189	Abnormality of the kidney / Intellectual disability / Tall stature	15q24 overgrowth syndrome [duplication]	15q26	(416,417)
16	10001	784372	Abnormality of erythrocytes / Intellectual disability	ATR-16 syndrome [deletion]	16p13.3	(418–420)
16	3725054	3880120	Intellectual disability / Microcephaly / Dysmorphic facial features	Rubinstein-Taybi Syndrome [deletion]	16p13.3	(421,422)
16	14892827	16392827	Intellectual disability / Microcephaly / Epilepsy / Short stature / Facial dysmorphism / Behavioral problems	16p13.11 recurrent microdeletion (neurocognitive disorder susceptibility locus)	16p13.11	(423–425)
16	14892827	16392827	Behavioral abnormalities / Intellectual disability / Congenital heart defects / Skeletal anomalies	16p13.11 recurrent microduplication (neurocognitive disorder susceptibility locus)	16p13.11	(423,424,426)
16	21463739	29272756	Autism / Abnormal facial shape / Global developmental delay / Microcephaly / Intellectual disability / Short Stature	16p11.2-p12.2 microduplication syndrome	16p11.2-p12.2	(427–433)
16	29595531	30188534	Autism / Schizophrenia / Thinness	16p11.2 microduplication syndrome	16p11.2	(434–437)
16	29595531	30188534	Autism / Severe obesity	16p11.2 microdeletion syndrome	16p11.2	(434,437,438)
17	150208	2685615	Abnormality of the heart / Frontal bossing / Lissencephaly / Microcephaly / Midface retrusion	Miller-Dieker syndrome [deletion]	17p13.3	(439–442)
17	14194598	15567589	Abnormality of the motor neurons / Decreased motor nerve conduction velocity / Hypertrophic nerve changes / Impaired pain sensation / Impaired proprioception / Impaired temperature sensation / Impaired vibratory sensation / Pes cavus	Charcot-Marie-Tooth syndrome type 1A [duplication]	17p11.2	(443–445)

17	14194598	15567589	Abnormality of the motor neurons / Motor conduction block	Hereditary Liability to Pressure Palsies [deletion]	17p11.2	(446–450)
17	16869758	20318836	Brachydactyly syndrome / Hoarse voice / Hyperactivity / Intellectual disability / Muscular hypotonia / Stereotypic behavior / Self-mutilation / Short stature / Sleep disturbance	Smith-Magenis syndrome [deletion]	17p11.2	(451–458)
17	16869758	20318836	Autism / Hyperactivity / Short attention span / Short stature	Potocki-Lupski syndrome [duplication]	17p11.2	(453,454,456–460)
17	30780079	31936302	Abnormality of the heart / Neurobromas / Cafe-au-lait spot / Hypertelorism / Coarse facial features / Intellectual disability	NF1-microdeletion syndrome	17q11.2	(461–464)
17	36459258	37856298	Abnormality of the liver / Diabetes mellitus / Multiple renal cysts	RCAD (renal cysts and diabetes) [deletion]	17q12	(465,466)
17	45627800	46217040	Feeding difficulties in infancy / Intellectual disability / Muscular hypotonia / Small for gestational age / Long face	17q21.31 recurrent microdeletion syndrome (Koolen de Vries syndrome)	17q21.31	(302,467–473)
22	1	16971860	Abnormality of the heart / Anal atresia / Palpebral fissures / Intellectual disability / Iris coloboma / Micrognathia / Preauricular pit / Preauricular skin tag / Short stature	Cat-Eye Syndrome (Type I) [duplication]	22q11	(474–477)
22	19009792	21452445	Abnormality of the heart / Hypocalcemia / Delayed speech and language development / Nasal speech / T lymphocytopenia	22q11 deletion syndrome (Velocardiofacial / DiGeorge syndrome)	22q11	(476,478–481)
22	19009792	21452445	Intellectual disability / Nasal speech / Telecanthus	22q11 duplication syndrome	22q11	(482–487)
22	21562828	23380258	Abnormality of the face / Intellectual disability / Short stature / Small for gestational age	22q11.2 distal deletion syndrome	22q11.2	(488)
22	50607088	50749416	Autism / Intellectual disability / Delayed speech and language development / Hyperactivity / Muscular hypotonia	22q13 deletion syndrome (Phelan-Mcdermid syndrome)	22q13	(489–492)

Toutes les coordonnées se basent sur le génome de référence en version hg38.

Table B. Données cliniques des participants.

<i>ID</i>	<i>Sex</i>	<i>Age</i>	<i>BMI</i> *	<i>IOTF</i> <i>BMI</i> **	<i>Isolated</i> <i>obesity</i>	<i>Consanguinity</i>	<i>Intellectual</i> <i>disability</i>	<i>Seizures</i>	<i>Behavioral</i> <i>disorder</i>	<i>Visual</i> <i>auditory</i> <i>impairment</i>	<i>Dysmorphic</i> <i>feature</i> <i>malformation</i>
OSV11	F	7.6	27.8	> 35			x		x		
OSV14	F	9.6	27.5	> 30	x		x				x
OSV17	F	12.1	32	> 35			x	x	x	x	x
OSV20	M	8.1	22.5	> 30		x	x				
OSV21	F	11.5	29	> 30		x	x				
OSV24	M	17.6	36.7	> 35			x		x		
OSV26	F	7.5	25.1	> 35			x		x		x
OSV3	M	10.6	29.5	> 30	x		x		x	x	x
OSV30	M	6.8	27.5	> 35	x		x	x	x		x
OSV33	M	8.4	31	> 35			x		x		x
OSV34	F	11.8	27.6	> 30	x		x				
OSV37	F	4.6	21.3	> 35			x		x	x	x
OSV41	F	9.3	23.1	> 30			x				x
OSV44	F	16.5	30.6	> 30			x				
OSV46	F	12.5	28.3	> 30				x	x	x	x
OSV48	M	8.8	25.8	> 30	x		x		x	x	
OSV52	M	10.3	26.2	> 30	x		x		x		x
OSV53	M	11.4	36	> 35			x			x	
OSV58	M	3.2	25	> 35		x					
OSV6	M	3.9	21.9	> 35			x		x		x
OSV65	F	7.5	21	> 30			x		x	x	
OSV68	M	11.1	31	> 35			x				
OSV70	M	12.5	37.5	> 35			x		x		x
OSV71	M	9.3	24.7	> 30					x	x	x
OSV74	F	9.1	23	> 30	x		x			x	
OSV77	M	4.4	22	> 35	x				x		x
OSV80	F	16.2	40	> 35			x		x		x

OSV82	F	7.2	22	> 30			x	x	x	x	x
OSV86	M	6.3	23	> 35	x		x		x	x	x
OSV9	F	7.5	26.9	> 35	x		x				x
U1	M	16	34.7	> 35		x	x			x	x
U2	F	29	37.0				x				
U3	M	5.7	23.9	> 35							
U4	F	7.5	22.2	> 30							
U5	M	10.7		< 30			x	x	x	x	
U6	F	2.2	33.1	> 35							
U7	M	14.5	48.0	> 35		x					
U8	M	10.5	29.4	> 35		x					
U9	M	10.5	44.0	> 35	x	x	x				x
U10	F	35	67.0								
U11	M	7.0	40.0	> 35							
U12	F	11.5	58.0	> 35							

BMI, body mass index; **F**, female; **IOTF BMI**, extended international (IOTF) body mass index; **M**, male.

*Obesity cut-off in adults (over 18 years of age): BMI = 30; Severe obesity cut-off: BMI = 35; Morbid obesity cut-off: BMI = 40

**Obesity cut-off in children (under 18 years of age): IOTF BMI = 30; Morbid obesity cut-off: IOTF BMI = 35

Table C. Liste des gènes dans lesquels des mutations ponctuelles ont été montré comme causant des formes Mendéliennes d'obésité, de déficience intellectuelle, d'hypercholesterolemie, de diabète et de lipodystrophie.

<i>Gene</i>	<i>Name</i>	<i>Location</i>	<i>Mendelian disorder</i>	<i>NM transcript</i>
ABCC8	ATP binding cassette subfamily C member 8	11p15.1	Diabetes	NM_000352.4
ACAN	aggrecan	15q26.1	Intellectual Disability	NM_013227.3
ACAT1	acetyl-CoA acetyltransferase 1	11q22.3	Intellectual Disability	NM_000019.3
ACSL4	acyl-CoA synthetase long chain family member 4	Xq23	Intellectual Disability	NM_004458.2
ACTB	actin beta	7p22.1	Intellectual Disability	NM_001101.3
ACY1	aminoacylase 1	3p21.2	Intellectual Disability	NM_000666.2
ADA	adenosine deaminase	20q13.12	Intellectual Disability	NM_000022.2
ADNP	activity dependent neuroprotector homeobox	20q13.13	Intellectual Disability	NM_015339.3
ADSL	adenylosuccinate lyase	22q13.1	Intellectual Disability	NM_000026.2
AFF2	AF4/FMR2 family member 2	Xq28	Intellectual Disability	NM_002025.3
AFF4	AF4/FMR2 family member 4	5q31.1	Intellectual Disability & Obesity	NM_014423.3
AGA	aspartylglucosaminidase	4q34.3	Intellectual Disability	NM_000027.3
AGO1	argonaute 1, RISC catalytic component	1p34.3	Intellectual Disability	NM_012199.2
AGPAT2	1-acylglycerol-3-phosphate O-acyltransferase 2	9q34.3	Lipodystrophy	NM_006412.3
AHDC1	AT-hook DNA binding motif containing 1	1p36.11-p35.3	Intellectual Disability	NM_001029882.3
AHI1	Abelson helper integration site 1	6q23.3	Intellectual Disability	NM_017651.4
AKT2	AKT serine/threonine kinase 2	19q13.2	Lipodystrophy	NM_001626.5
ALDH18A1	aldehyde dehydrogenase 18 family member A1	10q24.1	Intellectual Disability	NM_002860.3
ALDH3A2	aldehyde dehydrogenase 3 family member A2	17p11.2	Intellectual Disability	NM_000382.2
ALDH5A1	aldehyde dehydrogenase 5 family member A1	6p22.3	Intellectual Disability	NM_001080.3
ALDH7A1	aldehyde dehydrogenase 7 family member A1	5q23.2	Intellectual Disability	NM_001182.4
ALDOB	aldolase, fructose-bisphosphate B	9q31.1	Intellectual Disability	NM_000035.3
ALG1	ALG1, chitobiosyldiphosphodolichol beta-mannosyltransferase	16p13.3	Intellectual Disability	NM_019109.4
ALG12	ALG12, alpha-1,6-mannosyltransferase	22q13.33	Intellectual Disability	NM_024105.3
ALG13	ALG13, UDP-N-acetylglucosaminyltransferase subunit	Xq23	Intellectual Disability	NM_001099922.2
ALG3	ALG3, alpha-1,3- mannosyltransferase	3q27.1	Intellectual Disability	NM_005787.5
ALG6	ALG6, alpha-1,3-glucosyltransferase	1p31.3	Intellectual Disability	NM_013339.3
ALMS1	ALMS1, centrosome and basal body associated protein	2p13.1	Obesity & Syndromic features	NM_015120.4
AMT	aminomethyltransferase	3p21.31	Intellectual Disability	NM_000481.3
ANK2	ankyrin 2	4q25-q26	Intellectual Disability	NM_001148.4
ANK3	ankyrin 3	10q21.2	Intellectual Disability	NM_020987.3
ANKH	ANKH inorganic pyrophosphate transport regulator	5p15.2	Intellectual Disability	NM_054027.4
ANKRD11	ankyrin repeat domain 11	16q24.3	Intellectual Disability	NM_013275.5
ANOS1	anosmin 1	Xp22.31	Obesity & Syndromic features	NM_000216.2
AP1S2	adaptor related protein complex 1 sigma 2 subunit	Xp22.2	Intellectual Disability	NM_003916.4
AP4B1	adaptor related protein complex 4 beta 1 subunit	1p13.2	Intellectual Disability	NM_006594.3

AP4E1	adaptor related protein complex 4 epsilon 1 subunit	15q21.2	Intellectual Disability	NM_007347.4
AP4M1	adaptor related protein complex 4 mu 1 subunit	7q22.1	Intellectual Disability	NM_004722.3
AP4S1	adaptor related protein complex 4 sigma 1 subunit	14q12	Intellectual Disability	NM_007077.4
APOB	apolipoprotein B	2p24.1	Hypercholesterolemia	NM_000384.2
APOE	apolipoprotein E	19q13.32	Hypercholesterolemia	NM_000041.3
APPL1	adaptor protein, phosphotyrosine interacting with PH domain and leucine zipper 1	3p14.3	Diabetes	NM_012096.2
ARFGEF2	ADP ribosylation factor guanine nucleotide exchange factor 2	20q13.13	Intellectual Disability	NM_006420.2
ARHGEF6	Rac/Cdc42 guanine nucleotide exchange factor 6	Xq26.3	Intellectual Disability	NM_004840.2
ARHGEF9	Cdc42 guanine nucleotide exchange factor 9	Xq11.1	Intellectual Disability	NM_015185.2
ARID1A	AT-rich interaction domain 1A	1p36.11	Intellectual Disability	NM_006015.4
ARID1B	AT-rich interaction domain 1B	6q25.3	Intellectual Disability	NM_020732.3
ARIH1	ariadne RBR E3 ubiquitin protein ligase 1	15q24.1	Intellectual Disability	NM_005744.3
ARL6	ADP ribosylation factor like GTPase 6	3q11.2	Obesity & Syndromic features	NM_177976.2
ARSA	arylsulfatase A	22q13.33	Intellectual Disability	NM_000487.5
ARX	aristaless related homeobox	Xp21.3	Intellectual Disability	NM_139058.2
ASAH1	N-acylsphingosine amidohydrolase 1	8p22	Intellectual Disability	NM_177924.3
ASH1L	ASH1 like histone lysine methyltransferase	1q22	Intellectual Disability	NM_018489.2
ASL	argininosuccinate lyase	7q11.21	Intellectual Disability	NM_000048.3
ASPA	aspartoacylase	17p13.2	Intellectual Disability	NM_000049.2
ASPM	abnormal spindle microtubule assembly	1q31.3	Intellectual Disability	NM_018136.4
ASS1	argininosuccinate synthase 1	9q34.11	Intellectual Disability	NM_000050.4
ASXL1	additional sex combs like 1, transcriptional regulator	20q11.21	Intellectual Disability	NM_015338.5
ASXL3	additional sex combs like 3, transcriptional regulator	18q12.1	Intellectual Disability	NM_030632.1
ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	2q35	Intellectual Disability	NM_004044.6
ATP6AP2	ATPase H+ transporting accessory protein 2	Xp11.4	Intellectual Disability	NM_005765.2
ATP7A	ATPase copper transporting alpha	Xq21.1	Intellectual Disability	NM_000052.6
ATP7B	ATPase copper transporting beta	13q14.3	Intellectual Disability	NM_000053.3
ATR	ATR serine/threonine kinase	3q23	Intellectual Disability	NM_001184.3
ATRX	ATRX, chromatin remodeler	Xq21.1	Intellectual Disability & Obesity	NM_000489.4
AUTS2	AUTS2, activator of transcription and developmental regulator	7q11.22	Intellectual Disability	NM_015570.3
BANF1	barrier to autointegration factor 1	11q13.1	Lipodystrophy	NM_001143985.1
BBS1	Bardet-Biedl syndrome 1	11q13.2	Obesity & Syndromic features	NM_024649.4
BBS10	Bardet-Biedl syndrome 10	12q21.2	Obesity & Syndromic features	NM_024685.3
BBS12	Bardet-Biedl syndrome 12	4q27	Obesity & Syndromic features	NM_152618.2
BBS2	Bardet-Biedl syndrome 2	16q13	Obesity & Syndromic features	NM_031885.3
BBS4	Bardet-Biedl syndrome 4	15q24.1	Obesity & Syndromic features	NM_033028.4
BBS5	Bardet-Biedl syndrome 5	2q31.1	Obesity & Syndromic features	NM_152384.2

BBS7	Bardet-Biedl syndrome 7	4q27	Obesity & Syndromic features	NM_176824.2
BBS9	Bardet-Biedl syndrome 9	7p14.3	Obesity & Syndromic features	NM_198428.2
BCKDHA	branched chain keto acid dehydrogenase E1, alpha polypeptide	19q13.2	Intellectual Disability	NM_000709.3
BCKDHB	branched chain keto acid dehydrogenase E1 subunit beta	6q14.1	Intellectual Disability	NM_183050.2
BCKDK	branched chain ketoacid dehydrogenase kinase	16p11.2	Intellectual Disability	NM_005881.3
BCL11A	B-cell CLL/lymphoma 11A	2p16.1	Intellectual Disability	NM_022893.3
BCOR	BCL6 corepressor	Xp11.4	Intellectual Disability	NM_017745.5
BDNF	brain derived neurotrophic factor	11p14.1	Obesity & Syndromic features	NM_170735.5
BLK	BLK proto-oncogene, Src family tyrosine kinase	8p23.1	Diabetes	NM_001715.2
BRAF	B-Raf proto-oncogene, serine/threonine kinase	7q34	Intellectual Disability	NM_004333.4
BRD3	bromodomain containing 3	9q34.2	Intellectual Disability	NM_007371.3
BRPF1	bromodomain and PHD finger containing 1	3p25.3	Intellectual Disability	NM_001003694.1
BRWD1	bromodomain and WD repeat domain containing 1	21q22.2	Intellectual Disability	NM_018963.4
BRWD3	bromodomain and WD repeat domain containing 3	Xq21.1	Intellectual Disability	NM_153252.4
BSCL2	BSCL2, seipin lipid droplet biogenesis associated	11q12.3	Lipodystrophy	NM_032667.6
BTD	biotinidase	3p25.1	Intellectual Disability	NM_000060.3
C12orf57	chromosome 12 open reading frame 57	12p13.31	Intellectual Disability	NM_138425.3
CA8	carbonic anhydrase 8	8q12.1	Intellectual Disability	NM_004056.4
CACNA1C	calcium voltage-gated channel subunit alpha1 C	12p13.33	Intellectual Disability	NM_000719.6
CACNA1D	calcium voltage-gated channel subunit alpha1 D	3p21.1	Intellectual Disability	NM_000720.3
CACNA1F	calcium voltage-gated channel subunit alpha1 F	Xp11.23	Intellectual Disability	NM_005183.3
CACNA1G	calcium voltage-gated channel subunit alpha1 G	17q21.33	Intellectual Disability	NM_018896.4
CACNG2	calcium voltage-gated channel auxiliary subunit gamma 2	22q12.3	Intellectual Disability	NM_006078.3
CASK	calcium/calmodulin dependent serine protein kinase	Xp11.4	Intellectual Disability	NM_003688.3
CASR	calcium sensing receptor	3q13.33-q21.1	Intellectual Disability	NM_000388.3
CAV1	caveolin 1	7q31.2	Lipodystrophy	NM_001753.4
CAVIN1	caveolae associated protein 1	17q21.31	Lipodystrophy	NM_012232.5
CBS	cystathionine-beta-synthase	21q22.3	Intellectual Disability	NM_000071.2
CC2D1A	coiled-coil and C2 domain containing 1A	19p13.12	Intellectual Disability	NM_017721.4
CCDC22	coiled-coil domain containing 22	Xp11.23	Intellectual Disability	NM_014008.4
CDH10	cadherin 10	5p14.2-p14.1	Intellectual Disability	NM_006727.3
CDH15	cadherin 15	16q24.3	Intellectual Disability	NM_004933.2
CDK13	cyclin dependent kinase 13	7p14.1	Intellectual Disability	NM_003718.4
CDK5RAP2	CDK5 regulatory subunit associated protein 2	9q33.2	Intellectual Disability	NM_018249.5
CDK6	cyclin dependent kinase 6	7q21.2	Intellectual Disability	NM_001259.6
CDKL5	cyclin dependent kinase like 5	Xp22.13	Intellectual Disability	NM_003159.2
CEL	carboxyl ester lipase	9q34.13	Diabetes	NM_001807.4
CENPJ	centromere protein J	13q12.12-q12.13	Intellectual Disability	NM_018451.4
CEP152	centrosomal protein 152	15q21.1	Intellectual Disability	NM_014985.3

CEP19	centrosomal protein 19	3q29	Obesity & Syndromic features	NM_032898.4
CEP290	centrosomal protein 290	12q21.32	Obesity & Syndromic features	NM_025114.3
CHAMP1	chromosome alignment maintaining phosphoprotein 1	13q34	Intellectual Disability	NM_001164144.1
CHD2	chromodomain helicase DNA binding protein 2	15q26.1	Intellectual Disability & Obesity	NM_001271.3
CHD4	chromodomain helicase DNA binding protein 4	12p13.31	Intellectual Disability	NM_001273.3
CHD7	chromodomain helicase DNA binding protein 7	8q12.2	Intellectual Disability	NM_017780.3
CHD8	chromodomain helicase DNA binding protein 8	14q11.2	Intellectual Disability	NM_001170629.1
CHRNA2	cholinergic receptor nicotinic alpha 2 subunit	8p21.2	Intellectual Disability	NM_000742.3
CHRNA4	cholinergic receptor nicotinic alpha 4 subunit	20q13.33	Intellectual Disability	NM_000744.6
CHRN2	cholinergic receptor nicotinic beta 2 subunit	1q21.3	Intellectual Disability	NM_000748.2
CIDEC	cell death inducing DFFA like effector c	3p25.3	Lipodystrophy	NM_022094.3
CLCN2	chloride voltage-gated channel 2	3q27.1	Intellectual Disability	NM_004366.5
CLCN4	chloride voltage-gated channel 4	Xp22.2	Intellectual Disability	NM_001830.3
CLIC2	chloride intracellular channel 2	Xq28	Intellectual Disability	NM_001289.5
CLN3	CLN3, battenin	16p12.1	Intellectual Disability	NM_001042432.1
CLN5	CLN5, intracellular trafficking protein	13q22.3	Intellectual Disability	NM_006493.2
CLN6	CLN6, transmembrane ER protein	15q23	Intellectual Disability	NM_017882.2
CLN8	CLN8, transmembrane ER and ERGIC protein	8p23.3	Intellectual Disability	NM_018941.3
CLTC	clathrin heavy chain	17q23.1	Intellectual Disability	NM_001288653.1
CNKS2	connector enhancer of kinase suppressor of Ras 2	Xp22.12	Intellectual Disability	NM_014927.3
CNOT3	CCR4-NOT transcription complex subunit 3	19q13.42	Intellectual Disability	NM_014516.3
CNTN4	contactin 4	3p26.3-p26.2	Intellectual Disability	NM_175607.2
CNTNAP2	contactin associated protein like 2	7q35-q36.1	Intellectual Disability	NM_014141.5
COL4A3BP	collagen type IV alpha 3 binding protein	5q13.3	Intellectual Disability	NM_001130105.1
COQ8A	coenzyme Q8A	1q42.13	Intellectual Disability	NM_020247.4
CPA6	carboxypeptidase A6	8q13.2	Intellectual Disability	NM_020361.4
CPS1	carbamoyl-phosphate synthase 1	2q34	Intellectual Disability	NM_001875.4
CRBN	cereblon	3p26.2	Intellectual Disability	NM_016302.3
CREBBP	CREB binding protein	16p13.3	Intellectual Disability & Obesity	NM_004380.2
CSDE1	cold shock domain containing E1	1p13.2	Intellectual Disability	NM_001130523.2
CSNK2A1	casein kinase 2 alpha 1	20p13	Intellectual Disability	NM_001895.3
CSTB	cystatin B	21q22.3	Intellectual Disability	NM_000100.3
CTCF	CCCTC-binding factor	16q22.1	Intellectual Disability	NM_006565.3
CTNFB1	catenin beta 1	3p22.1	Intellectual Disability	NM_001904.3
CTS2	cathepsin F	11q13.2	Intellectual Disability	NM_003793.3
CUL3	cullin 3	2q36.2	Intellectual Disability	NM_003590.4
CUL4B	cullin 4B	Xq24	Intellectual Disability & Obesity	NM_003588.3
CUX2	cut like homeobox 2	12q24.11-q24.12	Intellectual Disability	NM_015267.3
CYFIP1	cytoplasmic FMR1 interacting protein 1	15q11.2	Intellectual Disability	NM_014608.3
CYP27A1	cytochrome P450 family 27 subfamily A member 1	2q35	Intellectual Disability	NM_000784.3

CYP2U1	cytochrome P450 family 2 subfamily U member 1	4q25	Intellectual Disability	NM_183075.2
D2HGDH	D-2-hydroxyglutarate dehydrogenase	2q37.3	Intellectual Disability	NM_152783.4
DBT	dihydrolipoamide branched chain transacylase E2	1p21.2	Intellectual Disability	NM_001918.3
DCX	doublecortin	Xq23	Intellectual Disability	NM_178153.2
DDHD2	DDHD domain containing 2	8p11.23	Intellectual Disability	NM_015214.2
DDX3X	DEAD-box helicase 3, X-linked	Xp11.4	Intellectual Disability	NM_001193416.2
DEAF1	DEAF1, transcription factor	11p15.5	Intellectual Disability	NM_021008.3
DEPDC5	DEP domain containing 5	22q12.2-q12.3	Intellectual Disability	NM_001242896.1
DHCR7	7-dehydrocholesterol reductase	11q13.4	Intellectual Disability	NM_001360.2
DISC1	disrupted in schizophrenia 1	1q42.2	Intellectual Disability	NM_018662.2
DKC1	dyskerin pseudouridine synthase 1	Xq28	Intellectual Disability	NM_001363.4
DLG2	discs large MAGUK scaffold protein 2	11q14.1	Intellectual Disability	NM_001142699.1
DLG3	discs large MAGUK scaffold protein 3	Xq13.1	Intellectual Disability	NM_021120.3
DNM1	dynamamin 1	9q34.11	Intellectual Disability	NM_004408.3
DNMT3A	DNA methyltransferase 3 alpha	2p23.3	Intellectual Disability	NM_175629.2
DNMT3B	DNA methyltransferase 3 beta	20q11.21	Intellectual Disability	NM_006892.3
DPAGT1	dolichyl-phosphate acetylglucosaminophosphotransferase 1	N-11q23.3	Intellectual Disability	NM_001382.3
DPM1	dolichyl-phosphate mannosyltransferase subunit 1, catalytic	20q13.13	Intellectual Disability	NM_003859.1
DPP6	dipeptidyl peptidase like 6	7q36.2	Intellectual Disability	NM_001936.4
DPYD	dihydropyrimidine dehydrogenase	1p21.3	Intellectual Disability	NM_000110.3
DSCAM	DS cell adhesion molecule	21q22.2	Intellectual Disability	NM_001389.3
DYNC1H1	dynein cytoplasmic 1 heavy chain 1	14q32.31	Intellectual Disability	NM_001376.4
DYRK1A	dual specificity tyrosine phosphorylation regulated kinase 1A	21q22.13	Intellectual Disability	NM_001396.3
DYRK1B	dual specificity tyrosine phosphorylation regulated kinase 1B	19q13.2	Obesity & Syndromic features	NM_004714.2
EBF3	early B-cell factor 3	10q26.3	Intellectual Disability	NM_001005463.2
EEF1A2	eukaryotic translation elongation factor 1 alpha 2	20q13.33	Intellectual Disability	NM_001958.3
EFTUD2	elongation factor Tu GTP binding domain containing 2	17q21.31	Intellectual Disability	NM_004247.3
EHMT1	euchromatic histone lysine methyltransferase 1	9q34.3	Intellectual Disability & Obesity	NM_024757.4
EIF2AK3	eukaryotic translation initiation factor 2 alpha kinase 3	2p11.2	Diabetes & Syndromic features	NM_004836.5
EIF2S3	eukaryotic translation initiation factor 2 subunit gamma	Xp22.11	Intellectual Disability, Diabetes & Obesity	NM_001415.3
EIF4E	eukaryotic translation initiation factor 4E	4q23	Intellectual Disability	NM_001130679.1
ELP2	elongator acetyltransferase complex subunit 2	18q12.2	Intellectual Disability	NM_001242875.1
EN2	engrailed homeobox 2	7q36.3	Intellectual Disability	NM_001427.3
EP300	E1A binding protein p300	22q13.2	Intellectual Disability & Obesity	NM_001429.3
EPB41L1	erythrocyte membrane protein band 4.1 like 1	20q11.23	Intellectual Disability	NM_012156.2
EPM2A	EPM2A, laforin glucan phosphatase	6q24.3	Intellectual Disability	NM_005670.3
ERLIN2	ER lipid raft associated 2	8p11.23	Intellectual Disability	NM_007175.6
FBN1	fibrillin 1	15q21.1	Lipodystrophy	NM_000138.4
FGD1	FYVE, RhoGEF and PH domain containing 1	Xp11.22	Intellectual Disability	NM_004463.2

FGFR1	fibroblast growth factor receptor 1	8p11.23	Obesity & Syndromic features	NM_023110.2
FLNA	filamin A	Xq28	Intellectual Disability	NM_001456.3
FMR1	fragile X mental retardation 1	Xq27.3	Intellectual Disability & Obesity	NM_002024.5
FOLR1	folate receptor 1	11q13.4	Intellectual Disability	NM_016725.2
FOXG1	forkhead box G1	14q12	Intellectual Disability	NM_005249.4
FOXP1	forkhead box P1	3p13	Intellectual Disability	NM_032682.5
FOXP2	forkhead box P2	7q31.1	Intellectual Disability	NM_014491.3
FOXP3	forkhead box P3	Xp11.23	Diabetes & Syndromic features	NM_014009.3
FRMPD4	FERM and PDZ domain containing 4	Xp22.2	Intellectual Disability	NM_014728.3
FTCD	formimidoyltransferase cyclodeaminase	21q22.3	Intellectual Disability	NM_006657.2
FTSJ1	FtsJ RNA methyltransferase homolog 1	Xp11.23	Intellectual Disability	NM_012280.3
FUCA1	alpha-L-fucosidase 1	1p36.11	Intellectual Disability	NM_000147.4
GABBR2	gamma-aminobutyric acid type B receptor subunit 2	9q22.33	Intellectual Disability	NM_005458.7
GABRA1	gamma-aminobutyric acid type A receptor alpha1 subunit	5q34	Intellectual Disability	NM_000806.5
GABRB2	gamma-aminobutyric acid type A receptor beta2 subunit	5q34	Intellectual Disability	NM_021911.2
GABRB3	gamma-aminobutyric acid type A receptor beta3 subunit	15q12	Intellectual Disability	NM_000814.5
GABRG2	gamma-aminobutyric acid type A receptor gamma2 subunit	5q34	Intellectual Disability	NM_000816.3
GALT	galactose-1-phosphate uridylyltransferase	9p13.3	Intellectual Disability	NM_000155.3
GAMT	guanidinoacetate N-methyltransferase	19p13.3	Intellectual Disability	NM_000156.5
GAP43	growth associated protein 43	3q13.31	Intellectual Disability	NM_001130064.1
GATA4	GATA binding protein 4	8p23.1	Diabetes & Syndromic features	NM_002052.3
GATA6	GATA binding protein 6	18q11.2	Diabetes & Syndromic features	NM_005257.5
GATAD2B	GATA zinc finger domain containing 2B	1q21.3	Intellectual Disability	NM_020699.2
GATM	glycine amidinotransferase	15q21.1	Intellectual Disability	NM_001482.2
GBA	glucosylceramidase beta	1q22	Intellectual Disability	NM_001005741.2
GCDH	glutaryl-CoA dehydrogenase	19p13.13	Intellectual Disability	NM_000159.3
GCH1	GTP cyclohydrolase 1	14q22.2	Intellectual Disability	NM_000161.2
GCK	glucokinase	7p13	Diabetes	NM_000162.3
GDI1	GDP dissociation inhibitor 1	Xq28	Intellectual Disability	NM_001493.2
GFPT2	glutamine-fructose-6-phosphate transaminase 2	5q35.3	Intellectual Disability	NM_005110.2
GHR	growth hormone receptor	5p13.1-p12	Obesity & Syndromic features	NM_000163.4
GK	glycerol kinase	Xp21.2	Intellectual Disability	NM_000167.5
GLDC	glycine decarboxylase	9p24.1	Intellectual Disability	NM_000170.2
GLIS3	GLIS family zinc finger 3	9p24.2	Diabetes & Syndromic features	NM_152629.3
GLUD1	glutamate dehydrogenase 1	10q23.2	Intellectual Disability	NM_005271.3
GNAI1	G protein subunit alpha i1	7q21.11	Intellectual Disability	NM_002069.5
GNAO1	G protein subunit alpha o1	16q13	Intellectual Disability	NM_020988.2
GNAS	GNAS complex locus	20q13.32	Obesity & Syndromic features	NM_000516.4
GOSR2	golgi SNAP receptor complex member 2	17q21.32	Intellectual Disability	NM_004287.3

GPC3	glypican 3		Xq26.2	Intellectual Disability	NM_004484.3
GRIA3	glutamate ionotropic receptor AMPA type subunit 3		Xq25	Intellectual Disability	NM_000828.4
GRID2	glutamate ionotropic receptor delta type subunit 2		4q22.1-q22.2	Intellectual Disability	NM_001510.3
GRIK2	glutamate ionotropic receptor kainate type subunit 2		6q16.3	Intellectual Disability	NM_021956.4
GRIN1	glutamate ionotropic receptor NMDA type subunit 1		9q34.3	Intellectual Disability	NM_007327.3
GRIN2A	glutamate ionotropic receptor NMDA type subunit 2A		16p13.2	Intellectual Disability	NM_000833.4
GRIN2B	glutamate ionotropic receptor NMDA type subunit 2B		12p13.1	Intellectual Disability	NM_000834.3
GRIP1	glutamate receptor interacting protein 1		12q14.3	Intellectual Disability	NM_021150.3
GUSB	glucuronidase beta		7q11.21	Intellectual Disability	NM_000181.3
HAL	histidine ammonia-lyase		12q23.1	Intellectual Disability	NM_002108.3
HCCS	holocytochrome c synthase		Xp22.2	Intellectual Disability	NM_005333.4
HCFC1	host cell factor C1		Xq28	Intellectual Disability	NM_005334.2
HCN1	hyperpolarization activated cyclic nucleotide gated potassium channel 1		5p12	Intellectual Disability	NM_021072.3
HCN2	hyperpolarization activated cyclic nucleotide gated potassium and sodium channel 2		19p13.3	Intellectual Disability	NM_001194.3
HDAC4	histone deacetylase 4		2q37.3	Intellectual Disability & Obesity	NM_006037.3
HDAC8	histone deacetylase 8		Xq13.1	Intellectual Disability & Obesity	NM_018486.2
HEPACAM	hepatic and glial cell adhesion molecule		11q24.2	Intellectual Disability	NM_152722.4
HEXA	hexosaminidase subunit alpha		15q23	Intellectual Disability	NM_000520.4
HGSNAT	heparan-alpha-glucosaminide acetyltransferase	N-	8p11.21-p11.1	Intellectual Disability	NM_152419.2
HIVEP2	human immunodeficiency virus type I enhancer binding protein 2		6q24.2	Intellectual Disability	NM_006734.3
HNF1A	HNF1 homeobox A		12q24.31	Diabetes	NM_000545.5
HNF1B	HNF1 homeobox B		17q12	Diabetes	NM_000458.3
HNF4A	hepatocyte nuclear factor 4 alpha		20q13.12	Diabetes	NM_175914.4
HNRNPU	heterogeneous nuclear ribonucleoprotein U		1q44	Intellectual Disability	NM_031844.2
HOXA1	homeobox A1		7p15.2	Intellectual Disability	NM_005522.4
HPRT1	hypoxanthine phosphoribosyltransferase 1		Xq26.2-q26.3	Intellectual Disability	NM_000194.2
HRAS	HRas proto-oncogene, GTPase		11p15.5	Intellectual Disability	NM_005343.2
HSD17B10	hydroxysteroid 17-beta dehydrogenase 10		Xp11.22	Intellectual Disability	NM_004493.2
HUWE1	HECT, UBA and WWE domain containing 1, E3 ubiquitin protein ligase		Xp11.22	Intellectual Disability	NM_031407.6
IDS	iduronate 2-sulfatase		Xq28	Intellectual Disability	NM_000202.6
IDUA	iduronidase, alpha-L-		4p16.3	Intellectual Disability	NM_000203.4
IER3IP1	immediate early response 3 interacting protein 1		18q21.1	Diabetes & Syndromic features	NM_016097.4
IGBP1	immunoglobulin binding protein 1		Xq13.1	Intellectual Disability	NM_001551.2
IGF1	insulin like growth factor 1		12q23.2	Intellectual Disability	NM_000618.3
IGF1R	insulin like growth factor 1 receptor		15q26.3	Intellectual Disability	NM_000875.4
IKBKG	inhibitor of nuclear factor kappa B kinase subunit gamma		Xq28	Intellectual Disability	NM_001099857.2
IL1RAPL1	interleukin 1 receptor accessory protein like 1		Xp21.3-p21.2	Intellectual Disability	NM_014271.3
INPPE	inositol polyphosphate-5-phosphatase E		9q34.3	Obesity & Syndromic features	NM_019892.4

INS	insulin	11p15.5	Diabetes	NM_000207.2
IQSEC2	IQ motif and Sec7 domain 2	Xp11.22	Intellectual Disability	NM_001111125.2
ITPR1	inositol 1,4,5-trisphosphate receptor type 1	3p26.1	Intellectual Disability	NM_002222.5
IVD	isovaleryl-CoA dehydrogenase	15q15.1	Intellectual Disability	NM_002225.3
KANSL1	KAT8 regulatory NSL complex subunit 1	17q21.31	Intellectual Disability	NM_001193466.1
KANSL2	KAT8 regulatory NSL complex subunit 2	12q13.11	Intellectual Disability	NM_017822.3
KAT6A	lysine acetyltransferase 6A	8p11.21	Intellectual Disability	NM_006766.3
KAT6B	lysine acetyltransferase 6B	10q22.2	Intellectual Disability	NM_012330.3
KATNAL2	katanin catalytic subunit A1 like 2	18q21.1	Intellectual Disability	NM_031303.2
KCNA1	potassium voltage-gated channel subfamily A member 1	12p13.32	Intellectual Disability	NM_000217.2
KCNB1	potassium voltage-gated channel subfamily B member 1	20q13.13	Intellectual Disability	NM_004975.2
KCNC1	potassium voltage-gated channel subfamily C member 1	11p15.1	Intellectual Disability	NM_001112741.1
KCNH1	potassium voltage-gated channel subfamily H member 1	1q32.2	Intellectual Disability	NM_172362.2
KCNJ10	potassium voltage-gated channel subfamily J member 10	1q23.2	Intellectual Disability	NM_002241.4
KCNJ11	potassium voltage-gated channel subfamily J member 11	11p15.1	Diabetes	NM_000525.3
KCNJ6	potassium voltage-gated channel subfamily J member 6	21q22.13	Lipodystrophy	NM_002240.3
KCNK9	potassium two pore domain channel subfamily K member 9	8q24.3	Intellectual Disability	NM_001282534.1
KCNMA1	potassium calcium-activated channel subfamily M alpha 1	10q22.3	Intellectual Disability	NM_002247.3
KCNQ2	potassium voltage-gated channel subfamily Q member 2	20q13.33	Intellectual Disability	NM_172107.2
KCNQ3	potassium voltage-gated channel subfamily Q member 3	8q24.22	Intellectual Disability	NM_004519.3
KCNT1	potassium sodium-activated channel subfamily T member 1	9q34.3	Intellectual Disability	NM_020822.2
KCTD13	potassium channel tetramerization domain containing 13	16p11.2	Intellectual Disability & Obesity	NM_178863.4
KCTD7	potassium channel tetramerization domain containing 7	7q11.21	Intellectual Disability	NM_153033.4
KDM5A	lysine demethylase 5A	12p13.33	Intellectual Disability	NM_001042603.2
KDM5B	lysine demethylase 5B	1q32.1	Intellectual Disability	NM_006618.3
KDM5C	lysine demethylase 5C	Xp11.22	Intellectual Disability	NM_004187.3
KDM6A	lysine demethylase 6A	Xp11.3	Intellectual Disability & Obesity	NM_021140.3
KIAA2022/ NEXMIF	neurite extension and migration factor	Xq13.3	Intellectual Disability	NM_001008537.2
KIF1A	kinesin family member 1A	2q37.3	Intellectual Disability	NM_004321.6
KIRREL3	kirre like nephrin family adhesion molecule 3	11q24.2	Intellectual Disability	NM_032531.3
KLF11	Kruppel like factor 11	2p25.1	Diabetes	NM_003597.4
KLHL15	kelch like family member 15	Xp22.11	Intellectual Disability	NM_030624.2
KMT2C	lysine methyltransferase 2C	7q36.1	Intellectual Disability	NM_170606.2
KMT2D	lysine methyltransferase 2D	12q13.12	Intellectual Disability & Obesity	NM_003482.3
KMT5B	lysine methyltransferase 5B	11q13.2	Intellectual Disability	NM_017635.4
L1CAM	L1 cell adhesion molecule	Xq28	Intellectual Disability	NM_000425.4
LAMC3	laminin subunit gamma 3	9q34.12	Intellectual Disability	NM_006059.3

LAMP2	lysosomal associated membrane protein 2	Xq24	Intellectual Disability	NM_002294.2
LAS1L	LAS1 like, ribosome biogenesis factor	Xq12	Intellectual Disability	NM_031206.4
LDLR	low density lipoprotein receptor	19p13.2	Hypercholesterolemia	NM_000527.4
LDLRAP1	low density lipoprotein receptor adaptor protein 1	1p36.11	Hypercholesterolemia	NM_015627.2
LEP	leptin	7q32.1	Obesity	NM_000230.2
LEPR	leptin receptor	1p31.3	Obesity	NM_002303.5
LGI1	leucine rich glioma inactivated 1	10q23.33	Intellectual Disability	NM_005097.2
LIPA	lipase A, lysosomal acid type	10q23.31	Hypercholesterolemia	NM_000235.3
LIPE	lipase E, hormone sensitive type	19q13.2	Lipodystrophy	NM_005357.3
LMNA	lamin A/C	1q22	Lipodystrophy	NM_170707.3
LRP2	LDL receptor related protein 2	2q31.1	Intellectual Disability	NM_004525.2
MAGEL2	MAGE family member L2	15q11.2	Intellectual Disability & Obesity	NM_019066.4
MAGI2	membrane associated guanylate kinase, WW and PDZ domain containing 2	7q21.11	Intellectual Disability	NM_012301.3
MAN1B1	mannosidase alpha class 1B member 1	9q34.3	Intellectual Disability	NM_016219.4
MAN2B1	mannosidase alpha class 2B member 1	19p13.13	Intellectual Disability	NM_000528.3
MANBA	mannosidase beta	4q24	Intellectual Disability	NM_005908.3
MAOA	monoamine oxidase A	Xp11.3	Intellectual Disability	NM_000240.3
MAST1	microtubule associated serine/threonine kinase 1	19p13.2	Intellectual Disability	NM_014975.2
MBD5	methyl-CpG binding domain protein 5	2q23.2	Intellectual Disability	NM_018328.4
MBTPS2	membrane bound transcription factor peptidase, site 2	Xp22.12	Intellectual Disability	NM_015884.3
MC4R	melanocortin 4 receptor	18q21.32	Obesity	NM_005912.2
MCPH1	microcephalin 1	8p23.1	Intellectual Disability	NM_024596.3
MECP2	methyl-CpG binding protein 2	Xq28	Intellectual Disability & Obesity	NM_004992.3
MED12	mediator complex subunit 12	Xq13.1	Intellectual Disability	NM_005120.2
MED13L	mediator complex subunit 13 like	12q24.21	Intellectual Disability	NM_015335.4
MED17	mediator complex subunit 17	11q21	Intellectual Disability	NM_004268.4
MED23	mediator complex subunit 23	6q23.2	Intellectual Disability	NM_015979.3
MEF2C	myocyte enhancer factor 2C	5q14.3	Intellectual Disability	NM_002397.4
MET	MET proto-oncogene, receptor tyrosine kinase	7q31	Intellectual Disability	NM_001127500.1
MFN2	mitofusin 2	1p36.22	Lipodystrophy	NM_014874.3
MFSD8	major facilitator superfamily domain containing 8	4q28.2	Intellectual Disability	NM_152778.2
MID1	midline 1	Xp22	Intellectual Disability	NM_000381.3
MKKS	McKusick-Kaufman syndrome	20p12.2	Obesity & Syndromic features	NM_018848.3
MKS1	Meckel syndrome, type 1	17q22	Obesity & Syndromic features	NM_017777.3
MMACHC	methylmalonic aciduria (cobalamin deficiency) cblC type, with homocystinuria	1p34.1	Intellectual Disability	NM_015506.2
MNX1	motor neuron and pancreas homeobox 1	7q36.3	Diabetes & Syndromic features	NM_005515.3
MRAP2	melanocortin 2 receptor accessory protein 2	6q14.2	Obesity	NM_138409.2
MTHFR	methylenetetrahydrofolate reductase	1p36.22	Intellectual Disability	NM_005957.4
MTR	5-methyltetrahydrofolate-homocysteine methyltransferase	1q43	Intellectual Disability	NM_000254.2

MUT	methylmalonyl-CoA mutase	6p12.3	Intellectual Disability	NM_000255.3
MYCN	MYCN proto-oncogene, bHLH transcription factor	2p24.3	Intellectual Disability	NM_005378.5
MYT1L	myelin transcription factor 1 like	2p25.3	Intellectual Disability & Obesity	NM_015025.3
NAA10	N(alpha)-acetyltransferase 10, NatA catalytic subunit	Xq28	Intellectual Disability	NM_003491.3
NAGLU	N-acetyl-alpha-glucosaminidase	17q21.2	Intellectual Disability	NM_000263.3
NDE1	nudE neurodevelopment protein 1	16p13.11	Intellectual Disability	NM_001143979.1
NDST1	N-deacetylase and N-sulfotransferase 1	5q33.1	Intellectual Disability	NM_001543.4
NDUFA1	NADH:ubiquinone oxidoreductase subunit A1	Xq24	Intellectual Disability	NM_004541.3
NECAP1	NECAP endocytosis associated 1	12p13.31	Intellectual Disability	NM_015509.3
NEDD4L	neural precursor cell expressed, developmentally down-regulated 4-like, E3 ubiquitin protein ligase	18q21.31	Intellectual Disability	NM_015277.5
NEUROD1	neuronal differentiation 1	2q31.3	Diabetes & Syndromic features	NM_002500.4
NEUROG3	neurogenin 3	10q22.1	Diabetes & Syndromic features	NM_020999.3
NF1	neurofibromin 1	17q11.2	Intellectual Disability	NM_000267.3
NFIX	nuclear factor I X	19p13.13	Intellectual Disability	NM_001271043.2
NGFR	nerve growth factor receptor	17q21.33	Intellectual Disability	NM_002507.3
NHLRC1	NHL repeat containing E3 ubiquitin protein ligase 1	6p22.3	Intellectual Disability	NM_198586.2
NHS	NHS actin remodeling regulator	Xp22.2-p22.13	Intellectual Disability	NM_198270.3
NIPBL	NIPBL, cohesin loading factor	5p13.2	Intellectual Disability	NM_133433.3
NKX2-2	NK2 homeobox 2	20p11.22	Diabetes & Syndromic features	NM_002509.3
NLGN3	neuroligin 3	Xq13.1	Intellectual Disability	NM_018977.3
NLGN4X	neuroligin 4, X-linked	Xp22.32-p22.31	Intellectual Disability	NM_020742.3
NPC1	NPC intracellular cholesterol transporter 1	18q11.2	Intellectual Disability	NM_000271.4
NPC2	NPC intracellular cholesterol transporter 2	14q24.3	Intellectual Disability	NM_006432.3
NRXN1	neurexin 1	2p16.3	Intellectual Disability	NM_001135659.1
NRXN2	neurexin 2	11q13.1	Intellectual Disability	NM_138732.2
NSD1	nuclear receptor binding SET domain protein 1	5q35.3	Intellectual Disability	NM_022455.4
NSDHL	NAD(P) dependent steroid dehydrogenase-like	Xq28	Intellectual Disability	NM_015922.2
NSUN2	NOP2/Sun RNA methyltransferase family member 2	5p15.31	Intellectual Disability	NM_017755.5
NTNG1	netrin G1	1p13.3	Intellectual Disability	NM_014917.2
NTRK2	neurotrophic receptor tyrosine kinase 2	9q21.33	Obesity & Syndromic features	NM_006180.4
OCRL	OCRL, inositol polyphosphate-5-phosphatase	Xq26.1	Intellectual Disability	NM_000276.3
OFD1	OFD1, centriole and centriolar satellite protein	Xp22.2	Intellectual Disability	NM_003611.2
OPHN1	oligophrenin 1	Xq12	Intellectual Disability	NM_002547.2
OTC	ornithine carbamoyltransferase	Xp11.4	Intellectual Disability	NM_000531.5
OXTR	oxytocin receptor	3p25.3	Intellectual Disability	NM_000916.3
PACS1	phosphofurin acidic cluster sorting protein 1	11q13.1-q13.2	Intellectual Disability	NM_018026.3
PAFAH1B1	platelet activating factor acetylhydrolase 1b regulatory subunit 1	17p13.3	Intellectual Disability	NM_000430.3
PAH	phenylalanine hydroxylase	12q23.2	Intellectual Disability	NM_000277.1
PAK3	p21 (RAC1) activated kinase 3	Xq23	Intellectual Disability	NM_002578.3

PAX4	paired box 4	7q32.1	Diabetes	NM_006193.2
PAX5	paired box 5	9p13.2	Intellectual Disability	NM_016734.2
PAX6	paired box 6	11p13	Diabetes	NM_000280.4
PCBD1	pterin-4 alpha-carbinolamine dehydratase 1	10q22.1	Diabetes	NM_000281.3
PCCA	propionyl-CoA carboxylase alpha subunit	13q32.3	Intellectual Disability	NM_000282.3
PCCB	propionyl-CoA carboxylase beta subunit	3q22.3	Intellectual Disability	NM_000532.4
PCDH19	protocadherin 19	Xq22.1	Intellectual Disability	NM_001184880.1
PCSK1	proprotein convertase subtilisin/kexin type 1	5q15	Obesity	NM_000439.4
PCSK9	proprotein convertase subtilisin/kexin type 9	1p32.3	Hypercholesterolemia	NM_174936.3
PCYT1A	phosphate cytidylyltransferase 1, choline, alpha	3q29	Lipodystrophy	NM_005017.2
PDHA1	pyruvate dehydrogenase E1 alpha 1 subunit	Xp22.12	Intellectual Disability	NM_000284.3
PDHX	pyruvate dehydrogenase complex component X	11p13	Intellectual Disability	NM_003477.2
PDX1	pancreatic and duodenal homeobox 1	13q12.2	Diabetes & Syndromic features	NM_000209.3
PEX1	peroxisomal biogenesis factor 1	7q21.2	Intellectual Disability	NM_000466.2
PEX12	peroxisomal biogenesis factor 12	17q12	Intellectual Disability	NM_000286.2
PEX6	peroxisomal biogenesis factor 6	6p21.1	Intellectual Disability	NM_000287.3
PEX7	peroxisomal biogenesis factor 7	6q23.3	Intellectual Disability	NM_000288.3
PGAP2	post-GPI attachment to proteins 2	11p15.4	Intellectual Disability	NM_001256240.1
PGAP3	post-GPI attachment to proteins 3	17q21.2	Intellectual Disability	NM_033419.4
PHF6	PHD finger protein 6	Xq26.2	Intellectual Disability & Obesity	NM_032458.2
PHF8	PHD finger protein 8	Xp11.22	Intellectual Disability	NM_015107.2
PHGDH	phosphoglycerate dehydrogenase	1p12	Intellectual Disability	NM_006623.3
PHIP	pleckstrin homology domain interacting protein	6q14.1	Intellectual Disability & Obesity	NM_017934.5
PIGA	phosphatidylinositol glycan anchor biosynthesis class A	Xp22.2	Intellectual Disability	NM_002641.3
PIGN	phosphatidylinositol glycan anchor biosynthesis class N	18q21.33	Intellectual Disability	NM_176787.4
PIGO	phosphatidylinositol glycan anchor biosynthesis class O	9p13.3	Intellectual Disability	NM_032634.3
PIGV	phosphatidylinositol glycan anchor biosynthesis class V	1p36.11	Intellectual Disability	NM_017837.3
PIK3R1	phosphoinositide-3-kinase regulatory subunit 1	5q13.1	Lipodystrophy	NM_181523.2
PLCB1	phospholipase C beta 1	20p12.3	Intellectual Disability	NM_015192.3
PLIN1	perilipin 1	15q26.1	Lipodystrophy	NM_002666.4
PLP1	proteolipid protein 1	Xq22.2	Intellectual Disability	NM_000533.4
PMM2	phosphomannomutase 2	16p13.2	Intellectual Disability	NM_000303.2
PNKP	polynucleotide kinase 3'-phosphatase	19q13.33	Intellectual Disability	NM_007254.3
PNPO	pyridoxamine 5'-phosphate oxidase	17q21.32	Intellectual Disability	NM_018129.3
POGZ	pogo transposable element derived with ZNF domain	1q21.3	Intellectual Disability	NM_015100.3
POLD1	DNA polymerase delta 1, catalytic subunit	19q13.3	Lipodystrophy	NM_002691.3
POLG	DNA polymerase gamma, catalytic subunit	15q26.1	Intellectual Disability	NM_002693.2
POMC	proopiomelanocortin	2p23.3	Obesity	NM_001035256.1
POMGNT1	protein O-linked mannose N-acetylglucosaminyltransferase 1 (beta 1,2-)	1p34.1	Intellectual Disability	NM_017739.3
PORCN	porcupine O-acyltransferase	Xp11.23	Intellectual Disability	NM_203475.2

POU1F1	POU class 1 homeobox 1	3p11.2	Intellectual Disability	NM_000306.3
PPARG	peroxisome proliferator activated receptor gamma	3p25.2	Lipodystrophy	NM_015869.4
PPP2R5D	protein phosphatase 2 regulatory subunit B'delta	6p21.1	Intellectual Disability	NM_006245.3
PPT1	palmitoyl-protein thioesterase 1	1p34.2	Intellectual Disability	NM_000310.3
PQBP1	polyglutamine binding protein 1	Xp11.23	Intellectual Disability	NM_005710.2
PRICKLE1	prickle planar cell polarity protein 1	12q12	Intellectual Disability	NM_153026.2
PRKRA	protein activator of interferon induced protein kinase EIF2AK2	2q31.2	Intellectual Disability	NM_003690.4
PRODH	proline dehydrogenase 1	22q11.21	Intellectual Disability	NM_016335.4
PROK2	prokineticin 2	3p13	Obesity & Syndromic features	NM_001126128.1
PROKR2	prokineticin receptor 2	20p12.3	Obesity & Syndromic features	NM_144773.2
PRPS1	phosphoribosyl pyrophosphate synthetase 1	Xq22.3	Intellectual Disability	NM_002764.3
PRRT2	proline rich transmembrane protein 2	16p11.2	Intellectual Disability	NM_145239.2
PRSS12	protease, serine 12	4q26	Intellectual Disability	NM_003619.3
PSAP	prosaposin	10q22.1	Intellectual Disability	NM_002778.2
PSMB8	proteasome subunit beta 8	6p21.32	Lipodystrophy	NM_148919.3
PTCHD1	patched domain containing 1	Xp22.11	Intellectual Disability	NM_173495.2
PTEN	phosphatase and tensin homolog	10q23.31	Intellectual Disability	NM_000314.5
PTF1A	pancreas specific transcription factor, 1a	10p12.2	Diabetes & Syndromic features	NM_178161.2
PTPN11	protein tyrosine phosphatase, non-receptor type 11	12q24.13	Intellectual Disability	NM_002834.3
PTS	6-pyruvoyltetrahydropterin synthase	11q22.3	Intellectual Disability	NM_000317.2
PUF60	poly(U) binding splicing factor 60	8q24.3	Intellectual Disability	NM_078480.2
PURA	purine rich element binding protein A	5q31.3	Intellectual Disability	NM_005859.4
QARS	glutamyl-tRNA synthetase	3p21.31	Intellectual Disability	NM_005051.2
QRICH1	glutamine rich 1	3p21.31	Intellectual Disability	NM_017730.2
RAB23	RAB23, member RAS oncogene family	6p12.1-p11.2	Obesity & Syndromic features	NM_183227.2
RAB39B	RAB39B, member RAS oncogene family	Xq28	Intellectual Disability	NM_171998.3
RAD21	RAD21 cohesin complex component	8q24.11	Intellectual Disability & Obesity	NM_006265.2
RAI1	retinoic acid induced 1	17p11.2	Intellectual Disability	NM_030665.3
RBM10	RNA binding motif protein 10	Xp11.3	Intellectual Disability	NM_005676.4
RELN	reelin	7q22.1	Intellectual Disability	NM_005045.3
RFX6	regulatory factor X6	6q22.1	Diabetes & Syndromic features	NM_173560.3
RLIM	ring finger protein, LIM domain interacting	Xq13.2	Intellectual Disability	NM_183353.2
RNF135	ring finger protein 135	17q11.2	Intellectual Disability	NM_032322.3
RPL10	ribosomal protein L10	Xq28	Intellectual Disability	NM_006013.4
RPS6KA3	ribosomal protein S6 kinase A3	Xp22.12	Intellectual Disability & Obesity	NM_004586.2
SATB2	SATB homeobox 2	2q33.1	Intellectual Disability	NM_015265.3
SCARB2	scavenger receptor class B member 2	4q21.1	Intellectual Disability	NM_005506.3
SCN1A	sodium voltage-gated channel alpha subunit 1	2q24.3	Intellectual Disability	NM_001165963.1
SCN1B	sodium voltage-gated channel beta subunit 1	19q13.11	Intellectual Disability	NM_001037.4

SCN2A	sodium voltage-gated channel alpha subunit 2	2q24.3	Intellectual Disability	NM_021007.2
SCN8A	sodium voltage-gated channel alpha subunit 8	12q13.13	Intellectual Disability	NM_014191.3
SDCCAG8	serologically defined colon cancer antigen 8	1q43-q44	Obesity & Syndromic features	NM_006642.3
SEMA5A	semaphorin 5A	5p15.31	Intellectual Disability	NM_003966.2
SET	SET nuclear proto-oncogene	9q34.11	Intellectual Disability	NM_001122821.1
SETBP1	SET binding protein 1	18q12.3	Intellectual Disability	NM_015559.2
SETD1A	SET domain containing 1A	16p11.2	Intellectual Disability	NM_014712.2
SETD2	SET domain containing 2	3p21.31	Intellectual Disability	NM_014159.6
SETD5	SET domain containing 5	3p25.3	Intellectual Disability	NM_001080517.2
SGSH	N-sulfoglucosamine sulfohydrolase	17q25.3	Intellectual Disability	NM_000199.3
SH2B1	SH2B adaptor protein 1	16p11.2	Obesity	NM_001145795.1
SHANK1	SH3 and multiple ankyrin repeat domains 1	19q13.33	Intellectual Disability	NM_016148.2
SHANK2	SH3 and multiple ankyrin repeat domains 2	11q13.3-q13.4	Intellectual Disability	NM_012309.4
SHANK3	SH3 and multiple ankyrin repeat domains 3	22q13.33	Intellectual Disability	NM_033517.1
SIM1	single-minded family bHLH transcription factor 1	6q16.3	Obesity & Syndromic features	NM_005068.2
SKI	SKI proto-oncogene	1p36.33-p36.32	Intellectual Disability	NM_003036.3
SLC13A5	solute carrier family 13 member 5	17p13.1	Intellectual Disability	NM_177550.4
SLC16A2	solute carrier family 16 member 2	Xq13.2	Intellectual Disability	NM_006517.4
SLC17A5	solute carrier family 17 member 5	6q13	Intellectual Disability	NM_012434.4
SLC19A2	solute carrier family 19 member 2	1q24.2	Diabetes & Syndromic features	NM_006996.2
SLC1A1	solute carrier family 1 member 1	9p24.2	Intellectual Disability	NM_004170.5
SLC25A13	solute carrier family 25 member 13	7q21.3	Intellectual Disability	NM_014251.2
SLC25A15	solute carrier family 25 member 15	13q14.11	Intellectual Disability	NM_014252.3
SLC25A22	solute carrier family 25 member 22	11p15.5	Intellectual Disability	NM_024698.5
SLC2A1	solute carrier family 2 member 1	1p34.2	Intellectual Disability	NM_006516.2
SLC2A2	solute carrier family 2 member 2	3q26.2	Diabetes & Syndromic features	NM_000340.1
SLC35A2	solute carrier family 35 member A2	Xp11.23	Intellectual Disability	NM_001042498.2
SLC35A3	solute carrier family 35 member A3	1p21.2	Intellectual Disability	NM_012243.2
SLC35C1	solute carrier family 35 member C1	11p11.2	Intellectual Disability	NM_018389.4
SLC45A1	solute carrier family 45 member 1	1p36.23	Intellectual Disability	NM_001080397.2
SLC46A1	solute carrier family 46 member 1	17q11.2	Intellectual Disability	NM_080669.5
SLC6A1	solute carrier family 6 member 1	3p25.3	Intellectual Disability	NM_003042.3
SLC6A8	solute carrier family 6 member 8	Xq28	Intellectual Disability	NM_005629.3
SLC9A6	solute carrier family 9 member A6	Xq26.3	Intellectual Disability	NM_006359.2
SMAD4	SMAD family member 4	18q21.2	Intellectual Disability	NM_005359.5
SMARCA2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	9p24.3	Intellectual Disability	NM_003070.4
SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	19p13.2	Intellectual Disability	NM_001128849.1
SMARCB1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	22q11.23	Intellectual Disability	NM_003073.3

SMARCE1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1	17q21.2	Intellectual Disability	NM_003079.4
SMC1A	structural maintenance of chromosomes 1A	Xp11.22	Intellectual Disability & Obesity	NM_006306.3
SMC3	structural maintenance of chromosomes 3	10q25.2	Intellectual Disability & Obesity	NM_005445.3
SMS	spermine synthase	Xp22.11	Intellectual Disability	NM_004595.4
SOBP	sine oculis binding protein homolog	6q21	Intellectual Disability	NM_018013.3
SOX10	SRY-box 10	22q13.1	Obesity & Syndromic features	NM_006941.3
SOX11	SRY-box 11	2p25.2	Intellectual Disability	NM_003108.3
SOX3	SRY-box 3	Xq27.1	Intellectual Disability	NM_005634.2
SOX5	SRY-box 5	12p12.1	Intellectual Disability	NM_006940.4
SOX7	SRY-box 7	8p23.1	Intellectual Disability	NM_031439.3
SPRTN	SprT-like N-terminal domain	1q42.2	Lipodystrophy	NM_032018.5
SPTAN1	spectrin alpha, non-erythrocytic 1	9q34.11	Intellectual Disability	NM_001130438.2
SRCAP	Snf2 related CREBBP activator protein	16p11.2	Intellectual Disability	NM_006662.2
SRD5A3	steroid 5 alpha-reductase 3	4q12	Intellectual Disability	NM_024592.4
SRGAP3	SLIT-ROBO Rho GTPase activating protein 3	3p25.3	Intellectual Disability	NM_014850.3
ST3GAL3	ST3 beta-galactoside alpha-2,3-sialyltransferase 3	1p34.1	Intellectual Disability	NM_006279.3
ST3GAL5	ST3 beta-galactoside alpha-2,3-sialyltransferase 5	2p11.2	Intellectual Disability	NM_003896.3
STAG1	stromal antigen 1	3q22.3	Intellectual Disability	NM_005862.2
STAMBP	STAM binding protein	2p13.1	Intellectual Disability	NM_006463.4
STAP1	signal transducing adaptor family member 1	4q13.2	Hypercholesterolemia	NM_012108.2
STIL	STIL, centriolar assembly protein	1p33	Intellectual Disability	NM_003035.2
STX1B	syntaxin 1B	16p11.2	Intellectual Disability	NM_052874.4
STXBP1	syntaxin binding protein 1	9q34.11	Intellectual Disability	NM_003165.3
SYN1	synapsin I	Xp11.3-p11.23	Intellectual Disability	NM_133499.2
SYNE1	spectrin repeat containing nuclear envelope protein 1	6q25.2	Intellectual Disability	NM_033071.3
SYNGAP1	synaptic Ras GTPase activating protein 1	6p21.32	Intellectual Disability	NM_006772.2
SYNJ1	synaptojanin 1	21q22.11	Intellectual Disability	NM_003895.3
SYP	synaptophysin	Xp11.23	Intellectual Disability	NM_003179.2
SZT2	SZT2, KICSTOR complex subunit	1p34.2	Intellectual Disability	NM_015284.3
TAB2	TGF-beta activated kinase 1/MAP3K7 binding protein 2	6q25.1	Intellectual Disability	NM_015093.5
TAF1	TATA-box binding protein associated factor 1	Xq13.1	Intellectual Disability	NM_004606.4
TBC1D24	TBC1 domain family member 24	16p13.3	Intellectual Disability	NM_001199107.1
TBL1XR1	transducin beta like 1 X-linked receptor 1	3q26.32	Intellectual Disability	NM_024665.4
TBR1	T-box, brain 1	2q24.2	Intellectual Disability	NM_006593.2
TBX1	T-box 1	22q11.21	Intellectual Disability & Obesity	NM_080647.1
TBX3	T-box 3	12q24.21	Obesity & Syndromic features	NM_005996.3
TCF20	transcription factor 20	22q13.2	Intellectual Disability	NM_005650.2
TCF4	transcription factor 4	18q21.2	Intellectual Disability	NM_001083962.1
TECR	trans-2,3-enoyl-CoA reductase	19p13.12	Intellectual Disability	NM_138501.5

THOC6	THO complex 6	16p13.3	Intellectual Disability	NM_024339.3
THRA	thyroid hormone receptor, alpha	17q21.1	Intellectual Disability	NM_199334.3
TIMM8A	translocase of inner mitochondrial membrane 8A	Xq22.1	Intellectual Disability	NM_004085.3
TMLHE	trimethyllysine hydroxylase, epsilon	Xq28	Intellectual Disability	NM_018196.3
TNK2	tyrosine kinase non receptor 2	3q29	Intellectual Disability	NM_001010938.1
TPP1	tripeptidyl peptidase 1	11p15.4	Intellectual Disability	NM_000391.3
TRAPPC9	trafficking protein particle complex 9	8q24.3	Intellectual Disability & Obesity	NM_031466.6
TRIM32	tripartite motif containing 32	9q33.1	Obesity & Syndromic features	NM_012210.3
TRIO	trio Rho guanine nucleotide exchange factor	5p15.2	Intellectual Disability	NM_007118.2
TRIP12	thyroid hormone receptor interactor 12	2q36.3	Intellectual Disability	NM_004238.2
TRMT10A	tRNA methyltransferase 10A	4q23	Diabetes, Intellectual Disability & Syndromic features	NM_152292.4
TRPC5	transient receptor potential cation channel subfamily C member 5	Xq23	Intellectual Disability	NM_012471.2
TSC1	TSC complex subunit 1	9q34	Intellectual Disability	NM_000368.4
TSC2	TSC complex subunit 2	16p13.3	Intellectual Disability	NM_000548.3
TSPAN7	tetraspanin 7	Xp11.4	Intellectual Disability	NM_004615.3
TTC8	tetratricopeptide repeat domain 8	14q31.3	Obesity & Syndromic features	NM_198309.3
TUB	tubby bipartite transcription factor	11p15.4	Obesity & Syndromic features	NM_003320.4
TUBA1A	tubulin alpha 1a	12q13.12	Intellectual Disability	NM_006009.3
TUBB2B	tubulin beta 2B class IIb	6p25.2	Intellectual Disability	NM_178012.4
TUSC3	tumor suppressor candidate 3	8p22	Intellectual Disability	NM_006765.3
UBE2A	ubiquitin conjugating enzyme E2 A	Xq24	Intellectual Disability & Obesity	NM_003336.3
UBE3A	ubiquitin protein ligase E3A	15q11.2	Intellectual Disability & Obesity	NM_130838.1
UBE3B	ubiquitin protein ligase E3B	12q24.11	Intellectual Disability	NM_130466.3
UBR1	ubiquitin protein ligase E3 component n-recognin 1	15q15.2	Intellectual Disability	NM_174916.2
UPB1	beta-ureidopropionase 1	22q11.23	Intellectual Disability	NM_016327.2
UPF3B	UPF3B, regulator of nonsense mediated mRNA decay	Xq24	Intellectual Disability	NM_080632.2
UROC1	urocanate hydratase 1	3q21.3	Intellectual Disability	NM_144639.2
USP27X	ubiquitin specific peptidase 27, X-linked	Xp11.23	Intellectual Disability	NM_001145073.2
USP9X	ubiquitin specific peptidase 9, X-linked	Xp11.4	Intellectual Disability	NM_001039590.2
VCP	valosin containing protein	9p13.3	Intellectual Disability	NM_007126.3
VLDLR	very low density lipoprotein receptor	9p24.2	Intellectual Disability	NM_003383.3
VPS13B	vacuolar protein sorting 13 homolog B	8q22.2	Intellectual Disability & Obesity	NM_017890.4
WAC	WW domain containing adaptor with coiled-coil	10p12.1	Intellectual Disability	NM_016628.4
WDPCP	WD repeat containing planar cell polarity effector	2p15	Obesity & Syndromic features	NM_015910.5
WDR45	WD repeat domain 45	Xp11.23	Intellectual Disability	NM_007075.3
WDR45B	WD repeat domain 45B	17q25.3	Intellectual Disability	NM_019613.3
WDR62	WD repeat domain 62	19q13.12	Intellectual Disability	NM_001083961.1
WFS1	wolframin ER transmembrane glycoprotein	4p16.1	Diabetes & Syndromic features	NM_006005.3

WNK3	WNK lysine deficient protein kinase 3	Xp11.22	Intellectual Disability	NM_020922.4
WRN	Werner syndrome RecQ like helicase	8p12	Lipodystrophy	NM_000553.4
WWOX	WW domain containing oxidoreductase	16q23.1-q23.2	Intellectual Disability	NM_016373.3
ZBTB18	zinc finger and BTB domain containing 18	1q44	Intellectual Disability	NM_205768.2
ZBTB20	zinc finger and BTB domain containing 20	3q13.31	Intellectual Disability	NM_001164342.2
ZC3H14	zinc finger CCCH-type containing 14	14q31.3	Intellectual Disability	NM_024824.4
ZC4H2	zinc finger C4H2-type containing	Xq11.2	Intellectual Disability	NM_018684.3
ZDHHC9	zinc finger DHHC-type containing 9	Xq26.1	Intellectual Disability	NM_016032.3
ZEB2	zinc finger E-box binding homeobox 2	2q22.3	Intellectual Disability	NM_014795.3
ZFP57	ZFP57 zinc finger protein	6p22.1	Diabetes	NM_001109809.2
ZMPSTE24	zinc metallopeptidase STE24	1p34.2	Lipodystrophy	NM_005857.4
ZMYM3	zinc finger MYM-type containing 3	Xq13.1	Intellectual Disability	NM_201599.2
ZMYND11	zinc finger MYND-type containing 11	10p15.3	Intellectual Disability	NM_006624.5
ZNF292	zinc finger protein 292	6q14.3	Intellectual Disability	NM_015021.1
ZNF526	zinc finger protein 526	19q13.31	Intellectual Disability	NM_133444.1
ZNF711	zinc finger protein 711	Xq21.1	Intellectual Disability	NM_021998.4
ZNF81	zinc finger protein 81	Xp11.23	Intellectual Disability	NM_007137.3

Table D. Liste des CNV détectés grâce à CoDE-seq (CS), les puces CGH ou le WES (Ex) chez les 40 participants de la cohorte de l'étude pilote, avec les tailles des CNV indiquées en kb, ainsi que le nombre de sondes ayant permis de les détecter.

<i>Patient</i>	<i>Chr</i>	<i>Start_CS</i>	<i>End_CS</i>	<i>Type</i>	<i>Detection</i>	<i>Taille CS</i>	<i>Taille aCGH</i>	<i>Taille Ex</i>	<i>Probes CS</i>	<i>Probes aCGH</i>	<i>Probes Ex</i>
28189	13	21155078	21176496	Gain	CS+Ex	21	.	21	8	.	8
28189	16	6029793	6069694	Loss	CS	40	.	.	5	.	.
28189	18	32160092	32220277	Loss	CS+aCGH+Ex	60	64	30	19	4	15
28189	22	18860192	19020828	Gain	CS	161	.	.	34	.	.
28189	8	85828677	86153000	Gain	CS+aCGH+Ex	324	312	104	42	7	11
28189	8	24323811	24335738	Loss	CS+Ex	12	.	12	6	.	6
29352	1	197089905	197105194	Loss	CS+Ex	15	.	15	12	.	12
29352	10	80042528	80202743	Gain	CS+aCGH+Ex	160	119	80	41	4	15
29352	17	73856209	74671695	Gain	CS+aCGH+Ex	815	409	467	126	13	97
29352	2	178672384	178683320	Loss	CS+Ex	11	.	11	21	.	21
29352	2	178745561	178753206	Loss	CS+Ex	8	.	8	9	.	9
29352	6	145645449	145856364	Gain	CS+aCGH	211	198	.	21	8	.
29352	6	56598452	56609377	Loss	CS+Ex	11	.	11	12	.	12
30028	13	22595208	23079332	Gain	CS+aCGH	484	377	.	51	6	.
30028	13	20489343	20822317	Gain	CS+aCGH+Ex	333	293	332	62	9	55
30028	16	59542596	59830719	Gain	CS	288	.	.	15	.	.
30028	22	22646288	23364631	Gain	CS+aCGH+Ex	718	605	669	170	17	127
30028	22	24270019	24495521	Gain	CS+Ex	226	.	193	22	.	19
30226	15	36667786	36950430	Loss	CS+aCGH	283	236	.	43	6	.
30226	22	21958243	22215308	Gain	CS+Ex	257	.	257	52	.	28
30226	5	8704478	9305400	Gain	CS+aCGH+Ex	601	501	222	79	12	50
30226	5	11266288	11521607	Gain	CS+aCGH	255	207	.	27	9	.
31425	1	71579368	71962893	Loss	CS+aCGH	384	272	.	44	5	.
31425	12	79522765	79872962	Gain	CS+aCGH+Ex	350	326	287	62	9	37
31425	12	15594901	15813060	Gain	CS+aCGH+Ex	218	191	137	39	6	22
32776	11	32931712	33068868	Gain	CS+aCGH+Ex	137	130	137	37	.	35
32776	22	25280850	25504968	Loss	CS	224	.	.	21	.	.
32776	7	101210302	101616335	Gain	CS+aCGH+Ex	406	480	108	84	.	33
32996	10	44731183	44855941	Gain	CS	125	.	.	7	.	.
32996	15	89573720	89652971	Gain	CS+Ex	79	.	77	42	.	41
32996	16	29568896	30206106	Gain	CS+aCGH+Ex	637	525	542	215	32	185
32996	3	157391680	158223677	Gain	CS+aCGH+Ex	832	730	793	108	8	30
32996	9	136429641	136439455	Loss	CS+Ex	10	.	10	10	.	10
35144	17	1252252	1317996	Loss	CS	66	.	.	8	.	.
35144	3	15309522	15568614	Gain	CS+aCGH+Ex	259	243	238	59	.	41

35144	3	14089868	14198145	Gain	CS+aCGH+Ex	108	129	85	40	.	37
35144	9	4354549	4526950	Loss	CS+aCGH	172	180	.	20	.	.
35361	1	237142228	237388211	Loss	CS+aCGH+Ex	246	238	117	35	6	9
35361	12	99648143	99744292	Loss	CS+aCGH	96	86	.	16	4	.
35871	1	16033579	16057303	Gain	CS+Ex	24	.	24	17	.	17
35871	13	69513474	70397597	Gain	CS+aCGH+Ex	884	750	406	98	18	11
35871	7	85960691	86416390	Loss	CS	456	.	.	54	.	.
35871	9	4718499	5535070	Gain	CS+aCGH+Ex	817	771	804	132	15	57
35871	9	3648286	3958614	Gain	CS+aCGH+Ex	310	273	108	40	10	7
35871	X	6533969	8147207	Gain	CS+aCGH+Ex	1613	1512	1076	198	21	25
36499	1	1452011	1495837	Loss	CS	44	.	.	26	.	.
36499	2	66390896	66817732	Gain	CS+aCGH+Ex	427	303	135	52	5	16
36499	5	94100544	94678990	Loss	CS+Ex	578	.	525	49	.	33
36499	6	113900368	113970939	Loss	CS+aCGH+Ex	71	111	30	19	9	14
36504	16	75880696	76581426	Loss	CS+aCGH+Ex	701	695	281	98	16	25
36504	17	70444033	70634347	Gain	CS	190	.	.	23	.	.
36504	3	43153	7993841	Loss	CS+aCGH+Ex	7951	7685	7414	1022	139	229
36504	5	41442	31162895	Gain	CS+aCGH+Ex	31121	31011	26875	3989	408	891
36578	3	195175598	195451820	Loss	CS+aCGH+Ex	276	269	267	49	13	33
36578	5	116172508	116242834	Loss	CS+aCGH	70	70	.	9	3	.
36578	7	133131697	133317413	Loss	CS+aCGH	186	161	.	21	4	.
36578	9	5100524	5228628	Gain	CS	128	.	.	24	.	.
36616	1	16033579	16056533	Gain	CS+Ex	23	.	23	16	.	16
36616	13	43585397	44221333	Gain	CS+Ex	636	.	528	39	.	16
36616	14	35266705	35317503	Gain	CS+Ex	51	.	51	13	.	12
36616	16	21501382	21827897	Loss	CS+Ex	327	.	254	65	.	40
36616	16	84529537	84768392	Gain	CS+aCGH+Ex	239	188	201	47	6	28
36616	2	44320188	44570056	Loss	CS+aCGH+Ex	250	225	229	38	8	21
36616	7	38264672	38330859	Loss	CS+Ex	66	.	66	12	.	12
36711	11	31520033	31707194	Loss	CS+aCGH+Ex	187	107	162	25	4	9
36711	15	22781870	23129536	Gain	CS+aCGH+Ex	348	57	342	68	6	63
36711	17	74877692	74881297	Loss	CS+Ex	4	.	4	4	.	4
36711	7	63120609	63374269	Gain	CS	254	.	.	11	.	.
36981	15	98507942	101861073	Loss	CS+aCGH+Ex	3353	3411	3169	513	90	272
36981	18	71138517	80257123	Gain	CS+aCGH+Ex	9119	8955	5489	1158	146	173
36981	9	123557538	123960973	Loss	CS+aCGH+Ex	403	382	371	47	9	13
37146	19	14028169	14042830	Gain	CS+Ex	15	.	15	9	.	9
37146	22	25280850	25504968	Loss	CS	224	.	.	21	.	.
37146	22	18906344	19020828	Gain	CS	114	.	.	30	.	.
37146	5	42633203	42810867	Gain	CS+aCGH+Ex	178	127	122	30	10	21

37146	7	143127984	143184474	Loss	CS+Ex	56	.	52	12	.	5
37146	8	8847017	9224537	Loss	CS+aCGH+Ex	378	265	265	45	6	11
37146	X	6489809	6525743	Loss	CS	36	.	.	7	.	.
37403	16	16492861	18175294	Gain	CS+aCGH+Ex	1682	1213	353	166	24	13
37403	7	72532657	72828116	Gain	CS+Ex	295	.	269	25	.	17
37465	11	61248200	61250040	Gain	CS+Ex	2	.	2	4	.	4
37465	13	91409277	91650157	Loss	CS+aCGH	241	212	.	30	7	.
37465	2	110091568	110392154	Gain	CS+Ex	301	.	113	35	.	24
37465	2	241912831	242092403	Loss	CS	180	.	.	22	.	.
37465	4	72955190	73177559	Gain	CS+aCGH+Ex	222	141	119	54	5	43
37465	9	131009179	131164184	Gain	CS+aCGH+Ex	155	105	155	59	5	59
37588	15	100048822	100341928	Gain	CS+aCGH+Ex	293	281	293	39	8	20
37588	8	2949276	3108782	Loss	CS+aCGH+Ex	160	134	160	30	3	25
37872	12	111744618	111871215	Gain	CS+aCGH+Ex	127	125	127	37	8	31
37872	8	2494944	2722597	Gain	CS	228	.	.	25	.	.
37872	X	6821874	7692693	Loss	CS+aCGH+Ex	871	689	300	108	10	17
38100	15	92191681	92670133	Gain	CS+aCGH+Ex	478	433	261	62	13	13
38100	15	93000485	93020341	Gain	CS+Ex	20	.	20	7	.	7
38100	4	85715	3992441	Loss	CS+aCGH+Ex	3907	3772	3682	939	81	729
38100	8	205215	7121027	Gain	CS+aCGH+Ex	6916	7624	6815	855	106	242
38114	15	55223860	55324856	Loss	CS+aCGH+Ex	101	113	101	19	9	9
38114	17	156185	733221	Gain	CS+Ex	577	.	577	97	.	62
38114	2	167742275	167811955	Gain	CS	70	.	.	9	.	.
38114	2	178432074	178451044	Gain	CS	19	.	.	8	.	.
38114	7	101324897	101480942	Gain	CS	156	.	.	22	.	.
38206	19	29798945	31173810	Gain	CS+aCGH+Ex	1375	1357	898	167	20	34
38206	19	4219558	4224547	Gain	CS+Ex	5	.	5	4	.	4
38206	20	14719244	14867459	Loss	CS+aCGH	148	124	.	16	3	.
38206	4	131023763	131427881	Gain	CS	404	.	.	52	.	.
38206	4	161605	172779	Gain	CS	11	.	.	6	.	.
38213	18	47535	7791187	Loss	CS+aCGH+Ex	7744	7760	7726	1142	149	436
38213	18	7810771	8293586	Gain	CS+aCGH+Ex	483	424	365	67	11	19
38213	20	59032190	64329015	Gain	CS+aCGH+Ex	5297	5187	5255	1180	151	789
38506	10	53619410	54881439	Loss	CS+aCGH+Ex	1262	1084	721	171	31	42
38506	11	87346164	87692873	Gain	CS	347	.	.	18	.	.
38506	6	57050752	57084591	Gain	CS	34	.	.	9	.	.
38506	7	80405243	80581524	Loss	CS+aCGH	176	71	.	24	3	.
38645	16	28811431	29052826	Loss	CS+aCGH+Ex	241	205	227	136	11	129
38645	4	89934788	90028366	Loss	CS	94	.	.	7	.	.
38645	6	167924983	168200387	Gain	CS+Ex	275	.	156	58	.	39

38645	6	72045966	72284159	Loss	CS+aCGH+Ex	238	215	187	46	5	25
38645	X	6841799	7253379	Gain	CS+aCGH+Ex	412	249	203	51	6	9
38785	16	29599186	29779652	Gain	CS	180	.	.	30	.	.
38785	9	28621496	28701352	Loss	CS+aCGH	80	64	.	10	3	.
38785	X	16590768	16702930	Gain	CS+aCGH+Ex	112	106	112	22	3	22
38818	10	24275739	24513481	Gain	CS+aCGH+Ex	238	222	238	33	6	11
38818	20	14719244	14887403	Loss	CS+aCGH	168	176	.	18	3	.
38818	6	89029179	89699753	Gain	CS+aCGH+Ex	671	670	618	154	21	115
38818	7	143127984	143184474	Loss	CS	56	.	.	12	.	.
38818	8	449960	613545	Gain	CS+Ex	164	.	44	26	.	12
38839	12	63551882	63715131	Loss	CS+Ex	163	.	100	30	.	22
38839	12	1840713	1875361	Loss	CS+Ex	35	.	35	10	.	10
38839	3	17214183	17470105	Loss	CS+aCGH	256	187	.	33	5	.
38839	7	108564156	108887149	Loss	CS+aCGH	323	316	.	41	8	.
38839	X	53231708	53878306	Gain	CS+aCGH+Ex	647	624	422	185	16	141
38991	10	66990109	67167854	Gain	CS+aCGH	178	182	.	22	9	.
38991	16	6982497	7333124	Loss	CS+aCGH	351	312	.	41	6	.
38991	18	67603449	67722907	Loss	CS	119	.	.	13	.	.
38991	18	1896184	1975890	Loss	CS	80	.	.	11	.	.
38991	8	8230670	8282256	Gain	CS	52	.	.	9	.	.
39147	14	39000163	39096148	Loss	CS+aCGH+Ex	96	90	62	28	5	21
39147	15	36516961	36895301	Gain	CS+aCGH	378	342	.	56	8	.
39268	14	61015882	61052435	Loss	CS+Ex	37	.	37	15	.	15
39268	4	3444014	3448310	Loss	CS+Ex	4	.	4	10	.	10
39268	7	124143633	124810634	Gain	CS+aCGH	667	582	.	77	13	.
39268	X	284166	2774766	Gain	CS+Ex	2491	.	2491	122	.	122
39268	X	155773850	156010442	Loss	CS+Ex	237	.	237	19	.	18
39268	Y	2786954	7091513	Gain	CS+aCGH+Ex	4305	.	4304	51	.	51
39268	Y	19481580	20781052	Gain	CS+aCGH+Ex	1299	.	1299	44	.	44
39402	10	536092	585858	Gain	CS	50	.	.	6	.	.
39402	20	14768964	14897332	Loss	CS+aCGH	128	99	.	14	3	.
39402	3	97610768	97899303	Gain	CS+aCGH+Ex	289	227	289	48	5	33
39409	1	1457101	1486699	Loss	CS	30	.	.	15	.	.
39409	10	51344735	51887971	Loss	CS+aCGH	543	507	.	67	21	.
39409	16	80104104	80528056	Gain	CS	424	.	.	21	.	.
39409	16	8495166	8572250	Loss	CS	77	.	.	5	.	.
39409	16	20537235	20565149	Gain	CS+Ex	28	.	28	16	.	15
39409	16	20460080	20486686	Loss	CS+Ex	27	.	27	14	.	14
39409	20	14176304	14564346	Loss	CS+aCGH	388	334	.	46	10	.
39409	22	25280850	25504968	Gain	CS	224	.	.	21	.	.

39409	7	73316540	74798284	Gain	CS+aCGH+Ex	1482	1412	1482	339	46	261
39465	15	83460498	84111657	Gain	CS+aCGH+Ex	651	687	547	91	17	39
39465	15	85014381	85167781	Gain	CS+Ex	153	.	73	33	.	22
39465	15	89573720	89652971	Gain	CS+aCGH+Ex	79	109	77	42	7	41
39465	5	12696753	13994038	Gain	CS+aCGH+Ex	1297	745	239	197	10	79
39465	8	144800458	144808433	Loss	CS+Ex	8	.	8	5	.	5
39509	11	55603419	55650942	Loss	CS	48	.	.	6	.	.
39509	15	31739875	32243158	Gain	CS+aCGH	503	434	.	58	8	.
39509	6	83253607	83385919	Loss	CS	132	.	.	16	.	.
39509	8	15693720	15951771	Loss	CS+aCGH+Ex	258	127	33	34	4	7
39612	10	101383666	101628102	Gain	CS+aCGH+Ex	244	228	197	51	10	35
39612	11	55603419	55650942	Loss	CS	48	.	.	6	.	.
39612	7	111278905	111445694	Loss	CS	167	.	.	24	.	.
39612	X	16581654	16702930	Gain	CS+aCGH+Ex	121	106	112	23	3	22
39622	1	12745877	13005656	Gain	CS	260	.	.	33	.	.
39622	2	44530376	45069527	Gain	CS+aCGH+Ex	539	460	459	55	8	10
39622	4	68815984	68840935	Loss	CS+Ex	25	.	15	10	.	9
39622	5	140482549	140702153	Gain	CS+aCGH+Ex	220	189	220	131	11	123
39622	7	40088117	40141973	Gain	CS+Ex	54	.	54	8	.	8
39622	8	144800458	144808433	Loss	CS+Ex	8	.	8	5	.	5
39638	1	5786233	6005923	Gain	CS+Ex	220	.	123	40	.	29
39638	12	82777014	82836812	Loss	CS	60	.	.	8	.	.
39638	15	31729880	32225059	Gain	CS+aCGH	495	434	.	60	8	.
39638	21	23129395	23254268	Loss	CS	125	.	.	6	.	.
39638	3	1107666	1308155	Gain	CS+aCGH+Ex	200	140	150	24	4	6
39638	X	141881367	141895574	Loss	CS	14	.	.	3	.	.

Table E. CNV de pathogénicité incertaine identifiés chez les 82 participants souffrant d'obésité ou de désordres neurologiques.

Patient	Chr	Start	End	Size	Probes	Type	Parental origin	Gene	Phenotype
28189	22	18860192	19020828	160.64	34	Gain	Unk	DGCR6, PRODH	Intellectual disability
29352	1	197089905	197105194	15.29	12	Loss	D	ASPM	Intellectual disability, hypotonia, arthrochhalasis
29352	10	80042528	80202743	160.22	41	Gain	F	TMEM254, PLAC9, ANXA11	Intellectual disability, hypotonia, arthrochhalasis
29352	6	145645449	145856364	210.92	21	Gain	M	EPM2A, FBXO3	Intellectual disability, hypotonia, arthrochhalasis
30028	22	22646288	23364631	718.34	170	Gain	M	GGTLC2	Intellectual disability, behavioral disorder, hypotonia, arthrochhalasis, dysmorphic feature
30028	22	24270019	24495521	225.5	22	Gain	M	SPECC1L, ADORA2A, UPB1	Intellectual disability, behavioral disorder, hypotonia, arthrochhalasis, dysmorphic feature
30226	22	21958243	22215308	257.07	52	Gain	Unk	TOP3B	Intellectual disability, overweight
30226	5	8704478	9305400	600.92	79	Gain	Unk	SEMA5A	Intellectual disability, overweight
30226	5	11266288	11521607	255.32	27	Gain	Unk	CTNND2	Intellectual disability, overweight
32996	9	136429641	136439455	9.81	10	Loss	F	INPP5E	Intellectual disability, short stature, thinness, dysmorphic feature
35144	17	1252252	1317996	65.75	8	Loss	Unk	BHLHA9, TUSC5	Intellectual disability, behavioral disorder, strabismus, sleepwalking
36499	1	1452011	1495837	43.83	26	Loss	F	ATAD3C, ATAD3B	Intellectual disability, dysmorphic feature
36616	16	21501382	21827897	326.52	65	Loss	Unk	METTL9, IGSF6, OTOA	Intellectual disability, behavioral disorder, insomnia
36616	2	44320188	44570056	249.87	38	Loss	Unk	SLC3A1, PREPL, CAMKMT	Intellectual disability, behavioral disorder, insomnia
37146	22	18906344	19020828	114.48	30	Gain	M	DGCR6, PRODH	Intellectual disability
37146	5	42633203	42810867	177.66	30	Gain	F	GHR, CCDC152	Intellectual disability
37146	8	8847017	9224537	377.52	45	Loss	M	MFHAS1, EIR1	Intellectual disability
37465	9	131009179	131164184	155.01	59	Gain	Unk	LAMC3, AIF1L	Intellectual disability, intrauterine growth restriction, short stature
37588	15	100048822	100341928	293.11	39	Gain	F	ADAMTS17	Intellectual disability, sleep disorder

38100	15	93000485	93020341	19.86	7	Gain	Unk	<i>CHD2</i>	Intellectual disability, neonatal hypotonia, intrauterine growth restriction, dysmorphic feature
38114	17	156185	733221	577.04	97	Gain	M	<i>DOC2B, RPH3AL, LOC100506388, C17orf97, RFLNB, VPS53, FAM57A</i>	Intellectual disability, behavioral disorder
38114	2	178432074	178451044	18.97	8	Gain	M	<i>PRKRA</i>	Intellectual disability, behavioral disorder
38206	20	14719244	14867459	148.22	16	Loss	F	<i>MACROD2</i>	Intellectual disability
38785	16	29599186	29779652	180.47	30	Gain	M	<i>SPN, QPRT, C16orf54, ZG16</i>	Intellectual disability, behavioral disorder
38818	20	14719244	14887403	168.16	18	Loss	Unk	<i>MACROD2</i>	Intellectual disability, dyspraxia, arachnoid cysts
38991	8	8230670	8282256	51.59	9	Gain	Unk	-	Intellectual disability, obesity, hypotonia, malformation
39402	20	14768964	14897332	128.37	14	Loss	Unk	<i>MACROD2</i>	Intellectual disability, behavioral disorder
39402	3	97610768	97899303	288.54	48	Gain	Unk	<i>EPHA6, ARL6, CRYBG3</i>	Intellectual disability, behavioral disorder
39409	1	1457101	1486699	29.6	15	Loss	Unk	<i>ATAD3C, ATAD3B</i>	Intellectual disability, behavioral disorder
39409	20	14176304	14564346	388.04	46	Loss	Unk	<i>MACROD2, FLRT3</i>	Intellectual disability, behavioral disorder
39509	8	15693720	15951771	258.05	34	Loss	M	<i>TUSC3</i>	Intellectual disability, seizures
39622	1	12745877	13005656	259.78	33	Gain	Unk	<i>C1orf158, PRAMEF12, PRAMEF1, PRAMEF11, HNRNPCL1, PRAMEF2, PRAMEF4, PRAMEF10, PRAMEF7, PRAMEF6</i>	Intellectual disability, behavioral disorder
39622	7	40088117	40141973	53.86	8	Gain	Unk	<i>CDK13</i>	Intellectual disability, behavioral disorder
39638	1	5786233	6005923	219.69	40	Gain	Unk	<i>KCNAB2, NPHP4</i>	Intellectual disability
OSV11	15	99847446	99956044	108.6	12	Loss	Unk	-	Intellectual disability, behavioral disorder, morbid obesity
OSV14	17	30944895	30988132	43.24	10	Gain	F	<i>ADAP2, RNF135</i>	Intellectual disability, isolated obesity, malformation
OSV17	22	20700678	21209085	508.41	156	Gain	F	<i>PI4KA, SERPIND1, SNAP29, CRKL, AIFM3, LZTR1, THAP7, P2RX6, SLC7A4, LRRC74B, GGT2</i>	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature, cross-eye, seizures
OSV33	2	178432074	178451894	19.82	10	Gain	M	<i>PRKRA</i>	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature
OSV44	2	178432074	178451044	18.97	8	Gain	Unk	<i>PRKRA</i>	Intellectual disability, obesity
OSV46	12	66965108	67044837	79.73	12	Loss	Unk	-	Seizures, behavioral disorder, hypermetropia, obesity, dysmorphic feature
OSV48	1	1450651	1495837	45.19	27	Loss	M	<i>ATAD3C, ATAD3B</i>	Intellectual disability, isolated obesity, behavioral disorder, astigmatism, myopia, dyspraxia

OSV48	1	145509900	145809184	299.29	57	Gain	M	<i>GPR89A, PDZK1, CD160, RNF115</i>	Intellectual disability, isolated obesity, behavioral disorder, astigmatism, myopia, dyspraxia
OSV6	20	14749069	14867459	118.39	13	Loss	F	<i>MACROD2</i>	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature
OSV82	22	22687018	22747958	60.94	19	Loss	F	-	Intellectual disability, seizures, obesity, acromicria, oculomotor apraxia, behavioral disorder
U2	22	21955538	22215308	259.77	54	Gain	Unk	<i>TOP3B</i>	Intellectual disability, severe obesity

Toutes les coordonnées se basent sur le génome de référence en version hg38.

Chr, Chromosome; **D**, de novo; **F**, father; **M**, mother; **Ref**, référence; **Unk**, unknown.

Table F. Découvertes fortuites identifiées chez les 82 participants souffrant d'obésité ou de désordres neurologiques.

<i>Patient</i>	<i>Sex*</i>	<i>Gene</i>	<i>Inheritance</i>	<i>Mutation</i>	<i>Status</i>	<i>Parental origin**</i>	<i>rs ID</i>	<i>Pathogenicity</i>	<i>ACMG criteria</i>	<i>Ref</i>
39465	F	<i>WFS1</i>	AD/AR	c.409_424dup p.Val142Glyfs*110	/ het	M	rs767046229	Pathogenic	PVS1, PS1, PP2, PP4	(493)
OSV17	F	<i>KCNJ11</i>	AD	c.151G>T / p.Glu51*	het	M	-	Pathogenic	PVS1, PS1, PM2, PP2, PP4	(494)
OSV24	M	<i>CPA6</i>	AD/AR	c.544C>T / p.Arg182*	het	M	rs773734224	Pathogenic	PVS1, PS1, PP2, PP4	ClinVar RCV000171414.1
OSV24	M	<i>STAP1</i>	AD	c.827-1del / -	het	F	-	Pathogenic	PVS1, PM2, PP2, PP4	-
OSV48	M	<i>PPARG</i>	AD/AR	c.419G>A / p.Arg140His	het	M	rs773089349	Pathogenic	PS1, PS3, PP2, PP4	(495)
OSV58	M	<i>POGZ</i>	AD	c.3026G>A / p.Arg1009His	het	Unk	-	Likely pathogenic	PM1, PM2, PP2, PP4	-
OSV71	M	<i>HRAS</i>	AD	c.367C>T / p.Arg123Cys	het	Unk	rs369106578	Likely pathogenic	PS1, PP2, PP3, PP4	ClinVar RCV000157922.2
OSV86	M	<i>HCN2</i>	AD	c.1269C>G / p.Ser423Arg	het	D	-	Likely pathogenic	PS2, PM2, PP2, PP4	-

ACMG, American College of Medical Genetics and Genomics; **AD**, autosomal dominant; **AR**, autosomal recessive; **Chr**, Chromosome; **D**, de novo; **F***, female; **F****, father; **het**, heterozygous; **M***, male; **M****, mother; **PM**, moderate pathogenic criterion; **PP**, supporting pathogenic criterion; **PS**, strong pathogenic criterion; **PVS**, very strong pathogenic criterion; **Ref**, reference; **Unk**, unknown.

Table G. Mutations ponctuelles de signification incertaine identifiées chez les 82 participants souffrant d'obésité ou de désordres neurologiques.

Patient	Sex*	Gene	Inheritance	Mutation	Status	Parental origin*	rs ID	ACMG criteria	Phenotype
28189	M	FLNA	XL	c.481G>A p.Glu161Lys	/ hem	Unk	-	PM2, PP2, PP4	Intellectual disability
28189	M	CHD7	AD	c.994T>G p.Leu332Val	/ het	Unk	-	PM2, PP2, PP4	Intellectual disability
28189	M	MRAP2	AD	c.140T>C / p.Ile47Thr	het	Unk	rs79682316	PP2, PP3, PP4	Intellectual disability
29352	M	MYT1L	AD	c.2657C>T p.Ala886Val	/ het	M	-	PM2, PP2, PP3, PP4	Intellectual disability, hypotonia, arthrochalis
30226	M	CLTC	AD	c.554T>C p.Met185Thr	/ het	Unk	-	PM2, PP2, PP3, PP4	Intellectual disability, overweight
31425	M	EP300	AD	c.1058G>A p.Arg353His	/ het	Unk	-	PP2, PP3, PP4	Intellectual disability, behavioral disorder, malformation
32776	M	ANKRD11	AD	c.1011G>C p.Gln337His	/ het	M	-	PM2, PP2, PP4	Intellectual disability, behavioral disorder
32996	F	ACAN	AD/AR	c.742G>A p.Ala248Thr	/ het	M	rs564289325	PP1, PP2, PP4	Intellectual disability, short stature, thinness, dysmorphic feature
32996	F	LRP2	AD/AR	c.5306A>T p.Asp1769Val	/ het	F	-	PM2, PP2, PP3, PP4	Intellectual disability, short stature, thinness, dysmorphic feature
36499	F	CHD8	AD	c.7304G>A p.Gly2435Asp	/ het	M	rs758415263	PM2, PP2, PP4	Intellectual disability, dysmorphic feature
36578	F	CHRNA2	AD	c.767C>A p.Thr256Asn	/ het	Unk	-	PM2, PP2, PP3, PP4	Intellectual disability, behavioral disorder, obesity, dysmorphic feature
36578	F	RELN	AD/AR	c.3431A>G p.Asn1144Ser	/ het	Unk	-	PM2, PP2, PP4	Intellectual disability, behavioral disorder, obesity, dysmorphic feature
36616	M	MC4R	AD	c.587G>A p.Cys196Tyr	/ het	Unk	-	PM1, PP2, PP3, PP4	Intellectual disability, behavioral disorder, insomnia
36616	M	CREBBP	AD	c.1000A>G p.Ile334Val	/ het	Unk	-	PM2, PP2, PP4	Intellectual disability, behavioral disorder, insomnia
36616	M	RELN	AD/AR	c.2245T>A p.Phe749Ile	/ het	Unk	-	PM2, PP2, PP3, PP4	Intellectual disability, behavioral disorder, insomnia
36711	M	HNF4A	AD	c.1012C>T p.Leu338Phe	/ het	F	-	PM2, PP2, PP4	Intellectual disability, behavioral disorder
37465	M	CHRNA4	AD	c.542C>T p.Ser181Phe	/ het	Unk	-	PM2, PP2, PP3, PP4	Intellectual disability, intrauterine growth restriction, short stature
37465	M	PAFAH1B1	AD	c.40G>A / p.Ala14Thr	het	Unk	rs770331294	PP2, PP3, PP4	Intellectual disability, intrauterine growth restriction, short stature
37588	M	CLCN2	AD/AR	c.2168C>T p.Ala723Val	/ het	M	rs200869847	PM2, PP2, PP4	Intellectual disability, sleep disorder

37588	M	CLCN2	AD/AR	c.2555C>T p.Pro852Leu	/	het	M	-	PM2, PP2, PP3, PP4	Intellectual disability, sleep disorder
38100	F	NRXN2	AD	c.3236C>G p.Thr1079Ser	/	het	Unk	-	PM2, PP2, PP3, PP4	Intellectual disability, neonatal hypotonia, intrauterine growth restriction, dysmorphic feature
38506	M	CACNA1D	AD	c.1105G>A p.Val369Met	/	het	Unk	rs148674323	PP2, PP3, PP4	Intellectual disability, behavioral disorder, overweight
38506	M	APPL1	AD	c.2009T>C p.Ile670Thr	/	het	Unk	-	PM2, PP2, PP4	Intellectual disability, behavioral disorder, overweight
38645	M	OFD1	XL	c.397A>G p.Lys133Glu	/	hem	Unk	rs759975626	PP2, PP3, PP4	Intellectual disability
38645	M	RAB39B	XL	c.627_629del p.Arg210del	/	hem	Unk	rs864309529	PM4, PP2, PP4	Intellectual disability
38785	M	DCX	XL	c.1064C>T p.Ser355Leu	/	hem	M	rs781679748	PM2, PP2, PP4	Intellectual disability, behavioral disorder
38839	M	ASH1L	AD	c.5861C>T p.Pro1954Leu	/	het	Unk	rs200062254	PM2, PP2, PP4	Intellectual disability, neonatal hypotonia, thinness
39268	M	RELN	AD/AR	c.6925G>A p.Asp2309Asn	/	het	Unk	rs138978280	PP2, PP3, PP4	Intellectual disability, behavioral disorder, heart disorder, seizures
39465	F	NSD1	AD	c.1765T>C p.Ser589Pro	/	het	M	-	PM2, PP2, PP4	Intellectual disability, behavioral disorder, dysmorphic features
39612	M	RELN	AD/AR	c.2615A>G p.Asp872Gly	/	het	Unk	-	PM2, PP2, PP4	Intellectual disability
39622	F	TCF20	AD	c.4217A>G p.Glu1406Gly	/	het	Unk	-	PM2, PP2, PP4	Intellectual disability, behavioral disorder
39638	F	TCF20	AD	c.3629A>G p.Lys1210Arg	/	het	Unk	-	PM2, PP2, PP4	Intellectual disability
OSV11	F	ASH1L	AD	c.7598G>A p.Arg2533His	/	het	Unk	rs138122556	PP2, PP3, PP4	Intellectual disability, behavioral disorder, morbid obesity
OSV11	F	SMARCB1	AD	c.158G>A p.Arg53Gln	/	het	Unk	rs779769475	PM5, PP2, PP4	Intellectual disability, behavioral disorder, morbid obesity
OSV11	F	MAGI2	AD	c.2515C>T p.Arg839Cys	/	het	Unk	rs769012473	PP2, PP3, PP4	Intellectual disability, behavioral disorder, morbid obesity
OSV17	F	MAST1	AD	c.994C>T p.Arg332Cys	/	het	M	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature, cross-eye, seizures
OSV17	F	AUTS2	AD	c.2082T>G p.Ser694Arg	/	het	F	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature, cross-eye, seizures
OSV17	F	PURA	AD	c.715A>G p.Lys239Glu	/	het	M	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature, cross-eye, seizures
OSV20	M	KCNC1	AD	c.1391C>T p.Pro464Leu	/	het	M	-	PM2, PP2, PP4	Intellectual disability, obesity
OSV20	M	SLC6A1	AD	c.1250G>A p.Arg417His	/	hom	F / M	rs191293931	PM2, PP2, PP4	Intellectual disability, obesity
OSV20	M	GRIN2A	AD	c.414+7C>T / -	/	het	M	-	PM2, PP2, PP3, PP4	Intellectual disability, obesity
OSV24	M	MAST1	AD	c.3092T>C p.Met1031Thr	/	het	M	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, depression, hypersomnia, behavioral disorder
OSV24	M	DDHD2	AR	c.1982A>G p.Tyr661Cys	/	het	M	rs201656753	PM1, PM2, PP2, PP3, PP4	Intellectual disability, morbid obesity, depression, hypersomnia, behavioral disorder

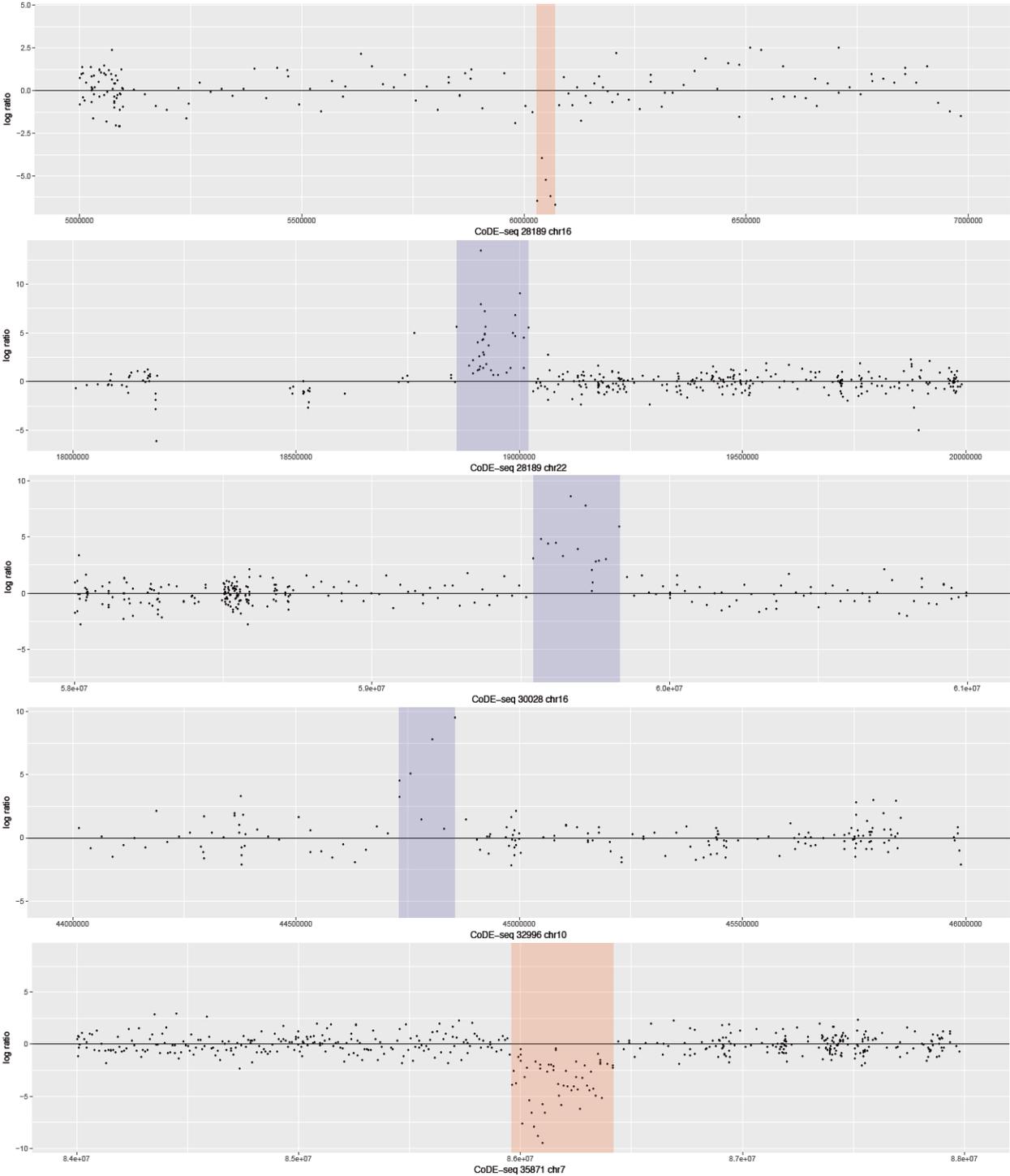
OSV24	M	<i>DDHD2</i>	AR	c.1461+1G>A / -	het	D	-	PVS1, PS2, PM2, PP2, PP4	Intellectual disability, morbid obesity, depression, hypersomnia, behavioral disorder
OSV26	F	<i>CPA6</i>	AD/AR	c.387A>T / p.Arg129Ser	het	Unk	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, dysmorphic features, behavioral disorder
OSV26	F	<i>CNTN4</i>	AD	c.77C>T / p.Pro26Leu	het	Unk	rs146175469	PP2, PP3, PP4	Intellectual disability, morbid obesity, dysmorphic features, behavioral disorder
OSV3	M	<i>PRRT2</i>	AD/AR	c.719G>T / p.Arg240Leu	het	M	-	PM2, PP2, PP4	Intellectual disability, dysmorphic features, behavioral disorder, isolated obesity, cross-eye
OSV30	M	<i>ABCC8</i>	AD	c.3191C>G / p.Ala1064Gly	het	M	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, seizures, behavioral disorder, dymorphic feature
OSV30	M	<i>MBD5</i>	AD	c.1487C>T / p.Thr496Ile	het	F	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, seizures, behavioral disorder, dymorphic feature
OSV30	M	<i>DNM1</i>	AD	c.2192T>C / p.Met731Thr	het	M	-	PM2, PP2, PP3, PP4	Intellectual disability, morbid obesity, seizures, behavioral disorder, dymorphic feature
OSV30	M	<i>DYNC1H1</i>	AD	c.782_783delinsGC / p.Lys261Ser	het	M	-	PM2, PP2, PP3, PP4	Intellectual disability, morbid obesity, seizures, behavioral disorder, dymorphic feature
OSV33	M	<i>EIF4E</i>	AD	c.58_60del / p.Glu20del	het	F	-	PM2, PM4, PP4	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature
OSV34	F	<i>SRCAP</i>	AD	c.5383G>T / p.Val1795Phe	het	Unk	-	PM2, PP2, PP4	Intellectual disability, isolated obesity
OSV41	F	<i>NRXN2</i>	AD	c.1223C>T / p.Pro408Leu	het	F	rs573733911	PP2, PP3, PP4	Intellectual disability, obesity, dysmorphic feature
OSV41	F	<i>BCOR</i>	XL	c.2072T>G / p.Leu691Arg	het	M	-	PM2, PP2, PP4	Intellectual disability, obesity, dysmorphic feature
OSV44	F	<i>SMC1A</i>	XL	c.1811A>C / p.His604Pro	het	Unk	-	PM2, PP2, PP4	Intellectual disability, obesity
OSV44	F	<i>PDX1</i>	AD/AR	c.14A>G / p.Glu5Gly	het	Unk	-	PM2, PP2, PP3, PP4	Intellectual disability, obesity
OSV46	F	<i>SLC35A2</i>	XL	c.623C>T / p.Ala208Val	het	Unk	-	PM2, PP2, PP4	Seizures, behavioral disorder, hypermetropia, obesity, dysmorphic feature
OSV46	F	<i>TSC2</i>	AD	c.2750G>A / p.Arg917Gln	het	Unk	rs397515046	PM1, PP2, PP4	Seizures, behavioral disorder, hypermetropia, obesity, dysmorphic feature
OSV48	M	<i>CASR</i>	AD/AR	c.2785A>G / p.Arg929Gly	het	M	-	PM2, PP2, PP4	Intellectual disability, isolated obesity, behavioral disorder, astigmatism, myopia, dyspraxia
OSV52	M	<i>CHD8</i>	AD	c.6734G>T / p.Gly2245Val	het	Unk	-	PM2, PP2, PP4	Intellectual disability, behavioral disorder, obesity, facial features
OSV52	M	<i>ZMYM3</i>	XL	c.3271C>T / p.Arg1091Cys	hem	Unk	rs773814463	PM2, PP2, PP4	Intellectual disability, behavioral disorder, obesity, facial features
OSV53	M	<i>CDK13</i>	AD	c.2048G>A / p.Cys683Tyr	het	Unk	-	PM2, PP2, PP3, PP4	Intellectual disability, morbid obesity, hypermetropia
OSV53	M	<i>USP9X</i>	XL	c.4417G>A / p.Val1473Ile	hem	Unk	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, hypermetropia
OSV53	M	<i>ARID1A</i>	AD	c.2615T>C / p.Met872Thr	het	Unk	rs140568310	PP2, PP3, PP4	Intellectual disability, morbid obesity, hypermetropia
OSV53	M	<i>ANK2</i>	AD	c.6143A>G / p.Gln2048Arg	het	Unk	rs201628725	PP2, PP3, PP4	Intellectual disability, morbid obesity, hypermetropia
OSV6	M	<i>TBX1</i>	AD	c.205G>A / p.Ala69Thr	het	Unk	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature

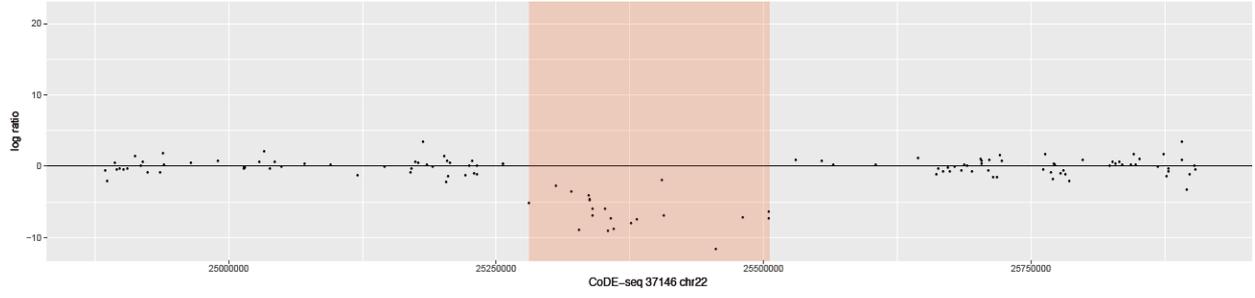
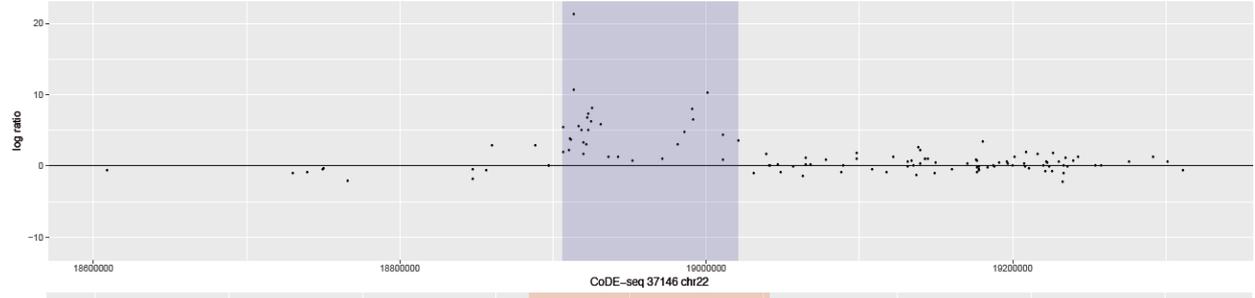
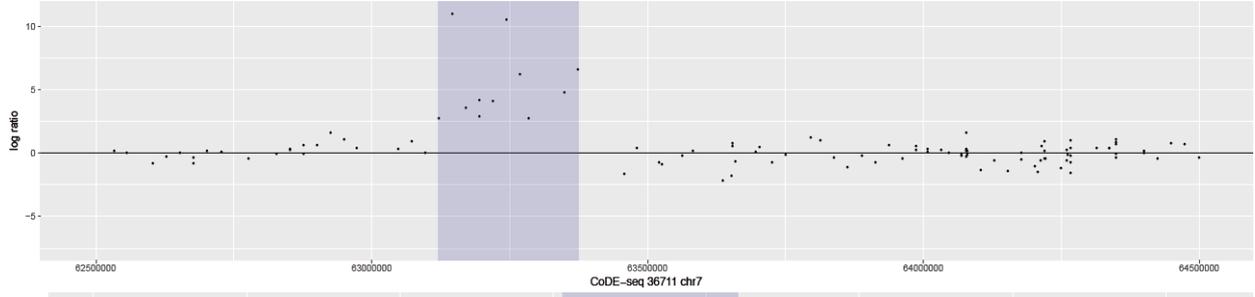
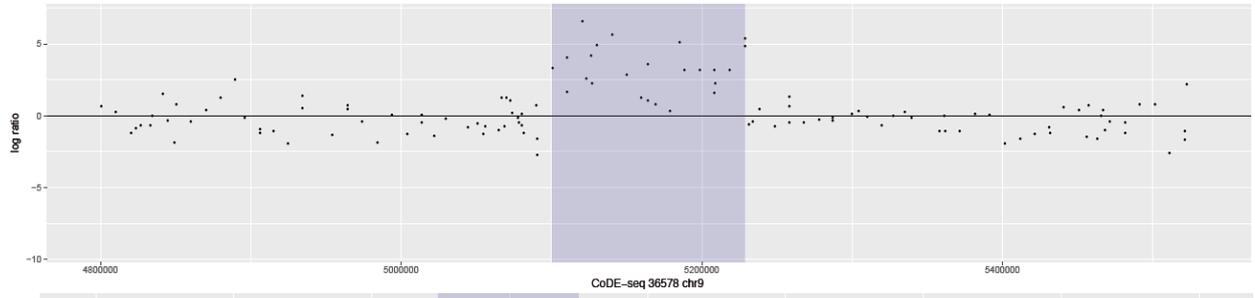
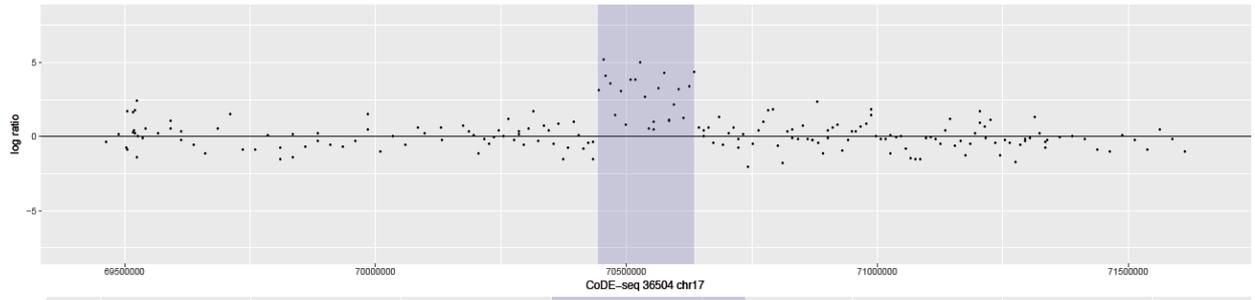
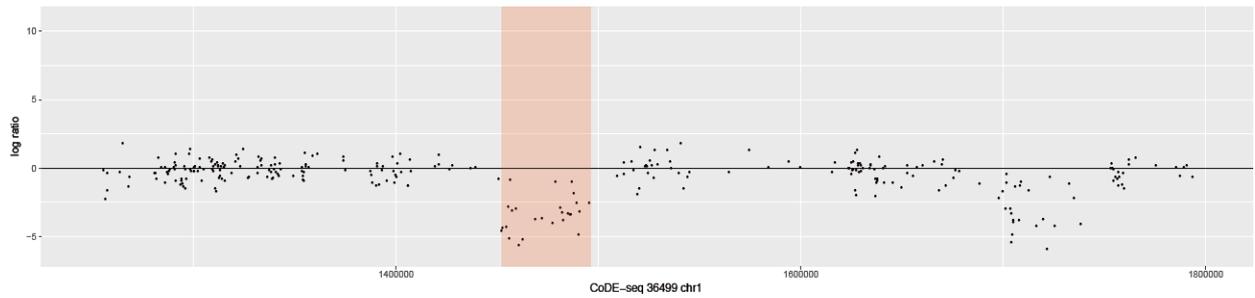
OSV65	F	<i>QRICH1</i>	AD	c.1334G>A p.Cys445Tyr	/	het	M	-	PM2, PP2, PP4	Intellectual disability, obesity, hypermetropia, behavioral disorder
OSV70	M	<i>GRIN2B</i>	AD	c.1087G>C p.Val363Leu	/	het	F	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature
OSV71	M	<i>MED12</i>	XL	c.1996A>G p.Met666Val	/	hem	Unk	-	PM2, PP2, PP3, PP4	Behavioral disorder, obesity, dysmorphic feature, malformation, hypermetropia
OSV74	F	<i>MAGEL2</i>	AD	c.2054C>T p.Ser685Phe	/	het	Unk	-	PM2, PP2, PP4	Intellectual disability, isolated obesity, delayed bone age, myopia, astigmatism, repetitive seromucous otitis
OSV74	F	<i>NRXN2</i>	AD	c.154G>A p.Glu52Lys	/	het	Unk	-	PM2, PP2, PP4	Intellectual disability, isolated obesity, delayed bone age, myopia, astigmatism, repetitive seromucous otitis
OSV77	M	<i>SLC17A5</i>	AD	c.617C>A p.Ala206Glu	/	het	F	-	PM2, PP2, PP3, PP4	Behavioral disorder, isolated morbid obesity, dysmorphic feature
OSV77	M	<i>ACSL4</i>	XL	c.545A>G p.Lys182Arg	/	hem	M	-	PM2, PP2, PP4	Behavioral disorder, isolated morbid obesity, dysmorphic feature
OSV77	M	<i>LRP2</i>	AD/AR	c.6986T>C p.Phe2329Ser	/	het	M	rs780651375	PP2, PP3, PP4	Behavioral disorder, isolated morbid obesity, dysmorphic feature
OSV80	F	<i>CDH10</i>	AD	c.2116G>C p.Asp706His	/	het	Unk	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature
OSV80	F	<i>ANK2</i>	AD	c.9853A>C p.Ile3285Leu	/	het	Unk	rs144603824	PP2, PP3, PP4	Intellectual disability, morbid obesity, behavioral disorder, dysmorphic feature
OSV82	F	<i>DPP6</i>	AD	c.366T>G p.Ile122Met	/	het	M	rs758478344	PM2, PP2, PP4	Intellectual disability, seizures, obesity, acromicria, oculomotor apraxia, behavioral disorder
OSV9	F	<i>NFIX</i>	AD	c.1304T>C p.Val435Ala	/	het	F	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, malformation
U1	M	<i>MYT1L</i>	AD	c.1400A>G p.Asp467Gly	/	hom	Unk	-	PM2, PP2, PP4	Intellectual disability, morbid obesity, neonatal hypotonia, one kidney, dysmorphic feature, acanthosis nigricans, visual impairment
U1	M	<i>CLTC</i>	AD	c.3262-3T>C / -	/	het	Unk	-	PM2, PP2, PP3, PP4	Intellectual disability, morbid obesity, neonatal hypotonia, one kidney, dysmorphic feature, acanthosis nigricans, visual impairment
U3	M	<i>TSC1</i>	AD	c.3311G>A p.Cys1104Tyr	/	het	M	-	PP1, PP2, PP4	Morbid obesity, macrocephaly, abdominal lipoma
U3	M	<i>ACAN</i>	AD/AR	c.1682G>A p.Arg561His	/	het	F	rs376881991	PP2, PP3, PP4	Morbid obesity, macrocephaly, abdominal lipoma
U5	M	<i>ZDHH9</i>	XL	c.1030C>G p.Pro344Ala	/	hem	M	-	PM2, PP2, PP4	Intellectual disability, behavioral disorder, seizures, neonatal hypotonia, kyphosis, cross-eye, ataxic gait, sleep disorder, no language
U6	F	<i>ZEB2</i>	AD	c.2752C>T p.Pro918Ser	/	het	Unk	-	PM2, PP2, PP4	Morbid obesity, tall stature, hyperlipidemia
U7	M	<i>SMARCA4</i>	AD	c.808C>T p.Pro270Ser	/	het	Unk	-	PM2, PP2, PP4	Morbid obesity, none adiposity rebound
U8	M	<i>FOXP2</i>	AD	c.1082G>A p.Gly361Glu	/	het	Unk	-	PM2, PP2, PP4	Morbid obesity, liver steatosis, cytolysis, hypertriglyceridemia, glucose intolerance
U9	M	<i>STAP1</i>	AD	c.521C>T p.Pro174Leu	/	het	Unk	-	PM2, PP2, PP4	Intellectual disability, brachydactyly, isolated morbid obesity
U9	M	<i>MED13L</i>	AD/AR	c.535A>G p.Thr179Ala	/	hom	Unk	-	PM2, PP2, PP4	Intellectual disability, brachydactyly, isolated morbid obesity

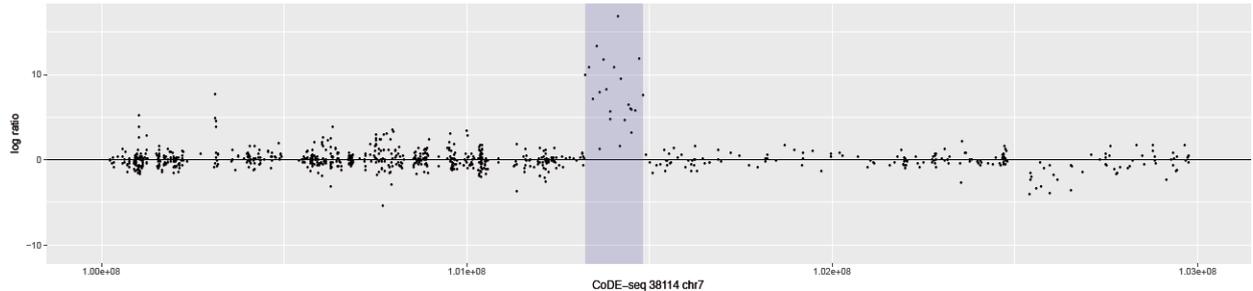
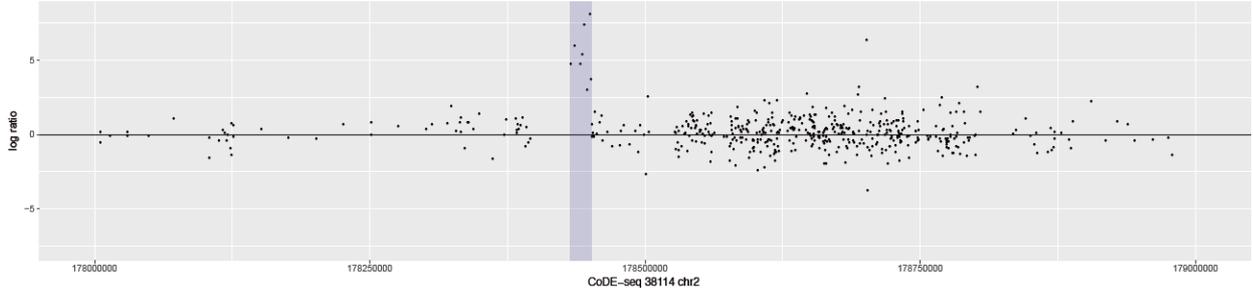
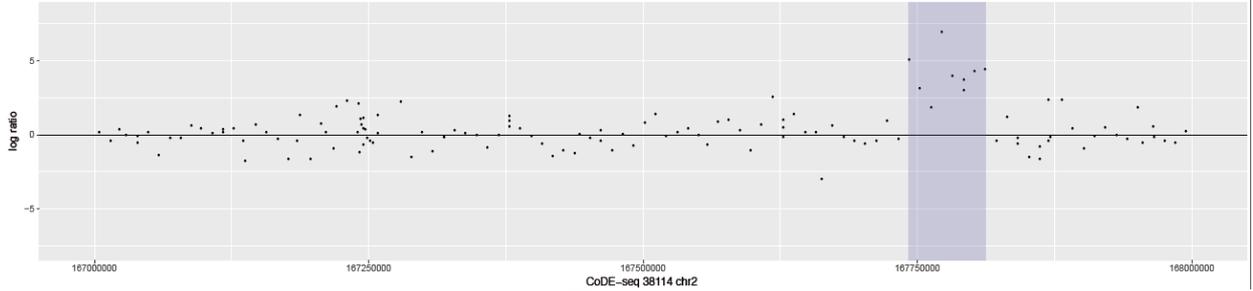
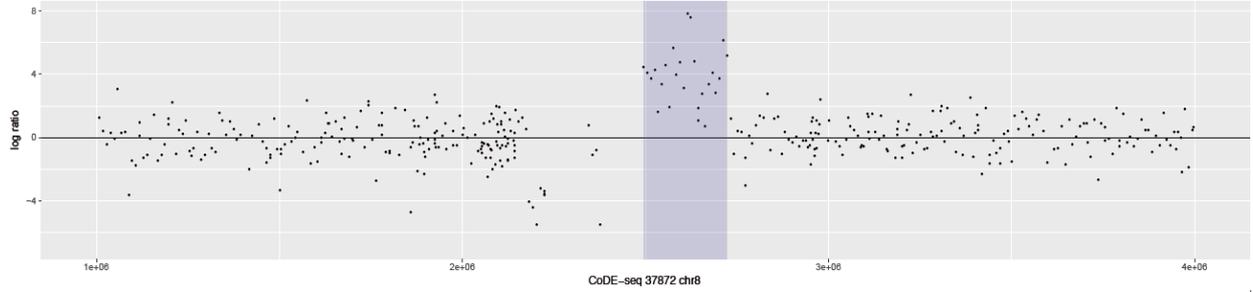
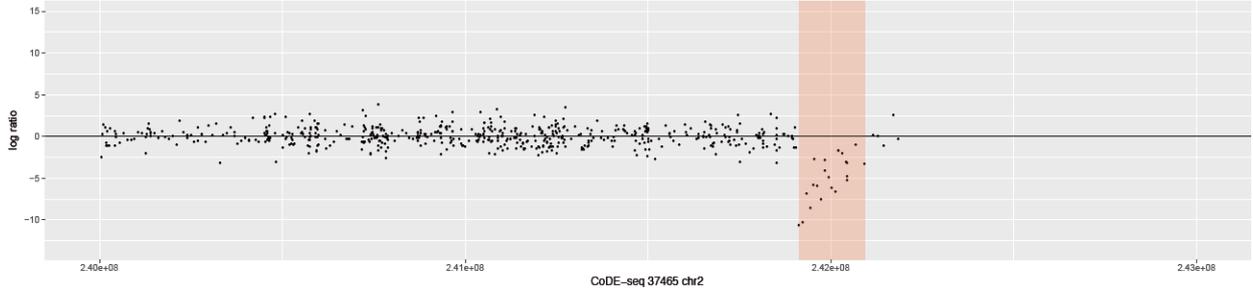
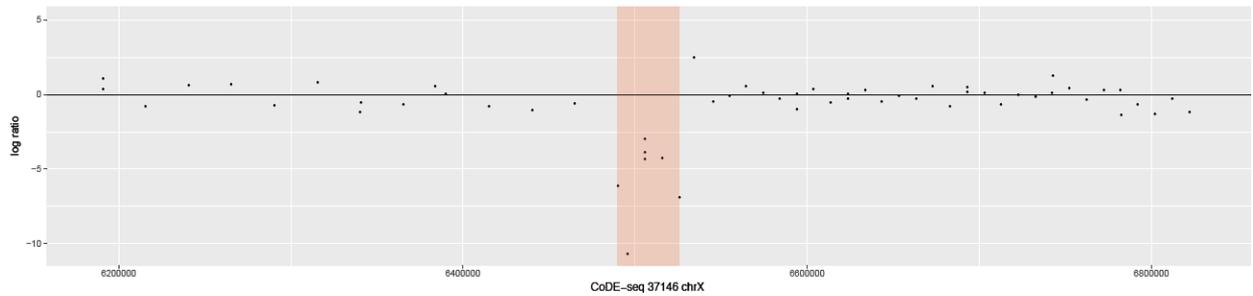
U9	M	<i>KMT2D</i>	AD	c.5324C>G p.Ala1775Gly	/	het	F	-	PM2, PP2, PP3, PP4	Intellectual disability, brachydactyly, isolated morbid obesity
U10	F	<i>CTCF</i>	AD	c.324A>G p.Ile108Met	/	het	Unk	-	PM2, PP2, PP4	Morbid obesity, early-onset severe type 2 diabetes, macrosomia at birth, dyslipidemia
U11	M	<i>SMARCA4</i>	AD	c.968C>T p.Ser323Leu	/	het	Unk	-	PM2, PP2, PP4	Morbid obesity, insulin resistance, enuresis
U12	F	<i>SMARCA4</i>	AD	c.968C>T p.Ser323Leu	/	het	Unk	-	PM2, PP2, PP4	Isolated morbid obesity, none adiposity rebound, high blood pressure, insulin resistance, sleep apnea, hyperuricemia

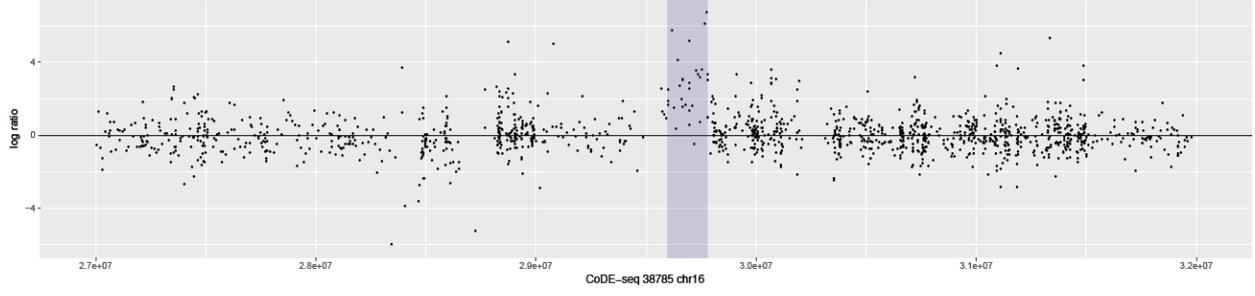
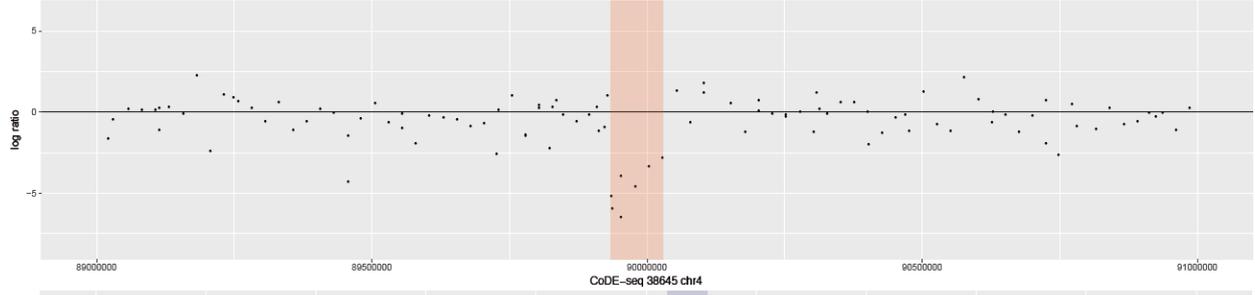
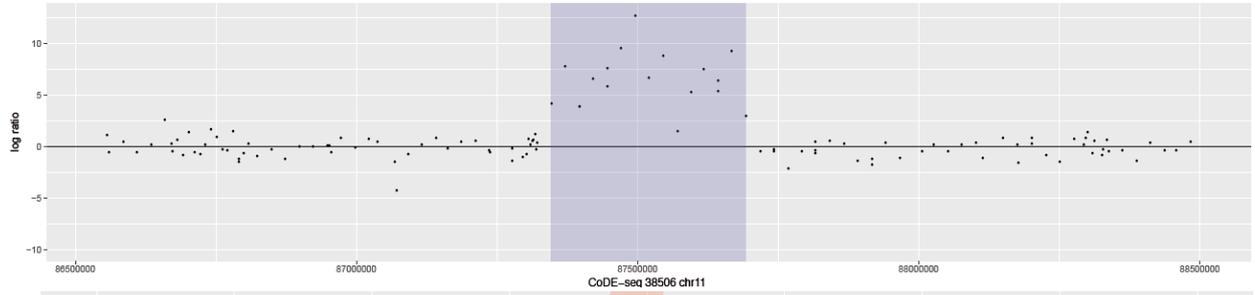
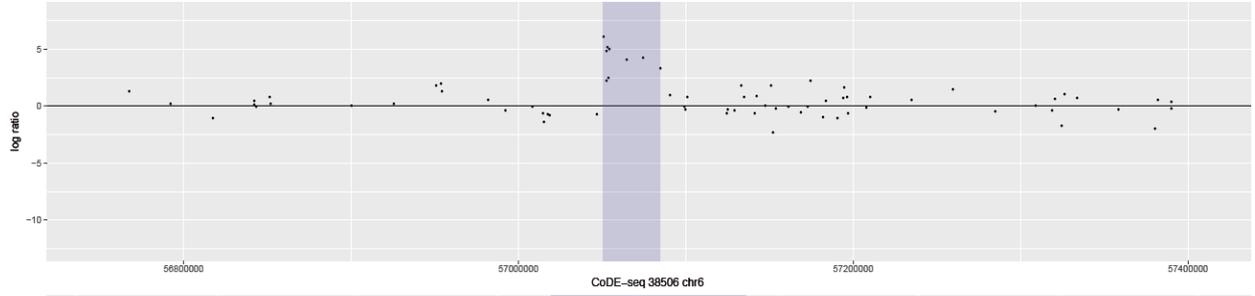
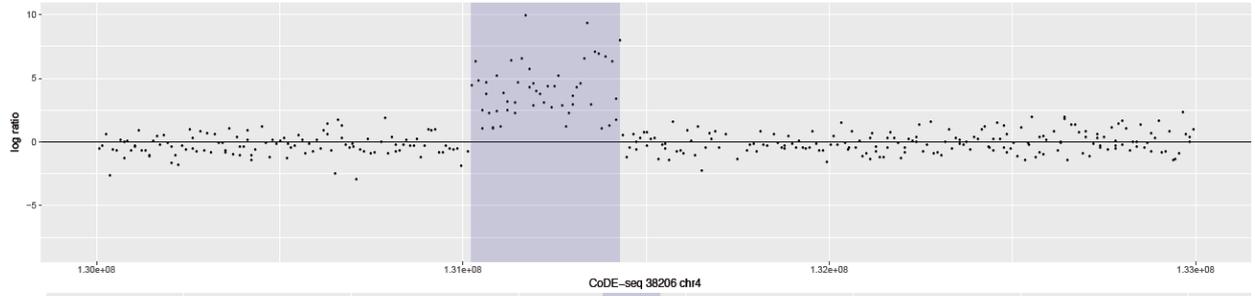
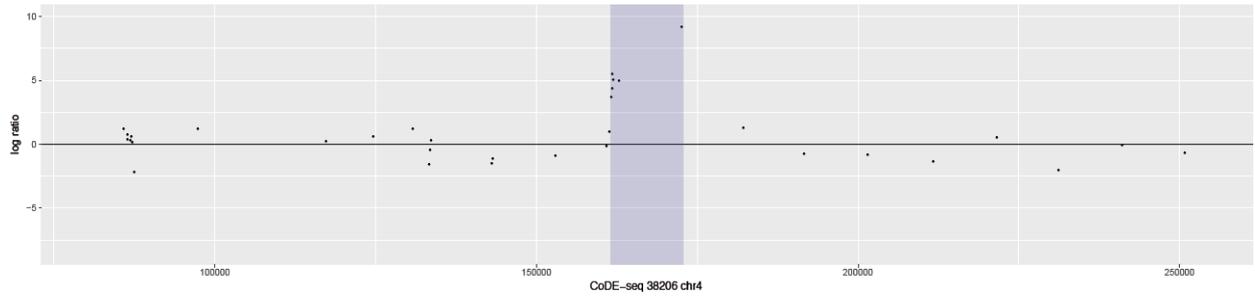
ACMG, American College of Medical Genetics and Genomics; **AD**, autosomal dominant; **AR**, autosomal recessive; **Chr**, Chromosome; **D**, de novo; **F***, female; **F****, father; **hem**, hemizygous; **het**, heterozygous; **hom**, homozygous; **M***, male; **M****, mother; **PM**, moderate pathogenic criterion; **PP**, supporting pathogenic criterion; **PS**, strong pathogenic criterion; **PVS**, very strong pathogenic criterion; **Ref**, reference; **Unk**, unknown; **XL**, X-linked.

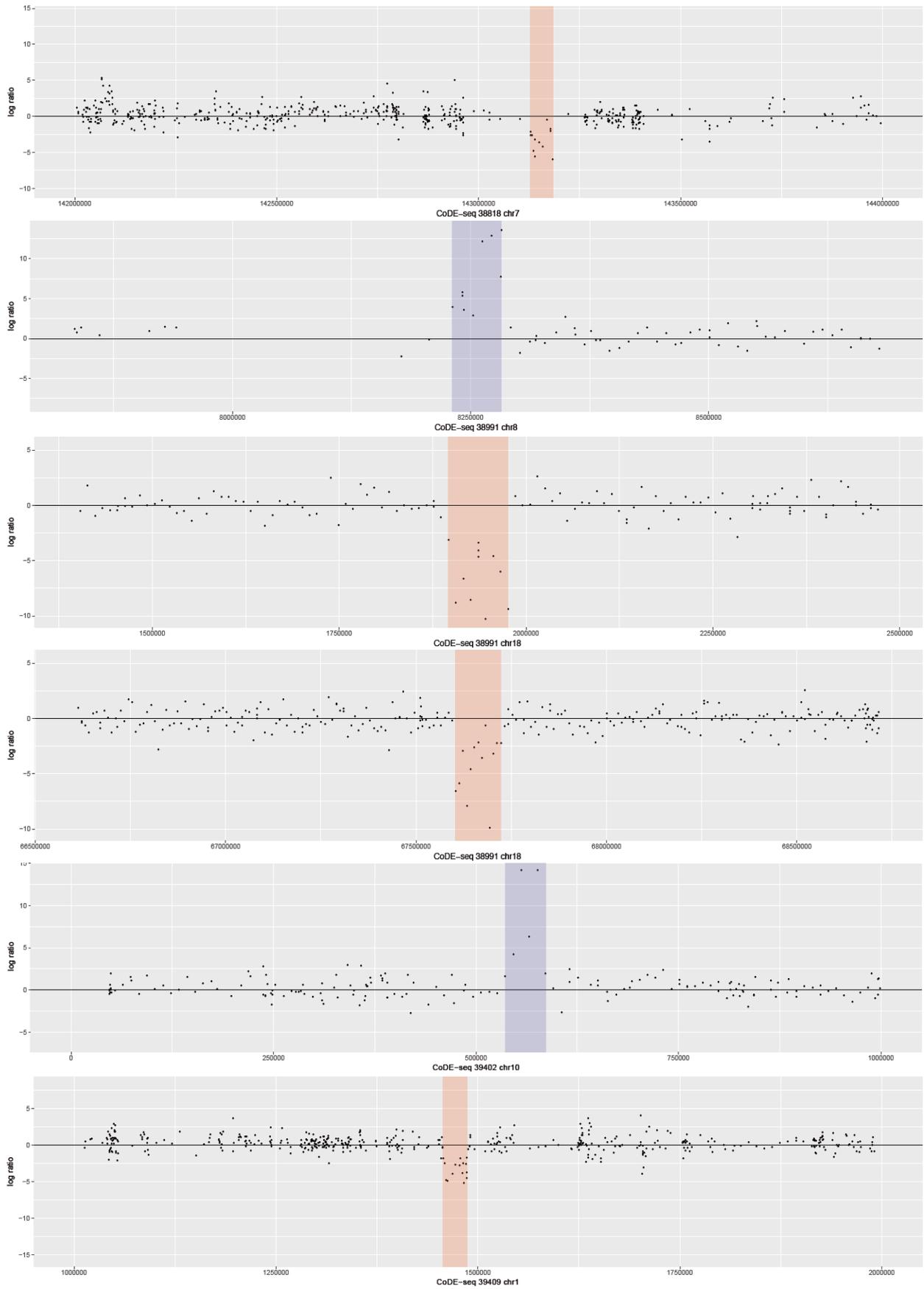
Figure A. Profils des 42 CNV (délétions en rouge / duplications en bleu) détectés par CoDE-seq uniquement chez les 40 participants de la cohorte de l'étude pilote. Les points représentent les sondes utilisées, avec en ordonnée le log-ratio de profondeur du patient analysé par rapport à la ligne de base de l'ensemble des autres patients de l'étude, en en abscisse les positions génomiques en bases.

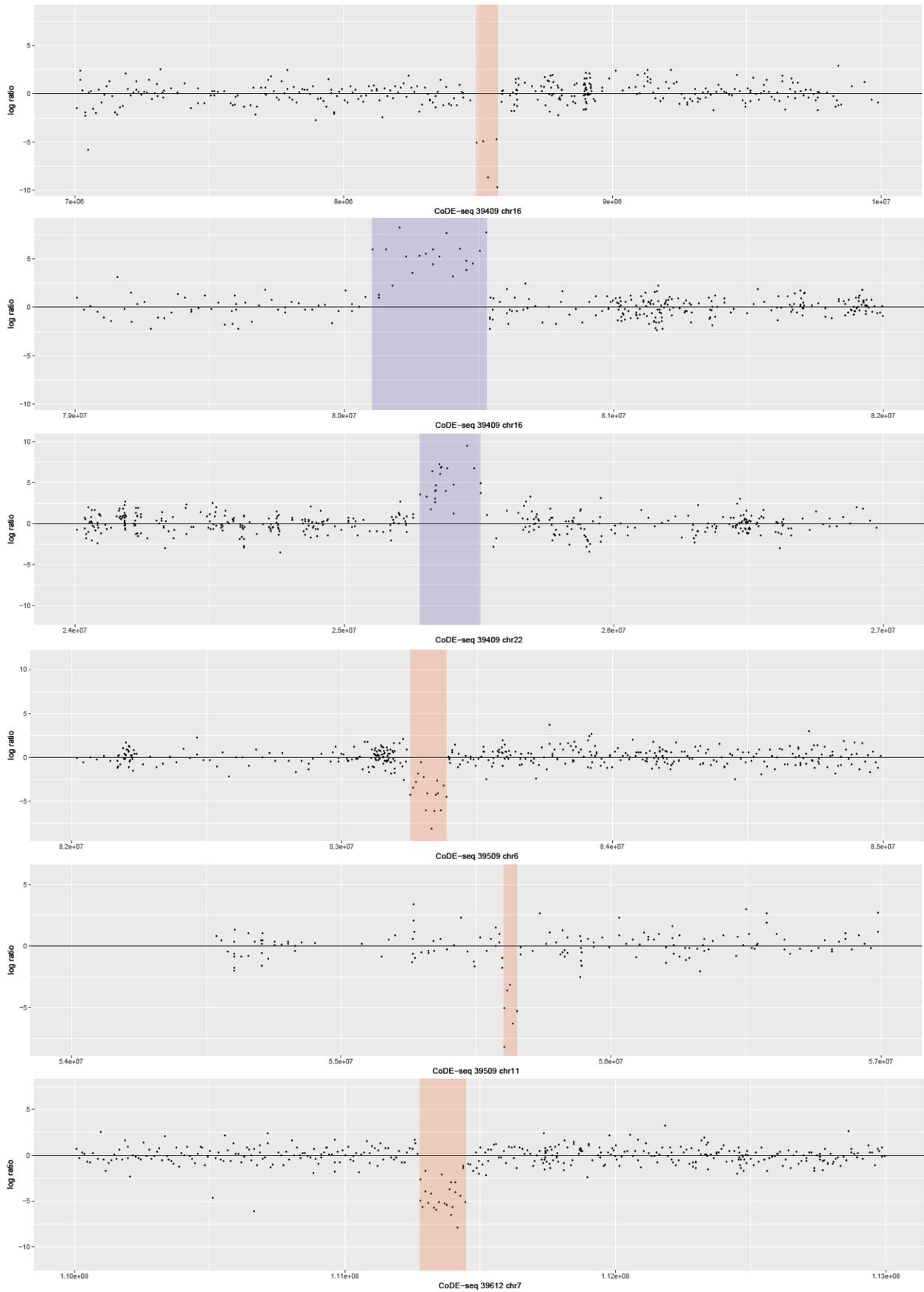


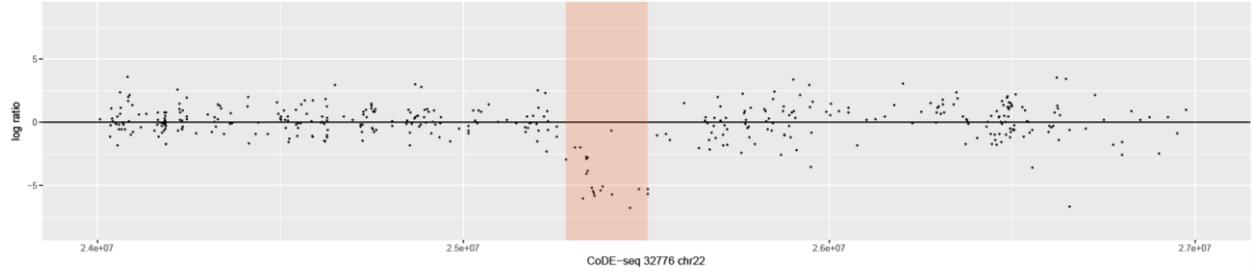
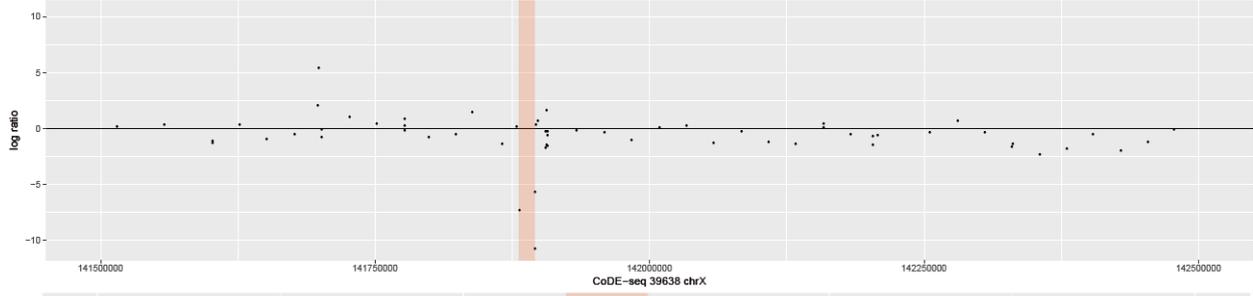
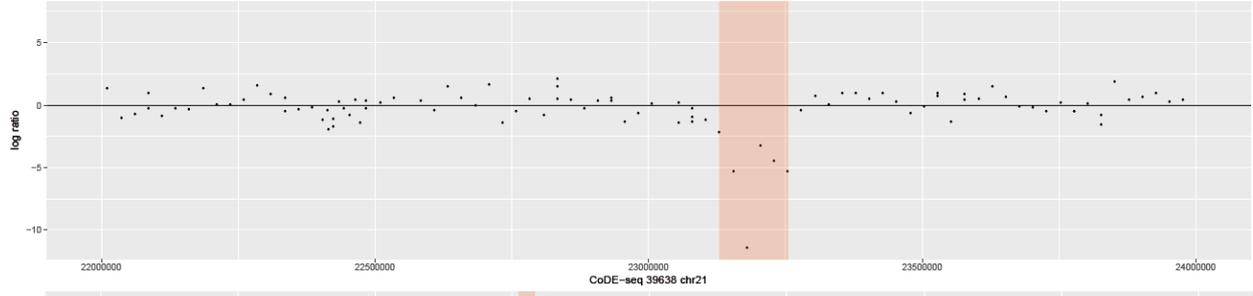
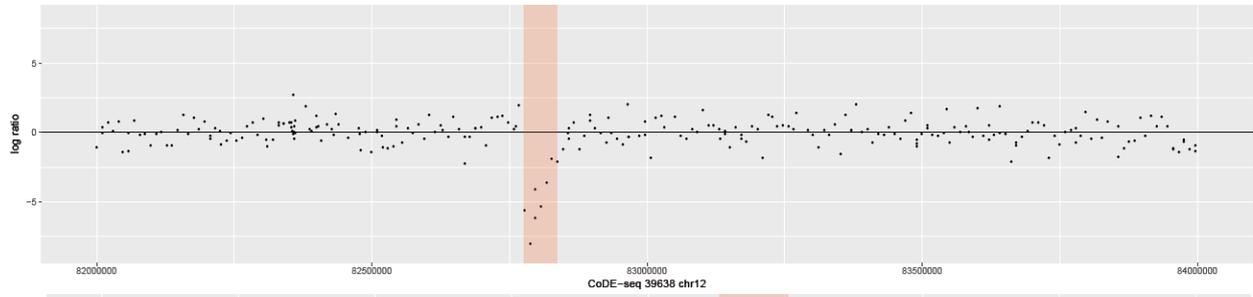
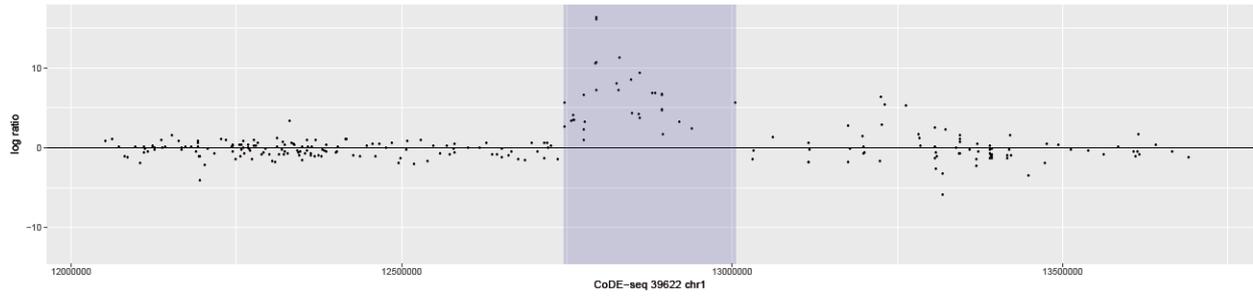
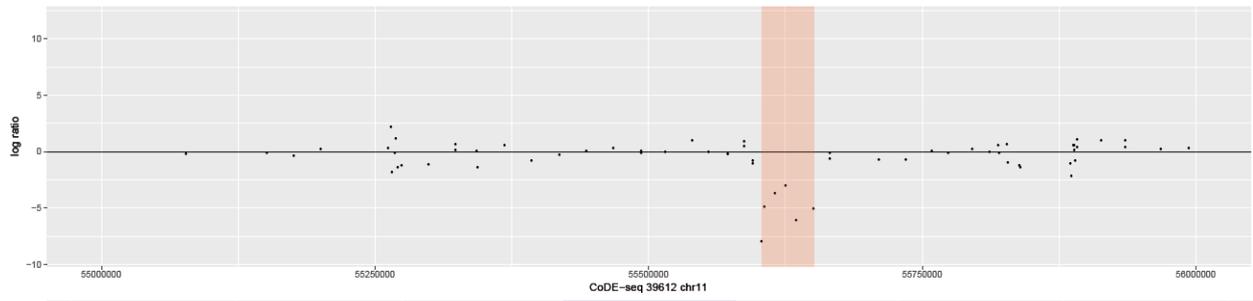


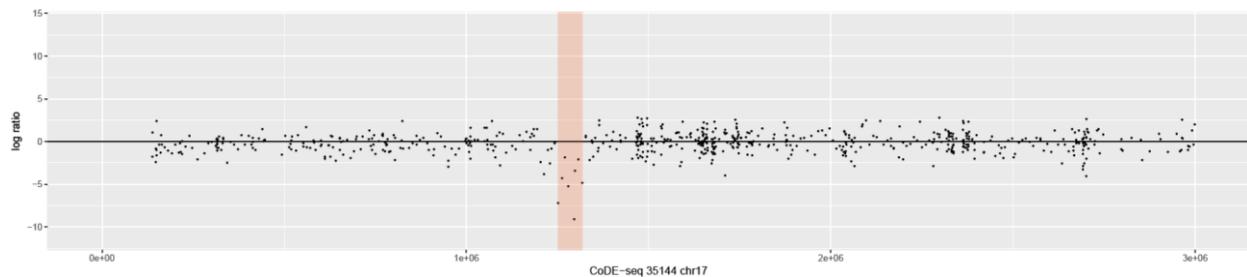












References Supplementary Data :

- [1] Slavotinek, A., Shaffer, L.G., Shapira, S.K., 1999. Monosomy 1p36. *Journal of Medical Genetics* 36(9): 657–63.
- [2] Knight-Jones, E., Knight, S., Heussler, H., Regan, R., Flint, J., Martin, K., 2000. Neurodevelopmental profile of a new dysmorphic syndrome associated with submicroscopic partial deletion of 1p36.3. *Developmental Medicine and Child Neurology* 42(3): 201–6.
- [3] Zenker, M., Rittinger, O., Grosse, K.P., Speicher, M.R., Kraus, J., Rauch, A., et al., 2002. Monosomy 1p36—a recently delineated, clinically recognizable syndrome. *Clinical Dysmorphology* 11(1): 43–8.
- [4] Heilstedt, H.A., Ballif, B.C., Howard, L.A., Lewis, R.A., Stal, S., Kashork, C.D., et al., 2003. Physical Map of 1p36, Placement of Breakpoints in Monosomy 1p36, and Clinical Characterization of the Syndrome. *The American Journal of Human Genetics* 72(5): 1200–12, Doi: 10.1086/375179.
- [5] Heilstedt, H.A., Ballif, B.C., Howard, L.A., Kashork, C.D., Shaffer, L.G., 2003. Population data suggest that deletions of 1p36 are a relatively common chromosome abnormality. *Clinical Genetics* 64(4): 310–6.
- [6] Gajecka, M., Mackay, K.L., Shaffer, L.G., 2007. Monosomy 1p36 deletion syndrome. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics* 145C(4): 346–56, Doi: 10.1002/ajmg.c.30154.
- [7] Jacquemont, M.-L., Sanlaville, D., Redon, R., Raoul, O., Cormier-Daire, V., Lyonnet, S., et al., 2006. Array-based comparative genomic hybridisation identifies high frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders. *Journal of Medical Genetics* 43(11): 843–9, Doi: 10.1136/jmg.2006.043166.
- [8] Kirov, G., Pocklington, A.J., Holmans, P., Ivanov, D., Ikeda, M., Ruderfer, D., et al., 2012. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Molecular Psychiatry* 17(2): 142–53, Doi: 10.1038/mp.2011.154.
- [9] Klopocki, E., Schulze, H., Strauß, G., Ott, C.-E., Hall, J., Trotier, F., et al., 2007. Complex Inheritance Pattern Resembling Autosomal Recessive Inheritance Involving a Microdeletion in Thrombocytopenia–Absent Radius Syndrome. *The American Journal of Human Genetics* 80(2): 232–40, Doi: 10.1086/510919.
- [10] Christiansen, J., 2004. Chromosome 1q21.1 Contiguous Gene Deletion Is Associated With Congenital Heart Disease. *Circulation Research* 94(11): 1429–35, Doi: 10.1161/01.RES.0000130528.72330.5c.
- [11] Sharp, A.J., Hansen, S., Selzer, R.R., Cheng, Z., Regan, R., Hurst, J.A., et al., 2006. Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nature Genetics* 38(9): 1038–42, Doi: 10.1038/ng1862.
- [12] Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., et al., 2006. Global variation in copy number in the human genome. *Nature* 444(7118): 444–54, Doi: 10.1038/nature05329.
- [13] Stone, J.L., O'Donovan, M.C., Gurling, H., Kirov, G.K., Blackwood, D.H.R., Corvin, A., et al., 2008. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 455(7210): 237–41, Doi: 10.1038/nature07239.
- [14] Stefansson, H., Rujescu, D., Cichon, S., Pietiläinen, O.P.H., Ingason, A., Steinberg, S., et al., 2008. Large recurrent microdeletions associated with schizophrenia. *Nature* 455(7210): 232–6, Doi: 10.1038/nature07229.
- [15] Mefford, H.C., Sharp, A.J., Baker, C., Itsara, A., Jiang, Z., Buysse, K., et al., 2008. Recurrent Rearrangements of Chromosome 1q21.1 and Variable Pediatric Phenotypes. *New England Journal of Medicine* 359(16): 1685–99, Doi: 10.1056/NEJMoa0805384.
- [16] Brunetti-Pierrri, N., Berg, J.S., Scaglia, F., Belmont, J., Bacino, C.A., Sahoo, T., et al., 2008. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nature Genetics* 40(12): 1466–71, Doi: 10.1038/ng.279.
- [17] Rosenfeld, J.A., Traylor, R.N., Schaefer, G.B., McPherson, E.W., Ballif, B.C., Klopocki, E., et al., 2012. Proximal microdeletions and microduplications of 1q21.1 contribute to variable abnormal phenotypes. *European Journal of Human Genetics* 20(7): 754–61, Doi: 10.1038/ejhg.2012.6.
- [18] Szatmari, P., Paterson, A.D., Zwaigenbaum, L., Roberts, W., Brian, J., Liu, X.-Q., et al., 2007. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nature Genetics* 39(3): 319–28, Doi: 10.1038/ng1985.
- [19] Christian, S.L., Brune, C.W., Sudi, J., Kumar, R.A., Liu, S., Karamohamed, S., et al., 2008. Novel Submicroscopic Chromosomal Abnormalities Detected in Autism Spectrum Disorder. *Biological Psychiatry* 63(12): 1111–7, Doi: 10.1016/j.biopsych.2008.01.009.

- [20] Parvari, R., Brodyansky, I., Elpeleg, O., Moses, S., Landau, D., Hershkovitz, E., 2001. A Recessive Contiguous Gene Deletion of Chromosome 2p16 Associated with Cystinuria and a Mitochondrial Disease. *The American Journal of Human Genetics* 69(4): 869–75, Doi: 10.1086/323624.
- [21] Parvari, R., Gonen, Y., Alshafee, I., Buriakovsky, S., Regev, K., Hershkovitz, E., 2005. The 2p21 deletion syndrome: Characterization of the transcription content. *Genomics* 86(2): 195–211, Doi: 10.1016/j.ygeno.2005.04.001.
- [22] Jaeken, J., Martens, K., François, I., Eyskens, F., Lecointre, C., Derua, R., et al., 2006. Deletion of PREPL, a Gene Encoding a Putative Serine Oligopeptidase, in Patients with Hypotonia-Cystinuria Syndrome. *The American Journal of Human Genetics* 78(1): 38–51, Doi: 10.1086/498852.
- [23] Chabrol, B., Martens, K., Meulemans, S., Cano, A., Jaeken, J., Matthijs, G., et al., 2008. Deletion of C2orf34, PREPL and SLC3A1 causes atypical hypotonia-cystinuria syndrome. *Journal of Medical Genetics* 45(5): 314–8, Doi: 10.1136/jmg.2007.055475.
- [24] Martens, K., Jaeken, J., Matthijs, G., Creemers, J.W.M., 2008. Multi-system disorder syndromes associated with cystinuria type I. *Current Molecular Medicine* 8(6): 544–50.
- [25] Rajcan-Separovic, E., Harvard, C., Liu, X., McGillivray, B., Hall, J.G., Qiao, Y., et al., 2007. Clinical and molecular cytogenetic characterisation of a newly recognised microdeletion syndrome involving 2p15-16.1. *Journal of Medical Genetics* 44(4): 269–76, Doi: 10.1136/jmg.2006.045013.
- [26] de Leeuw, N., Pfundt, R., Koolen, D.A., Neefs, I., Scheltinga, I., Mieloo, H., et al., 2007. A newly recognised microdeletion syndrome involving 2p15p16.1: narrowing down the critical region by adding another patient detected by genome wide tiling path array comparative genomic hybridisation analysis. *Journal of Medical Genetics* 45(2): 122–4, Doi: 10.1136/jmg.2007.054049.
- [27] FitzPatrick, D.R., 2003. Identification of SATB2 as the cleft palate gene on 2q32-q33. *Human Molecular Genetics* 12(19): 2491–501, Doi: 10.1093/hmg/ddg248.
- [28] Van Buggenhout, G., Van Ravenswaaij-Arts, C., MC Maas, N., Thoelen, R., Vogels, A., Smeets, D., et al., 2005. The del(2)(q32.2q33) deletion syndrome defined by clinical and molecular characterization of four patients. *European Journal of Medical Genetics* 48(3): 276–89, Doi: 10.1016/j.ejmg.2005.05.005.
- [29] Aldred, M.A., Sanford, R.O.C., Thomas, N.S., Barrow, M.A., Wilson, L.C., Brueton, L.A., et al., 2004. Molecular analysis of 20 patients with 2q37.3 monosomy: definition of minimum deletion intervals for key phenotypes. *Journal of Medical Genetics* 41(6): 433–9.
- [30] Shrimpton, A., Braddock, B., Thomson, L., Stein, C., Hoo, J., 2004. Molecular delineation of deletions on 2q37.3 in three cases with an Albright hereditary osteodystrophy-like phenotype: Molecular delineation of deletions on 2q37.3. *Clinical Genetics* 66(6): 537–44, Doi: 10.1111/j.1399-0004.2004.00363.x.
- [31] Roohi, J., Montagna, C., Tegang, D.H., Palmer, L.E., DeVincent, C., Pomeroy, J.C., et al., 2009. Disruption of contactin 4 in three subjects with autism spectrum disorder. *Journal of Medical Genetics* 46(3): 176–82, Doi: 10.1136/jmg.2008.057505.
- [32] Willatt, L., Cox, J., Barber, J., Cabanas, E.D., Collins, A., Donnai, D., et al., 2005. 3q29 Microdeletion Syndrome: Clinical and Molecular Characterization of a New Syndrome. *The American Journal of Human Genetics* 77(1): 154–60, Doi: 10.1086/431653.
- [33] Ballif, B.C., Theisen, A., Coppinger, J., Gowans, G.C., Hersh, J.H., Madan-Khetarpal, S., et al., 2008. Expanding the clinical phenotype of the 3q29 microdeletion syndrome and characterization of the reciprocal microduplication. *Molecular Cytogenetics* 1(1): 8, Doi: 10.1186/1755-8166-1-8.
- [34] Lisi, E.C., Hamosh, A., Doheny, K.F., Squibb, E., Jackson, B., Galczynski, R., et al., 2008. 3q29 interstitial microduplication: A new syndrome in a three-generation family. *American Journal of Medical Genetics Part A* 146A(5): 601–9, Doi: 10.1002/ajmg.a.32190.
- [35] Zollino, M., Di Stefano, C., Zampino, G., Mastroiacovo, P., Wright, T.J., Sorge, G., et al., 2000. Genotype-phenotype correlations and clinical diagnostic criteria in Wolf-Hirschhorn syndrome. *American Journal of Medical Genetics* 94(3): 254–61.
- [36] Marcelis, C., Schrandt-Stumpel, C., Engelen, J., Schoonbrood-Lenssen, A., Willemsse, A., Beemer, F., et al., 2001. Wolf-Hirschhorn (4P-) syndrome in adults. *Genetic Counseling (Geneva, Switzerland)* 12(1): 35–48.
- [37] Shannon, N.L., Maltby, E.L., Rigby, A.S., Quarrell, O.W., 2001. An epidemiological study of Wolf-Hirschhorn syndrome: life expectancy and cause of mortality. *Journal of Medical Genetics* 38(10): 674–9.
- [38] Zollino, M., Lecce, R., Fischetto, R., Murdolo, M., Faravelli, F., Selicorni, A., et al., 2003. Mapping the Wolf-Hirschhorn Syndrome Phenotype Outside the Currently Accepted WHS Critical Region and Defining a New Critical Region, WHSCR-2. *The American Journal of Human Genetics* 72(3): 590–7, Doi: 10.1086/367925.
- [39] Medina, M., Marinescu, R.C., Overhauser, J., Kosik, K.S., 2000. Hemizyosity of δ -Catenin (CTNND2) Is Associated with Severe Mental Retardation in Cri-du-Chat Syndrome. *Genomics* 63(2): 157–64, Doi: 10.1006/geno.1999.6090.
- [40] Van Buggenhout, G.J., Pijckels, E., Holvoet, M., Schaap, C., Hamel, B.C., Fryns, J.P., 2000. Cri du chat syndrome: changing phenotype in older patients. *American Journal of Medical Genetics* 90(3): 203–15.
- [41] Mainardi, P.C., Perfumo, C., Cali, A., Coucourde, G., Pastore, G., Cavani, S., et al., 2001. Clinical and molecular characterisation of 80 patients with 5p deletion: genotype-phenotype correlation. *Journal of Medical Genetics* 38(3): 151–8.
- [42] Mosca-Boidron, A.-L., Gueneau, L., Huguette, G., Goldenberg, A., Henry, C., Gigot, N., et al., 2016. A de novo microdeletion of SEMA5A in a boy with autism spectrum disorder and intellectual disability. *European Journal of Human Genetics* 24(6): 838–43, Doi: 10.1038/ejhg.2015.211.

- [43] Sieber, O.M., Lamlum, H., Crabtree, M.D., Rowan, A.J., Barclay, E., Lipton, L., et al., 2002. Whole-gene APC deletions cause classical familial adenomatous polyposis, but not attenuated polyposis or “multiple” colorectal adenomas. *Proceedings of the National Academy of Sciences* 99(5): 2954–8, Doi: 10.1073/pnas.042699199.
- [44] Michils, G., Tejpar, S., Thoelen, R., Cutsem, E. van., Vermeesch, J.R., Fryns, J.-P., et al., 2005. Large deletions of the APC gene in 15% of mutation-negative patients with classical polyposis (FAP): A Belgian study. *Human Mutation* 25(2): 125–34, Doi: 10.1002/humu.20122.
- [45] Kurotaki, N., Imaizumi, K., Harada, N., Masuno, M., Kondoh, T., Nagai, T., et al., 2002. Haploinsufficiency of NSD1 causes Sotos syndrome. *Nature Genetics* 30(4): 365–6, Doi: 10.1038/ng863.
- [46] Kurotaki, N., Harada, N., Shimokawa, O., Miyake, N., Kawame, H., Uetake, K., et al., 2003. Fifty microdeletions among 112 cases of Sotos syndrome: Low copy repeats possibly mediate the common deletion. *Human Mutation* 22(5): 378–87, Doi: 10.1002/humu.10270.
- [47] Kurotaki, N., 2004. Sotos syndrome common deletion is mediated by directly oriented subunits within inverted Sos-REP low-copy repeats. *Human Molecular Genetics* 14(4): 535–42, Doi: 10.1093/hmg/ddi050.
- [48] Tatton-Brown, K., 2005. Multiple mechanisms are implicated in the generation of 5q35 microdeletions in Sotos syndrome. *Journal of Medical Genetics* 42(4): 307–13, Doi: 10.1136/jmg.2004.027755.
- [49] Tatton-Brown, K., Douglas, J., Coleman, K., Baujat, G., Cole, T.R.P., Das, S., et al., 2005. Genotype-Phenotype Associations in Sotos Syndrome: An Analysis of 266 Individuals with NSD1 Aberrations. *The American Journal of Human Genetics* 77(2): 193–204, Doi: 10.1086/432082.
- [50] Tatton-Brown, K., Rahman, N., 2007. Sotos syndrome. *European Journal of Human Genetics* 15(3): 264–71, Doi: 10.1038/sj.ejhg.5201686.
- [51] Depienne, C., Heron, D., Betancur, C., Benyahia, B., Trouillard, O., Bouteiller, D., et al., 2007. Autism, language delay and mental retardation in a patient with 7q11 duplication. *Journal of Medical Genetics* 44(7): 452–8, Doi: 10.1136/jmg.2006.047092.
- [52] Somerville, M.J., Mervis, C.B., Young, E.J., Seo, E.-J., del Campo, M., Bamforth, S., et al., 2005. Severe Expressive-Language Delay Related to Duplication of the Williams–Beuren Locus. *New England Journal of Medicine* 353(16): 1694–701, Doi: 10.1056/NEJMoa051962.
- [53] Committee on Genetics, 2001. American Academy of Pediatrics: Health care supervision for children with Williams syndrome. *Pediatrics* 107(5): 1192–204.
- [54] Osborne, L.R., Li, M., Pober, B., Chitayat, D., Bodurtha, J., Mandel, A., et al., 2001. A 1.5 million-base pair inversion polymorphism in families with Williams-Beuren syndrome. *Nature Genetics* 29(3): 321–5, Doi: 10.1038/ng753.
- [55] Eronen, M., Peippo, M., Hiippala, A., Raatikka, M., Arvio, M., Johansson, R., et al., 2002. Cardiovascular manifestations in 75 patients with Williams syndrome. *Journal of Medical Genetics* 39(8): 554–8.
- [56] Bayés, M., Magano, L.F., Rivera, N., Flores, R., Pérez Jurado, L.A., 2003. Mutational mechanisms of Williams-Beuren syndrome deletions. *American Journal of Human Genetics* 73(1): 131–51.
- [57] Tassabehji, M., 2003. Williams-Beuren syndrome: a challenge for genotype-phenotype correlations. *Human Molecular Genetics* 12(suppl 2): R229–37, Doi: 10.1093/hmg/ddg299.
- [58] Morris, C.A., Mervis, C.B., Hobart, H.H., Gregg, R.G., Bertrand, J., Ensing, G.J., et al., 2003. GTF2I hemizyosity implicated in mental retardation in Williams syndrome: Genotype-phenotype analysis of five families with deletions in the Williams syndrome region. *American Journal of Medical Genetics* 123A(1): 45–59, Doi: 10.1002/ajmg.a.20496.
- [59] Tassabehji, M., 2005. GTF2IRD1 in Craniofacial Development of Humans and Mice. *Science* 310(5751): 1184–7, Doi: 10.1126/science.1116142.
- [60] Del Campo, M., Antonell, A., Magano, L.F., Muñoz, F.J., Flores, R., Bayés, M., et al., 2006. Hemizyosity at the NCF1 Gene in Patients with Williams-Beuren Syndrome Decreases Their Risk of Hypertension. *The American Journal of Human Genetics* 78(4): 533–42, Doi: 10.1086/501073.
- [61] Meyer-Lindenberg, A., Mervis, C.B., Faith Berman, K., 2006. Neural mechanisms in Williams syndrome: a unique window to genetic influences on cognition and behaviour. *Nature Reviews Neuroscience* 7(5): 380–93, Doi: 10.1038/nrn1906.
- [62] Sanders, S.J., Ercan-Sencicek, A.G., Hus, V., Luo, R., Murtha, M.T., Moreno-De-Luca, D., et al., 2011. Multiple Recurrent De Novo CNVs, Including Duplications of the 7q11.23 Williams Syndrome Region, Are Strongly Associated with Autism. *Neuron* 70(5): 863–85, Doi: 10.1016/j.neuron.2011.05.002.
- [63] Palmer, S.E., Scherer, S.W., Kukulich, M., Wijsman, E.M., Tsui, L.C., Stephens, K., et al., 1994. Evidence for locus heterogeneity in human autosomal dominant split hand/split foot malformation. *American Journal of Human Genetics* 55(1): 21–6.
- [64] Jarvik, G.P., Patton, M.A., Homfray, T., Evans, J.P., 1994. Non-Mendelian transmission in a human developmental disorder: split hand/split foot. *American Journal of Human Genetics* 55(4): 710–3.
- [65] Scherer, S.W., Poorkaj, P., Allen, T., Kim, J., Geshuri, D., Nunes, M., et al., 1994. Fine mapping of the autosomal dominant split hand/split foot locus on chromosome 7, band q21.3-q22.1. *American Journal of Human Genetics* 55(1): 12–20.
- [66] Genuardi, M., Pomponi, M.G., Sammito, V., Bellussi, A., Zollino, M., Neri, G., 1993. Split hand/split foot anomaly in a family segregating a balanced translocation with breakpoint on 7q22.1. *American Journal of Medical Genetics* 47(6): 823–31, Doi: 10.1002/ajmg.1320470606.
- [67] Crackower, M.A., Scherer, S.W., Rommens, J.M., Hui, C.C., Poorkaj, P., Soder, S., et al., 1996. Characterization of the split hand/split foot malformation locus SHFM1 at 7q21.3-q22.1 and analysis of a candidate gene for its expression during limb development. *Human Molecular Genetics* 5(5): 571–9.

- [68] Ignatius, J., Knuutila, S., Scherer, S.W., Trask, B., Kere, J., 1996. Split hand/split foot malformation, deafness, and mental retardation with a complex cytogenetic rearrangement involving 7q21.3. *Journal of Medical Genetics* 33(6): 507–10.
- [69] Tackels-Horne, D., Toburen, A., Sangiorgi, E., Gurrieri, F., de Mollerat, X., Fischetto, R., et al., 2001. Split hand/split foot malformation with hearing loss: first report of families linked to the SHFM1 locus in 7q21. *Clinical Genetics* 59(1): 28–36.
- [70] Robledo, R.F., 2002. The *Dlx5* and *Dlx6* homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. *Genes & Development* 16(9): 1089–101, Doi: 10.1101/gad.988402.
- [71] Faivre, L., Morichon-Delvallez, N., Viot, G., Narcy, F., Loison, S., Mandelbrot, L., et al., 1998. Prenatal diagnosis of an 8p23.1 deletion in a fetus with a diaphragmatic hernia and review of the literature. *Prenatal Diagnosis* 18(10): 1055–60.
- [72] Devriendt, K., Matthijs, G., Van Dael, R., Gewillig, M., Eyskens, B., Hjalgrim, H., et al., 1999. Delineation of the critical deletion region for congenital heart defects, on chromosome 8p23.1. *American Journal of Human Genetics* 64(4): 1119–26.
- [73] Pehlivan, T., Pober, B.R., Brueckner, M., Garrett, S., Slaugh, R., Van Rheeden, R., et al., 1999. GATA4 haploinsufficiency in patients with interstitial deletion of chromosome region 8p23.1 and congenital heart disease. *American Journal of Medical Genetics* 83(3): 201–6.
- [74] Shimokawa, O., Miyake, N., Yoshimura, T., Sosonkina, N., Harada, N., Mizuguchi, T., et al., 2005. Molecular characterization of del(8)(p23.1p23.1) in a case of congenital diaphragmatic hernia. *American Journal of Medical Genetics Part A* 136A(1): 49–51, Doi: 10.1002/ajmg.a.30778.
- [75] Slavotinek, A., 2005. Fryns syndrome phenotype caused by chromosome microdeletions at 15q26.2 and 8p23.1. *Journal of Medical Genetics* 42(9): 730–6, Doi: 10.1136/jmg.2004.028787.
- [76] Hollox, E.J., Barber, J.C.K., Brookes, A.J., Armour, J.A.L., 2008. Defensins and the dynamic genome: What we can learn from structural variation at human chromosome band 8p23.1. *Genome Research* 18(11): 1686–97, Doi: 10.1101/gr.080945.108.
- [77] Aldhous, M.C., Abu Bakar, S., Prescott, N.J., Palla, R., Soo, K., Mansfield, J.C., et al., 2010. Measurement methods and accuracy in copy number variation: failure to replicate associations of beta-defensin copy number with Crohn's disease. *Human Molecular Genetics* 19(24): 4930–8, Doi: 10.1093/hmg/ddq411.
- [78] Ballarati, L., Cereda, A., Caselli, R., Selicorni, A., Recalcati, M.P., Maitz, S., et al., 2011. Genotype–phenotype correlations in a new case of 8p23.1 deletion and review of the literature. *European Journal of Medical Genetics* 54(1): 55–9, Doi: 10.1016/j.ejmg.2010.10.003.
- [79] Barber, J.C.K., Rosenfeld, J.A., Graham, J.M., Kramer, N., Lachlan, K.L., Bateman, M.S., et al., 2015. Inside the 8p23.1 duplication syndrome; eight microduplications of likely or uncertain clinical significance. *American Journal of Medical Genetics Part A* 167(9): 2052–64, Doi: 10.1002/ajmg.a.37120.
- [80] Barber, J.C.K., Rosenfeld, J.A., Foulds, N., Laird, S., Bateman, M.S., Thomas, N.S., et al., 2013. 8p23.1 duplication syndrome; common, confirmed, and novel features in six further patients. *American Journal of Medical Genetics Part A* 161(3): 487–500, Doi: 10.1002/ajmg.a.35767.
- [81] Barber, J.C., Bunyan, D., Curtis, M., Robinson, D., Morlot, S., Dermitzel, A., et al., 2010. 8p23.1 duplication syndrome differentiated from copy number variation of the defensin cluster at prenatal diagnosis in four new families. *Molecular Cytogenetics* 3(1): 3, Doi: 10.1186/1755-8166-3-3.
- [82] Barber, J.C.K., Maloney, V.K., Huang, S., Bunyan, D.J., Cresswell, L., Kinning, E., et al., 2008. 8p23.1 duplication syndrome; a novel genomic condition with unexpected complexity revealed by array CGH. *European Journal of Human Genetics* 16(1): 18–27, Doi: 10.1038/sj.ejhg.5201932.
- [83] Barber, J.C.K., Maloney, V., Hollox, E.J., Stuke-Sontheimer, A., du Bois, G., Daumiller, E., et al., 2005. Duplications and copy number variants of 8p23.1 are cytogenetically indistinguishable but distinct at the molecular level. *European Journal of Human Genetics* 13(10): 1131–6, Doi: 10.1038/sj.ejhg.5201475.
- [84] Fellermann, K., Stange, D.E., Schaeffeler, E., Schmalzl, H., Wehkamp, J., Bevins, C.L., et al., 2006. A Chromosome 8 Gene-Cluster Polymorphism with Low Human Beta-Defensin 2 Gene Copy Number Predisposes to Crohn Disease of the Colon. *The American Journal of Human Genetics* 79(3): 439–48, Doi: 10.1086/505915.
- [85] Giglio, S., Broman, K.W., Matsumoto, N., Calvari, V., Gimelli, G., Neumann, T., et al., 2001. Olfactory Receptor–Gene Clusters, Genomic-Inversion Polymorphisms, and Common Chromosome Rearrangements. *The American Journal of Human Genetics* 68(4): 874–83, Doi: 10.1086/319506.
- [86] Hollox, E.J., Huffmeier, U., Zeeuwen, P.L.J.M., Palla, R., Lascorz, J., Rodijk-Olthuis, D., et al., 2008. Psoriasis is associated with increased β -defensin genomic copy number. *Nature Genetics* 40(1): 23–5, Doi: 10.1038/ng.2007.48.
- [87] Longoni, M., Lage, K., Russell, M.K., Loscertales, M., Abdul-Rahman, O.A., Baynam, G., et al., 2012. Congenital diaphragmatic hernia interval on chromosome 8p23.1 characterized by genetics and protein interaction networks. *American Journal of Medical Genetics Part A* 158A(12): 3148–58, Doi: 10.1002/ajmg.a.35665.
- [88] Páez, M.T., Yamamoto, T., Hayashi, K., Yasuda, T., Harada, N., Matsumoto, N., et al., 2008. Two patients with atypical interstitial deletions of 8p23.1: Mapping of phenotypic traits. *American Journal of Medical Genetics Part A* 146A(9): 1158–65, Doi: 10.1002/ajmg.a.32205.
- [89] Stuart, P.E., Huffmeier, U., Nair, R.P., Palla, R., Tejasvi, T., Schalkwijk, J., et al., 2012. Association of β -Defensin Copy Number and Psoriasis in Three Cohorts of European Origin. *Journal of Investigative Dermatology* 132(10): 2407–13, Doi: 10.1038/jid.2012.191.
- [90] Wat, M.J., Shchelochkov, O.A., Holder, A.M., Breman, A.M., Dagli, A., Bacino, C., et al., 2009. Chromosome 8p23.1 deletions as a cause of complex congenital heart defects and diaphragmatic hernia. *American Journal of Medical Genetics Part A* 149A(8): 1661–77, Doi: 10.1002/ajmg.a.32896.

- [91] Yu, S., Zhou, X.-G., Fiedler, S.D., Brawner, S.J., Joyce, J.M., Liu, H.-Y., 2011. Cardiac Defects Are Infrequent Findings in Individuals With 8p23.1 Genomic Duplications Containing GATA4. *Circulation: Cardiovascular Genetics* 4(6): 620–5, Doi: 10.1161/CIRCGENETICS.111.960302.
- [92] Yu, S., Fiedler, S., Stegner, A., Graf, W.D., 2010. Genomic profile of copy number variants on the short arm of human chromosome 8. *European Journal of Human Genetics* 18(10): 1114–20, Doi: 10.1038/ejhg.2010.66.
- [93] Palomares, M., Delicado, A., Mansilla, E., de Torres, M.L., Vallespín, E., Fernandez, L., et al., 2011. Characterization of a 8q21.11 Microdeletion Syndrome Associated with Intellectual Disability and a Recognizable Phenotype. *The American Journal of Human Genetics* 89(2): 295–301, Doi: 10.1016/j.ajhg.2011.06.012.
- [94] Cormier-Daire, V., Molinari, F., Rio, M., Raoul, O., de Blois, M.-C., Romana, S., et al., 2003. Cryptic terminal deletion of chromosome 9q34: a novel cause of syndromic obesity in childhood? *Journal of Medical Genetics* 40(4): 300–3.
- [95] Stewart, D.R., Huang, A., Faravelli, F., Anderlid, B.-M., Medne, L., Ciprero, K., et al., 2004. Subtelomeric deletions of chromosome 9q: A novel microdeletion syndrome: Subtelomeric 9q Deletion Syndrome. *American Journal of Medical Genetics Part A* 128A(4): 340–51, Doi: 10.1002/ajmg.a.30136.
- [96] Neas, K.R., Smith, J.M., Chia, N., Huseyin, S., Heaps, L.S., Peters, G., et al., 2005. Three patients with terminal deletions within the subtelomeric region of chromosome 9q. *American Journal of Medical Genetics Part A* 132A(4): 425–30, Doi: 10.1002/ajmg.a.30496.
- [97] Yatsenko, S.A., 2005. Deletion 9q34.3 syndrome: genotype-phenotype correlations and an extended deletion in a patient with features of Opitz C trigonocephaly. *Journal of Medical Genetics* 42(4): 328–35, Doi: 10.1136/jmg.2004.028258.
- [98] Kleefstra, T., Koolen, D.A., Nillesen, W.M., de Leeuw, N., Hamel, B.C.J., Veltman, J.A., et al., 2006. Interstitial 2.2 Mb deletion at 9q34 in a patient with mental retardation but without classical features of the 9q subtelomeric deletion syndrome. *American Journal of Medical Genetics Part A* 140A(6): 618–23, Doi: 10.1002/ajmg.a.31123.
- [99] Harada, N., Visser, R., Dawson, A., Fukamachi, M., Iwakoshi, M., Okamoto, N., et al., 2004. A 1-Mb critical region in six patients with 9q34.3 terminal deletion syndrome. *Journal of Human Genetics* 49(8): 440–4, Doi: 10.1007/s10038-004-0166-z.
- [100] Balciuniene, J., Feng, N., Iyadurai, K., Hirsch, B., Charnas, L., Bill, B.R., et al., 2007. Recurrent 10q22-q23 Deletions: A Genomic Disorder on 10q Associated with Cognitive and Behavioral Abnormalities. *The American Journal of Human Genetics* 80(5): 938–47, Doi: 10.1086/513607.
- [101] Crolla, J.A., Cawdery, J.E., Oley, C.A., Young, I.D., Gray, J., Fantes, J., et al., 1997. A FISH approach to defining the extent and possible clinical significance of deletions at the WAGR locus. *Journal of Medical Genetics* 34(3): 207–12.
- [102] Crolla, J.A., van Heyningen, V., 2002. Frequent Chromosome Aberrations Revealed by Molecular Cytogenetic Studies in Patients with Aniridia. *The American Journal of Human Genetics* 71(5): 1138–49, Doi: 10.1086/344396.
- [103] Scott, D.A., Cooper, M.L., Stankiewicz, P., Patel, A., Potocki, L., Cheung, S.W., 2005. Congenital diaphragmatic hernia in WAGR syndrome. *American Journal of Medical Genetics Part A* 134A(4): 430–3, Doi: 10.1002/ajmg.a.30654.
- [104] Gray, J., Yeo, G.S.H., Cox, J.J., Morton, J., Adlam, A.-L.R., Keogh, J.M., et al., 2006. Hyperphagia, severe obesity, impaired cognitive function, and hyperactivity associated with functional loss of one copy of the brain-derived neurotrophic factor (BDNF) gene. *Diabetes* 55(12): 3366–71, Doi: 10.2337/db06-0550.
- [105] Wakui, K., Gregato, G., Ballif, B.C., Glotzbach, C.D., Bailey, K.A., Kuo, P.-L., et al., 2005. Construction of a natural panel of 11p11.2 deletions and further delineation of the critical region involved in Potocki–Shaffer syndrome. *European Journal of Human Genetics* 13(5): 528–40, Doi: 10.1038/sj.ejhg.5201366.
- [106] Thevenon, J., Callier, P., Andrieux, J., Delobel, B., David, A., Sukno, S., et al., 2013. 12p13.33 microdeletion including ELKS/ERC1, a new locus associated with childhood apraxia of speech. *European Journal of Human Genetics* 21(1): 82–8, Doi: 10.1038/ejhg.2012.116.
- [107] Menten, B., Buysse, K., Zahir, F., Hellems, J., Hamilton, S.J., Costa, T., et al., 2007. Osteopoikilosis, short stature and mental retardation as key features of a new microdeletion syndrome on 12q14. *Journal of Medical Genetics* 44(4): 264–8, Doi: 10.1136/jmg.2006.047860.
- [108] Moncla, A., Malzac, P., Voelckel, M.-A., Auquier, P., Girardot, L., Mattei, M.-G., et al., 1999. Phenotype–genotype correlation in 20 deletion and 20 non-deletion Angelman syndrome patients. *European Journal of Human Genetics* 7(2): 131–9, Doi: 10.1038/sj.ejhg.5200258.
- [109] Moncla, A., Malzac, P., Livet, M.O., Voelckel, M.A., Mancini, J., Delaroziere, J.C., et al., 1999. Angelman syndrome resulting from UBE3A mutations in 14 patients from eight families: clinical manifestations and genetic counselling. *Journal of Medical Genetics* 36(7): 554–60.
- [110] Glenn, C.C., Deng, G., Michaelis, R.C., Tarleton, J., Phelan, M.C., Surh, L., et al., 2000. DNA methylation analysis with respect to prenatal diagnosis of the Angelman and Prader-Willi syndromes and imprinting. *Prenatal Diagnosis* 20(4): 300–6.
- [111] Kokkonen, H., Leisti, J., 2000. An unexpected recurrence of Angelman syndrome suggestive of maternal germ-line mosaicism of del(15)(q11q13) in a Finnish family. *Human Genetics* 107(1): 83–5.
- [112] Lossie, A.C., Whitney, M.M., Amidon, D., Dong, H.J., Chen, P., Theriaque, D., et al., 2001. Distinct phenotypes distinguish the molecular classes of Angelman syndrome. *Journal of Medical Genetics* 38(12): 834–45.
- [113] Clayton-Smith, J., Laan, L., 2003. Angelman syndrome: a review of the clinical and genetic aspects. *Journal of Medical Genetics* 40(2): 87–95.
- [114] Sahoo, T., 2006. Microarray based comparative genomic hybridization testing in deletion bearing patients with Angelman syndrome: genotype-phenotype correlations. *Journal of Medical Genetics* 43(6): 512–6, Doi: 10.1136/jmg.2005.036913.

- [115] Schulze, A., Mogensen, H., Hamborg-Petersen, B., Graem, N., Ostergaard, J.R., Brøndum-Nielsen, K., 2001. Fertility in Prader-Willi syndrome: a case report with Angelman syndrome in the offspring. *Acta Paediatrica* (Oslo, Norway: 1992) 90(4): 455–9.
- [116] Holland, A.J., Whittington, J.E., Butler, J., Webb, T., Boer, H., Clarke, D., 2003. Behavioural phenotypes associated with specific genetic disorders: evidence from a population-based study of people with Prader-Willi syndrome. *Psychological Medicine* 33(1): 141–53.
- [117] Chai, J.-H., Locke, D.P., Greally, J.M., Knoll, J.H.M., Ohta, T., Dunai, J., et al., 2003. Identification of Four Highly Conserved Genes between Breakpoint Hotspots BP1 and BP2 of the Prader-Willi/Angelman Syndromes Deletion Region That Have Undergone Evolutionary Transposition Mediated by Flanking Duplicons. *The American Journal of Human Genetics* 73(4): 898–925, Doi: 10.1086/378816.
- [118] Goldstone, A.P., 2004. Prader-Willi syndrome: advances in genetics, pathophysiology and treatment. *Trends in Endocrinology and Metabolism: TEM* 15(1): 12–20.
- [119] Schrandt-Stumpel, C.T.R.M., Curfs, L.M.G., Sastrowijoto, P., Cassidy, S.B., Schrandt, J.J.P., Fryns, J.-P., 2004. Prader-Willi syndrome: Causes of death in an international series of 27 cases. *American Journal of Medical Genetics* 124A(4): 333–8, Doi: 10.1002/ajmg.a.20371.
- [120] Kantor, B., 2004. Control elements within the PWS/AS imprinting box and their function in the imprinting process. *Human Molecular Genetics* 13(7): 751–62, Doi: 10.1093/hmg/ddh085.
- [121] Veltman, M.W.M., Thompson, R.J., Roberts, S.E., Thomas, N.S., Whittington, J., Bolton, P.F., 2004. Prader-Willi syndrome. *European Child & Adolescent Psychiatry* 13(1): 42–50, Doi: 10.1007/s00787-004-0354-6.
- [122] Butler, M.G., Bittel, D.C., Kibiryeva, N., Talebizadeh, Z., Thompson, T., 2004. Behavioral differences among subjects with Prader-Willi syndrome and type I or type II deletion and maternal disomy. *Pediatrics* 113(3 Pt 1): 565–73.
- [123] Sharp, A.J., Selzer, R.R., Veltman, J.A., Gimelli, S., Gimelli, G., Striano, P., et al., 2007. Characterization of a recurrent 15q24 microdeletion syndrome. *Human Molecular Genetics* 16(5): 567–72, Doi: 10.1093/hmg/ddm016.
- [124] Klopocki, E., Graul-Neumann, L.M., Grieben, U., Tönnies, H., Ropers, H.-H., Horn, D., et al., 2008. A further case of the recurrent 15q24 microdeletion syndrome, detected by array CGH. *European Journal of Pediatrics* 167(8): 903–8, Doi: 10.1007/s00431-007-0616-7.
- [125] Mefford, H.C., Rosenfeld, J.A., Shur, N., Slavotinek, A.M., Cox, V.A., Hennekam, R.C., et al., 2012. Further clinical and molecular delineation of the 15q24 microdeletion syndrome. *Journal of Medical Genetics* 49(2): 110–8, Doi: 10.1136/jmedgenet-2011-100499.
- [126] Okubo, Y., Siddle, K., Firth, H., O’Rahilly, S., Wilson, L.C., Willatt, L., et al., 2003. Cell Proliferation Activities on Skin Fibroblasts from a Short Child with Absence of One Copy of the Type 1 Insulin-Like Growth Factor Receptor (IGF1R) Gene and a Tall Child with Three Copies of the IGF1R Gene. *The Journal of Clinical Endocrinology & Metabolism* 88(12): 5981–8, Doi: 10.1210/jc.2002-021080.
- [127] Tatton-Brown, K., Pilz, D.T., Örstavik, K.H., Patton, M., Barber, J.C.K., Collinson, M.N., et al., 2009. 15q overgrowth syndrome: A newly recognized phenotype associated with overgrowth, learning difficulties, characteristic facial appearance, renal anomalies and increased dosage of distal chromosome 15q. *American Journal of Medical Genetics Part A* 149A(2): 147–54, Doi: 10.1002/ajmg.a.32534.
- [128] Lamb, J., Harris, P.C., Wilkie, A.O., Wood, W.G., Dauwerse, J.G., Higgs, D.R., 1993. De novo truncation of chromosome 16p and healing with (TTAGGG)_n in the alpha-thalassemia/mental retardation syndrome (ATR-16). *American Journal of Human Genetics* 52(4): 668–76.
- [129] Lindor, N.M., Valdes, M.G., Wick, M., Thibodeau, S.N., Jalal, S., 1997. De novo 16p deletion: ATR-16 syndrome. *American Journal of Medical Genetics* 72(4): 451–4.
- [130] Daniels, R.J., Peden, J.F., Lloyd, C., Horsley, S.W., Clark, K., Tufarelli, C., et al., 2001. Sequence, structure and pathology of the fully annotated terminal 2 Mb of the short arm of human chromosome 16. *Human Molecular Genetics* 10(4): 339–52.
- [131] Aradhya, S., Lewis, R., Bonaga, T., Nwokekeh, N., Stafford, A., Boggs, B., et al., 2012. Exon-level array CGH in a large clinical cohort demonstrates increased sensitivity of diagnostic testing for Mendelian disorders. *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 14(6): 594–603, Doi: 10.1038/gim.2011.65.
- [132] Rusconi, D., Negri, G., Colapietro, P., Picinelli, C., Milani, D., Spena, S., et al., 2015. Characterization of 14 novel deletions underlying Rubinstein-Taybi syndrome: an update of the CREBBP deletion repertoire. *Human Genetics* 134(6): 613–26, Doi: 10.1007/s00439-015-1542-9.
- [133] Ullmann, R., Turner, G., Kirchoff, M., Chen, W., Tonge, B., Rosenberg, C., et al., 2007. Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Human Mutation* 28(7): 674–82, Doi: 10.1002/humu.20546.
- [134] Hannes, F.D., Sharp, A.J., Mefford, H.C., de Ravel, T., Ruivenkamp, C.A., Breuning, M.H., et al., 2009. Recurrent reciprocal deletions and duplications of 16p13.11: the deletion is a risk factor for MR/MCA while the duplication may be a rare benign variant. *Journal of Medical Genetics* 46(4): 223–32, Doi: 10.1136/jmg.2007.055202.
- [135] de Kovel, C.G.F., Trucks, H., Helbig, I., Mefford, H.C., Baker, C., Leu, C., et al., 2010. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain* 133(1): 23–32, Doi: 10.1093/brain/awp262.
- [136] Mefford, H.C., Cooper, G.M., Zerr, T., Smith, J.D., Baker, C., Shafer, N., et al., 2009. A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. *Genome Research* 19(9): 1579–85, Doi: 10.1101/gr.094987.109.

- [137] Engelen, J.J.M., de Die-Smulders, C.E.M., Dirckx, R., Verhoeven, W.M.A., Tuinier, S., Curfs, L.M.G., et al., 2002. Duplication of chromosome region (16)(p11.2 ? p12.1) in a mother and daughter with mild mental retardation. *American Journal of Medical Genetics* 109(2): 149–53, Doi: 10.1002/ajmg.10287.
- [138] Finelli, P., Natacci, F., Bonati, M.T., Gottardi, G., Engelen, J.J.M., de Die-Smulders, C.E.M., et al., 2004. FISH characterisation of an identical (16)(p11.2p12.2) tandem duplication in two unrelated patients with autistic behaviour. *Journal of Medical Genetics* 41(7): e90.
- [139] Ballif, B.C., Hornor, S.A., Jenkins, E., Madan-Khetarpal, S., Surti, U., Jackson, K.E., et al., 2007. Discovery of a previously unrecognized microdeletion syndrome of 16p11.2–p12.2. *Nature Genetics* 39(9): 1071–3, Doi: 10.1038/ng2107.
- [140] Girirajan, S., Rosenfeld, J.A., Cooper, G.M., Antonacci, F., Siswara, P., Itsara, A., et al., 2010. A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nature Genetics* 42(3): 203–9, Doi: 10.1038/ng.534.
- [141] Tabet, A.-C., Pilorge, M., Delorme, R., Amsellem, F., Pinard, J.-M., Leboyer, M., et al., 2012. Autism multiplex family with 16p11.2p12.2 microduplication syndrome in monozygotic twins and distal 16p11.2 deletion in their brother. *European Journal of Human Genetics* 20(5): 540–6, Doi: 10.1038/ejhg.2011.244.
- [142] Barber, J.C.K., Hall, V., Maloney, V.K., Huang, S., Roberts, A.M., Brady, A.F., et al., 2013. 16p11.2–p12.2 duplication syndrome; a genomic condition differentiated from euchromatic variation of 16p11.2. *European Journal of Human Genetics* 21(2): 182–9, Doi: 10.1038/ejhg.2012.144.
- [143] Okamoto, N., Fujii, T., Tanaka, J., Saito, K., Matsui, T., Harada, N., 2014. A clinical study of patients with pericentromeric deletion and duplication within 16p12.2–p11.2. *American Journal of Medical Genetics Part A* 164(1): 213–9, Doi: 10.1002/ajmg.a.36217.
- [144] Weiss, L.A., Shen, Y., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., et al., 2008. Association between Microdeletion and Microduplication at 16p11.2 and Autism. *New England Journal of Medicine* 358(7): 667–75, Doi: 10.1056/NEJMoa075974.
- [145] McCarthy, S.E., Makarov, V., Kirov, G., Addington, A.M., McClellan, J., Yoon, S., et al., 2009. Microduplications of 16p11.2 are associated with schizophrenia. *Nature Genetics* 41(11): 1223–7, Doi: 10.1038/ng.474.
- [146] Shinawi, M., Liu, P., Kang, S.H.L., Shen, J., Belmont, J.W., Scott, D.A., et al., 2010. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *Journal of Medical Genetics* 47(5): 332–41, Doi: 10.1136/jmg.2009.073015.
- [147] Jacquemont, S., Reymond, A., Zufferey, F., Harewood, L., Walters, R.G., Kutalik, Z., et al., 2011. Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature* 478(7367): 97–102, Doi: 10.1038/nature10406.
- [148] Walters, R.G., Jacquemont, S., Valsesia, A., de Smith, A.J., Martinet, D., Andersson, J., et al., 2010. A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature* 463(7281): 671–5, Doi: 10.1038/nature08727.
- [149] Cardoso, C., Leventer, R.J., Ward, H.L., Toyo-oka, K., Chung, J., Gross, A., et al., 2003. Refinement of a 400-kb Critical Region Allows Genotypic Differentiation between Isolated Lissencephaly, Miller-Dieker Syndrome, and Other Phenotypes Secondary to Deletions of 17p13.3. *The American Journal of Human Genetics* 72(4): 918–30, Doi: 10.1086/374320.
- [150] Kato, M., Dobyns, W.B., 2003. Lissencephaly and the molecular basis of neuronal migration. *Human Molecular Genetics* 12 Spec No 1: R89-96.
- [151] Toyo-oka, K., Shionoya, A., Gambello, M.J., Cardoso, C., Leventer, R., Ward, H.L., et al., 2003. 14-3-3 ϵ is important for neuronal migration by binding to NUDEL: a molecular explanation for Miller–Dieker syndrome. *Nature Genetics* 34(3): 274–85, Doi: 10.1038/ng1169.
- [152] Yingling, J., Toyo-oka, K., Wynshaw-Boris, A., 2003. Miller-Dieker Syndrome: Analysis of a Human Contiguous Gene Syndrome in the Mouse. *The American Journal of Human Genetics* 73(3): 475–88, Doi: 10.1086/378096.
- [153] Boerkoel, C.F., Inoue, K., Reiter, L.T., Warner, L.E., Lupski, J.R., 1999. Molecular mechanisms for CMT1A duplication and HNPP deletion. *Annals of the New York Academy of Sciences* 883: 22–35.
- [154] Boerkoel, C.F., Takashima, H., Garcia, C.A., Olney, R.K., Johnson, J., Berry, K., et al., 2002. Charcot-Marie-Tooth disease and related neuropathies: mutation distribution and genotype-phenotype correlation. *Annals of Neurology* 51(2): 190–201.
- [155] Szigeti, K., Garcia, C.A., Lupski, J.R., 2006. Charcot-Marie-Tooth disease and related hereditary polyneuropathies: molecular diagnostics determine aspects of medical management. *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 8(2): 86–92, Doi: 10.109701.gim.0000200160.29385.73.
- [156] Chance, P.F., Alderson, M.K., Leppig, K.A., Lensch, M.W., Matsunami, N., Smith, B., et al., 1993. DNA deletion associated with hereditary neuropathy with liability to pressure palsies. *Cell* 72(1): 143–51.
- [157] Mouton, P., Tardieu, S., Gouider, R., Birouk, N., Maisonnobe, T., Dubourg, O., et al., 1999. Spectrum of clinical and electrophysiologic features in HNPP patients with the 17p11.2 deletion. *Neurology* 52(7): 1440–6.
- [158] Li, J., Krajewski, K., Shy, M.E., Lewis, R.A., 2002. Hereditary neuropathy with liability to pressure palsy: the electrophysiology fits the name. *Neurology* 58(12): 1769–73.
- [159] Koehler, P.J., 2003. Hereditary neuropathy with liability to pressure palsies: the first publication (1947). *Neurology* 60(7): 1211–3.
- [160] Hong, Y.-H., Kim, M., Kim, H.-J., Sung, J.-J., Kim, S.H., Lee, K.-W., 2003. Clinical and electrophysiologic features of HNPP patients with 17p11.2 deletion. *Acta Neurologica Scandinavica* 108(5): 352–8.
- [161] Chen, K.-S., Manian, P., Koeuth, T., Potocki, L., Zhao, Q., Chinault, A.C., et al., 1997. Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nature Genetics* 17(2): 154–63, Doi: 10.1038/ng1097-154.

- [162] Udwin, O., Webber, C., Horn, I., 2001. Abilities and attainment in Smith-Magenis syndrome. *Developmental Medicine and Child Neurology* 43(12): 823–8.
- [163] Shaw, C.J., Bi, W., Lupski, J.R., 2002. Genetic Proof of Unequal Meiotic Crossovers in Reciprocal Deletion and Duplication of 17p11.2. *The American Journal of Human Genetics* 71(5): 1072–81, Doi: 10.1086/344346.
- [164] Potocki, L., Shaw, C.J., Stankiewicz, P., Lupski, J.R., 2003. Variability in clinical phenotype despite common chromosomal deletion in Smith-Magenis syndrome [del(17)(p11.2p11.2)]. *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 5(6): 430–4, Doi: 10.109701.GIM.0000095625.14160.AB.
- [165] Slager, R.E., Newton, T.L., Vlangos, C.N., Finucane, B., Elsea, S.H., 2003. Mutations in RAI1 associated with Smith-Magenis syndrome. *Nature Genetics* 33(4): 466–8, Doi: 10.1038/ng1126.
- [166] Walz, K., 2003. Behavioral characterization of mouse models for Smith-Magenis syndrome and dup(17)(p11.2p11.2). *Human Molecular Genetics* 13(4): 367–78, Doi: 10.1093/hmg/ddh044.
- [167] Shaw, C.J., Shaw, C.A., Yu, W., Stankiewicz, P., White, L.D., Beaudet, A.L., et al., 2004. Comparative genomic hybridisation using a proximal 17p BAC/PAC array detects rearrangements responsible for four genomic disorders. *Journal of Medical Genetics* 41(2): 113–9.
- [168] Lupski, J.R., 2007. Genomic rearrangements and sporadic disease. *Nature Genetics* 39(7s): S43–7, Doi: 10.1038/ng2084.
- [169] Potocki, L., Chen, K.S., Park, S.S., Osterholm, D.E., Withers, M.A., Kimonis, V., et al., 2000. Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nature Genetics* 24(1): 84–7, Doi: 10.1038/71743.
- [170] Potocki, L., Bi, W., Treadwell-Deering, D., Carvalho, C.M.B., Eifert, A., Friedman, E.M., et al., 2007. Characterization of Potocki-Lupski Syndrome (dup(17)(p11.2p11.2)) and Delineation of a Dosage-Sensitive Critical Interval That Can Convey an Autism Phenotype. *The American Journal of Human Genetics* 80(4): 633–49, Doi: 10.1086/512864.
- [171] Dorschner, M.O., Sybert, V.P., Weaver, M., Pletcher, B.A., Stephens, K., 2000. NF1 microdeletion breakpoints are clustered at flanking repetitive sequences. *Human Molecular Genetics* 9(1): 35–46.
- [172] Kluwe, L., Siebert, R., Gesk, S., Friedrich, R.E., Tinschert, S., Kehrer-Sawatzki, H., et al., 2004. Screening 500 unselected neurofibromatosis 1 patients for deletions of the NF1 gene. *Human Mutation* 23(2): 111–6, Doi: 10.1002/humu.10299.
- [173] Venturin, M., Guarnieri, P., Natacci, F., Stabile, M., Tenconi, R., Clementi, M., et al., 2004. Mental retardation and cardiovascular malformations in NF1 microdeleted patients point to candidate genes in 17q11.2. *Journal of Medical Genetics* 41(1): 35–41.
- [174] Kehrer-Sawatzki, H., Kluwe, L., Sandig, C., Kohn, M., Wimmer, K., Krammer, U., et al., 2004. High Frequency of Mosaicism among Patients with Neurofibromatosis Type 1 (NF1) with Microdeletions Caused by Somatic Recombination of the JAZ1 Gene. *The American Journal of Human Genetics* 75(3): 410–23, Doi: 10.1086/423624.
- [175] Bellanne-Chantelot, C., Clauin, S., Chauveau, D., Collin, P., Daumont, M., Douillard, C., et al., 2005. Large Genomic Rearrangements in the Hepatocyte Nuclear Factor-1 (TCF2) Gene Are the Most Frequent Cause of Maturity-Onset Diabetes of the Young Type 5. *Diabetes* 54(11): 3126–32, Doi: 10.2337/diabetes.54.11.3126.
- [176] Mefford, H.C., Clauin, S., Sharp, A.J., Moller, R.S., Ullmann, R., Kapur, R., et al., 2007. Recurrent Reciprocal Genomic Rearrangements of 17q12 Are Associated with Renal Disease, Diabetes, and Epilepsy. *The American Journal of Human Genetics* 81(5): 1057–69, Doi: 10.1086/522591.
- [177] Shaw-Smith, C., Pittman, A.M., Willatt, L., Martin, H., Rickman, L., Gribble, S., et al., 2006. Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability. *Nature Genetics* 38(9): 1032–7, Doi: 10.1038/ng1858.
- [178] Koolen, D.A., Vissers, L.E.L.M., Pfundt, R., de Leeuw, N., Knight, S.J., Regan, R., et al., 2006. A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nature Genetics* 38(9): 999–1001, Doi: 10.1038/ng1853.
- [179] Lupski, J.R., 2006. Genome structural variation and sporadic disease traits. *Nature Genetics* 38(9): 974–6, Doi: 10.1038/ng0906-974.
- [180] Kirchhoff, M., Bisgaard, A.-M., Duno, M., Hansen, F.J., Schwartz, M., 2007. A 17q21.31 microduplication, reciprocal to the newly described 17q21.31 microdeletion, in a girl with severe psychomotor developmental delay and dysmorphic craniofacial features. *European Journal of Medical Genetics* 50(4): 256–63, Doi: 10.1016/j.ejmg.2007.05.001.
- [181] Koolen, D.A., Sharp, A.J., Hurst, J.A., Firth, H.V., Knight, S.J.L., Goldenberg, A., et al., 2008. Clinical and molecular delineation of the 17q21.31 microdeletion syndrome. *Journal of Medical Genetics* 45(11): 710–20, Doi: 10.1136/jmg.2008.058701.
- [182] Koolen, D.A., Kramer, J.M., Neveling, K., Nillesen, W.M., Moore-Barton, H.L., Elmslie, F.V., et al., 2012. Mutations in the chromatin modifier gene KANSL1 cause the 17q21.31 microdeletion syndrome. *Nature Genetics* 44(6): 639–41, Doi: 10.1038/ng.2262.
- [183] Zollino, M., Orteschi, D., Murdolo, M., Lattante, S., Battaglia, D., Stefanini, C., et al., 2012. Mutations in KANSL1 cause the 17q21.31 microdeletion syndrome phenotype. *Nature Genetics* 44(6): 636–8, Doi: 10.1038/ng.2257.
- [184] Footz, T.K., 2001. Analysis of the Cat Eye Syndrome Critical Region in Humans and the Region of Conserved Synteny in Mice: A Search for Candidate Genes at or near the Human Chromosome 22 Pericentromere. *Genome Research* 11(6): 1053–70, Doi: 10.1101/gr.154901.
- [185] Berends, M.J., Tan-Sindhunata, G., Leegte, B., van Essen, A.J., 2001. Phenotypic variability of Cat-Eye syndrome. *Genetic Counseling (Geneva, Switzerland)* 12(1): 23–34.

- [186] McDermid, H.E., Morrow, B.E., 2002. Genomic Disorders on 22q11. *The American Journal of Human Genetics* 70(5): 1077–88, Doi: 10.1086/340363.
- [187] Gentile, M., De Sanctis, S., Cariola, F., Spezzi, T., Di Carlo, A., Tontoli, F., et al., 2005. FISH approach to determine cat eye syndrome chromosome breakpoints of a patient with cat eye syndrome type II. *European Journal of Medical Genetics* 48(1): 33–9, Doi: 10.1016/j.ejmg.2005.01.020.
- [188] Ryan, A.K., Goodship, J.A., Wilson, D.I., Philip, N., Levy, A., Seidel, H., et al., 1997. Spectrum of clinical features associated with interstitial chromosome 22q11 deletions: a European collaborative study. *Journal of Medical Genetics* 34(10): 798–804.
- [189] Perez, E., Sullivan, K.E., 2002. Chromosome 22q11.2 deletion syndrome (DiGeorge and velocardiofacial syndromes). *Current Opinion in Pediatrics* 14(6): 678–83.
- [190] Greenhalgh, K.L., Aligianis, I.A., Bromilow, G., Cox, H., Hill, C., Stait, Y., et al., 2003. 22q11 deletion: a multisystem disorder requiring multidisciplinary input. *Archives of Disease in Childhood* 88(6): 523–4.
- [191] Rauch, A., 2005. Systematic assessment of atypical deletions reveals genotype-phenotype correlation in 22q11.2. *Journal of Medical Genetics* 42(11): 871–6, Doi: 10.1136/jmg.2004.030619.
- [192] Ensenauer, R.E., Adeyinka, A., Flynn, H.C., Michels, V.V., Lindor, N.M., Dawson, D.B., et al., 2003. Microduplication 22q11.2, an Emerging Syndrome: Clinical, Cytogenetic, and Molecular Analysis of Thirteen Patients. *The American Journal of Human Genetics* 73(5): 1027–40, Doi: 10.1086/378818.
- [193] Yobb, T.M., Somerville, M.J., Willatt, L., Firth, H.V., Harrison, K., MacKenzie, J., et al., 2005. Microduplication and Triplication of 22q11.2: A Highly Variable Syndrome. *The American Journal of Human Genetics* 76(5): 865–76, Doi: 10.1086/429841.
- [194] Portnoi, M.-F., Lebas, F., Gruchy, N., Ardalan, A., Biran-Mucignat, V., Malan, V., et al., 2005. 22q11.2 duplication syndrome: Two new familial cases with some overlapping features with DiGeorge/velocardiofacial syndromes. *American Journal of Medical Genetics Part A* 137A(1): 47–51, Doi: 10.1002/ajmg.a.30847.
- [195] Alberti, A., Romano, C., Falco, M., Cali, F., Schinocca, P., Galesi, O., et al., 2007. 1.5 Mb de novo 22q11.21 microduplication in a patient with cognitive deficits and dysmorphic facial features. *Clinical Genetics* 71(2): 177–82, Doi: 10.1111/j.1399-0004.2007.00750.x.
- [196] Ou, Z., Berg, J.S., Yonath, H., Enciso, V.B., Miller, D.T., Picker, J., et al., 2008. Microduplications of 22q11.2 are frequently inherited and are associated with variable phenotypes. *Genetics in Medicine* 10(4): 267–77, Doi: 10.1097/GIM.0b013e31816b64c2.
- [197] Wentzel, C., Fernström, M., Öhrner, Y., Annerén, G., Thuresson, A.-C., 2008. Clinical variability of the 22q11.2 duplication syndrome. *European Journal of Medical Genetics* 51(6): 501–10, Doi: 10.1016/j.ejmg.2008.07.005.
- [198] Ben-Shachar, S., Ou, Z., Shaw, C.A., Belmont, J.W., Patel, M.S., Hummel, M., et al., 2008. 22q11.2 Distal Deletion: A Recurrent Genomic Disorder Distinct from DiGeorge Syndrome and Velocardiofacial Syndrome. *The American Journal of Human Genetics* 82(1): 214–21, Doi: 10.1016/j.ajhg.2007.09.014.
- [199] Manning, M.A., Cassidy, S.B., Clericuzio, C., Cherry, A.M., Schwartz, S., Hudgins, L., et al., 2004. Terminal 22q deletion syndrome: a newly recognized cause of speech and language disability in the autism spectrum. *Pediatrics* 114(2): 451–7.
- [200] Koolen, D.A., Reardon, W., Rosser, E.M., Lacombe, D., Hurst, J.A., Law, C.J., et al., 2005. Molecular characterisation of patients with subtelomeric 22q abnormalities using chromosome specific array-based comparative genomic hybridisation. *European Journal of Human Genetics* 13(9): 1019–24, Doi: 10.1038/sj.ejhg.5201456.
- [201] Bonaglia, M.C., Giorda, R., Mani, E., Aceti, G., Anderlid, B.-M., Baroncini, A., et al., 2006. Identification of a recurrent breakpoint within the SHANK3 gene in the 22q13.3 deletion syndrome. *Journal of Medical Genetics* 43(10): 822–8, Doi: 10.1136/jmg.2005.038604.
- [202] Phelan, A.L., Kunselman, A.R., Chuang, C.H., Raja-Khan, N.T., Legro, R.S., 2016. Exclusion of Women of Childbearing Potential in Clinical Trials of Type 2 Diabetes Medications: A Review of Protocol-Based Barriers to Enrollment. *Diabetes Care*: dc152723, Doi: 10.2337/dc15-2723.
- [203] Gómez-Zaera, M., Strom, T.M., Rodríguez, B., Estivill, X., Meitinger, T., Nunes, V., 2001. Presence of a major WFS1 mutation in Spanish Wolfram syndrome pedigrees. *Molecular Genetics and Metabolism* 72(1): 72–81, Doi: 10.1006/mgme.2000.3107.
- [204] Kapoor, R.R., Flanagan, S.E., Arya, V.B., Shield, J.P., Ellard, S., Hussain, K., 2013. Clinical and molecular characterisation of 300 patients with congenital hyperinsulinism. *European Journal of Endocrinology* 168(4): 557–64, Doi: 10.1530/EJE-12-0673.
- [205] Majithia, A.R., Flannick, J., Shahinian, P., Guo, M., Bray, M.-A., Fontanillas, P., et al., 2014. Rare variants in PPARG with decreased activity in adipocyte differentiation are associated with increased risk of type 2 diabetes. *Proceedings of the National Academy of Sciences of the United States of America* 111(36): 13127–32, Doi: 10.1073/pnas.1410428111.