

Université de Montréal

**Contribution différentielle des variations du nombre de copies aux
troubles du spectre autistique et aux traits cognitifs**

Par Abderrahim Zeribi

**Département des sciences biomédicales
Faculté de médecine**

Mémoire présenté à la Faculté de médecine
En vue de l'obtention du grade de Maîtrise (Msc.) en Sciences biomédicales
Option médecine expérimentale

Janvier, 2018

© Abderrahim Zeribi, 2018

Université de Montréal

Faculté de médecine

Ce mémoire intitulé :

Contribution différentielle des variations du nombre de copies aux troubles du spectre
autistique et aux traits cognitifs

Présenté par :

Abderrahim Zeribi

A été évalué par un jury composé des personnes suivantes :

Dr Luis Barreiro, président rapporteur

Dr Sébastien Jacquemont, directeur de recherche

Dr Mayada Elsabbagh, membre du jury

Résumé

Les variations du nombre de copies (CNVs) représentent des altérations génétiques hautement étudiées dans les troubles du spectre de l'autisme (TSA). Dans notre étude nous utilisons une cohorte de TSA composée de 2591 individus autistes et 7282 apparentés au premier degré (non autistes) et 2 cohortes de population générale (imagen, sagueney) avec un N total = 4072. À l'aide de ces cohortes nous investiguons l'effet différentiel des CNVs rares sur le risque d'autisme et sur l'atteinte du QI non verbal (NVIQ).

Le niveau d'intolérance à la mutation des gènes contenus dans les CNVs nous permet de modéliser l'effet de ces derniers sur le risque d'autisme ainsi que sur l'atteinte du NVIQ. Nous retrouvons grâce à ce modèle, que les CNVs sont associés à une atteinte du NVIQ qui est deux fois plus importante pour les délétions (suppression d'une copie du génome) que pour les duplications (dédoublément d'une copie du génome).

Même si les CNVs larges semblent associés à un plus haut risque de TSA, notre étude montre que les délétions sont associées plus spécifiquement à la déficience intellectuelle qu'à un plus haut risque d'autisme. A l'opposé, les duplications sont plus spécifiques des TSA. Ces dernières augmentent le risque d'autisme indépendamment du NVIQ. Nos observations sont renforcées par le fait qu'on note un excès de duplications et non pas de délétions chez les apparentés au premier degré comparé à la population générale ce qui suggère une agrégation de ces variants dans ces familles.

Mots clé : Autisme, NVIQ, cognition, CNV, génétique.

Abstract

Copy Number Variations (CNVs) represent highly studied genetic mutations in Autism Spectrum Disorders (ASD). We use in our study, 1 ASD cohort composed from 2591 individuals with ASD (N=2591) and their first degree related (N=7282) and 2 general population cohorts (imagen, sagueneay) with a total N = 4072 individuals.

The level of intolerance to mutations of genes contained in CNVs allows us to model the effect of those CNVs on the risk of autism along with NVIQ impairment. We find, using this model, that CNVs are associated with a decrease of NVIQ which is 2 times more important for deletions than for duplications.

Even though large CNVs seems associated with a higher risk of autism, our study shows that deletions are associated more specifically with intellectual disability than specifically to ASD. Duplications in the other hand increase risk for ASD independently from NVIQ. Our observations are strengthened by the fact that we observe an excess of duplications and not deletions, in first degree related to ASD individuals compared to controls which suggest an aggregation of those CNVs in ASD families.

Keywords : Autism, NVIQ, cognition, CNV, genetic.

TABLE DES MATIERES

Résumé.....	i
Abstract.....	ii
Table des matières.....	iii
Liste des tableaux.....	vi
Liste des figures.....	vii
Liste des abréviations	viii
Dédicace	x
Remerciements	xi
Chapitre 1. Introduction	1
1. Caractéristiques cliniques	2
1.1 Historique.....	2
1.2 Critères diagnostiques.....	3
1.3 Comorbidités.....	8
a. Comorbidités psychiatriques.....	8
b. Comorbidités médicales.....	12
2. Epidémiologie.....	15
3. Facteurs de risque	18
3.1 Facteurs génétiques	18
a. Héritabilité de la TSA.....	18
b. Hétérogénéité génétique.....	18
3.2 Modèles génétiques.....	30
a. Modèle du gène majeur.....	30
b. Modèle polygénique.....	30
3.3 Facteurs environnementaux.....	32
Objectifs et hypothèses.....	35
Chapitre 2. Méthodologie.....	38
1. Population	39
2. Variables cliniques.....	40
3. Détection des CNVs.....	42

4. Définition des CNVs rares	44
5. Annotation des CNVs	44
6. Analyses statistiques	45
6.1 Association entre la fréquence des CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme	46
a. Comparaison entre les probands et la fratrie saine voire la population générale	46
b. Comparaison entre les apparentés aux premiers degrés (parents, fratrie saine) et la population générale	46
6.2 Prédiction du risque d'autisme et de la variation du NVIQ, grâce aux caractéristiques géniques intrinsèques aux CNVs	47
a. Prédiction du risque d'autisme	47
b. Prédiction de la variation du NVIQ	48
6.3 Association entre les CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme (proband- population générale) en ajustant pour le NVIQ	49
Chapitre 3. Résultats	52
1. Association entre la fréquence des CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme	53
1.1 Comparaison entre les probands et la fratrie saine voire la population générale	53
1.2 Comparaison entre les apparentés aux premiers degrés (parents, fratrie saine) et la population générale	53
2. Prédiction du risque d'autisme et de la variation du NVIQ, grâce aux caractéristiques géniques intrinsèques aux CNVs	56
2.1 Prédiction du risque d'autisme	56
2.2 Prédiction de la variation du NVIQ	59
3. Association entre les CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme (proband- population générale) en ajustant pour le NVIQ	61
Chapitre 4. Discussion	65
1. Association entre la fréquence des CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme	66

2. Prédiction du risque d'autisme et de la variation du NVIQ, grâce aux caractéristiques géniques intrinsèques aux CNVs	66
3. Association entre les CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme (proband-population générale) en ajustant pour le NVIQ	67
4. Limite de l'étude.....	68
Conclusion et perspective	70
Bibliographie	71

Liste des tableaux

Tableau 1 : Descriptif des cohortes utilisées dans nos analyses	39
Tableau 2 : Impact du nombre CNVs rares $\geq 250\text{Kb}$ sur la probabilité d’avoir le diagnostic de TSA dans une comparaison avec la fratrie ou la population générale	52
Tableau 3 : Impact du nombre de CNVs rares $\geq 250\text{Kb}$ sur la probabilité d’être apparenté au premier degré à une personne autiste comparé à la population générale	53
Tableau 4 : Association entre le nombre de gènes, les scores d’intolérance (pLI, Del/Dup, RVIS) des CNVs rares $\geq 250\text{Kb}$ et la probabilité d’avoir le diagnostic de TSA dans une comparaison avec la fratrie ou la population générale	56
Tableau 5 : Association entre le score pLI des CNVs (duplication et délétion, séparément) et le NVIQ chez les probands de la SSC.....	59
Tableau 6 : Impact du nombre de CNVs rares $\geq 250\text{Kb}$ sur la probabilité d’avoir le diagnostic de TSA dans une comparaison avec la population générale, après ajustement (modèle 1) puis appariement (modèle 2) pour le NVIQ	61
Tableau 7 : Association entre le nombre de gènes, les scores d’intolérance (pLI, Del/Dup, RVIS) des CNVs rares $\geq 250\text{Kb}$ et la probabilité d’avoir le diagnostic de TSA dans une comparaison avec la population générale, après ajustement (modèle 1) puis appariement (modèle 2) pour le NVIQ	62

Liste des figures

Figure 1 : Evolution de la nosologie de l'autisme à travers les éditions du DSM	7
Figure 2 : Organisation des catégories diagnostics du DSM IV au DSM V	14
Figure 3 : Comorbidités associées aux TSA dans le DSM V	15
Figure 4 : Évolution de la prévalence des TSA seules, TSA+DI et DI seule entre 1996-2010.....	17
Figure 5 : Groupes de gènes impliqués dans les TSA	19
Figure 6 : Différence d'apparition entre les CNVs hérités et <i>de novo</i>	20
Figure 7 : Principe de la puce d'hybridation génomique comparative.....	21
Figure 8 : Principe du génotypage de SNP de type Illumina.....	22
Figure 9 : Représentation de la région 16p11.2	24
Figure 10 : Architecture génétique des troubles du spectre de l'autisme	28
Figure 11 : Évolution de la génétique de l'autisme entre 1975 à 2015.....	30
Figure 12 : Processus d'appariement en fonction du NVIQ, des probands et de la population générale.....	47
Figure 13 : Distribution du QI non verbal (NVIQ) chez les probands et la population générale avant appariement pour le NVIQ	48
Figure 14 : Exemple de distribution du QI non verbal (NVIQ) chez les probands et la population générale après appariement pour le NVIQ.....	48
Figure 15 : Association entre le score pLI des CNV de type délétion et le NVIQ chez les probands de la SSC	58
Figure 16 : Association entre le score pLI des CNV de type duplication et le NVIQ chez les probands de la SSC	58

Liste des abréviations

ADDM : Autism and Developmental Disabilities Monitoring

ADI-R : Autism Diagnostic Interview-Revised

ADOS : Autisme Diagnostic Observation Schedule

ADOS css : Score de sévérité calibré de l'ADOS

AIC : Akaike information criterion

BAF : Fréquence de l'allèle B

BIC : Critère d'information bayésien

CDC's : Centers for Disease Control and Prevention's

CGH array : Puces d'Hybridation Génomique Comparative

CNG : Centre National de Génotypage

CNP : Polymorphismes de nombre de copies

CNV : Variation du nombre de copies DAS-

II : Differential Ability Scales

DGV : Données pour variants génomiques

DI : Déficience intellectuelle

DSM : Diagnostic and Statistical Manual of Mental Disorders

FIQ : Full scale Intelligence Quotient

FISH : Fluorescent In Situ Hybridisation

FMRI Fragility Mental Retardation 1

FMRP : Fragile X mental retardation protein

FXPOI : 'Insuffisance Ovarienne Précoce liée à l'X fragile

FXTAS : Syndrome du tremblement-ataxie lié à l'X fragile

HOE : HumanOmniExpress

LRR : Ratio du log R

MADDSP : Metropolitan Atlanta Developmental Disabilities Surveillance Program

MSEL : Mullen Scales of Early Learning

MTHFR96 : Methylene tetrahydrofolate reductase

NVIQ : Non Verbal Intelligence Quotient

pLI : Probabilité d'être intolérant à la perte de fonction du gène
QI : Quotient intellectuel
RUCDR : Rutgers University Cell and DNA Repository
RVIS : Score d'intolérance aux variations résiduelles
SNP : polymorphisme d'un seul nucléotide
SNV : Variants ponctuelles
SSC : Simon Simplex Collection
SYS : Saguenay Youth Study
TED : Troubles envahissants du développement
TED-NOS : Le trouble envahissant du développement non spécifié
TGD : Troubles globaux du développement
TDA-H : Déficit de l'attention/hyperactivité
TSA : Troubles du spectre autistique
VIQ : Verbal Intelligence Quotient
WF : Facteur d'ondulation
WISC-IV : Wechsler Intelligence Scale for Children
WASI : Wechsler Abbreviated Scale of Intelligence

Dédicace

A mes parents, mon épouse et mon fils

Remerciements

Mes remerciements les plus sincères,

À mon directeur de maîtrise, Dr Sebastien Jacquemont, pour son support et ses conseils durant mes deux années de maîtrise. Sa rigueur et son dévouement pour la recherche en génétique seront pour moi un exemple qui me guidera tout au long de ma carrière professionnelle.

Aux membres de l'équipe du Dr Jacquemont, notamment Dr Huguet, Dr Rodriguez-Herreros et Dr Schramm qui m'ont guidé et soutenu tout au long de cette maitrise par ses conseils avisés. Leurs disponibilités pour enseigner, leurs soutiens, leurs écoutes et leurs conseils judicieux ont été des piliers à la réussite de mon programme d'étude.

Aux autres étudiants gradués du laboratoire, Elise, Mari-pier et tous ceux et celles qui ont partagé mon parcours au sein du laboratoire du Dr Jacquemont.

Aux assistants de recherche pour toute la contribution qu'ils ont pu apporter à mon projet de maîtrise.

Au Dr Luis Barreiro et Dr Elsabbagh Mayada, qui me font l'honneur d'être président et membre du jury d'évaluation de ce travail

À Mes parents, pour m'ont permis de concrétiser ce projet d'étude, Mon épouse Manel, pour sa présence et sa patience; Mon fils Adam, à qui je dédie ce travail.

CHAPITRE 1. INTRODUCTION

1. Caractéristiques cliniques.

1.1 Historique.

Les premières descriptions des troubles neuro-développementaux ont émergé en 1887 avec les travaux du Dr. John Langdon Down qui a décrit pour la première fois le syndrome de Down ou trisomie 21. Dans ses travaux sur le retard mental, il décrit des patients souffrant de ce qu'il appelle « retard développemental ». Dans leurs premières années de vie, ces patients avaient un développement normal, puis vers l'âge de 5 ans sont apparus des mouvements répétitifs rythmiques, une isolation, un manque d'interaction avec autrui, ainsi qu'une régression du langage et du QI. Ces travaux sont les premiers à définir les troubles caractérisant l'autisme tel qu'on le connaît aujourd'hui¹.

Le terme « Autisme » a été utilisé pour la première fois par Eugen Bleuler en 1911 alors qu'il décrivait des patients atteints de schizophrénie. Ces derniers étaient décrits comme vivant dans leur propre monde, coupant ainsi tout contact avec l'extérieur. Cette dissociation de la réalité était considérée comme la forme la plus sévère de schizophrénie et fut appelée « Autisme »². Par la suite le concept d'autisme a beaucoup évolué, notamment avec les travaux effectués par le pédopsychiatre américain d'origine austro-hongroise, Leo Kanner, en 1943³ ainsi que les travaux du psychiatre autrichien Johann Hans Friedrich Karl Asperger en 1944⁴.

Dans son article « *Perturbations autistiques du contact affectif* »³, Dr. Kanner décrit onze enfants, huit garçons et trois filles, âgés de deux à onze ans souffrant d'une perturbation du contact affectif inné, d'isolement et d'immuabilité. Ces symptômes pouvaient apparaître à un âge très précoce (avant 2 ans) à la différence de ceux observés chez les patients schizophrènes pour qui le développement était essentiellement normal avant l'apparition d'une modification progressive du comportement à partir de l'âge adulte³.

En 1944, Dr. Asperger a utilisé le concept de « psychopathie autistique » pour décrire un « dérèglement fondamental [...] entraînant des difficultés d'intégration sociale considérables et très typiques », qui s'accompagne par contre d'« une originalité de la pensée et du vécu qui peut mener, par la suite, à des accomplissements exceptionnels »⁴. Ces travaux ont été repris en 1981 par la psychiatre Lorna Wing qui a défini cliniquement ces symptomatologies comme faisant

partie d'un syndrome qu'elle nomma « syndrome d'Asperger »^{4,5}. Dr. Kanner nommait cette condition « autisme infantile précoce » ou ultérieurement « autisme de Kanner » tandis que Dr. Asperger employait le terme « psychopathie autistique » puis « syndrome d'Asperger »³⁻⁵.

Les troubles décrits par les psychiatres Kanner, Asperger puis Wing étaient similaires mais pas identiques. En effet, tous ces patients souffraient des mêmes difficultés au niveau des interactions sociales mais les patients décrits par Johann Hans Friedrich Karl Asperger avaient de meilleures habiletés langagières accompagnées d'une compréhension au-dessus de la moyenne des tâches techniques. Actuellement, il est acquis que ces troubles font partie d'un même désordre caractérisé par une grande hétérogénéité clinique observée chez les patients.

Afin d'obtenir un consensus autour de la définition de l'autisme, sur lequel pourraient se baser les cliniciens et les chercheurs, l'Association Américaine de Psychiatrie a introduit le terme autisme dans sa troisième édition du DSM (*Diagnostic and Statistical Manual of Mental Disorders*) en 1980⁶. A travers les différentes éditions du DSM (DSM III, DSM IV, DSM V), l'autisme est défini sur la base de la triade clinique de Wing et Gould (Wing et Gould, 1979) incluant une diminution de :

- L'interaction sociale,
- La communication sociale,
- L'imagination sociale.

Dans la définition actuelle de l'autisme, le déficit d'imagination sociale est remplacé par des comportements répétitifs ou stéréotypés et des intérêts restreints, qui seraient une conséquence du déficit d'imagination préalablement décrit.

1.2 Critères diagnostiques.

Dans les premières et deuxièmes éditions du DSM, le terme autisme n'existait pas et les patients autistes étaient majoritairement diagnostiqués comme schizophrènes.

Ainsi, dans le DSM-I (1952)⁷, on nommait « Réactions Psychotiques » les symptômes apparus durant l'enfance chez des patients diagnostiqués ultérieurement comme schizophrènes. Ce n'est

qu'en 1968 que le terme « comportements autistiques » fut employé dans la deuxième édition du DSM⁸, pour décrire une certaine forme de schizophrénie infantile (Figure 1)^{9,10}.

Dans le DSM-III (1980)⁶, l'autisme a été finalement distingué de la schizophrénie et caractérisé par un ensemble de critères diagnostiques :

- Âge d'apparition précoce (avant 30 mois),
- Atteinte de la communication sociale et du langage,
- Intérêt étrange et peur du changement.

On retrouve l'introduction du terme « Troubles globaux du développement » (TGD) qui représentent l'autisme infantile (IA) tel que décrit par Kanner mais aussi des formes moins typiques nommées « Trouble global atypique du développement », qui ne remplissent pas parfaitement les critères diagnostiques de l'autisme infantile (Figure 1)⁹⁻¹¹.

Dans l'édition révisée du DSM-III (1987)¹², les troubles globaux du développement sont remplacés par « Troubles envahissants du développement » (TED) avec évolution du terme « Autisme infantile » vers « Troubles autistiques ». On retrouve aussi la triade clinique décrite préalablement par Wing et Gould. On y retrouve également une description plus large des particularités comportementales tirées d'exemples concrets rencontrés en clinique, facilitant ainsi le diagnostic et la compréhension de ce trouble par les cliniciens (Figure 1)^{9,10}.

Dans le DSM IV (1994)¹³ puis le DSM IV-R (2000)¹⁴, ces troubles autistiques font partie des TED, qui comportent en plus (Figure 1)⁹:

- Le syndrome d'Asperger qui se distingue de l'autisme par l'absence de troubles du langage et un QI normal ;
- Le trouble désintégratif de l'enfance qui se caractérise par l'apparition, à partir de 4 ans, de symptômes autistiques (désintérêt pour l'environnement social, perte du langage...);
- Le syndrome de Rett, une maladie génétique due principalement à une mutation dans le gène *MECP2* (Méthyl-CpG-binding protein 2) touchant principalement les filles et provoquant un handicap mental avec des atteintes motrices sévères ;
- Le trouble envahissant du développement non spécifié (TED-NOS) dont le diagnostic est proposé chez des patients souffrants d'une seule composante de la triade clinique

caractérisant l'autisme. Ils ne réunissent pas toutes les conditions nécessaires à un diagnostic de TED spécifiques. On retrouve chez ces patients des symptômes atypiques soit dans leur chronologie d'apparition soit dans leur présentation.

Le DSM V (2013)⁹ est actuellement considéré comme la référence sur laquelle se base les psychiatres pour effectuer le diagnostic d'autisme. Dans cette édition on note un remaniement des critères diagnostiques qui passent d'un modèle tridimensionnel à un modèle bidimensionnel comprenant (Source : CRETCD (2013): DSM- V nouveautés & changements, terminologie issue d'une traduction libre.):

- Des troubles des interactions sociales et de la communication,
- Des intérêts et comportements restreints et répétitifs.

On note aussi le remplacement des « Troubles Envahissants du Développement » par des « Troubles du spectre autistique » ou TSA. Ce terme proposé initialement par Doris Allen en 1988 regroupe Trouble autistique, Syndrome d'Asperger et les Troubles Envahissants du Développement non spécifiés. Le syndrome de Rett ainsi que le trouble désintégratif de l'enfance sont exclus du diagnostic (Figure 1)⁹.

Ainsi, le diagnostic de TSA est posé dès lors que les points a et b mentionnés ci-après sont établis pour un même individu ⁹.

- a L'individu présente une atteinte des trois éléments constitutifs de la communication et des interactions sociales, à savoir :
 - La réciprocité socio-émotionnelle, qui représente l'initiation et l'entretien de conversations normales ;
 - La communication non verbale, (qui peut aller d'une simple absence du contact visuel à un manque complet d'expression faciale et de communication non verbale) ;
 - Le développement, le maintien et la compréhension des relations sociales (difficulté à réagir face à des situations sociales voire absence complète d'intérêt pour autrui).

- b L'individu présente des intérêts et comportements restreints et répétitifs caractérisés par au moins deux des quatre éléments suivants :

- Des mouvements ou utilisation de langage de manière répétitif et stéréotypé (exemple : écholalie) ;
 - Une rigidité impliquant le maintien de routines et rituels verbaux ou non verbaux dans le quotidien (exemples : les repas, l'endormissement...) ;
 - Des intérêts limités et restreints voire atypiques (attachement à des objets inhabituels contrastant avec un manque d'intérêt envers des sujets voire des objets communs) ;
 - Réactivité excessive ou baissée à des stimuli sensoriels simulant un intérêt particulier pour des aspects de l'environnement (exemples : indifférence à la douleur, fascination pour la lumière, réaction négative à des stimulations sonore particulières).
- c. Les symptômes doivent être présents à partir de la petite enfance mais peuvent des fois se manifester ultérieurement (une fois que la capacité à absorber le stress social soit dépassée).
- d. Ces symptômes altèrent significativement le fonctionnement quotidien en entravant la réussite sociale, scolaire et/ou professionnelle.
- e. Les difficultés ne doivent pas être mieux expliquées par de la déficience intellectuelle tout en sachant que cette dernière peut accompagner un diagnostic de TSA engendrant impérativement une communication sociale plus basse que ce qui est prévue par le niveau de développement global.

La sévérité des symptômes n'est mesurée qu'en fonction des critères a et b et est cotée dépendamment du niveau de soutien nécessaire, de 1 à 3 :

- Niveau 1 : nécessite un soutien ;
- Niveau 2 : nécessite un soutien substantiel ;
- Niveau 3 : nécessite un soutien très substantiel.

De plus, lors de son évaluation, le clinicien doit préciser des spécificateurs. Ces derniers ne représentent pas une condition au diagnostic mais sont présents pour mieux cerner et comprendre l'hétérogénéité clinique caractéristique des TSA. Ils comprennent :

- La présence ou l'absence de déficience intellectuelle avec description précise des capacités cognitives ;
- L'évaluation détaillée des aptitudes langagières avec possibilité d'avoir un langage réceptif inférieur au langage expressif ;
- La présence ou l'absence d'une condition médicale telle que l'épilepsie par exemple ou encore d'une condition génétique tel que le Syndrome de Rett par exemple ;
- La présence de maladie et/ou de troubles fréquemment associés aux TSA tels que la dépression, les troubles anxieux, le déficit de l'attention/hyperactivité (TDA-H).
- Une présentation clinique très hétérogène qui peut inclure : auto mutilation, mutisme, négligence.... (Source : CRETCD (2013): DSM- V nouveautés & changements, terminologie issue d'une traduction libre.)

DSM-I	1952	Inclus dans: Schizophrenic reaction – childhood type
DSM-II	1968	Inclus dans: Schizophrenia – childhood type
DSM-III	1980	Troubles globaux du développement <ul style="list-style-type: none"> - Autisme infantile (syndrome complet ou en stade résiduel) - Trouble global du développement débutant dans l'enfance (syndrome complet ou en stade résiduel) - Trouble global atypique du développement
DSM-III-R	1987	Troubles envahissants du développement <ul style="list-style-type: none"> - Trouble autistique - Trouble envahissant du développement non-spécifié
DSM-IV	1994	Troubles envahissants du développement <ul style="list-style-type: none"> - Trouble autistique - Syndrome d'Asperger - Trouble envahissant du développement non-spécifié - Syndrome de Rett - Trouble désintégratif de l'enfance
DSM-5	2013	TSA

Figure 1 : Evolution de la nosologie de l'autisme à travers les éditions du DSM.

Tous les désordres retrouvés dans le DSM IV ont été regroupés dans le DSM V sous le nom de TSA, excepté le syndrome de Rett et le trouble désintégratif de l'enfance.

Source : CRETCD (2013): DSM- V nouveautés & changements, terminologie issue d'une traduction libre.

1.3 Comorbidités.

Les comorbidités liées aux TSA sont les maladies ou troubles, autres que les TSA, observés plus fréquemment chez les individus TSA que dans la population générale. Elles peuvent être d'ordre psychiatrique ou médical.

La fréquence élevée de ces comorbidités dans les TSA est en partie responsable de l'hétérogénéité clinique caractéristique de ce trouble et peut être le reflet d'une complexité génétique importante (Figure 2 et Figure 3), (Source : CRETCD (2013): DSM- V nouveautés & changements, terminologie issue d'une traduction libre.).

a. Comorbidités psychiatriques.

L'autisme tel que défini actuellement résulte d'un ensemble de troubles cliniquement très hétérogènes, qui forment ensemble un seul spectre connu sous le nom de TSA. Cette hétérogénéité se reflète par une importante variabilité clinique observée entre deux individus diagnostiqués autistes.

Plusieurs troubles psychiatriques peuvent survenir plus fréquemment chez les personnes TSA augmentant ainsi la diversité clinique et la complexité de ces troubles. Le diagnostic de ces comorbidités représente un défi étant donné le manque de réciprocité dans la conversation et la difficulté d'évaluer les émotions et les pensées chez les personnes autistes (Figure 3) (Source : CRETCD (2013): DSM- V nouveautés & changements, terminologie issue d'une traduction libre.)^{9,15}.

➤ La déficience intellectuelle.

Anciennement appelée retard mental, la déficience intellectuelle (DI) appartient à la catégorie des troubles neuro-développementaux au même titre que les TSA (Figure 2)⁹. Il a été longtemps pensé que la majorité des enfants autistes avaient une déficience intellectuelle et la barrière entre ces deux troubles n'a pas cessé de changer, notamment avec l'évolution des critères diagnostiques du DSM. Dans le DSM V, la déficience intellectuelle a été subdivisée en déficience intellectuelle réelle et retard global du développement⁹.

La déficience intellectuelle réelle se caractérise par la présence d'une atteinte significative du fonctionnement intellectuel. Elle est évaluée grâce au test du quotient intellectuel (QI), un test psychométrique permettant d'évaluer différents domaines d'intelligence (le raisonnement, la compréhension de tâches complexes, l'apprentissage...). En plus de l'atteinte du QI, on retrouve une limitation des comportements adaptatifs nécessaires au fonctionnement quotidien tel que la communication, la socialisation, l'indépendance dans les tâches quotidiennes, dans la vie professionnelle et les études. Le QI et le comportement adaptatif sont évalués grâce à des mesures standardisées⁹. Contrairement aux versions antérieures du DSM, qui se basaient uniquement sur un QI <70 pour poser le diagnostic de déficience intellectuelle, le DSM V accorde plus d'importance aux capacités d'adaptation, qui reflèteraient mieux l'étendu du soutien nécessaire à ces personnes et donc la sévérité de leur déficience intellectuelle⁹.

La prévalence de la déficience intellectuelle chez les enfants TSA a considérablement diminué avec l'évolution des critères diagnostiques du TSA pour se stabiliser aux alentours du tiers¹⁶.

Le diagnostic de retard global du développement est établi chez des enfants qui semblent présenter des déficiences au niveau de leur fonctionnement intellectuel global mais étant donné leur jeune âge, l'estimation formelle de leurs capacités intellectuelles n'est pas possible. Ainsi l'enfant sera réévalué ultérieurement pour clarifier la présence ou non d'un diagnostic de déficience intellectuelle avec des outils standardisés⁹.

➤ **Déficit de l'attention - hyperactivité (TDA-H).**

Ces déficits sont fréquents chez les personnes TSA, avec une prévalence allant de 30 à 80%¹⁷. Les TDA-H sont caractérisés par une inattention et/ou hyperactivité-impulsivité. Ces symptômes retentissent sur le fonctionnement quotidien^{9,15}.

Le diagnostic de TSA-H chez des personnes TSA est rendu possible dans le DSM V. Pour cela il faut pour les enfants âgés de moins de 17 ans⁹ :

- Une atteinte d'au moins six symptômes appartenant à la catégorie de l'inattention ou de l'hyperactivité-impulsivité.

Pour les enfants/adultes de plus de 17 ans :

- Une atteinte d'au moins cinq symptômes appartenant à la catégorie de l'inattention et/ou de l'hyperactivité-impulsivité

Les symptômes doivent persister au moins 6 mois et se manifestent obligatoirement avant 12 ans.

➤ **Troubles anxieux.**

Ces troubles sont parmi les plus observés chez les adolescents et les enfants autistes d'âge scolaire avec une fréquence de 41,9 à 43%^{18,19}. Une plus grande anxiété est associée à un QI plus bas et une moindre sévérité des symptômes autistiques²⁰. Ils sont subdivisés selon le DSM V⁹ en trouble d'anxiété de séparation, trouble d'anxiété sociale, trouble de personnalité obsessionnelle-compulsive et état de stress post-traumatique.

Les troubles d'anxiété de séparation sont déclenchés à la suite de la séparation avec des personnes auxquelles l'individu est attaché. La sévérité des manifestations est extrême comparé à ce qui est attendu contenu du niveau de développement de l'enfant⁹. Ils entraînent une détresse clinique importante chez l'individu avec un impact sur le fonctionnement quotidien, professionnel, scolaire et social.

Le trouble d'anxiété sociale (phobie sociale) est le trouble d'anxiété le plus observé chez les autistes avec un prévalence de 29,2%²¹. C'est une phobie déclenchée par des situations sociales particulières qui peut conduire à une réaction d'évitement et ainsi retentir considérablement sur la vie sociale, la performance scolaire et la réussite professionnelle.

Les troubles de personnalité obsessionnelle-compulsive sont caractérisés par la survenue d'images et de pensées récurrentes ainsi que des comportements répétitifs⁹. Ils sont observés plus fréquemment chez les personnes TSA²².

L'état de stress post-traumatique se développe suite à une exposition à un événement traumatique marquant. Il se caractérise par une détérioration des aptitudes sociales et de communications, des troubles du sommeil, de l'agitation et de l'hyperactivité²³.

Sa fréquence est plus élevée chez la population TSA (17%) comparée à la population générale (6-8%)²³.

➤ **Troubles des comportements et des conduites.**

Ces troubles regroupent les troubles oppositionnels avec provocation, les troubles de la conduite ainsi que les troubles explosifs intermittents.

Les troubles oppositionnels avec provocation sont observés plus souvent chez les personnes TSA (28.1%)²¹. Ils se définissent comme un ensemble de comportements négatifs, provocateurs voire hostiles dirigés envers une personne en dehors de la fratrie (humeur colérique, comportement opposant et comportement vindicatif)⁹. Ces comportements doivent être présents depuis au moins 6 mois, altèrent le fonctionnement social, scolaire et professionnel.

Le trouble de la conduite se caractérise par un comportement répétitif et persistant dans lequel la personne atteinte ne respecte pas les normes sociales correspondantes à son âge. La symptomatologie doit être présente pendant au moins 12 mois pour pouvoir poser le diagnostic. La fréquence est établie à 3,2% dans la population TSA²¹

Les troubles explosifs intermittents se caractérisent par des crises récurrentes d'agressivité verbale ou physique que l'individu n'arrive pas à contrôler. Ils se déclenchent subitement sans signes annonciateurs, durent moins de 30 minutes et altèrent considérablement le fonctionnement quotidien⁹.

➤ **Trouble disruptif de dysrégulation émotionnelle.**

Ce trouble représente une sous-catégorie des troubles dépressifs. Il est diagnostiqué chez les enfants âgés entre 6 et 18 ans, qui présentent une irritabilité persistante accompagnée de

crises de colère⁹. Il est fréquemment associé aux TDA-H, troubles de la conduite et les troubles de la consommation⁹.

b. Comorbidités médicales.

Certaines conditions médicales se produisent plus souvent chez les TSA comparé à la population générale. En plus de la difficulté de l'examen clinique rencontrée chez les autistes, la présentation clinique de certaines comorbidités peut différer de ce qui est observé traditionnellement chez la population générale. Cela implique une connaissance minutieuse de toutes les comorbidités médicales associées à l'autisme (Source : CRETCD (2013): DSM- V nouveautés & changements, terminologie issue d'une traduction libre.)⁹. Parmi ces comorbidités on a :

➤ **L'épilepsie.**

Les études de méta-analyse montrent que la prévalence de l'épilepsie chez les patients TSA est de 8% comparé à 2-3% dans la population générale. Cette prévalence atteint 21,4% chez les patients TSA avec retard mental²⁴. L'épilepsie est plus fréquente chez les TSA ayant des habiletés langagières plus basses, souffrants d' atteintes neurologiques telle que la paralysie cérébrale^{25 26} et de sexe féminin²⁴. Les individus atteints à la fois de TSA et d'épilepsie ont souvent une atteinte plus sévère dans tous les domaines de fonctionnement, une motricité fine et globale moins bonne, un déficit social ainsi que des comportements difficiles^{27,28}. Cette sévérité pourrait résulter soit des complications physiologiques engendrées à la suite de crises épileptiques récurrentes soit de l'origine génétique commune à ces deux désordres.

➤ **Troubles immunitaires.**

Plusieurs études se sont intéressées au possible lien entre les maladies immunitaires et les TSA. Zerbo et al ont montré que la présence d'une maladie immunitaire était globalement associée à un plus grand risque de TSA²⁹. Chen et al retrouvent que ce risque est aussi retrouvé chez les enfants qui ont eu des problèmes allergiques³⁰. Puig-Alcatraz et al ont montré qu'une augmentation de la concentration de l'homocystéine, un marqueur pour les maladies

immunitaires était corrélée positivement à la sévérité des troubles de la communication observés chez les enfants TSA³¹.

➤ **Troubles gastro-intestinaux.**

La prévalence des troubles gastro-intestinaux chez les personnes TSA est très variable allant de 9 à 91%³²⁻³⁵. À l'exception de Black et al³³, toutes les études s'accordent à dire que cette prévalence est plus élevée chez les individus TSA que dans la population générale avec des symptômes allant de la constipation, à la sélectivité alimentaire et à la diarrhée^{34,36}. Les patients TSA avec des symptômes gastro-intestinaux auraient un plus grand risque d'anxiété, d'hyperréactivité sensorielle, d'irritabilité, de retrait social et de régression du langage³⁷⁻³⁹.

L'association entre la sensibilité au gluten voire la maladie cœliaque et les TSA a été longuement suspectée. La plupart des études qui ont évalué ce lien n'ont retrouvé aucune association significative⁴⁰⁻⁴³. Ludvigsson et al⁴³ ont utilisé les données de biopsies provenant du registre suédois. Il inclue 26 995 individus ayant un diagnostic histologique de maladie cœliaque (stade 3), 12 304 souffrants d'inflammation à l'histologie (stades 1, 2) et 3 719 sans altérations histologiques mais avec un test sérologique positif (stade 0). Ces individus ont été comparés à 213 208 individus contrôles. Aucune association n'a été retrouvée entre la maladie cœliaque confirmée (stade 3), l'inflammation (stade 1, 2) et les TSA. Cependant, ils retrouvent une augmentation significative du risque de TSA chez les individus au stade 0⁴³. Les auteurs concluent qu'une sensibilité au gluten non cœliaque est associée aux TSA.

➤ **Troubles du sommeil.**

Rzepecka et al estiment que 77,2% des TSA présentent des problèmes de sommeil⁴⁴. Cela fait des troubles du sommeil une des comorbidités les plus fréquentes dans les TSA. Tsai et al retrouvent une fréquence plus grande des troubles du sommeil chez la population TSA comparé à la population générale⁴⁵. Les troubles du sommeil prédissent significativement la sévérité des troubles autistiques et le déficit social⁴⁶.

➤ **Troubles moteurs.**

Ces troubles peuvent toucher l'acquisition de la coordination, les mouvements stéréotypés et les tics. Ils ont été récemment mis en avant car ils sont considérés comme annonciateurs de troubles neuro-développementaux tels que les TSA.⁴⁷

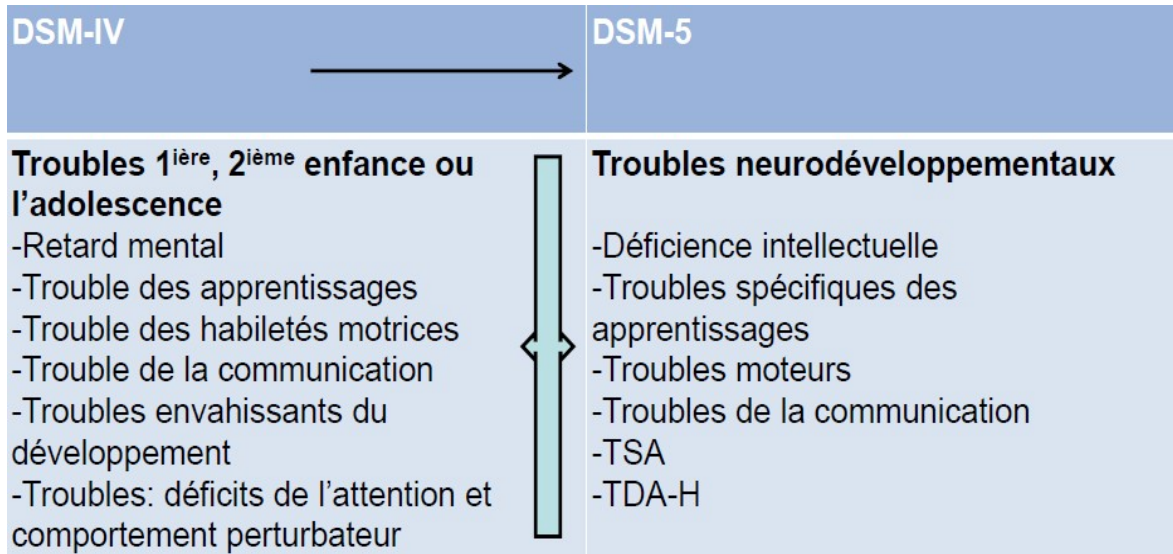


Figure 2 : Organisation des catégories diagnostics du DSM IV au DSM V.

Source : CRETCD (2013): DSM-V nouveautés & changements, terminologie issue d'une traduction libre.

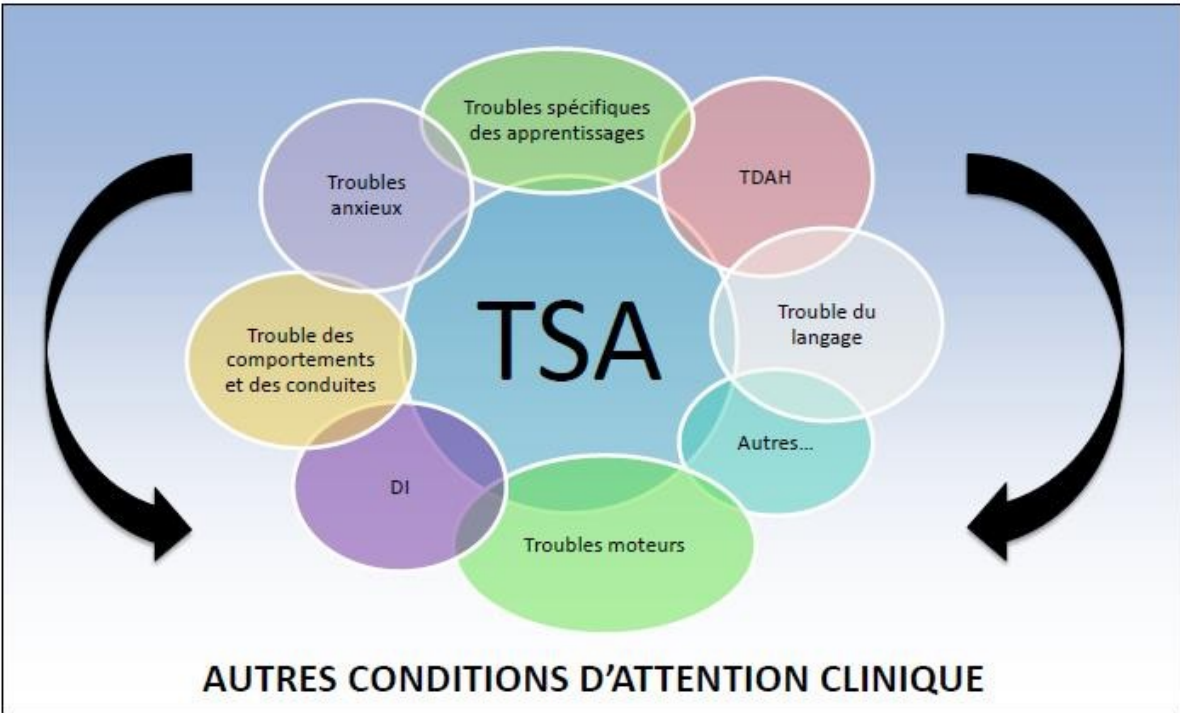


Figure 3 : Comorbidités associées aux TSA dans le DSM V.

Source : CRETCD (2013): DSM-V nouveautés & changements, terminologie issue d'une traduction libre.

2. Epidémiologie.

Les premières études épidémiologiques portant sur la prévalence de l'autisme ont été réalisées dans les années 1966 en Grande Bretagne ⁴⁸ et 1970 aux Etats-Unis ⁴⁹. Elles portaient sur une population d'enfants et estimaient que la prévalence de l'autisme était de 2 à 4 cas par 10 000 enfants ^{48,49} avec un ratio de 3 à 4 garçons pour 1 fille atteinte. Ces premières études laissaient penser que l'autisme était un trouble rare et dont la répercussion à l'échelle de la société était limitée (Figure 4).

A la suite de la de l'évolution de la définition de l'autisme avec les différentes éditions du DSM III-R (1987)¹², DSM IV (1994)¹³ et DSM V (2013)⁹, la prévalence de ce trouble a augmenté exponentiellement. Elle est passée de 2 à 4 cas pour 10 000 enfants à 6,7 pour 1000 enfant en l'an 2000. Cette augmentation concorde avec une diminution de la prévalence de la déficience intellectuelle (anciennement appelée retard mental) qui passe de 13 cas sur 1000 en 1996 à 8,6 cas sur 1000 en 2010⁵⁰. Cette augmentation considérable de la prévalence des TSA s'explique en

majeur partie par l'élargissement des critères diagnostics incluant désormais des formes cliniques moins sévères (Figure 4).

Le réseau Autism and Developmental Disabilities Monitoring (ADDM) du Centers for Disease Control and Prevention's (CDC's) a effectué tous les 2 ans une surveillance de l'autisme chez des enfants âgés de 8 ans, évalués dans différents sites aux Etats-Unis ^{16,50}.

Cette étude rapporte une prévalence estimée croissante de la période allant de 2000 à 2012. Ainsi en l'an 2000, on retrouve une prévalence de 6,7 cas pour 1000 enfants, en 2004 elle atteint 8 pour 1000 enfants, en 2008 elle est de 11,3 pour 1000 enfants et atteint finalement en 2012 une prévalence de 14,6 pour 1000 enfants. Cette augmentation ne peut être attribuée complètement à l'évolution des critères diagnostics de l'autisme car tous les enfants participants à cette étude ont été évalués par la même édition du DSM (DSM IV -R). Parmi les autres facteurs pouvant jouer un rôle dans cette augmentation notable de la prévalence des TSA, on peut citer ^{16,51} :

- L'amélioration de l'accessibilité et de la qualité des évaluations auxquelles les enfants sont soumis en lien avec une plus grande sensibilisation à ce trouble à travers les médias et les professionnels de la santé ;
- L'augmentation de certains facteurs connus voire suspectés pour favoriser l'apparition de l'autisme tels que la hausse de l'âge parental à la conception et l'obésité maternelle ;
- L'introduction des TSA comme diagnostic donnant droit à une éducation spéciale. Cette disposition est rentrée en vigueur en 1991 aux Etats-Unis, ce qui a mené à plus d'enfants évalués et diagnostiqués avec ce trouble au niveau des écoles. Ainsi des enfants qui auraient bénéficié d'une éducation spéciale sous d'autres catégories cliniques dont la déficience intellectuelle seraient de plus en plus servis sous la catégorie des TSA.

Si nous nous concentrons maintenant sur les données du Metropolitan Atlanta Developmental Disabilities Surveillance Program (MADDSP), des chercheurs ont démontré que parmi les individus souffrant de TSA en 1996, 61% avaient une déficience intellectuelle (anciennement appelée retard mental). Ce pourcentage a varié pour atteindre environ 40% en 2010 et cette tendance vers la diminution semble continuer ⁵⁰.

Les données collectées par le réseau ADDM retrouve une augmentation de la prévalence de l'autisme plus importante chez les garçons que chez les filles. Les données collectées en l'an 2000 montrent 3,5 fois plus de garçons atteints que de filles. Cette différence continue de se

creuser et atteint 4,5 en 2010. Ainsi l'augmentation de la prévalence de l'autisme observée entre 2000 et 2010 toucherait plus les garçons que les filles ¹⁶. On note que cette augmentation de la prévalence est plus importante pour les enfants de race blanche non hispaniques que les enfants hispaniques¹⁶ (Figure 4).

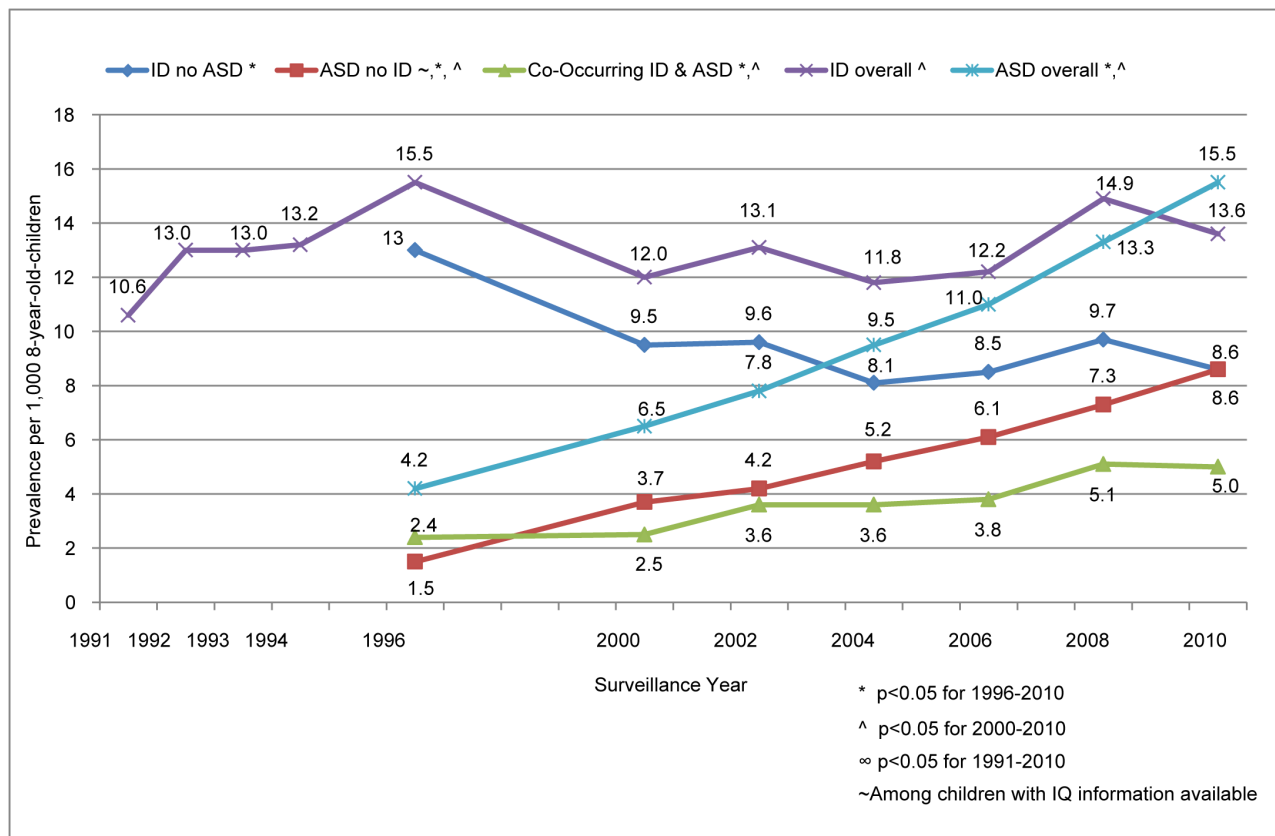


Figure 4 : Évolution de la prévalence de l'autisme (ASD) et de la déficience intellectuelle (ID) entre 1996-2010.

Le diagnostic d'autisme et de déficience intellectuelle est établi selon les critères présents dans les éditions antérieures au DSM 5. Ainsi la ID (antérieurement appelée retard mental) est définie par un $QI < 70$.

La ligne bleu foncé représente la prévalence de cas de déficience intellectuelle n'ayant pas le diagnostic d'ASD.

La ligne bleu clair représente la prévalence de cas d'autisme avec ou sans ID.

La ligne violette représente la prévalence de cas de déficience intellectuelle avec ou sans ASD.

La ligne rouge représente la prévalence de cas d'autisme n'ayant pas le diagnostic de ID.

La ligne verte représente la prévalence de cas d'autisme accompagné de déficience intellectuelle.

Source : Braun et al., 2015⁵⁰.

3. Facteurs de risque.

3.1 Facteurs génétiques.

a. Héritabilité des TSA.

L'Autisme est un trouble neuro-développemental complexe caractérisé par une forte composante génétique. Tick et al. ont démontré sur une méta-analyse regroupant sept études de jumeaux, que l'héritabilité de ce trouble variait de 64% à 91%⁵².

L'estimation de l'héritabilité effectuée grâce au calcul du taux de variations génétiques communs de type SNP (polymorphisme d'un seul nucléotide) partagées avec les individus non apparentés montre une héritabilité de ce trouble de ~59%⁵³.

Un autre argument en faveur de l'origine génétique des TSA est la fréquence élevée de survenue de ce trouble chez les individus atteints de Syndromes génétiques connus tel que le Syndrome de l'X fragile.

Ces données confirment la présence d'une grande contribution des facteurs génétiques dans les TSA mais ces derniers n'expliquent pas la totalité de ce trouble. En effet les TSA résulteraient de la contribution de facteurs génétiques, de facteurs environnementaux voire de l'interaction complexe des deux.

b. Hétérogénéité génétique.

L'architecture génétique de l'autisme est complexe. Elle impliquerait différents variants en terme de fréquence (communs et rares), de type (variants ponctuelles (SNV), variation du nombre de copies (CNV)) et de mode de transmission (récessif, dominant, additif, *de novo*)⁵⁴. On estime que jusqu'à 1000 le nombre de gènes impliqués dans l'autisme⁵⁵. Une centaine de ces gènes sont déjà connus (*SHANK3 NRXNI...*). Ils convergeraient sur différentes voies biologiques telles que la synaptogenèse, la guidance axonale, le développement et la régulation neuronale (gene.sfari.org), (Figure 5, Figure 11)⁵⁵⁻⁵⁷.

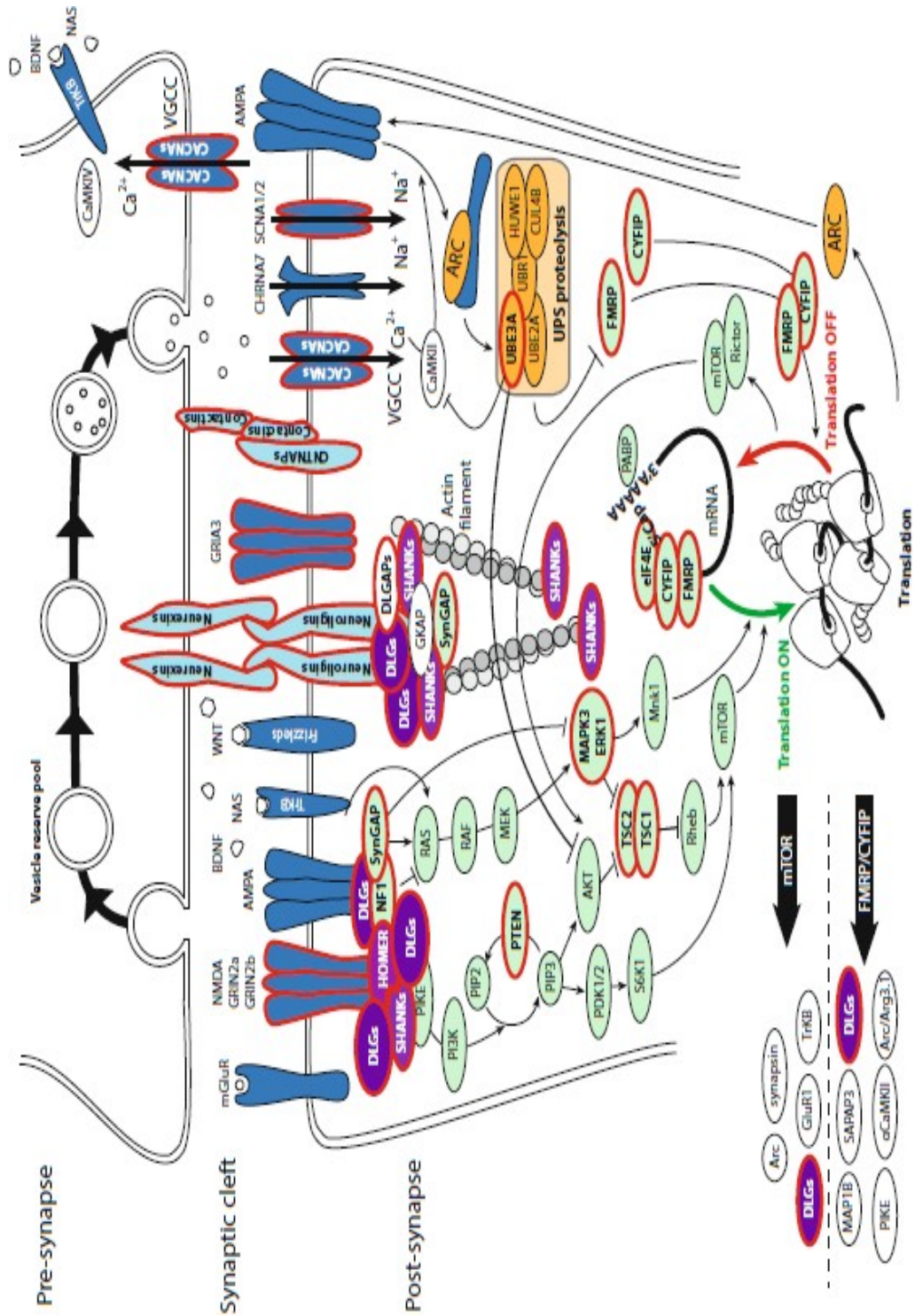
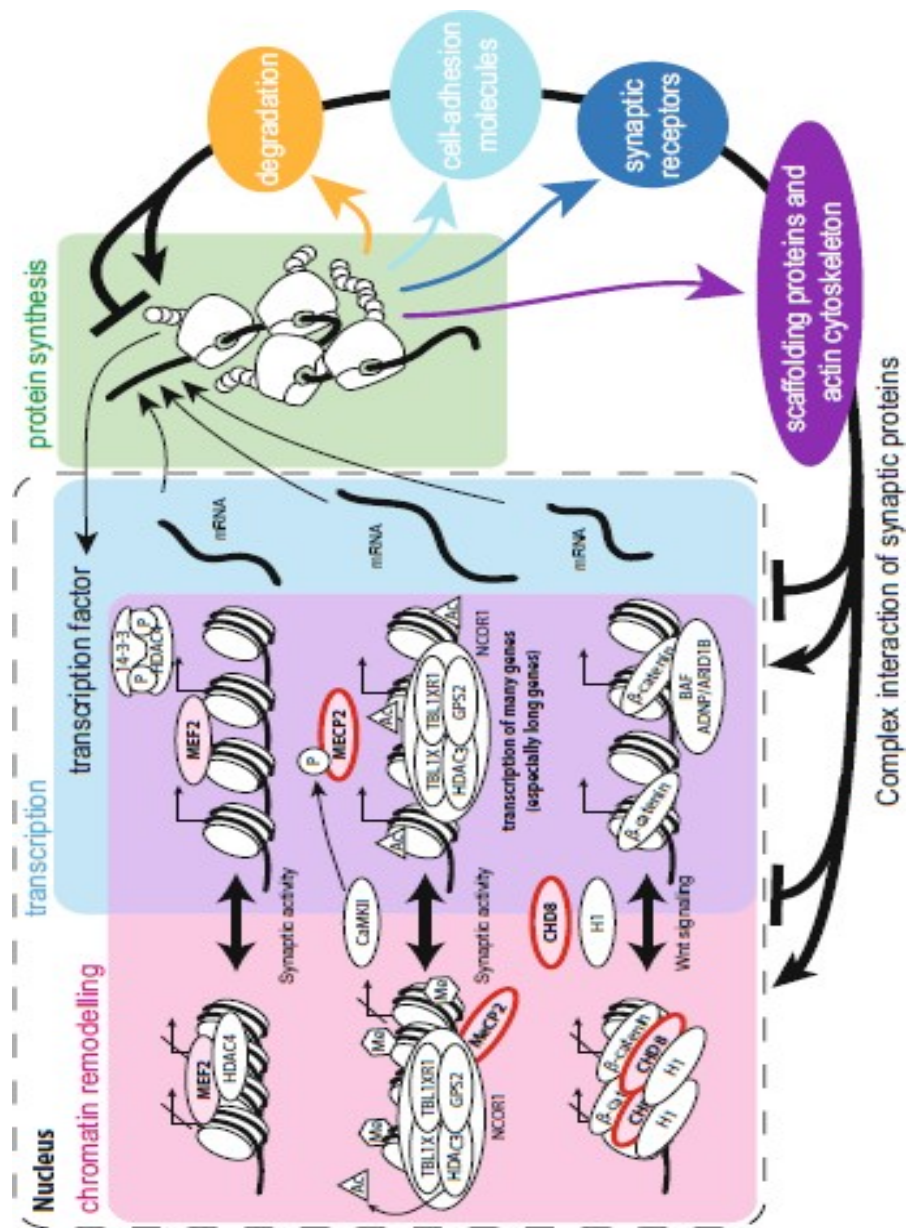


Figure 5 (a)



(b)

Figure 5 (a-b) : Groupes de gènes impliqués dans les TSA.

Les gènes impliqués dans l'autisme codent des protéines jouant un rôle dans le remodelage de la chromatine, la transcription, la synthèse et la dégradation protéique, le dynamisme du cytosquelette et les fonctions synaptiques. Les protéines associées à l'autisme sont encadrées de rouge.

Source : A time for metabolism and hormones⁵⁸.

➤ **Réarrangements chromosomique et CNVs.**

Les CNVs sont défini par une perte (délétion) ou un gain (duplication) d'une partie du génome d'au moins 1000 paires de base et pouvant aller jusqu'à un chromosome entier. La délétion d'un chromosome complet est appelée monosomie, elle est dans la plupart des cas non viable. La duplication d'un chromosome complet donne la trisomie dont la plus connue, trisomie 21 (Syndrome de Down) a été associée à une fréquence plus élevée de TSA (40%)^{59,60}.

Les CNVs apparaissent lors de la méiose grâce à un phénomène d'enjambement inégal, où des nucléotides sont échangés d'une manière disproportionnée entre deux chromosomes.

Il existe des segments dans l'ADN, qui pendant l'évolution ont été dupliqués de nombreuses fois. Ces segments sont présents communément chez les individus et sont sensible à l'apparition de CNVs. Ces derniers, quand ils sont flanqués par ces segments arrivent plus fréquemment. C'est ainsi qu'un même CNV peut être observé chez plusieurs individus «CNV récurrents». D'autre part, les CNV qui apparaissent en dehors de ces régions dupliquées sont produits aléatoirement dans le génome ce qui limiterait grandement la possibilité d'observer le même CNV chez deux individus différents «CNV non récurrents». Les CNVs peuvent être soit hérités (transmis) des parents soit apparus *de novo* (non présents chez les parents) (Figure 6).

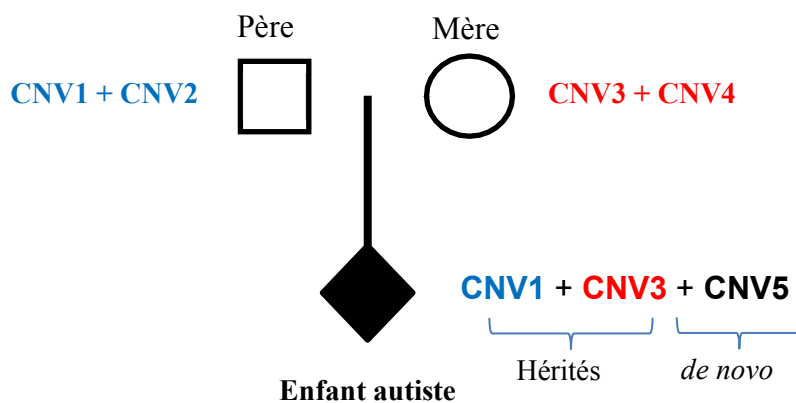


Figure 6 : Différence d'apparition entre les CNVs hérités et *de novo*.

Source : réalisation personnelle.

Les CNVs sont détectés grâce à plusieurs techniques, de cytogénétique, de puce de génotypage et de séquençages à haut débit.

Le caryotype est une technique de cytogénétique permettant de voir l'arrangement de tous les chromosomes d'une cellule. Elle est utile dans l'observation de gros réarrangements chromosomiques, tels que les réarrangements de chromosomes complets (monosomie, trisomie...). Son usage reste par contre limité pour les CNVs car ne peut détecter que ceux dont la taille est supérieure à 5 Mb. Une autre technique de cytogénétique plus résolutive est le FISH (*Fluorescent In Situ Hybridisation*), elle permet de cibler des régions chromosomiques spécifiques et prédéterminées avec des sondes fluorescentes afin de détecter la présence de CNVs dont la taille est d'un minimum de 200 Kb. Une dernière technique de cytogénétique permet de détecter les CNV grâce à des puces d'Hybridation Génomique Comparative (CGH array), qui permettent l'hybridation compétitive de l'ADN du patient et d'un ADN contrôle avec des milliers de fragments d'ADN représentant différentes parties du génome. Etant donné que l'ADN du patient et l'ADN contrôle sont marqués avec des fluorophores différents, on peut quantifier l'intensité de fluorescence émise par chaque fragment hybridé et calculer un ratio d'intensité qui reflèterait la différence du nombre de copies dans l'ADN patient comparé à l'ADN contrôle (Figure 7).

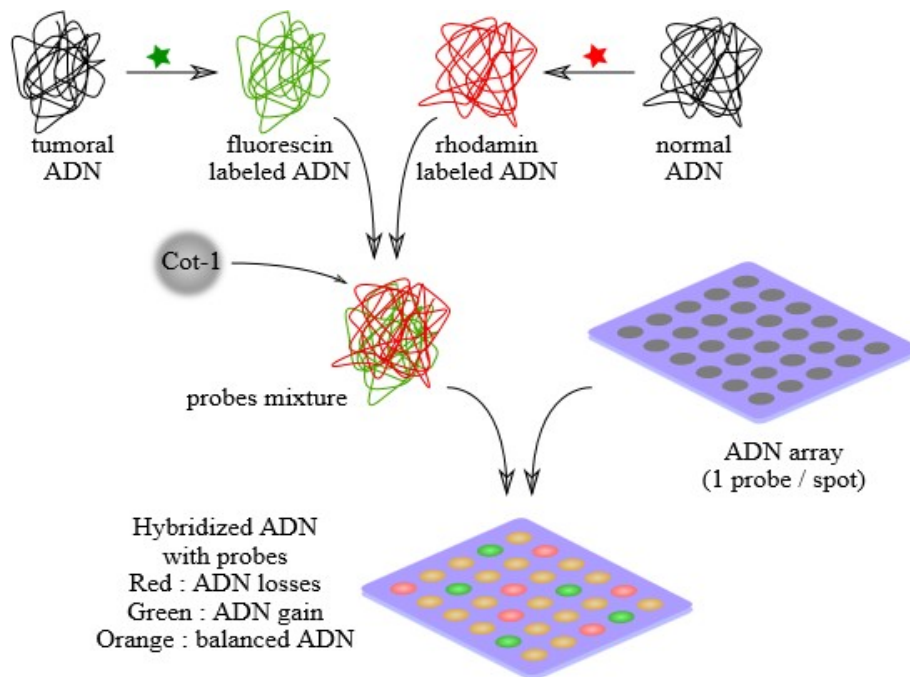


Figure 7 : Principe de la puce d'hybridation génomique comparative.

Source : commons.wikimedia.org

Dans les années 2000, on a assisté à la naissance des puces de génotypage de SNP (polymorphisme d'un seul nucléotide) qui ont été beaucoup utilisées depuis pour la détection des CNVs (Les CNVs utilisés dans notre projet ont été détectés grâce à ce type de puces). Elles sont également basées sur l'hybridation de l'ADN suivie par une détection de marquage fluorescent. Elles permettent non seulement d'obtenir l'information sur le génotype de chaque SNP, utile pour des études d'associations mais aussi de détecter les CNV grâce à l'analyse de l'intensité du signal d'hybridation ainsi que la recherche des régions de perte d'hétérozygotie. Les plus connus de ces puces sont Illumina et Affymetrix (Figure 8).

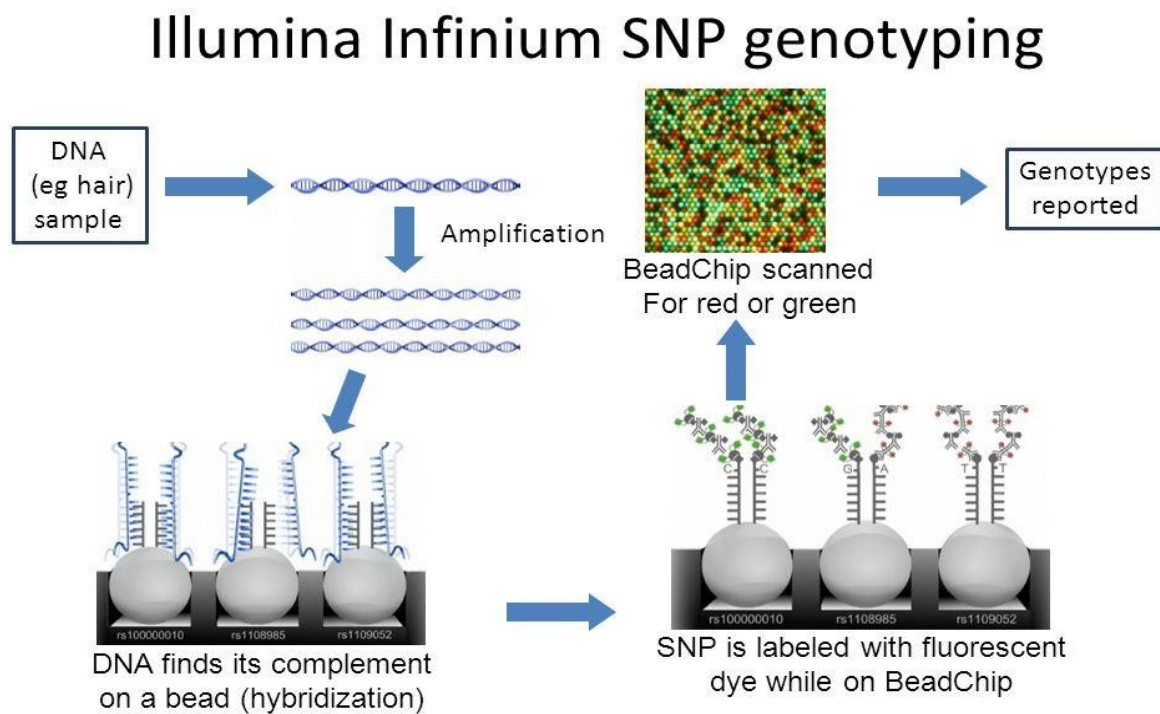


Figure 8 : Principe du génotypage de SNP de type Illumina.

Après préparation de l'ADN, une amplification jusqu'à 1000 fois est réalisée à l'aide de whole genome. Cet ADN est fragmenté et capté par des amorces (primers) sur la puce (BeadChip) constituant l'étape de l'hybridation. Ensuite on a une extension à l'aide de nucléotides fluorescents pour que finalement le signal soit scanné et analysé par un logiciel qui permet de donner le génotype.

Source : Understanding Conventional and Genomic EPDs (<http://slideplayer.com>).

Deux principales techniques de séquençages à haut débit sont utilisées pour la détection de CNV ; le séquençage d'exome qui permet de détecter tous les CNVs touchant la région codante du génome et le séquençage du génome complet qui lui permet de séquencer tout le génome et

ainsi détecter même des CNVs intergéniques non observés avec du séquençage d'exome. Le séquençage du génome complet est une promesse qui est de plus en plus utilisée car rendu moins coûteuse.

➤ **CNV associés aux TSA.**

Les études ont montré que les CNVs, qu'ils soient de type délétion ou duplication sont impliqués dans les TSA^{56,61,62}. Ces CNVs sont plus larges, contiennent plus de gènes comparé à ce qui est observé chez les contrôles^{55,61,62}. Cette association est principalement due aux CNVs apparus *de novo* qui sont jusqu'à 4 fois plus fréquents chez les individus autistes comparés aux contrôles et sont retrouvés chez environ 3% des autistes⁶¹. La contribution des CNV hérités reste quant à elle peu démontrée. Effectivement, à l'exception de certains CNV récurrents, qui peuvent être hérités et qui sont clairement associés à l'autisme, les études n'ont pas pu identifier une contribution claire des CNVs hérités dans les TSA, (Figure 11)⁶¹.

Plusieurs études se sont intéressées aux CNV récurrents et à leur implication dans les TSA. En étudiant 2591 autistes et 2100 frères ou sœurs non autistes, Sanders et al. ont confirmé l'association de six régions du génome (loci) aux TSA (1q21.1, 3q29, 7q11.23, 16p11.2, 15q11.2-13, et 22q11.2)⁶¹. Girirajan et al⁶³ retrouvent que les CNVs récurrents sont plus présents chez les autistes comparés aux contrôles (2,8% vs 1,3%, $p=7,2.10^{-7}$). 48% de ces CNVs étaient hérités des parents et plusieurs d'entre eux ont été préalablement associés aux TSA, notamment : 16p11.2, 1q21.1, 15q13.3, 17p12 et 17q12. Les auteurs observent que l'augmentation de la taille de ces délétions était corrélée à une diminution du QI mais n'avait aucun impact sur la sévérité des symptômes autistiques tel que mesuré par le score ADOS (Autisme Diagnostic Observation Schedule). Comparativement la taille des duplications n'est pas corrélée au QI mais prédit la sévérité des symptômes autistiques (score ADOS), (Figure 11)⁶³.

Des études ont essayé de caractériser l'impact des CNV récurrents sur la clinique. En voici quelques exemples :

Délétion/Duplication 16p11.2. Ces CNVs dont la fréquence avoisine 0,5% dans la population générale sont responsables de traits phénotypiques distincts entre les délétions et les duplications

^{64,65}. En effet Jacquemont et al ont montré que les délétions 16p11.2 étaient associées en plus de l'autisme à une macrocéphalie et une obésité tandis que les duplications de ce locus (région du chromosome) qui causerait elle la schizophrénie est associée à une microcéphalie et une diminution du poids. Ainsi la délétion et la duplication d'une même région du génome est responsable d'un phénotype clinique radicalement différent (en miroir), (Figure 9) ⁶⁶.

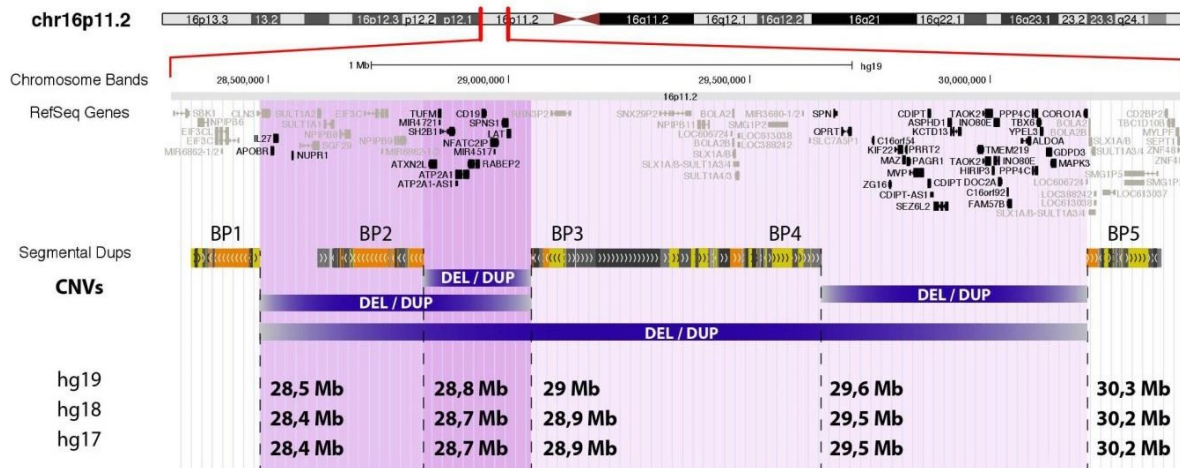


Figure 9 : Représentation de la région 16p11.2.

Cette figure montre en utilisant UCSC, le chromosome 16 et plus spécifiquement la bande 16p11.2 avec les différents CNVs récurrents présents au sein de cette région.

Source : <http://www.minds-genes.org>.

Délétion 22q11. Ce syndrome dont la prévalence est de 1 sur 4000 se caractérise par une hétérogénéité clinique importante. En effet, dans la délétion 22q11 on peut retrouver des malformations cardiaques congénitales voire palatines, des anomalies vertébrales, une dysmorphie faciale, un retard du développement et une immunodéficience. De plus, il a été démontré une association claire entre le syndrome 22q11.2 et l'autisme ⁶⁷. Au passé ce syndrome avait été divisé en trois sous syndrome cliniquement distincts (syndrome de Di George, syndrome cardio-facial, syndrome vélocardiofacial) mais cette catégorisation n'est plus d'actualité (ghr.nlm.nih.gov), (Figure 11).

➤ **Mutations ponctuelles.**

Les SNVs sont causés par la modification au niveau de l'ADN d'un seul nucléotide. Grace aux avancés des techniques de séquençage d'exome, tous les SNV touchant la région

codante peuvent être détectés et analysés. De plus, à l'aide des puces de génotypage on arrive à détecter les SNVs à travers tout le génome (région codante et non codante).

Plusieurs études se sont intéressées à l'impact des SNVs rares (fréquence observée dans la population < 1%) dans l'autisme. Elles retrouvent une contribution claire des SNVs rares qui altèrent complètement la fonction du gène (variants faux-sens, d'épissage ou induisant un décalage du cadre de lecture). Cette contribution concernerait les SNVs *de novo* mais aussi hérités. La contribution des SNV hérités proviendrait préférentiellement des variants délétères transmis de la mère à son enfant autiste et concernerait les gènes les plus intolérants à la mutation (haploinsuffisant) selon le score RVIS (score d'intolérance aux variations résiduelles)⁶⁸.

➤ **Variants communs.**

Les variants communs sont définis par une fréquence est $\geq 1\%$ dans la population et sont hérités des parents. Les études estiment que les variants communs sont des facteurs de risque importants de TSA mais leurs effets étant individuellement petits, cette contribution n'a pas encore pu être clairement démontrée ⁵³.

Différents chercheurs ont essayé d'investiguer l'impact de ces derniers dans les TSA grâce à des études d'association pangénomique portant sur les polymorphismes d'un seul nucléotide (SNP) mais aussi sur les polymorphismes de nombre de copies (CNP). Ces études explorent la variation à travers le génome de ces SNP voire CNP et leur possible association avec des phénotypes tel que l'autisme.

On notera une inconsistance des résultats obtenus qui serait attribuée soit à une absence de contribution des variants communs dans les TSA soit à des effets trop faibles individuellement pour être observés ^{53,54}.

Les dernières données suggèrent que les variants communs dans leur globalité, expliqueraient 49% de la variance des TSA mais que considérés individuellement, ces derniers n'ont qu'un petit effet. Ainsi il faudrait une taille d'échantillon importante pour pouvoir les observer. De plus, l'hétérogénéité clinique caractéristique des TSA diminuerait la puissance de ces études ^{53,54}.

Malgré ces limites, certains gènes ont été proposés comme candidats potentiels dans les TSA grâce à l'étude des variants communs. On retrouve ; le gène «*RELN*» codant la protéine reelin

qui joue un rôle dans la migration de différents types de cellules neuronales⁶⁹, le gène «*SLC6A4*» qui code la protéine qui transporte la sérotonine SERT vers l'espace pré-synaptique⁷⁰, le gène «*En-2*» qui code l'homéoprotéine «Engrailed», une protéine avec un rôle important dans le développement du système nerveux central^{71,72}.

➤ **TSA Syndromique.**

Une importante percée dans la génétique de l'autisme a été réalisée grâce à la découverte de certaines formes monogéniques ou dites syndromiques de TSA. Ces dernières sont cliniquement sévères avec des symptômes allant du retard développemental, malformations aux traits autistiques. Cette forme syndromique de TSA nécessite des mutations génétiques hautement délétères et pénétrantes. Elles n'expliquent globalement pas plus de 5% des TSA. Différents modes de transmission sont observés (Figure 10)⁵⁴ :

- Mendélien dominant : une mutation présente chez un des parents est transmise à l'enfant et provoque impérativement le syndrome ; exemple : Syndrome de Timothy, sclérose tubéreuse complexe....
- Mendélien récessif : nécessite la présence de la mutation chez les deux parents. La transmission d'une seule mutation ne suffit pas et l'enfant doit hériter la mutation des deux parents pour développer le syndrome ; exemple : Syndrome de Cohen, Syndrome de Smith-Lemli-Opitz
- Lié à l'X : Cela concerne des mutations touchant des gènes situés sur le chromosome X, transmise préférentiellement par la mère à sa descendance. Etant donné que les garçons sont porteurs d'une seule copie du chromosome X, la pénétrance de ces mutations et de leurs syndromes associés sera plus prononcée chez ces derniers.⁵⁴.
- *De novo* : La mutation n'est pas présente chez les parents et apparaît chez l'enfant qui développe le syndrome ; exemple : Syndrome d'Angelman. (Figure 6)⁵⁴.

Une douzaine de ces syndromes génétiques rares sont répertoriés dont (Figure 10)⁵⁴ :

Syndrome de l'X fragile. Il est causé par une expansion des trinuécléotides CGG situés en amont du gène *FMRI* (Fragility Mental Retardation 1), responsable d'une hyperméthylation du promoteur et ainsi une diminution de la synthèse de la protéine FMRP (fragile X mental retardation protein). Deux formes symptomatiques distinct du syndrome de l'X fragile peuvent être distinguées ⁷³⁻⁷⁵:

- La prémutation qui se caractérise par une expansion du trinuécléotide CGG allant de 55 à 200 fois. Les individus porteurs de cette prémutation ne développent pas le syndrome de l'X fragile mais sont à risque de développer le syndrome de l'Insuffisance Ovarienne Précoce liée à l'X fragile (FXPOI)⁷⁴ chez les femmes voire le syndrome du tremblement-ataxie lié à l'X fragile (FXTAS) chez les hommes et à moindre effet les chez les femmes.
- La mutation complète qui est induite par une expansion du trinuécléotide CGG de plus de 200 fois et qui cause le syndrome de l'X fragile. Ce dernier se caractérise par une grande variabilité clinique avec des symptômes tels que ; la déficience intellectuelle, l'anxiété, l'autisme, les troubles de l'attention.

Environ 5 à 60 % des individus avec un syndrome de l'X fragile ont de l'autisme et 90 % au moins présentent des traits autistiques. Le syndrome de l'X fragile représente la forme génétique qui explique le plus de cas de TSA 1.94%⁷⁶.

Le syndrome d'Angelman/ Prader-Willi. C'est un désordre génétique rare dont la fréquence est estimée à 1 sur 12000/20000. Il se manifeste par un retard du développement, déficience intellectuelle, retard du langage, troubles moteurs, épilepsie, microcéphalie et traits autistiques.

Il est provoqué majoritairement par une perte de la fonction du gène *UBE3A* produite lors d'une délétion de la région 15q11-13. Ce gène est exprimé dans le cerveau ou il est soumis à empreinte parentale ce qui signifie que le même gène engendre un effet différent quand la copie maternelle est touchée (Syndrome d'Angelman) comparé à quand la copie paternelle est touchée (Syndrome de Prader-Willi), (ghr.nlm.nih.gov).

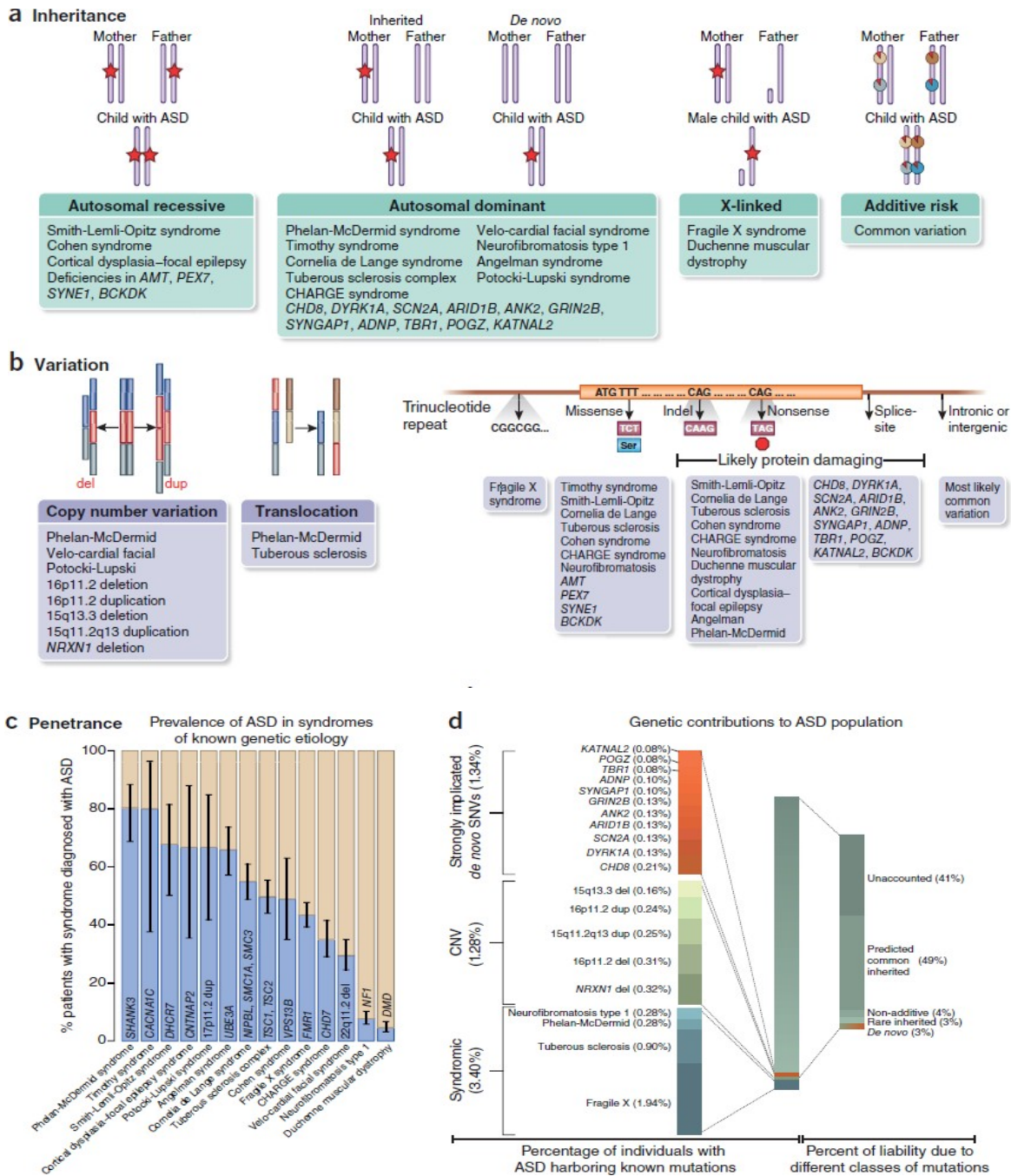


Figure 10. Architecture génétique des troubles du spectre de l'autisme.

(a) Mode d'héritabilité de TSA syndromiques ainsi que des gènes récemment associés à l'autisme. Les étoiles rouges indiquent des allèles. (b) Différents types de variations génétiques et de désordres neurodéveloppementaux retrouvés dans les TSA. (c) La pénétrance de variants syndromiques connus. L'intervalle de confiance à 95% basé sur le score Wilson est montré. (d) le pourcentage de contribution des différents types de variants connus dans l'autisme.

Source : Revue de la littérature ⁵⁴

3.2 Modèles génétiques.

Deux modèles génétiques s'opposent pour expliquer les TSA: Le premier se base sur des variants suffisamment délétères pour engendrer à eux seuls les TSA (modèle du gène majeur) alors que le second se base sur l'accumulation de variants ayant un faible effet mais qui ensemble entraîne un TSA (modèle polygénique) ⁵⁴.

a. Modèle du gène majeur.

Il stipule qu'une voire quelques mutations rares hautement pénétrantes sont suffisantes pour provoquer l'autisme. Ce modèle est renforcé par ^{54,68,77} :

- La découverte d'une douzaine de syndromes provoquant l'Autisme et qui sont dus à des mutations rares et très pénétrantes (Figure 11).
- L'excès de mutations *de novo* chez les personnes Autistes.
- La présence d'une transmission préférentielle de variants délétères de la mère à son enfant autiste comparé à sa fratrie non atteinte.

Ce modèle présente néanmoins des inconsistances à différents niveaux :

- On ne retrouve ces mutations *de novo* que chez environ 3% des personnes autistes⁵⁴:
- Si la majorité du risque d'Autisme est expliqué par les mutations *de novo*, on s'attendrait à avoir un risque plus bas chez la fratrie de ces Autistes, ce qui n'est pas le cas ^{54,78}.
- L'estimation de l'héritabilité effectuée grâce au calcul du taux de variations génétiques communes (SNP) partagées avec les individus non apparentés démontre que le pourcentage de variance expliqué par les variants génétiques communs est de 49%⁵³.

b. Modèle polygénique.

Il postule que l'autisme résulte d'une accumulation d'un nombre important de variants génétiques ayant chacun un petit effet pour l'autisme et qui en combinaison avec des facteurs environnementaux atteignent le seuil nécessaire pour un diagnostic d'autisme. Plusieurs données vont dans le sens de ce modèle⁵⁴ :

- La récurrence élevée de l'autisme dans les familles avec au moins un individu autiste suggère une importante composante génétique héritée^{54,79}.

- Les apparentés au premier degré de personnes autistes montrent des traits autistiques plus importants que dans la population générale^{54,80}.
- Un pourcentage de variance expliqué par les variants génétiques communs estimé à 49%^{53,54}.

La faiblesse de ce modèle réside dans l'absence d'association statistique entre ces variants communs et l'autisme. Cela est probablement dû à une taille d'échantillon insuffisante pour pouvoir capter les effets faibles de ces variants⁵⁴.

Les dernières données en génétique montrent qu'aucun de ces deux modèles n'est faux mais que ces derniers se complètent. Ainsi dans ce large spectre de l'autisme, on retrouve une contribution importante des variants communs dont l'effet est additif (49%), une contribution moindre mais importante des variants rares *de novo* (3%) et rare hérités (3%)^{53,54}.

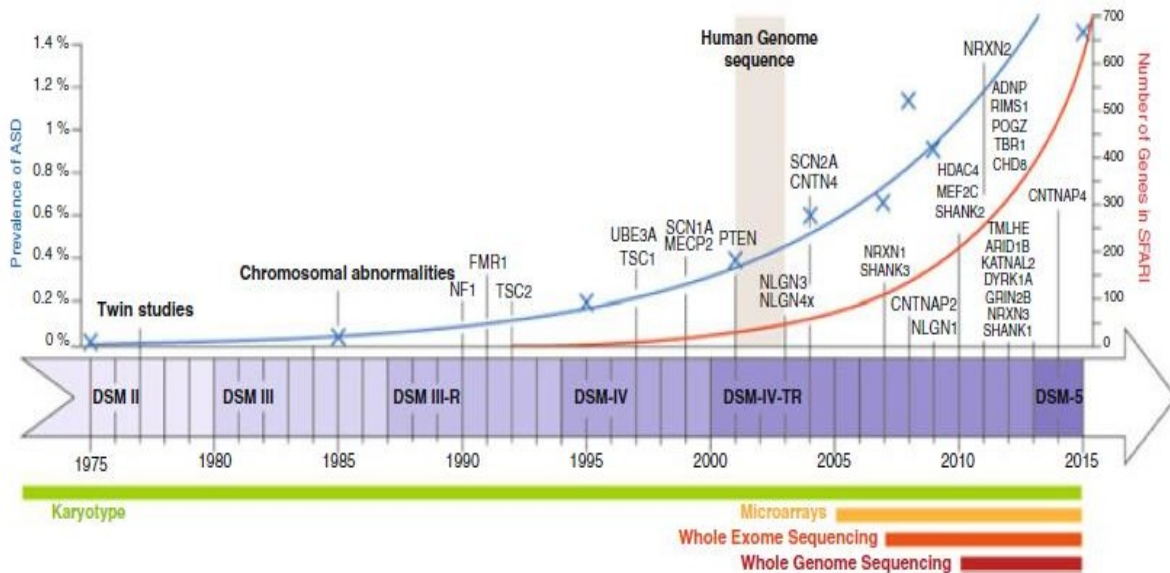


Figure 11 : L'évolution de la génétique de l'autisme entre 1975 à 2015.

L'information sur l'augmentation de la prévalence et du nombre de gènes associés à ce trouble est présente. Les différentes versions du DSM ainsi que le progrès des technologies de détection génétiques sont représentés.

Source : A time for metabolism and hormones⁵⁸.

3.3 Facteurs environnementaux.

Les études portant sur les facteurs environnementaux impliqués dans l'Autisme sont moins avancées que les études génétiques. Elles ont plusieurs limitations au niveau de la méthodologie avec une taille d'échantillon souvent petite, un design épidémiologique non adapté et des mesures cliniques peu robustes. Afin de clarifier la contribution des facteurs environnementaux dans l'autisme, Modabbernia et al. ont regroupé les données de revues systématiques et méta-analyses publiés jusqu'à Décembre 2016. On peut ainsi distinguer deux catégories de facteurs environnementaux selon si le mécanisme biologique est suggéré ou pas ⁸¹ :

- Facteurs environnementaux dont le mécanisme biologique est connu ou suggéré :

➤ Age parental.

L'association entre l'augmentation de l'âge parental à la conception et le risque d'autisme a été longuement évoquée et étudiée mais les résultats étaient parfois contradictoires. Une méta-analyse de 27 études portant sur cette problématique a montré qu'une augmentation de l'âge parental à la conception de 10 ans s'accompagnait d'une augmentation du risque d'autisme de 18% pour la mère et de 21% pour le père ⁸¹.

Les mécanismes sous-jacents à cette augmentation de risque seraient potentiellement différents chez la mère que chez le père. L'augmentation du risque liée à l'âge paternel serait principalement due à un plus grand nombre de mutations *de novo* voire de modifications épigénétiques accumulées avec l'âge, tandis que l'augmentation du risque liée à l'âge maternel serait liée non seulement aux altérations génomiques mais à une augmentation des complications périnatales et obstétricales. ⁸²⁻⁸⁴.

➤ Complications et états liés à la grossesse.

Gardener et al. ^{85,86} ont effectué deux revues de littérature avec méta-analyses afin d'investiguer la contribution des facteurs de risques prénataux, périnataux et néonataux dans l'autisme. Ils retrouvent une association significative pour certains de ces facteurs : type de présentation fœtale, complications liées au cordon ombilical, détresse fœtale, grossesse multiple, poids de naissance bas, malformations congénitales, hyperbilirubinémie, ordre de naissance.

Dans une autre revue systématique regroupant sept études, Conde-Agudelo et al⁸⁷. retrouvent qu'un espacement des naissances de moins de 12 mois voire de plus de 60 mois augmenterait le risque d'autisme. D'autres facteurs tels que l'accouchement par césarienne⁸⁸, l'infection pendant la grossesse⁸⁹, le diabète gestationnel⁹⁰, le diabète pré gestationnel⁹⁰ et l'obésité⁹¹ auraient une contribution mineure. A l'inverse, aucune association entre la reproduction in vitro et l'autisme n'a été observée⁹².

Ces complications liées à la grossesse agiraient à travers l'hypoxie et l'ischémie cérébrale, qui seraient responsable d'un état de stress oxydatif à l'origine d'un dommage voire une mort neuronale. Des auteurs ont montré une expression plus élevée de la protéine FMRP entre la 36 et la 39^{ème} semaine de grossesse. La dysfonction de cette protéine est responsable du syndrome du fragile X, une des formes syndromiques les plus fréquentes d'Autisme (vu précédemment, section 1.3.1.2). Aussi ils observent que le niveau d'expression de cette protéine est plus bas chez les nouveaux nés atteints d'encéphalopathie hypoxique-ischémique. Ils postulent que les complications ischémiques pendant la grossesse agiraient potentiellement à travers l'altération de l'expression de cette protéine dont le rôle est essentiel dans la plasticité synaptique.

Différents traitements pris pendant la grossesse ont été incriminés dans l'Autisme mais peu d'entre eux ont été statistiquement associés à ce trouble. On peut citer le valproate, un traitement de l'épilepsie et des troubles bipolaires, pour qui l'association (qui semble être dose dépendante) a été plusieurs fois démontrée⁸¹.

➤ **Immigration.**

Des études ont montré que les familles issues de l'immigration avaient plus de risque (2.69%) de donner naissance à des enfants atteints d'autisme. Plusieurs hypothèses pourraient expliquer cette association dont la perturbation du microbiote en lien avec un changement diététique⁸¹.

- **Facteurs environnementaux dont l'association n'est pas claire :**

Ici nous discutons deux types de facteurs environnementaux ; soit des facteurs dont l'association statistique est présente mais aucun mécanisme biologique n'a pu être proposé, soit des facteurs environnementaux longuement suspectés dans l'autisme mais dont les dernières études n'ont

montré aucune association statistique. Il est important de mettre la lumière sur ces facteurs environnementaux même si leur association à l'autisme n'est pas claire car ces derniers pourraient s'avérer très important dans la genèse de ce trouble. Effectivement, Les dernières données suggèrent une forte composante environnementale dans l'autisme (~ 40%)⁸¹. De plus, il y'a considérablement moins d'effort portant sur l'étude des facteurs environnementaux comparé aux facteurs génétiques. Parmi ces facteurs environnementaux, on a :

➤ **Facteurs nutritionnels**

Des études ont montré un taux plus faible de protéines⁹³, de calcium⁹³, de zinc⁹⁴ et de vitamine D⁹⁵ chez les autistes comparé aux contrôles. De plus, on retrouve une possible interaction entre la déficience en acide folique et la dysfonction du gène de la methylene tetrahydrofolate reductase (*MTHFR*⁹⁶).

➤ **Exposition aux toxines.**

L'exposition à des substances toxiques a été longuement incriminée dans l'autisme et différentes substances ont été suspectées dont le thiomersal, le mercure organique, les métaux lourds et la pollution de l'air. Dans une méta-analyse de trois études cas-contrôle, Yoshimasu et al retrouvent qu'une exposition prénatale au mercure inorganique s'accompagnait d'une augmentation de 60% du risque d'autisme. Par contre, aucun lien entre l'exposition au thiomersal et l'autisme n'a été observé⁹⁷.

➤ **Vaccination.**

Etant données les craintes que certains vaccins ont provoqué quant à leur lien avec l'Autisme, plusieurs études se sont penchées sur cette problématique. Ainsi Taylor et al ont fait une méta-analyse de ces dernières et montrent que non seulement la vaccination n'était pas associée à un risque plus élevé d'Autisme mais que cette dernière pourrait même protéger contre ce trouble⁹⁸.

➤ **Consommation maternelle de tabac.**

Dans une méta-analyse de 15 études Rosen et al ne retrouvent aucune association entre la consommation de tabac pendant la grossesse et à l'accouchement et l'autisme⁹⁹.

Mise en contexte, objectifs et hypothèses.

Des études se sont intéressées à l'association entre les CNVs et le diagnostic catégoriel d'autisme. Il est connu que les CNVs larges, qu'ils soient de type délétion ou duplication contribuent aux TSA¹⁰⁰. Par contre, même s'il est établi que les apparentés aux premiers degrés à des personnes autistes (fratrie saine, parents) présentent des traits autistiques plus importants que ce qui est observé dans la population générale⁸⁰, aucun excès de CNVs ne fut rapporté chez eux comparé à la population générale. **Notre première hypothèse est qu'il existe un excès de CNVs rares $\geq 250\text{Kb}$ chez les probands autistes mais aussi chez leurs apparentés aux premiers degrés (fratrie saine, parents) comparé à la population générale.**

L'autisme est un trouble dont l'origine est hautement génétique ($\sim 59\%$ d'après les dernières estimations)^{53,54}. Néanmoins, le risque de TSA expliqué par la génétique résulte que dans 4 à 5% des cas (figure 10 (d)) d'une seule mutation dont l'effet est assez important pour engendrer à lui seul une grande partie de la symptomatologie autistique (TSA syndromique, section 3.2 du chapitre 1). Une douzaine de ces syndromes existent dont le plus fréquent est le syndrome de l'X fragile. Les autistes ayant ce syndrome ont été exclus de la SSC (Simon Simplex Cohort). Ainsi, dans l'écrasante majorité des cas d'autisme d'origine génétique ($\sim 55\%$), l'individu ne développe le TSA qu'à la suite d'une accumulation de plusieurs mutations (communes et rares) dont les effets sont individuellement petits mais cumulés provoquent l'autisme^{53,54}.

Il est possible d'utiliser les caractéristiques des gènes touchés par ces mutations pour évaluer l'impact de ces dernières sur le risque de TSA. Effectivement, des études ont montré que les gènes associés à l'autisme et/ou à une diminution du QI non verbal avaient de plus grands scores d'intolérance à la mutation^{68,121}. Ces scores permettent de classer les gènes en fonction du nombre de mutations qu'elles peuvent accumuler dans la population générale. Ainsi, un gène avec peu ou pas de mutations observées dans la population générale aurait un score d'intolérance élevé et vice versa (section 5 du chapitre 2).

La majorité des CNVs retrouvés chez les autistes sont rares voire uniques (observés que chez un individu). De ce fait, l'impact individuel de chacun de ces CNVs ne peut être estimé grâce à des études d'associations. Un CNV qu'il soit de type délétion ou duplication, emporte une partie du génome affectant ainsi un voir plusieurs gènes. A l'instar de l'autisme, le QI est aussi un trait

hautement héritable. Ainsi, une étude a retrouvé chez la population générale qu'un modèle additif/polygénique permettait d'estimer l'effet des CNVs sur la variation du QI non verbal. Ils retrouvent que l'impact d'un CNV sur la diminution du QI non verbal résulte de la somme des impacts des gènes contenus dans ce dernier et à leur niveau d'intolérance à la mutation (effet additif/polygénique) ¹²¹. **Notre deuxième hypothèse est que des caractéristiques géniques tels que les scores d'intolérance permettraient d'estimer chez une population autiste, l'effet d'une mutation d'un gène sur le risque de TSA/NVQI et que l'impact global d'un CNV résulterait de la somme (modèle additif/polygénique) des effets des gènes présents dans ce dernier.**

Les troubles du spectre de l'autisme sont caractérisés par une hétérogénéité génétique et clinique importante, notamment en terme d'intelligence telle que mesurée par le QI. Cette dernière peut varier chez les individus autistes de la déficience intellectuelle (un tiers des autistes) à des QI au-dessus de la moyenne. La déficience intellectuelle est ainsi une comorbidité clinique fréquente dans l'autisme, de plus, la majorité des CNVs retrouvés dans le TSA diminuent le QI et sont impliqués dans la déficience intellectuelle ^{61,101}. Les CNVs *de novo*, un des variants les plus étudiés et les plus fortement impliqués dans l'autisme sont connus pour diminuer le QI ⁶¹. Néanmoins, une étude a montré que les autistes porteurs de ces CNVs, développent une forme muette d'autisme (pas d'atteinte de la communication sociale ni du langage) comparé aux autistes non porteurs (à QI égal) ¹²². Ces données réunies mettent en évidence une relation ambiguë entre les CNVs, le QI et les traits spécifiques de l'autisme. **Ainsi nous posons l'hypothèse que l'association entre les CNVs rares et le risque de TSA est partiellement expliquée par l'altération du NVIQ.**

L'Objectif générale dans cette étude est de quantifier et prédire l'effet des CNVs sur le risque d'autisme ainsi que sur le NVIQ.

Nos objectifs spécifiques consistent à :

- 1) Réévaluer l'association entre la fréquence des CNVs rares ≥ 250 Kb et le risque d'autisme.

- 2) Prédire grâce aux caractéristiques géniques intrinsèques aux CNVs, le diagnostic d'autisme et l'altération du NVIQ.
- 3) Notre dernier objectif était d'évaluer si la relation entre les CNVs rares $\geq 250\text{Kb}$ et l'autisme telle qu'estimait par leur fréquence et leurs caractéristiques géniques, était entraînée par une atteinte du NVIQ.

CHAPITRE 2. METHODOLOGIE

1. Populations.

1.1 Population TSA.

Les données génotype et phénotype sont issues de la Simon Simplex Collection (SSC)¹⁰², une cohorte multicentrique regroupant des familles avec un seul enfant atteint de TSA. Deux types de familles sont présentes : (i) les familles « trio » (N=491) composées de l'enfant atteint d'autisme (proband) et des deux parents et (ii) les familles « quads » (N=2100) composées du proband, des deux parents et d'un membre de la fratrie non atteint de TSA. Les critères d'inclusion d'une famille dans l'étude sont que l'âge du proband se situe entre 4 et 18 ans, qu'il ait un diagnostic établi de TSA basé sur l'Autisme Diagnostic Observation Schedule (ADOS) et sur l'Autisme Diagnostic Interview-Revised (ADI-R). Les critères d'exclusion sont la présence d'un apparenté de premier, deuxième voire troisième degré avec un diagnostic ou une suspicion de TSA, un âge mental inférieur à 18 mois, un déficit neurologique sévère, des complications périnatales ou encore la présence d'un TSA syndromique⁶¹.

Etant donné que nous utilisons un jeu de donnée nettoyé et publié par l'équipe du Dr Stephan Sanders ⁶¹, les individus dont la puce de génotypage ne passe pas les critères de contrôle de qualité ont été exclus préalablement. De plus, nous avons exclu les familles pour lesquels l'information génétique était incomplète (N = 16). Enfin, comme nous nous intéressons à la mesure du QI chez les probands, nous avons exclu deux familles ou les probands n'avaient pas de mesure du QI. Notre échantillon final comprend 2573 probands, 2085 Sibling et 5146 parents avec 488 familles trio et 2085 familles quater.

1.2 Population générale.

Nous avons aussi utilisé deux cohortes de population générale dont les données génétiques (CNV) et clinique (QI) sont accessibles :

1. Imagen, une cohorte européenne multicentrique avec 2090 adolescents¹⁰³.
2. Saguenay Youth Study (SYS), une cohorte de canadiens français incluant 1032 enfants et 951 parents (486 familles)¹⁰⁴.

De la cohorte Imagen, nous incluons dans notre étude tous les individus avec l'information génétique (qui passent les critères de qualité de la puce) et le QI (N = 1744). De la cohorte Saguenay, nous incluons dans notre étude uniquement les enfants ayant la même technologie de détection des CNVs que leurs parents (N = 212 enfants) et pour lesquels l'information génétique

et le QI sont disponibles (N = 212 enfants). Nous utilisons aussi les parents de la SYS ayant l'information génétique et pas de test de QI réalisé (N = 850).

Un descriptif des cohortes est détaillé dans le tableau 1.

2. Variables cliniques.

2.1 QI.

Dans la SSC, le QI a été évalué chez les probands grâce à des tests standardisés qui mesurent le quotient intellectuel général (FIQ), le quotient intellectuel verbal (VIQ) et le quotient intellectuel non verbal (NVIQ). Différents tests ont été réalisés dépendamment de l'âge et du niveau intellectuel des individus : la deuxième édition du test *Differential Ability Scales (DAS-II)*¹⁰⁵, la quatrième édition du test Wechsler Intelligence Scale for Children (WISC-IV)¹⁰⁶, le test Wechsler Abbreviated Scale of Intelligence (WASI)¹⁰⁷ ou le Mullen Scales of Early Learning (MSEL)¹⁰⁸. Un QI standard (deviation IQ), calculé dépendamment de combien un individu dévie d'un QI moyen de 100 est présent chez la plupart des probands (N=2195). Par contre pour ceux qui n'ont pas pu compléter ces QI standards étant donné leur âge ou leur niveau de développement, une estimation du ratio QI est générée en divisant l'âge mental par l'âge chronologique et en multipliant par 100.

Dans les populations générales en plus de l'âge et du sexe, nous avons pour les individus d'Imagen l'évaluation du QI via WISC-IV¹⁰⁶ tandis que pour les enfants de SYS nous l'avons via WISC- III¹⁰⁹. Les scores standards de QI sont présents pour tous les participants à la fois pour le VIQ et le NVIQ.

Le NVIQ reflète mieux les capacités cognitives réelles des individus autistes que le VIQ. Ainsi tout au long de notre projet nous utilisons le NVIQ pour mesurer le niveau d'intelligence non verbale des individus¹¹⁰.

2.2 Score de sévérité calibré de l'ADOS (ADOS css).

Cette variable a été générée à partir des scores totaux bruts de l'ADOS, qui furent standardisés pour créer une variable continue qui évalue la sévérité des symptômes autistiques indépendamment de l'âge et du niveau de langage (contrairement aux scores bruts de l'ADOS). L'ADOS css représente une mesure évaluant la globalité des symptômes spécifiques de l'autisme (communication et interaction sociale, mouvements répétés et stéréotypés)^{111,112}.

Variables	Imagen	SYS (parents)	SYS (enfants)	SSC (parents)	SSC (fratrie saine)	SSC (probands)
N individus (inclus dans l'analyse)	1,744	850	212	5146	2085	2573
Technologies	610Kq et 660Wq	Human omni express v12	Human omni express v12	Illumina 1Mv1, 1Mv3, Omni2,5M	Illumina 1Mv1, 1Mv3, Omni2,5M	Illumina 1Mv1, 1Mv3, Omni2,5M
N délétion total*, moyenne par individu	1,584	609	483	1570 (0,67)	621 (0,66)	826 (0,67)
N duplication total*, moyenne par individu	1,179	608	465	1768 (0,76)	732 (0,78)	986 (0,80)
N Délétions 250Kb, moyenne par individu	41 (0,02)	17 (0,02)	4 (0,02)	130 (0,02)	52(0,02)	130 (0,05)
N duplications ≥ 250Kb, moyenne par individu	112 (0,06)	48 (0,05)	18 (0,08)	421 (0,08)	171 (0,08)	274 (0,10)
Age (années); moyenne	14	NA	15	41	10	9
Sexe (male) N. (%)	853 (48%)	448 (52%)	103 (48%)	2573 (50%)	972 (46%)	2243 (86%)
NVIQ, moyenne	106	ND	103	ND	ND	84
ADOS css, moyenne	ND	ND	ND	ND	ND	7

Tableau 1: Descriptif des cohortes utilisées dans nos analyses. Légende : ND : non disponible, N : nombre, total* : le nombre total de CNVs correspond à tous les CNVs rares \geq 50Kb.

3. Détection des CNVs.

L'ensemble des données collectées dans ce travail de master sont issues de plusieurs travaux. Pour la cohorte SSC, les données proviennent de l'article publié par Stephan Sanders en 2015 dans le journal *Neuron*⁶¹. Pour les populations générales, les données sont issues des travaux réalisés au sein de notre laboratoire, détaillés dans l'article « Measuring and predicting the effect of copy number variants on general intelligence in 4 community-based samples. », publié en 2018 dans *JAMApsy*.

➤ SSC.

L'ADN a été extrait du sang complet et de cellules lymphoblastes au « Rutgers University Cell and DNA Repository » (RUCDR). Une fois obtenu, cet ADN a été analysé grâce à une puce de génotypage Illumina (1Mv1, 1Mv3 Duo, or Omni2.5M). Les CNVs ont été détectés par l'équipe du Dr. Stephan Sanders en utilisant PennCNV¹¹³, QuantiSNP¹¹⁴ et GNOSIS puis joints à l'aide de CNVision (www.CNVision.org)⁶¹. Les paramètres de contrôle de la qualité sont, un taux de call $\geq 98,5\%$ puis par algorithme⁶¹ :

- PennCNV: déviation standard du ratio du log R (LRR) > 0.28 , déviation de la fréquence de l'allèle B (BAF) > 0.01 , Facteur d'ondulation (WF) déviant du 0 par > 0.05 .
- QuantiSNP: valeurs extrême de la BAF < 0.1 , déviation standard de la BAF < 0.2 , déviation standard du LRR < 0.4 .
- GNOSIS: score de qualité > 10 .

CNVision permet d'identifier via la mesure du LRR, deux causes de mauvaise prédiction de CNV :

- Une valeur avec beaucoup d'ondulations.
- Un nombre important de sondes avec des valeurs très négatives.

De plus, une nouvelle approche a été développée par cette équipe pour améliorer le contrôle de la qualité des CNVs qui consiste en une valeur (pCNV), qui évalue, la variabilité par SNP en LRR ainsi que le nombre de SNP basé sur la BAF. Tous les CNVs *de novo* ont été validés par qPCR⁶¹. Une fois les CNVs détectés, un script a été utilisé au sein de l'équipe du Dr. Stephan Sanders pour vérifier que ces derniers ont été concaténés correctement. Chaque concaténation a été vérifiée visuellement⁶¹.

➤ **Imagen/Saguenay.**

L'ADN a été extrait du sang complet. Les critères utilisés pour garantir la qualité de l'échantillon sont : un taux d'appel 99%; un LRR - écart-type $<0,35$; une BAF - écart-type $<0,08$ et un | Facteur d'ondulation | $<0,05$.

Pour la cohorte Imagen, 2 090 individus ont été génotypés via deux technologies de puces Illumina : 610Kq (N de sondes = 620 901, N individus = 708) et 660Wq (N de sondes = 657 366, N individus = 1 385). Le génotypage a été réalisé au Centre National de Génotypage (CNG, Paris, France). Par souci d'homogénéité dans la détection des CNVs, seules les sondes communes entre les deux puces ont été utilisées. Ainsi 90,5% des sondes pour la puce 610Kq ont été gardées car présentes aussi sur la puce 660Wq (85.5%). Pour la cohorte SYS, 566 enfants ont été analysés via Illumina 610Kq (N de sondes = 620 901) puis 1,395 enfants et parents via HumanOmniExpress BeadChip -V12 (HOE-V12) (N de sondes = 730,525). Le génotypage a été effectué au CNG pour la 610Kq et au centre d'analyse de génome de Helmholtz Zentrum München (Munich, Allemagne) pour la HOE-V12. Etant donné l'importante différence de sensibilité de détection dans la technologie SYS (HOE-V12 et 610Kq), nous avons décidé de ne garder que les enfants qui partagent la même technologie que leurs parents (HOE-V12) et cela dans le but de pouvoir estimer correctement la fréquence de CNVs calculée chez les parents appartenant à la cohorte SYS.

Les critères de contrôle de la qualité par puce pour les deux cohortes sont : un taux d'appel de 99%; déviation standard du LRR $<0,35$; déviation standard de la BAF $<0,08$ et | Facteur d'ondulation | $<0,05$. Par la suite, PennCNV¹¹³ et QuantiSNP¹¹⁴ ont été utilisés pour détecter les CNVs selon trois paramètres : (i) nombre de sondes consécutives pour la détection des CNVs ≥ 3 ; (ii) taille du CNV $\geq 1\text{Kb}$ et (iii) scores de confiance ≥ 15 . Les CNVs détectés ont été fusionnés à l'aide de CNVision⁶¹. Une fois les CNVs détectés, nous avons concaténé les CNVs qui étaient incorrectement détectés via un script « python » développé par nos collaborateurs au sein du laboratoire du Dr Thomas Bourgeron (Institut Pasteur, Paris, France). Celui-ci regroupe tous les CNVs du même type avec une taille $\geq 500\text{Kb}$ pour ≥ 100 SNPs séparés par moins de 150Kb en un seul CNV. Chaque concaténation a été vérifiée visuellement.

Après ces critères de contrôle de la qualité, on se retrouve avec 1744 pour Imagen et 1062 pour SYS (N enfants= 212 et N parents= 850 pour un totale de N famille= 486). Il est à noter que trois individus ont été retirés de l'étude car ils avaient des CNVs de plus de 10Mb dans la population générale. De plus, on a exclu au sein de la population générale (Imagen, SYS) les CNVs dont le score de confiance est inférieur à 30 pour les deux algorithmes.

Au sein des trois cohortes (Imagen, SYS, SSC), Les CNVs dont la taille est < 50 Kb ou qui chevauchent des duplications segmentaires à plus de 50% ont été exclus. De plus, tous les CNVs supérieurs ou égaux à 250 Kb ont été vérifiés manuellement par un opérateur au sein de notre laboratoire.

4. Définition des CNVs rares.

Nous définissons les CNVs rares selon la base de données pour variants génomiques (DGV), (hg19, www.dgv.tcag.ca), (chevauchement minimum de 90%) et les observations intra cohortes (chevauchement minimum de 50%).

Un CNV est considéré comme rare si tous les critères listés ci-dessous sont remplis :

- Fréquence DGV $\leq 0.1\%$.
- Fréquence du CNV calculée chez les parents pour la SSC $\leq 0.1\%$, soit au moins 1 observation chez les parents pour SYS ou une fréquence intra cohorte $\leq 0.1\%$ dans Imagen.
- Un chevauchement <50% avec des régions de CNV communes (>1%) dans DGV.

5. Annotation des CNVs.

Les CNVs ont été annotés en utilisant ANNOVAR¹¹⁵. Dans une seconde étape d'annotation, les gènes sont annotés en utilisant des scores d'intolérance (transformés tel qu'expliqué ci-après) :

- la probabilité d'être intolérant à la perte de fonction du gène (pLI)¹¹⁶ : les valeurs sont comprises entre 0 et 1 où 1 signifie que le gène est complètement intolérant ;
- le score d'intolérance aux variations résiduelles (RVIS)⁶⁸ entre 0 et 100, celui-ci a été transformé par « 100-RVIS », ainsi 100 devient la valeur associée au gène le plus intolérant ;

- le score DEL (score CNV pour les délétions) ¹¹⁷ entre -2,62 et 3,81, celui-ci a été transformé par « score DEL + minimum (score DEL) » de telle sorte qu'il n'y ait plus de valeurs négatives. 6,43 représente la valeur associée au gène le plus intolérant.
- le score DUP (score CNV pour les duplications)¹¹⁷ entre -2,53 et 2,86, celui-ci a été transformé par « score DUP + minimum (score DUP) » de telle sorte qu'il n'y ait plus de valeurs négatives. 5,39 représente la valeur associée au gène le plus intolérant.

Pour les trois scores détaillés ci-dessus, la valeur par défaut associée à un gène n'ayant pas de score est 0. On génère un score CNV global qui résulte de la somme des scores des gènes complètement inclus dans le CNV. Si un individu présente plusieurs CNVs, on additionne les scores de chacun de ces CNVs pour générer un score global par individu. Etant donné que notre modèle se base sur une approche polygénique/additif ou l'impact global résulterait d'une accumulation de petits effets, nous avons décidé de traiter le pLI score comme une variable continue attribuant ainsi à chaque gène un poids. Ces petits poids s'additionnent au sein d'un CNV pour donner l'effet clinique observé. Dans une approche monogénique, le pLI pourrait être traité en binaire. Ainsi, l'individu aura un score pLI de 0 s'il ne possède pas de mutations touchant des gènes importants (selon le score pLI) et 1 si c'est le contraire.

Nous avons défini les CNVs récurrents dans chaque cohorte. Ceux-ci sont définis par plusieurs articles et représentent une liste de 121 régions de CNVs récurrents associés à des troubles neuro-développementaux ou psychiatriques ¹¹⁸. Ainsi si un CNV est chevauchant à plus de 40% avec l'une de ces régions et inclut le ou les gènes clés de cette région (si connu) alors il est considéré comme étant un CNV récurrent. Il est à noter que lorsque les coordonnées de la région ne comprenaient qu'un seul gène (par exemple, *NRXN1*), nous avons défini comme CNVs récurrents les délétions emportant au moins un exon du gène et les duplications couvrant la totalité du gène.

6. Analyses statistiques.

Les analyses incluent uniquement les CNVs rares $\geq 250\text{Kb}$ car du fait de la divergence de technologies de détection des CNVs entre les cohortes, nous travaillerons uniquement sur les CNVs de grande taille. La raison est qu'à une telle taille les CNVs sont détectables quelle que soit la technologie ou les paramètres de détection. La seule exception à cette règle consiste en nos

analyses de la section 6.1.b qui sont conduites uniquement sur les probands de la SSC ce qui élimine le biais de technologie présent dans nos autres analyses. Cela nous permet d'inclure tous les CNVs dont la taille est supérieure ou égale à 50Kb. On exclut par contre les petits CNVs (<50Kb) car des données de génotypage ne permettent pas de les détecter correctement.

6.1 Association entre la fréquence de CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme.

a. Comparaison entre les probands et la fratrie saine voire la population générale.

Tout d'abord, nous avons évalué si la fréquence des CNVs rares $\geq 250\text{Kb}$ prédisait significativement le diagnostic d'autisme en comparant les probands à leur fratrie saine puis puis à la population générale.

Pour la comparaison probands-fratrie saine, nous avons utilisé un modèle de régression logistique conditionnelle pour données appariées afin de tenir compte de la nature corrélée des données où des individus peuvent appartenir à la même famille. De plus, nous ajustons ce modèle sur le facteur sexe (modèle1). Ces analyses ont été conduites chez les 2085 familles quads, séparément pour les délétions et les duplications.

Pour la comparaison probands-population générale, nous avons utilisé un modèle de régression logistique permettant d'ajuster sur le facteur sexe (modèle2). Ces analyses ont été conduites chez les 2573 probands et 1956 individus de la population générale (N Imagen = 1744 et 212 enfants de SYS), séparément pour les délétions et les duplications.

Après correction de Bonferroni, on considère qu'une valeur $P < 0.005$ est significative.

b. Comparaison entre les apparentés au premier degré (parents, fratrie saine) et la population générale.

Nous avons évalué si la fréquence des CNVs rares $\geq 250\text{Kb}$ prédisait le fait d'être apparenté au premier degré à un enfant autiste (parents, fratrie saine) comparé à la population générale (Imagen + parent de SYS). Pour chaque variable, nous avons utilisé un modèle de régression logistique permettant d'ajuster sur le facteur sexe. Afin d'augmenter notre puissance statistique,

nous avons inclus pour la population générale, les parents de SYS (ne peuvent pas être rajoutés dans les autres analyses car pas de mesure de QI chez eux). Ces analyses ont été conduites chez les 5146 parents de la SSC, 2085 fratrie saine de la SSC et 2594 individus de la population générale (N Imagen=1744, N parents SYS = 850). Elles ont été réalisées séparément pour les délétions et les duplications.

Après correction de Bonferroni, on considère qu'une valeur $P < 0.025$ est significative.

6.2 Prédiction du risque d'autisme et de la variation du NVIQ, grâce aux caractéristiques géniques intrinsèques aux CNVs.

a. Prédiction du risque d'autisme.

Tout d'abord, nous avons évalué si parmi ces cinq variables fréquence des CNVs rares $\geq 250\text{Kb}$, leurs caractéristiques géniques (nombre de gènes, leurs scores d'intolérance [pLI, RVIS, DEL/DUP scores]), une prédisait significativement le diagnostic d'autisme en comparant probands à leur fratrie saine puis les probands à la population générale. De plus, à l'aide de l'AIC (Akaike information criterion), nous déterminons laquelle des quatre caractéristiques géniques décrite ci-dessus, expliquait le mieux le diagnostic d'autisme (proband-fratrie saine) puis (proband-population générale). Un AIC plus bas signifie un meilleur modèle statistique et donc une meilleure prédiction de la variable dépendante (risque d'autisme)

Pour la comparaison proband-fratrie saine, nous avons utilisé séparément pour les cinq variables décrites ci-dessus, un modèle de régression logistique conditionnelle pour données appariées afin de tenir compte la nature corrélée des individus appartenant à la même famille tout en ajustant sur le facteur sexe (modèle 1). Ces analyses ont été conduites chez les 2085 familles quads, séparément pour les délétions et les duplications.

Pour la comparaison probands-population générale, nous avons utilisé séparément pour les cinq variables, un modèle de régression logistique permettant d'ajuster sur le facteur sexe (modèle2). Ces analyses ont été conduites chez les 2573 probands et 1956 individus de la population générale (N Imagen = 1744 et 212 enfants de SYS)., séparément pour les délétions et les duplications.

Après correction de Bonferroni, on considère qu'une valeur $P < 0.005$ est significative.

b. Prédiction de la variation du NVIQ.

Ces analyses font suite aux résultats obtenus à l'aide d'un modèle statistique développé par une équipe au sein de notre laboratoire et qui a permis de déterminer parmi 10 caractéristiques génétiques des CNVs, lesquelles prédisaient le mieux le NVIQ au sein de la population générale (Imagen, SYS). Nous expliquerons sommairement le fonctionnement de ce modèle qui est détaillé dans l'article « Measuring and predicting the effect of copy number variants on general intelligence in 4 community-based samples. », publié en 2018 dans JAMApsy.

Ce modèle fonctionne en deux étapes:

- Ajustement du NVIQ par d'autres covariables (age, sex technologie de puce, corrélation familiale et la distance génétique) afin d'obtenir les résidus.
- Grâce à un modèle de régression multiple, où la variable dépendante est les résidus du NVIQ et les variables indépendantes sont les 10 caractéristiques génétiques des CNVs, le modèle sélectionne à l'aide du Critère d'information bayésien (BIC) et d'une procédure de sélection pas à pas (stepwise), les meilleures caractéristiques génétiques des CNVs, permettant d'expliquer la variation des résidus du NVIQ:

Ce modèle avait estimé que la meilleure caractéristique génétique pour les CNVs de type délétions, expliquant le NVIQ dans la population générale était le score d'intolérance pLI.

Par contre aucune caractéristique ne fut sélectionnée pour les CNVs de type duplication et aucun impact de ces derniers n'a pu être observé dans la population générale.

Nous utilisons sur les probands de la SSC les mêmes variables (score pLI, age, sex, technologie de puce et la distance génétique) qui ont servi pour modéliser l'impact des CNVs sur le NVIQ dans la population générale. Ainsi à l'aide d'une régression linéaire multiple, on tente de prédire le NVIQ grâce au score pLI (meilleur prédicteur génétique du NVIQ dans la population générale) tout en ajustant pour ; Age, sexe, technologie de puce et la distance génétique. Ce modèle a été appliqué pour les délétions et les duplications séparément. On ne corrige pas pour la corrélation familiale car les probands de la SSC ne sont pas apparentés.

On ajoute en second temps comme covariable dans notre modèle décrit ci-dessus, le score ADOS CSS.

Après correction de Bonferroni, on considère qu'une valeur $P < 0.01$ est significative.

6.3 Association entre les CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme (proband-population générale) en ajustant pour le NVIQ.

Nous avons utilisé deux approches différentes dans le but d'étudier indépendamment du NVIQ, la relation entre le diagnostic d'autisme (proband-populations générale) et la fréquence ou les caractéristiques géniques des CNVs rares $\geq 250\text{Kb}$.

- Dans une première approche, nous reprenons le modèle 2 développé précédemment (section 6.2) où en plus de d'ajuster sur le facteur sexe, nous avons ajusté les modèles sur le QI non verbal (NVIQ).

- Dans une deuxième approche, nous avons ré-échantillonné les probands et les individus appartenant à la population générale afin de générer un sous ensemble aléatoire composé de paires « proband-individu de la population générale » (ratio de 1:1) dont le NVIQ est semblable (maximum de 5 points de différence de QI non verbal), (Figure 12, 13, 14).

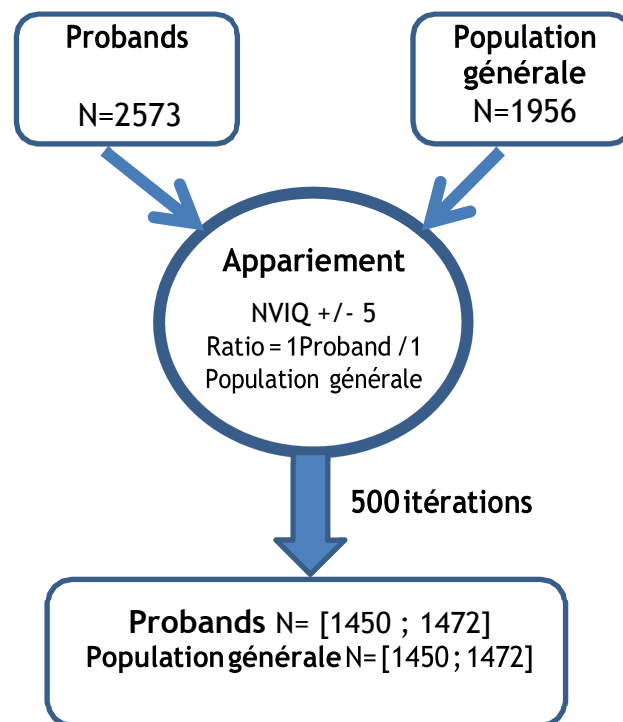


Figure 12 : processus d'appariement en fonction du NVIQ, des probands et de la population générale. Ce processus a été refait 500 fois avec pour chaque itération un nombre d'individus identique chez les probands et chez la population générale (ratio 1/1). Le nombre d'individus après chaque appariement peut varier $N = [1450 ; 1472]$ pour chaque groupe (probands, population générale) pour un N total qui peut varier $= [2900 ; 2944]$. Source : réalisation personnelle.

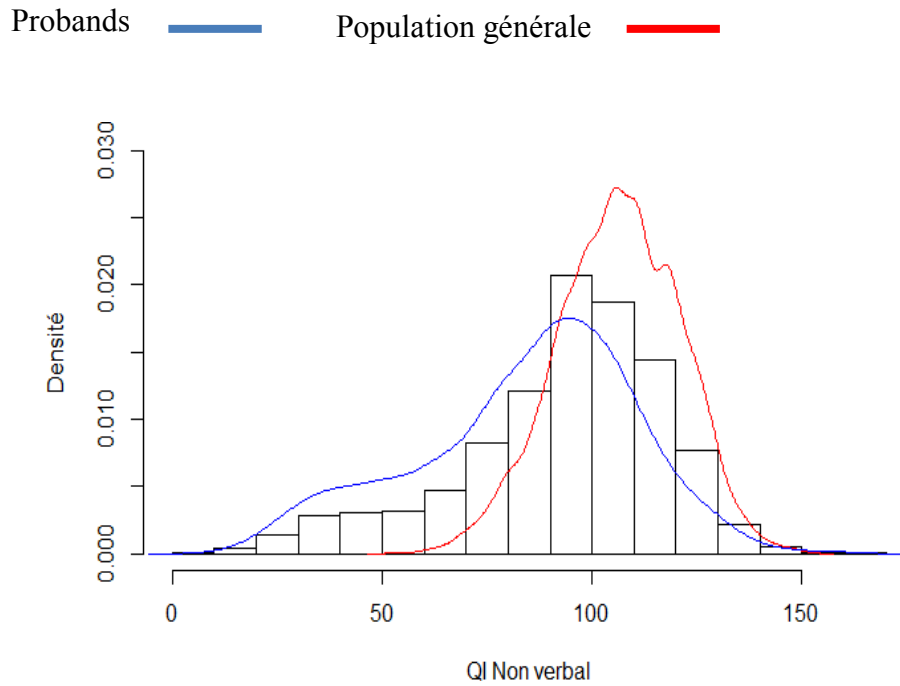


Figure 13 : Distribution du QI non verbal (NVIQ) chez les probands et la population générale avant appariement pour le NVIQ.

La ligne bleu représente la distribution des probands de la SSC tandis que la ligne rouge la population générale.

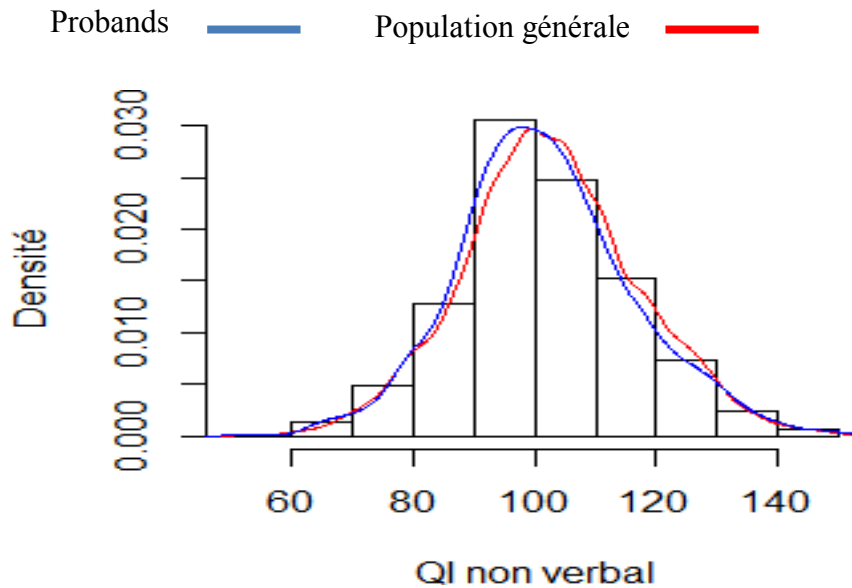


Figure 14 : Exemple de distribution du QI non verbal (NVIQ) chez les probands et la population générale après appariement pour le NVIQ. Cet exemple représente une distribution parmi 500 itérations réalisées.

En utilisant cet échantillon de données appariées, nous avons évalué l'impact de la fréquence et des caractéristiques géniques des CNVs rares $\geq 250\text{Kb}$ sur le diagnostic d'autisme (proband-population générale) à l'aide d'une régression logistique conditionnelle (pour chaque variable séparément) permettant de prendre en compte la nature corrélée de ces données appariées tout en ajustant sur le facteur sexe. Encore une fois, les modèles ont été réalisés séparément pour les délétions et les duplications. Etant donné le caractère aléatoire de cet appariement, nous avons effectué ce processus 500 fois et les résultats statistiques seront rapportés en pourcentage de fois que le test est significatif (valeur $P < 0.05$), médiane des valeurs P , médiane de l'odds-ratio, obtenus lors des 500 itérations.

Etant donné le fait que dans notre comparaison proband-population générale on ajuste pour le NVIQ, nous utilisons pour la population générale uniquement Imagen et les enfants de la SYS. Après correction de Bonferroni, on considère qu'une valeur $P < 0.005$ est significative.

Les analyses statistiques ont été réalisées à l'aide du logiciel R.3.2.4 (<https://cran.r-project.org/bin/windows/base/old/3.2.4/>). L'appariement a été réalisé grâce au package «matching» (<http://www.jstatsoft.org/v42/i07/>). La régression logistique conditionnelle a été réalisée grâce au package «survival» (<https://CRAN.R-project.org/package=survival>).

CHAPITRE 3. RESULTATS

1. Association entre la fréquence de CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme.

1.1 Comparaison entre les probands à la fratrie saine voire la population générale.

Tel que rapporté précédemment par Girirajan et al., 2011⁶², on observe que la fréquence des CNVs rares $\geq 250\text{Kb}$ est associée positivement au fait d'être autiste. En effet, les délétions sont plus fréquentes chez les probands autistes que chez la fratrie saine (valeur $P = 1,43 \text{ e-}4$, $\text{OR} = 2,79$, $\text{IC } 95\% = [1,64 ; 4,73]$) ou la population générale (valeur $P = 9,17\text{e-}06$, $\text{OR} = 2,31$, $\text{IC } 95\% = [1,61 ; 3,38]$). Les duplications sont aussi plus fréquentes chez les probands autistes comparé à la fratrie saine (valeur $P = 0,001$, $\text{OR} = 1,63$, $\text{IC } 95\% = [1,20 ; 2,20]$) et à la population générale (valeur $P = 4,26\text{e-}06$, $\text{OR} = 1,69$, $\text{IC } 95\% = [1,35 ; 2,11]$), (Tableau 2).

1.2 Comparaison entre les apparentés au premier degré (parents, fratrie saine) et la population générale.

On retrouve un excès de CNVs rares $\geq 250\text{Kb}$ de type duplication, chez les parents d'une personne autiste comparé à la population générale (Imagen + parents de SYS), (valeur $P = 2,44 \text{ e-}4$, $\text{OR} = 1,44$, $\text{IC } 95\% = [1,17 ; 1,67]$). Une tendance similaire est observée chez la fratrie saine (valeur $P = 0,02$, $\text{OR} = 1,35$, $\text{IC } 95\% = [1,02 ; 1,85]$). Par contre, aucun excès de CNVs rares $\geq 250\text{Kb}$ de type délétion n'est observé chez les apparentés aux premiers degrés à une personne autiste comparé à la population générale (parents : valeur $P = 0,30$, $\text{OR} = 1,17$, $\text{IC } 95\% = [0,87 ; 1,67]$; fratrie saine : valeur $P = 0,60$, $\text{OR} = 0,89$, $\text{IC } 95\% = [0,58 ; 1,37]$), (Tableau 3).

Variables		Probands/ Fratrie saine				Probands/Population générale			
		N=4170				N=4531			
		Probands N=2085	Fratrie saine N=2085	OR [IC 95%]	Valeur P	Probands N=2573	Population générale N=1956	OR [95% CI]	Valeur P
N CNVs rares de type délétion ≥ 250Kb	0	1989 (95,4%)	2034 (97,5%)	2,79 [1,64 ; 4,73]	1,43 e-4	2447 (95,1%)	1913 (97,7%)	2,31 [1,61 ; 3,38]	9.17e-06
	1	93 (4,5%)	49 (2,3%)			122 (4,7%)	45 (2,3%)		
	2	3 (0,14%)	1 (0,05%)			4 (0,1%)	0 (0%)		
	3	0 (0%)	1 (0,05%)			0 (0%)	0 (0%)		
N CNVs rares de type duplication ≥ 250Kb	0	1884 (90,3%)	1923 (92,2%)	1,63 [1,20 ; 2,20]	0,001	2315 (90%)	1837 (93,9%)	1,69 [1,35 ; 2,11]	4.26e-06
	1	188 (9%)	151 (7,2%)			242 (9,4%)	112 (5,7%)		
	2	13(0,6%)	10 (0,5%)			16 (0,6%)	9 (0,4%)		
	3	0 (0%)	1 (0,05%)			0 (0%)	0 (0%)		

Tableau 2: Impact du nombre CNVs rares ≥ 250Kb sur la probabilité d’avoir le diagnostic de TSA dans une comparaison avec la fratrie ou la population générale. Légende : N : nombre ; OR : odds-ratio, IC : intervalle de confiance. Les résultats sont obtenus à partir de modèles de régressions logistiques où la variable dépendante est le diagnostic et où la variable indépendante est le nombre de CNVs rares ≥ 250Kb. Les modèles ont été réalisés indépendamment pour les délétions et les duplications et tous ont été ajustés sur le facteur sexe. À noter que pour le modèle comparant les probands et leur fratrie saine, les données sont appariées et un modèle de régression logistique conditionnelle a donc été utilisé pour tenir compte du caractère corrélé de ces données.

Après correction de Bonferroni, on considère qu’une valeur P <0.005 est significative.

Variables		Fratie saine/Population générale N =4681				Parents SSC/ Population générale N=7740			
		Fratie saine N =2085	Population générale N =2594	OR [IC 95%]	valeur P	Parents SSC N =5146	Population générale N =2594	OR [IC 95%]	valeur P
N CNVs rares de type deletion ≥ 250Kb	0	2034 (97,5%)	2582 (97,8%)	0,89 [0,58 ; 1,37]	0,60	5016 (97,4%)	2582 (97,8%)	1,17 [0,87 ; 1,67]	0,30
	1	49 (2,3%)	56 (2,1%)			128 (2,5%)	56 (2,1%)		
	2	1 (0,05%)	1 (0,04%)			1 (0,02%)	1 (0,04%)		
	3	1 (0,05%)	0 (0%)			1 (0,02%)	0 (0%)		
N CNVs rares de type duplication ≥ 250Kb	0	1923 (92,2%)	2490 (94,3%)	1,35 [1,02 ; 1,85]	0,02	4731 (91,9%)	2490 (94,3%)	1,44 [1,17 ; 1,67]	2.44 e-4
	1	151 (7,2%)	138 (5,2)			391(7,6%)	138 (5,2%)		
	2	10 (0,5%)	11 (0,4%)			21 (0,4%)	11 (0,4%)		
	3	1 (0,05%)	0 (0%)			3 (0,06%)	0 (0%)		

Tableau 3 : Impact du nombre de CNVs rares ≥ 250Kb sur la probabilité d’être apparenté au premier degré à une personne autiste comparé à la population générale. Légende : N : nombre ; OR : odds-ratio, IC : intervalle de confiance. Les résultats sont obtenus à partir de modèles de régressions logistiques où la variable dépendante est le statut (apparenté au premier degré à une personne autiste – population générale) et où la variable indépendante est le nombre de CNVs rares ≥ 250Kb. Les modèles ont été réalisés indépendamment pour les délétions et les duplications et tous ont été ajustés sur le facteur sexe.

Après correction de Bonferroni, on considère qu’une valeur P <0.025 est significative.

2. Prédiction du risque d'autisme et de la variation du NVIQ, grâce aux caractéristiques géniques intrinsèques aux CNVs

2.1 Prédiction du risque d'autisme

Au-delà de cette association entre la fréquence de CNVs rares $\geq 250\text{Kb}$ et le diagnostic d'autisme, on observe qu'une augmentation du nombre de gènes inclus dans les délétions est associée à un risque d'autisme plus élevé (proband-fratrie saine : valeur $P = 1,78 \times 10^{-4}$, OR = 1,08, IC 95% = [1,04 ; 1,12] ; proband-population générale : valeur $P = 2,63 \times 10^{-5}$, OR = 1,11, IC 95% = [1,06 ; 1,17]). Cette même association est observée pour les duplications (proband-fratrie saine : valeur $P = 1,82 \times 10^{-6}$, OR = 1,09, IC 95% = [1,05 ; 1,13] ; proband-population générale : valeur $P = 9,28 \times 10^{-7}$, OR = 1,07, IC 95% = [1,04 ; 1,10]), (Tableau 4).

En faisant l'hypothèse d'un effet additif des gènes dans les CNVs, nous avons ensuite établi différents scores d'intolérance des CNVs (score pLI, score RVIS, score Del/Dup) obtenus en faisant la somme des scores d'intolérance de tous les gènes inclus en totalité dans les CNVs. Il existe pour les délétions, une association positive entre le score pLI et le risque d'autisme, que ce soit pour la comparaison proband -fratrie saine (valeur $P = 4,34 \times 10^{-5}$, OR = 1,40, IC 95% = [1,19 ; 1,64]) ou pour la comparaison proband-population générale (valeur $P = 4,12 \times 10^{-5}$, OR = 1,49, IC 95% = [1,26 ; 1,86]). L'utilisation des autres scores d'intolérance (score RVIS, score Del/Dup) donne le même niveau d'association. Pour les Duplication, le score pLI semble encore plus fortement associé au TSA que ce qui est observé pour les délétions (AIC plus bas), ce qui est à première vue paradoxal pour un score d'intolérance. Ce résultat est observé, que ce soit pour la comparaison proband -fratrie saine (valeur $P = 9,39 \times 10^{-7}$, OR = 1,45, IC 95% = [1,25 ; 1,69]) ou pour la comparaison proband-population générale (valeur $P = 9,40 \times 10^{-7}$, OR = 1,37, IC 95% = [1,22 ; 1,57]). L'utilisation des autres scores d'intolérance à la mutation (score RVIS, score Del/Dup) donne le même niveau d'association (Tableau 4).

Parmi les différentes caractéristiques géniques du CNV (nombre de gènes, leurs scores d'intolérance [pLI, RVIS, DEL/DUP scores]), le meilleur prédicteur du diagnostic d'autisme est le score pLI en se basant sur l'AIC (plus il est bas plus le modèle est juste). Ceci est vrai pour les

CNVs de type délétion mais aussi pour les CNVs de type duplication, ce qui peut sembler à première vue paradoxale.

Type	Variables	Probands/Fratrie saine N probands=2085 ; N fratrie saine=2085					Probands/Population générale N probands=2573; N population générale=1956				
		N probands avec score ≥ 0	N fratrie saine avec score ≥ 0	AIC	OR [IC 95%]	Valeur P	N probands avec score ≥ 0	N population générale avec score ≥ 0	AIC	OR [IC 95%]	Valeur P
		CNVs rares de type deletion ≥ 250Kb	N gènes	96	51	2125,93	1,08 [1,04 ; 1,12]	1,78 e-4	177	45	5392,3
score pLI	72		35	2120,96	1,40 [1,19 ; 1,64]	4,34 e-05	128	30	5391,2	1,49 [1,26 ; 1,86]	4,12 e-05
score Del	58		25	2127,59	1,07 [1,03 ; 1,13]	0,002	102	20	5398,5	1,09 [1,04 ; 1,16]	0,001
score RVIS	72		34	2123,77	1,002 [1,001 ; 1,003]	1e-04	127	31	5394,8	1,002 [1,001 ; 1,003]	4,84 e-05
CNVs rares de type duplication ≥ 250Kb	N gènes	201	162	2116,15	1,09 [1,05 ; 1,13]	1,82e-06	420	121	5390,4	1,07 [1,04 ; 1,10]	9,28 e-07
	score pLI	161	122	2108,934	1,45 [1,25 ; 1,69]	9,39e-07	326	83	5383,6	1,37 [1,22 ; 1,57]	9,40 e-07
	score Dup	142	96	2122,399	1,08 [1,04 ; 1,12]	7,38e-05	280	68	5394,5	1,07 [1,04 ; 1,10]	1,22 e-05
	score RVIS	162	122	2110,53	1,002 [1,001 ; 1,003]	5,82 e-07	328	81	5386,4	1,002 [1,001 ; 1,003]	7,50 e-07

Tableau 4: Association entre le nombre de gènes, les scores d'intolérance (pLI, Del/Dup, RVIS) des CNVs rares ≥ 250Kb et la probabilité d'avoir le diagnostic de TSA dans une comparaison avec la fratrie ou la population générale. Légende : N : nombre ; OR : odds-ratio, IC : intervalle de confiance. Les résultats sont obtenus à partir de modèles de régressions logistiques où la variable dépendante est le diagnostic et où la variable indépendante est soit le nombre de gènes présent dans les CNVs rares ≥ 250Kb soit un leurs scores d'intolérance (pLI, Del/Dup, RVIS). Les modèles ont été réalisés indépendamment pour les délétions et les duplications et tous ont été ajustés sur le facteur sexe. A noter que pour le modèle comparant les probands et leur fratrie, les données sont appariées et un modèle de régression logistique conditionnelle a donc été utilisé pour tenir compte du caractère corrélé de ces données.

Après correction de Bonferroni, on considère qu'une valeur P < 0.005 est significative.

2.2 Prédiction de la variation du NVIQ.

Tel que démontré dans la population générale (Huguet/Schramm et al.,2018, en cours de publication sur JAMApsy), le meilleur prédicteur génétique du NVIQ pour les CNVs de type délétion est le score d'intolérance pLI. Ils rapportent qu'une variation de 1 point de ce score diminuerait le NVIQ par 2,74 (IC 95% = [-4,07 ; -1,41]) pour les CNVs de type délétion. En appliquant ce même modèle sur les probands de la SSC, on retrouve pour les CNVs de type délétion qu'une augmentation du score pLI (meilleur prédicteur génétique du NVIQ dans la population générale) de 1 point diminue le NVIQ de 2.65 points (IC 95% = [- 3,71 ; -1,59], valeur P = 1,22 e-06), (Figure 15). Contrairement à la population générale où on ne retrouve aucun impact des CNVs de type duplication sur le NVIQ (Huguet/Schramm et al.,2018, en cours de publication sur JAMApsy), une augmentation de 1 point du score pLI s'accompagne d'une diminution de 1.22 point du NVIQ chez les probands de la SSC (IC 95% = [-2,03 ; -0,41], valeur P = 0,003), (Figure 16).

Tel que décrit précédemment dans la littérature, il existe une corrélation négative entre la sévérité des symptômes autistiques et le QI¹¹⁹. On estime chez les probands de la SSC que la corrélation entre le NVIQ et la sévérité des symptômes autistiques telle que mesurés par l'ADOS css est de -0,12 (IC 95% = [-0,19 ; -0,04], valeur p = 0,001), Afin d'estimer si la relation observée entre le NVIQ et les CNVs (délétion, duplication) dans cette population autiste n'était pas uniquement due à la variation de la sévérité des traits autistiques (ADOS css), nous avons corrigé en second temps notre modèle (appliqué précédemment) pour l'ADOS css. On retrouve à nouveau pour les CNVs de type délétion, qu'une augmentation du score pLI de 1 point équivaut toujours à une diminution de 2.65 points Du NVIQ (IC 95% = [-3,71 ; 1,59], valeur P = 1.25e-06). En effectuant cette même correction pour les CNVs de type duplication, on note qu'une augmentation du pLI score de 1 point diminue le NVIQ de 1.08 point (IC 95% = [-1,90 ; -0,28], valeur P = 0,008), (Tableau 5).

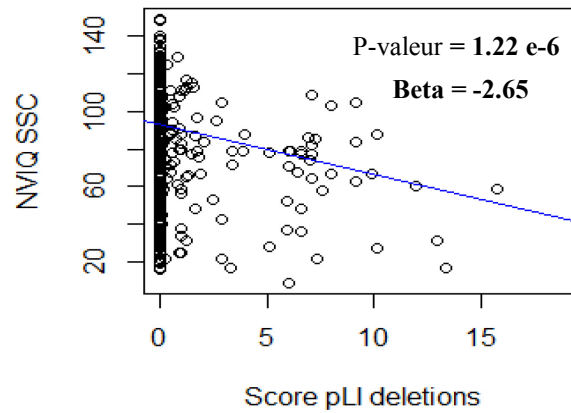


Figure 15 : association entre le score pLI des CNV de type délétion et le NVIQ chez les probands de la SSC. Nous avons utilisé une régression linéaire multiple pour prédire l'impact du score pLI des CNVs de type délétion sur le NVIQ tout en corrigeant pour l'âge, le sexe, la technologie et la distance génétique. Les variables confondantes ont été centrées. La ligne bleue représente la droite de régression. Après correction de Bonferroni, on considère qu'une valeur $P < 0.01$ est significative.

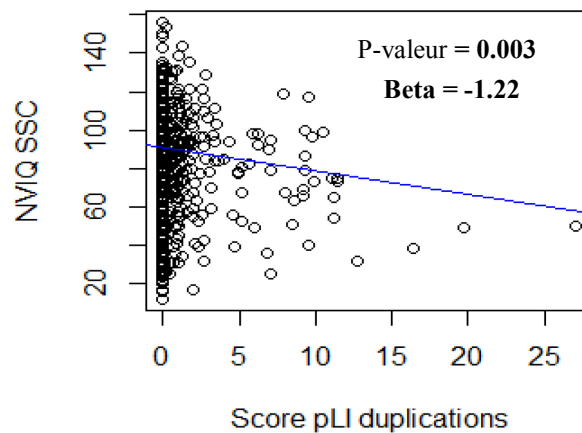


Figure 16 : association entre le score pLI des CNV de type duplication et le NVIQ chez les probands de la SSC. Nous avons utilisé une régression linéaire multiple pour prédire l'impact du score pLI des CNV de type duplication sur le NVIQ tout en corrigeant pour l'âge, le sexe, la technologie et la distance génétique. Les variables confondantes ont été centrées. La ligne bleue représente la droite de régression. Après correction de Bonferroni, on considère qu'une valeur $P < 0.01$ est significative.

Variables	NVIQ		NVIQ Ajusté pour ADOS css	
	OR [IC 95%]	Valeur P	OR [Q1 ; Q3]	Valeur P [Q1 ; Q3]
Score pLI pour les délétions*	-2,65 [-3,71 ; -1,59]	1,22 e-06	-2,65 [-3,71 ; -1,59]	1,25e-06
Score pLI pour les duplications*	-1,08 [-1,90 ; -0,28]	5,43 e-4	-1,08 [-1,90 ; -0,28]	0,008

Tableau 5 : association entre le score pLI des CNVs (duplication et délétion, séparément) et le NVIQ chez les probands de la SSC. Nous avons utilisé une régression linéaire multiple pour prédire l'impact du score pLI des CNVs (délétions, duplications séparément) sur le NVIQ tout en corrigeant pour l'âge, le sexe, la technologie et la distance génétique. Nous corrigeons ultérieurement pour le score ADOS css.

Après correction de Bonferroni, on considère qu'une valeur $P < 0.01$ est significative.

3. Association entre les CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme (proband-population générale) en ajustant pour le NVIQ.

Nous avons investigué si l'association entre les CNVs et le diagnostic d'autisme était entraînée par une diminution du NVIQ. A l'aide d'un modèle de régression logistique ou on corrige pour le NVIQ on retrouve que ni la fréquence de CNVs rares $\geq 250\text{Kb}$ de type délétion (valeur $P = 0,06$, OR = 1,51, IC 95% = [0,99 ; 2,34]), ni leur nombre de gènes (valeur $P = 0,03$, OR = 1,06, IC 95% = [1,02 ; 1,12]) ou leurs scores d'intolérance (valeur $P = 0,04$, OR = 1,25, IC 95% = [1,07 ; 1,59] pour le score pLI, valeur $P = 0,04$, OR = 1,001, IC 95% = [1,0003 ; 1,002] pour le score RVIS, valeur $P = 0,11$, OR = 1,04, IC 95% = [1,002 ; 1,11] pour le score Del) n'étaient associés significativement au diagnostic d'autisme (Tableau 6, Tableau 7). Contrairement au CNVs de type délétion, quand le même modèle est appliqué aux CNVs de type duplications, on retrouve une association significative entre la fréquence de CNV rares $\geq 250\text{Kb}$ (valeur $P = 5,43 \text{ e-}4$, OR = 1,56, IC 95% = [1,22 ; 2,02]), leur nombre de gènes (valeur $P = 1,50 \text{ e-}04$, OR = 1,06, IC 95% = [1,03 ; 1,09]) ou leurs scores d'intolérance (valeur $P = 6,34 \text{ e-}05$, OR = 1,31, IC 95% = [1,15 ; 1,51] pour le score pLI, valeur $P = 1,10 \text{ e-}04$, OR = 1,002, IC 95% = [1,0008 ; 1,002] pour le score RVIS, valeur $P = 0,001$, OR

= 1,05, IC 95% = [1,02 ; 1,09] pour le score Dup) et le diagnostic d'autisme (Tableau 6, Tableau 7).

Afin de confirmer que notre modèle logistique corrigeait adéquatement le NVIQ, nous avons ré-échantillonné notre groupe de probands autistes et de population générale afin d'obtenir une distribution de NVIQ identique. On retrouve à nouveau à l'aide d'une régression logistique conditionnelle, sur 500 itérations que ni la fréquence des CNVs rares $\geq 250\text{Kb}$ de type délétion, ni leur nombre de gènes ou leurs scores d'intolérance ne prédissent adéquatement le diagnostic d'autisme (% de significativité de la valeur P sur 500 itérations $<50\%$ pour toutes ces variables). Comparativement et tel qu'observé précédemment, on note que la fréquence des CNV de type duplication, leur nombre de gènes ou leurs scores d'intolérance prédisent significativement le diagnostic d'autisme (% de significativité de la valeur P sur 500 itérations $>99\%$ pour toutes ces variables), (Tableau 6, Tableau 7).

A l'aide du AIC, on retrouve à nouveau que le score pLI est le meilleur prédicteur du diagnostic d'autisme, indépendamment du NVIQ. Cette observation est consistante pour les CNVs de type délétion et duplications.

Variables		Probands/Population générale				Probands/Population générale		
		Ajustés pour le NVIQ				Appariés pour le NVIQ N probands = [1450 ; 1472] N population générale = [1450 ; 1472]		
		Probands N =2573	Population générale N =1956	OR [IC 95%]	Valeur P	OR [Q1 ; Q3]	Valeur P [Q1 ; Q3]	% significativité*
N CNVs rares de type deletion ≥ 250Kb	0	2447 (95,1%)	1913 (97,7%)	1,51 [0,99 ; 2,34]	0,06	1,27 [1,18 ; 1,27]	0,19 [0,12 ; 0,26]	5,4%
	1	122 (4,7%)	45 (2,3%)					
	2	4 (0,1%)	0 (0%)					
N CNVs rares de type duplications ≥ 250Kb	0	2315 (90%)	221837 (93,9%)	1,56 [1,22 ; 2,02]	5,43 e-4	1,46 [1,43 ; 1,49]	0,006 [0,002 ; 0,01]	99,8%
	1	242 (9,4%)	112 (5,7%)					
	2	16 (0,6%)	9 (0,4%)					

Tableau 6 : Impact du nombre de CNVs rares ≥ 250Kb sur la probabilité d’avoir le diagnostic de TSA dans une comparaison

avec la population générale, après ajustement (modèle 1) puis appariement (modèle 2) pour le NVIQ. Légende : N : nombre ; OR : odds-ratio, IC : intervalle de confiance, Q1 : premier quartile, Q3 : deuxième quartile. Les résultats sont obtenus à partir de modèles de régressions logistiques où la variable dépendante est le diagnostic et où la variable indépendante est le nombre de CNVs rares ≥ 250Kb. Les modèles ont été réalisés indépendamment pour les délétions et les duplications et tous ont été ajustés sur le facteur sexe. Dans notre modèle 1, on a ajusté notre régression logistique pour le NVIQ tandis que pour le modèle 2, nous avons utilisé une régression logistique conditionnelle pour tenir compte du caractère corrélé de ces données (appariement). Etant donné le caractère aléatoire de cet appariement, nous avons répété ce processus 500 fois et les résultats statistiques ont été rapportés en pourcentage de fois que le test est significatif (valeur P <0.05), médiane des valeurs P avec premier et troisième quartile, médiane de l’odds-ratio avec premier et troisième quartile, obtenus lors des 500 itérations.

*pourcentage de fois ou la valeur p est inférieure à 0,05 sur 500 itérations.

Pour le modèle ou on ajuste pour le NVIQ, on considère qu’une valeur P <0.005 est significative (correction de Benferroni).

Type	Variables	Probands/Population générale Ajustés pour le NVIQ N probands = 2573 N population générale = 1956				Probands/Population générale Appariés pour le NVIQ N probands = [1450 ; 1472] N population générale = [1450 ; 1472]			
		N probands avec score ≥ 0	N population générale avec score ≥ 0	AIC	OR [IC 95%]	valeur P	OR [Q1 ; Q3]	valeur P [Q1 ; Q3]	% Significativité
CNV's rares de type deletion $\geq 250Kb$	N gènes	177	45	4469,1	1,06 [1,02 ; 1,12]	0,03	1,04 [1 ; 1,09]	0,05 [0,01 ; 0,10]	45%
	score pLI	128	30	4469	1,25 [1,07 ; 1,59]	0,04	1,18 [0,99 ; 1,42]	0,05 [0,01 ; 0,09]	50%
	score Del	102	20	4473,4	1,04 [1,002 ; 1,11]	0,11	1,03 [0,98 ; 1,09]	0,18 [0,13 ; 0,24]	0%
	score RVIS	127	31	4470,7	1,001 [1,0003 ; 1,002]	0,04	1,008 [0,99 ; 1,02]	0,08 [0,04 ; 0,11]	20%
CNV's rares de type duplication $\geq 250Kb$	N gènes	420	121	4460,9	1,06 [1,03 ; 1,09]	1,50 e-04	1,05 [1,01 ; 1,08]	0,003 [0,001 ; 0,007]	100%
	score pLI	326	83	4455,9	1,31 [1,15 ; 1,51]	6,34 e-05	1,25 [1,098 ; 1,44]	0,002 [0,0001 ; 0,004]	100%
	score Dup	280	68	4463,8	1,05 [1,02 ; 1,09]	0,001	1,04 [1,01 ; 1,08]	0,008 [0,006 ; 0,01]	100%
	score RVIS	328	81	4459	1,002 [1,0008 ; 1,002]	1,10 e-04	1,001 [1 ; 1,002]	0,005 [0,003 ; 0,007]	100%

Tableau 7 : Association entre le nombre de gènes, les scores d'intolérance (pLI, Del/Dup, RVIS) des CNVs rares $\geq 250Kb$ et la probabilité d'avoir le diagnostic de TSA dans une comparaison avec la population générale, après ajustement (modèle 1) puis appariement (modèle 2) pour le NVIQ. Légende : N : nombre ; AIC : critère d'information akaïke ; OR : odds-ratio, IC : intervalle de confiance, Q1 : premier quartile, Q3 : troisième quartile. Les résultats sont obtenus à partir de modèles de régressions logistiques où la variable dépendante est le diagnostic et où la variable indépendante est soit le nombre de gènes présent dans les CNVs rares $\geq 250Kb$ soit un leurs scores d'intolérance (pLI, Del/Dup, RVIS). Les modèles ont été réalisés indépendamment pour les délétions et les duplications et tous, ont été ajustés sur le facteur sexe. Dans notre modèle 1, on a ajusté notre régression logistique pour le NVIQ tandis que pour le modèle 2, nous avons utilisé une régression logistique conditionnelle pour tenir compte du caractère corrélé de ces données (appariement). Etant donné le caractère aléatoire de cet appariement, nous avons répété ce processus 500 fois et les résultats statistiques ont été rapportés en pourcentage de fois que le test est significatif avec une valeur $P < 0.05$ (% Significativité*), médiane des valeurs P avec premier et troisième quartile, médiane de l'odds-ratio avec premier et troisième quartile, obtenus lors des 500 itérations. Pour le modèle 1, on considère qu'une valeur $P < 0.005$ est significative (correction de Benferroni).

CHAPITRE 4. DISCUSSION

1. Association entre la fréquence de CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme.

Nous faisons l'hypothèse qu'il y'a un excès de CNVs rares $\geq 250\text{Kb}$ chez les probands autistes ainsi que chez leurs apparentés aux premiers degrés (fratrie saine, parents) comparé à la population générale. Nous retrouvons effectivement une association positive entre la fréquence de CNVs et le risque d'autisme. Cette association tend à être plus importante pour les délétions que pour les duplications (bien que les intervalles de confiance restent chevauchants). Ces résultats sont observables pour nos deux groupes contrôles (population générale ou fratrie saine) et sont concordants avec ce qui a été montré dans la littérature⁶². Effectivement, Girirajan et al., 2011⁶² ont retrouvé une contribution des CNVs larges $\geq 500\text{Kb}$ dans les TSA. Notre étude montre pour la première qu'il existe un excès de CNVs larges $\geq 250\text{Kb}$ chez les apparentés aux premiers degrés sains (fratrie, parents) comparé à la population générale. Cet excès concerne uniquement les duplications. L'absence d'excès de délétions pourrait suggérer que ces derniers seraient trop délétères pour être présent chez des individus non autistes et qu'ils surviendraient essentiellement *de novo*. Nous confirmons ainsi notre première hypothèse.

2. Prédiction du risque d'autisme et de la diminution du NVIQ, grâce aux caractéristiques géniques intrinsèques aux CNVs

Notre deuxième hypothèse est que les caractéristiques géniques communes aux CNVs, permettent de prédire le risque d'autisme voire la diminution du NVIQ. Notre étude démontre que parmi les caractéristiques géniques des CNVs (nombre de gènes, leurs scores d'intolérance [pLI, RVIS, DEL/DUP scores]), celle qui explique le mieux le risque d'autisme est le score pLI. Les modèles suggèrent que les risques associés aux scores de chaque gène inclus dans le CNV, s'additionnent de façon linéaire sans interaction.

Cette observation est valable aussi bien pour les délétions que les duplications. Ceci peut sembler paradoxal car le pLI est un score d'intolérance à l'haploinsuffisance alors que les duplications entraînent presque toujours une surexpression des gènes impliqués.

Nos modèles pourraient servir à prédire dans le futur le risque d'autisme associé aux CNVs identifiés chez les patients. Même si certains gènes majeurs sont associés à des risques très importants d'autisme (formes syndromiques), ces gènes ne représentent pas plus de 5% des

TSA⁵⁴. Nos données suggèrent que le risque d'autisme associé à un CNV résulte d'une somme de petits risques associés aux gènes contenus dans ce dernier. Ceci est en faveur d'un modèle « omnigenic »¹²⁰ pour les TSA qui implique qu'une fraction importante du génome codant est associée au risque de TSA lorsqu'elle est mutée.

Un modèle similaire (basé sur l'effet additif) a été développé dans la population générale et a permis grâce au score d'intolérance à la mutation pLI, de prédire adéquatement la diminution du NVIQ associée aux CNVs de type délétion. À l'aide de ce modèle, nous retrouvons dans la population autiste que les CNVs de type délétion sont aussi associés à une diminution du NVIQ. De plus, leurs impacts chez les autistes semblent similaires à ce qui a été rapporté dans la population générale. Ce résultat suggère que la relation entre les délétions et la baisse du NVIQ n'est pas influencée par le diagnostic d'autisme. Contrairement à ce qui a été observé dans la population générale où les CNVs de type duplication n'impactaient pas le NVIQ, ce même modèle appliqué aux duplications retrouvés chez les autistes, retrouve que ces dernières diminuent le NVIQ. Cette diminution est néanmoins > 2 fois moins importante que ce qui est observé pour les délétions. Étant donné que les autistes ont tendance à moins bien scorer dans les tests de QI, nous avons tenté d'explorer si l'effet des CNVs sur la diminution du NVIQ n'était pas du uniquement à une augmentation de la sévérité des symptômes autistique. En corrigeant notre modèle pour l'ADOS css, nous retrouvons que la relation entre les délétions et le NVIQ reste sensiblement identique. Pour les duplications, même si une légère diminution de l'impact est observée, ces dernières restent associées à une diminution du NVIQ. Nous validons ainsi notre hypothèse en montrant qu'un modèle prédictif basé sur le score d'intolérance pLI, nous permet d'estimer non seulement l'impact des CNVs de type délétion et duplication sur le risque d'autisme mais aussi sur la variation du NVIQ.

3. Association entre les CNVs rares $\geq 250\text{Kb}$ et le risque d'autisme (proband-population générale) en ajustant pour le NVIQ

Nous posons l'hypothèse que l'association entre les CNVs rare $\geq 250\text{Kb}$ est partiellement expliquée par une diminution du NVIQ. Après correction et appariement

pour le NVIQ, nous ne retrouvons pas d'association significative (après correction de Bonferroni) entre les CNVs de type délétion (fréquence, caractéristiques géniques) et le diagnostic d'autisme. On retrouve par contre qu'une tendance persiste. Ces résultats sont en accord avec ce qui a été montré par Girirajan et al., 2011, qui sans différencier les CNVs par type (délétion, duplication), retrouvaient que les CNVs rares et larges ($\geq 500\text{Kb}$) étaient enrichis chez les individus souffrant de déficience intellectuelle comparé aux individus autistes ⁶². On retrouve à l'opposé, que les duplications restent aussi fortement associées qu'avant au risque d'autisme. Ces résultats suggèrent que les délétions sont associées majoritairement à une baisse du NVIQ qu'au risque spécifique d'autisme. Les duplications seraient au contraire plus spécifiques de l'autisme et leur effet passerait non pas par une diminution du NVIQ mais par une atteinte de domaines cliniques caractéristiques de l'autisme.

4. Limite de l'étude.

Dans cette étude nous n'explorons pas l'impact des CNVs communs. Même si la plupart des études se concentrent sur les CNVs rare, on sait que les CNVs communs jouent un rôle important dans l'autisme. Leurs effets étant faibles individuellement, l'étude de ces derniers nécessite une grande taille d'échantillon. Même si la modélisation de l'impact des délétions a été déjà répliquée dans la population générale. Néanmoins certaines limites relatives aux cohortes utilisées doivent être mentionnées :

Biais au sein de la SSC :

La SSC est une cohorte simplex avec seulement un enfant atteint d'autisme au sein de la famille. Cela favoriserait des formes sporadiques d'autisme avec une plus grande charge des variants *denovo*. Étant donnée cette particularité, nos résultats doivent être répliqués sur une cohorte multiplex pour s'assurer que ce que nous observons n'est pas induit par un excès de variants *denovo*. *Ce qui nous rassure quant à cette limite est que :*

- finalement moins de 5% des probands sont porteurs de CNV de novo. Effectivement, ~ 95% des probands ont uniquement des CNVs hérités des parents.
- Même si les parents n'ont pas de diagnostic d'autisme, on sait à travers les études que ces derniers présentent des traits autistiques plus importants que ce qui est observé dans la

population générale⁸⁰. De plus, nos résultats montrent clairement un excès de CNVs délétères chez les parents d'une personne autiste comparé à la population générale. Ces données indiquent que même au sein d'une structure simplex, les variants hérités contribuent fortement au TSA.

- Dans toutes nos analyses, nous ne regroupons en aucun cas les probands, les parents et la fratrie saine. Ainsi chacun de ces 3 groupes est comparé séparément à la population générale.

Biais au sein de la population générale

Une des limites qu'on se doit de considérer est le fait que nous regroupons en un seul groupe les parents de la cohorte Saguenay Youth Study (SYS) avec les adolescents de la cohorte Imagen et cela pour notre analyse comparant ces derniers aux parents de la SSC. Effectivement, même si les deux groupes (parents SYS, Imagen) appartiennent à une cohorte de population générale, on a assez de recul sur les parents de la SYS pour dire que ces derniers ont tous des enfants non autistes tandis que pour Imagen, on a un recrutement d'adolescent qui s'est fait sur une seule génération impliquant que ces derniers pourraient avoir dans le futur des enfants autistes et seraient ainsi comparable aux parents de la SSC. Ce qui nous rassure quant à cette limite est qu'étant donné que la prévalence de l'autisme est de ~1%, on s'attend à ce que sur ~ 1700 adolescents de la population générale qui auraient dans le futur 1 enfant, seulement 17 seraient diagnostiqués d'autisme tombant ainsi dans le même cas de figure que les parents de la SSC.

Biais dans la comparaison SSC à la population générale

- Génétique

Étant donnée la différence de technologies de génotypage utilisées dans au sein de la SSC comparé à la population générale, nous pourrions observer une différence qualitative voire quantitative en terme de CNVs. Pour essayer de remédier à ça, nous avons utilisé dans nos analyses comparant la charge mutationnelle entre ces deux cohortes, exclusivement les CNVs dont la taille est assez importante pour être observés de manière presque identique peu importe la technologie.

- Clinique

Les échelles utilisées pour calculer le QI diffèrent entre la SSC et la population générale. Étant

donné qu'on effectue un appariement sur le QI entre ces deux cohortes, un biais pourrait être induit. Cette limite doit être prise en considération dans l'interprétation des résultats obtenus. De plus, des analyses de sensibilité complémentaires doivent être fait dans le future pour déterminer si nos analyses portant sur le QI restent stables et valables peu importe la gamme de QI étudiée (QI bas, moyens ou élevés).

5. Conclusion et perspective

Les études d'association entre les CNVs et des traits phénotypiques complexes tels que l'autisme sont limitées par l'extrême rareté de ces variants. Notre approche permettrait de modéliser grâce au score d'intolérance à la mutation, l'impact de ces CNVs sur le risque d'autisme mais aussi sur la variation du NVIQ. Peu d'études ont essayé de dissocier l'impact des CNVs sur le risque d'autisme, de leurs effets sur l'altération de déficience intellectuelle. Nos résultats montrent une implication claire des CNVs qu'ils soient de type délétion ou duplication dans l'autisme, Cette implication semble, par contre, plus due à une diminution du NVIQ pour les délétions qu'à un risque spécifique d'autisme. Les duplications sont, à l'opposé, associées à un risque d'autisme plus élevé et cela indépendamment du NVIQ, suggérant que l'effet de ces derniers passerait par une atteinte plus spécifique des domaines cliniques caractéristiques de l'autisme.

Nous prévoyons de répliquer ces résultats dans une autre cohorte d'autistes ainsi que sur des CNVs pathogéniques, récurrents dont l'impact sur le risque d'autisme et sur la variation du NVIQ est connu. Etant donné que les CNV communs jouent un rôle considérable dans les TSA nous pensons qu'il serait intéressant d'évaluer leurs impacts dans nos prochaines analyses.

Finalement, Cette approche pourrait servir à prédire l'impact des CNVs sur des traits importants dans l'autisme et qui se trouvent être hautement héritables. A titre d'exemple, l'SRS (social responsiveness scale), une échelle qui évalue le niveau d'atteinte sociale au sein des autistes, a une héritabilité de 0,76. (Constantino & Todd, 2000). Cette dernière représentera une variable de choix pour de futures modélisations au sein de notre laboratoire.

Bibliographie

1. Down JL. Lettsomian Lectures on Some of the Mental Affections of Childhood and Youth. *British Medical Journal*. 1887;1(1362):256-259.
2. Bleuler E, Ey H. *Dementia praecox oder Gruppe der Schizophrenien*. Théraplix; 1969.
3. Kanner L. Autistic disturbances of affective contact. *Nervous Child*. 1943;2:217-250.
4. Asperger H. Die „Autistischen Psychopathen“ im Kindesalter. *Archiv für Psychiatrie und Nervenkrankheiten*. 1944;117(1):76-136.
5. Wing L. Asperger's syndrome: a clinical account. *Psychol Med*. 1981;11(1):115-129.
6. American psychiatric a, Spitzer RL. *Diagnostic and statistical manual of mental disorders. DSM-III. 3rd. ed.* Washington (USA): American Psychiatric Association; 1981.
7. DSM-I Diagnostic and Statistical Manual: Mental Disorders. 1952.
8. *DSM-II : Diagnostic and statistical manual of mental disorders*. Washington, D.C.: American Psychiatric Association; 1968.
9. Copyright. *Diagnostic and Statistical Manual of Mental Disorders*: American Psychiatric Association; 2013.
10. Manolova H, Achkova M. *Diagnosis "Autism" – from Kanner and Asperger to DSM-5*. Vol 22014.
11. Gillet P. *Neuropsychologie de l'autisme chez l'enfant*. De Boeck Supérieur; 2013.
12. *Diagnostic and statistical manual of mental disorders : DSM-III-R*. Washington, DC: American Psychiatric Association; 1987.
13. Bell CC. Dsm-iv: Diagnostic and statistical manual of mental disorders. *JAMA*. 1994;272(10):828-829.
14. American Psychiatric A, American Psychiatric A, Task Force on D-I. *Diagnostic and statistical manual of mental disorders : DSM-IV-TR*. Washington, DC: American Psychiatric Association; 2000.
15. Belardinelli C, Raza M. *Comorbid Behavioral Problems and Psychiatric Disorders in Autism Spectrum Disorders*. Vol 022016.
16. Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *Morbidity and mortality weekly report. Surveillance summaries (Washington, D.C. : 2002)*. 2014;63(2):1-21.
17. van der Meer JMJ, Oerlemans AM, van Steijn DJ, et al. Are Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder Different Manifestations of One Overarching Disorder? Cognitive and Symptom Evidence From a Clinical and Population-Based Sample. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2012;51(11):1160-1172.e1163.
18. Ghaziuddin M, Ghaziuddin N, Greden J. Depression in persons with autism: implications for research and clinical care. *J Autism Dev Disord*. 2002;32(4):299-306.
19. Sukhodolsky DG, Scahill L, Gadow KD, et al. Parent-rated anxiety symptoms in children with pervasive developmental disorders: frequency and association with core autism symptoms and cognitive functioning. *J Abnorm Child Psychol*. 2008;36(1):117-128.
20. van Steensel FJ, Bogels SM, Perrin S. Anxiety disorders in children and adolescents with autistic spectrum disorders: a meta-analysis. *Clinical child and family psychology review*. 2011;14(3):302-317.
21. Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, Baird G. Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2008;47(8):921-929.

22. Abramson RK, Ravan SA, Wright HH, et al. The relationship between restrictive and repetitive behaviors in individuals with autism and obsessive compulsive symptoms in parents. *Child psychiatry and human development*. 2005;36(2):155-165.
23. Mehtar M, Mukaddes NM. Posttraumatic Stress Disorder in individuals with diagnosis of Autistic Spectrum Disorders. *Research in Autism Spectrum Disorders*. 2011;5(1):539-546.
24. Amiet C, Gourfinkel-An I, Bouzamondo A, et al. Epilepsy in autism is associated with intellectual disability and gender: evidence from a meta-analysis. *Biological psychiatry*. 2008;64(7):577-582.
25. Steffenburg S, Gillberg C. Autism spectrum disorders in children with active epilepsy and learning disability: comorbidity, pre- and perinatal background, and seizure characteristics. *Developmental Medicine & Child Neurology*. 2003;45(11):724-730.
26. Bolton PF, Carcani-Rathwell I, Hutton J, Goode S, Howlin P, Rutter M. Epilepsy in autism: features and correlates. *The British journal of psychiatry : the journal of mental science*. 2011;198(4):289-294.
27. Turk J, Bax M, Williams C, Amin P, Eriksson M, Gillberg C. Autism spectrum disorder in children with and without epilepsy: impact on social functioning and communication. *Acta Pædiatrica*. 2009;98(4):675-681.
28. Matson JL, Neal D, Hess JA, Mahan S, Fodstad JC. The effect of seizure disorder on symptom presentation in atypically developing children and children with autism spectrum disorders based on the BDI-2. *Developmental Neurorehabilitation*. 2010;13(5):310-314.
29. Zerbo O, Leong A, Barcellos L, Bernal P, Fireman B, Croen LA. Immune mediated conditions in autism spectrum disorders. *Brain, behavior, and immunity*. 2015;46:232-236.
30. Chen MH, Su TP, Chen YS, et al. Is atopy in early childhood a risk factor for ADHD and ASD? a longitudinal study. *Journal of psychosomatic research*. 2014;77(4):316-321.
31. Puig-Alcaraz C, Fuentes-Albero M, Calderon J, Garrote D, Cauli O. Increased homocysteine levels correlate with the communication deficit in children with autism spectrum disorder. *Psychiatry research*. 2015;229(3):1031-1037.
32. Coury DL, Ashwood P, Fasano A, et al. Gastrointestinal Conditions in Children With Autism Spectrum Disorder: Developing a Research Agenda. *Pediatrics*. 2012;130(Supplement 2):S160-S168.
33. Black C, Kaye JA, Jick H. Relation of childhood gastrointestinal disorders to autism: nested case-control study using data from the UK General Practice Research Database. *BMJ*. 2002;325(7361):419-421.
34. Ibrahim SH, Voigt RG, Katusic SK, Weaver AL, Barbaresi WJ. Incidence of Gastrointestinal Symptoms in Children With Autism: A Population-Based Study. *Pediatrics*. 2009;124(2):680-686.
35. Mouridsen SE, Rich B, Isager T. A longitudinal study of gastrointestinal diseases in individuals diagnosed with infantile autism as children. *Child: care, health and development*. 2010;36(3):437-443.
36. Gorrindo P, Williams KC, Lee EB, Walker LS, McGrew SG, Levitt P. Gastrointestinal Dysfunction in Autism: Parental Report, Clinical Evaluation, and Associated Factors. *Autism Research*. 2012;5(2):101-108.
37. Valicenti-McDermott MD, McVicar K, Cohen HJ, Wershil BK, Shinnar S. Gastrointestinal Symptoms in Children with an Autism Spectrum Disorder and Language Regression. *Pediatric Neurology*. 2008;39(6):392-398.
38. Nikolov RN, Bearss KE, Lettinga J, et al. Gastrointestinal Symptoms in a Sample of Children with Pervasive Developmental Disorders. *Journal of Autism and Developmental Disorders*. 2009;39(3):405-413.

39. Mazurek MO, Vasa RA, Kalb LG, et al. Anxiety, Sensory Over-Responsivity, and Gastrointestinal Problems in Children with Autism Spectrum Disorders. *Journal of Abnormal Child Psychology*. 2013;41(1):165-176.
40. Pavone L, Fiumara A, Bottaro G, Mazzone D, Coleman M. Autism and celiac disease: failure to validate the hypothesis that a link might exist. *Biological psychiatry*. 1997;42(1):72-75.
41. Batista IC, Gandolfi L, Nobrega YKM, et al. Autism spectrum disorder and celiac disease: no evidence for a link. *Arquivos de Neuro-Psiquiatria*. 2012;70:28-33.
42. Chandler S, Carcani-Rathwell I, Charman T, et al. Parent-Reported Gastro-intestinal Symptoms in Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*. 2013;43(12):2737-2747.
43. Ludvigsson JF, Reichenberg A, Hultman CM, Murray JA. A nationwide study of the association between celiac disease and the risk of autistic spectrum disorders. *JAMA Psychiatry*. 2013;70(11):1224-1230.
44. Rzepecka H, McKenzie K, McClure I, Murphy S. Sleep, anxiety and challenging behaviour in children with intellectual disability and/or autism spectrum disorder. *Research in Developmental Disabilities*. 2011;32(6):2758-2766.
45. Tsai F-J, Chiang H-L, Lee C-M, et al. Sleep problems in children with autism, attention-deficit hyperactivity disorder, and epilepsy. *Research in Autism Spectrum Disorders*. 2012;6(1):413-421.
46. Schreck KA, Mulick JA, Smith AF. Sleep problems as possible predictors of intensified symptoms of autism. *Research in Developmental Disabilities*. 2004;25(1):57-66.
47. Flanagan JE, Landa R, Bhat A, Bauman M. Head lag in infants at risk for autism: a preliminary study. *The American journal of occupational therapy : official publication of the American Occupational Therapy Association*. 2012;66(5):577-585.
48. Lotter V. *Epidemiology of autistic conditions in young children: I. Prevalence*. Vol 11966.
49. Treffert DA. Epidemiology of infantile autism. *Archives of general psychiatry*. 1970;22(5):431-438.
50. Van Naarden Braun K, Christensen D, Doernberg N, et al. Trends in the Prevalence of Autism Spectrum Disorder, Cerebral Palsy, Hearing Loss, Intellectual Disability, and Vision Impairment, Metropolitan Atlanta, 1991–2010. *PLoS ONE*. 2015;10(4):e0124120.
51. Shattuck PT. The Contribution of Diagnostic Substitution to the Growing Administrative Prevalence of Autism in US Special Education. *Pediatrics*. 2006;117(4):1028-1037.
52. Tick B, Bolton P, Happé F, Rutter M, Rijdsdijk F. Heritability of autism spectrum disorders: a meta- analysis of twin studies. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. 2016;57(5):585-595.
53. Gaugler T, Klei L, Sanders SJ, et al. Most genetic risk for autism resides with common variation. 2014;46:881.
54. de la Torre-Ubieta L, Won H, Stein JL, Geschwind DH. Advancing the understanding of autism disease mechanisms through genetics. *Nature Medicine*. 2016;22:345.
55. Krishnan A, Zhang R, Yao V, et al. Genome-wide prediction and functional characterization of the genetic basis of autism spectrum disorder. *Nature neuroscience*. 2016;19(11):1454-1462.
56. Pinto D, Delaby E, Merico D, et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *American journal of human genetics*. 2014;94(5):677-694.
57. De Rubeis S, He X, Goldberg AP, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*. 2014;515(7526):209-215.
58. Sassone-Corsi P, Christen Y. *A Time for Metabolism and Hormones*. Springer International Publishing; 2016.

59. DiGiuseppi C, Hepburn S, Davis JM, et al. Screening for Autism Spectrum Disorders in Children With Down Syndrome: Population Prevalence and Screening Test Characteristics. *Journal of developmental and behavioral pediatrics : JDBP*. 2010;31(3):181-191.
60. Ji NY, Capone GT, Kaufmann WE. Autism spectrum disorder in Down syndrome: cluster analysis of Aberrant Behaviour Checklist data supports diagnosis. *Journal of intellectual disability research : JIDR*. 2011;55(11):1064-1077.
61. Sanders SJ, He X, Willsey AJ, et al. Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron*. 2015;87(6):1215-1233.
62. Girirajan S, Brkanac Z, Coe BP, et al. Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS genetics*. 2011;7(11):e1002334.
63. Girirajan S, Dennis MY, Baker C, et al. Refinement and discovery of new hotspots of copy-number variation associated with autism spectrum disorder. *American journal of human genetics*. 2013;92(2):221-237.
64. Weiss LA, Shen Y, Korn JM, et al. Association between Microdeletion and Microduplication at 16p11.2 and Autism. *New England Journal of Medicine*. 2008;358(7):667-675.
65. Walsh KM, Bracken MB. Copy number variation in the dosage-sensitive 16p11.2 interval accounts for only a small proportion of autism incidence: a systematic review and meta-analysis. *Genetics in medicine : official journal of the American College of Medical Genetics*. 2011;13(5):377-384.
66. Jacquemont S, Reymond A, Zufferey F, et al. Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature*. 2011;478(7367):97-102.
67. Angkustsiri K, Goodlin-Jones B, Deprey L, Brahmabhatt K, Harris S, Simon TJ. Social Impairments in Chromosome 22q11.2 Deletion Syndrome (22q11.2DS): Autism Spectrum Disorder or a Different Endophenotype? *Journal of autism and developmental disorders*. 2014;44(4):739-746.
68. Krumm N, Turner TN, Baker C, et al. Excess of rare, inherited truncating mutations in autism. *Nat Genet*. 2015;47(6):582-588.
69. Skaar DA, Shao Y, Haines JL, et al. Analysis of the RELN gene as a genetic risk factor for autism. *Mol Psychiatry*. 2005;10(6):563-571.
70. Sutcliffe JS, Delahanty RJ, Prasad HC, et al. Allelic heterogeneity at the serotonin transporter locus (SLC6A4) confers susceptibility to autism and rigid-compulsive behaviors. *American journal of human genetics*. 2005;77(2):265-279.
71. Zec N, Rowitch DH, Bitgood MJ, Kinney HC. Expression of the homeobox-containing genes EN1 and EN2 in human fetal midgestational medulla and cerebellum. *Journal of neuropathology and experimental neurology*. 1997;56(3):236-242.
72. Wang L, Jia M, Yue W, et al. Association of the ENGRAILED 2 (EN2) gene with autism in Chinese Han population. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*. 2008;147b(4):434-438.
73. Hunter J, Rivero-Arias O, Angelov A, Kim E, Fotheringham I, Leal J. Epidemiology of fragile X syndrome: A systematic review and meta-analysis. *American Journal of Medical Genetics Part A*. 2014;164(7):1648-1658.
74. Welt CK. Primary ovarian insufficiency: a more accurate term for premature ovarian failure. *Clinical endocrinology*. 2008;68(4):499-509.
75. Jacquemont S, Hagerman RJ, Leehey MA, et al. Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. *Jama*. 2004;291(4):460-469.
76. Mc Devitt N, Gallagher L, Reilly RB. Autism Spectrum Disorder (ASD) and Fragile X Syndrome (FXS): Two Overlapping Disorders Reviewed through

- Electroencephalography—What Can be Interpreted from the Available Information? *Brain Sciences*. 2015;5(2):92-117.
77. Sanders SJ, Ercan-Sencicek AG, Hus V, et al. Multiple recurrent de novo copy number variations (CNVs), including duplications of the 7q11.23 Williams-Beuren syndrome region, are strongly associated with autism. *Neuron*. 2011;70(5):863-885.
 78. Gratten J, Visscher PM, Mowry BJ, Wray NR. Interpreting the role of de novo protein-coding mutations in neuropsychiatric disease. *Nat Genet*. 2013;45(3):234-238.
 79. Sandin S, Lichtenstein P, Kuja-Halkola R, Larsson H, Hultman CM, Reichenberg A. The familial risk of autism. *JAMA*. 2014;311(17):1770-1777.
 80. Lyall K, Constantino JN, Weisskopf MG, Roberts AL, Ascherio A, Santangelo SL. Parental social responsiveness and risk of autism spectrum disorder in offspring. *JAMA Psychiatry*. 2014;71(8):936-942.
 81. Modabbernia A, Velthorst E, Reichenberg A. Environmental risk factors for autism: an evidence-based review of systematic reviews and meta-analyses. *Molecular Autism*. 2017;8:13.
 82. Abel KM, Dalman C, Svensson AC, et al. Deviance in Fetal Growth and Risk of Autism Spectrum Disorder. *American Journal of Psychiatry*. 2013;170(4):391-398.
 83. Buchmayer S, Johansson S, Johansson A, Hultman CM, Sparén P, Cnattingius S. Can Association Between Preterm Birth and Autism be Explained by Maternal or Neonatal Morbidity? *Pediatrics*. 2009;124(5):e817.
 84. Wu S, Wu F, Ding Y, Hou J, Bi J, Zhang Z. Advanced parental age and autism risk in children: a systematic review and meta-analysis. *Acta Psychiatrica Scandinavica*. 2017;135(1):29-41.
 85. Gardener H, Spiegelman D, Buka SL. Prenatal risk factors for autism: comprehensive meta-analysis. *The British journal of psychiatry : the journal of mental science*. 2009;195(1):7-14.
 86. Gardener H, Spiegelman D, Buka SL. Perinatal and neonatal risk factors for autism: a comprehensive meta-analysis. *Pediatrics*. 2011;128(2):344-355.
 87. Conde-Agudelo A, Rosas-Bermudez A, Norton MH. Birth Spacing and Risk of Autism and Other Neurodevelopmental Disabilities: A Systematic Review. *Pediatrics*. 2016;137(5).
 88. Curran EA, O'Neill SM, Cryan JF, et al. Research review: Birth by caesarean section and development of autism spectrum disorder and attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. *J Child Psychol Psychiatry*. 2015;56(5):500-508.
 89. Jiang HY, Xu LL, Shao L, et al. Maternal infection during pregnancy and risk of autism spectrum disorders: A systematic review and meta-analysis. *Brain, behavior, and immunity*. 2016;58:165-172.
 90. Xu G, Jing J, Bowers K, Liu B, Bao W. Maternal diabetes and the risk of autism spectrum disorders in the offspring: a systematic review and meta-analysis. *J Autism Dev Disord*. 2014;44(4):766-775.
 91. Wang Y, Tang S, Xu S, Weng S, Liu Z. Maternal Body Mass Index and Risk of Autism Spectrum Disorders in Offspring: A Meta-analysis. *Scientific reports*. 2016;6:34248.
 92. Conti E, Mazzotti S, Calderoni S, Saviozzi I, Guzzetta A. Are children born after assisted reproductive technology at increased risk of autism spectrum disorders? A systematic review. *Human reproduction (Oxford, England)*. 2013;28(12):3316-3327.
 93. Sharp WG, Berry RC, McCracken C, et al. Feeding problems and nutrient intake in children with autism spectrum disorders: a meta-analysis and comprehensive review of the literature. *J Autism Dev Disord*. 2013;43(9):2159-2173.
 94. Babaknejad N, Sayehmiri F, Sayehmiri K, Mohamadkhani A, Bahrami S. The Relationship between Zinc Levels and Autism: A Systematic Review and Meta-analysis. *Iranian journal of child neurology*. 2016;10(4):1-9.

95. Wang T, Shan L, Du L, et al. Serum concentration of 25-hydroxyvitamin D in autism spectrum disorder: a systematic review and meta-analysis. *Eur Child Adolesc Psychiatry*. 2016;25(4):341-350.
96. Pu D, Shen Y, Wu J. Association between MTHFR gene polymorphisms and the risk of autism spectrum disorders: a meta-analysis. *Autism research : official journal of the International Society for Autism Research*. 2013;6(5):384-392.
97. Yoshimasu K, Kiyohara C, Takemura S, Nakai K. A meta-analysis of the evidence on the impact of prenatal and early infancy exposures to mercury on autism and attention deficit/hyperactivity disorder in the childhood. *Neurotoxicology*. 2014;44:121-131.
98. Taylor LE, Swerdfeger AL, Eslick GD. Vaccines are not associated with autism: an evidence-based meta-analysis of case-control and cohort studies. *Vaccine*. 2014;32(29):3623-3629.
99. Rosen BN, Lee BK, Lee NL, Yang Y, Burstyn I. Maternal Smoking and Autism Spectrum Disorder: A Meta-analysis. *J Autism Dev Disord*. 2015;45(6):1689-1698.
100. Pinto D, Pagnamenta AT, Klei L, et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*. 2010;466(7304):368-372.
101. Mefford HC, Eichler EE. Duplication hotspots, rare genomic disorders, and common disease. *Current opinion in genetics & development*. 2009;19(3):196-204.
102. Levy D, Ronemus M, Yamrom B, et al. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron*. 2011;70(5):886-897.
103. Schumann G, Loth E, Banaschewski T, et al. The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Mol Psychiatry*. 2010;15(12):1128-1139.
104. Pausova Z, Paus T, Abrahamowicz M, et al. Cohort Profile: The Saguenay Youth Study (SYS). *International Journal of Epidemiology*. 2017;46(2):e19-e19.
105. Beran TN, Elliott, C. D. (2007). Differential Ability Scales (2nd ed.). San Antonio, TX: Harcourt Assessment. *Canadian Journal of School Psychology*. 2007;22(1):128-132.
106. Wechsler D. *WISC-IV: Wechsler Intelligence Scale for Children, 4th Edition Integrated: Technical and Interpretive Manual*. Harcourt Brace and Company; 2004.
107. Wechsler D. *Wechsler Abbreviated Scale of Intelligence WASI: Manual*. Pearson/PsychCorpl; 1999.
108. Mullen EM. *Infant MSEL Manual: Infant Mullen Scales of Early Learning*. American Guidance Service; 1989.
109. *WISC III (Wechsler Intelligence Scale for Children)*. Psychological Corporation; 1992.
110. Joseph RM, Tager-Flusberg H, Lord C. Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. *Journal of child psychology and psychiatry, and allied disciplines*. 2002;43(6):807-821.
111. Hus V, Gotham K, Lord C. Standardizing ADOS Domain Scores: Separating Severity of Social Affect and Restricted and Repetitive Behaviors. *Journal of autism and developmental disorders*. 2014;44(10):2400-2412.
112. Gotham K, Pickles A, Lord C. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *J Autism Dev Disord*. 2009;39(5):693-705.
113. Wang K, Li M, Hadley D, et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome research*. 2007;17(11):1665-1674.
114. Colella S, Yau C, Taylor JM, et al. QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic acids research*. 2007;35(6):2013-2025.
115. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research*. 2010;38(16):e164-e164.

116. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285-291.
117. !!! INVALID CITATION !!! ¹.
118. Stefansson H, Meyer-Lindenberg A, Steinberg S, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature*. 2014;505(7483):361-366.
119. Rao PA, Raman V, Thomas T, Ashok MV. IQ in Autism: Is there an Alternative Global Cognitive Index? *Indian Journal of Psychological Medicine*. 2015;37(1):48-52.
120. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*. 2017;169(7):1177-1186.
121. Huguet G, Schramm C, Douard E, et al. Measuring and Estimating the Effect Sizes of Copy Number Variants on General Intelligence in Community-Based Samples. *JAMA Psychiatry*. 2018.
122. Bishop SL, Farmer C, Bal V, et al. Identification of Developmental and Behavioral Markers Associated With Genetic Abnormalities in Autism Spectrum Disorder. *The American journal of psychiatry*. 2017;174(6):576-585