UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE Sciences de la Vie et de la Santé

UMR 7104 I.G.B.M.C

THÈSE présentée par :

Thomas ARBOGAST

soutenue le : **01 décembre 2014**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Neurosciences

Apport des modèles murins à la compréhension des maladies associées aux variations du nombre de copies : monosomie 21 partielle et délétions et duplications des régions 16p11.2 et 17q21.31

THÈSE dirigée par :

Yann HERAULT	Directeur de Recherche, Université de Strasbourg

RAPPORTEURS :

Jean-Maurice DELABAR	Directeur de Recherche, Université Paris Diderot, Paris
Pierre BILLUART	Chargé de Recherche, Université Paris Descartes, Paris

---

AUTRES MEMBRES DU JURY :

Jamel CHELLY	Professeur, Université de Strasbourg
Jacques S. BECKMANN	Professeur, Swiss Institute of Bioinformatics, Lausanne

Remerciements

Je tiens tout d'abord à exprimer ma reconnaissance au Dr Yann Hérault pour la confiance qu'il m'a accordée il y a quatre ans en m'accueillant dans son équipe. Merci de m'avoir confié ces projets scientifiques passionnants et de m'avoir appris à aborder les questions scientifiques de manière rigoureuse. Merci enfin de m'avoir permis de développer mon autonomie tout en étant disponible à chacune de mes interrogations.

Je remercie les membres du jury qui m'ont fait l'honneur d'accepter d'évaluer mon travail de thèse : Dr. Jean-Maurice Delabar, Dr. Pierre Billuart, Pr Jamel Chelly, Pr Jacques S. Beckmann, Pr Alexandre Reymond, Dr. Marika Nosten-Bertrand et Dr Abdel-Mouttalib Ouagazzal.

Un grand merci à toute l'équipe de m'avoir initié à la recherche dans le domaine des neurosciences. Merci Véronique Brault pour tes nombreuses explications notamment en ce qui concerne la biologie moléculaire. Merci Arnaud Duchon de m'avoir épaulé lors de mes premiers tests comportementaux sur le projet Ms5Yah et pour tes nombreux conseils comportementaux et statistiques qui ont suivi avec les nouveaux projets. Merci Claire Chevalier pour ta gestion des nombreuses lignées de l'équipe et pour la génération des cohortes. Bien évidemment, j'exprime mes sentiments les meilleurs à Damien Maréchal qui est en pleine ascension sociale : doctorant souris, post-doc souris et bientôt manager rat. Merci pour ton aide et pour tes réponses à mes innombrables questions. Merci Michel Roux pour ton implication vision sur le projet Ms5Yah. Merci Valérie Nalesso et Christelle Wagner pour vos conseils de bienséance et d'élégance. Un grand merci à tous les membres de l'équipe pour votre gentillesse et votre aide qui m'ont permis, tout comme la souris, de tirer mon épingle du jeu. Je renouvelle mes remerciements à Véronique Brault et à Michel Roux qui ont participé à la relecture de ce manuscrit.

Je tiens à exprimer ma profonde reconnaissance à Abdel-Moutallib Ouagazzal pour sa collaboration sur le projet 16p11.2. Un grand merci, Abdel, pour m'avoir fait partager tes connaissances et ta rigueur scientifique dans les domaines du comportement murin et des troubles neuropsychiatriques.

Je remercie Hamid Meziane et l'ensemble des membres de la plateforme comportementale notamment Aline Lux, Valérie Lalanne, le Christ Mittelhaeuser et le taulier Fabrice Riet pour votre disponibilité, vos nombreux conseils techniques et vos chronomètres. C'était un plaisir de travailler parmi vous, surtout le mercredi. Merci à Marie-Christine Birling pour la génération des lignées de souris floxées. Merci à Hugues Jacobs et Olivia Wendling pour l'étude histologique du projet Ms5Yah et pour leurs conseils. J'en profite pour saluer Patrice Goetz-Reiner et Bruno Weber. Merci à Marie-France Champy ainsi qu'à Laurent Pouilly et Aurélie Aubertin pour les études métaboliques en alimentation enrichie pour le projet 16p11.2. Merci à Alain Guimond pour l'évaluation auditive des animaux 16p11.2. Merci également à nos collaborateurs externes, Alexandre Reymond et Eugenia Migliavacca pour les analyses transcriptomiques du projet 16p11.2 et Maksym Kopanitsa pour les études électrophysiologiques des projets 16p11.2 et 17q21.31.

Je terminerai en remerciant ma mère Catherine, ma sœur Tiphaine, et Marine. Un grand merci à toutes les trois pour votre soutien bienfaisant.

Sommaire

Liste des figures

Abréviations

ABR	Auditory Brainstem Response
aCGH	Microarray-based Comparative Genomic Hybridization
ADHD	Attention Deficit Hyperactivity Disorder
ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
AS	Angelman Syndrome
ASD	Autism Spectrum Disorder
AAV	Adeno-Associated Virus
B6N	C57BL/6N genetic background
B6NC3B	Mixed C57BL/6N-C3H/HeH genetic background
BIR	Break-Induced Replication
BMI	Body Mass Index
BP	Breakpoint
C3B	C3H/HeH genetic background
CA	Circadian Activity
CNV	Copy Number Variation
Del/+	Animal porteur d'une délétion pour la région étudiée à l'état hétérozygote
Del/Dup	Animal porteur d'une délétion et d'une duplication pour la région étudiée
Dup/+	Animal porteur d'une délétion pour la région étudiée à l'état hétérozygote
EE	Energy Expenditure
EEG	Electroencéphalogramme
EMG	Electromyogramme
FC	Fold Change / variation de niveau d'expression
fEPSP	Field Excitatory PostSynaptic Potential
FoSTeS	Fork Stalling and Template Switching
g	Gramme
GO	Gene Ontology
GWAS	Genome Wide Association Study
HapMap	Haplotype Map
HSA21	Chromosome humain 21
ID	Intellectual Disability
Indel	Insertion/Délétion
IEG	Idiopathic Generalized Epilepsy
IPGTT	IntraPeriotoneal Glucose Tolerance Test
Kansl1$^{+/-}$	Animal knock-out à l'état hétérozygote pour le gène *Kansl1*
kb	Kilobase

KO	Knock-Out
LCRs	Low Copy Repeats
LTP	Long-Term Potentiation
Mb	Mégabase
min	Minute
MMBIR	Microhomology-Mediated Break-Induced Replication
MMEJ	Microhomology-Mediated End Joining
MMU10, 16, 17	Chromosome murin 10, 16, 17
MWM	Test de piscine de Morris
NAHR	Non-Allelic Homologous Recombination
NGS	Next-Generation Sequencing
NH	Non-Homologous Recombination
NHEJ	Non-Homologous End Joining
NPR	Novel Place Recognition test
NOR	Novel Object Recognition test
OF	Open Field
OMIM	Online Mendelian Inheritance in Man
pb	Paire de base
PCR	Réaction de polymérisation en chaîne
PM21	Monosomie 21 partielle
PPI	PrePulse Inhibition
PPF	Paired Pulse Facilitation
PTZ	Pentylenetetrazole
QI	Quotient Intellectuel
RMA	Robust Multi Array
RNA-seq	Séquençage nouvelle génération de l'ARN
rpm	Rotation par minute
RS	Replication Slippage
s	Seconde
SD	Segment génomique Dupliqué
SDG	DiGeorge Syndrome
SEM	Erreur standard de la moyenne
SNP	Single Nucleotide Polymorphism
SNParray	Miccroarray-based Single Nucleotide Polymophism genotyping
SS	Sotos Syndrome
TAMERE	Targeted Meiotic Recombination
vs	Versus
wt	Wild-type / animal sauvage
WS	Williams syndrome

Introduction Générale

I Les variations du nombre de copies

1. Amélioration des techniques de détection

Jusqu'à la fin du 20ème siècle, notre connaissance des variations génétiques humaines se limitait principalement aux aneuploïdies, aux variations structurales assez larges pour être visibles au microscope, ainsi qu'aux polymorphismes nucléotidiques (SNPs pour Single Nucleotide Polymorphisms) identifiés par le traditionnel séquençage par amplification en chaîne par polymérase. Le développement des techniques d'hybridation in situ en fluorescence (Pinkel et al., 1988), d'hybridation génomique comparative (Kallioniemi et al., 1992) permit d'augmenter la résolution de détection des variations structurales. La technique de puce d'hybridation génomique comparative (aCGH pour Array Comparative Genomic Hybridization) aboutit à la découverte et l'étude des variations du nombre de copies (CNVs pour Copy Number Variations) submicroscopiques (Albertson and Pinkel, 2003). L'amélioration des technologies de puces d'hybridation (aCGH et SNParray) et surtout le développement du séquençage de nouvelle génération (NGS pour Next-Generation Sequencing) ont révolutionné l'analyse génomique (Fromer et al., 2012; Handsaker et al., 2011; MacDonald et al., 2014) (Figure 1). Ces techniques permettent la détection de CNVs de taille variant de quelques paires de base (pb) à plusieurs Mb. Les loci associés aux CNVs englobent une fraction génomique supérieure aux SNPs et sont donc aujourd'hui considérés comme la source principale de variabilité entre les individus (Conrad et al., 2010b; Redon et al., 2006). L'analyse des CNVs a profondément modifié notre appréhension des variations structurelles du génome et notre compréhension des maladies complexes et des maladies syndromiques associées aux anomalies de structure chromosomique (Manolio et al., 2009).

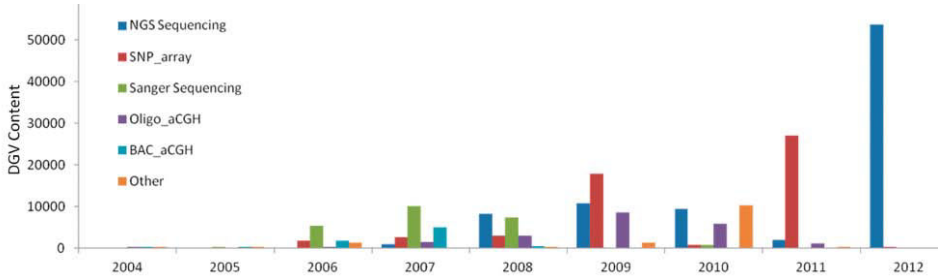


Figure 1: Evolution du nombre de polymorphismes référencés dans les bases de données

Augmentation considérable du nombre de polymorphismes référencés dans les bases de données à partir de 2008 avec le développement des techniques de puces d'hybridation (aCGH, SNParray) et de séquençage de nouvelle génération (NGS) prédominant depuis 2012 (Tiré de MacDonald *et al.*, 2014).

2. Catégories

Les CNVs incluent des variations de structure chromosomique associées à une perte ou un gain de matériel génétique. Les CNVs incluent les délétions, duplications, triplications, insertions et translocations d'une taille variant de quelques dizaines de pb à plusieurs Mb. En addition d'une large classe de CNVs de petite taille générés par des mécanismes de transpositions et d'erreurs de réplication ou de réparation pour les segments répétés en tandem, il existe deux catégories principales de CNVs qui divergent dans leur structure et leur origine cellulaire : les CNVs récurrents et les CNVs non-récurrents (Figure 2).

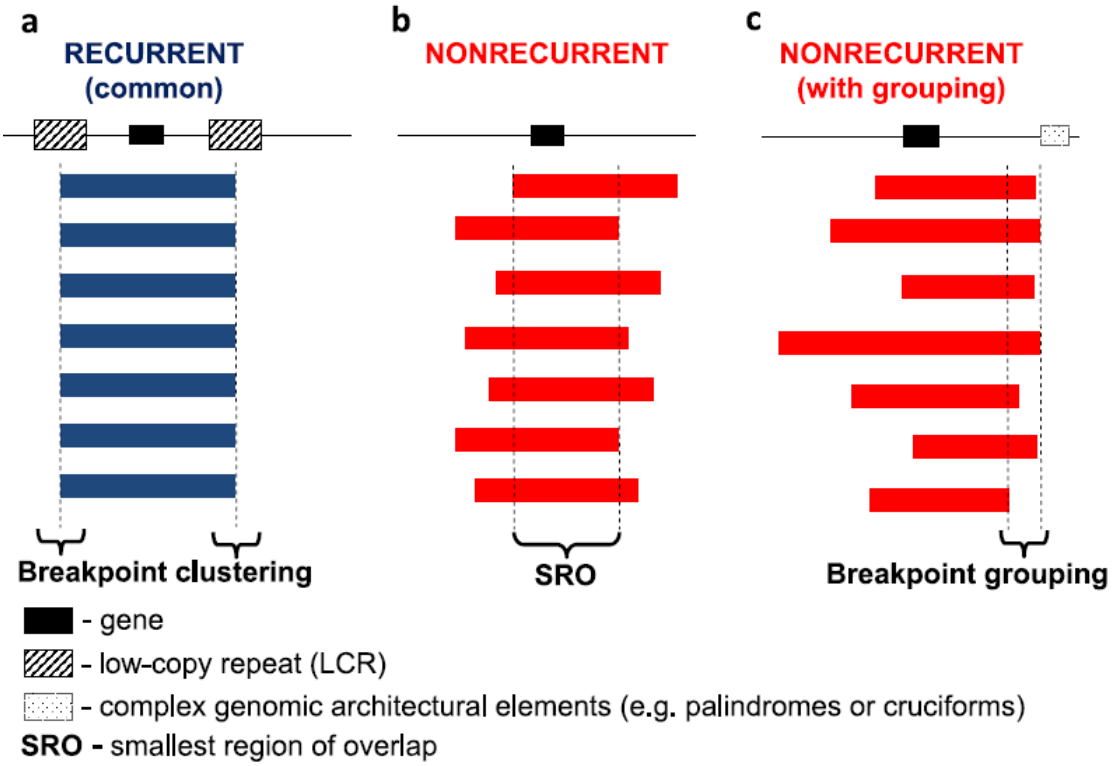


Figure 2 : CNVs récurrents et non-récurrents

a. Les délétions et duplications récurrentes ou communes ont deux points de cassure définis par des segments dupliqués ou LCR. La vaste majorité des réarrangements récurrents est le résultat de recombinaison homologue non-allélique (NAHR).
b. Les délétions et duplications non-récurrentes ont des points de cassure dispersés. Une région critique peut être définie comme la région commune des différents réarrangements.
c. Une séquence d'ADN à l'architecture particulière peut favoriser un point de cassure.
(Tiré de Gu *et al.*, 2008).

2.1. CNVs récurrents

De 20 à 40% des CNVs sans incidence phénotypique ainsi qu'une majorité de CNVs *de novo* pathogènes présentent des points de cassure récurrents au niveau de segments dupliqués (SDs) (Abecasis et al., 2010; Conrad et al., 2010b; Iafrate et al., 2004; Korbel et al., 2007; Mills et al., 2011) (Figure 2a). Les SDs ou LCRs (Low Copy Repeats) sont des segments d'ADN d'une taille généralement comprise entre 10 et 300 kb présents à plus d'un site dans le génome et présentant plus de 90% d'identité de séquence (Bailey et al., 2002; Eichler, 2001). Les SDs constituent environ 5% du génome et présentent une distribution largement hétérogène avec un enrichissement (32%) en périphérie des centromères chromosomiques (Koszul and Fischer, 2009). Les CNVs récurrents sont donc prédominants dans les régions péricentromériques qui présentent généralement des structures complexes et enrichies en séquences répétées (She et al., 2004; Stankiewicz and Lupski, 2002) (Figure 3).

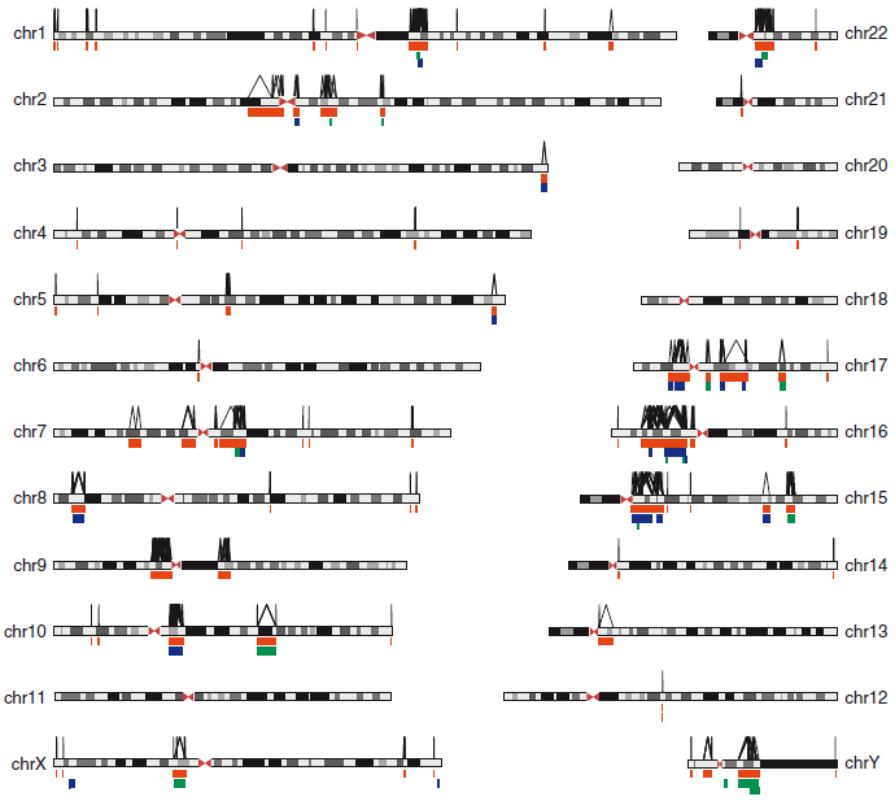


Figure 3 : Carte génomique des loci associés aux CNVs récurrents de taille supérieure à 50kb

Les rectangles rouges correspondent aux CNVs récurrents prédictifs, les rectangles verts aux délétions récurrentes pathogènes et les rectangles bleus aux délétions et duplications récurrentes pathogènes. Les lignes au-dessus des chromosomes correspondent aux NAHR potentiels entre SDs séparés d'une distance supérieure à 50kb. (Adapté de Liu *et al.*, 2012).

2.2. CNVs non-récurrents

Les CNVs non-récurrents présentent des points de cassure uniques (Figure 2b). Les points de jonction montrent généralement des micro-homologies de séquence. Une minorité de ces polymorphismes présente un caractère pathogène (Conrad et al., 2010a). La plupart des CNVs non-récurrents sont de simples délétions ou des duplications en tandem mais certains peuvent être plus complexes et interrompus par des séquences normales ou des inversions (Liu et al., 2011b). Il est probable que les CNVs non-récurrents complexes soient sous-estimés en conséquence de la difficulté d'identification de ces évènements.

3. Mécanismes de génération

Plusieurs mécanismes majeurs de génération des CNVs ont été proposés. Ces mécanismes incluent notamment la recombinaison homologue non-allélique, la recombinaison non-homologue et la transposition d'éléments mobiles (Hastings et al., 2009b). Globalement, les CNVs récurrents sont générés par des crossing-over entre séquences homologues stables. Au contraire, les CNVs non-récurrents sont la résultante d'insertions d'éléments transposables, d'erreurs de réplication, de mécanismes de réparation de l'ADN double brin et de recombinaisons entre séquences homologues instables et dispersées.

3.1. Recombinaison homologue non-allélique

La recombinaison homologue non-allélique (NAHR pour Non-Allelic Homologous Recombination) a été le premier mécanisme à être proposé et associé à des désordres génétiques (Lupski, 1998). Il correspond au crossing-over inégal entre régions d'homologie, notamment les SDs (Stankiewicz and Lupski, 2002), les éléments *Alu* (de Smith et al., 2008) et les éléments L1(Bi et al., 2009) .

Le mécanisme NAHR correspond à l'unique mécanisme de génération de variations structurelles récurrentes. Cette récurrence est la résultante d'une architecture génomique commune entre les individus qui correspond à la distribution génomique des segments dupliqués (SDs). En fonction de l'orientation des SDs (directe ou indirecte) et du type de recombinaison (en *cis* : intra-chromatide, ou en *trans* : inter-chromosomique et intra-chromosomique), le mécanisme NAHR aboutit à différents réarrangements structuraux (Liu et al., 2012) (Figure 4).

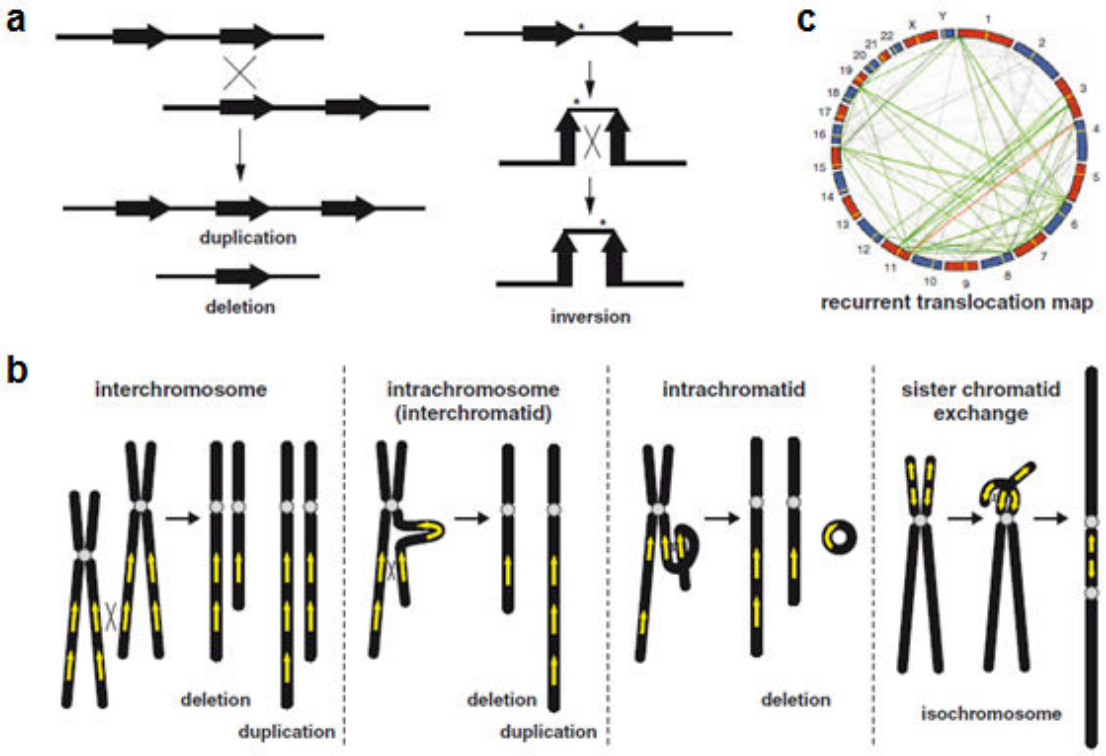


Figure 4 : Mécanisme NAHR et variations structurales récurrentes

a. Le crossing-over en *trans* entre SDs d'orientation directe conduit à une délétion et une duplication réciproque. Le crossing-over en *cis* entre SDs d'orientation indirecte conduit à une inversion.
b. Le mécanisme NAHR entre SDs d'orientation directe conduit à des délétions ou duplications par des crossing-over inter-chromosomique, intra-chromosomique et intra-chromatide. La recombinaison intra-chromosomique conduit uniquement à une délétion. La recombinaison entre des SDs d'orientation inversée de chromatides sœurs conduit à la formation d'un isochromosome.
c. Carte génomique des translocations récurrentes par le mécanisme NAHR observées expérimentalement (ligne rouge) et prédictives (lignes vertes et grises) selon la distribution des SDs.
(Tiré de Liu *et al.*, 2012).

Un crossing-over entre des SDs de chromosomes non-homologues aboutit à des translocations réciproques (Ou et al., 2011). Un crossing-over entre des SDs qui se trouvent dans une orientation inversée aboutit à des inversions pouvant éventuellement induire la formation d'isochromosome (Barbouti et al., 2004). Ces réarrangements n'engendrent ni perte ni gain de matériel génétique et constituent donc des variations structurelles neutres.

Un crossing-over entre des SDs qui se trouvent dans une orientation directe aboutit à des délétions et duplications. Le mécanisme NAHR favorise néanmoins les délétions par rapport aux duplications. Effectivement, si les crossing-over en *trans* inter-chromosomiques et intra-chromosomiques engendrent des délétions et des duplications, les crossing-over intra-

chromatide en *cis* n'engendrent que des délétions. Une étude indique que durant la méiose mâle, les crossing-over entre chromosomes autosomes engendrent deux fois plus de délétions que de duplications (Turner et al., 2008).

Certains rares NAHR peuvent également avoir lieu entre des LINEs et particulièrement entre des éléments *Alu* (Gu et al., 2008). Dans ce cas, le mécanisme NAHR aboutit à la génération de CNVs non-récurrents.

3.2. Mécanismes de réparation de l'ADN

Lors d'une cassure double brin de l'ADN, un des mécanismes de réparation homologue est l'appariement simple brin (SSA) (Figure 5c). Ce mécanisme a été décrit il y a trente ans chez la souris (Lin et al., 1984). La résection d'ADN aux deux extrémités est réalisée jusqu'à apparition d'une région d'homologie. Il y a alors appariement des répétitions et clivage des portions simple-brin non appariées puis jonction de l'ADN. Ce mécanisme de réparation implique la perte de la séquence d'ADN comprise entre les deux répétitions. Il est fréquemment observé lorsqu'il y a cassure double brin de l'ADN entre des séquences *Alu* séparés de quelques centaines de pb (Elliott et al., 2005).

Lors d'un effondrement de la fourche réplicative, la cassure double brin présente une extrémité disponible. La réparation peut alors être effectuée par le mécanisme BIR (Break-Induced Replication). L'extrémité envahit le brin homologue, initie la synthèse à partir du site d'invasion et réplique le chromosome homologue jusqu'au télomère. Si le processus de réparation implique des séquences répétées, des réarrangements complexes impliquant délétions, duplications et translocations peuvent apparaître (Sheen et al., 2007).

Il existe deux mécanismes de réparation des cassures double brin d'ADN qui ne requièrent pas de matrice homologue : les mécanismes NHEJ (Non-Homologous End Joining) (Figure 5a) et MMEJ (Microhomology-Mediated End Joining) (Figure 5b). Ces deux types de réparation peuvent occasionner la perte de matériel génétique et l'insertion d'ADN transposable (Haviv-Chesner et al., 2007; Lieber, 2008; McVey and Lee, 2008). Le mécanisme NHEJ fait intervenir la protéine Ku dans la ligature des extrémités et engendre des délétions ou insertions d'une taille de 1 à 4 pb. Le mécanisme MMEJ nécessite l'hybridation des séquences de micro-homologie d'une taille de 5 à 25 pb mais comme pour le mécanisme SSA, il y a délétion du segment compris entre les régions de micro-homologie et la possible insertion d'ADN transposable.

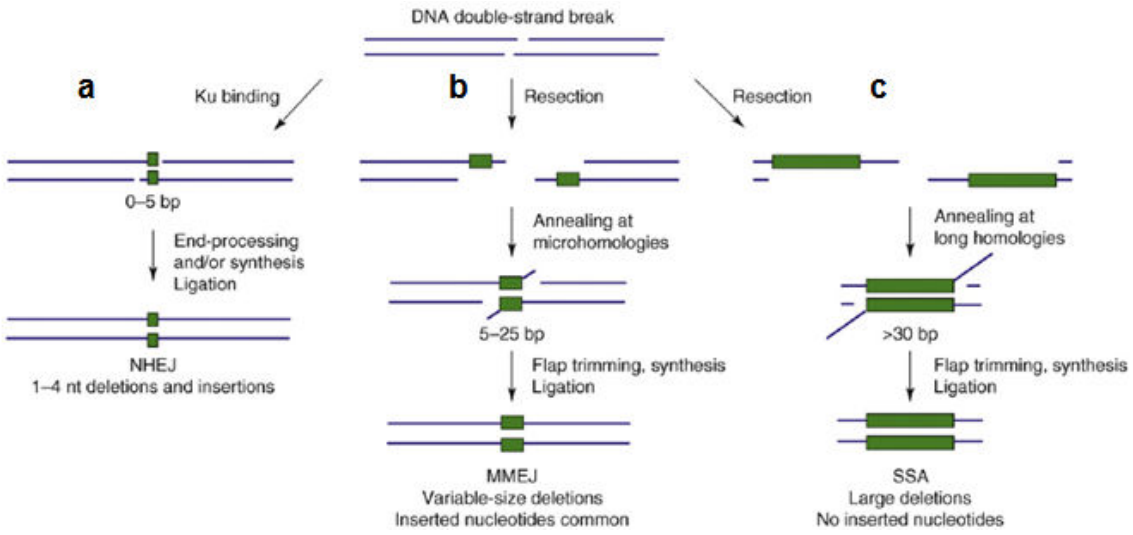


Figure 5 : Mécanismes de réparation des cassures d'ADN double brin

a. Le mécanisme NHEJ induit de courtes délétions à proximité des régions de micro-homologie.
b. Le mécanisme MMEJ induit la délétion du fragment compris entre les régions de micro-homologie.
c. Le mécanisme SSA aboutit à de larges délétions entre les régions d'homologie.
(Tiré de McVey et Lee, 2008).

3.3. Erreurs de réplication de l'ADN

De nombreuses évidences suggèrent que les mécanismes de réplication génèrent une part importante de CNVs non-récurrents (Hastings et al., 2009a). Un stress réplicatif généré par des agents qui perturbent la réplication normale de l'ADN, notamment l'aphidicoline et l'hydro-urée, augmentent la probabilité d'erreurs de réplication (Arlt et al., 2012). Les erreurs de réplication associées à des régions de micro-homologie incluent les mécanismes RS (Replication Slippage), FoSTeS (Fork Stalling and Template Switching) et MMBIR (Microhomology-Mediated Break-Induced Replication) (Figure 6).

Le mécanisme RS se produit lorsque deux régions de micro-homologie sont sous forme d'ADN simple brin durant la réplication. Des structures secondaires peuvent se former entre régions de micro-homologie inversées et aboutir à la délétion ou la duplication du segment d'ADN compris entre ces deux régions (Bzymek and Lovett, 2001). La taille de ces recombinaisons n'excède pas la longueur d'un fragment d'Okazaki, soit entre 100 et 200 pb.

Le mécanisme FoSTeS a été découvert chez *Escherichia coli* soumis à un stress réplicatif. Comme pour le mécanisme RS, le mécanisme FoSTeS peut apparaître lors d'un ralentissement de la fourche réplicative par la formation de structures secondaires. Dans le cas du FoSTeS, le changement de matrice a lieu entre régions de micro-homologie situées sur

deux brins retardés différents provenant de deux fourches réplicatives (Slack et al., 2006). Ce mécanisme peut engendrer des délétions et des duplications de plusieurs kb.

Le modèle MMBIR apparaîtrait lors de l'effondrement d'une fourche réplication en condition de stress cellulaire (Hastings et al., 2009a). Ce modèle propose l'appariement de l'extrémité disponible avec une région de micro-homologie sur un ADN simple brin disponible à proximité, amorçant une synthèse de faible processivité avec de multiples changements de matrices et éventuellement le ré-établissement d'une fourche réplicative. Ce modèle est proposé pour les CNVs impliquant des réarrangements complexes.

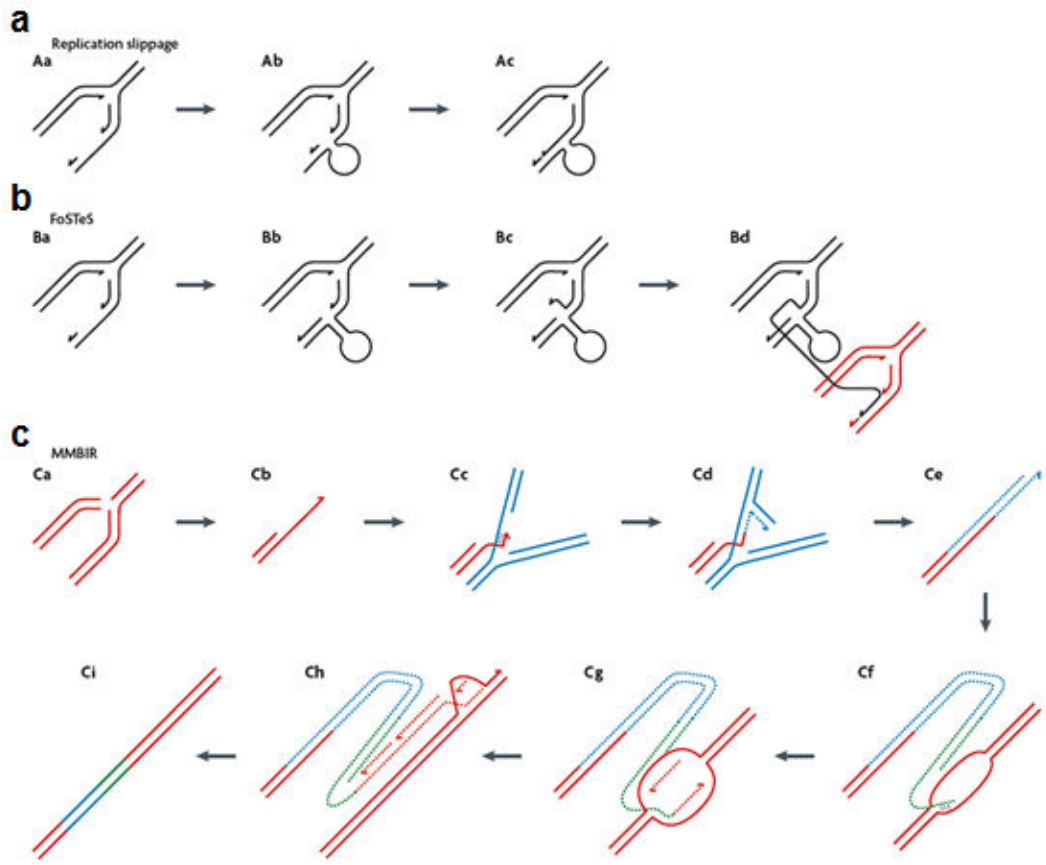


Figure 6 : Erreurs de réplication de l'ADN

a. Le mécanisme RS apparaît lors de la formation de structures secondaires entre deux fragments d'Okazaki. La synthèse du brin retardé se poursuit entre les régions de micro-homologie ce qui engendre la perte de la structure secondaire. Les délétions induites sont de taille inférieure à 200 pb.
b. Le mécanisme FoSTeS apparaît également lors de la formation de structures secondaires. Le changement de matrice a lieu entre régions micro-homologues de brins retardés provenant de deux fourches réplicatives. Ce mécanisme peut engendrer des délétions et duplications de plusieurs kb.
c. Le modèle MMBIR est proposé pour la génération de CNVs impliquant des réarrangements complexes. Lors de l'effondrement de la fourche réplicative, il y aurait appariement de l'extrémité disponible avec n'importe quel ADN simple brin disponible, multiples changements de matrices et ré-établissement d'une fourche réplicative. (Tiré de Hastings *et al.*, 2009).

4. Proportion et distribution dans la population normale

4.1 Premières études

La technique NGS est aujourd'hui majoritairement utilisée dans les études de variations structurales et de polymorphismes nucléotidiques à l'échelle génomique. Historiquement, les techniques aCGH et SNParray sont associées aux premières études de proportion et distribution des CNVs décrites dans les paragraphes qui suivent.

En 2004, une première étude de la distribution des CNVs fut effectuée en utilisant la technique aCGH (Sebat et al., 2004). Cette étude réalisée sur 20 individus non apparentés aboutit à la détection de 76 loci présentant des déséquilibres génomiques d'une taille supérieure à 100 kb. En moyenne, chaque individu présentait un nombre de 11 CNVs.

La même année, une étude aCGH réalisée sur 55 individus non apparentés aboutit à l'identification de 255 CNVs (Iafrate et al., 2004). Chaque individu présentait une moyenne de 12,4 CNVs. Les polymorphismes présentaient une distribution sur l'ensemble du génome humain à l'exception du chromosome Y. De nombreux loci coïncidait avec des segments dupliqués impliquant le mécanisme NAHR et la génération de CNVs récurrents.

Un an plus tard, une étude fut réalisée sur 47 individus non apparentés (Sharp et al., 2005). Sur la base de l'architecture des segments dupliqués du génome humain, une série de 130 sites de réarrangements potentiels ont été testés. 119 régions présentant des CNVs furent identifiées incluant des fréquences similaires pour les délétions et les duplications. Un important enrichissement des CNVs fut observé dans les régions flanquées par des segments dupliqués, confirmant l'implication de ces séquences homologues dans les mécanismes de génération des variations structurales de l'ADN.

En 2006, les données de génotypage de ~ 1,3 millions de SNPs provenant de 60 trios du consortium international HapMap (ou Haplotype Map) (Consortium, 2003) furent analysées par différents algorithmes conçus pour détecter les délétions (Conrad et al., 2006). Après confirmation par la technique aCGH, l'analyse aboutit à l'identification de 586 loci associés à des délétions d'une taille supérieure à 5 kb. Chaque individu présentait entre 30 et 50 délétions d'une taille moyenne de 10 kb. L'ensemble des polymorphismes couvrait un total de 267 gènes référencés. De manière générale, les loci détectés dans l'étude étaient relativement pauvres en gènes. Un léger déficit du nombre de délétions a été retrouvé sur le chromosome X pouvant refléter une pression de sélection concernant les individus mâles. Parmi les 267 gènes chevauchés par les CNVs, 23 gènes associés à des maladies ont été retrouvés dans la base de

données “Online Mendelian Inheritance in Man” ou OMIM (Hamosh et al., 2005). La base de données PANTHER (Thomas et al., 2003) permit d’étudier la fonction des gènes situés aux loci présentant des déséquilibres génomiques. L’analyse en gène ontologie (GO) révéla que les gènes impliqués dans la défense immunitaire, l’adhésion cellulaire et la transduction du signal étaient enrichis dans les délétions. Au contraire, les gènes codant des protéines de liaison aux acides nucléiques et des protéines impliquées dans le métabolisme des acides nucléiques étaient sous-représentés et subissaient donc une importante pression de sélection.

La même année, combinant les techniques aCGH et SNParray, une carte de la distribution globale des CNVs dans le génome humain fut réalisée à partir de l’ADN de 270 individus non apparentés provenant du consortium HapMap (Consortium, 2005; Redon et al., 2006). L’étude aboutit à l’identification de 1447 loci présentant des déséquilibres génomiques couvrant un ensemble de 360 Mb soit 12% du génome humain. Près de 60% des CNVs chevauchaient des gènes référencés dont 285 sont référencés dans la base de données OMIM. Une proportion plus faible de délétions par rapport aux duplications recouvrait ces 285 gènes. Ces résultats indiquent que les délétions induisent un effet particulièrement délétère sur l’expression génique. Des CNVs ont également été retrouvés dans les régions associées à de nombreux syndromes neurodéveloppementaux incluant syndromes de DiGeorge, Smith-Magenis, Williams-Beuren, Prader-Willi et Angelman. Ces résultats indiquent que l’identification des CNVs présente un intérêt majeur pour l’étude des maladies.

4.2 Etudes récentes

L’amélioration considérable des techniques de détection aCGH ainsi que le développement des séquençages NGS permettent aujourd’hui de détecter les CNVs d’une taille de quelques dizaines de pb. Les études récentes ont tiré avantage des évolutions technologiques pour générer des cartes de distribution des CNVs toujours plus précises.

En 2010, la distribution des CNVs de 40 individus non apparentés a été caractérisée en utilisant la méthode aCGH incluant 42 millions de sondes (espacement moyen de 56 bp). L’étude a permis l’identification de 11700 CNVs d’une taille supérieure à 443 pb (Conrad et al., 2010b). Parmi les 8599 CNVs validés, 70% des loci associés n’avaient pas encore été caractérisés dans des études précédentes. Ces données indiquent la prépondérance des CNVs de petite taille (quelques kb) qui n’étaient pas détectables par les méthodes utilisées dans les études précédentes. En moyenne, chaque individu présentait 1098 CNVs. Globalement, l’ensemble des CNVs chevauchaient 2698 gènes référencés (13,4%) et altéraient la séquence

codante de 1519 (5,5%) ARNm. Un pourcentage très faible de délétions englobant des gènes entiers a été observé confirmant un effet sur l'expression génique particulièrement pathogène pour les délétions en comparaison aux duplications (Figure 7).

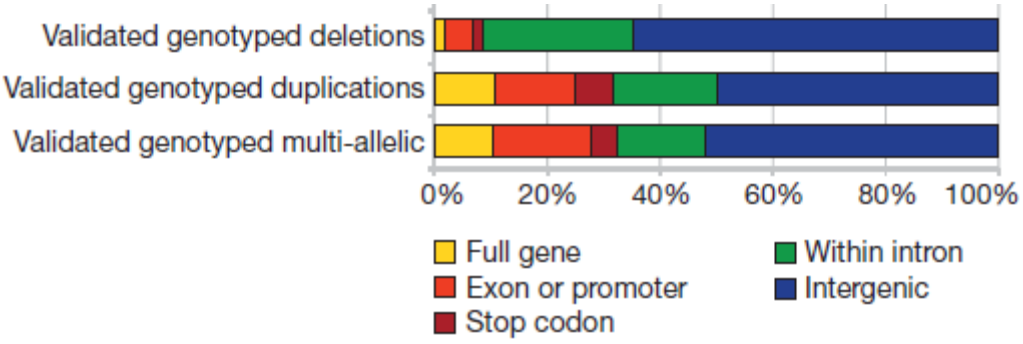


Figure 7 : Impact fonctionnel des différentes classes de CNVs

Impact sur l'expression génique des délétions, duplications, et CNVs multi-alléliques. La majorité des loci présentant des déséquilibres génomiques sont localisés dans des régions non-codantes. L'effet est exacerbé pour les délétions dont moins de 10% chevauchent des régions codantes.
(Tiré de Conrad *et al.*, 2010).

L'analyse GO indiquait un enrichissement des gènes impliqués dans les processus extracellulaires comme l'adhésion cellulaire et la communication extracellulaire ainsi qu'un appauvrissement des gènes impliqués dans certains processus intracellulaires comme les voies métaboliques de biosynthèse. Dans un deuxième temps, les auteurs ont déterminé le génotype de 10819 des loci identifiés en plus des 375 loci identifiés dans des études précédentes en utilisant 450 échantillons d'ADN provenant d'individus européens, africains et asiatiques du consortium HapMap. 5 238 CNVs ont été identifiés dans au moins une population HapMap et 4 978 concernaient les loci identifiés dans l'étude décrite. Ces résultats indiquent qu'un nombre importants de CNVs sont des polymorphismes récurrents identifiés chez des individus de différentes populations.

La même année, le séquençage de 185 individus européens provenant du consortium international "1000 Genomes Project" ou 1000GP (Abecasis et al., 2010) a été analysé. L'étude portait principalement sur la détection des délétions, ces dernières étant plus pathogène que les duplications réciproques (Mills et al., 2011). Les différentes analyses ont permis l'identification de 22025 délétions, 5371 insertions et 501 duplications en tandem d'une taille supérieure à 50 pb. Ces résultats mettent en lumière l'importante résolution du séquençage génome entier en comparaison à la technique aCGH dans l'identification des

variations structurelles du génome. La taille médiane des CNVs était de 729 pb ce qui confirme encore une fois la prépondérance des CNVs de petite taille variant de quelques centaines de pb à quelques kb.

Un avantage majeur du séquençage est la résolution nucléotidique des CNVs. Effectivement, les points de cassure d'environ 15000 CNVs ont été identifiés, permettant l'analyse de leur origine et de leur impact fonctionnel. La majorité des CNVs détectés d'une taille inférieure à 10 kb était des délétions générées par des mécanismes NH de type NHEJ, MMEJ et MMBIR (Figure 8a). La quasi-totalité des insertions était générée par l'insertion d'éléments mobiles, majoritairement des éléments *Alu* (Figure 8b). Enfin, la majorité des duplications d'une taille inférieure à 10 kb était générée par le mécanisme NAHR entre des éléments *Alu*. Cette étude confirme l'importante des mécanismes NH et des éléments *Alu* dans la génération de CNVs de taille inférieure à 10 kb.

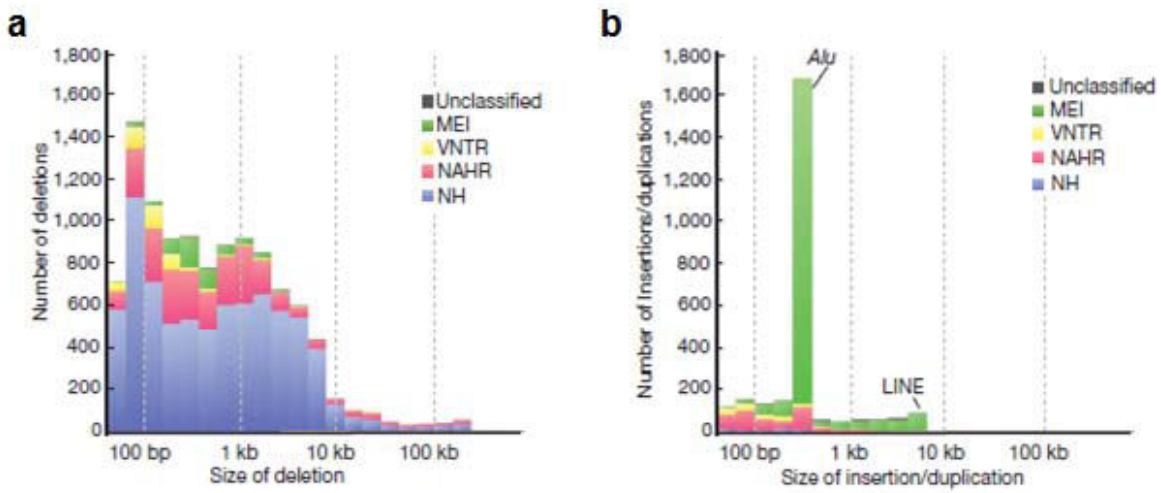


Figure 8 : Distribution de la taille des CNVs classés selon leur mécanisme de génération

a. Taille des délétions en fonction de leur mécanisme de génération. Le graphique indique que le génome humain contient une majorité de délétions d'une taille inférieure à 10kb générés par des mécanismes non-homologues de type NHEJ et MMBIR.
b. Taille des insertions et des duplications en fonction de leur mécanisme de génération. Le graphique indique une majorité d'insertions d'éléments *Alu* d'une taille de 300 pb ainsi qu'un nombre important d'éléments L1 d'une taille de 6 kb. Une majorité de duplications d'une taille inférieure à 5kb générée par des mécanismes NAHR par recombinaison entre éléments *Alu*.
MEI : insertion d'éléments mobiles. VNTR : répétition en tandem de nombre variable.
(Adapté de Mills et *al.*, 2011).

5. Association aux traits phénotypiques

5.1 Du CNV au phénotype

Les études de proportion et de distribution dans la population normale ont identifié des dizaines de milliers de CNVs chevauchant plusieurs milliers de gènes référencés (Abecasis et al., 2010; Conrad et al., 2010b). Une proportion non négligeable de CNVs aboutit donc à des variations d'expression génique (Lupski and Stankiewicz, 2005; Stankiewicz and Lupski, 2010). Les études d'association pangénomique ont permis de relier certaines variations génétiques à des traits phénotypiques. L'enrichissement de CNVs a notamment été observé dans des cohortes de patients pour un grand nombre de pathologies.

Les CNVs peuvent aboutir à des phénotypes par différents mécanismes moléculaires (Figure 9). Le plus commun d'entre eux est l'altération du nombre de gènes sensibles au dosage génique (Lupski et al., 1992). L'altération d'une séquence codante par interruption ou fusion peut aboutir à la synthèse d'un produit protéique non-fonctionnel ou pathogène. Les CNVs peuvent avoir un effet de position par la perte ou l'altération d'une séquence régulatrice et générer une variation d'expression d'un gène situé jusqu'à 1Mb en amont ou en aval de leur position (Kleinjan and van Heyningen, 2005). Enfin, les délétions peuvent aboutir à un phénotype pathogène quand une mutation récessive est présente dans l'allèle hémizygote (Kurotaki et al., 2005).

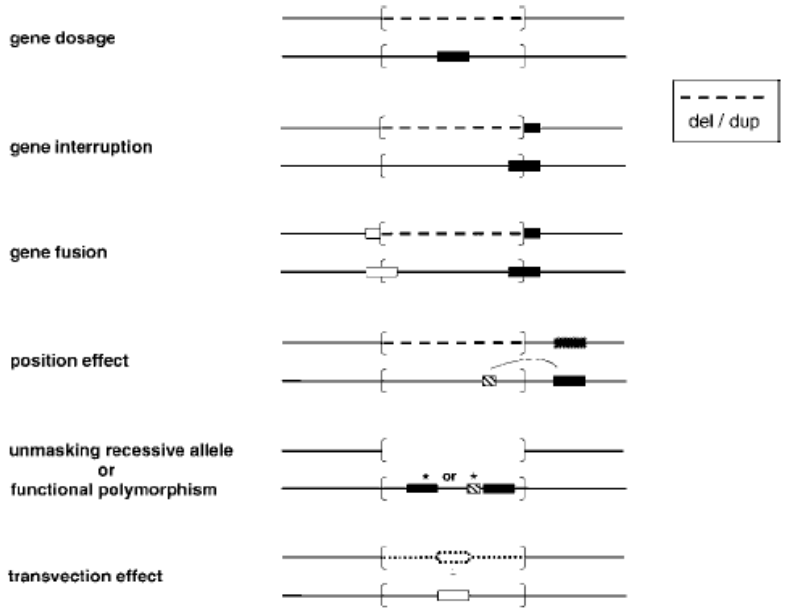


Figure 9 : Mécanismes moléculaires de désordres génétiques associés aux CNVs

Représentation des mécanismes aboutissant à une altération de l'expression génique ou de la synthèse protéique par interruption des séquences codante et régulatrice. (Tiré de Lupski *et al.*, 2005).

5.2 Les études d'association

Les études d'association consistent à déterminer si un polymorphisme génétique est associé à un ou plusieurs trait(s) phénotypique(s), principalement les symptômes d'une pathologie. Le but est de mettre en évidence une différence de fréquence allélique pour un marqueur génétique entre une population d'individus présentant un phénotype particulier et une population d'individus sains (les témoins). Un polymorphisme est associé à un trait phénotypique s'il est en excès ou en déficit dans la population considérée. Les études d'association offrent une grande puissance pour la détection de polymorphismes à effet mineur (Risch and Merikangas, 1996) mais requièrent un grand nombre d'échantillons afin d'éviter les faux positifs. Pour l'étude des pathologies, deux stratégies peuvent être utilisées. La stratégie la plus utilisée et la stratégie cas/témoins. La deuxième stratégie est le déséquilibre de transmission par analyse de trios (père, mère, enfant atteint). Les études d'association peuvent cibler des gènes candidats déterminés par des études de liaison ou impliqués dans les fonctions moléculaires ou cellulaires étudiées. La deuxième approche qui s'est largement développée avec le développement des technologies de puces d'hybridation est l'étude d'association sur génome entier (GWAS pour Genome-Wide Association Study). Une comparaison de la répartition des polymorphismes situés sur l'ensemble du génome est réalisée entre patients et témoins.

Depuis 2005, les GWAS ont permis d'associer certains CNVs à de nombreux traits phénotypiques observés entre les individus. Certains CNVs ont notamment été associés aux variations de taille corporelle (Visscher, 2008), de masse corporelle (Wheeler et al., 2013; Willer et al., 2009), de glycémie (Prokopenko et al., 2009), etc. Les CNVs peuvent engendrer des maladies syndromiques décrites dans la troisième partie de l'introduction et conférer des prédispositions à certaines maladies complexes, notamment la maladie de Crohn (McCarroll et al., 2008), le psoriasis (de Cid et al., 2009), la maladie de Parkinson (Pankratz et al., 2011), la maladie d'Alzheimer (Chapman et al., 2013) et de nombreux cancers (Hopman et al., 2013; Lu et al., 2014). Certains CNVs sont également associés aux désordres neuro-développementaux caractérisés par des anomalies du développement cérébral aboutissant à des dysfonctionnements cognitifs, neurologiques ou psychiatriques. Les GWAS ont permis l'identification de nombreux CNVs associés à des conditions neuropsychiatriques incluant la déficience intellectuelle, les troubles du spectre autistique, la schizophrénie, l'épilepsie, les troubles du déficit de l'attention avec hyperactivité (ADHD pour Attention Deficit Hyperactivity Disorder), le trouble bipolaire, etc.

6. Association aux troubles neuropsychiatriques

6.1 Association à la déficience intellectuelle

La déficience intellectuelle (ID pour Intellectual Disability) se caractérise par un quotient intellectuel (QI) inférieur ou égal à 70 (Chelly et al., 2006). L'ID est légère pour un QI de 50-70, modérée pour un IQ de 35-50, sévère pour un QI de 20-35 et profonde pour un QI<20 (Katz and Lazcano-Ponce, 2008). L'ID a une prévalence de 1 à 3% dans la population générale et montre un excès d'environ 30% chez les hommes (Leonard and Wen, 2002; Maulik et al., 2011). L'ID légère, majoritaire, représente 80 à 90% des cas d'ID dont une proportion importante est d'ordre familial et donc associée à des mutations héritées. Au contraire, l'ID associée à un QI inférieur à 50 est associée à des mutations de *novo*. Les causes génétiques compteraient pour environ 40% des cas d'ID et présentant une hétérogénéité considérable. Plus de 450 gènes ont déjà été associés à l'ID mais on estime que plus de 2000 gènes seraient associés aux formes monogéniques (Kleefstra et al., 2014; van Bokhoven, 2011). Les facteurs génétiques comprennent des aneuploïdies (notamment le syndrome de Down dont la prévalence est de 1/800), des CNVs (délétion/duplication 15q11.2, 22q11.2, 2q13, 16p11.2, etc.) ainsi que des SNPs et indels (pour insertions/délétions) dans des centaines de gènes impliquant des maladies monogéniques syndromiques (gènes *FMR1, MECP2, TCF4, NSD1, ARX*, etc.) et non-syndromiques. Les CNVs seraient responsables d'environ 15 à 20% des cas d'ID (Koolen et al., 2009; Miller et al., 2010) et les SNPs et indels pour 10% des cas d'ID (Ropers, 2010). Une étude récente indique néanmoins que les SNPs seraient la principale cause génétique pour les cas d'ID présentant un QI inférieur à 50 (Gilissen et al., 2014).

L'ID étant le plus souvent associée à d'autres troubles du développement incluant facteurs de comorbidité psychiatrique et anomalies congénitales (notamment cardiopathies et malformations cranio-faciales), un nombre important de GWAS ont porté sur des cohortes de patients présentant un retard développemental.

En 2011, une étude aCGH a permis l'étude du profil CNVs de 15767 enfants atteints de retard développemental et de 8329 adultes sains (Cooper et al., 2011). Les auteurs ont identifié 59 loci pathogènes incluant 14 nouvelles régions ainsi que 940 gènes candidats sensibles au dosage génique. 1492 CNVs incluant 954 délétions et 538 duplications ont été détectées parmi 1400 individus au niveau de 45 loci associés à des désordres génétiques. Ces résultats confirment encore une fois l'effet particulièrement pathogène des délétions par

rapport aux duplications. Un important enrichissement de larges CNVs a été identifié dans un grand nombre de patients déficients intellectuels (Figure 10a). Environ 25,7% des enfants atteints d'ID étaient porteurs d'au moins un CNV d'une taille supérieure à 400 kb contre 11,5% des témoins suggérant qu'environ 14,2% des cas d'ID sont causés par des CNVs d'une taille supérieure à 400 kb. Les individus atteints d'ID porteurs de CNVs d'une taille supérieure à 1,5 Mb ne représentaient que 11,3% contre 0,6% des témoins. Un enrichissement particulièrement important des larges CNVs a été retrouvé chez les individus présentant une ID associée à des anomalies cranio-faciales ou des défauts cardiaques (Figure 10b).

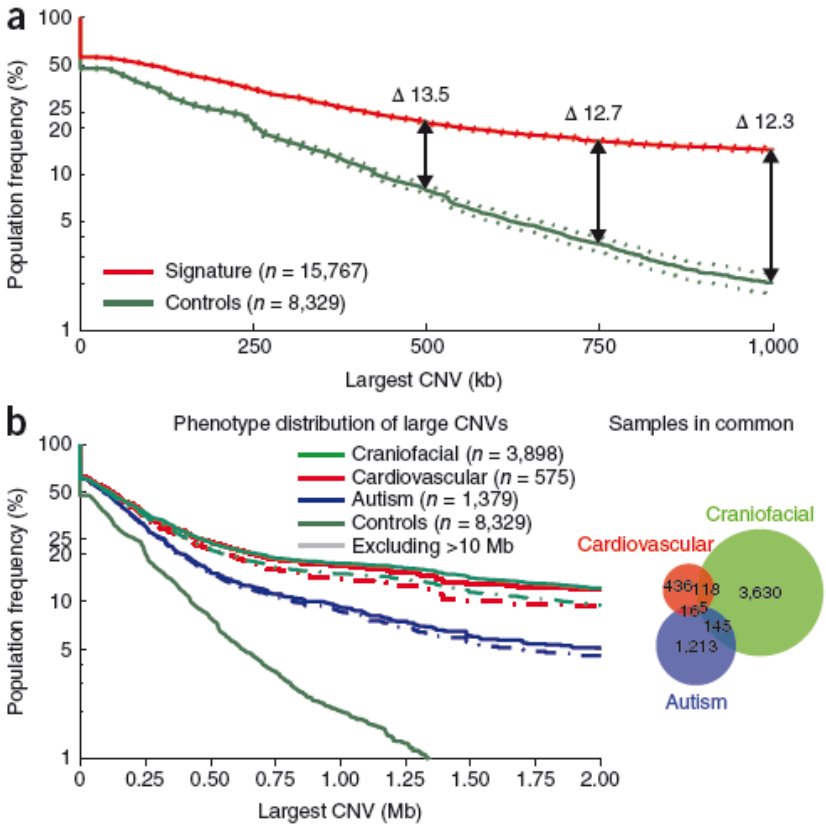


Figure 10 : Distribution de la taille des CNVs pour les individus témoins et les individus atteints de retard développemental

a. Enrichissement des larges CNVs chez les patients atteints de retard développemental. Plus la taille du CNVs est importante, plus l'écart entre les fréquences d'individus porteurs affectés et non-affectés se creuse, ce qui indique un effet particulièrement pathogène des larges CNVs.
b. La charge de larges CNVs est particulièrement importante pour les patients présentant une ID associée à des phénotypes cardiovasculaires ou cranio-faciaux par rapport aux patients atteints de troubles du spectre autistique.
(Tiré de Cooper et *al.*, 2011).

Une étude aCGH récente a été réalisée sur 29085 enfants atteints de retard développemental et 19584 individus témoins (Coe et al., 2014). L'analyse a permis l'identification de 2184 CNVs incluant 1348 délétions et 836 duplications dans 55 loci associés à des désordres génétiques incluant 40 régions flanquées par des SDs. Parmi ces loci, 19 étaient suspectées pathogènes et ont atteint une significativité dans cette étude. Sept loci étaient associés à des délétions, sept loci à des duplications, et cinq loci aux deux réarrangements. Les CNVs correspondaient notamment à des CNVs extrêmement rares comme la délétion 2q11.2 ou encore à des CNVs peu pathogènes comme les duplications 15q24 et 16p13.11. Une analyse comparée de l'ensemble des CNVs enrichis chez les patients et des données de séquençage d'exome provenant de nombreuses GWAS a permis d'identifier une liste de gènes candidats dont 26 furent séquencés chez 4716 patients atteints de retard développemental et de troubles du spectre autistique ainsi que 2193 témoins additionnels. L'analyse comparée des CNVs et des polymorphismes nucléotidiques a permis l'identification de dix gènes enrichis en mutations prédictives perte de fonction incluant les gènes *ADNP, DYRK1A, NRXN1, NRG3, SETBP1, ZMYND11, DNM3, CYFIP1, FOXP1* et *SCN2A*. Cette analyse a abouti notamment à l'identification du gène *ZMYND11* causatif du syndrome de délétion 10p15.3 (DeScipio et al., 2012) et de cinq mutations tronquées pour le gène *SETBP1* confirmant que l'haplo-insuffisance de ce gène suffit à causer le syndrome de délétion 18q12.3 (Marseglia et al., 2012).

6.2 Association aux troubles du spectre autistique

Les troubles du spectre autistique (ASD pour Autism Spectrum Disorder) incluent un ensemble hétérogène d'anomalies comportementales caractérisées par une réduction de l'interaction sociale, des déficits de la communication verbale et non verbale, ainsi que des comportements et activités répétitifs et restreints. Ces symptômes apparaissent avant l'âge de trois ans et sont regroupés sous l'appellation « triade autistique ». L'ASD inclut cinq conditions cliniques répertoriées : le désordre désintégratif de l'enfance, le syndrome de Rett, le trouble autistique, le trouble envahissant du développement non spécifié et le syndrome d'Asperger.

La prévalence de l'ASD a considérablement augmenté ces dernières années. En 2005, elle était estimée à 1/160 (Chakrabarti and Fombonne, 2005; Fombonne, 2005). Elle est estimée aujourd'hui à 1% de la population générale avec 4 fois plus d'hommes diagnostiqués que de femmes (Lai et al., 2014). L'anomalie génétique la plus commune est la duplication

maternelle du locus 15q11-q13 qui concerne 1 à 3% des cas (Abrahams and Geschwind, 2008; Marshall et al., 2008). La deuxième anomalie correspond aux réarrangements de la région 16p11.2 qui concerne environ 1% des cas (Fernandez et al., 2010; Weiss et al., 2008). Des mutations dans les gènes *FMR1*, *MECP2*, *TSC1*, *NF1* et *PTEN* ont été reportés chez des patients présentant un ASD syndromique. Les gènes participant à l'activité neuronale, notamment *CNTNAP2*, *SHANK3*, *NLGN3* et *NLGN4* contribuent également à l'étiologie de l'ASD (Abrahams and Geschwind, 2008). Globalement, la cause génétique de l'ASD n'est reconnue que dans 10 à 20% des cas ce qui est extrêmement faible et correspond à une grande variabilité génique. Bien que des centaines de gènes aient été associés à l'ASD, la majorité des candidats incluent de rares variations de pénétrance incomplète ce qui suggère que le développement de l'ASD serait dans la majorité des cas lié à l'accumulation de plusieurs polymorphismes.

En 2007, une étude aCGH fut réalisée sur 165 familles comprenant au moins un enfant atteint d'ASD et 99 familles témoins (Sebat et al., 2007). 17 CNVs *de novo* d'une taille variant de 0,1 à 12,2 Mb furent identifiés chez 16 individus incluant 14 cas d'ASD et 2 témoins. Les CNVs pathogènes incluaient trois duplications et douze délétions notamment une délétion 15q11-q13, une délétion 22q13.31-33 renfermant le gène *SHANK3*, une délétion 20p13 incluant le gène *OXT* codant pour l'ocytocine, ainsi qu'une délétion 16p11.2. L'étude permit de déterminer que les CNVs apparus *de novo* étaient fortement associés à l'autisme. Effectivement, la fréquence de mutations spontanées était de 10% pour les cas sporadiques, 3% pour les familles multiples et 1% pour les individus contrôles.

L'année suivante, une étude SNParray effectuée sur 427 patients ASD non apparentés confirma la présence de CNVs *de novo* chez 7% des cas sporadiques et chez 2% des cas de familles multiplex (Marshall et al., 2008). L'analyse permit l'identification de loci incluant de nouveaux gènes candidats pour l'ASD, notamment *ANKRD11*, *DLGAP2*, *DPP6*, *DPP10*, *DPYD*, *PCDH9* et *PTCHD*. Les auteurs identifièrent également deux loci précédemment associés à la déficience intellectuelle : le locus 15q24 (Sharp et al., 2007) et le locus 16p11.2 (Ballif et al., 2007; Ghebranious et al., 2007). Six cas d'ASD non apparentés furent notamment identifiés comme porteurs de délétions et de duplications pour la région 16p11.2 confirmant la présence de gènes critiques pour le développement du langage et de la sociabilité dans cette région.

Plus récemment, une étude SNParray effectuée sur 1084 familles incluant un ou plusieurs cas d'autisme et sur 896 contrôles a permis d'observer un enrichissement des larges délétions chez les cas d'ASD (Griswold et al., 2012). Cette étude confirme les résultats de Cooper *et al.* sur l'implication des larges CNVs dans les cas de désordres neurodéveloppementaux incluant des phénotypes autistiques. De nouveaux loci présentant des délétions d'une taille supérieure à 1Mb incluant les loci 2q22.1, 3p26.3, 4q12 et 14q23 ont été identifiés. De nouveaux gènes candidats impliqués dans la transmission gabaergique notamment *DBI, GABARAPL1* et *SLC6A11* ont été retrouvés dans différents CNVs. Les auteurs ont également identifié des gènes associés à d'autres conditions neuropsychiatriques incluant les gènes *GRIK2* et *GRID1* impliqués dans la transmission glutamatergique et associés à la schizophrénie (Bah et al., 2004; Shibata et al., 2002; Treutlein et al., 2009).

6.3 Association à la schizophrénie

La schizophrénie constitue un désordre psychiatrique sévère caractérisé par des hallucinations, des délires, un déficit cognitif et une apathie évoluant par phases aiguës alternées de phases quasi asymptomatiques (Sullivan et al., 2003; van Os and Kapur, 2009). La schizophrénie présente une prévalence de 0,4 à 0,7% dans la population générale (Saha et al., 2005). Les études épidémiologiques suggèrent une influence génétique majoritaire (73 à 90%), bien que l'héritage soit variable et ne suive pas une distribution mendélienne (Lohmueller et al., 2003). De nombreuses anomalies chromosomiques ont été détectées chez les individus affectés et les familles (MacIntyre et al., 2003). L'étude de ces CNVs aboutit à l'identification de gènes associés à un risque élevé de développer une schizophrénie incluant notamment *DISC1, TBX1, PDE4B* et *NPAS3* (Chubb et al., 2008; Kamnasaran et al., 2003).

A partir de 2006, des GWAS utilisant la technique aCGH identifièrent de nouveaux CNVs associés à la schizophrénie incluant des délétions et duplications Xq23, des délétions 3q13.12 (Moon et al., 2006), des délétions 7q34-36.1 comprenant le gène *CNTNAP2* (Friedman et al., 2008), une duplication 15q13.1 comprenant le gène *APBA2* et une délétion 2p16.3 comprenant le gène *NRXN1* (Kirov et al., 2008). Ce dernier gène codant pour la neuréxine 1-alpha avait précédemment été associé à l'autisme (Szatmari et al., 2007)et à la déficience intellectuelle (Friedman et al., 2006) confirmant l'interconnexion entre différentes conditions neuropsychiatriques.

En 2008, une étude aCGH réalisée sur 150 individus atteints de schizophrénie et 268 témoins permit d'étudier la distribution de CNVs d'une taille supérieure à 100 kb (Walsh et

al., 2008). Des délétions et duplications *de novo* chevauchant des séquences codantes furent retrouvées chez 5% des contrôles contre 15% des individus schizophrènes. La délétion 2p16.3 incluant le gène *NRXN1* a été identifiée chez deux jumeaux schizophrènes. De manière intéressante, la duplication de la région 16p11.2 précédemment associée à l'autisme (Kumar et al., 2008; Sebat et al., 2007; Weiss et al., 2008) fut retrouvée chez deux cas schizophrènes. Un enrichissement de gènes impliqués dans le développement cérébral (signalisation par les neurégulines, intégrines et ERK/MAPK, guidage axonal, LTP synaptique) fut retrouvé dans les CNVs spécifiques aux patients schizophrènes. Ces gènes incluaient notamment *ERBB4, MAGI2, DLG2, PRKCD, PRKAG2, PTK2, CAV1, GRM7, SLC1A3, PTPRM* et *LAMA1*.

Deux études utilisant la technique SNParray réalisée sur 3391 individus schizophrènes et 3181 témoins pour le consortium international de la schizophrénie (Consortium, 2008) et sur 1433 + 3285 individus schizophrènes ainsi que 33250 + 7951 témoins pour Stefansson *et al.*(Stefansson et al., 2008) identifièrent des délétions aux loci 15q13.3 et 1q21.1 associées de manière significative à la schizophrénie. Les deux groupes confirmèrent également l'implication de la duplication 15q11.2 et de la délétion 22q11.2 associée précédemment au syndrome de DiGeorge ainsi qu'au syndrome vélocardiofacial pour lequel 30% des patients présentent des symptômes psychotiques (Karayiorgou et al., 1995; Williams et al., 2006).

6.4 Association à l'épilepsie

L'épilepsie est un désordre neuropsychiatrique présentant une prévalence de 0,4 à 1% (Hauser et al., 1996). L'épilepsie présente un large spectre de symptômes cliniques et comprend plus de cinquante syndromes épileptiques qui peuvent être divisés en deux grandes catégories : les épilepsies idiopathiques et symptomatiques. Les épilepsies symptomatiques concernent 20 à 30% des cas et sont conséquentes à un facteur identifié comme une anomalie métabolique, un trauma ou une tumeur (Hauser et al., 1993; Jallon et al., 2001). Les épilepsies idiopathiques généralisées (IGEs pour Idiopathic Generalized Epilepsies) apparaissent en l'absence de facteur causatif identifié (Helbig et al., 2008; Kjeldsen et al., 2003).

Les IGEs sont caractérisées par la récurrence de crises généralisées spontanées sans lésion cérébrale détectable et sans anomalie métabolique. L'absence épileptique, l'épilepsie myoclonique juvénile et les épilepsies avec des crises toniques-cloniques généralisées sont les IGEs les plus communes (1989). Lors des IGEs, l'électroencéphalogramme (EEG) est marquée par des pointes-ondes sur la totalité du cortex cérébral ce qui reflète un état d'hyperexcitation synchrone des boucles thalamo-corticales (Blumenfeld, 2005). Les facteurs

génétiques ont un rôle prédominant dans l'étiologie des IGEs et l'héritabilité est estimée à plus de 80%. Les approches de génétique moléculaire ont identifié les gènes associés à de nombreuses IGEs mendéliennes dont la majorité code pour des sous-unités de canaux ioniques incluant *CHRNA4, CHRNA2, CHRNB2, GABRG2, GABRA1, KCNQ2, KCNQ3, SCN1B, SCN1A, STXBP1* mais également *ARX, CDKL5, LGI1, PCDH19, SLC2A1, SPTAN1* et *STXBP1* (Hildebrand et al., 2013; Reid et al., 2009).

En 2008, une étude aCGH menée sur 757 patients présentant une déficience intellectuelle et/ou des anomalies congénitales aboutit à l'identification de deux individus non-apparentés présentant une délétion *de novo* au locus 15q13.3 (Sharp et al., 2008). Les patients présentaient une déficience intellectuelle légère à modérée, une dysmorphie faciale et un électroencéphalogramme anormal. Le génotypage de la région par PCR quantitative de 1040 individus supplémentaires présentant une déficience intellectuelle d'étiologie inconnue aboutit à la détection de neuf individus porteurs de délétions au locus 15q13.3 dont sept présentaient une IGE et/ou un EEG anormal. Une région critique de 1,5 Mb fut identifiée entre les points de cassure BP4 et BP5 incluant six gènes dont le gène candidat *CHRNA7* codant pour la sous-unité α7 du récepteur nicotinique de l'acétylcholine.

L'implication de la région 15q13.3 fut confirmée en 2009 par une étude aCGH qui aboutit à l'identification de douze délétions 15q13.3 sur 1223 individus présentant une IGE et d'aucun réarrangement du locus 15q13.3 parmi les 3699 contrôles (Helbig et al., 2009). La même année, une deuxième étude indépendante identifia sept délétions 15q13.3 sur 539 individus présentant une IGE et aucun CNV parmi les 194 contrôles (Dibbens et al., 2009). La délétion 15q13.3, associée également à la schizophrénie (Consortium, 2008; Stefansson et al., 2008) concerne environ 1% des cas de déficience intellectuelle associée à des IGEs.

Plus récemment, une étude SNParray a évalué la proportion des délétions aux loci 1q21.1, 15q11.2, 15q13.3, 16p11.2, 16p13.11 et 22q11.2 associées aux troubles neuro-psychiatriques (déficience intellectuelle, ASD et schizophrénie) chez des patients atteints d'IGEs (de Kovel et al., 2010). Sur 1234 patients présentant une IGE, vingt-deux étaient porteurs d'une des cinq délétions (1,8%) contre cinq porteurs sur 3022 témoins (0,3%). Un enrichissement significatif des délétions 15q11.2 (douze patients porteurs et six témoins), 15q13.3 (neuf patients porteurs et aucun témoin) et 16p13.11 (six patients porteurs et deux témoins) fut observé chez les patients présentant une IGE. En 2013, une étude confirma que les délétions pour ces trois loci concernent environ 3% des cas de déficience intellectuelle associée à l'IGE (Mullen et al., 2013).

6.5 Prépondérance des facteurs de risque génétique non-spécifiques

Les résultats des études GWAS décrites ci-dessus mettent en lumière la grande complexité des désordres neuropsychiatriques dont les facteurs génétiques incluent une multitude de rares variations à la pénétrance variable. Les analyses indiquent néanmoins que les larges délétions sont fortement associées aux maladies développementales. De nombreux polymorphismes aboutissant à l'haplo-insuffisance mais aussi à la surexpression de gènes spécifiques peuvent augmenter le risque de développer plusieurs conditions neuropsychiatriques et constituent donc des facteurs de risque génétique non-spécifiques (Zhu et al., 2014) (Figure 11). Dans des études récentes, les relations entre ASD, schizophrénie, trouble de l'attention avec hyperactivité, trouble bipolaire et dépression ont clairement été démontrées par l'étude de SNPs communs (Consortium, 2013; Lee et al., 2013).

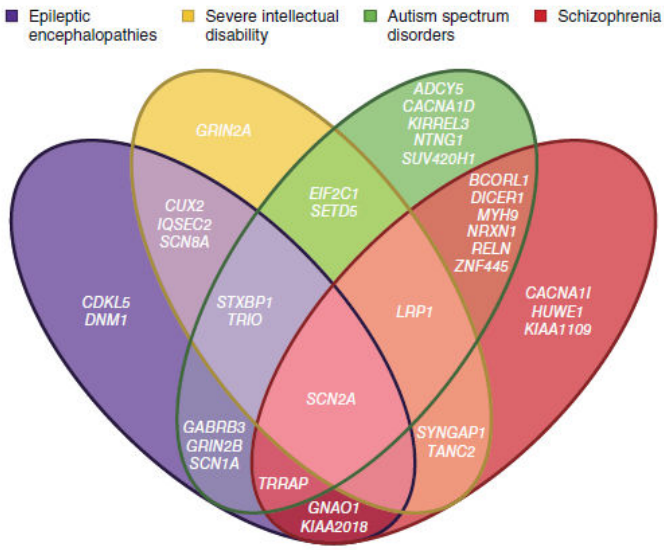


Figure 11 : Chevauchement des gènes associés aux désordres neuropsychiatriques

Diagramme de Venn représentant le chevauchement de gènes affectés par des mutations *de novo* associées aux quatre désordres neuropsychiatriques principaux. (Tiré de Zhu *et al.*, 2013).

La plupart des CNVs associés à plusieurs désordres neuropsychiatriques, notamment les délétions et duplications des loci 2q13, 13q12, 15q11.2, 15q13.3, 16p11.2, 16p13.11 et 22q11.2 sont responsables de maladies syndromiques complexes caractérisées par une variabilité et une pénétrance incomplète des phénotypes associés (Cooper et al., 2011). Nous allons nous intéresser plus particulièrement à ces maladies syndromiques dans la troisième partie de cette introduction.

II Les syndromes de délétions et de duplications associés à la déficience intellectuelle

1. Premiers syndromes caractérisés

Un grand nombre de CNVs incluant des régions codantes est associé à un risque accru de développer des maladies syndromiques. La plus grande proportion de ces syndromes est associée à de larges CNVs récurrents, majoritairement des délétions, d'une taille comprise entre plusieurs centaines de kb et quelques Mb générés par des mécanismes NAHR entre SDs (Cooper et al., 2011; Stankiewicz and Lupski, 2010). Une proportion plus faible de maladies syndromiques est associée à des CNVs non-récurrents principalement induits par les mécanismes NHEJ, MMEJ et MMBIR (Hastings et al., 2009b). Ces CNVs comprennent majoritairement des délétions télomériques (Bonaglia et al., 2011; Rodríguez-Caballero et al., 2010; Wu et al., 1999).

Les premiers syndromes de délétions et de duplications découverts dans les années 1970 étaient facilement identifiables par des traits phénotypiques caractéristiques. Ces syndromes étaient associés à des CNVs assez larges pour être détectés par les méthodes cytogénétiques de l'époque. Ils incluent entre autres les syndromes de Jacobsen (Jacobsen et al., 1973), de délétion 9p (Alfi et al., 1976), de duplication 10q (Yunis and Sanchez, 1974) et de duplication 15q (Fujimoto et al., 1974).

Néanmoins, de nombreux syndromes ont été décrits avant l'identification de leur cause génétique et portent le nom des cliniciens ayant établi la caractérisation d'un petit nombre de patients présentant une symptomatique similaire. Les symptômes incluaient systématiquement une déficience intellectuelle ainsi qu'une dysmorphie faciale caractéristique permettant une identification des malades plus facile. On peut citer le syndrome de Williams (WS pour Williams Syndrome) caractérisé en 1961 chez 4 enfants présentant une sténose aortique supravalvulaire, le syndrome de Sotos (SS) décrit en 1964 chez 5 enfants présentant une croissance rapide et une acromégalie, le syndrome d'Angelman (AS) caractérisé en 1965 chez 3 enfants présentant des positions anormales des membres ou encore le syndrome de DiGeorge (DGS) décrit en 1965 chez 4 enfants avec une hypoplasie du thymus, une cardiopathie et une endocrinopathie avec hypocalcémie (Greenberg, 1993; Hart, 2008; SOTOS et al., 1964; WILLIAMS et al., 1961).

Plusieurs années plus tard, l'amélioration des techniques de cytogénétique et de biologie moléculaire a permis d'aboutir à l'identification des différents CNVs associés à de nombreux syndromes décrits dans la littérature médicale. Le DGS a été associé à la délétion 22q11.2 au début des années 1980 (de la Chapelle et al., 1981; Kelley et al., 1982). En 1989, le syndrome de Prader-Willi (PWS) et le syndrome AS ont été associés à une délétion commune du locus 15q11q13 d'origine maternelle pour le syndrome AS et d'origine paternelle pour le syndrome PWS (Knoll et al., 1989). Le syndrome SS a été relié à la délétion de la région 5q35 à partir de 1994 (Maroun et al., 1994). Quant au syndrome WS, il a été associé à la délétion de la région 7q11.23 en 1995 (Lowery et al., 1995; Nickerson et al., 1995). De nombreux autres syndromes associés aux CNVs ont été décrits dans la littérature comme les syndromes de Potocki-Lupski causé par la duplication de la région 17p11.2 (Potocki et al., 2000) et de Smith-Magenis dû à la délétion 17p11.2 (Smith et al., 1986), de Miller-Dieker causé par la délétion de la région 17p13.3 (Stratton et al., 1984). Les régions associées à ces syndromes de délétions et de duplications chevauchent toutes des régions géniques.

Les syndromes décrits ci-dessus sont caractérisés par une pénétrance totale de la déficience intellectuelle et par une faible variabilité des autres phénotypes associés ce qui a facilité leur description précoce. A partir de 2005, les GWAS aboutirent à la découverte de nombreux autres syndromes de pénétrance complète comme le syndrome de Koolen-de Vries (Koolen et al., 2006). Les GWAS permirent également l'identification de syndromes de pénétrance incomplète à la symptomatique plus variable comme les syndromes de délétion et de duplication de la région 16p11.2 (Marshall et al., 2008; Weiss et al., 2008). Au contraire des syndromes de pénétrance totale dont les CNVs apparaissent majoritairement *de novo*, les CNVs associés aux syndromes de pénétrance incomplète sont majoritairement hérités de l'un des deux parents et plus particulièrement de la mère. L'implication de polymorphismes secondaires modulant la pathogénicité des CNVs primaires a été démontrée pour certains de ces syndromes (Girirajan and Eichler, 2010; Girirajan et al., 2012).

Dans les paragraphes qui suivent, nous nous intéresserons plus particulièrement aux délétions partielles du chromosome 21, ainsi qu'aux syndromes de délétion et de duplication des régions 16p11.2 et 17q21.31 qui font l'objet du travail de thèse présenté dans ce manuscrit.

2. La monosomie 21 partielle

La monosomie 21 est une anomalie chromosomique due à une délétion totale ou partielle du chromosome 21. Cette anomalie est connue depuis 1964, année où Jérôme Lejeune décrit le premier cas clinique (Lejeune et al., 1964). Durant de nombreuses années, des cas cliniques de monosomie 21 complète ont été décrits dans la littérature scientifique (Chang et al., 1992; Dziuba et al., 1976; Garzicić et al., 1988; Kaneko et al., 1975). Néanmoins, des études récentes indiquent que la monosomie totale du chromosome 21 sans translocation chromosomique ni mosaïcisme conduit à des anomalies cardiaques létales (Burgess et al., 2014; Toral-Lopez et al., 2012). La monosomie 21 partielle (PM21) se caractérise par la présence d'une délétion partielle variable d'un segment du bras long du chromosome 21 (HSA21). Moins de 50 cas ont été décrits dans la littérature ce qui a nécessité leur rassemblement sous une classification syndromique commune.

Selon la taille et la localisation des délétions, la PM21 est associée à une large hétérogénéité de signes cliniques. Les symptômes les plus caractéristiques incluent un retard de croissance intra-utérin, une diminution de la taille, des dysmorphies cranio-faciales (microcéphalie, nez proéminent, micrognathie, fente labio-palatine, fente des commissures buccales inversées), une déficience intellectuelle, des déficits musculaires, des malformations cardiaques et squelettiques (Lindstrand et al., 2010; Lyle et al., 2009; Roberson et al., 2011).

En 1995, une première étude comparative de six patients présentant des délétions de différentes régions du HSA21 permit de constituer la première cartographie génétique de vingt-trois signes cliniques associés à la monosomie 21 (Chettouh et al., 1995). Cette étude aboutit à l'identification de la région *APP-SOD1* contribuant fortement aux signes cliniques de la PM21. Effectivement, la région regroupe dix des vingt-trois signes cliniques incluant la déficience intellectuelle et l'hypertonie potentiellement associés à l'haplo-insuffisance d'un ou plusieurs gène(s) de la région *APP-SOD1*.

En 2009, une étude aCGH menée sur onze patients détermina trois grandes régions du bras long du HSA21, chacune associée à une symptomatique particulière (Lyle et al., 2009) (Figure 12). La région 1, du centromère à 31,2Mb (~50 gènes), est associée à une déficience intellectuelle, des défauts musculaires et des malformations cranio-faciales. La région 2, de 31,2 à 36Mb (~80 gènes), est associée à des phénotypes extrêmement sévères incluant des malformations cardiaques le plus souvent létales. La région 3, de 36Mb au télomère (~130 gènes), est associée à des phénotypes plus nuancés incluant une déficience intellectuelle légère et des dysmorphies faciales.

Plus récemment, une méta-analyse a été effectuée à partir de cinq cohortes différentes englobant un nombre total de trente-six patients porteurs de PM21 (Roberson et al., 2011) (Figure 12). L'analyse a confirmé les symptomatiques associées aux régions précédemment décrites. Différents patients présentant des délétions partielles pour la région 2 ont été identifiés, notamment trois cas provenant de la banque de données DECIPHER (Firth et al., 2009). Néanmoins, aucun patient ne présentait de délétion recouvrant l'ensemble de la région 2 ce qui suggère que l'haplo-insuffisance des 80 gènes situés dans cette région est incompatible avec la survie humaine.

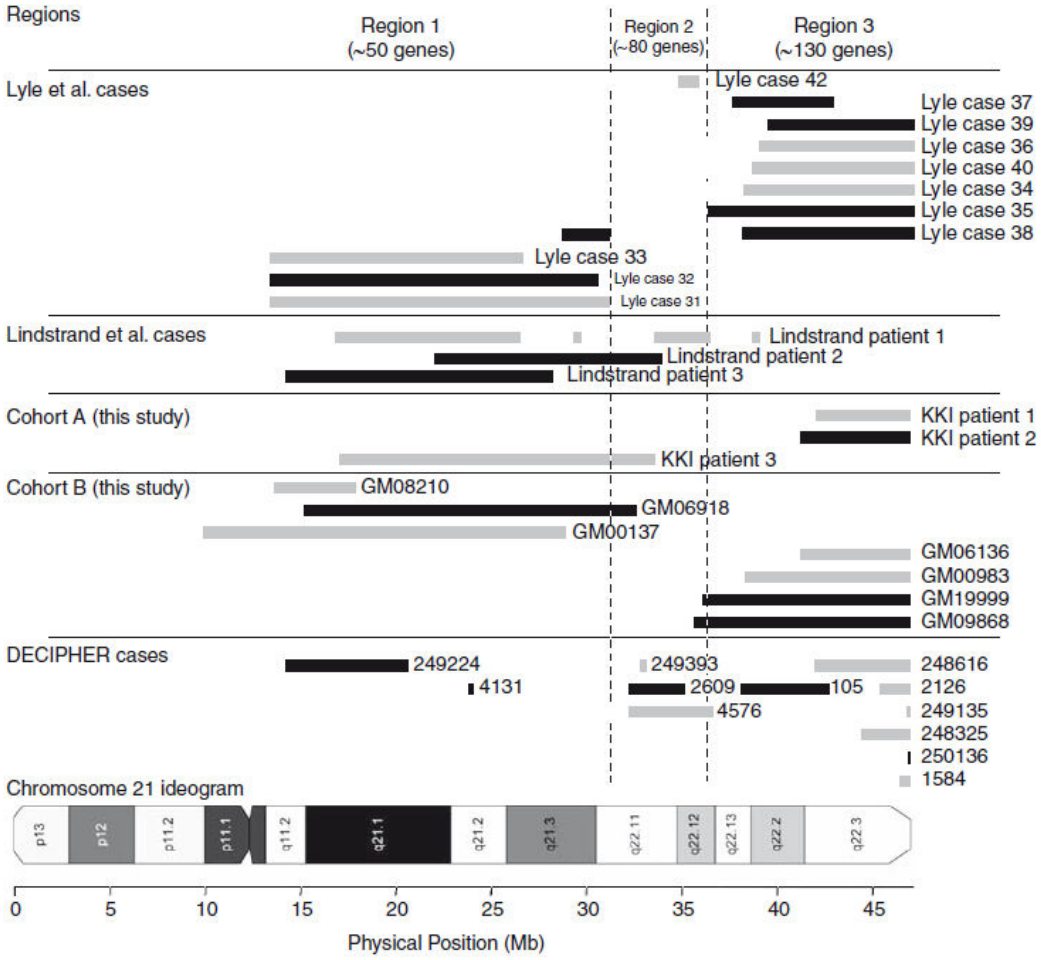


Figure 12 : Régions génomiques d'association avec la symptomatique de la monosomie 21 partielle et délétions des patients caractérisés

Délimitation des trois régions du bras long du HSA21 proposé par Lyle *et al*. et schématisation des délétions des patients atteints de PM21 caractérisés dans différentes études. Les bandes grises représentent les délétions des patients présentant d'autres anomalies chromosomiques. Les bandes noires représentent les délétions des patients ne présentant pas d'autre anomalie. La majorité des patients présentent des délétions dans la région 3 associée à une symptomatique modérée. Au contraire, une minorité de patients présente des délétions dans la région 2, associée à de sévères phénotypes cardiaques. (Tiré de Roberson *et al.*, 2011).

3. Les syndromes de délétions/duplications 16p11.2

3.1 Présentation des syndromes

Le locus 16p11.2 est l'un des plus fréquemment associé aux troubles neuro-développementaux (Grayton et al., 2012). Cette région péricentromérique du chromosome 16 contient de nombreux SDs prédisposant aux NAHR (Figure 13a). Les réarrangements les plus fréquemment identifiés se produisent entre les SDs proximaux délimitant une région de ~ 600 kb définie par les points de cassure BP4 et BP5 contenant 34 gènes : *BOLA2, SLX1B, SULT1A4, SPN, QPRT, C16orf54, ZG16, KIF22, MAZ, PRRT2, PAGR1, MVP, CDIPT, CDIPT-AS, SEZ6L2, ASPHD1, KCTD13, TMEM219, TAOK2, HIRIP3, INO80E, DOC2A, C16orf92, FAM57B, ALDOA, PPP4C, TBX6, YPEL3, GDPD3, MAPK3, CORO1A, BOLA2B, SLX1A, SULT1A3*. (Figure 13b).

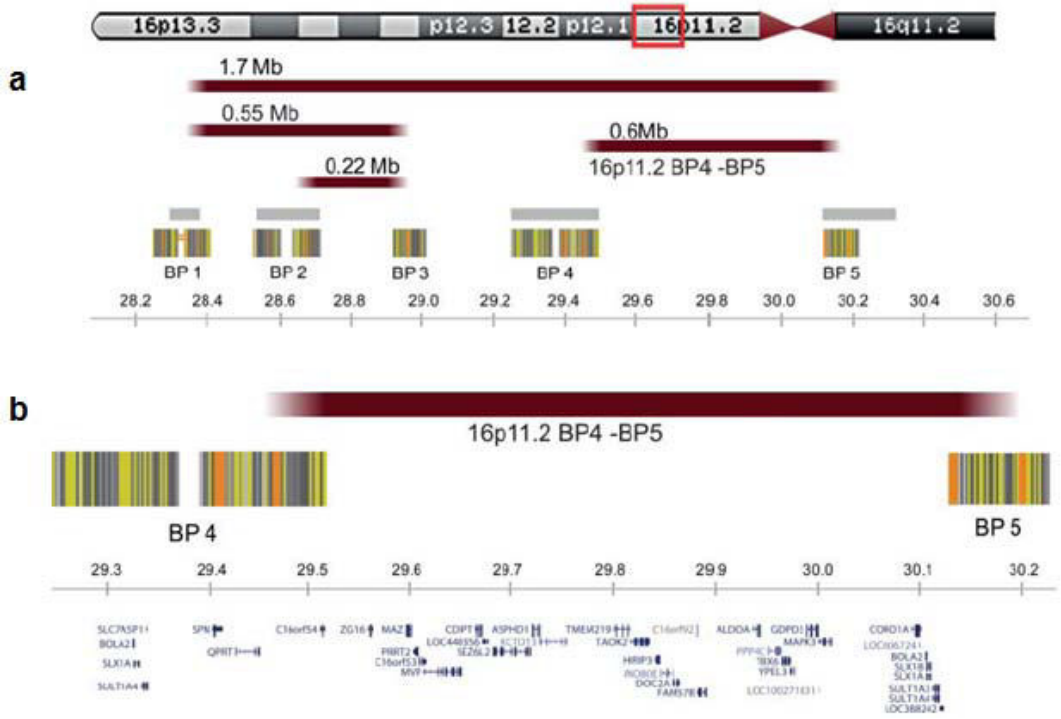


Figure 13 : Le locus 16p11.2

a. Représentation du locus 16p11.2 et des différents SDs.prédisposant aux réarrangements NAHR. Les quatre réarrangements récurrents de la région sont schématisés par des bandes bordeaux. Les réarrangements les plus fréquents ont lieux entre les points de cassure proximaux BP4 et BP5.
b. La région BP4-BP5, d'une taille de 600 kb comprend 34 gènes. Les positions génomiques proviennent du browser hg18/NCBI36 (Adapté de Zufferey *et al.*, 2012).

Les délétions et duplications de la région 16p11.2 BP4-BP5 ont une prévalence estimée à 1/1500 dans la population générale (Jacquemont et al., 2011) et concernent 1 à 1,5% des patients atteints d'ID (Cooper et al., 2011; Weiss et al., 2008). Les réarrangements de la région 16p11.2 sont associés à un large spectre de désordres neuropsychiatriques, à des variations de l'indice de masse corporelle (BMI pour Body Mass Index) associées à des troubles de la conduite alimentaire ainsi que des anomalies de croissance de la boîte crânienne. De manière intéressante, alors que la délétion et la duplication de la région sont associées à la déficience intellectuelle et à l'autisme, des phénotypes opposés sont observés pour le volume crânien et l'indice de masse corporelle : macrocéphalie et obésité pour les porteurs de la délétion, microcéphalie et insuffisance pondérale pour les porteurs de la duplication. Ces résultats suggèrent l'implication de gènes sensibles au dosage génique dans les phénotypes associés aux réarrangements de la région BP4-BP5.

3.2 Caractérisation Clinique

Les réarrangements de la région 16p11.2 BP4-BP5 furent initialement associés à l'ID et à l'ASD. En 2007, Sebat *et al.* identifièrent un porteur de la délétion sur une cohorte de 118 patients avec ASD. En 2008, une étude similaire identifia deux porteurs de la délétion dans un groupe de 180 patients (Kumar et al., 2008). La même année, deux études indépendantes associèrent la délétion mais aussi la duplication réciproque à une fréquence similaire de 1% des cas d'ASD (Marshall et al., 2008; Weiss et al., 2008). Une étude portant sur les polymorphismes des gènes de la région mit en évidence une association significative entre le polymorphisme R386H du gène *SEZ6L2* et l'ASD (Kumar et al., 2009). De manière intéressante, une délétion atypique de 118 kb incluant les gènes *MVP, CDIPT, SEZ6L2, ASPHD1* et *KCTD13* fut identifiée parmi les membres d'une famille de trois générations présentant un ASD ou des traits autistiques (Crepel et al., 2011). Sans exclure la potentielle implication des autres gènes de la région, *MVP, CDIPT, SEZ6L2, ASPHD1* et *KCTD13* sont donc des gènes candidats particulièrement fort expliquer l'ASD associée aux réarrangements de la région 16p11.2.

En addition de l'ASD, la duplication 16p11.2 BP4-BP5 est également associée à la schizophrénie et à d'autres désordres neuropsychiatriques. En 2008, Walsh *et al.* identifièrent deux porteurs de la duplication dans un groupe de 150 patients schizophrènes. L'année suivante, une étude focalisée sur les réarrangements de la région 16p11.2 mit en évidence l'association significative de la duplication et de la schizophrénie (McCarthy et al., 2009). Les

auteurs identifièrent 21 porteurs de la duplication sur 4551 cas (0,46%) et seulement 2 porteurs sur 6391 contrôles (0,03%). Dans la même étude, une méta-analyse de données GWAS pour de multiples désordres psychiatriques permit l'association de la duplication 16p11.2 à la schizophrénie, à l'autisme et au trouble bipolaire alors que la délétion était associée à l'autisme et aux désordres développementaux. Une augmentation significative du périmètre crânien fut également observée chez les patients porteurs de la délétion par rapport aux patients porteurs de la duplication.

Les réarrangements de la région 16p11.2 BP4-BP5 sont également associés à l'épilepsie. Dès 2007, la délétion 16p11.2 fut associée à l'épilepsie chez un patient présentant une anomalie de la valve aortique, une déficience intellectuelle et des crises épileptiques (Ghebranious et al., 2007). Plus récemment une étude menée sur 285 porteurs de la délétion a indiqué que 24% des patients présentaient des crises épileptiques (Zufferey et al., 2012). En 2010, la duplication réciproque fut associée aux IGEs (Bedoyan et al., 2010). Plus récemment, la duplication a également été associée à l'épilepsie Rolandique (Reinthaler et al., 2014).

La caractérisation clinique de vingt-sept patients porteurs de la délétion 16p11.2 BP4-BP5 et dix-huit patients porteurs de la duplication réciproque a été réalisée en 2010 (Shinawi et al., 2010).

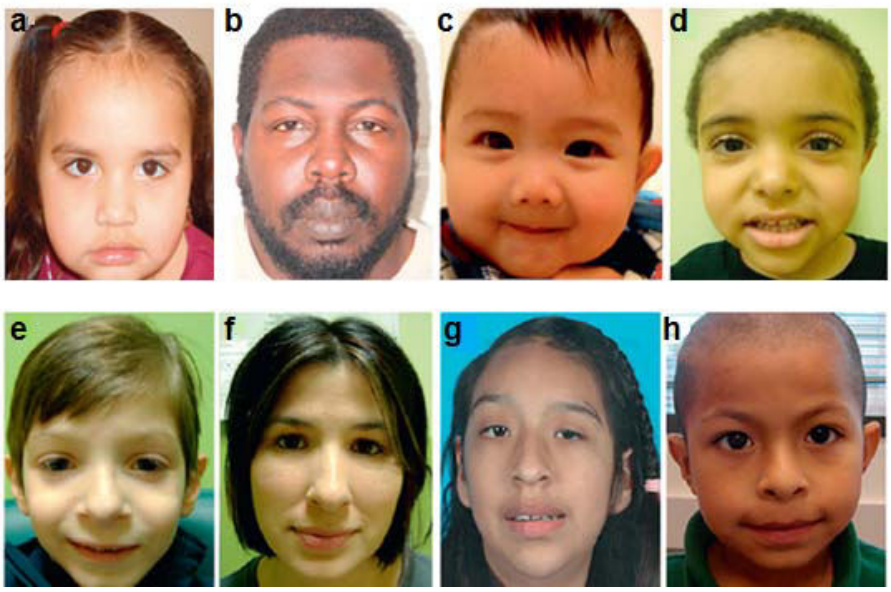


Figure 14 : Patients porteurs de réarrangements pour la région 16p11.2 BP4-BP5

a-d. Patients porteurs de la délétion 16p11.2 présentant une macrocéphalie.
e-h. Patients porteurs de la duplication 16p11.2 présentant une microcéphalie.
(Adapté de Shinawi *et al.*, 2010).

Pour les deux réarrangements, des cas récurrents de retard psychomoteur, de désordres comportementaux (principalement ADHD), d'épilepsie et d'anomalies congénitales ont été observés. Les porteurs de la délétion présentaient une macrocéphalie alors que les porteurs de la duplication présentaient une microcéphalie (Figure 14). Une diminution de la taille des patients porteurs de la délétion était également observée.

En plus des désordres neuropsychiatriques et des changements du volume crânien, les réarrangements de la région 16p11.2 engendrent également d'importantes altérations du BMI des patients. La délétion fut dans un premier temps associée à l'obésité. En 2010, une étude aCGH réalisée sur 312 patients présentant une obésité associée à un retard développemental a identifié neuf patients (2,9%) porteurs d'une délétion 16p11.2 (Walters et al., 2010). Une étude similaire a été réalisée sur 3947 patients présentant un retard développemental mais sans sélection pour l'obésité et seulement vingt-deux cas (0,6%) porteurs d'une délétion similaire ont été identifiés ce qui est significativement plus faible que pour le groupe précédent. L'analyse clinique des vingt-deux patients conclut que la délétion 16p11.2 était associée à un fort phénotype d'obésité chez l'adulte avec un phénotype plus variable à l'enfance. Au moins neuf cas sur vingt-deux présentaient une hyperphagie identifiée. Les auteurs analysèrent les données de GWAS de 16053 individus provenant de huit cohortes européennes incluant des témoins et des cas d'obésité/obésité morbide. L'analyse a permis l'identification de dix-neuf porteurs de la délétion 16p11.2. A l'exception d'un individu diabétique, tous les porteurs adultes étaient obèses ou obèses morbides. Aucun des dix-neuf individus porteurs ne présentait de déficience intellectuelle mais quatre sujets étaient atteints d'hyperphagie.

En 2011, une étude similaire a été réalisée à partir d'un groupe de 106 individus non apparentés porteurs de la duplication 16p11.2 (Jacquemont et al., 2011). Cette cohorte incluait soixante-treize patients présentant un retard intellectuel ou développemental, quatre patients atteints de troubles psychiatriques et vingt-neuf individus provenant de la population générale. Onze patients médicalisés présentaient un comportement alimentaire sélectif et restrictif. Les données de masses, tailles, BMIs et volumes crâniens des porteurs de la duplication ont été comparées aux données de référence pour une population d'âge et de sexe similaires. La duplication fut associée à une réduction du volume crânien ainsi qu'à une diminution de la masse corporelle et du BMI.

Les délétions et duplications de la région 16p11.2 BP4-BP5 sont associées à un large spectre de désordres neuropsychiatriques incluant l'ID, l'ASD, l'épilepsie, la schizophrénie et le trouble bipolaire. La caractérisation des patients indique que les réarrangements de la

région 16p11.2 BP4-BP5 sont également associés à un volume crânien et un BMI opposés : macrocéphalie et obésité pour les porteurs de la délétion, microcéphalie et insuffisance pondérale pour les porteurs de la duplication ce qui suggère l'implication de gènes sensibles au dosage génique dans l'apparition des phénotypes décrits.

3.3 Gènes de la région

La région 16p11.2 BP4-BP5 contenant 34 gènes, nous nous sommes limités à la description des principaux gènes candidats pour les phénotypes associés aux délétions et duplications. Les résultats présentés ci-dessous intègrent des études de génétique humaine aboutissant à l'identification de mutations pour des gènes de la région chez des patients mais aussi des études transrationnelles réalisées chez la souris et le poisson zèbre.

Le gène *PRRT2* (Proline-Rich Transmembrane Protein 2) code pour une protéine transmembranaire à la fonction mal connue. Des études ont néanmoins montré l'interaction de PRRT2 avec la protéine SNAP25 impliquée dans la fusion des vésicules synaptiques et la libération de neurotransmetteurs (Jarvis and Zamponi, 2005). Des mutations pour le gène *PRRT2* ont été associées à la dyskinésie kinésigénique paroxystique, à l'épilepsie infantile et à la migraine hémiplégique (Chen et al., 2011; de Vries et al., 2012; Marini et al., 2012; Nobile and Striano, 2014; Pelzer et al., 2014). Ces résultats suggèrent une implication du gène *PRRT2* dans la modulation fonctionnelle de canaux ioniques. *PRRT2* est donc un gène candidat important pour expliquer les phénotypes épileptiques.

Le gène *TBX6* (T-Box 6) code pour un facteur de transcription impliqué dans la formation du mésoderme (Papapetrou et al., 1999; Takemoto et al., 2011; Yasuhiko et al., 2008). Des mutations pour le gène *TBX6* on été associées à la scoliose et à la dystose spondylo-costale (Fei et al., 2010; Sparrow et al., 2013). *TBX6* est un gène candidat pour expliquer les malformations vertébrales associées à la délétion (Shen et al., 2011; Shimojima et al., 2009) ainsi que la syringomyélie associée à la délétion et à la duplication de la région 16p11.2 BP4-BP5 (Schaaf et al., 2011).

Le gène *KIF22* (Kinesin Family member 22) code pour une protéine motrice associée aux microtubules. Des mutations pour ce gène sont impliquées dans la dysplasie spondylo-métaphysaire (Boyden et al., 2011; Min et al., 2011). *KIF22* pourrait également être impliqué dans les malformations vertébrales associées à la délétion 16p11.2 BP4-BP5.

Le gène *SEZ6L2* (Seizure related 6 homolog (mouse)-Like 2) code pour une protéine de surface qui présente une importante identité de séquence avec le gène SRPX2 (Sushi-Repeat-

containing Protein, X-linked 2) associé à l'épilepsie et aux désordres du langage (Roll et al., 2006). De plus, le SNP R386H du gène *SEZ6L2* est associé l'ASD (Kumar et al., 2009). Ce gène est potentiellement impliqué dans le développement de l'ID, de l'ASD et de l'épilepsie des patients porteurs des réarrangements pour la région 16p11.2 BP4-BP5.

Le gène *DOC2A* (Double C2-like domains Alpha) code pour une protéine cytoplasmique qui se lie aux phospholipides en réponse au calcium (Ca^{2+}) (Kojima et al., 1996). DOC2A interagit avec les protéines MUNC13 et MUNC18 impliquées dans l'exocytose des vésicules de sécrétion et la libération des neurotransmetteurs (Mochida et al., 1998). Etant impliqué dans la transmission synaptique, *DOC2A* est un gène candidat pour les troubles neuropsychiatriques des syndromes 16p11.2.

Le gène *MVP* (Major Vault Protein) code pour une protéine impliquée dans la particule ribonucléoprotéique cytoplasmique Vault de fonction mal connue (Lara et al., 2011). Des SNPs et une surexpression du gène *MVP* ont été associés à la résistance au traitement de nombreux cancers (Izquierdo et al., 1996; Laurençot et al., 1997; Mossink et al., 2003) et de l'épilepsie réfractaire (Liu et al., 2011a; Sisodiya et al., 2003; van Vliet et al., 2004).

Le gène *MAPK3* (Mitogen-Activated Kinase 3) ou *ERK1* (Extracellular signal-Regulated Kinase 1) code pour une protéine sérine/thréonine kinase impliquée dans la voie Ras/MAPK. Cette voie de signalisation intracellulaire joue un rôle important dans la régulation de la prolifération, de la survie, de la différenciation, de la migration cellulaire et de l'angiogenèse. La suractivation de cette voie est impliquée dans de nombreux cancers. Bien que les cancers ne soient pas associés aux réarrangements de la région 16p11.2, il nous paraissait important de signaler la dérégulation de ce gène primordial pour de nombreuses fonctions biologiques.

Le gène *KCTD13* (Potassium Channel Tetramerization Domain containing 13) code pour la protéine PDIP1 (pour Polymerase Delta-Interacting Protein 1) interagissant avec PCNA (pour Proliferating Cell Nuclear Antigen) qui est le facteur de processivité de l'ADN polymerase delta. PCNA est impliqué dans la régulation du cycle cellulaire durant la neurogenèse (He et al., 2001). Une étude menée chez le poisson zèbre a récemment permis d'associer le gène *KCTD13* à des anomalies neuroanatomiques semblables aux altérations volumiques crâniennes des patients (Golzio et al., 2012). L'injection embryonnaire d'ARNm humain induit une diminution de la taille de la tête de l'animal alors que l'injection de morpholino dirigé contre l'orthologue du poisson induit une augmentation de la taille de la tête. Les auteurs ont également montré que la sous-expression du gène induisait une

augmentation du la prolifération cellulaire des neurones de poisson zèbre et de souris alors que sa surexpression induisait un ralentissement de la prolifération cellulaire couplé à une augmentation de l'apoptose. A ce jour, aucune mutation spécifique au gène *KCTD13* n'a été associé à une condition pathologique humaine. C'est néanmoins un gène très intéressant pour les phénotypes de macrocéphale et de microcéphalie des patients porteurs respectivement de la délétion et de la duplication 16p11.2 BP4-BP5.

Le gène *TAOK2* (Thousand-And-One-amino acid Kinase 2) code pour une protéine sérine/thréonine kinase. TAOK2 active les protéines MAPK (Mitogen-Activated Protein Kinase) et régule la transcription génique. TAOK2 interagit avec NRP1 (Neuropilin 1) impliqué dans le guidage axonal et l'arborisation dendritique (Chen et al., 2014). Des expériences de répression et de surexpression du gène dans des cultures de neurones corticaux de souris ont montré l'implication de la protéine Taok2 dans la formation dendritique à travers l'activation de la protéine sérine/thréonine kinase Jnk1 (c-Jun N-terminal Kinase 1) (de Anda et al., 2012). Comme pour *KCTD13*, aucune mutation spécifique au gène *TAOK2* n'a été associé à une condition pathologique humaine. Le gène *TAOK2* est néanmoins potentiellement impliqué dans la formation synaptique et représente donc un gène candidat pour les troubles neuropsychiatriques associés aux réarrangements de la région 16p11.2.

La région 16p11.2 BP4-BP5 contient donc un grand nombre de gènes candidats pour les troubles neuropsychiatriques associés aux réarrangements étudiés. Ces gènes incluent notamment *DOC2A*, *KCTD13*, *PRRT2*, *SEZ6L2* ainsi que *TAOK2.*

4. Les syndromes de délétion/duplication 17q21.31

4.1 Présentation des syndromes

Le syndrome de délétion de la région 17q21.31 ou syndrome de Koolen-De Vries est causé par la délétion récurrente d'un fragment de 500 à 650 kb comprenant cinq gènes codants : *CRHR1, SPPL2C, MAPT, STH* et *KANSL1* (Figure 15a). Sa prévalence est estimée à 1/16000 dans la population générale. La délétion 17q21.31 a initialement été identifiée en 2006 chez trois individus présentant une déficience intellectuelle, un comportement hypersociable ainsi qu'une dysmorphie faciale caractéristique (Koolen et al., 2006) (Figure 15c-e). La duplication réciproque, beaucoup plus rare, est associée à une grande variabilité symptomatique.

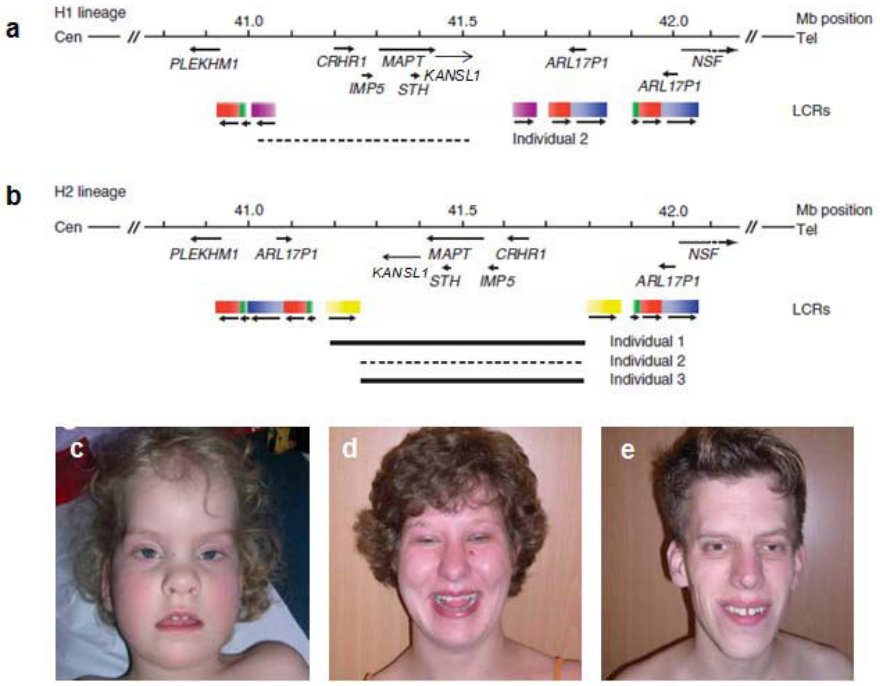


Figure 15 : Haplotypes de la région 17q21.31 et dysmorphie faciale caractéristique des individus atteints du syndrome de Koolen-de Vries

a-b. Haplotype 1 et 2 de la région 17q21.31. L'haplotype H2 place deux SDs indiqués par des bandes jaunes dans la même orientation et favorisent la perte ou gain d'une région contenant cinq gènes.
c-e. Individus 1 (c), 2 (d) et 3 (e) porteurs des délétions 17q21.31 et présentant une dysmorphie faciale caractéristique du syndrome de Koolen-de Vries. (Adapté de Koolen *et al.*, 2006).

Les délétions et les duplications apparaissent *de novo* et sont associées à l'haplotype H2 de la région 17q21.31 consistant en une inversion de 900 kb (Figure 15b). Ce polymorphisme présente une fréquence allélique de 20% dans la population européenne et place deux SDs dans la même orientation, favorisant les délétions et duplications par les recombinaisons NAHR (Steinberg et al., 2012; Zody et al., 2008).

4.2 Caractérisation clinique

En 2008, une première caractérisation clinique de 22 patients porteurs de la délétion 17q21.31 fut publiée (Koolen et al., 2008). Pendant l'enfance, tous les patients présentaient un retard développemental psychomoteur. Une pénétrance totale de la déficience intellectuelle était également notée. Une hypotonie associée à des difficultés alimentaires dans la période néonatale et durant l'enfance était observée pour 21 des 22 patients. 36% des cas présentaient une épilepsie généralisée. Une dysmorphie faciale caractéristique était retrouvée pour la majorité des patients. Celle-ci se caractérise par un front haut et large, un visage long, un

épicanthus, un nez bulbeux en forme de poire (80% des cas) et de larges oreilles proéminentes. L'évaluation ophtalmologique indiquait un strabisme chez 12 des 22 individus. Une déformation de la colonne vertébrale incluant scoliose, lordose et cyphose était observée dans 36% des cas. Des anomalies du septum cardiaque et des reins étaient également observées chez 27% et 32% des cas respectivement. Un cryptorchidisme était reporté pour 7 des 9 patients de sexe masculin. Enfin, un comportement de type hypersociable était retrouvé chez 20 des 22 patients.

Deux caractérisations cliniques ultérieures incluant onze (Tan et al., 2009) et quatorze (Dubourg et al., 2011) patients porteurs de délétion 17q21.31 ont abouti à des conclusions similaires. De plus, des anomalies récurrentes du système nerveux central ont été observées incluant un élargissement des ventricules, des agénésies totales ou partielles du corps calleux ainsi que des altérations de la substance grise périaqueducale (Tan et al., 2009).

La duplication réciproque de la région 17q21.31 a été identifiée dès 2007 chez une fille de dix ans présentant un retard psychomoteur sévère, une dysmorphie faciale, une microcéphalie et un hirsutisme (Kirchhoff et al., 2007). En 2009, la caractérisation de quatre cas supplémentaires n'indiquait pas de malformation cranio-faciale ni de déficience intellectuelle mais un retard psychomoteur léger et une interaction sociale réduite (Grisart et al., 2009). Un sixième cas fut identifié en 2012. L'individu adulte présentait une déficience intellectuelle, des troubles de l'humeur, une ataxie ainsi que de nombreuses anomalies congénitales (Kitsiou-Tzeli et al., 2012). Cette caractérisation clinique indique une importante variabilité des phénotypes associés à la duplication de la région 17q21.31. Le nombre extrêmement faible de patients identifiés depuis 2007 en comparaison aux porteurs de la délétion réciproque suggère une faible pathogénicité de la duplication 17q21.31. L'implication potentielle de polymorphismes secondaires pourrait expliquer le rare nombre de cas et l'hétérogénéité de la symptomatique. Néanmoins, une récurrence de deux phénotypes qui sont un retard psychomoteur et des troubles du spectre autistique a été observée chez les porteurs de la duplication 17q21.31. Les réarrangements de la région 17q21.31 sont donc associés à des phénotypes sociaux opposés : hypersociabilité pour les porteurs de la délétion et troubles du spectre autistique pour les porteurs de la duplication.

En 2012, des délétions atypiques et des mutations perte de fonction restreintes au gène *KANSL1*ont été identifiées chez six patients présentant une symptomatique identique aux patients porteurs de la délétion 17q21.31 (Koolen et al., 2012; Zollino et al., 2012). Aucun autre polymorphisme spécifique aux autres gènes de la région *CRHR1, SPPL2C, MAPT, STH*

n'a été identifié chez des patients présentant des symptômes associés à la délétion 17q21.31. Ces résultats suggèrent que l'haplo-insuffisance du gène *KANSL1* suffit à causer le syndrome de Koolen-de Vries.

4.3 Gènes de la région

Le gène *CRHR1* (Corticotropin Releasing Hormone Receptor 1) code pour un récepteur couplé à une protéine G qui lie l'hormone de libération de la corticotropine (CRH), neuropeptide impliqué dans l'axe hypothalamo-hypophysaire. Le récepteur est donc essentiel pour l'activation du signal de transduction qui régule le stress, la reproduction, la réponse immunitaire et l'obésité (Bornstein et al., 1998; Elenkov et al., 1999). Des SNPs pour le gène *CRHR1* ont été associés à l'anxiété, à la dépression et à la dépendance alcoolique (Chen et al., 2010; Rogers et al., 2013; Treutlein et al., 2006). *CRHR1* est donc un gène candidat intéressant pour les troubles sociaux des patients porteurs de la duplication 17q21.31.

Le gène *SPPL2C* (Signal Peptide Peptidase-Like 2C) ou *IMP5* (Intra Membrane Protease 5) code pour une peptidase de peptide signal localisée au niveau de la membrane plasmique du réticulum endoplasmique (Voss et al., 2013).

Le gène *MAPT* (Microtubule-Associated Protein Tau) code pour la protéine Tau, protéine associée aux microtubules qui contribue à l'assemblage des microtubules et à la stabilisation des axones (Kempf et al., 1996). Un grand nombre d'études a reporté l'effet toxique des agrégats de protéines Tau hyperphosphorylées dans l'apparition de conditions neurodégénératives (Spillantini and Goedert, 2013).

Le gène *STH* (Saitohin) est un gène spécifique à l'homme et aux primates à la fonction inconnue (Wang et al., 2011). Le gène situé dans un intron du gène *MAPT* contient un unique SNP (Q7R) qui confère une susceptibilité aux maladies neurodégénératives (Conrad et al., 2004; Gao et al., 2005) et aux troubles neuropsychiatriques (Bosia et al., 2012).

Enfin, le gène *KANSL1* (KAT8 regulatory NSL complex subunit 1) code pour une protéine d'interaction du complexe NSL (Non-Specific Lethal) (Dias et al., 2014) impliquée dans la régulation de la transcription et dans la maintenance et la reprogrammation des cellules souches (Mendjan et al., 2006; Raja et al., 2010). Le complexe NSL contient notamment la protéine MOF (Males absent On the First) codé par le gène KAT8 (K AcetylTransferase 8). MOF est une acétyltransférase d'histone qui permet l'acétylation des lysines K5, K8 et K16 de l'histone H4 facilitant l'activation de la transcription (Cai et al., 2010; Dou et al., 2005).

5. La souris comme animal modèle

Les syndromes de délétion et de duplication sont associés à une symptomatique de pénétrance très variable. Si certains syndromes présentent une forte pénétrance et une faible variabilité phénotypiques (syndromes de DiGeorge, de Williams ou de Prader-Willi, etc.), d'autres sont associés à un large spectre de troubles neuropsychiatriques et de malformations congénitales de pénétrance incomplète (syndromes de délétion/duplication pour les régions 2q13, 15q11.2 ou 16p11.2, etc.). La taille et la densité génique des loci présentant des déséquilibres génomiques sont également extrêmement variables. Le syndrome de Koolen-de Vries est associé à la délétion d'une région de 500 à 650 kb contenant cinq gènes codants alors que la monosomie 21 partielle peut concerner la délétion d'un segment d'une taille supérieure à 10 Mb contenant plus de cent gènes. En fonction de la densité génique de la région d'intérêt et de la variabilité des phénotypes associés, l'étude de la physiopathologie des syndromes de délétion et de duplication peut être extrêmement complexe.

Parallèlement aux études de génétique humaine, les modèles animaux ont été développés afin de faciliter l'étude des relations entre génotype et phénotype. Par ses similitudes du point de vue génétique, physiologique et moléculaire, la souris constitue un modèle de choix pour l'étude des gènes et de leurs fonctions et tient un rôle majeur dans l'étude des pathologies humaines. Au niveau génétique, plus de 90% du génome humain et du génome murin peuvent être fractionnés en régions de synténie (segments d'ADN pour lesquels la séquence génique du plus récent ancêtre commun est conservée entre les deux espèces) et environ 80% des gènes de la souris ont leur orthologue dans le génome humain (Waterston et al., 2002). Sur le plan comportemental, les rongeurs présentent des comportements d'anxiété, d'apprentissage et de mémoire, de sociabilité et autres réponses émotives. Beaucoup d'études emploient des modèles souris pour mimer des réponses comportementales humaines dans des conditions physiologiques et pathologiques, par exemple pour développer des traitements thérapeutiques ou pour des criblages toxicologiques. Sur le plan squelettique, il y a une grande conservation entre les deux espèces, notamment au niveau du crâne et de la mandibule (Richtsmeier et al., 2000) (Figure 16).

Il existe un grand nombre de lignées consanguines, c'est-à-dire génétiquement identiques, homozygotes à tous les locus de leur génome. L'intérêt de ces lignées est la reproductibilité des résultats d'un animal à l'autre. La composition génétique étant figée, dans des conditions environnementales strictement contrôlées (condition d'élevage, nourriture, cycle jour nuit, etc.), les données biologiques sont stables.

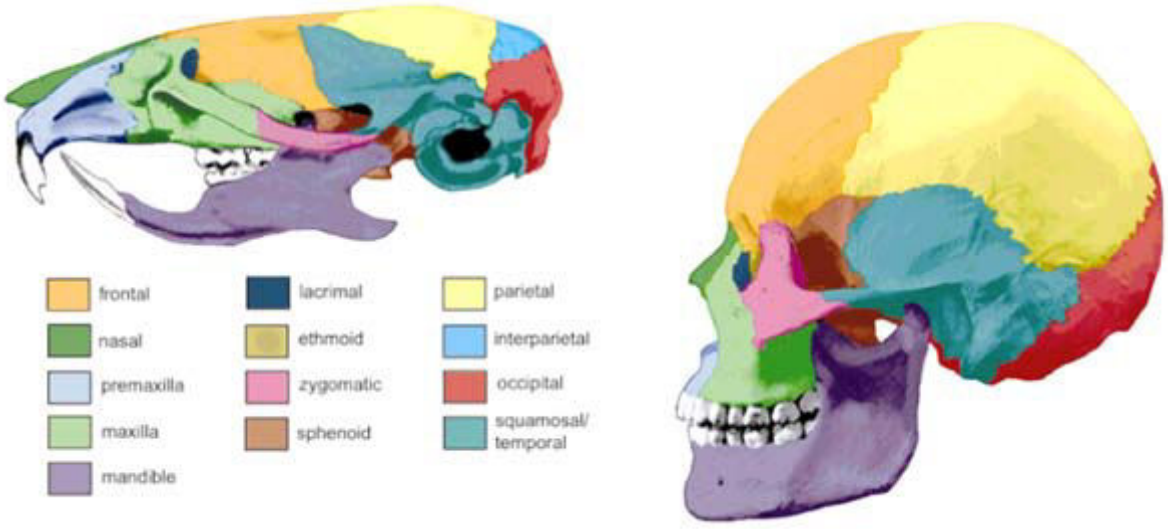


Figure 16 : Conservation des os du crâne et de la mandibule entre l'homme et la souris

(Adapté de Richtsmeier *et al.*, 2000).

Les technologies de manipulation du génome murin permettent la génération de délétion et de duplication d'une taille variant de quelques kb à plusieurs Mb (Brault et al., 2006). L'usage de la souris permet donc la modélisation des maladies monosomiques, mais aussi la modélisation des syndromes de délétions et de duplications. Un large panel de tests comportementaux permet d'évaluer les paramètres murins associés aux troubles neuropsychiatriques. De nombreuses techniques permettent l'évaluation anatomique et métabolique des animaux. La morphologie cranio-faciale, les fonctions cardiaques, la composition corporelle, ainsi que la structure cérébrale des animaux sont notamment étudiées par les techniques de microtomographie à rayon X, d'échocardiographie et de résonance magnétique nucléaire. L'usage de vecteur dérivé du virus adéno-associé permet la modulation de l'expression de gènes spécifiques localement par injection stéréotaxique dans certaines régions cérébrales (Lowery and Majewska, 2010), mais aussi de manière ubiquitaire (Foust et al., 2010). Le modèle murin offre donc de nombreux avantages pour l'étude des syndromes de délétions et de duplications notamment pour l'identification des gènes candidats, la compréhension des mécanismes moléculaires associés aux phénotypes et le développement de futurs traitements thérapeutiques. Effectivement, les modélisations murines des principaux syndromes de CNVs associés à l'ID incluant les syndromes de Williams, de Smith-Magenis et de Potocki-Lupski ont abouti à l'observation de nombreux déficits cognitifs (Nomura and Takumi, 2012; Ricard et al., 2010). Ces altérations comportementales associées à la symptomatique humaine offrent des perspectives très intéressantes.

III. Projet de recherche

Les travaux présentés dans ce manuscrit consistent en la caractérisation des modèles murins pour cinq syndromes neurodéveloppementaux associés aux variations du nombre de copies : la monosomie 21 partielle, ainsi que les délétions et duplications des régions 16p11.1 et 17q21.31. La caractérisation anatomique, métabolique et comportementale des modèles étudiés nous a permis d'évaluer un grand nombre des paramètres pouvant être associés aux désordres neuropsychiatriques et anomalies congénitales observés chez les patients. En collaboration avec le Dr. Maksym Kopanitsa et le Pr. Alexandre Reymond, nous avons également réalisé des analyses électrophysiologiques et transcriptomiques en ciblant nos investigations sur l'hippocampe, structure cérébrale qui, entre autres, joue un rôle central dans les processus de mémoire et d'apprentissage. Ce projet de recherche s'inscrit dans une perspective plus large qui est l'identification des gènes candidats et des mécanismes moléculaires impliqués dans la physiopathologie des différents syndromes étudiés dans le but de développer les premières stratégies thérapeutiques pouvant potentiellement aboutir à l'amélioration des capacités cognitives des patients.

La première partie de ce manuscrit décrit la caractérisation de la lignée Ms5Yah porteuse d'une délétion pour la région *App-Runx1*, modèle de monosomie 21 partielle pour la région 21q21.3q22.11. La région intègre les orthologues de la région *APP-SOD1* associée aux principaux signes cliniques de la PM21 (Chettouh et al., 1995) ainsi que les orthologues de la région *SOD1-RUNX1* localisée sur la bande 21q22.11 recouvrant la majeure partie de la région 2 associée aux symptômes les plus sévères de la PM21 (Lindstrand et al., 2010; Lyle et al., 2009; Roberson et al., 2011). Chez l'homme, la région *APP-RUNX1* constitue une région extrêmement critique associée à des malformations cranio-faciales, une déficience intellectuelle, un important retard psychomoteur, des anomalies du tonus musculaire et des malformations cardiaques. La caractérisation du modèle Ms5Yah a révélé une grande similarité phénotypique avec la symptomatique humaine.
Cette étude est présentée sous la forme d'un manuscrit en révision pour le journal « Disease Models and Mechanisms » intitulé "Major impact of the *App-Runx1* region in a mouse model of partial monosomy 21" ; T. Arbogast, M. Raveau, A. Duchon, C. Chevalier, V. Nalesso, M. Roux, D. Dembele, H. Jacobs, O. Wendling, Y. Herault.

Les deuxième et troisième parties de ce manuscrit ont nécessité la génération des modèles murins pour les réarrangements des régions 16p11.2 et 17q21.31. Nous avons utilisé le système Cre-Lox dans la stratégie de recombinaison méiotique trans-allélique ciblée ou stratégie TAMERE (pour Targeted Meiotic Recombination) (Brault et al., 2006; Herault et al., 1998) pour générer les animaux porteurs de la délétion (*Del/+*) et de la duplication (*Dup/+*) des régions de synténie 16p11.2 et 17q21.21 (Figure 17). Les croisements ont été effectués avec des animaux maintenus sur un fond génétique consanguin C57BL\6N (B6N).

La région de synténie du locus 16p11.2 est située sur le chromosome murin 7F3. Les sites loxP ont été insérés en amont du gène *Sult1a1* et en aval du gène *Spn*. Les croisements ont été initiés en novembre 2010 et les premiers animaux *Del/+* et *Dup/+* pour la région *Sult1a1-Spn* ont été obtenus respectivement en mars et mai 2012.

La région de synténie du locus 17q21.31 est localisée sur le chromosome murin 11E1. Les sites loxP ont été insérés en amont du gène *Arf2* et en aval du gène *Kansl1*. Les croisements ont été initiés en novembre 2010 et les premiers animaux *Del/+* et *Dup/+* pour la région *Arf2-Kansl1* ont été obtenus respectivement en septembre et décembre 2011.

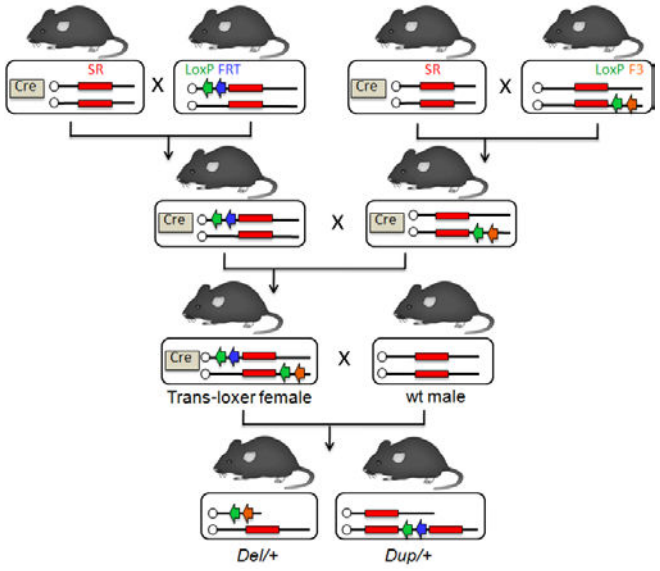


Figure 17 : Stratégie de croisement par recombinaison méiotique trans-allélique

La méthode de recombinaison méiotique trans-allélique se base sur deux croisements successifs aboutissant à l'obtention d'une femelle exprimant la Cre recombinase (dans notre cas sous le contrôle du promoteur ubiquitaire Hprt) et porteur de deux sites loxP dans une configuration trans en amont et en aval de la région de synténie (SR) ciblée. La Cre permet la recombinaison entre loxP durant la prophase méiotique et la génération d'ovocytes porteurs de la délétion ou de la duplication de la région délimitée par les sites LoxP. La dernière étape consiste à croiser la femelle avec un mâle sauvage pour générer des individus porteurs de la délétion ou de la duplication de la région d'intérêt.

La deuxième partie de ce manuscrit concerne la caractérisation des modèles murins pour les réarrangements de la région 16p11.2. Les animaux *Del*/+ et *Dup*/+ pour la région *Sult1a1-Spn* ont dans un premier temps été amplifiés séparément sur un fond consanguin B6N. En collaboration étroite avec le Dr. Abdel-Mouttalib Ouagazzal, nous avons développé des protocoles d'analyse comportementale permettant d'étudier les paramètres murins associés aux nombreux troubles neuropsychiatriques décrits chez l'homme. Nous avons également entrepris une large caractérisation incluant un suivi de la masse corporelle sans diète standard, une étude métabolique sous diète enrichie, ainsi qu'une analyse cranio-faciale. Des altérations importantes de la masse corporelle, de l'adipogenèse, de l'activité et de la mémoire de reconnaissance ont été observées. Au sacrifice, nous avons prélevé le foie, le cervelet, le striatum et l'hippocampe à partir desquels nous avons réalisé des analyses transcriptomiques. Dans le but de confirmer les phénotypes observés, nous avons généré une cohorte *Del-Dup* contenant les 4 génotypes *Del*/+, wt, *Del/Dup*, *Dup*/+ (Figure 18). Enfin, dans le but d'étudier l'influence du fond génétique, nous avons caractérisé les animaux *Del*/+ et *Dup*/+ sur un fond génétique hybride B6NC3B par croisement des animaux mutants de fond B6N avec des animaux sauvages de fond C3H/HeH (C3B). Notre choix s'est porté sur la souche murine C3B pour ses bonnes capacités cognitives et pour son activité plus faible en comparaison à la souche B6N (Mandillo et al., 2008).

Cette étude est présentée sous la forme d'un manuscrit intitulé "Opposite effects and genetic dosage of mouse models of 16p11.2 deletion and duplication syndromes" ; T. Arbogast, A.M. Ouagazzal, C. Chevalier, M. Kopanitsa, N.O. Afinowi, E. Migliavacca, M.C. Birling, M.F. Champy, A. Reymond, Y. Herault.

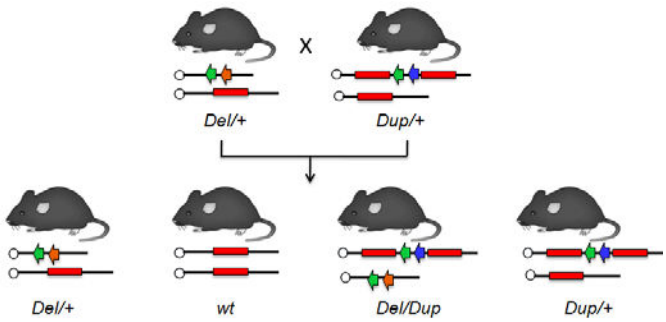


Figure 18 : Stratégie de croisement pour la génération de cohortes *Del-Dup*

Le croisement d'animaux hétérozygotes *Del*/+ et *Dup*/+ permet de générer des individus de 4 génotypes *Del*/+, wt, *Del/Dup* et *Dup*/+. Cette stratégie permet de comparer les capacités des animaux *Del*/+ et *Dup*/+ à un seul groupe d'animaux wt et d'étudier les animaux *Del/Dup* porteurs des deux régions génétiques d'intérêt sur un seul chromosome.

La troisième partie de ce manuscrit intègre la caractérisation des animaux modèles pour les réarrangements de la région 17q21.31. Les animaux *Del*/+ et *Dup*/+ pour la région *Arf2-Kansl1* ont dans un premier temps été amplifiés séparément sur un fond consanguin B6N. Pour chaque réarrangement, une analyse cranio-faciale et une caractérisation comportementale ont été réalisées. Nous avons confirmé les phénotypes observés par la caractérisation d'une cohorte *Del-Dup* de mâles (Figure 18). Nous avons évalué l'activité, l'anxiété, les capacités d'apprentissage et de mémoire, les comportements sociaux et la coordination motrice des animaux. Dans un deuxième temps, dans le but d'étudier les conséquences de l'haplo-insuffisance du gène *Kansl1*, nous avons généré une cohorte d'animaux knock-out (KO) hétérozygotes *Kansl1*$^{+/-}$. De manière très intéressante, des phénotypes comportementaux similaires ont été observés pour les animaux Del/+ et *Kansl1*$^{+/-}$. Cette étude est présentée sous la forme d'un manuscrit en préparation intitulé "Mouse models of Kansl1 haploinsuffisancy and 17q21.31 rearrangements resume major phenotypes found in patients" ; T. Arbogast, C. Chevalier, M. Kopanitsa, N.O. Afinowi, M.C. Birling, M.F. Champy, H. Meziane, Y. Herault.

Partie 1

Modélisation de la monosomie 21 partielle

Introduction

La monosomie 21 partielle (PM21) est une aberration chromosomique congénitale se caractérisant par la présence d'une délétion partielle variable d'un segment du bras long du chromosome 21 (HSA21). Selon la taille et la localisation des délétions, la PM21 peut être associée à un retard de croissance intra-utérin, une diminution de la taille, des dysmorphies cranio-faciales, une déficience intellectuelle, des déficits musculaires, des malformations cardiaques et squelettiques (Lindstrand et al., 2010; Lyle et al., 2009; Roberson et al., 2011).

Chez la souris, les orthologues du HSA21 sont localisés sur trois grandes régions de synténie au niveau des chromosomes murins (MMU) 10, 16 et 17. De nombreux modèles murins porteurs de délétions pour les trois régions de synténie du HSA21 ont été décrits dans la littérature. Les régions MMU 10 et 17 correspondent à la partie distale 21q22.3 du bras long du HSA21 dont les délétions sont associées, chez l'homme, à une symptomatique modérée. Alors que la délétion de l'ensemble de la région de synténie MMU10 induit un léger déficit de mémoire spatiale, la délétion de la région *Abcg1-Rpr1b* sur le MMU17 induit des déficits de mémoire associative (Yu et al., 2010). Une étude récente associe la délétion de la région *Abcg1-U2af1* de la région MMU17 à une réduction de l'interaction sociale, ainsi qu'à une altération de la LTP (Sahún et al., 2014).

A ce jour, seuls deux modèles murins de PM21 pour la région de synténie MMU16 ont été caractérisés. Le modèle Ms1Rhr porteur de la délétion *Cbr1-Orf9* correspondant à la région 21q22.12q22.2 ne présente aucun phénotype comportemental mais montre une augmentation du volume cérébelleux et une diminution du volume hippocampique (Olson et al., 2007). Le modèle Ms1Dja porteur de la délétion *Lipi-Usp25* correspondant à la région centromérique 21q11.2q21.1 montre une augmentation de l'adipogenèse, ainsi qu'un déficit modéré de préférence sociale (Migdalska et al., 2012). Ces deux modèles présentent de légers phénotypes morphologiques et comportementaux en comparaison à la sévérité des symptômes associés aux délétions des régions 1 et 2 du HSA21 décrites par Lyle et *al*.

Dans le but d'étudier l'influence de la région de synténie 21q21.3q22.11 qui constitue une région extrêmement critique de la PM21 (Figure 19), le modèle Ms5Yah porteur d'une délétion pour la région *App-Runx1* sur la région de synténie MMU16 a été généré dans notre équipe. Dans un premier temps, la lignée Ms5Yah a été croisée avec la lignée Ts65Dn, premier modèle murin viable de trisomie 21 (Davisson et al., 1993; Reeves et al., 1995). Les

études menées au sein de notre équipe ont révélé l'importance des effets de dose des gènes de la région *App-Runx1* dans les anomalies cardiaques du modèle Ts65Dn (Raveau et al., 2012).

Dans la présente étude, nous avons réalisé la caractérisation morphologique et comportementale du modèle Ms5Yah. Dans le but d'étudier les causes d'une importante létalité observée pour les animaux Ms5Yah, nous avons réalisé des tests de viabilité (Deans et al., 2000) ainsi qu'une étude anatomique des fœtus présentant une incapacité respiratoire. La masse corporelle et la motricité des animaux survivants ont été suivies dès les premiers jours post-partum. Lors de la caractérisation comportementale, nous avons évalué l'activité, l'anxiété, les capacités d'apprentissage et de mémoire, la coordination motrice et la force d'agrippement des animaux adultes. Des analyses transcriptomiques à partir de prélèvements d'hippocampe ont permis d'identifier les gènes de la région *App-Runx1* sensibles aux effets de dose et d'étudier l'impact de la délétion sur la régulation du génome murin.

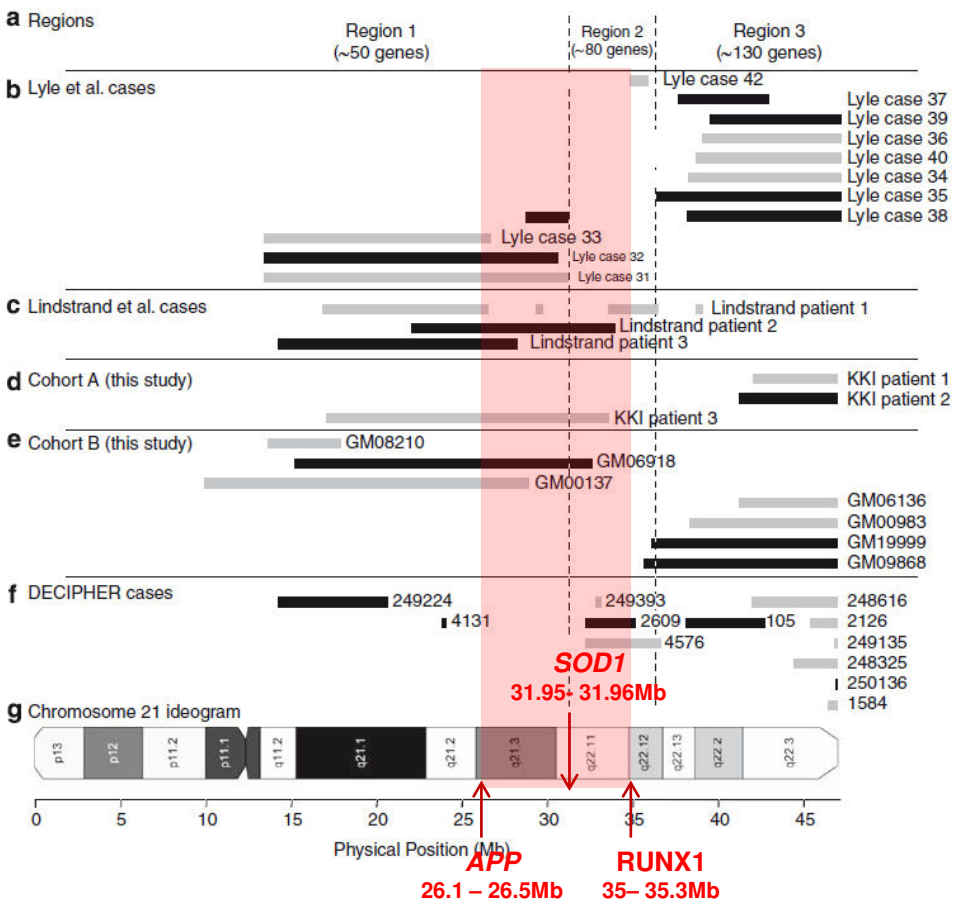


Figure 19 : Région *APP-RUNX1*

Localisation des gènes *APP, SOD1* et *RUNX1* au niveau des trois régions d'association du HSA21 avec la symptomatique de la monosomie 21 partielle définies en 2009 par Lyle et *al.*
(Adapté de Roberson *et al.*, 2011).

Major impact of the App-Runx1 region in mouse model of partial monosomy 21

Thomas ARBOGAST[1], Matthieu RAVEAU[1], Arnaud DUCHON[1], Claire CHEVALIER[1], Valérie NALESSO[1] Michel ROUX[1], Doulaye DEMBELE[1], Hugues JACOBS[2], Olivia WENDLING[2], Yann HERAULT[1,2]

Affiliations

[1]IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), Dpt of Translational Medicine and Neurogenetics; CNRS, UMR7104; INSERM, U964; Strasbourg University; 1 rue Laurent Fries, F-67404 Illkirch-Graffenstaden, France

[2]Institut Clinique de la Souris, ICS; PHENOMIN; GIE CERBM; 1 rue Laurent Fries, F-67404 Illkirch-Graffenstaden, France

Corresponding author: Yann Hérault

Keywords: Mouse model, aneuploidy, learning and memory, motor coordination

Running title

Deletion of the *App-Runx1* region

ABSTRACT

Partial monosomy 21 (PM21) is a rare chromosomal abnormality characterized by the loss of a variable segment along human chromosome 21 (Hsa21). Clinical phenotypes are heterogeneous and range from mild alterations to lethal consequence depending on the affected region of the Hsa21. Most common features include intellectual disabilities, craniofacial dysmorphology, short stature, muscular and cardiac defects. As a complement to human genetic approaches, our team has developed new monosomic mouse models carrying deletions on Hsa21 syntenic regions in order to identify dosage-sensitive genes responsible for the symptoms. We focused here on the Ms5Yah mouse model, deleted for a 7.7 Mb region from *App* to *Runx1* genes. Ms5Yah mice display high postnatal lethality with a few surviving individuals showing growth retardation, motor coordination deficits and spatial learning and memory impairments. Further studies confirmed gene dosage effect in Ms5Yah hippocampus and pinpoints disruptions of pathways related to cell adhesion (*App, Cntnap5b, Lgals3bp, Mag, Mcam, Npnt, Pcdhb2/3/4/6/7/8/16 and Vwf*). Our PM21 mouse model is the first one to display morphological abnormalities and behavioral phenotypes similar to those found in patients, and therefore, it demonstrates the major contribution of the *App-Runx1* region to the pathophysiology of the partial monosomy 21.

INTRODUCTION

Segmental aneuploidy, defined as an abnormal copy number of a genomic region, is a common cause of human genetic disorders leading to intellectual disabilities (ID). Human chromosome 21 (Hsa21; Hsa for *Homo sapiens*) aneuploidies are often associated with trisomy 21 or Down syndrome (DS), the principal genetic cause of ID (Megarbane et al., 2009). Although extremely rare, different cases of partial Hsa21 monosomies (PM21) have also been reported since 1964 when Lejeune described the first PM21 case for a small acrocentric chromosome (Lejeune et al., 1964). Modern techniques have confirmed that complete monosomy of Hsa21 without any translocation to another chromosome is incompatible with life (Burgess et al., 2014; Toral-Lopez et al., 2012). Partial deletions, depending on their size and location on Hsa21 are associated with a large heterogeneity of clinical phenotypes. Some patient's present severe phenotypes such as brain dysgenesis and heart defects not compatible with survival; others show milder phenotypes such as slight dimorphic features or no symptom at all. Most common features of PM21 include intellectual disability, craniofacial malformations, short stature, muscular and cardiac defects (Chettouh et al., 1995; Lindstrand et al., 2010; Lyle et al., 2009; Roberson et al., 2011; Theodoropoulos et al., 1995; Valero et al., 1999).

A first molecular mapping of features associated with PM21 was done in 1995 by comparing phenotypes and karyotypes of six patients (Chettouh et al., 1995). The analysis pinpointed a 5.3Mb region from *APP* to *SOD1* involved in intellectual disability, hypotonia and cranio-facial malformations. However, high-resolution mapping of pathogenic partial aneuploidies and unbalanced translocations involving Hsa21 argue against a single critical region and reveal susceptibility regions for the different phenotypes of PM21 (Lindstrand et al., 2010; Lyle et al., 2009; Roberson et al., 2011). The long arm of Hsa21 can roughly be divided into three regions (Lyle et al., 2009). The first region, from the centromere to ~ 31.2Mb, covers a gene-poor region of Hsa21 (~ 50 genes). Only large deletions are found in patients presenting intellectual disability, muscular defects and several cranio-facial malformations. The second region, from 31.2 to 36Mb, has a high gene density (~ 80 genes). Only few patients carrying a partial deletion have been diagnosed with severe phenotypes indicating that the haploinsufficiency of the entire region may not be compatible with survival. In the last region, from 36Mb to the telomere (~ 130 genes), deletions induce relatively mild phenotypes. Given the rarity of patients, it is very difficult to identify genes

responsible for the different PM21 symptoms. Complementary to the genetic analysis, mouse models have been developed in order to study the correlation between phenotype and genotype. Almost all protein coding genes found on the Hsa21 long arm have homologues carried by mouse chromosomes 16 (Mmu16 for *Mus musculus*, 23.3 Mb, 166 genes between *Lipi* and *Zfp295*), 17 (Mmu17, 1.1 Mb, 22 genes between *Umodl1* and *Hsf2bp*) and 10 (Mmu10, 2.3 Mb, 45 genes between *Pdxk* and *Prmt2*). This synteny allows the use of mouse models to facilitate the identification of genes responsible for the symptoms of PM21 patients.

To date, several mouse models carrying deletions on Mmu17 and Mmu10 syntenic regions corresponding to human 21q22.3 distal part have already been generated and characterized. Ms1Yah mice carrying a 0.5Mb deletion between *Prmt2* and *Col61a* genes on Mmu10 present no behavioral phenotype but exhibit an increased inflammatory reaction after intranasal lipopolysaccharide administration (Besson et al., 2007); Ms4Yah mice deleted for the 2.2Mb *Prmt2-Cstb* region on Mmu10 present no gross dimorphism or behavioral defects (Duchon et al., 2011). Df(10)1Yey carrying a 2.3Mb deletion for the entire Mmu10 syntenic region between *Prmt2* and *Pdxk* which is directly upstream from *Cstb* show moderate spatial memory impairment (Yu et al., 2010). Df(17)1Yey mice deleted for a 1.1Mb region on Mmu17 between *Abcg1* and *Rrp1b* present deficits in associative memory (Yu et al., 2010). A recent study show that Ms2Yah mice deleted for the 0.59 Mb *Abcg1-U2af1* region display decreased social novelty interaction and hippocampic long-term potentiation (LTP) alterations ((Sahún et al., 2014). Two mouse models carrying deletions in the Mmu16 syntenic region have been generated and characterized. Ms1Rhr mice deleted for the 4.2Mb *Cbr1-ORF9* region syntenic to Hsa21 so called “DS critical region” (Epstein, 1990; Korenberg et al., 1994) display no behavioral phenotype but show brain and hippocampus volume decrease and cerebellum volume increase (Olson et al., 2007). Ms1Dja mice harboring a 1.6Mb deletion between *Lipi* and *Usp25* corresponding to Hsa21 centromeric region 21q11.2-q21.1 exhibit increased of fat deposition and moderate social memory impairment (Migdalska et al., 2012).

In order to study the influence of the Hsa21q21.3-22.11 syntenic region, the Ms5Yah mouse model carrying a deletion for the 7.7 Mb *App-Runx1* genetic interval was engineered in our lab. In 2012, Raveau et al. demonstrated how the region influence lethality and cardiac defect seen in the Ts65Dn DS mouse model (Raveau et al., 2012). In the present study, we characterized the Ms5Yah mouse model and explored its impact on post-natal viability and its consequences on behavior and cognition. An important lethality was observed at birth and in

the first days post-partum. At weaning, only 22.4% of animals were carrying the *App-Runx1* deletion. In comparison with control littermates, Ms5Yah mice showed decreased body size, weight reduction and platelet deficit. Defect were identified in the rotarod, the notched bar test and in the Morris water maze. To understand the possible underlying mechanisms in the hippocampus, we made transcriptomic analysis and observed a whole genome effect from the decrease in gene dosage associated with the *App-Runx1* region. Overall, our observations demonstrate that Ms5Yah mouse model recapitulates major symptoms observed in PM21 human syndrome.

RESULTS

Loss of one copy of the App-Runx1 region impacts on early post-natal viability, basic behavior and general morphology

During breeding of Ms5Yah mice, we had major difficulties in establishing the monosomic model on a pure C57BL/6N genetic background. Thus, we maintained the line on a B6NC3B mixed genetic background. Despite a marked improvement in comparison to C57BL/6N background, we noticed a low transmission of the Ms5Yah allele at weaning. Indeed, over 389 mice generated from heterozygote carrier with wild-type, 87 Ms5Yah individuals were observed at weaning. The 22.4% allele frequency is far below the 50% that could be expected from a Mendelian ratio (Chi-squared test, $p < 0.001$).

To determine whether Ms5Yah animals died in utero or postnatally, the fetuses of 7 pregnant wild-type females impregnated by Ms5Yah males were collected at embryonic day 18.5 (E18.5), few hours before natural delivery, in order to perform viability tests (Figure 1). Over 52 fetuses extracted by caesarean section and analyzed, we genotyped 27 wild-type and 25 Ms5Yah fetuses which is the normal genotypic Mendelian ratio. While a majority of neonates including all wild-type animals turned to a pink color and breathed actively (Fig. 1A), 40% (10/25 mice) of Ms5Yah showed severe respiratory distress, remained in a cyanotic color and died during the first 30 min after delivery (Fig. 1B). 2 cases of craniorachischisis were also observed and genotyped as Ms5Yah (Fig. 1C). Whereas all 27 wild-type neonates survived without difficulty, only 52% (13/25 mice) of mutant E18.5 neonates were able to breathe but were clearly underweight and needed more time to turn pink in comparison with wild-type littermates. Histological analysis revealed that Ms5Yah fetuses which were unable to breathe appeared essentially normal when compared to wilt-type littermate at the exception of the lungs. Whereas control littermates showed opened alveoli (Fig. 1D and F), intrapulmonary bronchi and alveoli of Ms5Yah fetuses remained collapsed (Fig. 1E and G). However, thoracic muscles appeared normally developed, as well as thoraco-abdominal diaphragm muscle, larynx and extrapulmonar airway.

Surviving new-born mice were evaluated for righting reflex from day 2 to 10 postpartum (Figure 2). When placed on the back, latency to recover natural posture was longer for Ms5Yah animals compared to wild-type littermate (Repeated Measures ANOVA "genotype" $F_{(1,176)} = 11.982$, $p = 0.002$; Fig. 2A). Body weights were recorded once a week

(on the same day at the same time) from the age of 1 to 13 weeks. Ms5Yah mice were consistently underweight (Repeated Measures ANOVA "genotype" $F_{(1,84)} = 29.434$, $p < 0.001$; Fig. 2B) and smaller in size (Fig. 2C) compared to controls. Blood analysis done on adult animals revealed that Ms5Yah mice had platelet decrease (Ms5Yah: $774x10^3 \pm 68$ cells/µL, wt: $1113x10^3 \pm 79$ cells/µL; $F_{(1,17)} = 10.361$, $p = 0.005$; Supplemental Table S1). We measured bleeding time of animals and found out that time needed for wound coagulation was twice longer for Ms5Yah animals compared to controls (Ms5Yah: 5.50 ± 0.76 min, wt: 2.99 ± 0.33 min; $F_{(1,20)} = 9.086$, $p = 0.007$).

Ms5Yah mice show motor coordination deficit without alteration of exploration activity

Exploratory pattern and anxiety were evaluated in the open field and elevated plus maze tests. In open field, no significant difference was observed between Ms5Yah and wild-type mice in distance traveled ($F_{(1,22)} = 0.677$, $p = 0.420$; Supplemental Fig. S1A) and time spent in the center area ($H_{(1,22)} = 0.121$, $p = 0.728$; Fig. S1B). In the elevated plus maze, Ms5Yah mice explored the same number of arms ($H_{(1,22)} = 0.071$, $p = 0.790$; Fig. S1C) and spent the same time in open arms than controls ($H_{(1,22)} = 0.304$, $p = 0.581$; Fig. S1D).

The rotarod test was performed for general motor coordination and balance evaluation (Figure 3). During the training period made of 3 daily sessions during which mice were placed on an accelerating rotarod, latencies to fall off the rod (Fig. 3A) and corresponding speed (Fig. 3B) were calculated. Compared to controls, Ms5Yah mice showed a global deficit to stay on the rod (Repeated Measures ANOVA "genotype" $F_{(1,44)} = 15.967$, $p < 0.001$). Whereas wild-type animals were able to improve their coordination capacities from day 1 (D1) to day 3 (D3), Ms5Yah presented impairments to stay on the rod on the first day and showed no amelioration on the third day (wt: D1 vs D3, $q = 6.029$, $p < 0.001$; Ms5Yah: D1 vs D3: $q = 2.085$, $p = 0.313$). The phase test consisted of 6 trials of 2 min at constant speed, with increasing speed between each trial (4, 10, 16, 22, 28, 34 and 40 rpm). Two sessions were done on the same day. As for training, Ms5Yah mice showed difficulties to maintain on the rod with a significant difference compared to controls from speeds 16 to 34 rpm (16 rpm: $H_{(1,22)} = 3.889$, $p = 0.049$; 22 rpm: $H_{(1,22)} = 8.159$, $p = 0.004$; 28 rpm: $H_{(1,22)} = 9.255$, $p = 0.002$; 34 rpm: $H_{(1,22)} = 6.317$, $p = 0.012$; Fig. 3C). The notched bar test was used to evaluate hind limb coordination. Animals had to cross a notched bar twenty times and each time back paw went through the gap was counted as an error. Ms5Yah mice committed more errors than

wild-type mice ($H_{(1,22)}$ = 4.678, p = 0.031; Fig. 3D). We then evaluated Ms5Yah muscle strength using the grip test. The strength index was calculated by the ratio of the force by the body weight. No difference was observed between mutants and controls ($F_{(1,22)}$ = 1.411, p = 0.247; Fig. 3E). Histological analysis did not reveal any anatomical or thickness change in cerebellum (Fig. S2A-B). All layers of mutant cerebella including the molecular, granular and Purkinje cell layers appeared normal (Fig. S2C-D).

Loss of one copy of the *App-Runx1* affects spatial learning and memory

Working memory was evaluated by recording spontaneous alternation in the Y-maze test. No difference in the number of arm entries ($F_{(1,22)}$ = 2.125, p = 0.159; Fig. S1E) and in the percentage of alternation was found ($F_{(1,22)}$ = 0.189, p = 0.668; Fig. S1F). We then assessed recognition memory of mice using the novel object recognition (NOR) test. The choice to explore the novel object reflects the use of learning and recognition memory processes. Mice were placed in the open field arena for 10 min with a first object. Interestingly, Ms5Yah mice showed a higher exploration time comparing to controls (Ms5Yah: 12.7 ± 1.2 s, wt: 9.2 ± 0.9 s; $F_{(1,17)}$ = 4.776, p = 0.043; Fig. S1G). One hour after the first session, mice were placed in the same arena with the first object (considered as "known") and a novel object. Both genotypes had similar exploration levels for the former object ($F_{(1,17)}$ = 1.695, p = 0.210) and the new object ($F_{(1,17)}$ = 3.336, p = 0.085) and showed similar discrimination indexes ($F_{(1,17)}$ = 0.415, p = 0.528; Fig. S1H) indicating no alteration of recognition memory.

We further investigated spatial learning and memory in the Morris water maze (Figure 4). In this task, animals learned to locate the position of a submerged platform by using extra-maze spatial cues. In the acquisition phase, both genotypes were able to increase their capacities and travelled a lower distance to find the platform between day 1 (D1) and day 6 (D6) indicating that mice were able to learn (Repeated Measures ANOVA "day", $F_{(1,110)}$ = 25.709, $p < 0.001$; wt D1 vs D6: q = 10.963, $p < 0.001$; Ms5Yah D1 vs D6: q = 7.400, $p < 0.001$; Fig. 4A). But Ms5Yah mice showed a global delay in the acquisition of the task and travelled a greater distance than controls to find the platform, with significant difference from D3 to D6 (Repeated Measures ANOVA "genotype", $F_{(1,110)}$ = 18.333, $p < 0.001$; D3: q = 3.077, p = 0.030; D4: q = 3.771, p = 0.008; D5: q = 3.661, p = 0.010; D6: q = 3.156, p = 0.026). No difference in swimming speed was detected (Repeated Measures ANOVA "genotype", $F_{(1,110)}$ = 2.633, p = 0.119; Fig. 4B) indicating normal coordination in water environment for mutant mice. A probe test was performed on day 7. The platform was

removed and a unique trial was done in order to examine the duration spent in target quadrant (SO) and the number of time that mice crossed the exact spatial position of the platform (annulus crossing). Both genotypes spent similar time in the quadrant of interest ($F_{(1,22)}$ = 1.895, p = 0.182; Fig. 4C); but Ms5Yah mice showed a clear diminution of annulus crossing ($F_{(1,22)}$ = 8.910, p = 0.007; Fig. 4D) indicating a deficit in spatial memory. In the reversal sessions made on day 12 and day 13, animals had to learn new positions of a platform made visible by an indicator. Mice from both groups needed shorter time to find the flagged platform than in the hidden version but Ms5Yah still showed a delay compared to control performances (Repeated Measures ANOVA "genotype", $F_{(1,22)}$ = 28.753, $p < 0.001$; D12: q = 6.267, $p < 0.001$; D13: q = 5.462, $p < 0.001$, Fig. 4A). When we looked at the duration spent in the learning target quadrant before mice reached the flagged platform, we saw a clear increase for Ms5Yah mice at day 2 (Ms5Yah: 9.3 ± 1.0 s, wt: 6.4 ± 0.8 s; $H_{1,22)}$ = 7.573, p = 0.006) and day 13 (Ms5Yah: 7.4 ± 0.5 s, wt: 4.8 ± 0.5 s; $H_{1,22)}$ = 11.487, $p < 0.001$). This indicates that mutant mice showed difficulties to switch from a defined task to a new one.

In the optomotor response test, performed in the same lighting conditions as the Morris water maze test, both wild-type and Ms5Yah mice tracked rotating stripes for spatial frequencies of 0.13, 0.51 and 0.25 cycles/degree, but failed to respond at 1.25 cycles/degree, suggesting that wild-type and Ms5Yah mice had comparable visual acuity, sufficient to detect the flag place on the Morris water maze platform.

Transcriptomic expression analysis

Whole genome expression arrays were performed on adult mice hippocampi to determine dosage-sensitive genes and to observe the deregulations on the whole genome using Affymetrix Gene Chip technology (Figure 5). Over 23,332 probe set ID, 12,870 were found expressed in the hippocampi. After normalization, we obtained a subset of 192 deregulated genes with a fold change (FC) higher than 1.2 or lower than 0.8 (Table 1). Clustering analysis revealed two distinct groups with respect to genotype-associated expression patterns (Fig. 5A). The first group made of 96 genes down-regulated in Ms5Yah mice included 33 of the 45 transcripts of the *App-Runx1* region referenced on chips (Fig. 5B). As expected *App* (FC = 0.68, $p < 0.001$), *Olig1* (FC = 0.75, $p < 0.001$) and *Olig2* (FC=0.78, $p < 0.001$) from the monosomic interval were found under-expressed (Fig. 5B). The second group contained 96 up-regulated genes dispersed throughout the genome including *Cbs* (FC = 1.34, $p < 0.001$), a

strong candidate gene for DS cognitive traits. In our analysis, a majority of deregulated genes showed a FC varying from 0.4 to 1.6 except for 2 genes, *Klk6* (FC = 1.70, $p < 0.001$) and *Myoc* (FC = 2.49, $p < 0.001$). We analyzed the enrichment of functional annotation using DAVID software (Huang et al., 2009b). The table 2 describes the principal entities affected by *App-Runx1* deletion. We found an enrichment for cell adhesion transcripts encoded by *App, Cntnap5b, Lgals3bp, Mag, Mcam, Npnt, Pcdhb2/3/4/6/7/8/16* and *Vwf* (Fold enrichment (FE) = 3.82; $p = 7.8.10^{-5}$; False discovery rate (FDR) = 0.099). Concerning molecular functions, the analysis pinpointed a clear enrichment for protocadherins (*Pcdhb2/3/4/6/7/8/16*; FE = 35.8; $p = 2.9.10^{-8}$; FDR = $4.1.10^{-5}$) and for calcium ion binding (*Efhd1, Itsn1, Npnt, Pcdhb2/3/4/6/7/8/16*, *Pla2g5, Pnck, Pon3, Ppp2r3a, Slc24a4, Smoc2 , Sulf1*; FE = 2.08; $p = 7.1.10^{-3}$; FDR = 8.73).

DISCUSSION

We report here the characterization of Ms5Yah mouse model monosomic for the 7.7 Mb *App–Runx1* region. This genetic interval overlaps about 50 known protein-coding genes located on the Mmu16, homologous to Hsa21q21.3-22.11 genes. Previous studies of monosomic mice for Mmu17 and Mmu10 syntenic regions corresponding to telomeric part of Hsa21q revealed no morphological abnormality and mild behavioral phenotypes (Besson et al., 2007; Duchon et al., 2011; Sahún et al., 2014; Yu et al., 2010). Two mouse models deleted in the Mmu16 syntenic region of Hsa21 have been characterized. Ms1Rhr mice deleted in the Hsa21q22.2 syntenic region show brain volume alterations (Olson et al., 2007) and Ms1Dja mice deleted in the centromeric Hsa21q11.2-q21 homologous region present an increase of fat deposition (Migdalska et al., 2012). All these mouse models present relatively incomplete phenotypes in regard to the severity of PM21 human symptoms. On the contrary, the phenotypic analysis of Ms5Yah mice revealed dramatic changes on viability, morphology, motor coordination and cognition.

Ms5Yah mice were not viable on pure C57BL/6N genetic background. The line was maintained on an F1 B6NC3B genetic background to improve the survival of newborns. Viability tests revealed few cases of craniorachischisis and a clear breathing incapacity for 40% of Ms5Yah neonates. Histological evaluation failed to reveal significant malformation in any organ of Ms5Yah died fetuses. The only abnormality observed was pulmonary collapse indicating incapacity to breathe. Surviving animals were underweight and had coordination deficit. Put on their back, Ms5Yah pups needed more time to recover their normal position. This locomotion impairment could limit the feeding capacities of pups, potentially worsening the size deficit already observed at birth. With a longer bleeding time and decreased level of platelets, these deficits could explain that only 22.4% of animals which survived at weaning were carrying the *App-Runx1* deletion. With the other analysis carried out for all the syntenic regions covering the long arm of Hsa21 in mouse (Besson et al., 2007; Duchon et al., 2011; Migdalska et al., 2012; Olson et al., 2007; Sahún et al., 2014; Yu et al., 2010), the *App-Runx1* interval is the only critical region for embryonic development and viability, reinforcing the hypothesis of lethality due to full monosomy 21 observed in human (Schinzel, 1976; Toral-Lopez et al., 2012).

Adult behavioral phenotyping confirmed dramatic changes in locomotor coordination and spatial learning and memory defects of Ms5Yah mice. During the open field and elevated plus maze tests, Ms5Yah had normal exploratory behavior and did not show anxiety. In the rotarod test, mutant mice showed severe deficits in motor skills. Ms5Yah animals were unable to improve their performances during the training phase and failed to stay on the rod even from lower speeds of the test phase. This coordination deficit was confirmed by the notched bar test. Grip strength test did not reveal any muscular weakness and no gross anatomical change was observed in Ms5Yah cerebellum histological analysis. All these data pinpoint a motor coordination deficit of mutant mice without alteration of muscular strength and general activity. Y maze and novel object recognition test revealed no perturbation of short-term working memory and recognition memory. In the water maze test, Ms5Yah animals showed a global delay in the acquisition of the task and travelled a greater distance to find the platform without affectation of swimming speed. In the removal session, mutant mice spent the same amount of time in the target quadrant than controls but showed difficulty to remember the exact position of the platform, indicating a spatial memory deficit. Finally during reversal sessions, Ms5Yah showed difficulty to switch from a defined task to a new one.

Hippocampus-dependent spatial learning and locomotor learning deficits in mice can be directly related to intellectual disabilities of PM21 patients. In order to evaluate the impact of *App-Runx1* gene dosage reduction on whole genome, we choose to perform transcriptomic expression analysis on hippocampus. After enhancing the stringency and the power of the statistical tests, the analysis highlighted 192 deregulated protein-coding genes, half of them down-regulated, the other half up-regulated. Functional annotation analysis revealed enrichment for genes related to cell adhesion pathways (*App, Cntnap5b, Lgals3bp, Mag, Mcam, Npnt, Pcdhb2/3/4/6/7/8/16 and Vwf*). An important number of protocadherins, part of a gene cluster located on Mmu18, appeared up-regulated in Ms5Yah hippocampi. Some of this cell-cell adhesion proteins have been involved in intracellular signaling pathways associated with neuropsychiatric disorders and cognitive impairments (Redies et al., 2012). Notably, de novo mutations in *Pcdhb4*, found in our deregulated list, have been associated with sporadic autism spectrum disorders (O'Roak et al., 2012). Over the 45 genes of the *App-Runx1* region referenced on chips, 33 were found down-regulated including *App* which is implicated in many neuronal processes like synaptogenesis (Wang et al., 2009), neural plasticity and memory (Turner et al., 2003). Little is known about heterozygotes but *App*-deficient mice are viable, fertile, present a 15-20% decrease in weight, reduced forelimb grip

strength and impaired locomotor activity (Zheng et al., 1995) which is highly similar to Ms5Yah phenotypes. Aged *App*-knockout aged mice also show impairments in learning and memory associated with a deficit in hippocampal LTP (Ring et al., 2007). The two basic helix-loop-helix oligodendrocyte transcription factors *Olig1* and *Olig2* were also found down-regulated. These genes are implicated in the generation of motoneurons and oligodendrocytes (Lu et al., 2002). *Olig1/2* triplication is associated with developmental brain defects in DS (Chakrabarti et al., 2010). In *Olig1/2*-deficient mice, motoneurons are converted to V2 interneurons in the spinal cord and oligodendrocytes fail to differentiate throughout the nervous system (Zhou and Anderson, 2002). Under-expression of *Olig1* and *Olig2* may probably have an implication in Ms5Yah abnormal motor coordination. Our transcriptomic study revealed an important up-regulation of *Klk6* gene encoding for kallikrein-related peptidase 6, a biomarker of Alzheimer's disease (Diamandis et al., 2000). Interestingly, in vitro substrates of this serine protease include the amyloid precursor protein (Ashby et al., 2010), reinforcing the idea that *App* is an interesting candidate gene for Ms5Yah phenotypes. *Cbs* encoding for cystathionine β-synthase was also found over-expressed in Ms5Yah hippocampi. In human, this gene located on a different interval of Hsa21 is a good candidate for playing a role in DS cognitive traits (Pogribna et al., 2001), as shown by several studies in mice (Lopes Pereira et al., 2009; Regnier et al., 2012; Sahún et al., 2014). Considering the detailed expression analysis, we believe that further molecular and electrophysiological studies could give important information about impaired hippocampal functions of Ms5Yah mice seen in the Morris water maze test.

Overall the results presented here in the mouse confirmed the critical importance of *APP-RUNX1* region in partial monosomy 21 (Chettouh et al., 1995; Lyle et al., 2009). When PM21 patients present post-natal growth retardation, short stature, cranio-facial malformations, psychomotor retardation and intellectual disability, Ms5Yah mice show developmental delay, reduction of size and weight, thrombocytopenia, motor coordination deficit, spatial learning and memory impairments. To date, no patient case carrying a deletion for the entire *APP-RUNX1* region has been described in literature suggesting that this monosomic state is possibly incompatible with human survival. A meta-analysis of PM21 cases revealed that patients carrying large deletions in the first region of Hsa21q (from the centromere to ~ 31.2Mb) show severe phenotypes described above (Roberson et al., 2011). A majority of these large deletions encompass a part or the total *APP-SOD1* region. Less than 5 cases carrying partial deletions in the *SOD1-RUNX1* region have been identified. Combined

information identified a 0.56 Mb critical region containing four genes, *KCNE1, RCAN1, CLIC6* and *RUNX1*, associated with severe congenital heart defects (Click et al., 2011; Lindstrand et al., 2010; Lyle et al., 2009). PM21 deletion encompassing *RUNX1* (Shinawi et al., 2008) and *RUNX1* haplodeficiency (Jalagadugula et al., 2010; Liew and Owen, 2011) are also commonly associated with thrombocytopenia which is consistent our hematology analysis. By using genetic approaches to restore the expression of genes included in *App-Runx1* region, we think that our mouse model will help in the identification of candidate genes responsible for partial monosomy 21 symptoms and for the development of therapeutic treatments.

MATERIALS AND METHODS

Mouse lines, genotyping and ethical statement

Ms5Yah mice, also named Del(16App-Runx1)5Yah, were kept and bred on an F1 B6NC3B background. The model was generated through Cre-LoxP in vivo recombination (Brault et al., 2006) using a mouse line carrying two loxP sites inserted at *App* and *Runx1* in a cis configuration as described previously (Raveau et al., 2012). Ms5Yah allele was identified by PCR using one Fwd primer (5'-ATCCGGGAATGGTCCCTA-3') specific for the wt allele, one Fwd primer (5'-CAAGCACTGGCTATGCATGT-3') specific for the Ms5Yah allele and a Ms5Yah/wt Rev (5'-GTTCGTTGCCTGAAGGAGAG-3') primer common to both alleles. PCR reactions gave wt and mutant products of 482 bp and 328 bp respectively. Experimental procedures were approved by the local ethical committee Com'Eth under accreditation number (2012-069). YH is the principal investigator of this study, (accreditation 67-369).

Viability and early post-natal tests

Viability tests were done in order to study causes of early post-natal lethality. 52 fetuses at E18.5 embryonic stage were isolated from 7 pregnant females, one day prior to natural delivery, to monitor their capacities to survive at birth. Fetuses were placed on a warm plate at 37°C and rolled gently to stimulate them to breathe. 30 minutes after extraction, numbers of pink and moving animals versus cyanotic animals unable to breathe were counted. Tail samples were collected for genotyping.

Litters of 4 females including 8 Ms5Yah and 16 wild-type animals were evaluated for righting reflex. Briefly, animals were put on the back and time needed to recover their posture was measured. The test was done every day from P2 to P10 post-partum. In order to study weight evolution, body weights were recorded once a week from the age of 1 to 13 weeks.

Behavioral analysis

A pipeline of behavioral experiments was conducted in order to evaluate activity, memory, and locomotor coordination of mice. Majority of experimental procedures have been described previously in Duchon et al. (Duchon et al., 2011). To produce experimental groups, only animals coming from litters containing a minimum of two male pups were selected. A

cohort including 11 Ms5Yah and 13 controls was generated. After weaning, animals were gathered by litters in 39 x 20 x 16 cm cage (Green Line, Techniplast, Italy) where animals had free access to water and food (D04 chow diet, Safe). Temperature was maintained at 23±1°C and the light cycle was controlled on a 12 hour light and 12 hour dark cycle. Mice were transferred from animal facility to phenotyping area at the age of 12 weeks and were screened into behavioral tests from 14 to 19 weeks. Animals were transferred to experimental room antechambers 30 min before each experiment. All the experiments were done between 8:00 AM and 1:00 PM. Tests were conducted in the following order: elevated plus maze (at age 14 weeks), open field (at age 14 weeks), novel object recognition (at age 15 weeks), Y maze (at age 15 weeks), rotarod (at age 16 weeks), notched bar (at age 17 weeks), grip strength (at age 17 weeks) and Morris water maze (at age 18-19 weeks).

Elevated plus maze: The apparatus consists of two opposed open arms (30 x 5 cm) crossed by two enclosed arms (30 x 5 x 15 cm), and elevated 66 cm from the floor. The light intensity at the extremity of the open arms was kept at 50 Lux. Each mouse was tested for 5 min after being placed in the central platform and allowed to explore freely the apparatus. The number of entries and time spent in the open arms were used as an anxiety index. Closed arm entries and rears in the closed arms were used as measures of general motor activity.

Open-field (OF): The test is used to evaluate exploration behavior. Mice were tested in automated open fields (44.3 x 44.3 x 16.8 cm) made of PVC with transparent walls and a black floor, and covered with translucent PVC (Panlab, Barcelona, Spain). The open field arena is divided into central and peripheral regions and is homogeneously illuminated at 150 Lux. Each mouse was placed in the periphery of the open field and allowed to explore freely the apparatus for 30 min. The distance travelled, the number of rears and time spent in the central and peripheral parts of the arena were recorded over the test session.

Novel object recognition task (NOR): This test evaluates object recognition memory and is based on the innate preference of rodents to explore novelty. The test was carried in an open field arena as previously described (Goeldner et al., 2008). On the first day, mice were habituated to the arena for 30 min at 60 Lux. The following day, animals were submitted to a first 10-min acquisition trial during which they were individually placed in the presence of the object A (marble or dice) placed at 10cm of one corner of the box. The exploration time of object A (when the animal's snout was directed toward the object at a distance $\leq$1 cm) was recorded. A 10-min retention trial (second trial) was conducted 1 hour later. The familiar

object (object A) and novel object (object B) were placed at 10 cm of two open field corner (the distance between the two objects was about 20 cm) and the exploration time of the two objects was recorded. A discrimination index was defined as $(t_B/(t_A + t_B)) \times 100$. All mice which did not explore the first object more than 3 seconds during the acquisition trial were excluded from the analysis.

Y-maze: This test which is used to evaluate short-term working memory is based on the innate preference of animals to explore arms that has not been previously explored, a behavior that, if occurring with a frequency greater than 50%, is called spontaneous alternation. The apparatus consists in a Y-shaped maze with three white, opaque plexiglas arms of equivalent length forming a 120° angle with each other. The arms have walls with specific motifs allowing distinction from each other. After introduction at the center of the maze at 60 Lux, animal were allowed to freely explore the maze for 6 minutes. The number of arm entries and the number of triads were recorded in order to calculate the percentage of alternation.

Rotarod: The test assesses sensorimotor coordination and balance. The apparatus (Biosed, France) is a rotating bar of 5 cm diameter (hard plastic materiel covered by grey rubber foam) on which mice are placed facing the direction of rotation. Animals were first habituated to stay on the rod for 30 seconds at a constant speed of 4 rotations per minute (rpm). This was followed by 3 training days with 4 trials per day. Mice were placed on an accelerating rod increasing from 4 rpm to 40 rpm in 5 min of time at 100 Lux. The test was stopped when the mouse fell down from the rod or when they wade more than one passive rotation. The latency to fall and the maximal speed before falling was recorded. On the fourth day, mice had two test sessions consisting of 6 consecutive trials of 2 min at constant speed, with increasing speed between each trial (4, 10, 16, 22, 28, 34 and 40 rpm). For each trial, latencies to fall off the rod were recorded. A maximum latency of 2 minutes was attributed for each animal which did not fall during the test.

Notched bar: The test is used for testing the hindlimb coordination. Testing was performed as previously described (Duchon et al., 2011). Mice were trained with a bar consisting of a natural wooden piece 1.7 mm large and 50 cm long bearing terminal platforms of 6 cm × 6 cm. The second day, mice were tested at 100 Lux with a notched bar of same dimension constituted of 12 platforms of $2cm^2$ spaced out by 13 gaps of $2cm^3$. Animals had to cross the notched bar twice for training and 15 times for test. Each time a back paw went through the gap was considered as an error and global error percentage was calculated.

Grip strength: The test is used to evaluate muscular strength. Mice were first weighted and tested with a handy force gauge (Bioseb, France). Animals were placed on the instrument grid and pulled by the tail until letting go. The force (g) was related to animal weight (g).

Morris water maze paradigm (MWM): The test is a paradigm for spatial learning and memory. The protocol is adapted from the already established protocol (Duchon et al., 2011). The apparatus consists in a circular pool (150-cm diameter, 60-cm height) filled to a depth of 40 cm with water maintained at 20°C–22°C and made opaque using a white aqueous emulsion (Acusol OP 301 opacifier). An escape platform, made of 6 cm diameter rough plastic, is submerged 1 cm below the water surface. The test began with 6 days of acquisition, 4 trials per day, at 120 Lux. Each trial started with the mice facing the interior wall of the pool and ended when animals climb on the platform or after a maximum searching time of 90 sec. The platform was at the same position for all the four trials but starting positions changed randomly between each trial with departures from each cardinal point. Travelled distances to find the platform and swimming speeds were analyzed each day. On 7^{th} day, mice were given a single trial of 60 seconds trial during the probe test or removal session in which the platform has been removed. The distance traveled and duration spent in each quadrant (NW, NE, SW, SE) were recorded. Annulus crossing index was calculated as the number of times that animals crossed the exact platform position. On 12^{th} and 13^{th} days, mice were given reversal sessions with 4 trials of 90 seconds per day. The platform was made visible by a small dark ball placed 12cm on top of the platform, while the external cues were hidden by surrounding the pool with a black curtain. In order to be sure that the mouse used the platform cue, starting position and platform position were changed for each trial.

Optomotor response

Mice (3 wild-type and 2 Ms5Yah) were placed on a platform (11.5 cm diameter grid, 19.0 cm above the bottom), on the axis of a rotating drum (29.0 cm diameter, 2 rotations/minute), with inside surface covered with alternating black and white stripes. The stripe spatial frequency was set alternatively at 0.25, 0.51 and 1.27 and 0.13 cycles/degree. The light intensity was adjusted to 120 Lux, as in the Morris water maze test. Mice were visualized through a digital video camera (Sony, DCR-TRV24E) placed above. For each spatial frequency, the drum was rotated alternately clockwise and anticlockwise, for one minute or until two clear head-tracking movements were scored by two observers.

Blood hematology and bleeding time

Blood was collected from 25-weeks old males (11 Ms5Yah and 11 wt littermates) by retro orbital puncture under isoflurane anesthesia at 12.00 am after 4 hours fasting. 120 microliters of blood was collected into EDTA-coated pediatric Microvette tubes for hematological analysis. A complete blood count including total erythrocytes, leukocytes and platelets count; differential leukocytes count (granulocytes: neutrophils, eosinophils and basophils, lymphocytes, and monocytes); hemoglobin and hematocrit measurement; and calculation of blood indexes (mean cell volume, mean corpuscular hemoglobin concentration) was performed on total blood on an Advia 120 Vet. (Siemens). To evaluate if platelet deficit of Ms5Yah mice have an influence on coagulation process, bleeding time was measured in the same animals at 27 weeks of age. After having 5 mm distal tips transected, the tail was immersed into phosphate buffered saline at 37°C. Bleeding time was measured from incision to first stop of bleeding.

Histology

To evaluate E18.5 fetuses anatomy, 1 Ms5Yah fetus unable to breathe, 1 Ms5Yah fetus with craniorachischisis and 1 wild-type fetus were fixed at the end of viability test and subjected to histological analysis. Fetuses were decalcified in DC3 solution before being embedded in paraffin and then preceded in 7µm-thick sections stained by modified Mallory's tetrachrome (hematoxylin, acid fuschin, aniline blue, and Orange G). For cerebellum histological and anatomical studies, 5 Ms5Yah and 4 controls brains were collected, fixed in 10% formalin for 48 hours and embedded in paraffin. 5µm-thick transversal sections through the middle of the cerebellum were performed and were stained with hematoxylin & eosin. Stained sections were digitalized using a slide scanner (Nanozoomer 2.0-HT , *Hamamatsu* , Japan) and measurements of the cerebellar vermis and hemisphere thickness relative to whole thickness were performed using the NDPview software of the digital scanner.

Total RNA extraction

For Affymetrix arrays, hippocampus were isolated from Ms5Yah (n=9) and wt control (n=8) 4-months animals and flash frozen. Total RNA was prepared using Trizol (Invitrogen) and purified with the RNeasy Mini Kit (Quiagen) according to the manufacturer's instructions. Samples quality was checked using an Agilent 2100 Bioanalyzer (Agilent Technologies).

Whole-genome expression arrays

In conjunction with the Affymetrix GeneChip WT Terminal Labeling Kit, the Ambion WT Expression Kits was designed to generate amplified and biotylated sense-strand DNA targets from the entire expressed genome and hybridized onto GeneChip Mouse Exon 1.0ST arrays (Affymetrix). Chips were washed and scanned using the GeneChip Scanner 3000. Digitized images were generated with AGCC v3.2 (Affymetrix) software and data files were generated with the Expression Console v1.1 (Affymetrix) software. Raw data were processed with the Robust Multiarray Average (RMA) algorithm developed by Irizarry et al. (Irizarry et al., 2003) and values log transformed using Partek (Partek Inc.) and GeneSpring (Agilent Technologies) software. Statistical analysis was performed using GeneSpring (one way ANOVA) and the 192 genes with a fold change above 1.2 or bellow 0.8 and a *p*-value below 0.05 were selected for clustering analysis. Hierarchical clustering was carried out with Cluster3.0 software (de Hoon et al., 2004) using Euclidian distances to calculate the distances between the genes and between the samples. Calculated distances were then clustered by complete linkage clustering. Post-hoc analysis using GeneSpring gave a list of statistically deregulated genes. Known mammalian phenotypes database (Mouse Genome informatics, Jackson Laboratory) and functional annotation clustering using Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatics were performed to estimate the potential impact of deregulated genes in transgenic mice. This latter tool mainly provides typical batch annotation and gene-GO term enrichment analysis to highlight the most relevant Gene Ontology (GO) terms associated with a given genes list (Huang et al., 2009a). The microarray data have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and are accessible through GEO Series accession number GSE58639.

Statistical analyses

Results were processed for statistical analysis using the Sigma Plot software. All acquired behavioral data were analyzed using a one-way or a two-way ANOVA for repeated measures followed by Student's t-test or Tukey's post-hoc test whenever it was appropriate. Otherwise, non-parametric Kruskal-Wallis analysis and Mann-Whitney *U* test were done. One group Student t-tests was used to compare recognition index value to chance level (50%). Pearson's chi-squared test was used to evaluate mutant allele transmission. Data are represented as the mean ± SEM and the significant threshold was $p < 0.05$ or otherwise indicated.

ACKNOWLEDGEMENTS

We thank Veronique Brault and Damien Marechal for their helpful comments and the animal care-takers and services of the ICS.

COMPETING UNTERESTS STATMENT

All the authors declare no conflict of interest regarding the material discussed in the manuscript.

AUTHOR CONTRIBUTIONS

Y.H. was responsible for the study design and conception and for obtaining funding for the study. A.D., M.R., C.C., V.N. and Y.H. made the Ms5Yah mouse model. T.A. did viability tests and behavioral analyses. C.C. did righting reflex test and weight evolution study. H.J. and O.W. performed histological analysis. M.R. performed and analyzed optomotor response test. T.A. did transcriptional analysis with the assistance of D.D. T.A. and Y.H wrote the paper. All authors reviewed and contributed to the report. The study was supervised by Y.H.

FUNDINGS

This work was supported by the National Centre for Scientific Research (CNRS), the French National Institute of Health and Medical Research (INSERM), the University of Strasbourg (UDS), the “Centre Européen de Recherche en Biologie et en Médecine”, the European commission (AnEUploidy project to YH, LSHG-CT-2006-037627) with a fellowship from the “Fondation pour la Recherche Médicale” to TA.

REFERENCES

Ashby, E.L., Kehoe, P.G., and Love, S. (2010). Kallikrein-related peptidase 6 in Alzheimer's disease and vascular dementia. Brain Research *1363*, 1-10.

Besson, V., Brault, V., Duchon, A., Togbe, D., Bizot, J.C., Quesniaux, V.F.J., Ryffel, B., and Herault, Y. (2007). Modeling the monosomy for the telomeric part of human chromosome 21 reveals haploinsufficient genes modulating the inflammatory and airway responses. Human Molecular Genetics *16*, 2040-2052.

Brault, V., Pereira, P., Duchon, A., and Herault, Y. (2006). Modeling chromosomes in mouse to explore the function of genes, genomic disorders, and chromosomal organization. Plos Genetics *2*, 911-919.

Burgess, T., Downie, L., Pertile, M.D., Francis, D., Glass, M., Nouri, S., and Pszczola, R. (2014). Monosomy 21 seen in live born is unlikely to represent true monosomy 21: a case report and review of the literature. Case Rep Genet *2014*, 965401.

Chakrabarti, L., Best, T.K., Cramer, N.P., Carney, R.S.E., Isaac, J.T.R., Galdzicki, Z., and Haydar, T.F. (2010). Olig1 and Olig2 triplication causes developmental brain defects in Down syndrome. Nature Neuroscience *13*, 927-U939.

Chettouh, Z., Croquette, M.F., Delobel, B., Gilgenkrants, S., Leonard, C., Maunoury, C., Prieur, M., Rethore, M.O., Sinet, P.M., Chery, M.*, et al.* (1995). MOLECULAR MAPPING OF 21 FEATURES ASSOCIATED WITH PARTIAL MONOSOMY-21 - INVOLVEMENT OF THE APP-SOD1 REGION. American Journal of Human Genetics *57*, 62-71.

Click, E.S., Cox, B., Olson, S.B., Grompe, M., Akkari, Y., Moreau, L.A., Shimamura, A., Sternen, D.L., Liu, Y.J.J., Leppig, K.A.*, et al.* (2011). Fanconi Anemia-Like Presentation in an Infant with Constitutional Deletion of 21q Including the RUNX1 Gene. American Journal of Medical Genetics Part A *155A*, 1673-1679.

de Hoon, M.J.L., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. Bioinformatics *20*, 1453-1454.

Diamandis, E.P., Yousef, G.M., Petraki, C., and Soosaipillai, A.R. (2000). Human kallikrein 6 as a biomarker of Alzheimer's disease. Clinical Biochemistry *33*, 663-667.

Duchon, A., Pothion, S., Brault, V., Sharp, A.J., Tybulewicz, V.L.J., Fisher, E.M.C., and Herault, Y. (2011). The telomeric part of the human chromosome 21 from Cstb to Prmt2 is not necessary for the locomotor and short-term memory deficits observed in the Tc1 mouse model of Down syndrome. Behavioural Brain Research *217*, 271-281.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research *30*, 207-210.

Epstein, C.J. (1990). THE CONSEQUENCES OF CHROMOSOME IMBALANCE. American Journal of Medical Genetics, 31-37.

Goeldner, C., Reiss, D., Wichmann, J., Meziane, H., Kieffer, B.L., and Ouagazzal, A.M. (2008). Nociceptin receptor impairs recognition memory via interaction with NMDA receptor-dependent mitogen-activated protein kinase/extracellular signal-regulated kinase signaling in the hippocampus. J Neurosci *28*, 2190-2198.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research *37*, 1-13.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols *4*, 44-57.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics *4*, 249-264.

Jalagadugula, G., Mao, G.F., Kaur, G., Goldfinger, L.E., Dhanasekaran, D.N., and Rao, A.K. (2010). Regulation of platelet myosin light chain (MYL9) by RUNX1: implications for thrombocytopenia and platelet dysfunction in RUNX1 haplodeficiency. Blood *116*, 6037-6045.

Korenberg, J.R., Chen, X.N., Schipper, R., Sun, Z., Gonsky, R., Gerwehr, S., Carpenter, N., Daumer, C., Dignan, P., Disteche, C.*, et al.* (1994). DOWN-SYNDROME PHENOTYPES - THE CONSEQUENCES OF CHROMOSOMAL IMBALANCE. Proceedings of the National Academy of Sciences of the United States of America *91*, 4997-5001.

Lejeune, J., Berger, R., Rethore, M.O., Archamba.L, Jerome, H., Thieffry, S., Aicardi, J., Broyer, M., Lafourca.J, Cruveill.J*, et al.* (1964). MONOSOMIE PARTIELLE POUR UN PETIT ACROCENTRIQUE. Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences *259*, 4187-&.

Liew, E., and Owen, C. (2011). Familial myelodysplastic syndromes: a review of the literature. Haematologica-the Hematology Journal *96*, 1536-1542.

Lindstrand, A., Malmgren, H., Sahlen, S., Schoumans, J., Nordgren, A., Ergander, U., Holm, E., Anderlid, B.M., and Blennow, E. (2010). Detailed molecular and clinical characterization of three patients with 21q deletions. Clinical Genetics *77*, 145-154.

Lopes Pereira, P., Magnol, L., Sahún, I., Brault, V., Duchon, A., Prandini, P., Gruart, A., Bizot, J., Chadefaux-Vekemans, B., Deutsch, S.*, et al.* (2009). A new mouse model for the trisomy of the Abcg1-U2af1 region reveals the complexity of the combinatorial genetic code of down syndrome. Hum Mol Genet *18*, 4756-4769.

Lu, Q.R., Sun, T., Zhu, Z.M., Ma, N., Garcia, M., Stiles, C.D., and Rowitch, D.H. (2002). Common developmental requirement for Olig function indicates a motor neuron/oligodendrocyte connection. Cell *109*, 75-86.

Lyle, R., Bena, F., Gagos, S., Gehrig, C., Lopez, G., Schinzel, A., Lespinasse, J., Bottani, A., Dahoun, S., Taine, L.*, et al.* (2009). Genotype-phenotype correlations in Down syndrome identified by array CGH in 30 cases of partial trisomy and partial monosomy chromosome 21. European Journal of Human Genetics *17*, 454-466.

Megarbane, A., Ravel, A., Mircher, C., Sturtz, F., Grattau, Y., Rethore, M.O., Delabar, J.M., and Mobley, W.C. (2009). The 50th anniversary of the discovery of trisomy 21: The past, present, and future of research and treatment of Down syndrome. Genetics in Medicine *11*, 611-616.

Migdalska, A.M., van der Weyden, L., Ismail, O., White, J.K., Sanchez-Andrade, G., Logan, D.W., Arends, M.J., Adams, D.J., and Sanger Mouse Genetics, P. (2012). Modeling Partial Monosomy for Human Chromosome 21q11.2-q21.1 Reveals Haploinsufficient Genes Influencing Behavior and Fat Deposition. Plos One *7*.

O'Roak, B.J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B.P., Levy, R., Ko, A., Lee, C., Smith, J.D.*, et al.* (2012). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature *485*, 246-U136.

Olson, L.E., Roper, R.J., Sengstaken, C.L., Peterson, E.A., Aquino, V., Galdzicki, Z., Siarey, R., Pletnikov, M., Moran, T.H., and Reeves, R.H. (2007). Trisomy for the Down syndrome 'critical region' is necessary but not sufficient for brain phenotypes of trisomic mice. Human Molecular Genetics *16*, 774-782.

Pogribna, M., Melnyk, S., Pogribny, I., Chango, A., Yi, P., and James, S.J. (2001). Homocysteine metabolism in children with Down syndrome: In vitro modulation. American Journal of Human Genetics *69*, 88-95.

Raveau, M., Lignon, J.M., Nalesso, V., Duchon, A., Groner, Y., Sharp, A.J., Dembele, D., Brault, V., and Hérault, Y. (2012). The app-runx1 region is critical for birth defects and electrocardiographic dysfunctions observed in a down syndrome mouse model. PLoS Genet *8*, e1002724.

Redies, C., Hertel, N., and Hubner, C.A. (2012). Cadherins and neuropsychiatric disorders. Brain Research *1470*, 130-144.

Regnier, V., Billard, J.M., Gupta, S., Potier, B., Woerner, S., Paly, E., Ledru, A., David, S., Luilier, S., Bizot, J.C.*, et al.* (2012). Brain Phenotype of Transgenic Mice Overexpressing Cystathionine beta-Synthase. Plos One *7*.

Ring, S., Weyer, S.W., Kilian, S.B., Waldron, E., Pietrzik, C.U., Filippov, M.A., Herms, J., Buchholz, C., Eckman, C.B., Korte, M.*, et al.* (2007). The secreted beta-amyloid precursor protein ectodomain APPs alpha is sufficient to rescue the anatomical, behavioral, and electrophysiological abnormalities of APP-deficient mice. Journal of Neuroscience *27*, 7817-7826.

Roberson, E.D.O., Wohler, E.S., Hoover-Fong, J.E., Lisi, E., Stevens, E.L., Thomas, G.H., Leonard, J., Hamosh, A., and Pevsner, J. (2011). Genomic analysis of partial 21q monosomies with variable phenotypes. European Journal of Human Genetics *19*, 235-238.

Sahún, I., Marechal, D., Lopes Pereira, P., Nalesso, V., Gruart, A., Delgado Garcia, J.M., Antonarakis, S.E., Dierssen, M., and Herault, Y. (2014). Cognition and Hippocampal Plasticity in the Mouse Is Altered by Monosomy of a Genomic Region Implicated in Down Syndrome. Genetics.

Schinzel, A. (1976). DOES FULL MONOSOMY-21 EXIST - COMMENT TO PAPER - MALE INFANT WITH MONOSOMY-21. Human Genetics *32*, 105-106.

Shinawi, M., Erez, A., Shardy, D.L., Lee, B., Naeem, R., Weissenberger, G., Chinault, A.C., Cheung, S.W., and Plon, S.E. (2008). Syndromic thrombocytopenia and predisposition to acute myelogenous leukemia caused by constitutional microdeletions on chromosome 21q. Blood *112*, 1042-1047.

Theodoropoulos, D.S., Cowan, J.M., Elias, E.R., and Cole, C. (1995). PHYSICAL FINDINGS IN 21Q22 DELETION SUGGEST CRITICAL REGION FOR 21Q- PHENOTYPE IN Q22. American Journal of Medical Genetics *59*, 161-163.

Toral-Lopez, J., Gonzalez-Huerta, L.M., and Cuevas-Covarrubias, S.A. (2012). Complete monosomy mosaic of chromosome 21: Case report and review of literature. Gene *510*, 175-179.

Turner, P.R., O'Connor, K., Tate, W.P., and Abraham, W.C. (2003). Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory. Progress in Neurobiology *70*, 1-32.

Valero, R., Marfany, G., Gil-Benso, R., Ibanez, M.D., Lopez-Pajares, I., Prieto, F., Rul-lan, G., Sarret, E., and Gonzalez-Duarte, R. (1999). Molecular characterisation of partial chromosome 21 aneuploidies by fluorescent PCR. Journal of Medical Genetics *36*, 694-699.

Wang, Z.L., Wang, B.P., Yang, L., Guo, Q.X., Aithmitti, N., Zhou, S.Y., and Zheng, H. (2009). Presynaptic and Postsynaptic Interaction of the Amyloid Precursor Protein Promotes Peripheral and Central Synaptogenesis. Journal of Neuroscience *29*, 10788-10801.

Yu, T., Clapcote, S.J., Li, Z.Y., Liu, C.H., Pao, A., Bechard, A.R., Carattini-Rivera, S., Matsui, S.I., Roder, J.C., Baldini, A.*, et al.* (2010). Deficiencies in the region syntenic to human 21q22.3 cause cognitive deficits in mice. Mammalian Genome *21*, 258-267.

Zheng, H., Jiang, M.H., Trumbauer, M.E., Sirinathsinghji, D.J.S., Hopkins, R., Smith, D.W., Heavens, R.P., Dawson, G.R., Boyce, S., Conner, M.W.*, et al.* (1995). BETA-AMYLOID PRECURSOR PROTEIN-DEFICIENT MICE SHOW REACTIVE GLIOSIS AND DECREASED LOCOMOTOR-ACTIVITY. Cell *81*, 525-531.

Zhou, Q., and Anderson, D.J. (2002). The bHLH transcription factors OLIG2 and OLIG1 couple neuronal and glial subtype specification. Cell *109*, 61-73.

LEGENDS TO FIGURES

Figure 1. Viability tests and anatomical characterization of E18.5 fetuses. (**A-C**) Pictures of neonates, 30 min after caesarean section. Healthy fetuses moved and turned pink in the first minutes after delivery. (**A**) Wild-type and (**B**) cyanotic Ms5Yah littermates. (**C**) Ms5Yah fetus presenting craniorachischisis. (**D-G**) Anatomical characterization. Lungs from a wild-type fetus at low (**D**, scale = 3 mm) and high magnifications (**F**, scale = 400 µm) present opened alveoli as fetus was breathing before sacrifice. Lungs from a Ms5Yah fetus unable to breathe at low (**E**) and high (**G**) magnifications present collapsed alveoli.

Figure 2. Coordination capacities of newborn animals and evolution of the weight. (**A**) Righting reflex of new-born mice. Graphs plot the time (s) to recover natural posture when animals are putted on their back. (**B**) Evolution of body weight (g) from 1 to 13 weeks of age. (**C**) 10 weeks-old wild-type and Ms5Yah (behind) littermates. All graphs depict mean ± SEM. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Figure 3. Locomotor coordination and grip strength capacities. (**A-B**) Training phase on rotarod test. (**A**) Results are expressed as time (s) mice remained on an accelerating rod (4-40 rpm over 5 min) before falling. (**B**) Mean rotational velocity (rpm) at the time of falling. (**C**) Test phase. The graphs plot the time (s) mice stayed on the rod when tested at constant speeds between 4 and 40 rpm. (**D**) Notched bar test. Results are expressed percentage of hind paw errors made by mice when crossing the bar. (**E**) Four-paw grip test. The conclusion of these tests is that Ms5Yah show impaired motor coordination without alteration of grip strength. All graphs depict mean ± SEM. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Figure 4. Spatial learning and memory performances in the Morris water maze (MWM) test. **(A)** Distance travelled to find the platform along acquisition (A1 – A6) and reversal (Cued1 and Cued2) sessions. **(B)** Corresponding swimming speeds. **(C-D)** Removal session. Mice are scored for the percentage of time spent in the different quadrant of the MWM **(C)** and the annulus crossing **(D)** during one single trial where the platform was removed. Ms5Yah mice present delay in the acquisition of the task and travelled a higher distance to find the platform. In addition to spatial learning impairment, removal session indicates a deficit in accurate spatial memory as Ms5Yah mice had difficulty to remind the exact position of the platform. All graphs depict mean ± SEM. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Figure 5. Microarray expression analyses on hippocampi. (**A**) Clustering derived from statistically deregulated genes in Ms5Yah mice. Under-expression and over-expression are represented in green and red respectively, expression levels were calculated by comparison to mean expression level of all arrays for each gene. (**B**) Expression profile of *App-Runx1* genes referenced on chips. Most of genes have a fold change (FC) ranging from 0.5 to 0.8. Some genes of the *Krtap* cluster present a FC above threshold which is probably due to the fact that several Krtap RNAs hybridize to the same probes.

Supplementary Figure 1. Exploration activity, short-term working memory and recognition memory performances. (**A-B**) Open field test. Results are expressed as mean distance travelled (m) (**A**), and time percentage spent in the central area (**B**) over a 30 min test. (**C-D**) Elevated plus maze test. Number of arm entries (**C**) and percentage of time spent in the open arms (**D**). No difference in anxiety and exploration pattern was noticed between Ms5Yah and controls. (**E-F**) Y maze test. Number of arm entries (**E**) and percentage of alternation (**F**) over a 6 min test. (**G-H**) Novel object recognition (NOR) test. (**G**) Exploration time (s) of the first object. (**H**) Discrimination index. Data indicate that Ms5Yah mice present similar recognition memory capacities than controls. All graphs depict mean ± SEM. $^{*}p < 0.05$.

Supplementary Figure 2. Transversal sections of the cerebellum stained with hematoxylin and eosin. Comparison of wild-type (**A**, **C**) and Ms5Yah (**B**, **D**) sections at low (**A**, **B**, scale = 2 mm) and high magnification (**C**, **D**, scale = 200 µm) through the middle of the cerebellum revealed no anatomical abnormality. At the histological level, all layers of the mutant cerebellum including the molecular, granular and Purkinje cell layers appeared normal.

TABLES

Upregulated genes			**Downregulated genes**		
Gene	**Chr**	**mean ± sem**	**Gene**	**Chr**	**mean ± sem**
Myoc	1	2.49 ± 0.32 ***	*Slc12a2*	18	0.80 ± 0.02 ***
Klk6	7	1.70 ± 0.11 ***	*3110052M02Rik*	17	0.80 ± 0.04 **
Lct	1	1.57 ± 0.20 **	*Rhog*	7	0.79 ± 0.03 ***
Gpr115	17	1.55 ± 0.22 *	*Bach1*	16	0.79 ± 0.03 ***
Olfr690	7	1.53 ± 0.12 **	*Mag*	7	0.79 ± 0.02 ***
Fpr-rs3	17	1.53 ± 0.13 **	*Cd74*	18	0.79 ± 0.05 *
Olfr46	7	1.49 ± 0.12 *	*Kcne2*	16	0.79 ± 0.05 *
Ccl25	8	1.48 ± 0.05 ***	*Rbm43*	2	0.79 ± 0.04 **
Btg2	1	1.48 ± 0.14 *	*Pde8a*	7	0.79 ± 0.03 **
Pisd-ps3	11	1.48 ± 0.10 *	*Itih3*	14	0.78 ± 0.04 **
Pcdhb6	18	1.41 ± 0.06 ***	*Dgka*	10	0.78 ± 0.02 ***
Pcdhb2	18	1.40 ± 0.07 ***	*Prpf40b*	15	0.78 ± 0.01 ***
Cd180	13	1.39 ± 0.08 **	*Isyna1*	8	0.78 ± 0.03 ***
D930046H04Rik	7	1.38 ± 0.10 *	*Rnf5*	17	0.78 ± 0.06 **
Npas4	19	1.38 ± 0.10 *	*Scd3*	19	0.78 ± 0.03 **
Enpp6	8	1.35 ± 0.04 ***	*Pon3*	6	0.78 ± 0.03 *
Dnahc1	14	1.35 ± 0.04 ***	*Cryab*	9	0.78 ± 0.03 ***
Scn3a	2	1.34 ± 0.05 ***	*Glp1r*	17	0.78 ± 0.05 *
Cbs	17	1.34 ± 0.05 ***	*Dpp7*	2	0.78 ± 0.02 ***
Tdo2	3	1.34 ± 0.12 *	*Dhrs11*	11	0.78 ± 0.02 ***
Thrsp	7	1.33 ± 0.03 ***	*Alox8*	11	0.78 ± 0.03 ***
Pcdhb3	18	1.33 ± 0.11 *	*Olig2*	16	0.78 ± 0.03 ***
Gas5	1	1.32 ± 0.10 *	*Lgals3bp*	11	0.77 ± 0.02 ***
Tas2r107	6	1.31 ± 0.07 **	*Bmp2*	2	0.77 ± 0.05 **
Lce1b	3	1.31 ± 0.07 **	*Irf3*	7	0.77 ± 0.02 ***
A130014H13Rik	12	1.31 ± 0.04 ***	*Mcam*	9	0.77 ± 0.02 **
Pgam2	11	1.30 ± 0.05 **	*Krt73*	15	0.77 ± 0.03 ***
Ddc	11	1.30 ± 0.05 ***	*Rbm8a*	3	0.77 ± 0.08 *
Sema5a	15	1.30 ± 0.07 **	*Cd59a*	2	0.76 ± 0.07 *
Naaa	5	1.30 ± 0.04 ***	*Ifit1*	19	0.76 ± 0.05 *
Ccl6	11	1.30 ± 0.07 **	*Mal*	2	0.76 ± 0.02 ***
Cntnap5b	1	1.30 ± 0.08 **	*Mrpl48*	7	0.76 ± 0.07 *
Pnet-ps	18	1.29 ± 0.06 *	*Cdkl3*	11	0.76 ± 0.02 ***
Penk	4	1.29 ± 0.08 **	*Gjc3*	5	0.76 ± 0.02 ***
Gimap6	6	1.28 ± 0.11 *	*Tmem98*	11	0.75 ± 0.04 **
Hebp1	6	1.28 ± 0.03 **	*2900052N01Rik*	9	0.75 ± 0.03 ***

Upregulated genes			Downregulated genes		
Gene	**Chr**	**mean ± sem**	**Gene**	**Chr**	**mean ± sem**
Upk1b	16	1.28 ± 0.07 **	*Olig1*	16	0.75 ± 0.03 ***
Isoc2b	7	1.28 ± 0.05 ***	*H2-Aa*	17	0.74 ± 0.04 *
Spink10	18	1.28 ± 0.09 *	*Tspan2*	3	0.74 ± 0.02 ***
Heatr5b	17	1.27 ± 0.09 *	*Ptgds*	2	0.74 ± 0.03 ***
Lmcd1	6	1.26 ± 0.05 **	*Vamp1*	6	0.74 ± 0.02 ***
Rpsa	9	1.26 ± 0.08 **	*Slc4a2*	5	0.74 ± 0.03 **
Serpind1	16	1.26 ± 0.04 ***	*Aspa*	11	0.74 ± 0.03 **
Gpr126	10	1.26 ± 0.02 ***	*Ppp2r3a*	9	0.73 ± 0.08 *
5330431K02Rik	13	1.26 ± 0.03 ***	*Cdk10*	8	0.73 ± 0.05 **
Unc13c	9	1.25 ± 0.09 *	*Cyp4f15*	17	0.72 ± 0.03 **
Olfr266	3	1.25 ± 0.09 *	*Efhd1*	1	0.72 ± 0.03 ***
Kcng3	17	1.25 ± 0.09 *	*1500015L24Rik*	19	0.72 ± 0.04 **
Pcdhb16	18	1.24 ± 0.08 *	*Cldn1*	16	0.72 ± 0.06 *
Cbr3	16	1.24 ± 0.03 ***	*Plekhf1*	7	0.71 ± 0.06 *
Npnt	3	1.24 ± 0.04 ***	*Car14*	3	0.71 ± 0.03 **
Smoc2	17	1.24 ± 0.06 **	*Enpp2*	15	0.71 ± 0.07 *
Sspn	6	1.23 ± 0.05 **	*2610039C10Rik*	16	0.71 ± 0.02 ***
Tiaf2	15	1.23 ± 0.07 *	*Il33*	19	0.71 ± 0.03 ***
P2ry13	3	1.23 ± 0.03 **	*Fam165b*	16	0.70 ± 0.02 ***
9530091C08Rik	9	1.23 ± 0.05 *	*Sulf1*	1	0.69 ± 0.06 *
Foxf1a	8	1.23 ± 0.07 *	*App*	16	0.68 ± 0.01 ***
Pcdhb7	18	1.23 ± 0.07 *	*Tceb2*	17	0.68 ± 0.07 *
Pycard	7	1.23 ± 0.04 **	*Opalin*	19	0.67 ± 0.05 **
Slc24a4	12	1.23 ± 0.04 **	*Ace*	11	0.67 ± 0.06 *
Gpc4	X	1.23 ± 0.06 **	*Selk*	14	0.67 ± 0.10 *
Olfr149	9	1.23 ± 0.06 *	*Donson*	16	0.66 ± 0.02 ***
Hvcn1	5	1.23 ± 0.05 **	*5930434B04Rik*	2	0.65 ± 0.07 **
Fst	13	1.23 ± 0.08 *	*Il10rb*	16	0.64 ± 0.03 ***
Pdyn	2	1.22 ± 0.04 **	*Tiam1*	16	0.63 ± 0.01 ***
Syt10	15	1.22 ± 0.04 *	*Snx32*	19	0.63 ± 0.01 ***
Nrp2	1	1.22 ± 0.05 **	*Abca8a*	11	0.62 ± 0.02 ***
Pcdhb4	18	1.22 ± 0.06 **	*Son*	16	0.62 ± 0.01 ***
Dmrtc1a	X	1.22 ± 0.04 ***	*Srsf15*	16	0.62 ± 0.01 ***
Hist2h4	3	1.22 ± 0.07 *	*Hunk*	16	0.61 ± 0.02 ***
Gypc	18	1.22 ± 0.04 **	*Rwdd2b*	16	0.61 ± 0.03 ***
Hopx	5	1.21 ± 0.04 **	*Itsn1*	16	0.61 ± 0.01 ***

Upregulated genes			Downregulated genes		
Gene Name	**Chr**	**mean ± sem**	**Gene Name**	**Chr**	**mean ± sem**
Pcdhb8	18	1.21 ± 0.05 *	Rcan1	16	0.61 ± 0.02 ***
Prox1	1	1.21 ± 0.07 *	Ifngr2	16	0.61 ± 0.01 ***
Apcdd1	18	1.21 ± 0.08 *	Pla2g5	4	0.60 ± 0.06 *
Olfr411	11	1.21 ± 0.06 *	1110004E09Rik	16	0.58 ± 0.01 ***
Lrrc20	10	1.21 ± 0.04 ***	Gart	16	0.58 ± 0.01 ***
Vwf	6	1.21 ± 0.06 **	Gcfc1	16	0.57 ± 0.03 ***
Ccnd1	7	1.21 ± 0.05 **	Ifnar1	16	0.57 ± 0.02 ***
Wnt2	66	1.21 ± 0.03 **	Usp16	16	0.57 ± 0.01 ***
Glis3	19	1.21 ± 0.07 *	Pnck	X	0.57 ± 0.02 ***
Rfx3	19	1.21 ± 0.04 **	Tmem50b	16	0.56 ± 0.01 ***
Krt9	11	1.21 ± 0.07 *	N6amt1	16	0.56 ± 0.03 ***
1810029B16Rik	8	1.21 ± 0.06 *	Grik1	16	0.56 ± 0.03 ***
Aym1	5	1.21 ± 0.06 *	Dnajc28	16	0.55 ± 0.02 ***
BC049635	4	1.21 ± 0.05 **	Cryzl1	16	0.55 ± 0.02 ***
Fam50a	X	1.20 ± 0.03 ***	Cyyr1	16	0.54 ± 0.02 ***
Zfp212	6	1.20 ± 0.04 **	Adamts1	16	0.54 ± 0.03 ***
Wfs1	5	1.20 ± 0.03 **	Cct8	16	0.53 ± 0.03 ***
Igsf3	3	1.20 ± 0.05 **	Mrps6	16	0.52 ± 0.02 ***
Dhdh	7	1.20 ± 0.04 **	H2-D1	17	0.51 ± 0.08 **
Anp32a	9	1.20 ± 0.05 **	Atp5o	16	0.51 ± 0.02 ***
Micalcl	7	1.20 ± 0.03 ***	Ifnar2	16	0.50 ± 0.02 ***
Tmem158	9	1.20 ± 0.05 **	Slc5a3	16	0.49 ± 0.02 ***
Aox1	1	1.20 ± 0.04 **	Sostdc1	12	0.47 ± 0.07 *
Taf13	3	1.20 ± 0.04 **	Clic6	16	0.40 ± 0.05 **

Table 1. Genes deregulated in Ms5Yah hippocampus. Only the genes with a fold change above 1.2 and under 0.8 for which the ANOVA was below 0.05 are listed. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

GO category	%	Gene Number	Fold enrichment	P-value	FDR	List of genes
Biological process						
cell adhesion	7.45	14	3.82	$7.8x10^{-5}$	0.099	App, Cntnap5b, Lgals3bp, Mag, Mcam, Npnt, Pcdhb2, Pcdhb3, Pcdhb4, Pcdhb6, Pcdhb7, Pcdhb8, Pcdhb16, Vwf
Molecular function						
Proto cadherin	3.72	7	35.8	$2.9x10^{-8}$	$4.1x10^{-5}$	Pcdhb2, Pcdhb3, Pcdhb4, Pcdhb6, Pcdhb7, Pcdhb8, Pcdhb16
calcium ion binding	9.04	17	2.08	$7.1x10^{-3}$	8.73	Efhd1, Itsn1, Npnt, Pcdhb2, Pcdhb3, Pcdhb4, Pcdhb6, Pcdhb7, Pcdhb8, Pcdhb16, Pla2g5, Pnck, Pon3, Ppp2r3a, Slc24a4, Smoc2 , Sulf1

Table 2. Functional annotation clustering of genes deregulated in Ms5Yah hippocampus using DAVID software.

TRANSLATIONAL IMPACT

Clinical issue

Partial monosomy 21 (PM21) syndrome is a rare chromosomal abnormality characterized by the loss of a variable segment on long arm of chromosome 21 (Hsa21q). Depending on size and location of the deletions, PM21 is associated with a large heterogeneity of clinical phenotypes. Common symptoms include intellectual disability, craniofacial malformations, muscular, skeletal and cardiac abnormalities. There is currently no treatment to alleviate affected individual condition. The development of therapeutic strategies will require a broader understanding of the disease. In order to identify candidate genes, several mouse models deleted for different Hsa21q homologous regions have been established. However, all these models show relatively mild phenotypes. In this context, the value of a new animal model that recapitulates most severe PM21 symptoms cannot be understated.

Results

The authors characterized Ms5Yah mouse model deleted for the *App-Runx1* region. In previous work, they reported the importance of this interval in cardiac defect phenotypes of a Down syndrome mouse model. Here, they characterized the Ms5Yah model and reported developmental delay affecting viability, size and weight. Viability tests and histological analyses indicated that the majority of mutant neonates showed breathing incapacity. Hematology analysis revealed platelet deficit which was reported in some PM21 cases. Behavioral studies revealed severe impairments in motor coordination and spatial learning and memory deficits. Hippocampus transcriptomic analysis pinpointed a disruption of cell adhesion pathways. The amyloid precursor protein gene (*App*) was downregulated, whereas the cystathionine beta synthase gene (*Cbs*) not located in the *App-Runx1* region was found upregulated.

Implications and future directions

Anatomical and behavioral characterization of Ms5Yah mice suggests that the *App-Runx1* region has a major impact on PM21 most severe phenotypes. The present mouse model represents a novel genetic tool to identify pathways implicated in the pathophysiology of PM21. Behavioral characterization of Ms5Yah mice indicates locomotor learning deficits in addition to spatial learning and memory deficits which can be directly related to intellectual disabilities of patients. Future studies dedicated to the search for therapeutic agents that rescue lethality and/or learning impairments of Ms5Yah mice could provide the first treatment strategy for partial monosomy 21 syndrome. Thus, the Ms5Yah mouse model has implications for both researchers and clinicians.

Figure 1. Viability tests and anatomical characterization of E18.5 fetuses

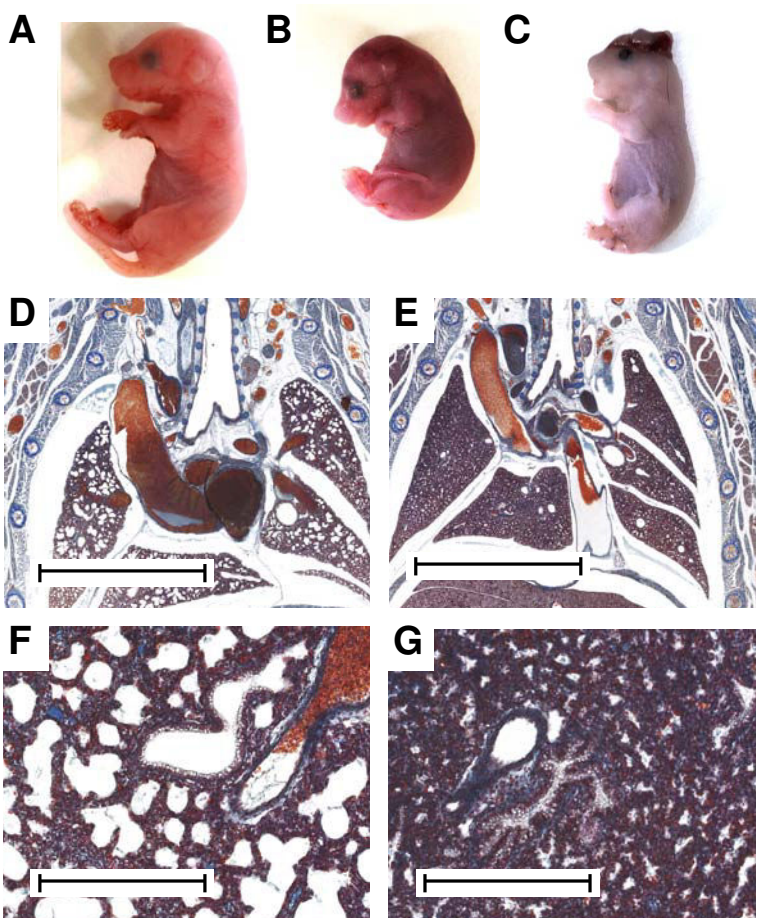


Figure 2. Coordination capacities of newborn animals and evolution of the weight

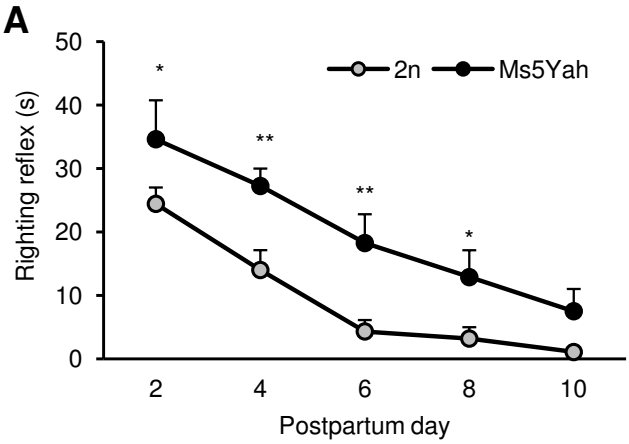


Figure 3. Locomotor coordination and grip strength capacities

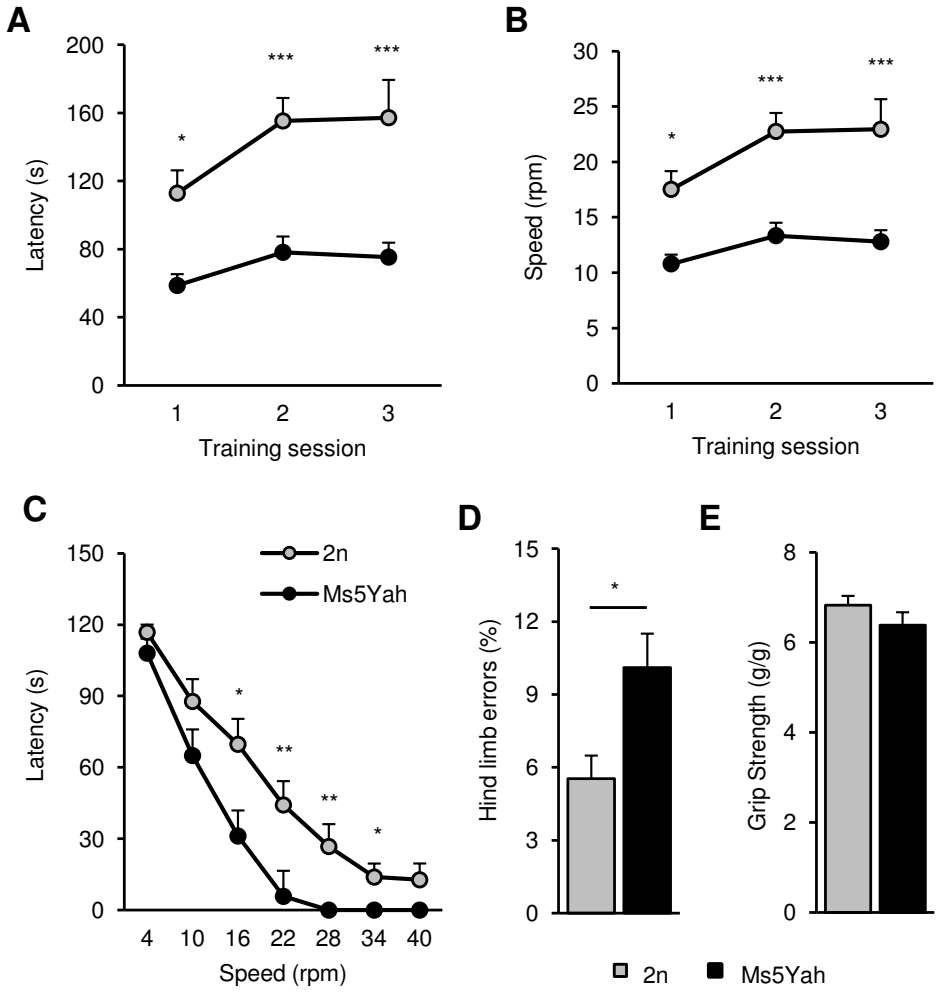


Figure 4. Spatial learning and memory capacities

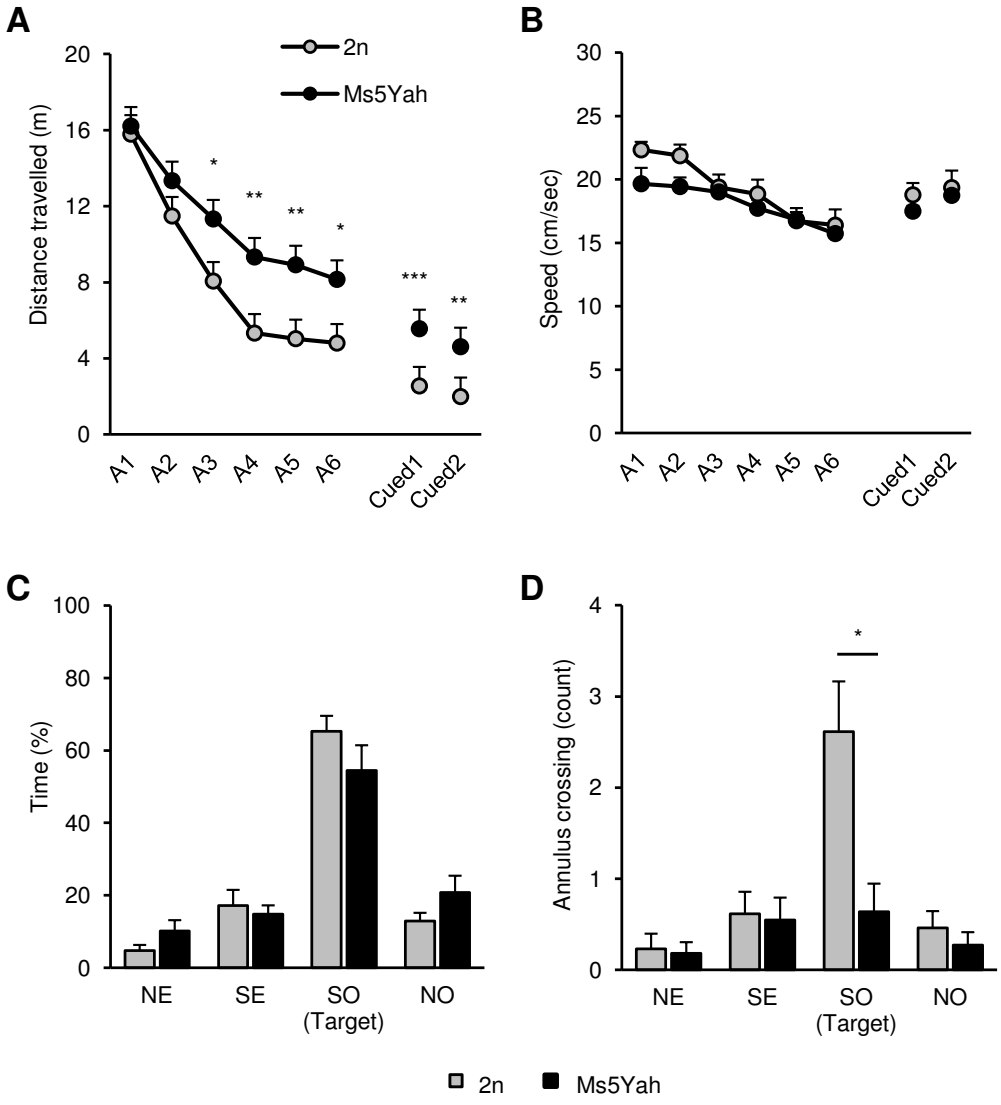


Figure 5. Miccroarray expression analyses on hippocampi

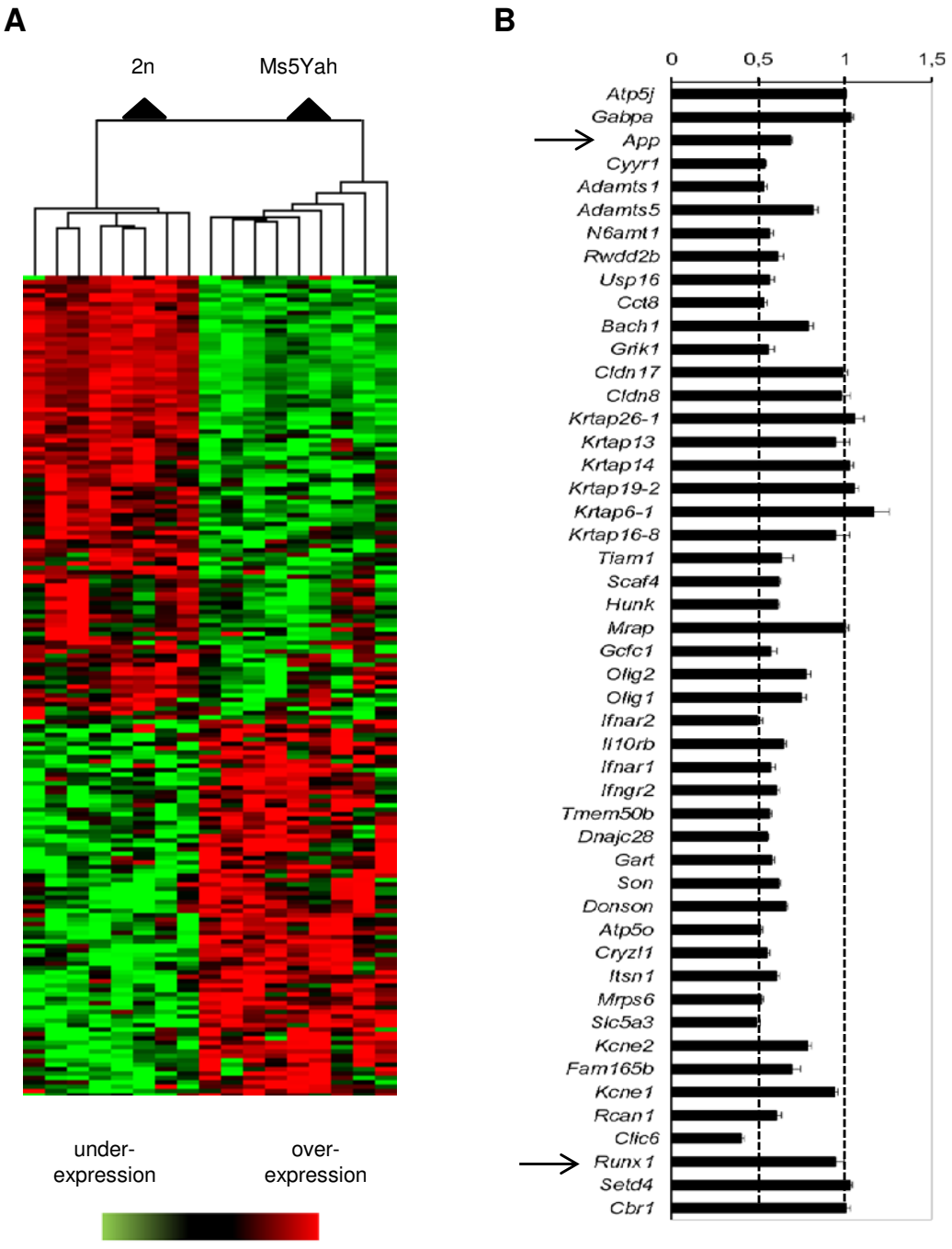


Figure S1. Exploratory activity, short-term working memory and recognition memory performances

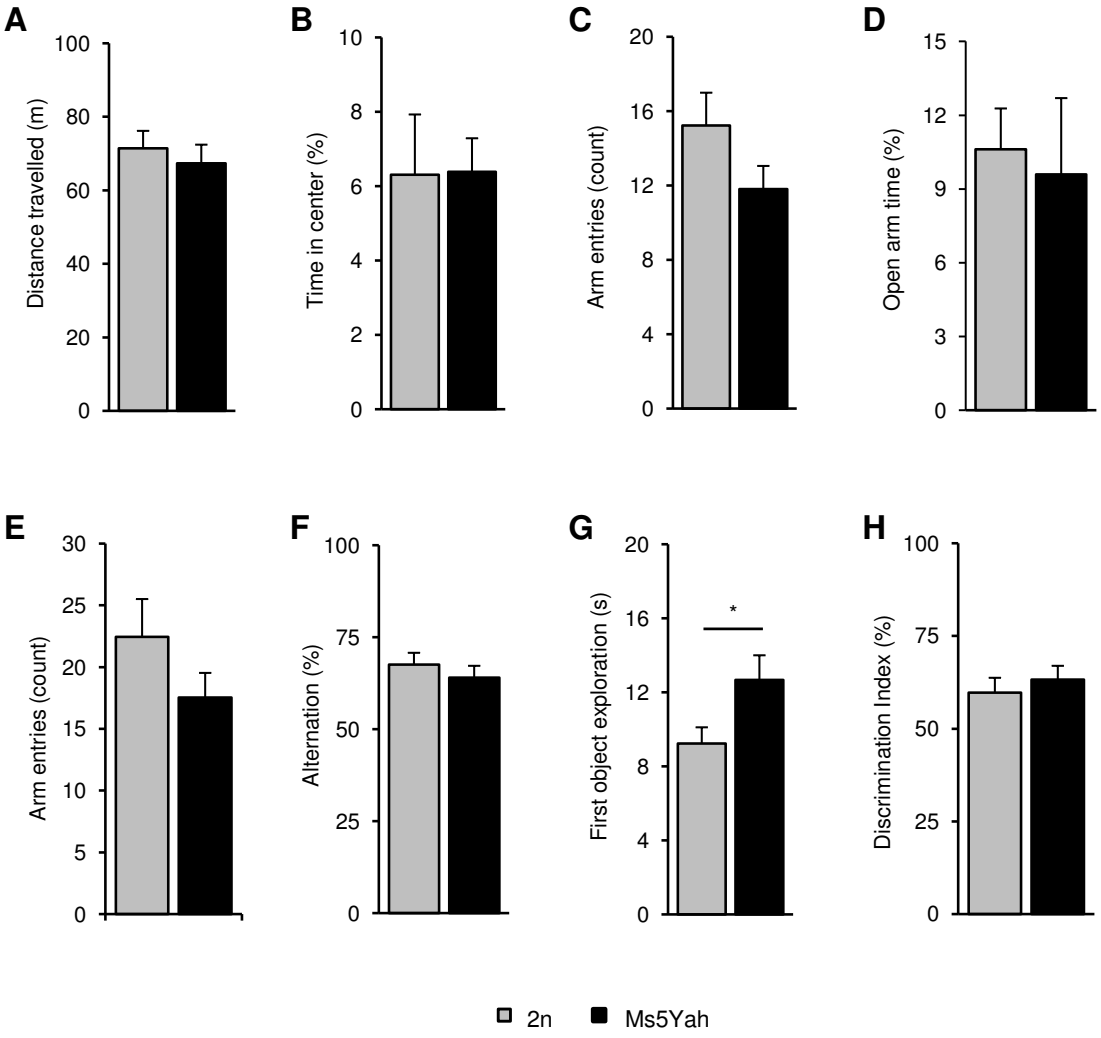
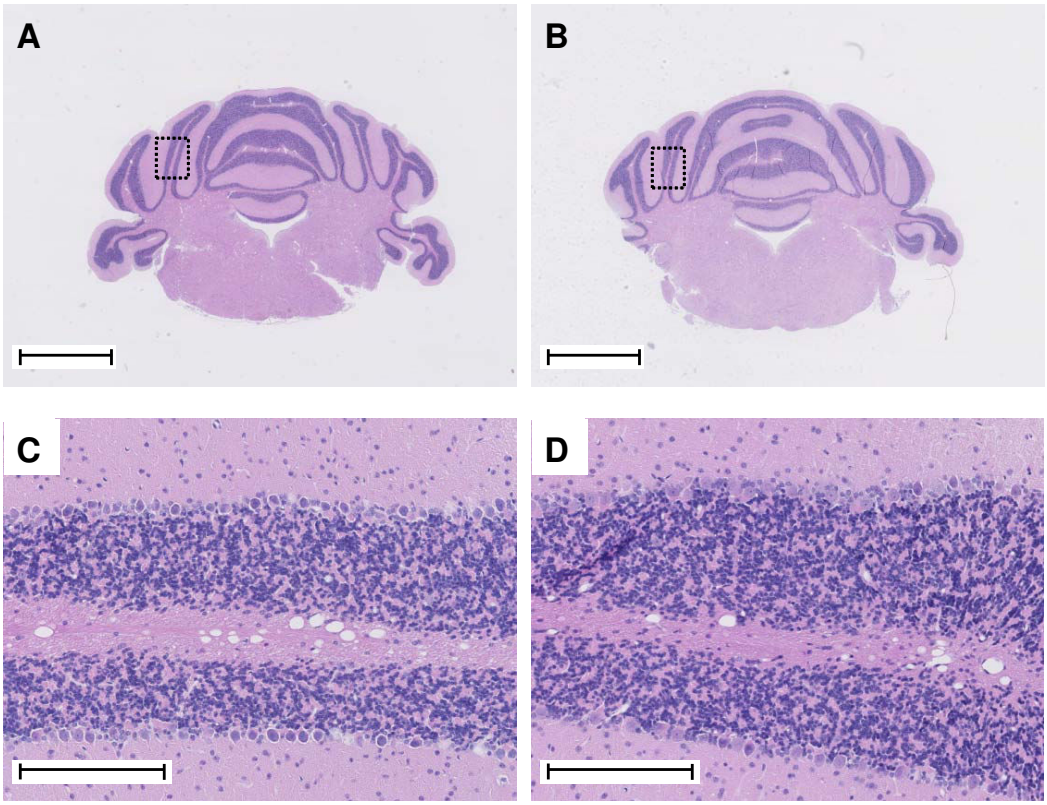


Figure S2. Transversal sections of the cerebellum stained with hematoxylin and eosin



Discussion et conclusion

Les animaux porteurs de la délétion pour la région *App-Runx1* n'étant pas viables sur un fond génétique consanguin B6N, nous avons maintenu la lignée Ms5Yah sur un fond hybride B6NC3B. Lors de notre amplification, aucune différence de transmission n'a été observée pour le sexe du parent porteur de la mutation. Une importante mortalité a été observée pour les animaux mutants. Au sevrage, seuls 22,4% des individus étaient porteurs de la délétion *App-Runx1*. Au stade embryonnaire E18.5, quelques heures avant la naissance, nous avons observé 48% des fœtus Ms5Yah indiquant l'absence de mortalité prénatale. Des tests de viabilité ont indiqué une importante diminution du poids des fœtus mutants ainsi qu'une incapacité respiratoire observée chez 52% d'entre eux. Hormis l'absence d'ouverture des alvéoles pulmonaires, aucune anomalie probante n'a été observée chez les fœtus Ms5Yah présentant une incapacité respiratoire. Les animaux survivants présentaient une importante diminution de la masse corporelle et de la taille, une démarche irrégulière, ainsi qu'un déficit de la coordination motrice observable chez les nouveau-nés en test de retournement. La caractérisation comportementale des animaux Ms5Yah adultes a révélé un important déficit de coordination locomotrice lors des tests de rotarod et de barre crantée. Cependant, la force d'agrippement et l'activité locomotrice en champ ouvert sont conservées. Le test de la piscine de Morris a révélé un important déficit de l'apprentissage et de la mémoire spatiale chez les animaux mutants. Aucune altération de la mémoire de travail n'a été observée dans le test du labyrinthe en Y. De même, la mémoire de reconnaissance des animaux Ms5Yah est conservée dans le test de reconnaissance du nouvel objet avec un délai d'une heure.

L'hippocampe ayant un rôle central dans l'établissement de la mémoire spatiale (Morris et al., 1982; Riedel et al., 1999), nous avons choisi cette structure pour notre étude transcriptomique sur puces Affymetrix Mouse Exon 1.0 ST. L'analyse a révélé que la majeure partie des gènes de la région *App-Runx1* est sensible à la dose. L'haplo-insuffisance des gènes *App* (codant pour la protéine précurseur de l'amyloïde), *Olig1 et Olig2* (codant pour les facteurs de transcription 1 et 2 régulant les oligodendrocytes) est potentiellement impliquée dans les déficits cognitifs et locomoteurs des animaux Ms5Yah. En dehors de la région *App-Runx1*, nous avons noté la surexpression du gène *Cbs* codant pour la cysthationine β-synthase qui est un gène candidat pour les phénotypes cognitifs du syndrome de Down.

En comparaison aux deux autres modèles de PM21 pour la région de synténie MMU16 qui montrent des phénotypes relativement modérés, le modèle Ms5Yah présente des phénotypes sévères. La caractérisation du modèle Ms5Yah corrobore les études de génétique humaine sur l'importance de la région 21q21.3q22.11 dans l'apparition des phénotypes les plus délétères associés à la PM21 (Lindstrand et al., 2010; Lyle et al., 2009). Effectivement, les animaux Ms5Yah présentent de nombreux phénotypes similaires à la symptomatique de la PM21. Le retard de croissance intra-utérin, la mortalité post-natale, la diminution de la taille et la déficience intellectuelle observés chez les patients sont à rapprocher du retard développemental affectant la viabilité post-natale, de la diminution de la taille et du poids, des déficits d'apprentissage locomoteur et d'apprentissage et de mémoire spatiale observés chez les souris Ms5Yah. Les déficits d'apprentissage des animaux mutants observés lors des tests de rotarod et de piscine de Morris sont particulièrement intéressants car ils peuvent être directement associés à la déficience intellectuelle des patients porteurs de délétion pour la région 21q21.3q22.11. Le modèle Ms5Yah représente un outil génétique très intéressant pour l'étude des mécanismes moléculaires associés à ces déficits cognitifs.

Partie 2

Modélisation des syndromes de délétion et de duplication 16p11.2

Introduction

Dans cette étude, nous nous intéressons au locus 16p11.2 qui constitue l'une des régions chromosomiques la plus associée aux syndromes neurodéveloppementaux (Grayton et al., 2012). Alors que la délétion et la duplication de la région sont associées à la déficience intellectuelle et à l'autisme, des phénotypes opposés sont observés pour le volume crânien et l'indice de masse corporelle : macrocéphalie et obésité pour les porteurs de la délétion, microcéphalie et insuffisance pondérale pour les porteurs de la duplication (Jacquemont et al., 2011; McCarthy et al., 2009; Shinawi et al., 2010; Walters et al., 2010; Weiss et al., 2008; Zufferey et al., 2012). L'étude des syndromes de délétion et de duplication de la région 16p11.2 BP4-BP5 présente un intérêt majeur pour la compréhension des troubles neuropsychiatriques, des anomalies de croissance de la boîte crânienne et de la variabilité du BMI. Néanmoins, la grande densité génique de la région 16p11.2 BP4-BP5 (34 gènes pour 600 kb) et la pénétrance incomplète des phénotypes rendent l'étude de ces deux syndromes extrêmement complexe.

Chez la souris, la région de synténie 16p11.2 BP4-BP5 a été retrouvée au niveau du chromosome 7F3. La caractérisation des premiers modèles des réarrangements 16p11.2 BP4-BP5 porteurs de la délétion et/ou de la duplication de la région *Slx1b-Sept1* a été publiée en 2011 (Horev et al., 2011). La caractérisation comportementale des animaux maintenus sur un fond hybride B6N129Sv a révélé une hyperactivité locomotrice en tests d'activité circadienne et de champ ouvert, et un comportement stéréotypique d'escalade de la grille en cage d'hébergement pour les animaux *Del/+*, ainsi qu'une hypoactivité locomotrice en test d'activité circadienne pour les animaux *Dup/+*. Les auteurs ont également observé des augmentations volumiques relatives de certaines régions cérébrales pour les animaux *Del/+* et des tendances opposées pour les animaux *Dup/+*. Aucun phénotype n'a été observé pour les animaux *Del/Dup* porteurs de la délétion et de la duplication. L'interprétation de ces résultats est néanmoins problématique. Effectivement, l'intervalle *Slx1b-Sept1* inclut 4 gènes (*Cd2bp2*, *Tbc1d10b*, *Mylpf* et *Sept1*) qui ne sont pas associés à la région de synténie 16p11.2 BP4-BP5 ce qui induit donc un biais dans la modélisation de la pathologie humaine.

Récemment, l'analyse d'un deuxième modèle murin pour la délétion de la région 16p11.2 porteur d'une délétion pour la région *Coro1a-Spn* sur un fond hybride B6N129Mo a montré des altérations neuronales et volumétriques des ganglions de la base ainsi que de nombreux phénotypes comportementaux (Portmann et al., 2014) incluant une hyperactivité

locomotrice, un déficit de mémoire de reconnaissance, une incapacité de nage en piscine de Morris, ainsi qu'une absence de réponse à des stimuli sonores. Cependant, les auteurs ont indiqué que les animaux mutants présentaient une surdité sévère d'origine indéterminée mais a priori non liée à la délétion *Coro1a-Spn*.

Dans ce contexte, nous avons généré et caractérisé de nouveaux modèles murins porteurs de la délétion et/ou de la duplication de la région *Sult1a1-Spn*. Nos modèles sont les premiers à avoir été générés sur un fond génétique consanguin B6N ce qui nous a permis d'étudier les réarrangements de la région 16p11.2 dans un contexte génétique spécifique. Les animaux *Del/+* et *Dup/+* pour la région *Sult1a1-Spn* ont dans un premier temps été amplifiés séparément. Une étude cranio-faciale, un suivi métabolique sous alimentation enrichie en graisse, ainsi qu'une analyse comportementale ont été réalisés. D'importantes altérations de la masse corporelle, de l'activité et de la mémoire de reconnaissance des animaux ont été observées. Au sacrifice, nous avons prélevé foie, cervelet, striatum et hippocampe à partir desquels nous avons effectué des analyses transcriptomiques. Dans le but de confirmer les phénotypes comportementaux observés, nous avons généré une cohorte de mâles *Del-Dup* contenant les 4 génotypes *Del/+*, wt, *Del/Dup*, *Dup/+*. Afin d'étudier des changements de transmission synaptique pouvant expliquer les altérations de mémoire de reconnaissance des animaux, une deuxième cohorte *Del-Dup* a été générée pour réaliser une étude électrophysiologique sur coupes aiguës d'hippocampe. Enfin, dans le but d'étudier l'influence du fond génétique, nous avons répété la caractérisation comportementale et l'analyse métabolique avec des animaux *Del/+* et *Dup/+* de fond génétique hybride B6NC3B obtenus par croisement des animaux mutants de fond B6N avec des animaux sauvages de fond C3H/HeH (C3B). Encore une fois, nous avons choisi ce fond génétique pour ses bonnes capacités cognitives, mais également pour son activité spontanée plus faible en comparaison à la souche B6N (Mandillo et al., 2008) ce qui nous a permis d'atténuer les altérations d'activité des animaux mutants et d'étudier l'impact de cette atténuation sur les autres phénotypes observés.

Opposite effects and genetic dosage in mouse models of 16p11.2 deletion and duplication syndromes

Thomas ARBOGAST[1], Abdel-Mouttalib OUAGAZZAL[1], Claire CHEVALIER[1], Maksym KOPANITSA[3], Nurudeen O. AFINOWI[3], Eugenia MIGLIAVACCA[4], Marie-Christine BIRLING[2], Marie-France CHAMPY[2], Alexandre REYMOND[4], Yann HERAULT[1,2]

Affiliations

[1]IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), Dpt of Translational Medicine and Neurogenetics; CNRS, UMR7104; INSERM, U964; Strasbourg University; 1 rue Laurent Fries, F-67404 Illkirch-Graffenstaden, France

[2]Institut Clinique de la Souris, ICS; PHENOMIN; GIE CERBM; 1 rue Laurent Fries, F-67404 Illkirch-Graffenstaden, France

[3]Synome Ltd, Moneta Building, Babraham Research Campus, Cambridge CB22 3AT, UK

[4]Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland

Corresponding author: Yann Hérault

Keywords: Copy number variation, neurodevelopmental disorder, autism spectrum disorders, body mass index, mouse model, stereotypy, recognition memory, social behaviors

Running title

Rearrangements of the *Sult1a1-Spn* region

ABSTRACT

During the last decade, it has become increasingly clear that many neurodevelopmental disorders are associated with gene copy number variations (CNVs). Here we focused on the 16p11.2 deletion and duplication syndromes. Whereas both rearrangements have been associated with developmental delay and autism spectrum disorders, a reciprocal effect of 16p11.2 gene dosage on body mass index and head size has been noted, as deletion is associated with obesity and macrocephaly and duplication is associated with being underweight and microcephaly. Mice carrying deletion (*Del/+*) and duplication (*Dup/+*) for the 16p11.2 homologous region displayed several opposite phenotypes on pure genetic background. *Del/+* animals showed weight and adipogenesis reductions, hyperactivity and recognition memory impairments whereas *Dup/+* animals showed weight and adipogenesis increases, hypoactivity and recognition memory improvements. *Del/+* and *Dup/+* mice on mixed genetic background also displayed social interaction deficits. Interestingly, if memory capacities and social behavior alterations observed in mice were similar than those observed in patients, weight and adipogenesis phenotypes were mirroring the human effects. Animal modeling of 16p11.2 CNV syndromes confirmed the implication of dosage-sensitive gene(s) and the potential implication of other(s) polymorphism(s) which differ between mice and humans.

INTRODUCTION

Neurodevelopmental diseases can be caused by inherited or *de novo* mutations including copy number variants (CNVs), defined as deletions and duplications of DNA segments that are 50bp or larger. CNVs are widely distributed throughout the human genome and substantially impact several human traits and diseases (Conrad et al., 2010; Cook and Scherer, 2008; Redon et al., 2006). Main mechanism of CNV generation is a non-allelic homologous recombination (NAHR) between segmental duplications (SDs) (Hastings et al., 2009). The pericentromeric locus 16p11.2 contains several SDs predisposing to genomic structural variants. Most common rearrangements occur between proximal LCRs at breakpoints 4 (BP4) and 5 (BP5) delineating a ~600-kb region which contains 34 genes.

The 16p11.2 BP4-BP5 rearrangements have a population prevalence of approximately 1/1500 (Jacquemont et al., 2011) and reach 1% in intellectual disability (ID) (Cooper et al., 2011) and autism spectrum disorders (ASD) cohorts (Fernandez et al., 2010; Marshall et al., 2008; Sanders et al., 2011; Weiss et al., 2008). In addition to ID and ASD, both rearrangements have also been associated with epilepsy (Ghebranious et al., 2007; Reinthaler et al., 2014; Shinawi et al., 2010; Zufferey et al., 2012) whereas only the duplication has been linked to schizophrenia, bipolar disorder and depression (Bergen et al., 2012; McCarthy et al., 2009; Steinberg et al., 2014). Body mass index (BMI) phenotypes and abnormal head size have also been reported in 16p11.2 CNVs carriers. Whereas the deletion has been linked with obesity and macrocephaly (Shinawi et al., 2010; Walters et al., 2010; Zufferey et al., 2012), the duplication has been linked with being underweight and microcephaly (Jacquemont et al., 2011; Shinawi et al., 2010). Reciprocal impacts on BMI and head size indicate that the phenotypes could have mirror etiologies depending on changes in gene transcript levels.

The investigation of the interplay between genes, brain activity and behavior is crucial for the understanding of neurocognitive processes. In the case of 16p11.2 CNV syndromes, candidate gene discovery is particularly challenging as the region shows a high density of genes, the majority of which are expressed in the brain and potentially important for neurodevelopment. Interestingly, a 118-kb deletion encompassing *MVP, CDIPT, SEZ6L2, ASPHD1 and KCTD13* segregating with ASD has been recently reported in a three-generation family (Crepel et al., 2011), pinpointing a potential key role of these genes in neuropsychiatric traits associated with 16p11.2 BP4-BP5 rearrangements.

Complementary to human genetic analysis, animal models have been developed in order to study the correlation between phenotype and genotype. Overexpression and suppression studies in zebrafish embryos and mouse cortical neurons have highlighted two genes, *KCTD13* and *TAOK2*, controlling head size and dendrite morphogenesis, respectively (de Anda et al., 2012; Golzio et al., 2012). Genes of the 16p11.2 region have been found to be highly conserved on mouse chromosome 7F3. In 2011, the first mouse models carrying deletion and duplication for the *Slx1b-Sept1* region were reported to display locomotor activity alterations and ventral midbrain volume changes (Horev et al., 2011). Nevertheless, these models encompass four genes (*Cd2bp2, Tbc1d10b, Mylpf and Sept1*) which are not related to the 16p11.2 BP4-BP5 homologous region. A recent study demonstrated circuit defects in the basal ganglia in a novel 16p11.2 deletion mouse model with deleted *Coro1a-Spn* region (Portmann et al., 2014). Several behavioral abnormalities were reported including activity alterations, circling behavior, memory impairments, motor coordination deficits, swimming inabilities, and an absence of startle response. However, severe hearing deficit was also revealed in this mouse line, which could account for some of the observed behavioral phenotypes.

In this context, we engineered and investigated novel 16p11.2 CNV mouse models for the *Sult1a1-Spn* region. We used a comprehensive test battery to assess the impact of the genetic rearrangements on behavior, synaptic function and metabolism of mice which can be related to neuropsychiatric disorders and BMI alterations of human patients. We found that *Sult1a1-Spn* CNVs produce several phenotypes in mice bred on pure (B6N) and on mixed (B6NC3B) genetic background. Compared to wild-types, mice carrying the deletion (*Del/+*) showed weight and adipogenesis deficits, locomotor hyperactivity, exaggerated repetitive behaviors, and recognition memory deficits whereas mice carrying the duplication (*Dup/+*) showed weight and adipogenesis increases, locomotor hypoactivity, reduced repetitive behaviors, and recognition memory improvements. All these behavioral phenotypes were normalized in the compound *Del/Dup* mutant mice, except the memory improvement. Deficits in social interaction were also observed in *Del/+* and *Dup/+* animals on a mixed genetic background. Transcriptomic analysis performed on hippocampus, striatum, cerebellum and liver revealed that the majority of genes located on the *Sult1a1-Spn* were dosage-sensitive genes and potentially implicated in the opposite phenotypes described above. Overall, our observations extend previous studies and demonstrate the recurrence of opposite metabolism and behavioral phenotypes in mouse models for the 16p11.2 BP4-BP5 rearrangements.

RESULTS

Generation and transmission of mice carrying rearrangements of the *Sult1a1–Spn* genetic interval

To generate mouse models of the 16p11.2 BP4-BP5 rearrangements, we introduced LoxP sites flanking syntenic region located on chromosome 7F3 (Fig. 1A). Mutant mice were obtained by in vivo TAMERE strategy (Brault et al., 2006; Herault et al., 1998). LoxP sites were first introduced by homologous recombination in C57BL/6N (B6N) embryonic stem cells at *Sult1a1* and *Spn* loci in the same orientation using specific vector and the corresponding mouse line were generated. Selection cassettes were excised then mice were crossed with Hprt<tm1(cre)Mnn> mouse line (Tang et al., 2002), expressing the Cre recombinase under the control of the X-linked hypoxanthine guanine phosphoribosyl transferase gene promoter active in oocytes. Females born from this mating and bearing both the Hprt<tm1(cre)Mnn> transgene and loxP sites in a trans configuration were mated with wild-type B6N males (Fig. 1B). We recovered mice carrying the deletion (*Del/+*) and the duplication (*Dup/+*) for the *Sult1a1-Spn* region with a recombination frequency of respectively 3 and 4 animals out of 97 newborns.

Del/+ and *Dup/+* mice were first amplified separately on pure B6N genetic background. The segregation of the *Del/+* allele (30.7%) was considerably reduced compared to the *Dup/+* allele (45.8%; Table1). In order to confirm results and to characterize mice carrying the deletion and the duplication, we crossed *Del/+* with *Dup/+* animals. Results confirmed the low transmission of the *Del/+* allele which was compensated in the *Del/Dup* carriers demonstrating that lethality is associated with the deletion in the B6N background. Finally, to evaluate the influence of the genetic background, we crossed B6N *Del/+* and *Dup/+* mice with sighted C3H/HeH (C3B) wild-type mice (Hoelter et al., 2008) and found normal segregation for *Del* (49.1%) and *Dup* (49.2%) alleles in the B6NC3B background (Table 1).

To determine whether B6N *Del/+* died in utero or postnatally, the fetuses of pregnant wild-type females impregnated by *Del/+* males were collected at embryonic day 18.5 (E18.5), few hours before natural delivery. Over 69 fetuses extracted by caesarean section and analyzed, we genotyped 37 wt and 32 *Del/+* fetuses which is the normal genotypic Mendelian ratio. Overall, *Del/+* fetuses needed more time to oxygenize and 3 mutant animals died few minutes after caesarean. Fetuses were weighted and we found that *Del/+* were underweight in

comparison with wt littermates (*Del/+*: 1.08 ± 0.02 g, wt: 1.18 ± 0.01 g; $F_{(1,67)} = 27.318$, $p < 0.001$; Supplemental Fig. S7A). We did the same experiment with *Dup/+* animals and found no weight difference between the 24 wt and 27 *Dup/+* fetuses extracted (*Dup/+*: 1.22 ± 0.02 g, wt: 1.22 ± 0.01 g; $F_{(1,49)} = 0.047$, $p = 0.829$). Thus, the *Del/+* allele on B6N genetic background induce a developmental delay and a specific weakness of carriers which leads to the death of ~40% of *Del/+* neonates between birth and weaning.

Rearrangements of the *Sult1a1–Spn* region induce opposite phenotypes on locomotor activity, repetitive behaviors and memory performances

We first studied the behavioral effect of *Sult1a1-Spn* deletion and duplication in mice bred on pure B6N genetic background. For each rearrangement, separate cohorts of mutants and their wild-type littermates were generated and subjected to the same battery of behavioral tests (Figure S1, Tables S1-2). These initial characterization studies revealed an opposite effect of the deletion and the duplication of *Sult1a1-Spn* region on several behavioral traits including locomotor activity (Fig. S1A and B), repetitive behaviors (Fig. S1C), and learning and memory performances (Fig. S1E and F).

To confirm these findings and ensure that the observed phenotypes were caused by gene-dosage alterations, we analyzed the behavior of a compound *Del-Dup* cohort with littermates of the 4 genotypes: *Del/+*, wt (+/+), *Del/Dup*, and *Dup/+* (Tables S3-5). Consistent with the previous results, *Del/+* and *Dup/+* mice had a normal pattern of circadian locomotor activity (Fig. S2A and B), but their baseline levels of spontaneous locomotor activity and of rears were distinct from those of wild-types (Fig. 2A). *Del/+* mice showed increased locomotor activity and rearing behavior during the dark phase, but only the later measure reached statistical significance compared to wild-types ($H_{(3, 51)} = 15.803$, $p = 0.001$; *Del/+* vs wt: $p = 0.039$). Conversely, in comparison with wild-types, *Dup/+* mice manifested reduced locomotor activity during light phase ($F_{(3,51)} = 5.724$, $p = 0.002$; *Dup/+* vs wt: $p = 0.003$) and reduced rears activity during dark ($H_{(3, 51)} = 15.803$, $p = 0.001$, *Dup/+* vs wt: $p = 0.015$) and light ($H_{(3, 51)} = 11.119$, $p = 0.011$; *Dup/+* vs wt: $p = 0.008$) phases. *Del/Dup* mice carrying both rearrangements behaved similarly than wild-types ($p > 0.05$ for all measures), confirming that the abnormal phenotypes of *Del/+* and *Dup/+* mice were linked to an alteration in gene dosage. When tested in novel open field arena, *Del/+* and *Dup/+* mice displayed again opposite locomotor activity phenotypes, but only *Dup/+* mice had a

significantly different scores from the wild-types ($F_{(3,58)} = 8.920$, $p < 0.001$; *Del/+* vs wt: $p = 0.089$, *Dup/+* vs wt: $p = 0.013$; Fig. 2B and S3). The time spent in the center of the arena, which is used as a measure of emotionality, was also increased in *Del/+* mice and decreased in *Dup/+* mice ($H_{(3, 58)} = 23.758$, $p < 0.001$; *Del/+* vs wt: $p = 0.001$, *Dup/+* vs wt: $p = 0.029$). No difference in exploratory activity was observed between *Del/Dup* and wild-type mice ($p > 0.05$ for all measures). Visual observations of the animals in a home cage revealed a range of abnormal repetitive behaviors, especially in *Del/+* mice (Fig. 2C). Marked increase in rearing ($F_{(3,41)} = 6.156$, $p = 0.001$; *Del/+* vs wt: $p = 0.010$) and jumping ($H_{(3, 41)} = 14.394$, $p = 0.002$; *Del/+* vs wt: $p = 0.016$) behaviors were noticed in these mice compared to wild-types. Climbing behavior tended to be increased in *Del/+* mice and reduced in *Dup/+* mice but the effect just failed to reach statistical significance ($H_{(3, 41)} = 9.168$, $p = 0.027$; *Del/+* vs wt: $p = 0.059$, *Dup/+* vs wt: $p = 0.074$). No behavioral abnormality was noted in *Del/Dup* mice, indicating that the motor stereotypies of the mutant mice were linked to a gene dosage alteration ($p > 0.05$ for all measures). In the social interaction test, no difference of sniffing and following time was noticed between genotypes. Both *Del/+* and *Dup/+* mutant mice behave similarly than wild-types ($p > 0.05$ for all measures; Fig. 2D).

Learning and memory functions were evaluated using the novel object recognition task, the most common task for assessing the various facets of recognition memory in rodents. We first investigated whether *Del/+* and *Dup/+* mice could discriminate a novel from a previously explored object after a short retention delay of 30 min (Fig. 2E). During the acquisition session, all genotypes spent an equal amount of time exploring the sample object (Table S4). In the subsequent choice session, *Del/+* mice displayed a significant memory impairment compared to wild-types, while *Dup/+* mice tended to display a memory improvement ($F_{(3,49)} = 8.080$, $p < 0.001$; *Del/+* vs wt: $p = 0.004$). To confirm this cognitive phenotype, we extended the retention delay to 3 hours (Fig. 2E). *Del/+* mice displayed again a poor recognition performance while *Dup/+* mice showed a clear-cut memory improvement compared to wild-type mice ($H_{(3, 51)} = 16.014$, $p = 0.001$; *Del/+* vs wt: $p = 0.043$, *Dup/+* vs wt: $p = 0.021$, *Del/Dup* vs wt: $p = 0.048$). Interestingly, *Del/Dup* also displayed memory improvement performances like *Dup/+* mice.

Potential alterations in sensory and motor functions were assessed in series of assays. No statistical differences in motor coordination and motor learning were detected between mutants and wild-types mice in the rotarod test (Fig. 2F). In the notched bar test, *Del/+* mice

made a significantly higher number of errors compared to wild-types, while *Dup/+* and *Del/Dup* mice performed normally ($H_{(3, 60)} = 14.044$, $p = 0.003$; *Del/+* vs wt: $p = 0.007$; Fig. 2F). In the grip test, *Del/+* and *Dup/+* mice showed respectively stronger and weaker grip strength compared to wild-types ($F_{(3,60)} = 23.598$, $p < 0.001$; *Del/+* vs wt: $p = 0.002$, *Dup/+* vs wt: $p < 0.001$; Table S4). No obvious alteration in auditory abilities was detected in *Del/+*, *Dup/+* and *Del/Dup* mice compared to wild-types (Fig. S4).

Influence of genetic background on behavioral phenotypes associated with *Sult1a1–Spn* interval rearrangements

The above studies show that *Sult1a1-Spn* CNVs induce a range of behavioral abnormalities in mice. To study the influence of genetic background on the expression of these phenotypes, we conducted the same behavioral analysis with mutant animals on a mixed B6NC3B genetic background (Figure S5, Tables S6-7). Separate cohorts of *Del/+*, *Dup/+* and corresponding wild-type littermates were used for the present study. As expected, *Del/+* and *Dup/+* mice on B6NC3B mixed genetic background showed a range of behavioral abnormalities compared to their wild-type counterparts. Opposite phenotypes were noted for rears in the circadian activity test (Fig. S5A), climbing repetitive behavior (Fig. S5C), and recognition memory capacities (Fig. S5E). With a retention delay of 3 hours, we confirmed the recognition memory deficits of *Del/+* mice with novel object and object-place recognition memory tasks (Table S7). Recognition memory improvements were also confirmed for *Dup/+* mice in novel object recognition task with a 3 hours retention delay. A clear-cut phenotype for social behaviors was also revealed in the mutant mice when animals were tested using the standard open field (Fig. S5D). Compared to wild-type mice, social interaction reduction was observed for both *Del/+* ($F_{(1,9)} = 6.799$, $p = 0.028$) and *Dup/+* ($F_{(1,10)} = 5.594$, $p = 0.040$) mutant mice. In the three-chambered social approach procedure, mutant and wild-type mice displayed similar exploration time of first stranger (Table S7). Nevertheless, in comparison to wild-type littermates, *Del/+* mice showed deficits for social preference of the second stranger ($F_{(1,14)} = 12.849$, $p = 0.003$). To validate these social phenotypes, we assessed olfactory capacities of animals. No alteration of olfactory functions was detected in *Del/+* and *Dup/+* mice following exposure to social and non-social odors (Fig. S6).

Effets of rearrangements of the *Sult1a1–Spn* region on synaptic transmission in hippocampal slices

To investigate potential electrophysiological underpinnings of observed behavioral phenotypes, we studied synaptic transmission and its plasticity in synapses between Schäffer collaterals and apical dendrites of CA1 pyramidal neurons in hippocampal slices of wt, *Del*/+, *Dup*/+ and *Del*/*Dup* mice on pure B6N genetic background. Analysis of input-output relationships (Fig. 3A) shows that there was no significant interaction between genotype and stimulus (Repeated Measures ANOVA $F_{(27, 468)} = 0.98$; $p = 0.498$). Although we observed nominally decreased slopes of field excitatory postsynaptic potentials (fEPSPs) in *Del*/+ and *Dup*/+ mutants, especially in response to higher stimulus strengths (Fig. 3A), a separate two-way nested ANOVA performed just on maximum values of fEPSPs also failed to reveal a significant effect ($F_{(3, 35)} = 1.229$, $p = 0.314$).

Likewise, genotype did not significantly influence values of paired-pulse facilitation, a model of short-term synaptic plasticity, recorded in hippocampal slices ($F_{(3, 35)} = 0.0298$; $p = 0.993$; Fig. 3B). To investigate long-term synaptic plasticity, we induced long-term potentiation (LTP) of fEPSPs in Schäffer collaterals-CA1 synapses by the theta-burst stimulation (Fig. 3C). Two way nested ANOVA demonstrated a significant effect of genotype on LTP values ($F_{(3, 35)} = 3.43$; $p = 0.027$). *Post hoc* Dunnett's Multiple Comparison test done on individual slice values demonstrated that LTP was significantly smaller ($q = 2.441$, $P < 0.05$) in slices from Dup/+ mice (140 ± 9%, $n = 10$, $N = 4$) compared to WT slices (177 ± 7%, $n = 18$; $N = 8$), while levels of LTP in *Del*/+ (160 ± 12%) and *Del*/*Dup* slices (182 ± 11%) did not differ significantly from WT levels.

Thus, we conclude that long-term synaptic plasticity was sensitive to the duplication of the of the *Sult1a1–Spn* region, while neither short- nor long-term plasticity were affected by deletion or deletion/duplication rearrangements. We also observed a trend to nominally lower fEPSP slopes in all three mutants compared to WT values. The latter observation may require further experiments with larger cohorts in order to determine if basal synaptic transmission is indeed affected by rearrangements of the Sult1a1-Spn region.

Rearrangements of the *Sult1a1–Spn* region induce opposite phenotypes on weight and adiposity

Body weights of adult animals were recorded once a week during behavioral characterization. Concerning B6N *Del/+* and *Dup/+* separate cohorts, *Del/+* mice were extremely underweight (Two-way ANOVA genotype effect $F_{(1,120)} = 88.115$, $p < 0.001$; Fig. S7B) whereas B6N *Dup/+* mice were overweight (Two-way ANOVA genotype effect $F_{(1,99)} = 8.391$, $p = 0.009$; Fig. S7B) compared to wt littermates. These results were confirmed with the *Del-Dup* cohort. In comparison with wt and *Del/Dup* mice, *Del/+* littermates were significantly underweight whereas *Dup/+* mice showed overweight trends (Two-way ANOVA genotype effect $F_{(3,780)} = 9.954$ $p < 0.001$; *Del/+* vs wt: $p = 0.002$, *Dup/+* vs wt: $p = 0.211$; Fig. 4A). Feeding behaviors evaluated during circadian activity test were unchanged in mutant animals (Fig. S2). In addition, no correlation was observed between activity and body mass for the different genotypes. We evaluated body fat percentage of animals by qNMR and found adiposity decrease for *Del/+* mice ($H_{(3, 60)} = 22.770$, $p < 0.001$; *Del/+* vs wt: $p < 0.001$; Fig. 4B) which was confirmed at autopsy by an important lack of epididymal fat pads (Fig. 4D). A body size reduction was also found for *Del/+* mice ($F_{(3,35)} = 6.834$, $p < 0.001$; *Del/+* vs wt: $p = 0.027$; Fig. 4C). Interestingly, we observed a clear correlation between body weight and body size of Del/+ animals (Pearson correlation coefficient, $\rho = 0.893$, $p = 0.001$). No weight and size phenotype was noted in *Del/Dup* mice ($p > 0.05$ for all measures).

Concerning *Del/+* and *Dup/+* separate cohorts maintained on a mixed B6NC3B genetic background, whereas *Del/+* animals were still underweight in comparison with wild-type littermates (Two-way ANOVA genotype effect $F_{(1,456)} = 16.620$, $p < 0.001$; Fig. S7D), no body weight phenotype was observed for *Dup/+* animals (Two-way ANOVA genotype effect $F_{(1,276)} = 2.023$, $p = 0.168$; Fig. S7D). Adiposities were similar between mutant and control mice (Fig. S7E). A body size reduction was found again for *Del/+* mice ($F_{(1,33)} = 17.137$, $p < 0.001$; Fig. S7F) but no correlation between body size and body weight was observed in the B6NC3B genetic background.

To further characterize weight and adiposity changes, we performed a metabolic analysis of B6N *Del/+* and *Dup/+* separate cohorts challenged with high-fat diet from 5 to 15 weeks of age (Fig. 4E-I, Table S8). Again, in comparison with wild-types, *Del/+* mice were extremely underweight (Two-way ANOVA genotype effect $F_{(1,102)} = 48.671$, $p < 0.001$; Fig.

4E) whereas *Dup/+* mice were overweight (Two-way ANOVA genotype effect $F_{(1,112)}$ = 8.674, $p = 0.010$). Body size was decreased for *Del/+* mice ($F_{(1,17)} = 51.771$, $p < 0.001$; Table S8) and increased for *Dup/+* mice ($F_{(1,16)} = 29.229$, $p < 0.001$). Body composition analysis revealed fat percentage decrease in *Del/+* mice ($F_{(1,16)} = 32.138$, $p < 0.001$; Fig. 4F) and increase in *Dup/+* mice ($F_{(1,16)} = 8.621$, $p = 0.010$). No correlation was found between weight and adiposity for the different genotypes. Energy expenditure (EE) of animal was analyzed by indirect calorimetry during dark and light phases (Fig. 4G). *Del/+* mice showed higher EE during dark phase ($F_{(1,16)} = 5.313$, $p = 0.035$) whereas *Dup/+* mice showed lower EE during dark ($F_{(1,15)} = 12.994$, $p = 0.003$) and light ($F_{(1,15)} = 13.997$, $p = 0.002$) phases. These results are in accordance with endogenous activity phenotypes of mutant mice observed in circadian activity test. Intraperitoneal glucose-tolerance test (IPGTT) indicated faster glucose clearance for *Del/+* mice ($F_{(1,15)} = 23.396$, $p < 0.001$; Fig. 4H) and hyperglycemia for *Dup/+* animals ($F_{(1,16)} = 8.220$, $p = 0.011$). In accordance with body fat composition data, endocrinology analysis (Fig. 4I) revealed decrease of leptin ($F_{(1,16)} = 7.059$, $p = 0.017$) and adiponectin ($F_{(1,17)} = 6.790$, $p = 0.018$) blood level for *Del/+* mice and increase of leptin ($F_{(1,15)} = 5.350$, $p = 0.035$) blood level for *Dup/+* mice. An increase for insulin blood level near from significance was also found for *Dup/+* mice. Blood chemistry analysis did not reveal gross hematology changes except a slight decrease of free fatty acids level for *Del/+* mice ($F_{(1,17)} = 5.333$, $p = 0.034$; Table S8).

Finally, we repeated the similar high-fat diet protocol on *Del/+* animals on a mixed B6NC3B genetic background (Fig. S8, Table S9). We found that B6NC3B *Del/+* mice showed similar phenotypes than B6N *Del/+* mice for weight, fat, energy expenditure, glucose clearance and endocrinology parameters. B6NC3B also present basal hyperglycemia, higher blood level of calcium and lower blood levels of total cholesterol and glucose. We did also a calorimetric bomb analysis on mouse feces and found similar level of food energy usage between wt and *Del/+* mice.

Rearrangements of the *Sult1a1–Spn* region induce craniofacial features

In order to study the influence of 16p11.2 CNVs on craniofacial structure, we performed an analysis of computed tomography (CT) cranial scans of animal heads with ulterior reconstruction of 3D skull images using 39 cranial landmarks (Fig. 5A). Separate *Del/+* and *Dup/+* cohorts of females were used for the analysis. A global skull effect was observed in

Del/+ females with a significantly reduced skull size (p = 0.009; Fig. 5B) and altered skull shape (p = 0.046; Fig. 5C). Regarding size reduction, most affected regions correspond to premaxilla, maxilla and zygomatic human bones indicating an important alteration of *Del/+* facial features. Concerning Dup/+ animals, we noticed no alteration of skull size (p = 0.297; Fig. 5D) but a significant alteration of skull shape (p = 0.037; Fig. 5E).

Transcriptomic analysis

We analyzed the transcriptome profiles of *Del/+*, *Dup/+* and wt animals in 3 brain regions, hippocampus, striatum, cerebellum, and in one peripheral tissue, liver. We investigated if the expression of genes harbored in the *Sult1a1-Spn* region was proportional to the copy number of the CNV or under dosage compensation (Fig. 6A-D). The expression of the genes in the *Sult1a1-Spn* region is largely affected by the CNVs in all three brain regions and in liver. Principal component analysis further suggested that the effect of the CNV on the expression of the genes in the *Sult1a1-Spn* interval is stronger in the Del/+ animals (Fig. S9).

In brain tissues the probesets associated to CNVs are: for cerebellum 29 (38) at 5% FDR (10%FDR), of which 22 (23) within the engineered region; for striatum 31 (35) at 5% FDR (10%FDR), of which 24 (26) within the engineered region; for hippocampus 37 (45) at 5% FDR (10%FDR), of which 26 (27) within the engineered region. Between 70 and 78% of the genes associated to CNVs are within the *Sult1a1–Spn* region when considering a cutoff at 5%FDR, more than 60% for 10%FDR. In liver 29 (47) probesets are significantly associated to the CNV with a FDR < 0.05 (FDR < 0.10); 15 (15) among the significant probesets are within the engineered region.

We performed gene set enrichment analysis on the lists of ranked genes resulted by the differential expression analysis to identify sets of genes that are positively associated with the CNVs (the expression of these genes increases with the CNVs) or negatively (the expression of these genes decreases with the CNVs). In case of cerebellum 4 (22) gene sets are negatively associated to the CNVs at FDR < 10% (FDR < 25%), no set is positively associated at FDR <25%. In case of striatum 38 (242) gene sets are positively associated to the CNVs at FDR < 10% (FDR < 25%); while 3 (6) gene sets. In case of hippocampus 0 (2) gene sets are positively associated to the CNVs at FDR < 10% (FDR < 25%); while 2 (6) gene sets negatively associated. In case of liver 31 (163) gene sets are positively associated to the CNVs at FDR < 10% (FDR < 25%); 1 (6) gene sets negatively associated.

DISCUSSION

The 16p11.2 BP4-BP5 region is one of the most frequent known single locus etiologies of neurodevelopmental disorders (Grayton et al., 2012). Deletion and duplication of the region are associated with several neuropsychiatric symptoms of incomplete penetrance including intellectual disability, autism spectrum disorders, schizophrenia and epilepsy. In addition, deletion is associated with macrocephaly and obesity and whereas duplication is associated with microcephaly and being underweight. These skull size and BMI mirror phenotypes could indicate the presence of dosage-sensitive genes in the 16p11.2 region.

In order to dissect the molecular mechanisms underlying the syndromes and to identify candidate genes for the different phenotypes, mouse models have been generated including two mouse models for the 16p11.2 deletion (*Del/+*) and one mouse model for the 16p11.2 duplication (*Dup/+*). Nevertheless, the first models include 4 genes outside of the BP4-BP5 interval (Horev et al., 2011) and the second *Del/+* model display dramatic hearing deficit which can alter behavior of animals (Portmann et al., 2014). To better achieve modeling of 16p11.2 BP4-BP5 rearrangement syndromes, we engineered and investigated novel mouse models carrying deletion and/or duplication for the *Sult1a1-Spn* region in two genetic backgrounds. We used a comprehensive test battery to assess the impact of the genetic rearrangements on behavior, synaptic function and metabolism of mice. Our results shade light on the recurrence of opposite phenotypes associated with the two rearrangements of the 16p11.2 syntenic region.

Del/+ and *Dup/+* animals were first amplified separately on a pure B6N genetic background. Animals were passed into two independent behavioral pipelines in order to evaluate every parameter which can be related to neuropsychiatric disorders of human patients. *Del/+* displayed hyperactivity, recognition memory deficits whereas *Dup/+* mice displayed hypoactivity and recognition memory improvements. Interestingly, *Del/+* mice also showed vertical stereotypic behaviors commonly found in mouse models of autism (Crawley, 2007). No obvious sign of anxiety was detected in mutant mice. *Del/+* and *Dup/+* mice displayed similar social behaviors, prepulse inhibition of the startle response and seizure susceptibility than wild-types indicating no alteration of mouse parameters which can be related to ASD, schizophrenia and epilepsy found in patients carrying rearrangements for the 16p11.2 BP4-BP5 region.

To confirm phenotypes, we generated a second cohort by crossing *Del/+* with *Dup/+* animals in order to compare *Del/+* and *Dup/+* mutant mice with a single group of wt animals and also to characterize *Del/Dup* animals carrying the two *Sult1a1-Spn* region on a single chromosome. We confirmed previous phenotypes for activity and recognition memory of mutant mice (Table 2). We were particularly interested in modulations of learning and memory functions found in mutant mice which can be linked to intellectual disabilities found in human patients. Thus, we performed the novel object recognition task using two retention delays: 30 min and 3 hours. With the shorter delay, *Del/+* mice showed deficits in recognition memory. With the longer delay, similarly to former results, *Del/+* mice and *Dup/+* mice displayed deficits and improvements respectively for novel object discrimination. Interestingly, *Del/Dup* animals had similar phenotypes than *Dup/+* animals and showed recognition memory improvements in comparison with wild-types. This suggests that this memory improvement phenotype may be linked to DNA structure changes associated with the duplication of the *Sult1a1-Spn* region.

In order to study the influence of genetic background, we generated *Del/+* and *Dup/+* animals on a mixed B6NC3B genetic background. Due to C3B alleles, we observed an important attenuation of activity phenotypes in *Del/+* and *Dup/+* mice. Indeed, circadian activity alterations were still found for mutant mice but no alteration of exploratory activity in the novel open field area was noticed. In the same way, we only found higher and lower level of climbing for *Del/+* and *Dup/+* mice respectively. We did not observed vertical stereotypy for *Del/+* mice. Horev *et al.* found similar activity phenotypes with their *Del/+* and *Dup/+* models maintained on a mixed B6N129Sv background (Horev et al., 2011) (Table 3). We found no phenotypes of motor coordination but we still observed higher and lower level of grip strength for *Del/+* and *Dup/+* mice respectively. With a retention delay of 3 hours, we found again recognition memory deficits for *Del/+* mice and improvements for *Dup/+* mice. Portmann *et al.* did the same experiment with a retention delay of 1 hour. The authors also found an important deficit for their *Del/+* model maintained on a mixed B6N129Mo (Portmann et al., 2014) (Table 3). Very interestingly, in comparison with wild-types, we found a diminution of social behaviors for *Del/+* and *Dup/+* mice in the social interaction test. These phenotypes are directly linked with ASD found in patients carrying the deletion and the duplication of the 16p11.2 BP4-BP5 region (Fernandez et al., 2010; Weiss et al., 2008).Reduction of social interaction observed for mutant mice in a mixed B6NC3B genetic background is the first phenotype going in the same way for *Del/+* and *Dup/+* animals. These

phenotypes which were not observed in B6N animals confirm that the genetic background is prominent for neuropsychiatric disorder modeling (Kerr et al., 2013; Pietropaolo et al., 2011).

Under standard diet, we observed an important reduction of body mass and size and adipogenesis deficits in *Del/+* animals and an increase of body mass in *Dup/+* animals. These phenotypes were exacerbated under high-fat diet. Surprisingly, body mass and adipogenesis mouse phenotypes were mirroring those observed in 16p11.2 CNVs human carriers. Differences between human and mouse can be partly explained by food intake. In human, 16p11.2 BP4-BP5 deletion is associated with obsesity and hyperphagia (Walters et al., 2010) whereas reciprocal duplication is associated with being underweight and restrictive eating behavior (Jacquemont et al., 2011). In comparison with wild-type littermates, *Del/+* and *Dup/+* mice displayed normal food and water consumption. Opposite body mass and adipogenesis between humans and mice can be explained by the deregulation of gene(s) located outside the 16p11.2 region. An interesting candidate to explain feeding and BMI alteration of patients is *SH2B1* implicated in leptin and insulin signaling controlling the regulation of appetite and energy expenditure. *SH2B1* is located in the 16p11.2 BP2-BP3 region 800 kb upstream of BP4-BP5 region. Interestingly, deletions of the 16p11.2 BP2-BP3 are also associated with obesity and developmental delay (Bachmann-Gagescu et al., 2010; Perrone et al., 2010). This gene was not deregulated in lymphoblastoid cell lines of deletion and duplication carriers (Jacquemont et al., 2011; Walters et al., 2010) but this does not mean that *SH2B1* is not deregulated in the brain of patients.

These several opposite phenotypes found with two different genetic backgrounds suggest the implication of gene(s) sensitive to dosage which could impact on mouse metabolism, activity and recognition memory. In order to investigate this hypothesis, we analyzed the transcriptome profiles of *Del/+*, *Dup/+* and wt animals in brain and peripheral tissues. Analysis revealed that the majority of *Sult1a1-Spn* region was sensitive to gene dosage. Nevertheless, rearrangements of *Sult1a1-Spn* region had a low impact on whole genome transcriptome suggesting that dosage sensitive gens leading to opposite phenotypes may be located in the 16p11.2 syntenic region. Consistent with normal feeding behavior of mutant mice, no deregulation of *Sh2b1* gene was found in the tissues analyzed. Principal component analysis further suggested that the effect of the deletion on the expression of the genes in the *Sult1a1-Spn* interval was stronger, as already reported by Horev *et al*. This may

also indicate why anatomic, metabolic and behavioral phenotypes were stronger in *Del/+* animals compared to opposite phenotypes of *Dup/+* animals.

Importantly, our results suggest that the pathways through which 16p11.2 deletion and duplication result in intellectual disability might include alteration of the expression of genes included in the 16p11.2 BP4-BP5 region. Alterations of social behaviors were dependent of the genetic background. Thus, we suggest that potential secondary polymorphisms may impact on neuropsychiatric disorders associated with 16p11.2 CNVs. Finally, body mass index alterations could result from deregulation of gene(s) located outside the 16p11.2 BP4-BP5 region which can differ(s) from human to mouse, leading to opposite body mass and adipogenesis phenotypes.

MATERIALS AND METHODS

Mouse lines, genotyping and ethical statement

16p11.2Yah mice were generated through Cre-LoxP in vivo recombination using a mouse line carrying two loxP sites inserted upstream upstream *Sult1a1* and downstream *Spn* genes in a trans configuration. Deletion of *Sult1a1–Spn* region was identified by PCR using primers Fwd1 (5'-CCTGTGTGTATTCTCAGCCTCAGGATG-3') and Rev2 (5'-GGACACACAGG AGAGCTATCCAGGTC-3'). Duplication of the same region was identified using primers Fwd2 (5'-ACTGCAGCCCGTCACCTAACTTCTT-3') and Rev1 (5'-GGACACACAGGAG AGCTATCCAGGTC-3'). Wild-type allele was identified using Fwd1 and Rev1 primers. PCR reactions gave deletion, duplication and wild-type products of 500 bp, 461 bp and 330 bp long respectively (Figure 1C). All mice were genotyped by PCR reaction following the program: 95°C /5min; 35x (95°C/30sec, 65°C/30sec, 70°C/1min), 70°C/5min. Experimental procedures were approved by the local ethical committee Com'Eth under accreditation number (2012-069). YH, is the principal investigator of this study, (accreditation 67-369).

Viability test

Viability tests were done in order to study causes of Del/+ lethality and to evaluate weight of animals at birth. Fetuses were collected at embryonic stage 18.5 (E18.5) to monitor their capacities to survive at birth. Fetuses were weighted and placed on a warm plate at 37°C and rolled gently to stimulate them to breathe. 30 minutes after extraction, numbers of breathing animals versus cyanotic and lethargic animals were counted. Tail samples were collected for genotyping. Over 69 B6N fetuses isolated from 7 pregnant wild-type females impregnated by Del/+ males, we genotyped 37 wt and 32 *Del/+* fetuses. Over 54 B6C3 fetuses isolated from 7 pregnant wild-type C3B females impregnated by Del/+ B6N males, we genotyped 24 wt and 30 del/+ fetuses. Over 51 B6N fetuses isolated from 6 pregnant wild-type females impregnated by *Dup/+* males, we genotyped 24 wt and 27 *Dup/+* fetuses.

Behavioral analysis

Only male mice were used for behavioral studies. Cohorts of experimental animals were formed by selecting mice coming from litters containing a minimum of two male pups were selected. After weaning, animals were gathered by litters in 39 x 20 x 16 cm cage (Green Line, Techniplast, Italy) where they had free access to water and food (D04 chow diet, Safe).

Temperature was maintained at 23±1°C and the light cycle was controlled as 12 hour light and 12 hour dark (lights on at 7 am). Mice were transferred from the animal facility to the phenotyping area at the age of 10 weeks. In the testing days, animals were transferred to experimental room antechambers 30 min before the start of the experiment. All the experiments were done between 8:00 AM and 2:00 PM. Behavioral experiments were conducted between 12 and 20 weeks of age. Body weights of animals were recorded once a week (same day at the same time) from the age of 12 weeks to sacrifice.

Del/+ and *Dup*/+ separate cohorts on B6N background were first analyzed. Two cohorts (n=15 wt, 11 *Del*/+; n= 13 wt, 11 *Dup*/+) were first analyzed separately at the same age. Tests were administered in the following order : plus maze, open field, Y maze, marble burying, forced swim, PPI, novel object recognition, three-chamber sociability, and repetitive behavior observation. Two other cohorts (n=9 wt, 8 *Del*/+; n= 10 wt, 12 *Dup*/+) were passed under a second behavioral pipeline : circadian activity, new location recognition, social interaction, rotarod, grip test, Morris water maze, PTZ sensitivity. Resting period of 2 days to 1 week was used between two consecutives tests.

To confirm first phenotypes and to study the behavior of animals carrying the deletion and the duplication, we crossed *Del*/+ with *Dup*/+ animals and generated two *Del-Dup* cohorts of males (n total = 11 *Del*/+, 19 wt, 14 *Del/Dup*, 19 *Dup*/+). Were passed animals at the same age under the same behavioral pipeline and we pooled the data together. Tests were administered in the following order: open field, sucrose preference, novel object recognition with 3h retention delay, circadian activity, repetitive behavior observation, object recognition with 30min retention delay, grip test, rotarod and notched bar. Resting period of 2 days to 1 week was used between two consecutives tests.

Finally, in order to study the influence of genetic background, we generated *Del*/+ and *Dup*/+ separate cohorts on B6NC3B background (n=12 wt, 14 *Del*/+; n= 13 wt, 12 *Dup*/+) by crossing B6N *Del*/+ and *Dup*/+ animals with C3B wild-type animals. Tests were administered in the following order: circadian activity, open field, new object recognition with 3 hour delay, Y maze, social interaction, repetitive behavior observation, grip test, rotarod, novel location recognition with 3 hour delay for Del/+ cohort and novel object recognition with 24 hour delay for Dup/+ cohort, and three-chamber sociability. Resting period of 2 days to 1 week was used between two consecutives tests.

Circadian activity (AC): The test allows the evaluation of endogenous activity and feeding behavior over a complete light/dark cycle. Testing was done in individual cages (11 x 21 x 18 cm^3) fitted with infrared captors linked to an electronic interface (Imetronic, France) which provide automated measures of position and locomotor activity were used for the experiment. Mice were put in cages at 11 am the first day and removed the second day at 7 pm. The light cycle was controlled as 12 h light and 12 h dark (lights on at 7 am). The 32 hours of test consist in three different phases: the habituation phase (from 11am to 7pm on the first day), the night/dark phase (from 7pm on the first day to 7am on the second day), and the day/light phase (from 7am to 7pm on the second day). Feeding behaviors were evaluated using an automated lickometer and a 20mg pellet feeder (Test Diet, Hoffman La-Roche).

Open-field (OF): The test is used to evaluate exploration behavior. Mice were tested in automated open fields (44.3 x 44.3 x 16.8 cm) made of PVC with transparent walls and a black floor, and covered with translucent PVC (Panlab, Barcelona, Spain). The open field arena is divided into central and peripheral regions and is homogeneously illuminated at 150 Lux. Each mouse was placed in the periphery of the open field and allowed to explore freely the apparatus for 30 min. The distance travelled, the number of rears and time spent in the central and peripheral parts of the arena were recorded over the test session.

Repetitive behavior: Mice were put individually in clean home-cages dimly lighted (60 lux) without the pellets and water bottle. The occurrence of repetitive behaviors (rearing, jumping, climbing, digging, grooming) was observed for 10 min and scored using an ethological keyboard (Viewpoint, Labwatcher, France).

Elevated plus maze: The apparatus consists of two opposed open arms (30 x 5 cm) crossed by two enclosed arms (30 x 5 x 15 cm), and elevated 66 cm from the floor. The light intensity at the extremity of the open arms was kept at 50 Lux. Each mouse was tested for 5 min after being placed in the central platform and allowed to explore freely the apparatus. The number of entries and time spent in the open arms were used as an anxiety index. Closed arm entries and rears in the closed arms were used as measures of general motor activity.

Y-maze: This test which is used to evaluate short-term working memory is based on the innate preference of animals to explore arms that has not been previously explored, a behavior that, if occurring with a frequency greater than 50%, is called spontaneous alternation. The apparatus consists in a Y-shaped maze with three white, opaque plexiglas arms of equivalent

length forming a 120° angle with each other. The arms have walls with specific motifs allowing distinction from each other. After introduction at the center of the maze at 60 Lux, animal were allowed to freely explore the maze for 6 minutes. The number of arm entries and the number of triads were recorded in order to calculate the percentage of alternation.

Novel object recognition task (NOR): This test evaluates object recognition memory and is based on the innate preference of rodents to explore novelty. The test was carried in an open field arena as previously described (Goeldner et al., 2008). On the first day, mice were habituated to the arena for 30 min at 60 Lux. The following day, animals were submitted to a first 10-min acquisition trial during which they were individually placed in the presence of the object A (marble or dice) placed at 10cm of one corner of the box. The exploration time of object A (when the animal's snout was directed toward the object at a distance ≤ 1 cm) was recorded. A 10-min retention trial (second trial) was conducted 30min, 3 hours or 24 hours later. The familiar object (object A) and novel object (object B) were placed at 10 cm of two open field corner (the distance between the two objects was about 20 cm) and the exploration time of the two objects was recorded. A discrimination index was defined as $(t_B/(t_A + t_B)) \times 100$. All mice which did not explore the first object more than 3 seconds during the acquisition trial were excluded from the analysis.

Object location recognition task (NLR): The test evaluates object location memory and is based on similar principles as NOR test. In the first 10-min acquisition trial, animals were exposed to two objects, a green marble and a dice. 3 hours later, one of the familiar object was displaced to a novel location (object B) and the exploration time of the two objects was recorded for 10 min. A discrimination index was defined as $(t_B/(t_A + t_B)) \times 100$. All mice which did not explore both objects more than 4 seconds during the acquisition trial were excluded from the analysis.

Morris water maze paradigm (MWM): The test is a paradigm for spatial learning and memory. The protocol is adapted from the already established protocol (Duchon et al., 2011). The apparatus consists in a circular pool (150-cm diameter, 60-cm height) filled to a depth of 40 cm with water maintained at 20°C–22°C and made opaque using a white aqueous emulsion (Acusol OP 301 opacifier). An escape platform, made of 6 cm diameter rough plastic, is submerged 1 cm below the water surface. The test began with 6 days of acquisition, 4 trials per day, at 120 Lux. Each trial started with the mice facing the interior wall of the pool and ended when animals climb on the platform or after a maximum searching time of 90 sec. The

platform was at the same position for all the four trials but starting positions changed randomly between each trial with departures from each cardinal point. Travelled distances to find the platform and swimming speeds were analyzed each day. On 7th day, mice were given a single trial of 60 seconds trial during the probe test or removal session in which the platform has been removed. The distance traveled and duration spent in each quadrant (NW, NE, SW, SE) were recorded. Annulus crossing index was calculated as the number of times that animals crossed the exact platform position. On 12th and 13th days, mice were given reversal sessions with 4 trials of 90 seconds per day. The platform was made visible by a small dark ball placed 12cm on top of the platform, while the external cues were hidden by surrounding the pool with a black curtain. In order to be sure that the mouse used the platform cue, starting position and platform position were changed for each trial.

Social interaction task: The test allows the evaluation of social behaviors. In each session, 2 mice of same genotype and similar body weight housed in different cages were put in open field area during 10 min. The duration of sniffing and following social behaviors were recorded during the session.

Three-chamber sociability: The test allows the evaluation of social preference and discrimination using a specific apparatus (Stoelting, Dublin) with three successive and identical chambers (20cm x 40cm x 22(height) cm with 5cm x 8cm openings allowing access between the chambers). The testing protocol is performed as previously described (Moy et al., 2004). In the habituation, mice were allowed to explore the three chambers freely for 10 min. In the second phase, the test mouse was places in the central box, while an unfamiliar mouse (stranger 1) was put in one of the wire cages in a random and balanced manner. The doors were reopened and the test mouse was allowed to explore the chambers for 10 min. Time spent to explore empty cage and cage with stranger 1 was recorded. In the third phase, the social discrimination was evaluated with a new mouse placed into the empty wire cage and the test mouse was allowed to explore again the entire arena for 10 min, having the choice between the familiar mouse (stranger 1 or 1) and the novel mouse (stranger 2 or 2). A social discrimination index or social preference was defined as the percentage of time spent exploring the novel mouse as $(t_2/(t_1 + t_2)) \times 100$.

Sucrose preference: The test allows assessment of hedonic behavior by measuring preference for sucrose when the animal has free access to two bottles containing either a 0.8% sucrose solution or water. Mice were first habituated to sucrose in their home cages where water was

replaced with sucrose solution (from 5pm on day 1 to 9am on day 2). From the second to the fourth days, mice were evaluated for sucrose preference in testing cages (30 x 15 x 12 cm^3). At 5pm of each day, mice were put in individual cages with food *ad libitum*. One hour later, two bottles of water and 0.8% sucrose were provided for 15 hours. At 9am the next day, mice were put back to their home cage and the two bottles were weighted.

Forced swim: The test allows induction of a depressed state by forcing mice to swim in a narrow cylinder from which they cannot escape. Mice were placed in a Plexiglas cylinder containing water (21°C-23°C) to a depth of 15 cm. After a brief period of vigorous activity, mice adopt a characteristic floating posture. Each animal was submitted to a forced swim session of 6 min and the total duration of floating and swimming behaviors were measured.

Acoustic startle reflex and Prepulse Inhibition (PPI): Testing was conducted in eight startle devices (SRLAB, San Diego Instruments, San Diego, CA), each consisting of a Plexiglas cylinder (5.1 cm outside diameter) mounted on a Plexiglas platform in a ventilated, sound-attenuated cubicle with a high-frequency loudspeaker (28 cm above the cylinder) producing both a continuous background noise and the various acoustic stimuli. The background noise of each chamber was set at 65 dB. Movements within the cylinder were detected and transduced by a piezoelectric accelerometer attached to the Plexiglas base, digitized, and stored by a computer. Beginning at the stimulus onset, 65 readings of 1ms were recorded to obtain the animal's startle amplitude. Auditory stimuli are burst of white noise (0–20KHz and 0 ms rise–decay). The optimal PPI parameters (eg, stimuli intensity and duration, prepulse–pulse intervals) were defined based on our previous validation studies (Aubert et al., 2006). Stimuli levels and piezoaccelerometer sensitivity were calibrated before each PPI session. Testing is adapted from the already established protocol (Ouagazzal et al., 2006). The session was initiated with 5 min acclimation period followed by 10 different trial types: acoustic startle pulse alone (white noise, 110 dB/40 ms); 8 different prepulse trials in which either 20 ms long 70, 80, 85 or 90 dB stimuli were presented alone or preceded the pulse by 50 ms, and finally 1 trial in which only the background noise was presented to measure the baseline movement in the Plexiglas cylinder. The test session begun with five presentations of the startle pulse alone trial, which were excluded from statistical analysis. Then, each acoustic or BN trial was presented 10 times in random order. The average ITI was 15 s (10–20 s).

PTZ sensitivity: Pentylenetetrazol (PTZ)-induced seizure test is used to determine the seizure threshold of animals. PTZ acts as a non-competitive GABA-A receptor antagonist, reducing neuronal inhibition (Ramanjaneyulu and Ticku, 1984). This substance induces myoclonic, clonic or tonic seizures depending on the dose injected (Meldrum, 2002). PTZ solution was freshly prepared (30 mg in 10 ml of 0.9% NaCl) and inject intraperitoneally at a 30 mg/kl dosage (300µL of solution for a 30g mouse). Immediately after PTZ adimistration, mice were put in a clean translucent arena. Events of myoclonic seizures were noted. If mice had clonic seizure, the latence after PTZ injection and the duration of event was noted.

Rotarod: The test assesses sensorimotor coordination and balance. The apparatus (Biosed, France) is a rotating bar of 5 cm diameter (hard plastic materiel covered by grey rubber foam) on which mice are placed facing the direction of rotation. Animals were first habituated to stay on the rod for 30 seconds at a constant speed of 4 rotations per minute (rpm). This was followed by 3 training days with 4 trials per day. Mice were placed on an accelerating rod increasing from 4 rpm to 40 rpm in 5 min of time. The test was stopped when mice fell down from the rod or when they made more than one passive rotation. The latency to fall and the maximal speed before falling was recorded.

Notched bar: The test is used for testing the hindlimb coordination. Testing was performed as previously described (Duchon et al., 2011). Mice were trained with a bar consisting of a natural wooden piece 1.7 mm large and 50 cm long bearing terminal platforms of 6 cm × 6 cm. The second day, mice were tested at 100 Lux with a notched bar of same dimension constituted of 12 platforms of $2cm^2$ spaced out by 13 gaps of $2cm^3$. Animals had to cross the notched bar twice for training and 15 times for test. Each time a back paw went through the gap was considered as an error and global error percentage was calculated.

Grip strength: The test is used to evaluate muscular strength. Mice were first weighted and tested with a handy force gauge (Bioseb, France). Animals were placed on the instrument grid and pulled by the tail until letting go. The force (g) was related to animal weight (g).

Social odor discrimination: This test is undertaken to validate sociability results based principally on olfaction. The test has been adapted from a previously described procedure (Yang and Crawley, 2009). Mice were assessed for water versus social odor discrimination. Odor exploration was measured in a clean cage with fresh bedding. Odors were presented on two cotton-tipped wooden applicators. Animals were first habituated to the novel cage for 3

min, and then tested during 3 consecutive sessions of 3 min each with 2 min inter-trial interval (ITI). During each session, two applicators are presented: (session 1) water-water, (session 2) water-social odor, and (session 3) water-social odor. The position of the odor for the sessions 2 and 3 was randomized. Cotton sniffing durations (when the animal's snout was directed toward the cotton at a distance ≤1 cm) were recorded for each session.

Nonsocial odor discrimination: The protocol is adapted from the already established protocol (Ferguson et al., 2000). Odor exploration was measured in a clean cage with fresh bedding. Odors were presented on a small piece of Whatman paper placed in a perforated tube (H: 4cm, diameter: 3cm). Animals were first habituated to the novel cage for 3 min, and then tested during 5 consecutive sessions of 2 min each with 8 min inter-trial interval (ITI). In the 4 first sessions, the perforated tube contained Whatman paper soaked with orange flower water. On the last session, Whatman paper was replaced with one soaked with vanilla extract. Odor sniffing duration (when the animal's snout was directed toward the perforated tube at a distance ≤1 cm) was recorded for each session.

Auditory brain response (ABR) recordings

The auditory brainstem response (ABR) is a method for determination of hearing sensitivity (Kuhn et al., 2012). =4 *Del/+*, 4 wt, 4 *Del/Dup* and 2 *Dup/+* animals were anaesthetized using ketamine/xylazine (ip) and placed on a heating blanket inside a sound attenuating booth. Sub-cutaneous needle electrodes were inserted in the skin on the vertex (active) and overlying the ventral region of the left (reference) and right (ground) bullae to record responses of the left ear. Stimuli were presented as free-field sounds from a speaker (Tucker Davis Technologies, FF1) whose leading edge was 10 cm in front of the mouse's interaural axis at an elevation of 30°. The sound delivery system was calibrated using an ACO Pacific 7017 microphone. For threshold determination, custom software and Tucker Davis Technologies hardware were used to deliver click (0.01 ms duration) and tone pip (6, 12, 18, 24 and 30kHz of 5 ms duration, 1 ms rise/fall time) stimuli at a range of intensity levels from 10–100+ dB SPL in 3 dB steps. Averaged responses to 256 stimuli, presented at 42.2/s, were analyzed and thresholds established as the lowest sound intensity giving a visually-detectable ABR response. For clicks, responses were also recorded from the right ear.

Hippocampal slices electrophysiology

Acute hippocampal slices were used to record field excitatory post synaptic potentials (fEPSPs), by the MEA60 electrophysiological suite (Multi Channel Systems, Reutlingen, FRG) as previously described (Coba et al., 2012; Kopanitsa et al., 2006). Eight set-ups consisting of a MEA1060-BC pre-amplifier and a filter amplifier (gain 550x) were run simultaneously by data acquisition units operated by MC_Rack software. Raw electrode data were digitized at 10 kHz and stored on a PC hard disk for subsequent analysis. To record fEPSPs, a hippocampal slice was placed into the well of 5x13 3D multi electrode array (MEA) biochip (Qwane Biosciences, Lausanne, Switzerland). The slice was guided to a desired position with a fine paint brush and gently fixed over MEA electrodes by a silver ring with attached nylon mesh lowered vertically by a one-dimensional U-1C micromanipulator (You Ltd, Tokyo, Japan). MEA biochips were fitted into the pre-amplifier case and fresh ACSF was delivered to the MEA well through a temperature-controlled perfusion cannula that warmed perfused media to 32°C. Monopolar stimulation of Schäffer collateral /commissural fibers through array electrodes was performed by STG4008 stimulus generator (Multi Channel Systems, Reutlingen, FRG). Biphasic (positive/negative, 100 µs/a phase) voltage pulses were used. Amplitude, duration and frequency of stimulation were controlled by MC_Stimulus II software. All experiments were performed using two-pathway stimulation of Schäffer collateral/commissural fibers. Our previous experiments that utilized MEAs, demonstrated that largest LTP was recorded in proximal part of apical dendrites of CA1 pyramidal neurons (Kopanitsa et al., 2006). We have therefore picked a single principal recording electrode in the middle of proximal part of CA1 and assigned two electrodes for stimulation of the control and test pathways on the subicular side and on the CA3 side of *stratum radiatum* respectively. The distance from the recording electrode to the test stimulation electrode was 400-510 µm and to the control stimulation electrode 316-447 µm. To evoke orthodromic fEPSPs, test and control pathways were activated in succession at a frequency of 0.02 Hz. Baseline stimulation strength was adjusted to evoke a response that corresponded to 40% of the maximal attainable fEPSP at the recording electrode located in proximal *stratum radiatum*. Slope of the negative part of fEPSPs was used as a measure of the synaptic strength. Paired stimulation with an interpulse interval of 50 ms was used to observe paired-pulse facilitation (PPF) in baseline conditions in the test pathway before LTP induction. PPF was calculated by dividing the negative slope of fEPSP obtained in response to the second pulse by the amplitude of fEPSP amplitude evoked by the preceding pulse. To

induce LTP, 10 bursts of baseline strength stimuli were administered at 5 Hz to test pathway with 4 pulses given at 100 Hz per burst (total 40 stimuli). LTP plots were scaled to the average of the first five baseline points. Normalisation of LTP values was performed by dividing the fEPSP amplitude in the tetanised pathway by the amplitude of the control fEPSP at corresponding time points. Normalised LTP values averaged across the period of 61-65 min after theta-burst stimulation were used for statistical comparison.

Metabolic and biochemistry analysis under high-fat diet

We only used females in order to avoid the use of a greater number of animals. *Del/+* and *Dup/+* separate cohorts on B6N genetic background (n = 11 wt and 9 Del/+; n = 8 wt and 10 Dup/+) and a *Del/+* cohort on B6C3B background (n= 8 wt and 10 Del/+) were put under high-fat diet for metabolism and biochemical analysis.

Mice were transferred from animal facility to phenotyping area at the age of 4 weeks. Animals were housed 2 to 4 per cage and fed with a standard chow diet (D04, Safe, USA) until the age of 5 weeks where the diet was switched to a high fat/ high carbohydrate diet (HFHCD, RD 12492, Research Diet, USA) until the end of the study. Body weights were recorded once a week from the age of 6 to 13 weeks. At the age of 11 weeks mice were put individually in the TSE cages for 24 hours for the measurement of energy expenditure by indirect calorimetry. At the age of 12 weeks an intra-peritoneal glucose tolerance test (IPGTT) was performed after overnight fast. Body composition (lean, fat and free body fluid content) was evaluated on conscious mice by quantitative nuclear magnetic resonance (qNMR) at the age of 13 weeks. At the age of 15 weeks blood was collected by retro orbital puncture under isoflurane anaesthesia for biochemistry, hematology and endocrine analysis. One week later mice were put in individual cages for 48h in order to collect feces for the measurement of energy content by calorimetric bomb.

Body composition by qNMR: Analysis was performed at 10 weeks and 15 weeks of age. This procedure was used to give precise analysis of the body composition for fat content, lean tissues and free body fluid by nuclear magnetic resonance apparatus and Minispec+ analyser (Bruker, Germany). The test was conducted during light period on conscious fed mice.

Indirect calorimetry: Energy expenditure was evaluated through indirect calorimetry by measuring oxygen consumption with an open flow respirometric system (TSE system,

Labmaster, Germany). Sensors measured the difference in CO_2 and O_2 concentrations in air volumes flowing through control or animal cages. The amount of oxygen consumed over a given period of time was calculated, as far as the air flow through the cage is known. Data from gas exchange are expressed as Kcal/h/kg$^{0.75}$. The system also monitored CO_2 production, thus the respiratory exchange ratio (RER = VCO_2/VO_2; which define fuel preference between glucose vs. lipid metabolism) and finally, heat production (Kcal/h/kg$^{0.75}$) was calculated. Following a 3h-acclimatization period, the experiment was performed over 21 hours: from 17:00 the first day to 11:00 the day after, at ambient temperature (21°C ± 2) under 7:00/19:00 light/dark cycle.

Intraperitoneal glucose tolerance test (IPGTT): IPGTT was performed at the age of 14 weeks after 9 weeks of HFHCD feeding. The glucose tolerance test was used to value the regulation of the glycemia after an induced hyperglycemia by injection of a standardized glucose bolus (2g/kg). Glucose solution was administered by intraperitoneal injection. Blood glucose was measured at different time points over 120 minutes after the injection on a drop of blood collected at the tail using blood glucose monitor and glucose test strips (Roche Diagnostics, Accu-Chek, France). The test was conducted during the light period, after 16 hours of overnight fasting.

Blood analyses: Blood was collected at the age of 16 weeks by retro orbital puncture under isoflurane anesthesia after 4 hours fasting. Blood chemistry was performed using a chemistry analyzer (Olympus AU-400, USA) using commercial reagent (Olympus Diagnostica GmbH, Lismeehan, Ireland). The following parameters were measured: total cholesterol, HDL and LDL cholesterol, triglycerides, free fatty acids, total proteins, albumin, calcium, phosphorus, glucose, urea, creatinine, total proteins and albumin. Plasma insulin and leptin levels were measured on a BioPlex analyser using Milliplex beads Panel (Millipore, USA). Plasma adiponectin levels were measured using Mouse adiponectin ELISA kit f (R&D system, USA).

Feces energy content by calorimetric bomb: The feces were collected during 48 hours on mice housed individually. The energy content of stools was evaluated in a bomb calorimeter (C503 control, IKA, USA). The sample was burned in an oxygen-rich atmosphere inside a sealed chamber surrounded by a jacket containing a known volume of water. The rise in temperature of the water was recorded and used to calculate the amount of heat produced. The assay was performed on stools to valuate the energy digested by mice and then indirectly the

intestinal function. The energy digested was calculated by the difference between the total calories ingested and excreted in faeces.

Craniofacial analysis

Craniums of 13 weeks old females were obtained from *Del/+* cohort (n = 10 wt and 10 Del/+) and *Dup/+* cohort (n = 8 wt and 8 Dup/+) on B6N genetic background were stored in 100% ethanol. 3D coordinates of 39 cranial relevant landmarks were recorded using Landmark software and posterior comparisons were performed using the Euclidean distance matrix analysis (EDMA) with the software WinEDMA (version 1.0.1 beta). 3D data were converted into linear distances compiling into a matrix. Both the form (size of the skull) difference matrix (FDM) and the shape difference matrix (SDM) were analyzed. A ratio different from 1 (FDM) or 0 (SDM) for any linear distance indicates that the two samples are not similar for that measure.

Confidence intervals were estimated using a non-parametric bootstrapping algorithm. For each linear distance, the null hypothesis was rejected if the 90% confidence interval does not include 1 (FDM) or 0 (SDM): rejection of the null hypothesis enables localization of differences to specific landmarks and linear distances. Bootstrap Distributions of T (FDM) and Z (SDM) were calculated as follow: for each FDM is calculated a T value. The location of the T observed from the FDM, allows to calculate the probability (*p*) in this distribution.

Transcriptome analysis

Expression profiling after RNA extraction was performed using Affymetrix Mouse Gene *2.1* ST 24-Array Plate following the manufacturer protocol. All arrays passed the standard Affymetrix quality control checks. Expression signals from 12 liver samples and 44 brain samples (12 cerebellum, 16 striatum, and 16 hippocampus) were quantile normalized separately. We applied a non-specific filter to discard probesets with low signals (maximum signal lower than the median signal for all probesets) or with low variance (lowest 25%ile of variance). Differentially expressed genes were defined using linear models as implemented in Limma v3.18.13 (Smyth, 2004). We fitted a dosage model where the 16p11.2 copy number variants (CNVs) were considered as a numerical variable, i.e. we assumed a dosage effect of the CNVs: gene_i ~ b0+ b1CNV CNV=-1,0,1. Gene Ontology (GO) analysis was performed using the library topGO v2.14.0. Gene Set Enrichment Analysis (GSEA) was conducted using

the Broad Institute algorithm v2.1.0 (Mootha et al., 2003; Subramanian et al., 2005). We used the preranked gene lists defined according to the limma results based on the dosage model and we selected MSigDB C2 (c2.all.v4.0.symbols.gmt) that collects curated gene sets from online pathways, publications in PubMed, knowledge of domain experts.

Statistical analyses

Results were processed for statistical analysis using the Sigma Plot software (Sigma). All acquired behavioral data were analyzed using a one-way or a two-way ANOVA with a repeated measures followed by Student's t-test or Tukey's post-hoc test whenever it was appropriate. Otherwise, the non-parametric Kruskal-Wallis analysis and Mann-Whitney *U* test was done. One group Student t-tests were used to compare recognition index values to the chance level (50%). A Pearson's chi-squared test was used to evaluate mutant allele transmission. Data are represented as the mean ± SEM and the significant threshold was $p < 0.05$ or otherwise indicated.

In electrophysiological experiments input-output relationships were initially compared by mixed model repeated-measures ANOVA and Bonferroni *post hoc* test implemented in Prism 5 (GraphPad Software, Inc., San Diego, CA, USA) using individual slice data as independent observations. Since several slices were routinely recorded from every mouse, fEPSPmax, PPF and LTP values between wild-type and mutant mice were also compared using two-way nested ANOVA design with genotype (group) and mice (sub-group) as factors (STATISTICA v. 10, StatSoft, Inc., Tulsa, OK, USA). To contrast data obtained in *Del*/+, *Dup*/+ and *Del*/*Dup* mice to their common, litter-matched WTs, we used post hoc Dunnett's Multiple Comparison test on data from individual slices. Statistical effects were considered significant if $P < 0.05$. Graph plots and normalisation were performed using OriginPro 8.5 (OriginLab, Northampton, MA, USA). Throughout the text, electrophysiological data are presented as mean ± s.e.m. with *n* and *N* indicating number of slices and mice respectively.

REFERENCES

Aubert, L., Reiss, D., and Ouagazzal, A.M. (2006). Auditory and visual prepulse inhibition in mice: parametric analysis and strain comparisons. Genes Brain Behav *5*, 423-431.

Bachmann-Gagescu, R., Mefford, H.C., Cowan, C., Glew, G.M., Hing, A.V., Wallace, S., Bader, P.I., Hamati, A., Reitnauer, P.J., Smith, R.*, et al.* (2010). Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. Genet Med *12*, 641-647.

Bergen, S.E., O'Dushlaine, C.T., Ripke, S., Lee, P.H., Ruderfer, D.M., Akterin, S., Moran, J.L., Chambert, K.D., Handsaker, R.E., Backlund, L.*, et al.* (2012). Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. Molecular Psychiatry *17*, 880-886.

Brault, V., Pereira, P., Duchon, A., and Herault, Y. (2006). Modeling chromosomes in mouse to explore the function of genes, genomic disorders, and chromosomal organization. Plos Genetics *2*, 911-919.

Coba, M.P., Komiyama, N.H., Nithianantharajah, J., Kopanitsa, M.V., Indersmitten, T., Skene, N.G., Tuck, E.J., Fricker, D.G., Elsegood, K.A., Stanford, L.E.*, et al.* (2012). TNiK is required for postsynaptic and nuclear signaling pathways and cognitive function. J Neurosci *32*, 13987-13999.

Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y.J., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P.*, et al.* (2010). Origins and functional impact of copy number variation in the human genome. Nature *464*, 704-712.

Cook, E.H., and Scherer, S.W. (2008). Copy-number variations associated with neuropsychiatric conditions. Nature *455*, 919-923.

Cooper, G.M., Coe, B.P., Girirajan, S., Rosenfeld, J.A., Vu, T.H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V.*, et al.* (2011). A copy number variation morbidity map of developmental delay. Nature Genetics *43*, 838-U844.

Crawley, J.N. (2007). Mouse behavioral assays relevant to the symptoms of autism. Brain Pathology *17*, 448-459.

Crepel, A., Steyaert, J., De la Marche, W., De Wolf, V., Fryns, J.P., Noens, I., Devriendt, K., and Peeters, H. (2011). Narrowing the Critical Deletion Region for Autism Spectrum Disorders on 16p11.2. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics *156B*, 243-245.

de Anda, F.C., Rosario, A.L., Durak, O., Tran, T., Gräff, J., Meletis, K., Rei, D., Soda, T., Madabhushi, R., Ginty, D.D.*, et al.* (2012). Autism spectrum disorder susceptibility gene TAOK2 affects basal dendrite formation in the neocortex. Nat Neurosci *15*, 1022-1031.

Duchon, A., Pothion, S., Brault, V., Sharp, A.J., Tybulewicz, V.L.J., Fisher, E.M.C., and Herault, Y. (2011). The telomeric part of the human chromosome 21 from Cstb to Prmt2 is not necessary for the locomotor and short-term memory deficits observed in the Tc1 mouse model of Down syndrome. Behavioural Brain Research *217*, 271-281.

Ferguson, J.N., Young, L.J., Hearn, E.F., Matzuk, M.M., Insel, T.R., and Winslow, J.T. (2000). Social amnesia in mice lacking the oxytocin gene. Nat Genet *25*, 284-288.

Fernandez, B.A., Roberts, W., Chung, B., Weksberg, R., Meyn, S., Szatmari, P., Joseph-George, A.M., MacKay, S., Whitten, K., Noble, B.*, et al.* (2010). Phenotypic spectrum associated with de novo and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder. Journal of Medical Genetics *47*, 195-203.

Ghebranious, N., Giampietro, P.F., Wesbrook, F.P., and Rezkana, S.H. (2007). A novel microdeletion at 16p11.2 harbors candidate genes for aortic valve development, seizure disorder, and mild mental retardation. American Journal of Medical Genetics Part A *143A*, 1462-1471.

Goeldner, C., Reiss, D., Wichmann, J., Meziane, H., Kieffer, B.L., and Ouagazzal, A.M. (2008). Nociceptin receptor impairs recognition memory via interaction with NMDA receptor-dependent mitogen-activated protein kinase/extracellular signal-regulated kinase signaling in the hippocampus. J Neurosci *28*, 2190-2198.

Golzio, C., Willer, J., Talkowski, M.E., Oh, E.C., Taniguchi, Y., Jacquemont, S., Reymond, A., Sun, M., Sawa, A., Gusella, J.F.*, et al.* (2012). KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. Nature *485*, 363-U111.

Grayton, H.M., Fernandes, C., Rujescu, D., and Collier, D.A. (2012). Copy number variations in neurodevelopmental disorders. Progress in Neurobiology *99*, 81-91.

Hastings, P.J., Lupski, J.R., Rosenberg, S.M., and Ira, G. (2009). Mechanisms of change in gene copy number. Nature Reviews Genetics *10*, 551-564.

Herault, Y., Rassoulzadegan, M., Cuzin, F., and Duboule, D. (1998). Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). Nature Genetics *20*, 381-384.

Hoelter, S.M., Dalke, C., Kallnik, M., Becker, L., Horsch, M., Schrewe, A., Favor, J., Klopstock, T., Beckers, J., Ivandic, B.*, et al.* (2008). "Sighted C3H" mice--a tool for analysing the influence of vision on mouse behaviour? Front Biosci *13*, 5810-5823.

Horev, G., Ellegood, J., Lerch, J.P., Son, Y.E.E., Muthuswamy, L., Vogel, H., Krieger, A.M., Buja, A., Henkelman, R.M., Wigler, M.*, et al.* (2011). Dosage-dependent phenotypes in models of 16p11.2 lesions found in autism. Proceedings of the National Academy of Sciences of the United States of America *108*, 17076-17081.

Jacquemont, S., Reymond, A., Zufferey, F., Harewood, L., Walters, R.G., Kutalik, Z., Martinet, D., Shen, Y.P., Valsesia, A., Beckmann, N.D.*, et al.* (2011). Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. Nature *478*, 97-U111.

Kerr, T.M., Muller, C.L., Miah, M., Jetter, C.S., Pfeiffer, R., Shah, C., Baganz, N., Anderson, G.M., Crawley, J.N., Sutcliffe, J.S.*, et al.* (2013). Genetic background modulates phenotypes of serotonin transporter Ala56 knock-in mice. Mol Autism *4*, 35.

Kopanitsa, M.V., Afinowi, N.O., and Grant, S.G. (2006). Recording long-term potentiation of synaptic transmission by three-dimensional multi-electrode arrays. BMC Neurosci *7*, 61.

Kuhn, S., Ingham, N., Pearson, S., Gribble, S.M., Clayton, S., Steel, K.P., and Marcotti, W. (2012). Auditory function in the Tc1 mouse model of down syndrome suggests a limited region of human chromosome 21 involved in otitis media. PLoS One *7*, e31433.

Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y.*, et al.* (2008). Structural variation of chromosomes in autism spectrum disorder. American Journal of Human Genetics *82*, 477-488.

McCarthy, S.E., Makarov, V., Kirov, G., Addington, A.M., McClellan, J., Yoon, S., Perkins, D.O., Dickel, D.E., Kusenda, M., Krastoshevsky, O.*, et al.* (2009). Microduplications of 16p11.2 are associated with schizophrenia. Nature Genetics *41*, 1223-U1285.

Meldrum, B. (2002). Do preclinical seizure models preselect certain adverse effects of antiepileptic drugs. Epilepsy Res *50*, 33-40.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E.*, et al.* (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet *34*, 267-273.

Moy, S.S., Nadler, J.J., Perez, A., Barbaro, R.P., Johns, J.M., Magnuson, T.R., Piven, J., and Crawley, J.N. (2004). Sociability and preference for social novelty in five inbred strains: an approach to assess autistic-like behavior in mice. Genes Brain Behav *3*, 287-302.

Ouagazzal, A.M., Reiss, D., and Romand, R. (2006). Effects of age-related hearing loss on startle reflex and prepulse inhibition in mice on pure and mixed C57BL and 129 genetic background. Behav Brain Res *172*, 307-315.

Perrone, L., Marzuillo, P., Grandone, A., and del Giudice, E.M. (2010). Chromosome 16p11.2 deletions: another piece in the genetic puzzle of childhood obesity. Ital J Pediatr *36*, 43.

Pietropaolo, S., Guilleminot, A., Martin, B., D'Amato, F.R., and Crusio, W.E. (2011). Genetic-background modulation of core and variable autistic-like symptoms in Fmr1 knock-out mice. PLoS One *6*, e17073.

Portmann, T., Yang, M., Mao, R., Panagiotakos, G., Ellegood, J., Dolen, G., Bader, P.L., Grueter, B.A., Goold, C., Fisher, E.*, et al.* (2014). Behavioral abnormalities and circuit defects in the basal ganglia of a mouse model of 16p11.2 deletion syndrome. Cell Rep *7*, 1077-1092.

Ramanjaneyulu, R., and Ticku, M.K. (1984). Interactions of pentamethylenetetrazole and tetrazole analogues with the picrotoxinin site of the benzodiazepine-GABA receptor-ionophore complex. Eur J Pharmacol *98*, 337-345.

Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W.W.*, et al.* (2006). Global variation in copy number in the human genome. Nature *444*, 444-454.

Reinthaler, E.M., Lal, D., Lebon, S., Hildebrand, M.S., Dahl, H.H., Regan, B.M., Feucht, M., Steinböck, H., Neophytou, B., Ronen, G.M., *et al.* (2014). 16p11.2 600 kb Duplications confer risk for typical and atypical Rolandic epilepsy. Hum Mol Genet.

Sanders, S.J., Ercan-Sencicek, A.G., Hus, V., Luo, R., Murtha, M.T., Moreno-De-Luca, D., Chu, S.H., Moreau, M.P., Gupta, A.R., Thomson, S.A., *et al.* (2011). Multiple Recurrent De Novo CNVs, Including Duplications of the 7q11.23 Williams Syndrome Region, Are Strongly Associated with Autism. Neuron *70*, 863-885.

Shinawi, M., Liu, P.F., Kang, S.H.L., Shen, J., Belmont, J.W., Scott, D.A., Probst, F.J., Craigen, W.J., Graham, B.H., Pursley, A., *et al.* (2010). Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. Journal of Medical Genetics *47*, 332-341.

Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol *3*, Article3.

Steinberg, S., de Jong, S., Mattheisen, M., Costas, J., Demontis, D., Jamain, S., Pietilainen, O.P.H., Lin, K., Papiol, S., Huttenlocher, J., *et al.* (2014). Common variant at 16p11.2 conferring risk of psychosis. Molecular Psychiatry *19*, 108-114.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., *et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Tang, S.H.E., Silva, F.J., Tsark, W.M.K., and Mann, J.R. (2002). A Cre/loxP-deleter transgenic line in mouse strain 129S1/SvImJ. Genesis *32*, 199-202.

Walters, R.G., Jacquemont, S., Valsesia, A., de Smith, A.J., Martinet, D., Andersson, J., Falchi, M., Chen, F., Andrieux, J., Lobbens, S., *et al.* (2010). A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. Nature *463*, 671-U104.

Weiss, L.A., Shen, Y.P., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M.A.R., Green, T., *et al.* (2008). Association between microdeletion and microduplication at 16p11.2 and autism. New England Journal of Medicine *358*, 667-675.

Yang, M., and Crawley, J.N. (2009). Simple behavioral assessment of mouse olfaction. Curr Protoc Neurosci *Chapter 8*, Unit 8.24.

Zufferey, F., Sherr, E.H., Beckmann, N.D., Hanson, E., Maillard, A.M., Hippolyte, L., Mace, A., Ferrari, C., Kutalik, Z., Andrieux, J., *et al.* (2012). A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. Journal of Medical Genetics *49*, 660-668.

LEGENDS TO FIGURES

Figure 1. Mouse models for 16p11.2 rearrangements. **(A)** Top: human 16p11.2 region and proximal SGs prone to generate BP4-BP5 CNVs by non allelic homologous recombination. All genomic positions are given according to UCSC human genome browser GRCh38/hg38. Bottom: 16p11.2 syntenic region on mouse chromosome 7F3. Scares delineate genetic intervals studied in precedent publications and in the present study. All genomic positions are given according to UCSC mouse genome browser GRCm38/mm10. **(B)** Strategy for in vivo cre-mediated recombination and targeted meiotic recombination (TAMERE) crossing strategy. LoxP sites were inserted upstream of *Sult1a1* and downstream of *Spn*. Breeding were done in order to have in one trans-loxer female expressing the Hprt<tm1(cre)Mnn> transgene and carrying the two loxP sites in a *trans* configuration. Last step consisted in mating trans-loxer females with wild-type males to generate, in the progeny, individuals carrying the deletion or the duplication of the *Sult1a1-Spn* region. Del/+ and Dup/+ animals were crossed with wild-type animals to generate *Del/+* and *Dup/+* cohorts. In the case of *Del-Dup* cohorts, we crossed *Del/+* with *Dup/+* in order to generate *Del/+*, wt, *Dup/+* and *Del/Dup* animals. **(C)** Molecular validation. PCR products specific for the *Dup/+* and *Del/+* alleles are 500-bp and 461-bp long respectively.

Figure 2. Behavioral characterization of the *Del-Dup* cohort. **(A)** Circadian activity test. The light-light cycle was controlled as 12-hours light and 12-hours dark (lights on at 7 am). Graphs plot spontaneous locomotor activity (count) and rearing behavior (count) during dark and light phases. **(B)** Open field test. Exploratory locomotor activity (distance travelled in m) and percentage of time spent in the central area over 30 min of test. **(C)** Repetitive behaviors. The graph plots the occurrences of rearing, jumping, climbing and digging behaviors during 10 min of observation in a novel home-cage. **(D)** Social interaction test. The graph plots the duration of active social interaction (sniffing and following behaviors) between pairs of unfamiliar mice of the same genotype and equivalent body weight tested in a familiar open field area during 10 min. **(E)** Novel object recognition task. Recognition index reflects the ability of mice to distinguish the novel object from the familiar object after a short (30 min) and a long (3 hours) retention delays. The dashed line materializes a chance level of 50%. All genotypes performed significantly above the chance level. **(F)** Motor coordination evaluation. The first graph plots the latency (s) that mice stayed on the rod before falling under acceleration speed (4-40 rpm in 5 minutes) over 3 consecutive days of test. The second graph

plots hindlimb errors during notched bar test. Animals had to cross a notched bar and each time hind paw went through a gap was count as an error. Data are mean ± SEM. One-way ANOVA analysis, Tukey's test otherwise Kruskal-Wallis analysis, Mann-Whitney *U* test. $^{*}p <$ 0.05, $^{\#}$significantly different from all other groups.

Figure 3. Effects of *Sult1a1–Spn* region rearrangements on electrophysiological parameters measured in the CA1 area of hippocampal slices. **(A)** Baseline synaptic transmission was slightly decreased in mice with *Del/+, Dup/+* and *Del/Dup* mutations, but no significant overall genotype effect was observed ($F_{(3, 52)} = 1.37$; $p = 0.261$; RM ANOVA of data in individual slices). Input-output relationships illustrate averaged fEPSP slopes in slices from *Del/+* ($n = 10$; $N = 4$), *Dup/+* ($n = 10$; $N = 4$), *Del/Dup* ($n = 19$; $N = 6$) and wt mice ($n = 18$; $N = 8$) in response to stimulation of Schäffer collaterals by biphasic voltage pulses of 0.1 – 4.2 V. **(B)** Paired-pulse facilitation in slices from mice with *Sult1a1–Spn* region rearrangements was normal and no significant genotype effect was revealed ($F_{(3, 35)} = 0.0298$; $p = 0.993$, two way nested ANOVA). **(C)** Theta-burst stimulation elicited pathway-specific long-term potentiation of synaptic transmission in hippocampal CA1 area. Normalized magnitude of this potentiation 60-65 min after LTP induction was affected by genotype ($F_{(3, 35)} = 3.43$; $p = 0.027$, two way nested ANOVA). *Post hoc* Dunnett's Multiple Comparison test on values in individual slices demonstrated that LTP was significantly smaller ($q = 2.441$, $P < 0.05$) in slices from *Dup/+* mice ($n = 10$; $N = 4$) compared to wt slices ($n = 18$; $N = 8$), while levels of LTP in *Del/+*($n = 10$; $N = 4$) and *Del/Dup* slices ($n = 19$; $N = 6$) did not differ significantly from wt levels. Data are expressed as mean ± SEM. Overall effect of genotype was determined, as appropriate, by repeated measures (A) or two-way nested ANOVA (B, C) and *post hoc* comparisons to wt values were done using Dunnett's multiple comparison test.

Figure 4. Weight and adiposity characterization under standard and high-fat diets. **(A-D)** Weight, adiposity and body length of the *Del-Dup* littermates under standard diet. **(A)** Evolution of body weight (g) of adult animals. **(B)** Body fat percentage of 20-weeks old animals measured by qNMR. **(C)** Body length (distance from snout to tail basis) of 20-weeks animals. **(D)** Adult male *Del/+* (21.9 grams), +/+ (31 g), *Del/Dup* (33.9 g) and *Dup/+* (35.8 g) littermates at 24-weeks of age. *Del/+* mouse was shorter and showed an important lack of epididymal adipose tissues (white arrows). **(E-I)** Metabolic and endocrinology analysis of B6N *Del/+* and *Dup/+* separate cohorts put under high-fat diet at 5 weeks of age. **(E)** Evolution of body weight (g). **(F)** Body fat percentage of 13-weeks old animals measured by

qNMR. **(G)** Energy expenditure (Kcal/h per kg^0.75 of body weight) during 12-hours dark and 12-hours light phases. **(H)** Intraperitoneal glucose tolerance test. Evolution of blood glucose (mg/dl) and glucose area under the curve (AUC) index (min*mg/dl*0.001) **(I)** Endocrinology analysis. Blood insulin (µg/l), leptin (ng/ml) and adiponectin (µg/ml) levels. Data are mean ± SEM. **(A, E)** Repeated Measures ANOVA "genotype" analysis, Tukey's test. **(B)** Kruskal-Wallis analysis, Mann-Whitney *U* test. **(C)** One-way ANOVA analysis, Tukey's test. **(F-I)** Student's t-test. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{\#}$significantly different from all other groups.

Figure 5. Craniofacial analysis of separate *Del/+* and *Dup/+* cohorts. **(A)** Representative reconstructed 3D skull images and landmark distribution. Euclidian distances between the different landmaks allowed calculation of both the form (or size) difference (FD) and the shape difference (SD). Analysis revealed size reduction of the skull of *Del/+* animals **(B)** and no alteration of skull size of *Dup/+* animals **(D)**. An alteration of skull shape was found for *Del/+* **(C)** and *Dup/+* **(E)** mice.

Figure 6. Expression levels of genes within the Sult1a1-Spn region for brain and peripheral tissues. Stripcharts of the expression level for 30 probesets within the engineered region in hippocampus **(A)**, striatum **(B)**, cerebellum **(C)**, and liver **(D)**. Vertical axis represents log2 fold change normalized expression ratio. Blue dots *Del/+* animals, black circles wt animals, orange dots *Dup/+* animals.

Supplementary Figure 1. Behavioral characterization of B6N *Del/+* and *Dup/+* separated cohorts. **(A)** Circadian activity test. Graphs plot spontaneous locomotor activity (count) and rearing behavior (count) during dark and light phases. **(B)** Open field test. Exploratory locomotor activity (distance travelled in m), vertical activity (count) and time percentage spent in the central area over 30 min of test. **(C)** Repetitive behaviors. Occurrences of rearing, jumping, climbing and digging during 10 min of observation in a novel home-cage. **(D)** Social interaction test. Graph plots the duration of social behaviors (sniffing and following) during 10 min of test in the open field. **(E)** Novel object recognition task. Recognition index reflects the ability of mice to distinguish the novel object from the familiar object after 3 hour retention delay. The dashed line materializes a chance level of 50%. All genotypes performed significantly above the chance level. **(F)** Object location recognition test. Recognition index reflects the ability of mice to detect spatial changes and distinguish the displaced familiar object from non-displaced familiar one. 3 hour delay was used between the acquisition and

choice trials. All genotypes performed significantly above the chance level. **(G)** Motor coordination evaluation. Graphs plot the latency (s) that mice stayed on the rod before falling under acceleration speed over 3 consecutive days of test. Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$.

Supplementary Figure 2. Spontaneous locomotor activity and feeding behavior of the *Del-Dup* cohort during the circadian activity test. Patterns of locomotor activity **(A)** and rearing behavior **(B)** profiles during 32-hours of test. **(C-D)** Feeding behaviors. Water **(C)** and pellet **(D)** consumption during the 32-hours of test. **(E)** Pellets loss by animals which passed through the holed ground. *Del/+* mice loss significantly more pellet compared to the other genotypes. Data are mean ± SEM. One-way ANOVA analysis otherwise Kruskal-Wallis analysis. **(E)** Kruskal-Wallis analysis, Mann-Whitney *U* test. $^{\#}$ significantly different from all other groups.

Supplementary Figure 3. Exploratory activity of the *Del-Dup* cohort during the open field. **(A)** Locomotor exploratory activity (distance travelled in m). Whereas wt and *Dup/+* mice showed a habituation and decreased their activity at the end of the test, *Del/+* and *Del/Dup* mice showed a constant activity. **(B)** Rearing behavior (count). All genotype showed similar pattern of vertical activity without habituation. **(C)** Time in center (%). *Del/+* mice showed a huge increase of time spent in the area center. Data are mean ± SEM. One-way ANOVA analysis, $^{*}p < 0.05$, $^{**}p < 0.01$.

Supplementary Figure 4. Audition capacities of the B6N *Del-Dup* cohort in the auditory brain response (ABR) experiment. **(A)** ABR threshold curve between 6 and 30 kHz. **(B)** ABR threshold in click test before and after the test. Data are mean ± SEM.

Supplementary Figure 5. Behavioral characterization of B6NC3B *Del/+* and *Dup/+* separated cohorts. **(A)** Circadian activity test. Graphs plot spontaneous locomotor activity (count) and rearing behavior (count) during dark and light phases. **(B)** Open field test. Exploratory locomotor activity (distance travelled in m), vertical activity (count) and time percentage spent in the central area over 30 min of test. **(C)** Repetitive behaviors. Occurrences of rearing, jumping, climbing and digging during 10 min of observation in a novel home-cage. **(D)** Social interaction test. Graph plots the duration of social behaviors (sniffing and following) during 10 min of test in the open field. **(E)** Novel object recognition task. Recognition index reflects the ability of mice to distinguish the novel object from the familiar

object after 3 hour retention delay. The dashed line materializes a chance level of 50%. All genotypes performed significantly above the chance level. **(F)** Motor coordination evaluation. Graphs plot the latency (s) that mice stayed on the rod before falling under acceleration speed over 3 consecutive days of test. Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Supplementary Figure 6. Olfaction capacities of the B6NC3B *Del*/+ and *Dup*/+ cohorts. **(A)** Social odor discrimination. Graph plots the sniffing time of cottons swabs soaked with water and social odors during 3 sessions of 3 min with 2 min inter-trial interval. In the first session (S1), the two cottons swabs were soaked with water. In the S2 and S3 sessions, one of the two cottons swabs was soaked with social odors. All mice showed change in the amount of time spent sniffing the social stimuli during S2 session and then showed habituation in S3 session. **(B)** Nonsocial odor discrimination. Graph plots the sniffing time of Whatman paper soaked with nonsocial odor during 5 sessions of 2 min with 8 min inter-trial interval. From S1 to S4, Whatman paper was soaked with orange flower water. On the last test session (TS), Whatman paper was replaced with one soaked with vanilla extract. All mice showed a clear change in the amount of time spent sniffing the new olfactory stimuli in the test session. Data are mean ± SEM.

Supplementary Figure 7. Body mass and size and adiposity of *Del*/+ and *Dup*/+ separated cohorts maintained on B6N and B6NC3B genetic background feed under standard diet. **(A-B)** Weight evaluation of B6N *Del*/+ and *Dup*/+ animals at E18.5 embryonic stage **(A)** and during adulthood **(B)** from 14 to 24 weeks. **(C-F)** Evaluation of B6C3B *Del*/+ and *Dup*/+ animals. **(C)** Weight of fetuses (g) at E18.5 embryonic stage. **(D)** Weight (g) evolution of adult mice. **(E)** Body fat percentage of 20-weeks old animals measured by qNMR. **(F)** Body length (distance from snout to tail basis) of 20-weeks animals. Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Supplementary Figure 8. Body mass and size, adiposity and metabolic characterization of *Del*/+ cohort maintained on B6NC3B genetic background feed under high-fat diet. **(A)** Evolution of body weight (g). **(B)** Body fat percentage measured by qNMR. **(C)** Energy expenditure (Kcal/h per kg^0.75 of body weight) during 12-hours dark and 12-hours light phases. **(D)** Intraperitoneal glucose tolerance test. Evolution of blood glucose (mg/dl) and glucose area under the curve (AUC) (min*mg/dl) **(E)** Endocrinology analysis. Leptin (ng/ml) and adiponectin (µg/ml) blood levels. **(F)** Blood chemistry analysis. Total cholesterol (µg/ml),

glucose (mmol/l) and calcium (mmol/l) blood levels. **(G)** Calorimetric bomb. Total daily energy excreted (Kcal). Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Supplementary Figure 9. Principal component analysis of the *Sult1a1-Spn* region. The first component (horizontal axis) explaining 58.2% of the variance is mainly capturing the CNV effect, the second component (vertical axis) explaining 28.7% of the variance in mainly capturing the regional effect, especially cerebellum (+) versus striatum (*) and hippocampus (-). Blue dots *Del/+* animals, black circles wt mice, orange dots *Dup/+* animals.

TABLES

Genetic backround	Crossings	Genotype	Observed number	Observed ratio	χ^2	*p*
B6N	*Del*/+ x wt	wt	230	69.3%	24.7	$6.79x10^{-07}$
		Del/+	102	30.7%		
B6N	*Dup*/+ x wt	wt	129	54.2%	0.84	0.36
		Dup/+	109	45.8%		
B6N	*Del*/+ x *Dup*/+	wt	117	34.6%	12.5	$4.07x10^{-04}$
		Del/+	39	11.6%	24.5	$7.43x10^{-07}$
		Dup/+	90	26.6%	0.36	0.55
		Del/Dup	92	27.2%	0.67	0.41
B6NC3B	*Del*/+ x wt	wt	80	49.1%	0.03	0.87
		Del/+	83	50.9%		
B6NC3B	*Dup*/+ x wt	wt	30	50.8%	0.01	0.93
		Dup/+	29	49.2%		

Table 1. Transmission rates of the *Sult1a1-Spn* deletion (*Del*) and duplication (*Dup*) alleles observed at weaning. On B6N genetic background, *Del* allele showed a reduced transmission rate. Combination of both *Del* and *Dup* alleles leads to rescue for this lethality. On B6NC3B genetic background, no lethality was associated with the *Del/* allele.

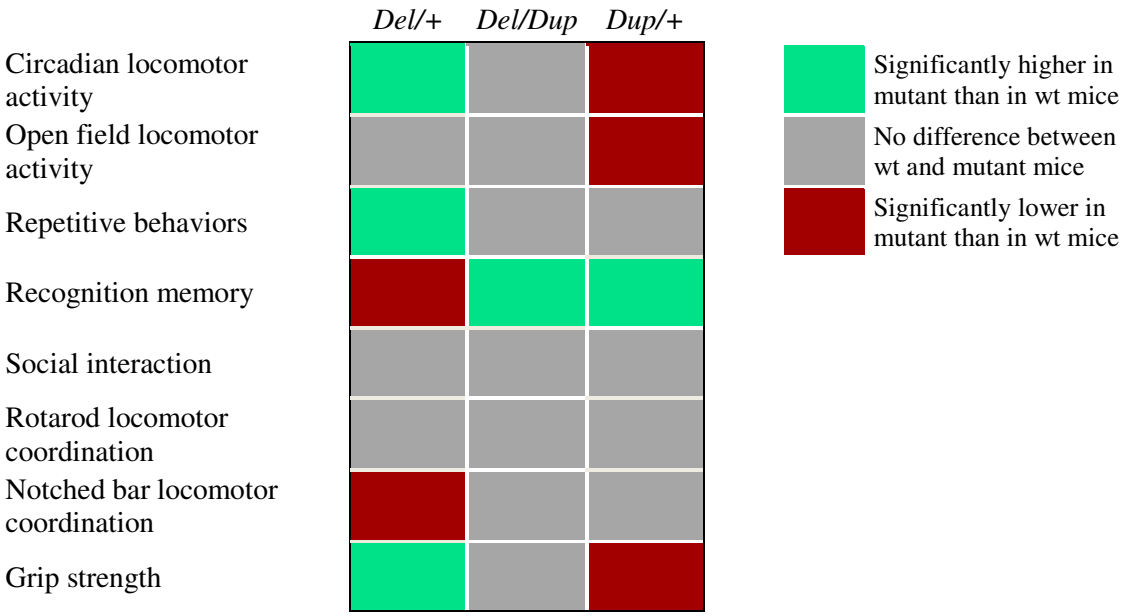


Table 2. Behavioral map of phenotypes observed in mice carrying deletion and/or duplication for the *Sult1a1-Spn* genetic interval on a pure B6N genetic background. *Del/+* mice displayed locomotor hyperactivity, recognition memory deficits, motor coordination impairments in notched bar test and grip strength improvements. *Dup/+* mice displayed locomotor hypoactivity, recognition memory improvements and grip strength deficits. Only one behavioral phenotype was observed for *Del/Dup* mice consisting in recognition memory improvements.

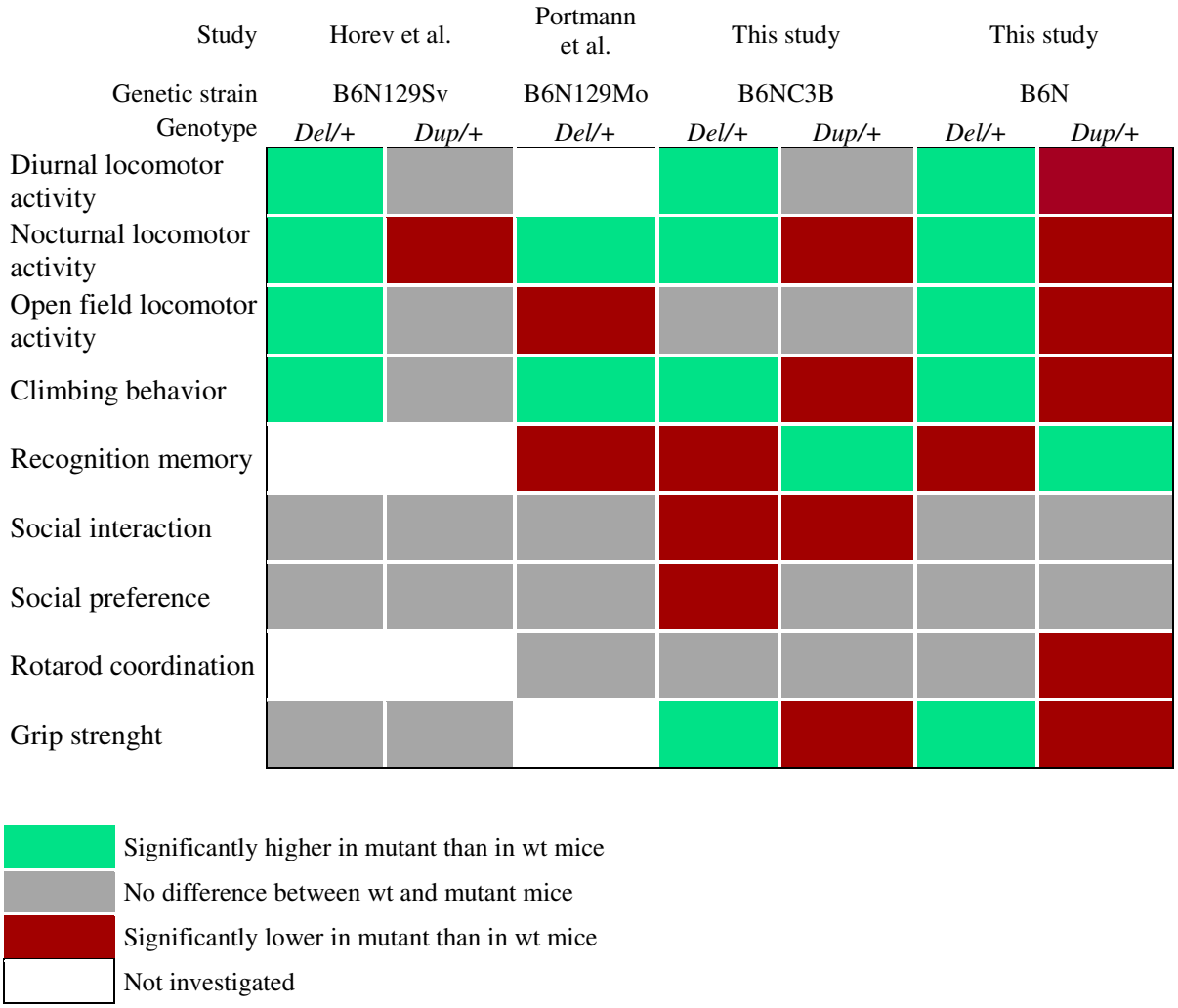


Table 3. Behavioral map of phenotypes observed in mice carrying deletion and duplication for *Slx1b-Sept1* (Horev *et al.*), *Coro1a-Spn* (Portmann *et al.*) and *Sult1a1-Spn* (this study) genetic intervals on mixed and pure genetic backgrounds. Common phenotypes are found for circadian locomotor activity and repetitive behaviors. Recognition memory was not investigated for *Slx1b-Sept1* models. Mice carrying *Coro1a-Spn* deletion displayed an impairment of object recognition memory (1 h retention delay) similar to *Sult1a1-Spn Del/+* mice (30 min and 3 hour retention delays). Changes in open field locomotor activity levels of *Del/+* mice were inconsistent between the different studies. In the present study, mice carrying rearrangements for the *Sult1a1-Spn* region displayed lower level of social interaction specifically on B6N background. *Del/+* and *Dup/+* mice also showed grip strength alteration for both mixed B6NC3B and pure B6N genetic backgrounds whereas coordination impairment was found for *Dup/+* mice only on B6N background.

FIGURES

Figure 1. Mouse models for 16p11.2 rearrangements

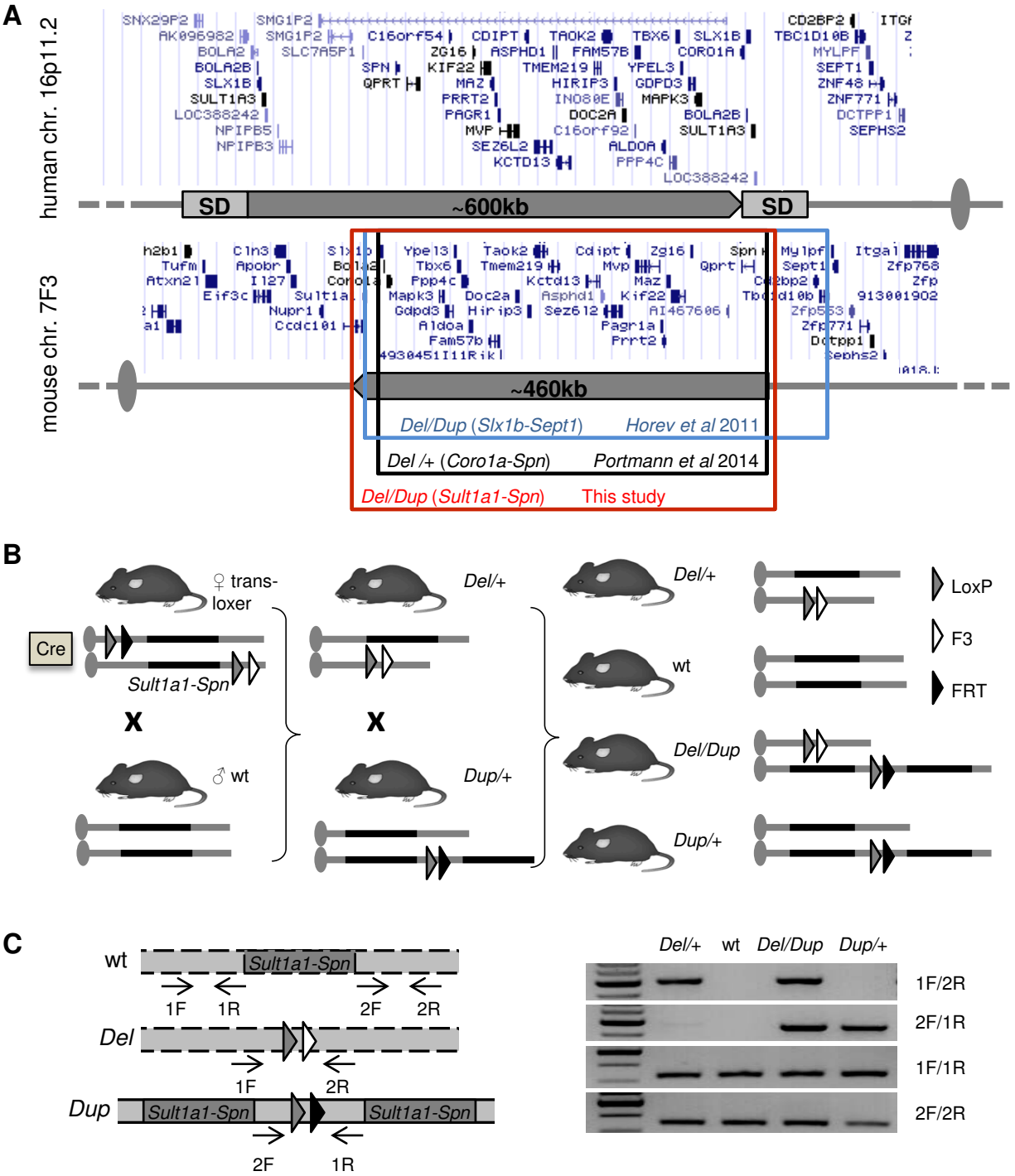


Figure 2. Behavioral characterization of the *Del-Dup* cohort

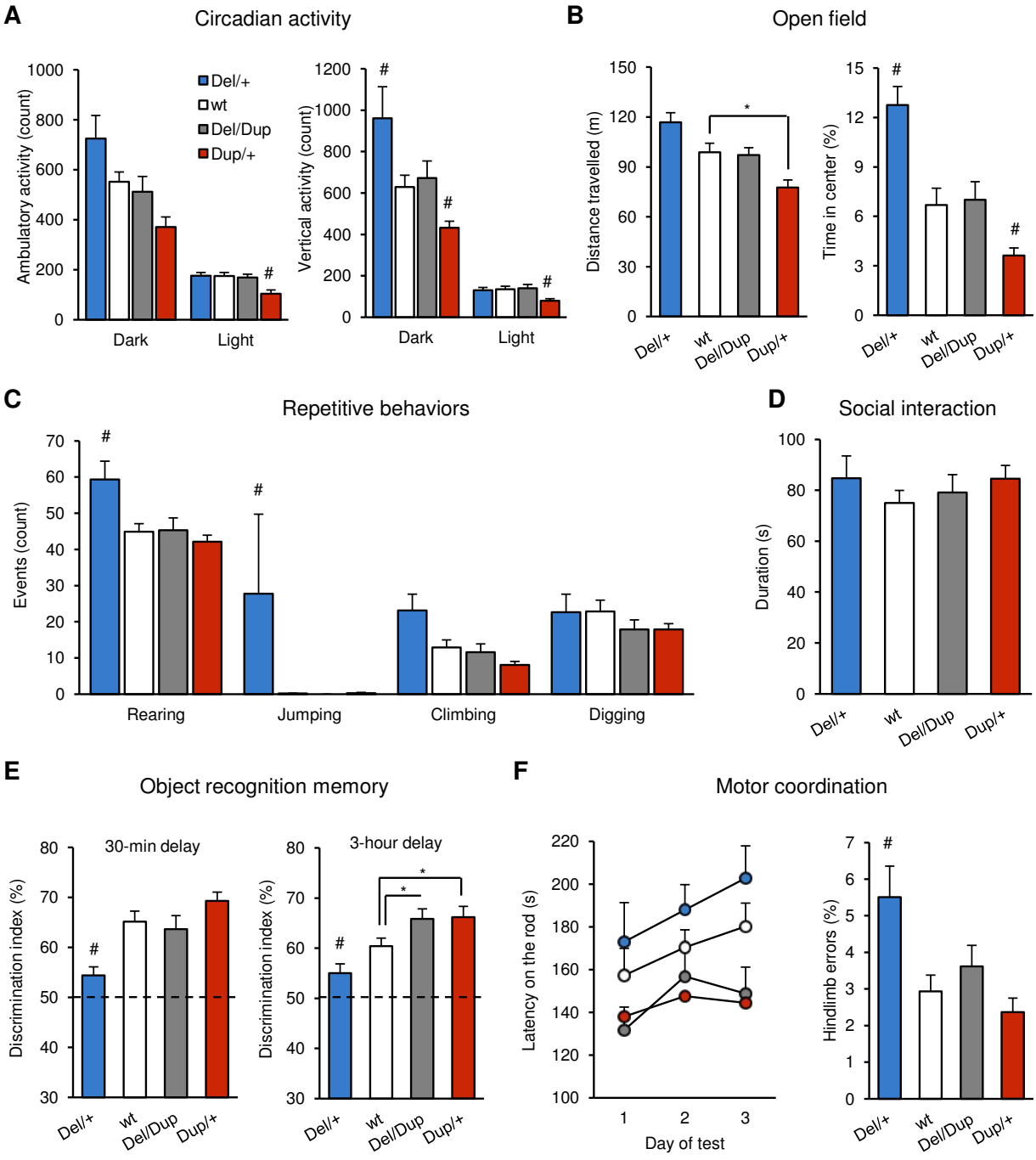


Figure 3. Effects of *Sult1a1–Spn* rearrangements on electrophysiological parameters measured in the CA1 area of hippocampal slices

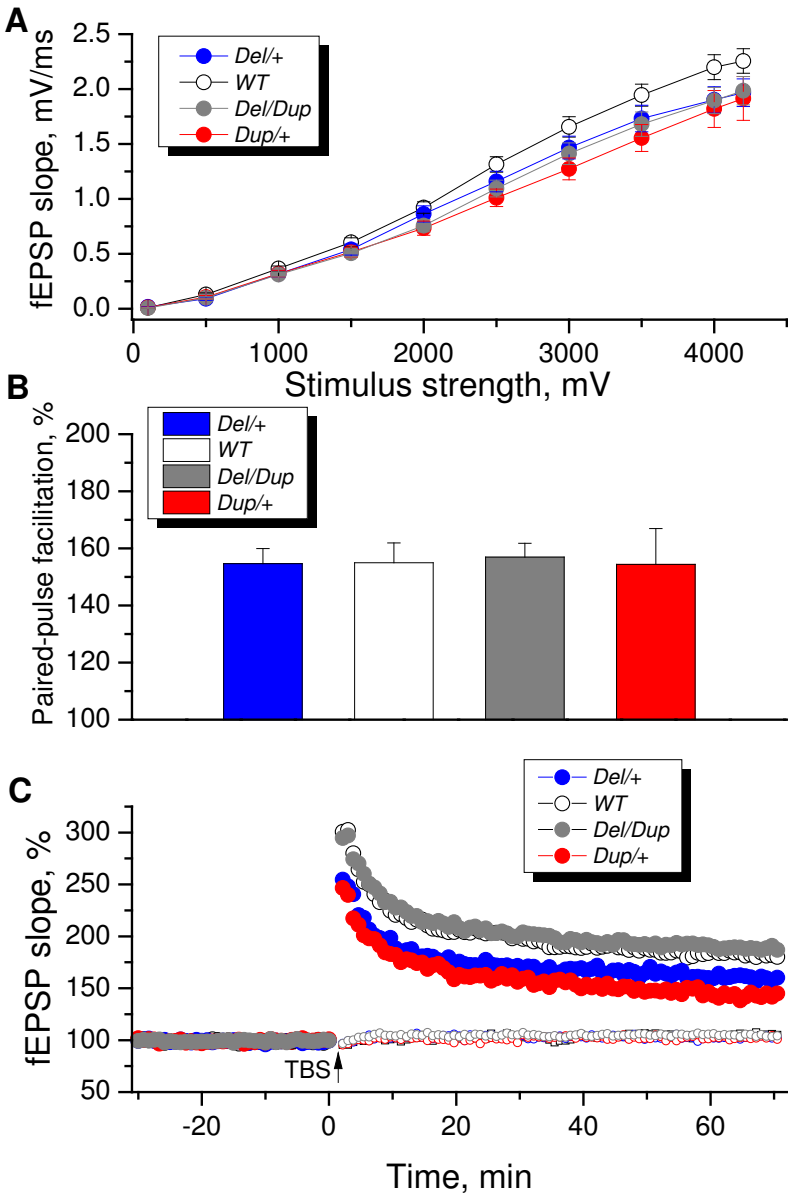


Figure 4. Body mass and size, adiposity and metabolic characterization under standard and high-fat diets

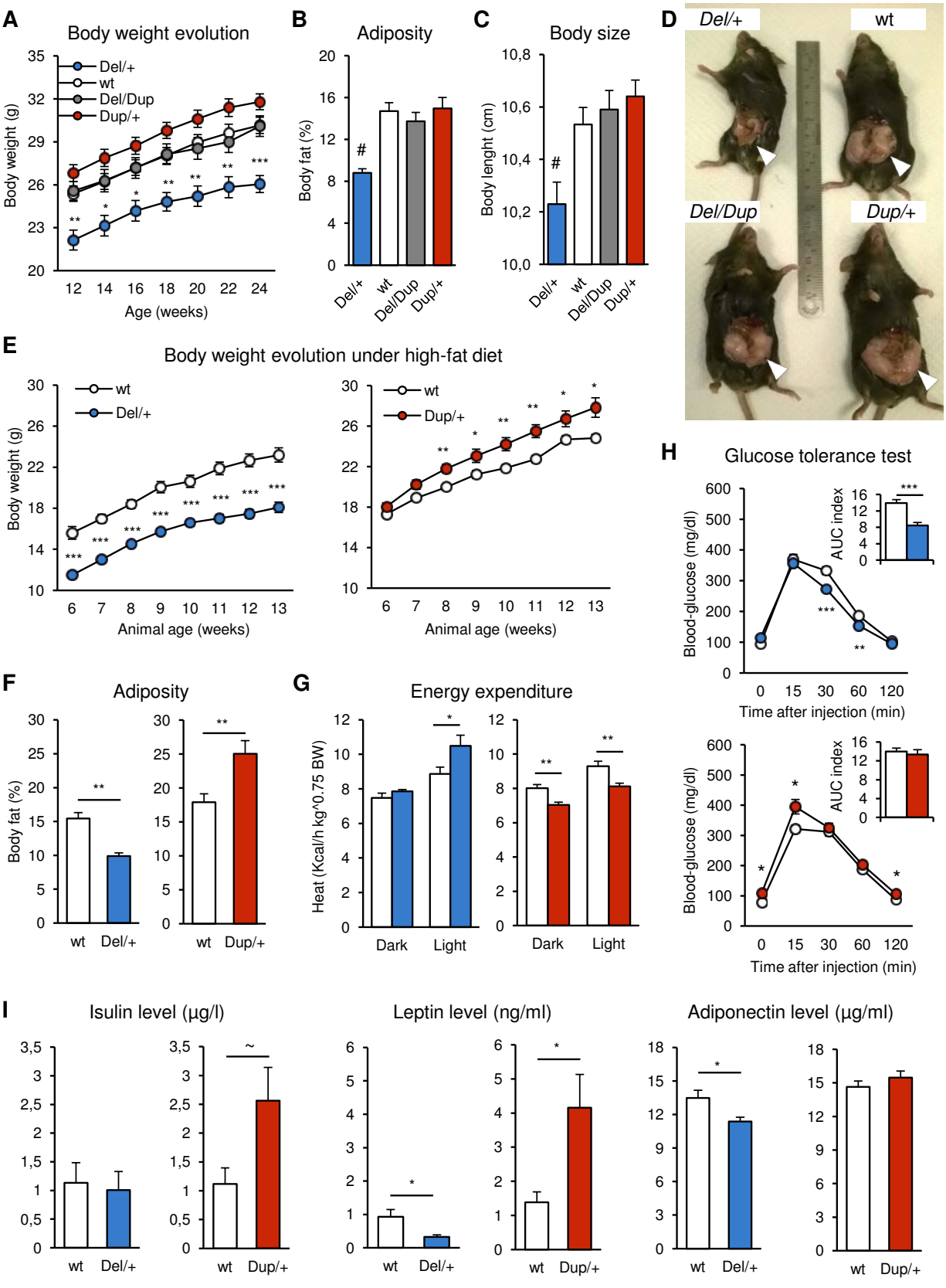


Figure 5. Craniofacial analysis of separate *Del*/+ and *Dup*/+ cohorts

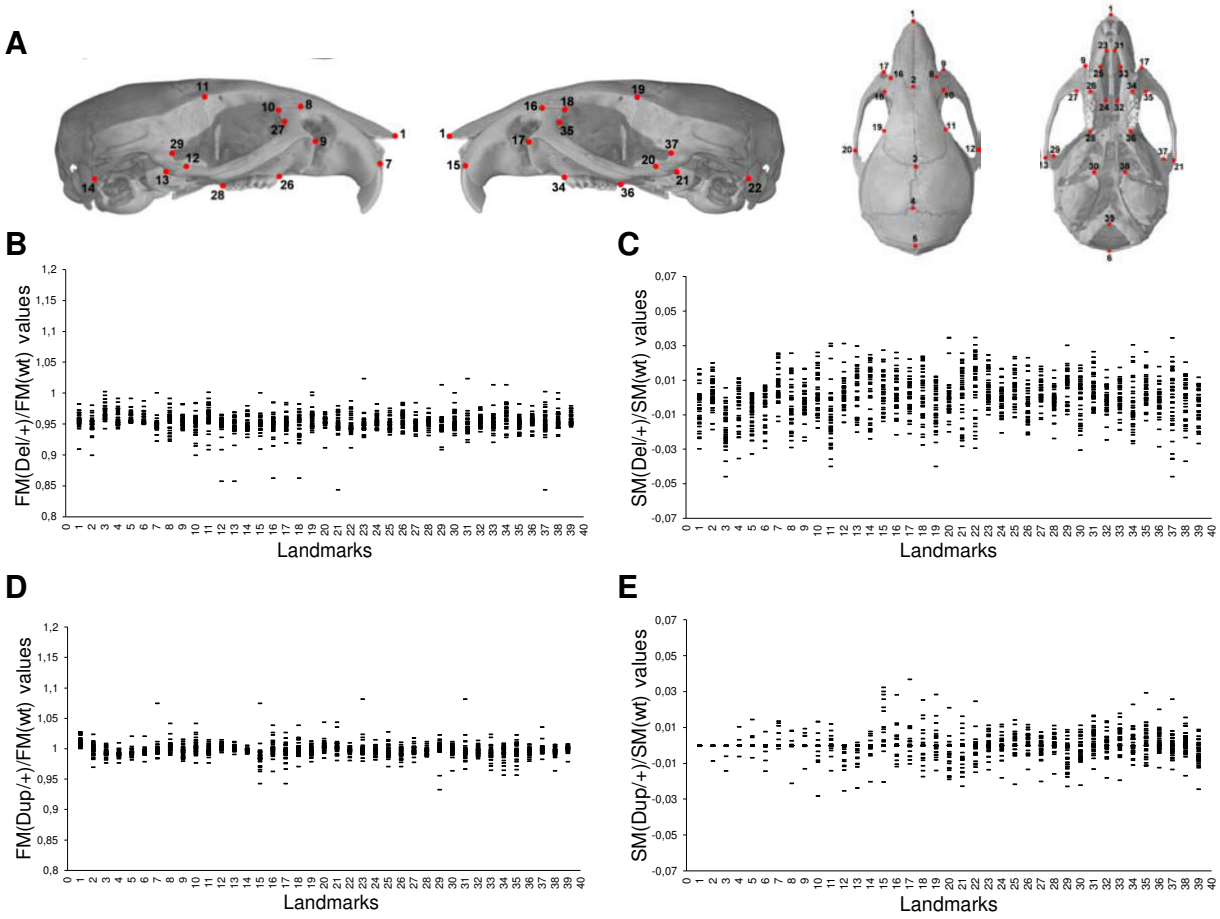
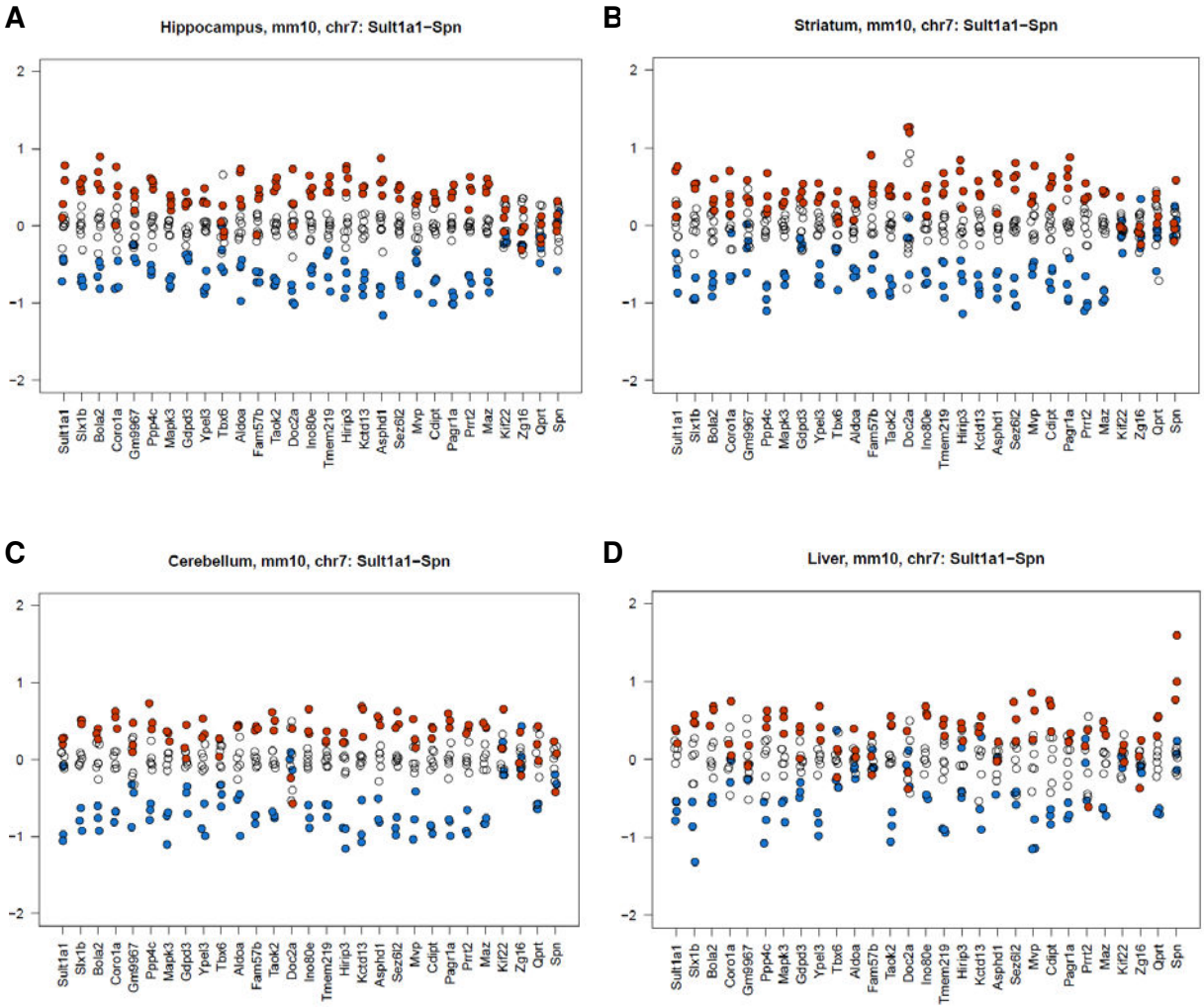
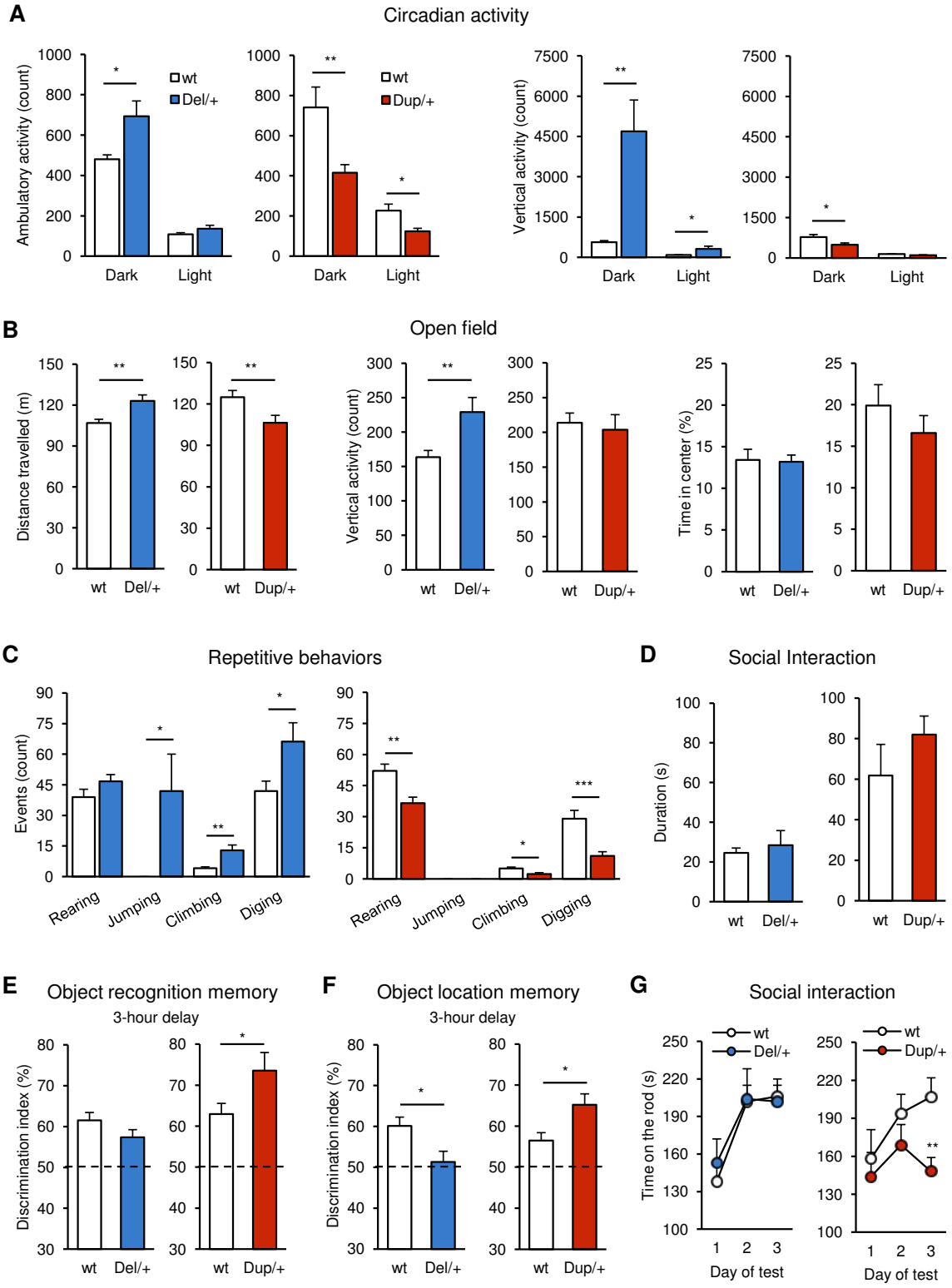


Figure 6. Expression levels of genes within the Sult1a1-Spn region for brain and peripheral tissues

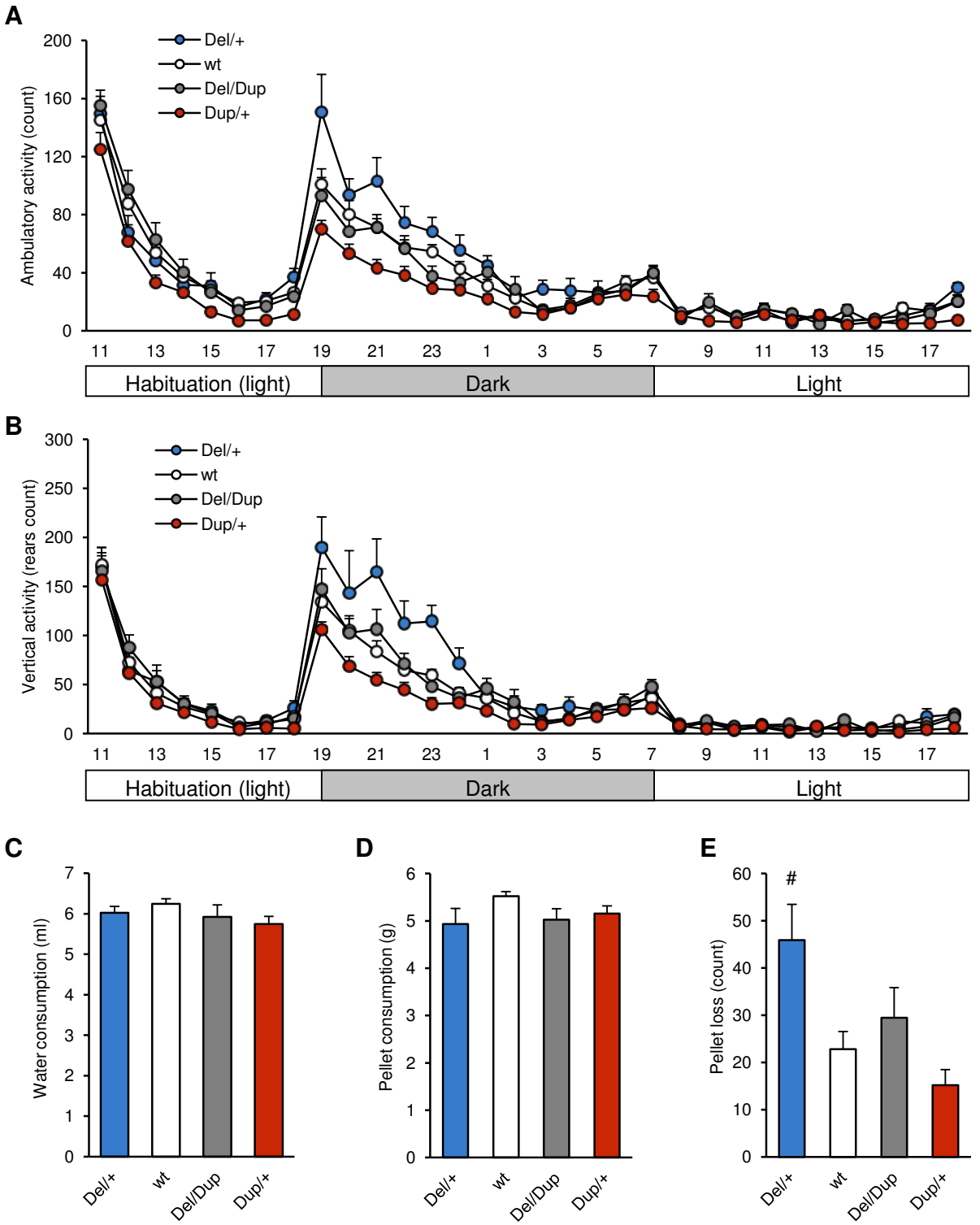


SUPPLEMENTARY INFORMATIONS

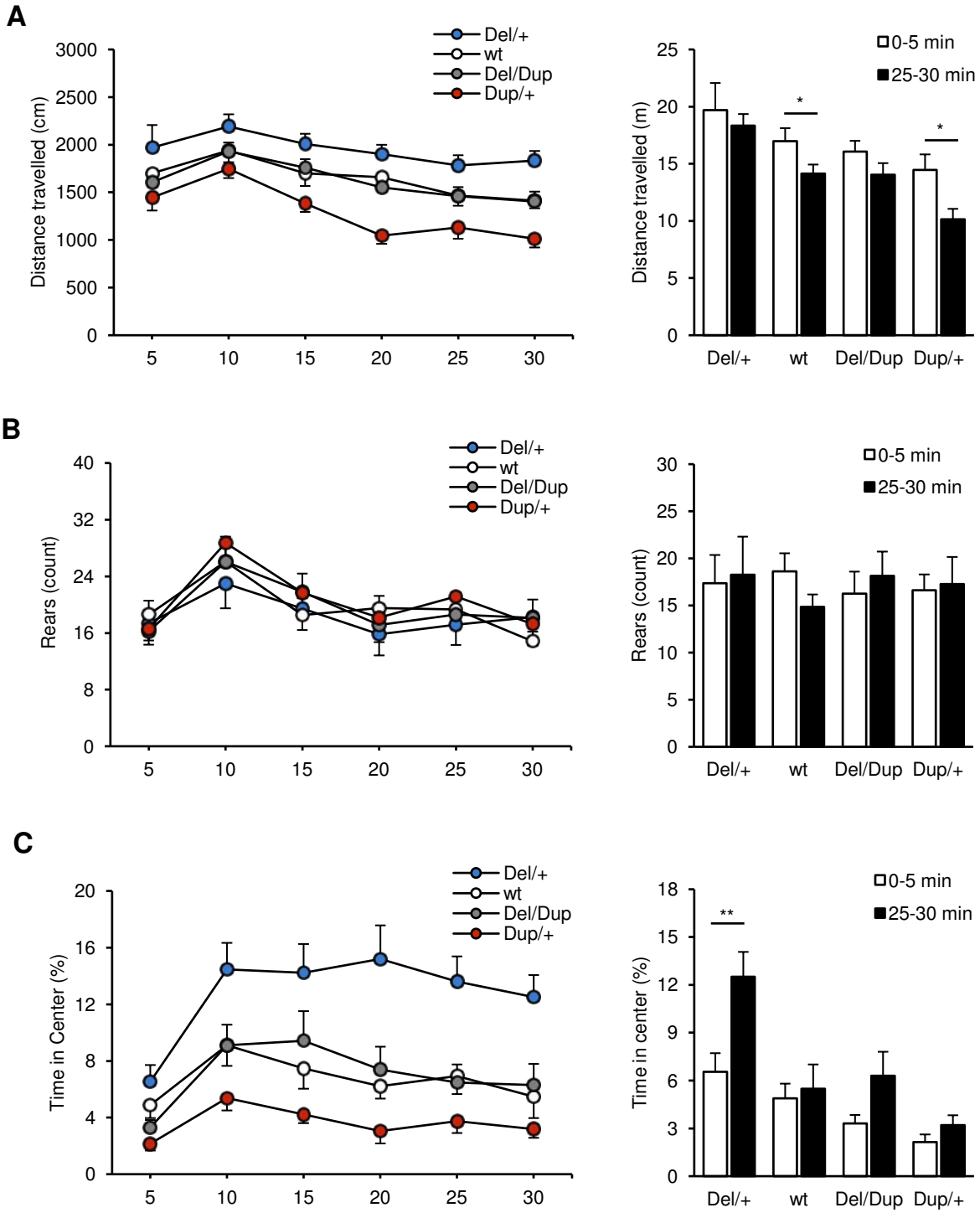
Supplementary Figure 1. Behavioral characterization of *Del/+* and *Dup/+* separated cohorts maintained on a pure B6N genetic background



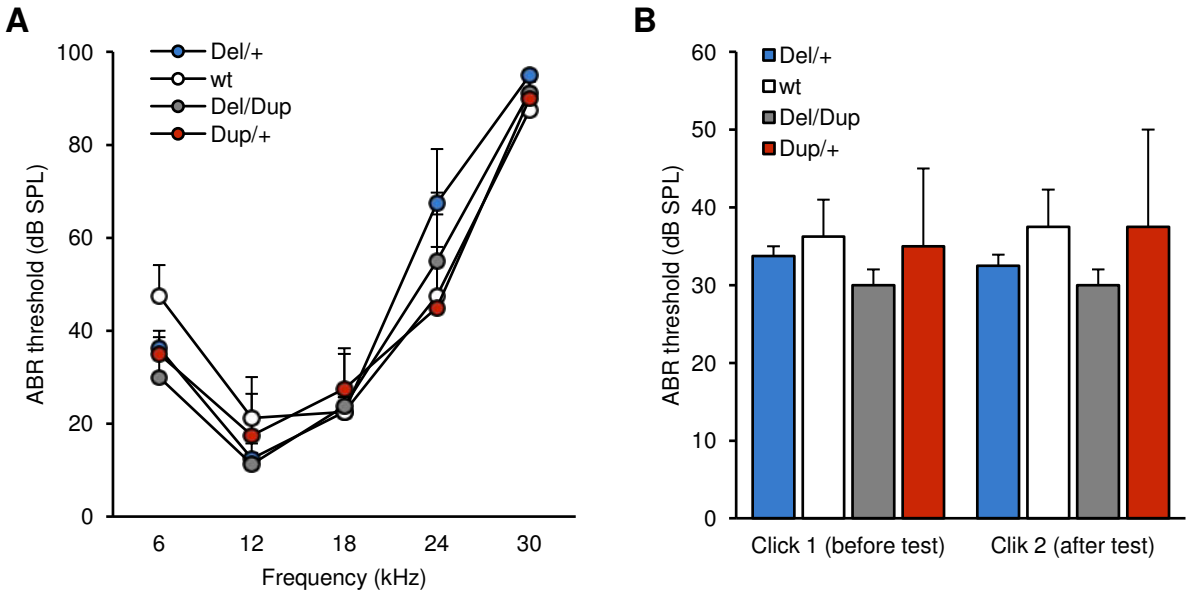
Supplementary Figure 2. Spontaneous locomotor activity and feeding behavior of the *Del-Dup* cohort during the circadian activity test



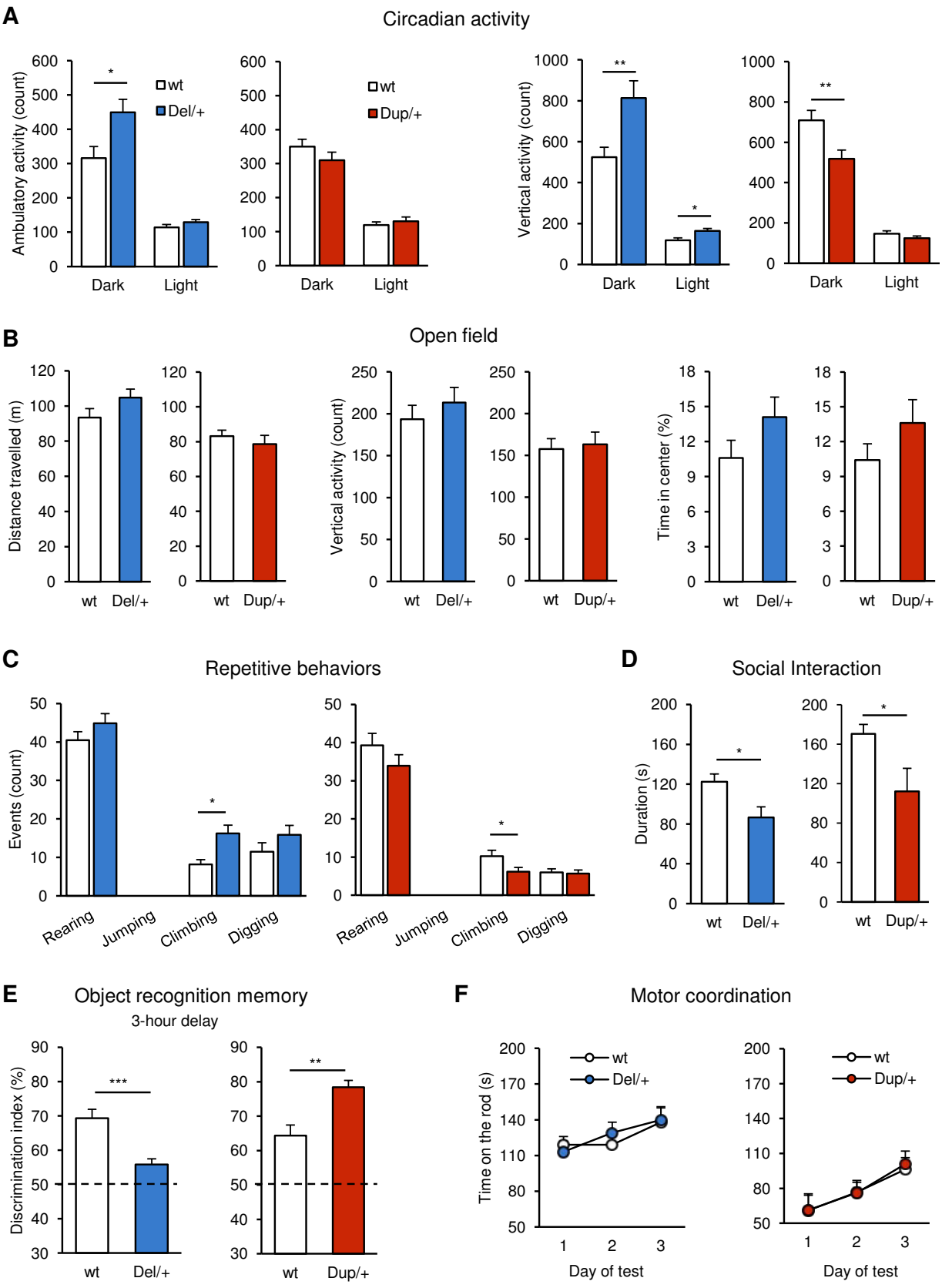
Supplementary Figure 3. Exploratory activity of the *Del-Dup* cohort during the open field test



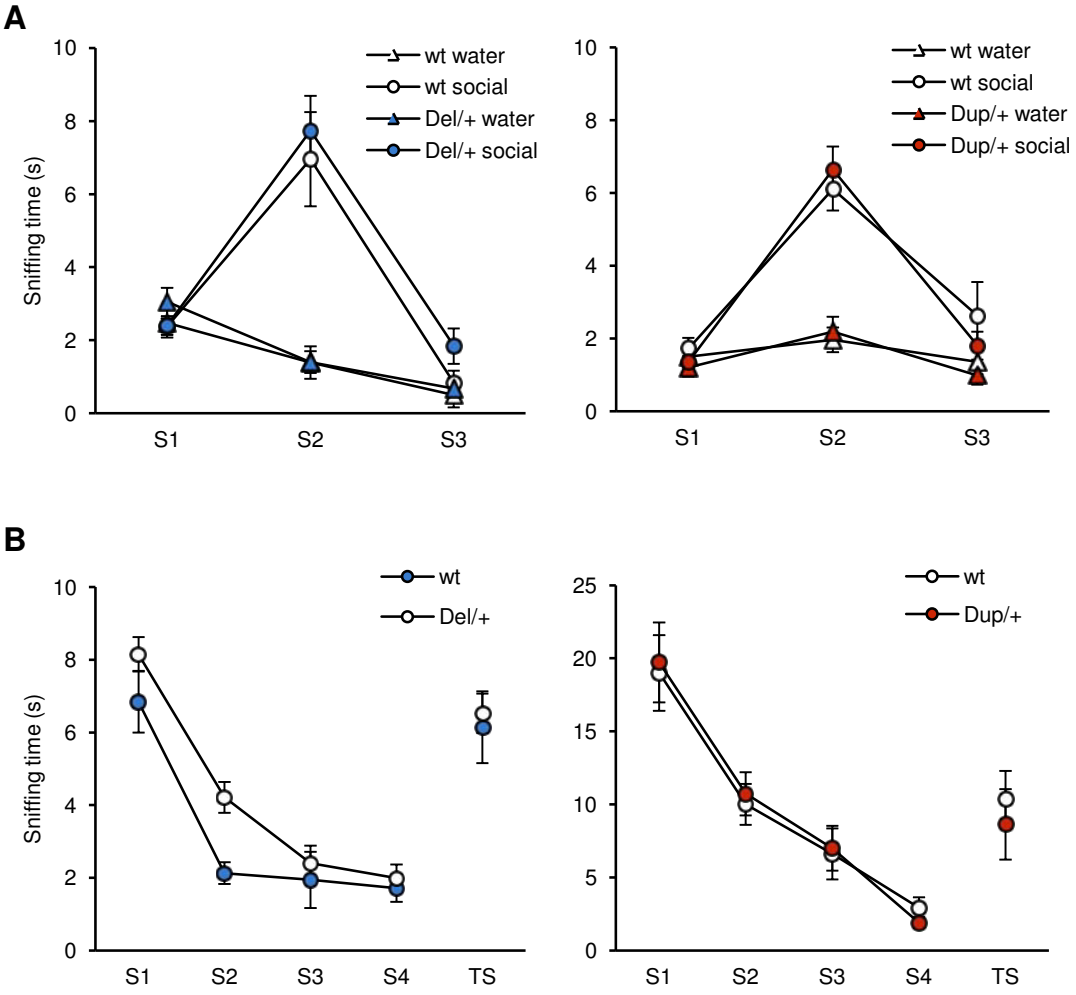
Supplementary Figure 4. Audition capacities of the B6N *Del-Dup* cohort in the auditory brain response experiment



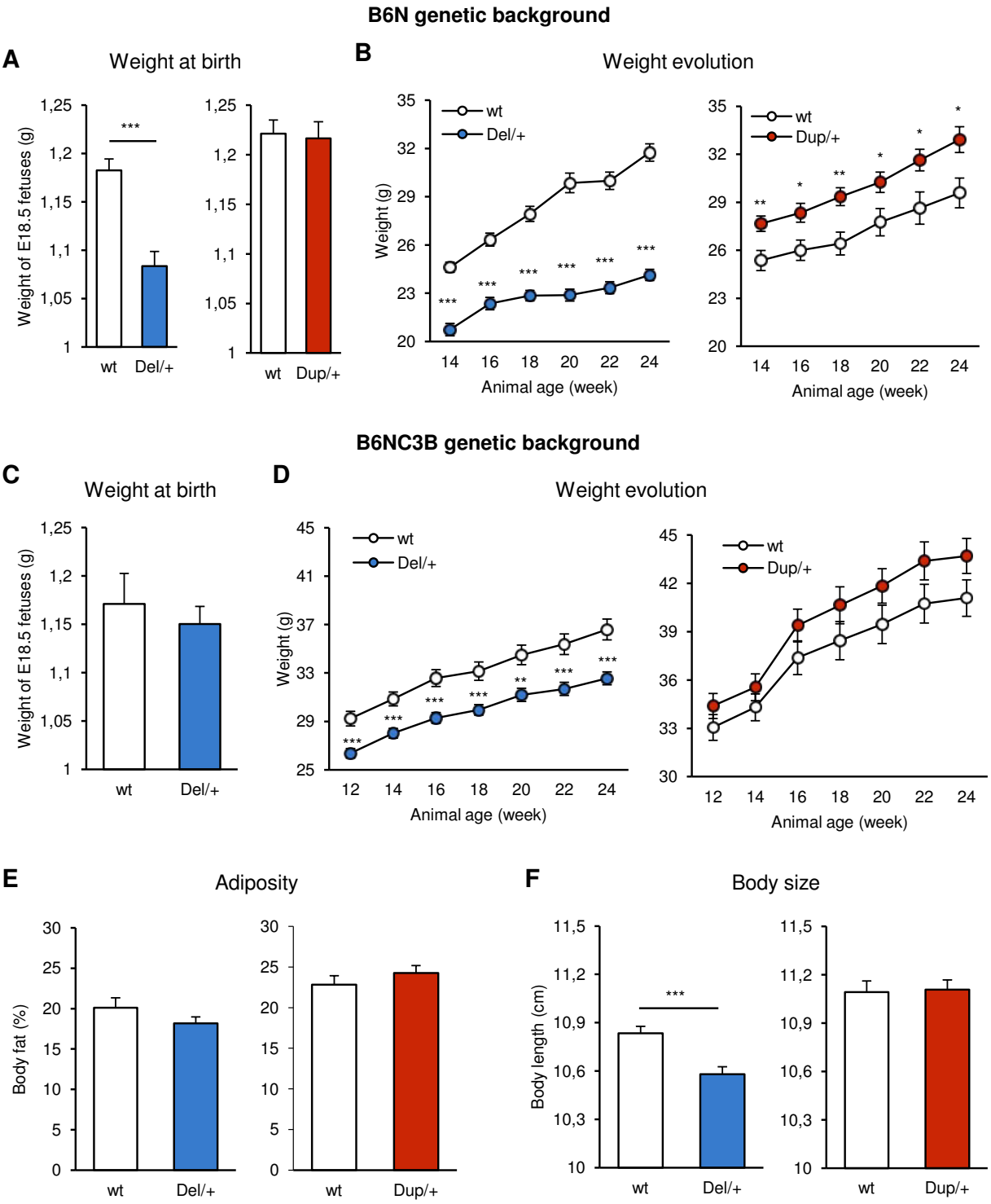
Supplementary Figure 5. Behavioral characterization of *Del/+* and *Dup/+* separated cohorts maintained on a mixed B6NC3B genetic background



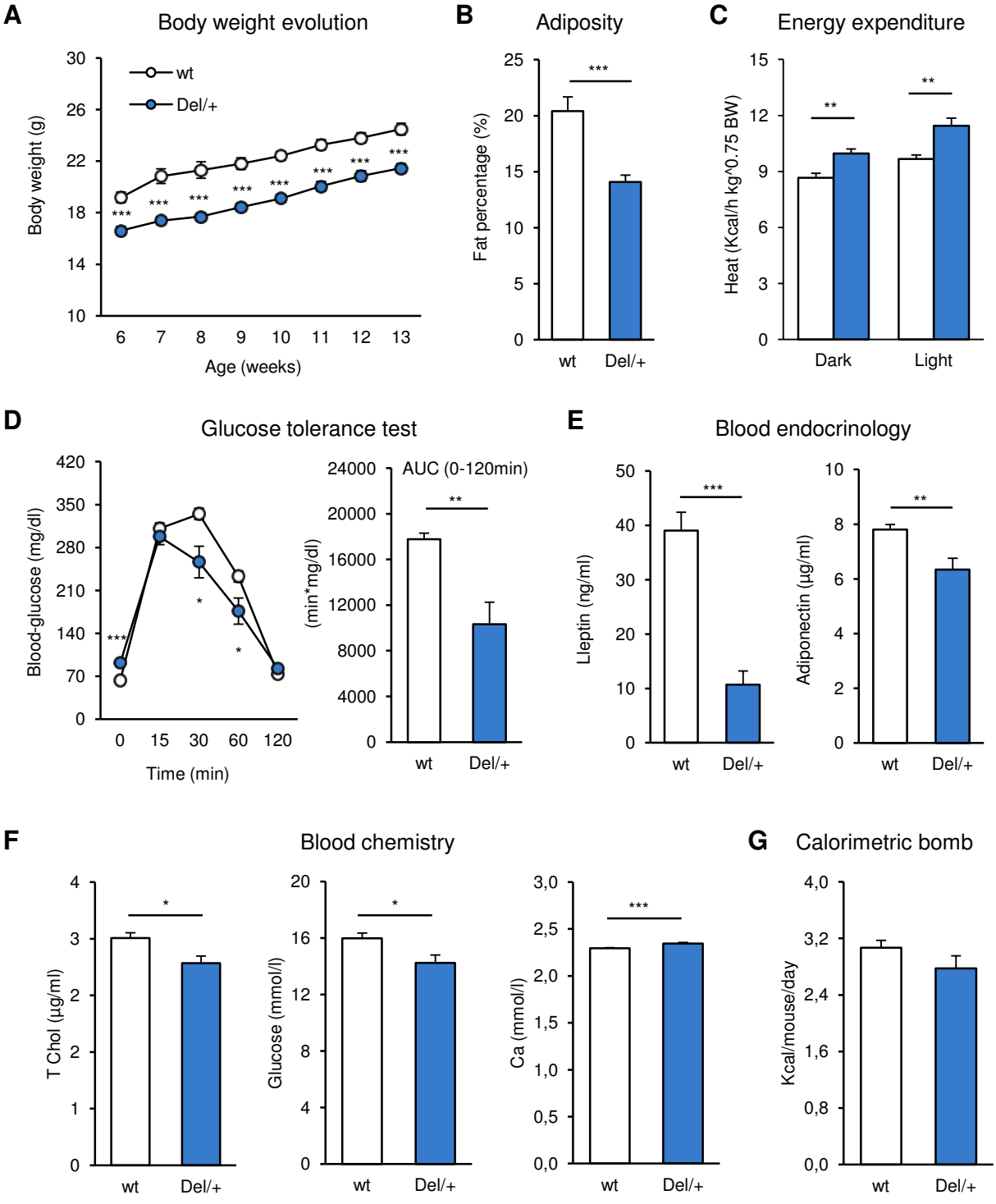
Supplementary Figure 6. Olfaction capacities of *Del*/+ and *Dup*/+ separated cohorts maintained on a mixed B6NC3B genetic background



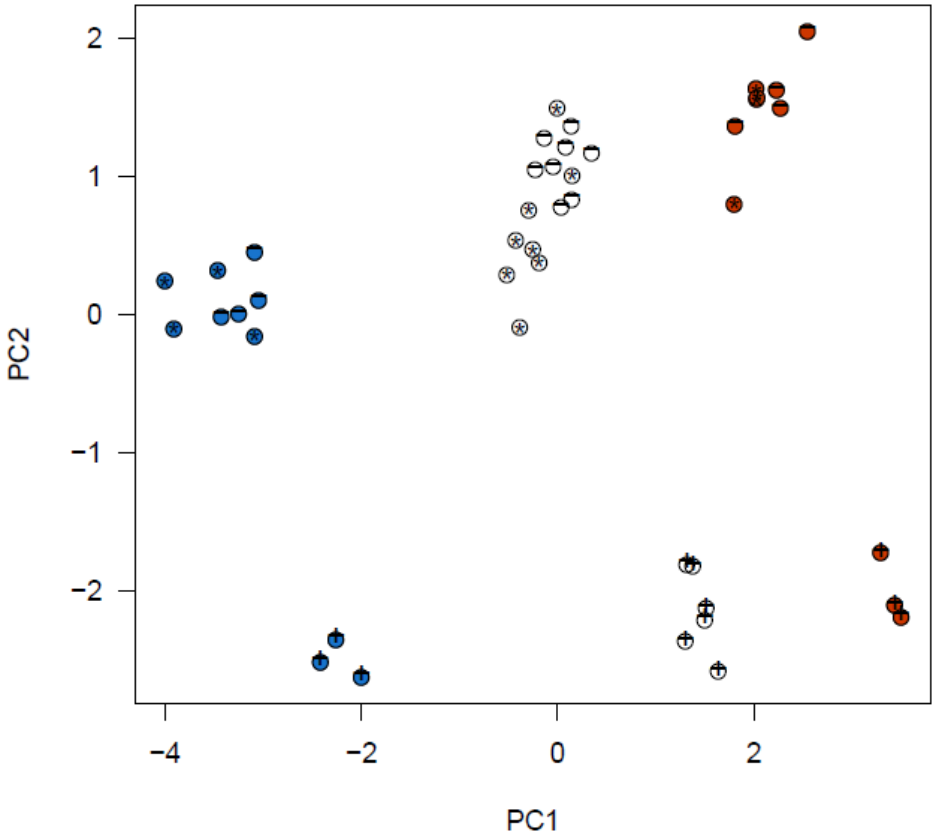
Supplementary Figure 7. Body mass and size and adiposity of *Del/+* and *Dup/+* separated cohorts maintained on B6N and B6NC3B genetic background feed under standard diet



Supplementary Figure 8. Body mass and size, adiposity and metabolic characterization of *Del/+* cohort maintained on B6NC3B genetic background feed under high-fat diet



Supplementary Figure 9. Principal component analysis of the *Sult1a1-Spn* region



Test	Parameter	B6N *Del*/+ cohort results		B6N *Dup*/+ cohort results	
		wt	*Del*/+	wt	*Dup*/+
Circadian Activity	Hab ambulatory activity (count)	301 ± 31	407 ± 46	412 ± 24	272 ± 32 **
	Hab vertical activity (count)	305 ± 50	718 ± 137 **	510 ± 50	304 ± 145 **
	Dark ambulatory activity (count)	481 ± 22	694 ± 76 *	741 ± 103	415 ± 41 **
	Dark vertical activity (count)	568 ± 59	4692 ± 1168 **	772 ± 94	489 ± 69 *
	Light ambulatory activity (count)	108 ± 9	136 ± 18	226 ± 32	124 ± 14 **
	Light vertical activity (count)	87 ± 9	312 ± 106 *	139 ± 13	101 ± 14
	Total food consumption (g)	5.9 ± 0.1	5.8 ± 0.3	5.4 ± 0.2	5.4 ± 0.4
	Total water consumption (ml)	5.3 ± 0.2	5.6 ± 0.2	6.1 ± 0.3	5.8 ± 0.3
Open Field	Locomotor exploration (m)	107 ± 3	123 ± 5 **	125 ± 5	106 ± 5 *
	Rears (count)	164 ± 10	229 ± 21 **	214 ± 14	204 ± 22
	Time in center (%)	13.4 ± 1.3	13.2 ± 0.8	19.9 ± 2.5	16.6 ± 2.1
Elevated Plus Maze	Arm entries (count)	17.1 ± 1.1	17.3 ± 1.2	15.6 ± 1.1	15.1 ± 1.6
	Open arm time (%)	10.1 ± 1.6	13.3 ± 3.0	9.4 ± 1.7	7.6 ± 2.2
Repetitive Behaviors	Digging (count)	20.1 ± 2.6	43.6 ± 6.5 **	29.0 ± 4.1	11.1 ± 1.9 ***
	Climbing (count)	4.1 ± 0.7	12.9 ± 2.6 **	5.0 ± 0.7	2.3 ± 0.6 *
	Rearing (count)	39.0 ± 3.8	46.7 ± 3.3	52.1 ± 3.2	36.5 ± 3.0 **
	Jumping (count)	0	42.0 ± 18.0 *	0	0

Supplementary Table 1. Behavioral characterization of *Del*/+ and *Dup*/+ cohorts on B6N genetic background. Activity was first assessed with the circadian activity test. *Del*/+ mice showed global hyperactivity whereas *Dup*/+ mice showed hypoactivity during habituation (hab), dark and light phases. A considerable increase of vertical activity was observed for *Del*/+ mice during the dark phase. No feeding behavior phenotype was noticed during the test. Exploratory activity was assessed with the open field test. In comparison with wild-type littermates, *Del*/+ and *Dup*/+ mice displayed respectively higher and lower exploratory activity during the 30 min of test. Rearing behavior was also increased for *Del*/+ mice. Time spent in the scared arena center was similar between mutants and controls. The elevated plus maze did not reveal any phenotype. Observation of repetitive behaviors in cages during 10 min showed more climbing, rearing, and jumping in *Del*/+ mice and less digging, climbing, and rearing for *Dup*/+ mice. Overall, *Del*/+ and *Dup*/+ mice showed spontaneous and exploratory hyperactivity and hypoactivity respectively. No sign of anxiety was detected in mutant mice. Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Test	Parameter	B6N *Del*/+ cohort results		B6N *Dup*/+ cohort results	
		wt	*Del*/+	wt	*Dup*/+
Y Maze	Arm entries (count)	16.2 ± 1.8	18.7 ± 1.5	21.2 ± 2.2	18.8 ± 2.7
	Alternation (%)	68.0 ± 3.5	65.6 ± 3.7	58.3 ± 4.1	64.3 ± 4.0
Novel Object Recognition 3 hour delay	S1 object A exploration (s)	6.9 ± 0.6	7.2 ± 0.4	7.3 ± 0.8	7.1 ± 0.5
	S2 object A exploration (s)	2.8 ± 0.4	4.9 ± 0.9 *	2.2 ± 0.4	1.8 ± 0.5
	S2 object B exploration (s)	4.4 ± 0.5	6.6 ± 1.3	3.6 ± 0.6	4.1 ± 0.7
	Discrimination index (%)	61.5 ± 1.9	57.4 ± 1.9	63.0 ± 2.6	73.5 ± 4.4 *
Novel Location Recognition 3 hour delay	S1 object A exploration (s)	3.9 ± 0.6	3.9 ± 0.5	5.8 ± 0.5	5.2 ± 0.6
	S1 object B exploration (s)	4.2 ± 0.8	3.9 ± 0.5	6.5 ± 0.7	5.4 ± 0.7
	S2 non-displaced object explo (s)	2.2 ± 0.5	3.7 ± 0.9	3.8 ± 0.5	2.0 ± 0.3 **
	S2 displaced object explo. (s)	3.1 ± 0.5	3.5 ± 0.4	5.0 ± 0.7	3.7 ± 0.6
	Discrimination index (%)	60.1 ± 2.2	51.3 ± 2.6 *	56.5 ± 1.9	65.2 ± 2.7 *
Morris watermaze	D6 distance to platform (m)	6.7 ± 1.1	6.3 ± 0.9	6.3 ± 1.1	8.8 ± 2.1
	D6 latency to platform (s)	49.7 ± 8.1	42.0 ± 8.0	42.6 ± 10.2	43.6 ± 9.4
	Probe - time in target quad. (%)	44.8 ± 8.4	42.2 ± 9.6	41.3 ± 7.9	42.0 ± 5.3
Social Interaction	Sniffing time (s)	21.1 ± 3.4	18.4 ± 3.7	50.3 ± 8.9	67.4 ± 6.5
	Following time (s)	3.4 ± 1.6	7.6 ± 3.9	11.5 ± 7.3	7.3 ± 3.7
Three-Chamber Sociability	S1 First animal exploration (s)	139 ± 13	129 ± 9	124 ± 13	113 ± 11
	S2 Former animal exploration (s)	76.0 ± 7.8	81.9 ± 8.7	43.1 ± 4.1	35.8 ± 3.0
	S2 Novel animal exploration (s)	110 ± 11	112 ± 11	75.4 ± 11.2	62.2 ± 7.8
	Discrimination index (%)	58.5 ± 3.6	58.1 ± 4.2	61.9 ± 2.3	62.2 ± 2.8
Forced Swim	Floating time (%)	74.5 ± 1.2	57.6 ± 5.7 **	75.9 ± 1.0	75.5 ± 1.9
Prepulse Inhibition (PPI)	120dB/40ms startle reflex	113 ± 15	135 ± 11	136 ± 16	110 ± 20
	Global prepulse inhibition (%)	68.0 ± 2.3	68.3 ± 1.7	74.9 ± 1.8	69.2 ± 4.1
PTZ sensitivity	Clonic seizure latency (s)	227 ± 33	255 ± 61	411 ± 77	368 ± 15
	Clonic seizure duration (s)	48.6 ± 8.0	65.6 ± 15.5	30.7 ± 5.6	55.3 ± 13.2
Rotarod	D1 Time on the rod (s)	138 ± 16	153 ± 19	158 ± 23	144 ± 19
	D2 Time on the rod (s)	202 ± 13	204 ± 24	194 ± 15	169 ± 16
	D3 Time on the rod (s)	206 ± 14	202 ± 13	207 ± 15	148 ± 11 **
Grip Test	Grip stength (g/g body weight)	7.7 ± 0.2	9.7 ± 0.3 ***	7.2 ± 0.4	5.8 ± 0.3 *

Supplementary Table 2. Behavioral characterization (following table) of *Del*/+ and *Dup*/+ cohorts on B6N genetic background. Short-term working memory was first assessed with the Y maze test which did not reveal phenotype in mutant mice. Recognition memory was evaluated with novel object and object location recognition tasks. In the first session (S1) of test, no difference in object exploration was noticed. After a 3-hour retention delay, *Del*/+ mice showed a deficit for displaced object discrimination whereas *Dup*/+ mice showed improvements for novel object and displaced object discriminations. In the Morris water maze, no change of spatial learning and memory was observed. Mutant and control mice

travelled the same distance and needed the same duration to find the platform from the first day to the sixth and last day (D6) of acquisition. In the probe test (PT) on day 7, when the platform is removed, all mice spent the same percentage of time in the target quadrant. We did not observe any phenotype which can be related to neuropsychiatric human symptoms. No phenotype was noticed in social interaction and three-chamber sociability tests. No sign of depression was seen in the forced swim test. On the contrary, hyperactive *Del/+* floated less than wild-type mice. Prepulse inhibition (PPI) test did not reveal any schizophrenic-like behavior and pentylenetetrazole (PTZ) sensitivity test did not reveal seizure susceptibility. In the rotarod test, *Dup/+* mice were not able to improve their coordination capacities from the first day (D1) to the third day (D3) of test. Finally, grip test indicated that *Del/+* and *Dup/+* mice showed stronger and weaker grip strength respectively. Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Test	Parameter	B6N *Del-Dup* cohort results			
		Del/+	**wt**	***Del/Dup***	***Dup/+***
Circadian Activity	Hab ambulatory activity (count)	367 ± 44	393± 36	414 ± 51	275 ± 32
	Hab vertical activity (count)	722 ± 363	353 ± 40	375 ± 46	292 ± 61
	Dark ambulatory activity (count)	725 ± 92	552 ± 39	512 ± 61	371 ± 40
	Dark vertical activity (count)	2178 ± 1225	629 ± 57	672 ± 82	432 ± 32
	Light ambulatory activity (count)	176 ± 13	175 ± 14	169 ± 13	104 ± 15
	Light vertical activity (count)	128 ± 12	135 ± 15	140 ± 19	79.3 ± 10.0
	Total food consumption (g)	4.9 ± 0.3	5.5 ± 0.1	5.0 ± 0.2	5.2 ± 0.2
	Total water consumption (ml)	6.0 ± 0.2	6.3 ± 0.1	5.9 ± 0.3	5.7 ± 0.2
Open Field	Locomotor exploration (m)	117 ± 6	98.8 ± 5.5	97.1 ± 4.4	79.3 ± 4.5
	Rears (count)	111 ± 15	117 ± 9	118 ± 14	123 ± 13
	Time in center (%)	12.8 ± 1.1	6.7 ± 1.0	7.0 ± 1.1	4.0 ± 0.5
Repetitive Behaviors	Digging (count)	22.7 ± 4.9	22.8 ± 3.1	17.9 ± 2.6	17.9 ± 1.6
	Climbing (count)	23.1 ± 4.5	12.9 ± 2.1	11.6 ± 2.3	8.1 ± 0.9
	Rearing (count)	59.3 ± 5.1	44.9 ± 2.2	45.3 ± 3.4	42.1 ± 1.8
	Jumping (count)	23.3 ± 22.1	0.2 ± 0.1	0	0.3 ± 0.2

Supplementary Table 3. Behavioral characterization of the *Del-Dup* cohort. In the circadian activity test, *Del/+* mice showed vertical hyperactivity during dark phase whereas *Dup/+* mice showed vertical hypoactivity during light and dark phases and ambulatory hypoactivity during light phase. During open field sessions, *Dup/+* mice displayed a lower exploration activity and spent less time in arena center whereas *Del/+* mice spent more time in the arena center. Observation of repetitive behaviors revealed more rearing and jumping in *Del/+* mice. Data are mean ± SEM. Statistical values are resumed in Supplementary Table 5.

Test	Parameter	B6N *Del-Dup* cohort results			
		Del/+	**wt**	***Del/Dup***	***Dup/+***
New Object Recognition 30 min delay	S1 object A exploration (s)	6.7 ± 0.7	6.8 ± 0.7	6.6 ± 0.8	6.5 ± 0.4
	S2 object A exploration (s)	5.4 ± 0.7	4.4 ± 0.6	4.7 ± 0.6	3.7 ± 0.6
	S2 object B exploration (s)	6.3 ± 0.8	8.0 ± 0.8	8.3 ± 1.0	7.7 ± 1.1
	Discrimination index (%)	54.4 ± 1.7	65.1 ± 2.1	63.7 ± 2.7	69.3 ± 1.7
New Object Recognition 3 hour delay	S1 object A exploration (s)	10.7 ± 1.1	12.1 ± 0.6	12.3 ± 1.3	10.2 ± 1.0
	S2 object A exploration (s)	6.8 ± 1.4	6.3 ± 1.0	4.6 ± 0.6	3.6 ± 0.4
	S2 object B exploration (s)	8.4 ± 1.9	9.2 ± 1.1	8.6 ± 0.8	8.0 ± 1.3
	Discrimination index (%)	55.0 ± 1.8	60.4 ± 1.6	65.8 ± 2.0	66.2 ± 2.2
Sucrose Preference	D1 sucrose preference (%)	65.7 ± 4.4	70.8 ± 4.1	71.3 ± 3.6	74.6 ± 3.6
	D2 sucrose preference (%)	68.7 ± 5.6	76.6 ± 4.0	71.7 ± 2.7	71.5 ± 4.7
	D3 sucrose preference (%)	81.5 ± 2.6	78.2 ± 4.1	76.1 ± 2.1	71.7 ± 2.7
Social Interaction	Sniffing time (s)	73.8 ± 7.9	60.4 ± 4.4	61.0 ± 3.6	73.9 ± 7.5
	Following time (s)	18.2 ± 3.9	11.4 ± 2.8	18.1 ± 5.3	16.0 ± 4.7
Rotarod	D1 Time on the rod (s)	173 ± 18	157 ± 13	132 ± 11	138 ± 14
	D2 Time on the rod (s)	188 ± 12	170 ± 8	157 ± 13	147 ± 12
	D3 Time on the rod (s)	203 ± 15	180 ± 11	149 ± 12	144 ± 14
Notched Bar	Hind limb errors (%)	5.5 ± 0.8	2.9 ± 0.4	3.6 ± 0.6	2.4 ± 0.4
Grip Test	Grip stength (g/g body weight)	10.8 ± 0.2	9.3 ± 0.2	9.2 ± 0.4	7.6 ± 0.2

Supplementary Table 4. Behavioral characterization (following table) of the *Del-Dup* cohort. Recognition memory of mice was assessed in the NOR test with two retention delays. *Del/+* mice showed recognition memory deficits for 30-min and 3-hours retention delays. *Dup/+* and *Del/Dup* mice showed improvements for a specific retention delay of 3 hour. No anhedonia phenotype was observed in the sucrose preference test. No phenotype was noticed in the social interaction test. During rotarod test, *Del/+* mice showed trends for locomotor coordination improvement whereas *Dup/+* and *Del/Dup* mice showed similar trends for locomotor coordination impairments trends. *Del/+* animals also showed impairments in the notched bar test. *Del/+* and *Dup/+* mice showed respectively stronger and weaker grip strength in the grip test. Data are mean ± SEM. Statistical values are resumed in Supplementary Table 5.

Circadian activity				
	Del/+ vs wt	*Dup*/+ vs wt	*Del/Dup* vs wt	*Del*/+ vs *Dup*/+
Dark horizontal activity	q=3.108, p=0.138	q=3.609, p=0.064	q=0.763, p=0.949	q=6.002, p<0.001
Dark vertical activity	U=43, p=0.039	U=66, p=0.015	U=105, p=0.730	U=5, p<0.001
Light horizontal activity	q=0.122, p=0.997	q=5.227, p=0.003	q=0.424, p=0.991	q=4.562, p=0.012
Light vertical activity	U=85, p=1.000	U=59.5, p=0.008	U=109.5, p=0.871	U=20.5, p=0.008
Open field				
	Del/+ vs wt	*Dup*/+ vs wt	*Del/Dup* vs wt	*Del*/+ vs *Dup*/+
Distance travelled	q=3.388, p=0.089	q=4.474, p=0.013	q=0.341, p=0.995	q=7.178, p<0.001
Time in center	U=29, p=0.001	U=92, p=0.029	U=140, p=0.945	U=2.5, p=0.004
Repetitive behaviors				
	Del/+ vs wt	*Dup*/+ vs wt	*Del/Dup* vs wt	*Del*/+ vs *Dup*/+
Climbing	U=27, p=0.059	U=53, p=0.074	U=47, p=0.642	U=26.5, p=0.015
Rearing	q=4.682, p=0.010	q=1.029, p=0.886	q=0.135, p=1.000	q=5.842, p=0.001
Jumping	U=24, p=0.016	U=88, p=0.906	U=45, p=0.235	U=30, p=0.008
NOR test (30 min delay)				
	Del/+ vs wt	*Dup*/+ vs wt	*Del/Dup* vs wt	*Del*/+ vs *Dup*/+
Discrimination index	q=5.054, p=0.004	q=2.107, p=0.451	q=0.711, p=0.958	q=6.826, p<0.001
NOR test (3 hour delay)				
	Del/+ vs wt	*Dup*/+ vs wt	*Del/Dup* vs wt	*Del*/+ vs *Dup*/+
Discrimination index	U=38, p=0.043	U=66, p=0.021	U=54, p=0.048	U=22, p=0.002
Rotarod				
	Del/+ vs wt	*Dup*/+ vs wt	*Del/Dup* vs wt	*Del*/+ vs *Dup*/+
D1 Time on the rod	q=1.128, p=0.855	q=1.656, p=0.647	q=2.088, p=0.456	q=2.878, p=0.182
D2 Time on the rod	q=1.275, p=0.804	q=1.950, p=0.516	q=1.101, p=0.864	q=2.923, p=0.172
D3 Time on the rod	q=1.631, p=0.658	q=3.065, p=0.140	q=2.563, p=0.274	q=4.221, p=0.019
Notched bar				
	Del/+ vs wt	*Dup*/+ vs wt	*Del/Dup* vs wt	*Del*/+ vs *Dup*/+
Hind limb errors (%)	U=41.5, p=0.007	U=144.5, p=0.298	U=113, p=0.313	U=29, p=0.001
Grip test				
	Del/+ vs wt	*Dup*/+ vs wt	*Del/Dup* vs wt	*Del*/+ vs *Dup*/+
Grip strength	q=5.447, p=0.002	q=7.175, p<0.001	q=0.321, p=0.996	q=11.591, p<0.001

Supplementary Table 5. Statistical values for *Del-Dup* behavioral characterization. The table summarizes two genotype comparisons with parametric Tukey's test (q and p values) or non-parametric Mann-Whitney *U* test (U and p values) when analysis of variance revealed significant differences between genotypes.

Test	Parameter	B6C3B *Del*/+ cohort results		B6C3B *Dup*/+ cohort results	
		wt	*Del*/+	wt	*Dup*/+
Circadian Activity	Hab ambulatory activity (count)	206 ± 20	220± 16	170 ± 13	172 ± 14
	Hab vertical activity (count)	331 ± 57	393 ± 45	319 ± 28	300 ± 34
	Dark ambulatory activity (count)	316 ± 34	449 ± 38 *	350 ± 21	310 ± 24
	Dark vertical activity (count)	524 ± 49	814 ± 84 **	709 ± 49	518 ± 43 **
	Light ambulatory activity (count)	114 ± 8	129 ± 7	120 ± 9	131 ± 12
	Light vertical activity (count)	118 ± 12	164 ± 11 **	147 ± 14	123 ± 11
	Total food consumption (g)	5.1 ± 0.1	4.8 ± 0.1	5.2 ± 0.3	5.6 ± 0.3
	Total water consumption (ml)	5.8 ± 0.2	6.1 ± 0.2	5.8 ± 0.2	6.1 ± 0.2
Open Field	Locomotor exploration (m)	93.4 ± 5.1	105 ± 5	83.2 ± 3.4	79.6 ± 5.0
	Rears (count)	194 ± 17	213 ± 18	158 ± 12	163 ± 15
	Time in center (%)	10.6 ± 1.5	14.1 ± 1.7	10.4 ± 1.4	13.6 ± 2.0
Repetitive Behaviors	Digging (count)	11.4 ± 2.3	15.9 ± 2.5	6.0 ± 0.9	5.7 ± 0.9
	Climbing (count)	8.2 ± 1.2	16.2 ± 2.2 **	10.3 ± 1.5	6.2 ± 1.1 *
	Rearing (count)	40.5 ± 2.2	44.9 ± 2.5	39.3 ± 3.2	33.9 ± 2.9

Supplementary Table 6. Behavioral characterization of *Del*/+ and *Dup*/+ cohorts on B6C3B genetic background. In the circadian activity test, *Del*/+ mice showed ambulatory hyperactivity during dark phase and vertical hyperactivity during dark and light phases whereas *Dup*/+ mice only showed vertical hypoactivity during dark phase. No exploration activity change was noticed in the open field test. Observation of repetitive behaviors revealed increased and decreased level of climbing for *Del*/+ and *Dup*/+ mice, respectively. Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Test	Parameter	B6C3B *Del*/+ cohort results		B6C3B *Dup*/+ cohort results	
		wt	*Del*/+	wt	*Dup*/+
Y Maze	Arm entries (count)	26.4 ± 1.5	30.1 ± 1.9	29.7 ± 1.4	23.4 ± 1.1 **
	Alternation (%)	65.9 ± 2.2	66.7 ± 2.6	65.3 ± 2.4	73.1 ± 2.2 *
Novel Object Recognition 3 hour delay	S1 object A exploration (s)	10.0 ± 1.1	10.5 ± 1.0	7.1 ± 0.9	5.1 ± 0.7
	S2 object A exploration (s)	4.1 ± 1.1	4.6 ± 0.5	4.2 ± 0.5	1.8 ± 0.6 **
	S2 object B exploration (s)	9.1 ± 2.0	5.8 ± 0.7	7.3 ± 0.6	6.2 ± 0.6
	Discrimination index (%)	69.3 ± 2.6	55.8 ± 1.6 ***	64.4 ± 3.0	78.4 ± 1.9 **
Novel Object Recognition 24 hour delay	S1 object A exploration (s)	NT		14.0 ± 2.0	13.5 ± 2.5
	S2 object A exploration (s)			4.8 ± 0.6	4.9 ± 0.9
	S2 object B exploration (s)			9.7 ± 1.6	12.8 ± 3.1
	Discrimination index (%)			64.0 ± 3.8	68.5 ± 2.7
Novel Location Recognition 3 hour delay	S1 object A exploration (s)	4.6 ± 0.6	5.5 ± 0.6	NT	
	S1 object B exploration (s)	4.3 ± 0.8	5.8 ± 0.7		
	S2 non-displaced object explo (s)	1.7 ± 0.2	2.5 ± 0.4		
	S2 displaced object explo. (s)	2.8 ± 0.5	2.9 ± 0.6		
	Discrimination index (%)	62.8 ± 2.9	51.8 ± 2.4 *		
Social Interaction	Sniffing time (s)	106 ± 9	76 ± 8 *	147 ± 7	99.1 ± 18.9 *
	Following time (s)	16.7 ± 1.8	10.6 ± 3.6	23.8 ± 2.5	12.3 ± 4.3 *
Three-Chamber Sociability	S1 First animal exploration (s)	167 ± 26	171 ± 21	155 ± 20	135 ± 13
	S2 Former animal exploration (s)	35.4 ± 8.4	82.4 ± 10.3 **	64.2 ± 10.1	62.7 ± 5.5
	S2 Novel animal exploration (s)	218 ± 51	113 ± 11 *	86.6 ± 6.3	97.2 ± 11.4
	Discrimination index (%)	81.3 ± 5.3	57.7 ± 4.0 **	59.2 ± 4.0	59.3 ± 3.9
Rotarod	D1 Time on the rod (s)	119 ± 7	113 ± 9	60.9 ± 14.4	61.4 ± 12.6
	D2 Time on the rod (s)	119 ± 10	129 ± 9	76.7 ± 8.5	76.1 ± 10.7
	D3 Time on the rod (s)	138 ± 12	140 ± 11	96.2 ± 10.1	101 ± 11
Notched Bar	Hind limb errors (%)	4.5 ± 0.8	5.2 ± 0.6	5.3 ± 0.6	5.6 ± 1.0
Grip Test	Grip stength (g/g body weight)	8.9 ± 0.5	10.7 ± 0.5 *	8.5 ± 0.5	7.2 ± 0.4 *

Supplementary Table 7. Behavioral characterization (following table) of *Del*/+ and *Dup*/+ cohorts on B6C3B genetic background. Short-term working memory was first assessed with the Y maze test which revealed improvements for *Dup*/+ animals. Recognition memory was first evaluated with the novel object recognition test. In the first session (S1) of test, no difference in object exploration was noticed. After a 3-hour retention delay, *Del*/+ and *Dup*/+ mice show respectively impairment and improvement for novel object discrimination. We did the same experience for *Dup*/+ cohort with a 24-hour retention delay and found no phenotype. With *Del*/+ cohort, we did object location recognition test with 3-hour retention delay and also found deficit for displaced object discrimination. In the social interaction test, we found a decrease of sniffing behavior for *Del*/+ mice and decrease of sniffing and following behaviors for *Dup*/+ animals. In the three-chamber sociability test, we did not

found difference in social interaction but novel animal discrimination was significantly decreased for *Del/+* animals. No locomotor coordination phenotype was seen in rotarod and notched bar tests but *Del/+* and *Dup/+* mice still presented respectively stronger and weaker grip strength in the grip test. NT: not tested. Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Test	Parameter	B6N *Del*/+ cohort results		B6N *Dup*/+ cohort results	
		wt	*Del*/+	wt	*Dup*/+
Body Weight Evolution	6-week weight	15.6 ± 0.6	11.5 ± 0.1 ***	17.3 ± 0.2	18.1 ± 0.4
	10-week weight	20.6 ± 0.6	16.6 ± 0.3 ***	21.9 ± 0.3	24.2 ± 0.6 **
	14-week weight	23.5 ± 0.8	18.5 ± 0.5 ***	25.6 ± 0.6	29.4 ± 1.2 *
Body size	snout/tail basis distance	9.61 ± 0.05	9.01 ± 0.07 ***	9.43 ± 0.04	9.84 ± 0.06 ***
qNMR	Fat body (%)	17.4 ± 2.1	9.9 ± 0.5 **	17.9 ± 1.2	25.0 ± 1.9 **
	Lean body (%)	72.8 ± 1.7	78.5 ± 0.7 **	73.1 ± 1.0	67.3 ± 1.6 *
	Free body fluids (%)	7.17 ± 0.80	8.54 ± 0.89	6.15 ± 0.26	5.52 ± 0.20
Indirect Calorimetry (TSE)	Light EE (Kcal/kg^0.75/h)	7.48 ± 0.27	7.86 ± 0.10	8.01 ± 0.21	7.04 ± 0.20 **
	Dark EE (Kcal/kg^0.75/h)	8.87 ± 0.39	10.49 ± 0.62 *	9.29 ± 0.28	8.11 ± 0.19 **
	Light OC (ml/kg^0.75/h)	1355 ± 47	1429 ± 18	1458 ± 37	1283 ± 28 **
	Dark OC (ml/kg^0.75/h)	1600 ± 70	1893 ± 111 *	1678 ± 51	1468 ± 32 **
Intraperitoneal Glucose tolerance test (IPGTT)	T_0 blood glucose (mg/dl)	94.0 ± 4.0	114 ± 8	79.0 ± 4.5	109 ± 8 *
	T_{60} blood glucose (mg/dl)	186 ± 6	154 ± 6 **	187 ± 10	203 ± 10
	T_{120} blood glucose (mg/dl)	103 ± 3	95.0 ± 7.8	87.1 ± 4.1	106 ± 6
	AUC (min*mg/dl)	13900 ± 855	8465 ± 739 ***	13995 ± 718	13352 ± 1038
Blood chemistry	Glucose (mmol/l)	15.0 ± 0.5	15.5 ± 0.4	16.5 ± 0.6	16.4 ± 0.5
	T Cholesterol (mmol/l)	2.40 ± 0.15	2.41 ± 0.09	2.51 ± 0.09	2.66 ± 0.17
	Free fatty acids (mEq/l)	0.65 ± 0.02	0.57 ± 0.03 *	0.54 ± 0.06	0.56 ± 0.02
Encocrinology	Insulin (µg/l)	1.13 ± 0.35	1.40 ± 0.49	1.12 ± 0.28	2.56 ± 0.58
	Leptin (ng/ml)	0.93 ± 0.22	0.32 ± 0.06 *	1.38 ± 0.31	4.16 ± 0.97 *
	Adiponectin (µg/l)	13.5 ± 0.7	11.4 ± 0.4 *	14.7 ± 0.5	15.5 ± 0.6

Supplementary Table 8. High-fat diet analysis of *Del*/+ and *Dup*/+ cohorts on B6N genetic background. Animals were put under high-fat diet at the age of 5 weeks. In comparison with controls, *Del*/+ animals were underweight, shorter in size and less fat whereas *Dup*/+ mice were overweight, longer in size and fatter. Indirect calorimetry showed increase level of energy expenditure (EE) and oxygen consumption (OC) for *Del*/+ mice during dark phase. *Dup*/+ mice presented a diminution of EE and OC during dark and light phases. In intraperitoneal glucose tolerance test (IPGTT), *Del*/+ mice showed faster glucose clearance whereas *Dup*/+ mice presented same glucose clearance than controls but had an increased glycemia before injection of glucose (T_0 blood-glucose). Blood chemistry analysis did not reveal gross hematology changes except a diminution of free fatty acids level of *Del*/+ mice. Finally, in consistence with qNMR results, endocrinology analysis showed diminution of leptin and adiponectin level for *Del*/+ mice and an increase of leptin level for *Dup*/+ mice. Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Test	Parameter	B6N *Del*/+ cohort results	
		wt	*Del*/+
Body Weight Evolution	6-week weight	19.2 ± 0.4	16.6 ± 0.3 ***
	10-week weight	22.4 ± 0.3	19.1 ± 0.3 ***
	14-week weight	25.2 ± 0.3	22.0 ± 0.4 ***
qNMR	Fat body (%)	20.4 ± 1.3	14.1 ± 0.6 ***
	Lean body (%)	71.7 ± 1.4	76.5 ± 0.5 **
	Free body fluids (%)	5.99 ± 0.28	7.23 ± 0.30 **
Indirect Calorimetry (TSE)	Light EE (Kcal/kg^0.75/h)	8.66 ± 0.23	9.96 ± 0.25 **
	Dark EE (Kcal/kg^0.75/h)	9.66 ± 0.22	11.45 ± 0.41 **
	Light OC (ml/kg^0.75/h)	1570 ± 40	1800 ± 45 **
	Dark OC (ml/kg^0.75/h)	1311 ± 33	1584 ± 57 **
Calorimetric Bomb	Feces energy content (cal/g feces)	3835 ± 46	3782 ± 70
	Energy excreted (Kcal/mouse/day)	3.07 ± 0.10	2.78 ± 0.20
Intraperitoneal Glucose tolerance test (IPGTT)	T_0 blood glucose (mg/dl)	62.7 ± 4.0	91.4 ± 4.7 ***
	T_{60} blood glucose (mg/dl)	232 ± 9	176 ± 21 *
	T_{120} blood glucose (mg/dl)	72.8 ± 1.6	82.1 ± 4.3
	AUC (min*mg/dl)	17778 ± 519	10334 ± 1908 **
Blood chemistry	Glucose (mmol/l)	16.0 ± 0.4	14.2 ± 0.6 *
	T Cholesterol (mmol/l)	3.21 ± 0.07	2.86 ± 0.10 *
	Free fatty acids (mEq/l)	0.95 ± 0.07	0.95 ± 0.03
	Ca (mmol/l)	2.29 ± 0.01	2.35 ± 0.01***
Encocrinology	Insulin (µg/l)	< 0.1	< 0.1
	Leptin (ng/ml)	39.0 ± 3.4	10.7 ± 2.5 ***
	Adiponectin (µg/l)	7.80 ± 0.18	6.34 ± 0.42 **

Supplementary Table 9. High-fat diet analysis of *Del*/+ cohort on B6C3B genetic background. Animals were put under high-fat diet at the age of 5 weeks. *Del*/+ animals underweight and less fat than wild-type littermates. Indirect calorimetry showed increase level of energy expenditure (EE) and oxygen consumption (OC) for *Del*/+ mice during dark and light phases. No difference of energy excreted in feces was found in calorimetric bomb analysis. During intraperitoneal glucose tolerance test (IPGTT), *Del*/+ mice showed faster glucose clearance but had an increased glycemia before injection of glucose (T_0 blood-glucose). Blood chemistry analysis revealed an increase level of calcium and decrease levels of glucose and cholesterol for mutant mice. Finally, in consistence with qNMR results, endocrinology analysis showed diminution of leptin and adiponectin level for *Del*/+ mice. Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Discussion et conclusion

Caractérisation des cohortes Del/+ et Dup/+ maintenues sur un fond génétique B6N

Lors de notre amplification, aucune différence de transmission n'a été observée pour le sexe du parent porteur de la mutation. Une importante létalité a été observée chez les animaux *Del/+*. Des tests de viabilité effectués au stade embryonnaire E18.5 ont permis de détecter une diminution de la masse corporelle, ainsi qu'un retard d'oxygénation des fœtus *Del/+*. Aucune anomalie n'a été observée en test de viabilité pour les individus *Dup/+*. Sous diète standard, les animaux adultes *Del/+* et *Dup/+* présentaient respectivement une diminution et une augmentation de la masse corporelle en comparaison aux animaux sauvages.

Dans le but de potentialiser les disparités de masse corporelle observées chez les animaux mutants, une analyse métabolique sous diète enrichie en graisses a été réalisée. Les animaux *Del/+* présentaient un déficit d'adipogenèse, une diminution de masse et de taille corporelles. Au contraire, les animaux *Dup/+* présentaient une augmentation d'adipogenèse, de masse et de taille corporelles. Aucune corrélation n'a été observée entre les différents paramètres. La technique de calorimétrie indirecte a révélé une importante perturbation de l'homéostasie énergétique. En comparaison aux animaux sauvages, la consommation d'oxygène et la chaleur dégagée étaient augmentées pour les animaux *Del/+* durant la phase nocturne et diminuées pour les animaux *Dup/+* durant les phases nocturnes et diurnes. Ces résultats indiquent une importante altération du niveau d'activité circadienne des animaux mutants. L'analyse hématologique n'a montré qu'une légère diminution des acides gras libres des animaux *Del/+*. L'analyse endocrinienne a révélé une diminution des niveaux de leptine et d'adiponectine pour les animaux *Del/+*, ainsi qu'une augmentation des niveaux de leptine pour les animaux *Dup/+*. L'ensemble de ces résultats indique une altération des fonctions métaboliques aboutissant à des phénotypes d'adipogenèse et de masse et de taille corporelles opposées chez les animaux *Del/+* et *Dup/+*. Ceci suggère la présence dans la région *Sult1a1-Spn* d'un ou plusieurs gène(s) sensible(s) au dosage génique codant pour un régulateur du métabolisme énergétique. Ces résultats obtenus chez la souris sous diète standard et diète enrichie en graisses sont opposés aux phénotypes retrouvés chez l'homme. Effectivement, la délétion de la région 16p11.2 est associée à l'obésité et à l'hyperphagie alors que la duplication réciproque est associée à l'insuffisance pondérale et à l'anorexie (Jacquemont et al., 2011; Walters et al., 2010). Néanmoins, les animaux *Del/+* et *Dup/+* n'ont présenté aucune différence de consommation d'eau et de nourriture en comparaison aux animaux

sauvages. Chez la souris, les altérations de la masse corporelle et d'adipogenèse ne sont donc pas associées à une perturbation de la prise alimentaire comme pour les patients.

L'analyse cranio-faciale a révélé une altération de la forme du crâne pour les animaux *Del/+* et *Dup/+*, ainsi qu'une diminution de la taille des crânes *Del/+*. Néanmoins, une importante diminution de la taille fémorale a également été observée pour les animaux *Del/+* qui indique un retard développemental affectant l'ensemble du squelette. Ces résultats sont à mettre en relation avec la diminution de la taille des patients porteurs de la délétion 16p11.2, ainsi qu'avec les altérations de taille du volume crânien observées chez les patients pour les deux réarrangements.

Lors de l'analyse comportementale des animaux, nous avons évalué l'ensemble des paramètres murins pouvant être associés aux nombreux troubles neuropsychiatriques observés chez les patients porteurs des délétions et duplications de la région 16p11.2. Nous avons observé pour les animaux *Del/+* une hyperactivité, des stéréotypies d'activité verticale ainsi qu'un déficit de la mémoire de reconnaissance. A l'opposé, les animaux *Dup/+* présentaient une hypoactivité, une diminution du niveau des mouvements répétés, ainsi qu'une amélioration de la mémoire de reconnaissance. Nous avons également observé des altérations de la coordination motrice et de la force d'agrippement pouvant être conséquents aux importantes altérations d'activité des animaux. Les tests de labyrinthe en Y et de piscine de Morris n'ont indiqué aucune altération de la mémoire de travail, ainsi que de l'apprentissage et de la mémoire spatiale. Les tests d'interaction sociale, de sensibilité au pentylènetétrazole et d'inhibition du sursaut n'ont révélé aucun phénotype pouvant être associé aux symptômes neuropsychiatriques décrits chez l'homme (McCarthy et al., 2009; Shinawi et al., 2010).

La récurrence de phénotypes opposés observés chez les animaux mutants et chez les patients suggère l'implication de gènes sensibles aux effets de dose. Des études transcriptomiques sur puces Affymetrix Mouse Gene 2.1 ST ont été réalisées à partir de prélèvements de foie, de cervelet, de striatum et d'hippocampe. L'analyse a révélé que la majorité des gènes de la région *Sult1a1-Spn* sont sensibles à la dose, particulièrement dans les structures cérébrales étudiées. L'analyse en composantes principales a indiqué un effet plus fort de la délétion sur l'expression des gènes de la région ce qui pourrait expliquer les phénotypes plus prononcés pour les animaux *Del/+*. Un impact très limité des réarrangements de la région *Sult1a1-Spn* a été observé pour la régulation du génome. Pour les nombreux phénotypes observés chez les animaux *Del/+* et *Dup/+*, notre étude transcriptomique suggère donc l'implication de gènes sensibles aux effets de dose localisés dans la région *Sult1a1-Spn*.

Caractérisation des cohortes Del-Dup maintenues sur un fond génétique B6N

Afin de valider nos résultats et d'étudier les animaux porteurs de la délétion et de la duplication (*Del/Dup*) de la région *Sult1a1-Spn*, nous avons généré une cohorte d'animaux *Del-Dup* comportant les 4 génotypes *Del/+*, wt, *Del/Dup* et *Dup/+*. Les phénotypes décrits précédemment ont été confirmés. Aucune altération des capacités auditives des souris en test d'ABR n'a été observée indiquant l'absence de surdité de nos animaux d'expérience.

Des phénotypes très intéressants ont été obtenus pour le test de reconnaissance du nouvel objet. Après un délai de 30min, les animaux *Del/+* présentaient un important déficit de discrimination, alors que les autres génotypes montraient des capacités similaires. Pour un délai de 3h, les animaux *Del/+* et *Dup/+* présentaient respectivement un déficit et une amélioration de la mémoire de reconnaissance, ce qui corrobore les résultats précédents. En addition, les animaux *Del/Dup* montraient également une amélioration de la discrimination du nouvel objet. Ces résultats suggèrent que la structure de l'ADN de la duplication *Sult1a1-Spn* pourrait impacter sur l'expression génique par des effets de position ou des perturbations de la chromatine pouvant potentiellement aboutir à l'amélioration de la mémoire de reconnaissance des animaux. De futures analyses transcriptomiques à partir de prélèvements d'hippocampe d'animaux *Del/Dup* permettront de répondre à notre hypothèse.

A la fin de la caractérisation comportementale, les animaux ont été hébergés de manière individuelle. Des fèces ont été récoltées pour chaque animal après trois semaines d'isolation pour éviter toute contamination. L'ADN bactérien a été isolé à et une analyse du microbiote intestinal a été réalisée. Néanmoins, aucune différence de distribution des populations bactériennes pouvant expliquer les disparités de masse corporelle et d'adipogenèse des animaux de différents génotypes n'a été observée.

Afin d'étudier des changements de transmission synaptique pouvant expliquer les altérations de mémoire de reconnaissance des animaux, des analyses électrophysiologiques sur coupes d'hippocampe aiguës ont été réalisées à partir d'une deuxième cohorte *Del-Dup*. Une diminution de la potentialisation à long terme (LTP) a été observée pour les animaux 16p11.2 *Del/+* et *Dup/+* dont seul le dernier groupe présente une significativité statistique en comparaison aux animaux sauvages. Aucune altération n'a été observée pour les animaux *Del/Dup*. Ces anomalies de transmission synaptiques de l'hippocampe ne suffisent pas à expliquer les phénotypes observés en test de reconnaissance du nouvel objet, notamment les améliorations des capacités de discrimination du nouvel objet observées pour les animaux *Dup/+* et *Del/Dup*.

Caractérisation des cohortes Del/+ et Dup/+ maintenues sur un fond génétique B6NC3B

Dans le but d'étudier l'influence du fond génétique, nous avons caractérisé les animaux *Del/+* et *Dup/+* maintenus sur un fond génétique hybride B6NC3B. Une complète restauration de la létalité associée à la délétion a été observée pour ce fond. En comparaison aux animaux sauvages, alors que les diminutions de masse et de la taille corporelles sont conservées, les animaux *Del/+* ne présentent plus de déficit d'adiposité. Quant aux animaux *Dup/+*, ils ne présentent aucun phénotype de masse, de taille et d'adiposité.

Nous avons observé une atténuation pour les phénotypes d'activité circadienne des animaux *Del/+* et *Dup/+*. En test de champ ouvert, les animaux mutants ne présentaient aucune altération de l'activité exploratoire en nouvel environnement. L'observation des mouvements répétés en cage d'hébergement n'a révélé qu'une augmentation et une diminution de l'activité d'escalade de la grille chez les animaux *Del/+* et *Dup/+* respectivement. Avec un délai de 3h, les phénotypes de reconnaissance du nouvel objet des animaux *Del/+* et *Dup/+* sont conservés. Nous avons également observé une amélioration des capacités de mémoire de travail en test de labyrinthe en Y pour les animaux *Dup/+*.

De manière intéressante, une diminution du niveau des comportements sociaux a été observée pour les animaux *Del/+* et *Dup/+* en test d'interaction sociale. Les animaux sauvages de souche B6NC3B présentent un niveau d'exploration sociale plus élevé que les animaux de souche pure B6N ce qui a permis de révéler les déficits d'exploration sociale des animaux mutants. Dans le test de sociabilité à trois chambres, aucun phénotype d'interaction sociale n'a été observé, mais les animaux *Del/+* ont présenté un important déficit de préférence sociale pour le nouvel individu. Pour valider ces résultats, nous avons évalué les capacités olfactives des animaux en test de reconnaissance d'odeurs sociales et non-sociales et n'avons observé aucune altération chez les animaux mutants. Ces premiers phénotypes allant dans le même sens pour les animaux *Del/+* et *Dup/+* sont à associer aux troubles du spectre autistiques observés chez les patients porteurs des délétions et duplications de la région 16p11.2 (Weiss et al., 2008). Nos résultats confirment l'influence du fond génétique sur la modulation du comportement émotionnel murin (Kerr et al., 2013; Pietropaolo et al., 2011). Le fond génétique est donc particulièrement important dans la modélisation murine des troubles neuropsychiatriques.

Conclusion de l'étude

Dans cette étude, nous proposons de nouveaux modèles murins pour les réarrangements de la 16p11.2 BP4-BP5. Contrairement aux études précédentes, la région ciblée dans notre étude correspond bien aux points de cassure BP4-BP5 et aucune altération des capacités auditives des animaux mutants n'a été observée. De plus, les animaux porteurs de la délétion et/ou de la duplication pour la région *Sult1a1-Spn* sont les premiers à avoir été générés sur un fond génétique consanguin B6N, ce qui nous a permis d'étudier les réarrangements de la région 16p11.2 dans un contexte génétique spécifique.

Un grand nombre de phénotypes opposés ont été observés pour les animaux *Del/+* et *Dup/+*, notamment pour les paramètres de masse corporelle, d'adipogenèse, d'activité, d'apprentissage et de mémoire. Nos études transcriptomiques suggèrent que la région de synténie 16p11.2 contiendrait un ou plusieurs gène(s) sensible(s) aux effets de dose. L'étude des animaux *Del/Dup* indique que la structure d'ADN de la duplication *Sult1a1-Spn* pourrait potentiellement impacter sur l'expression génique aboutissant à l'amélioration de la mémoire de reconnaissance. Dans le but d'étudier l'influence du fond génétique, nous avons répliqué notre étude avec des animaux maintenus sur un fond génétique B6NC3B. Malgré l'importante atténuation des phénotypes d'activité, nous avons retrouvé les phénotypes opposés de masse corporelle et de discrimination du nouvel objet. De manière intéressante, nous avons observé un déficit d'interaction sociale pour les animaux *Del/+* et *Dup/+* de fond hybride B6NC3B.

L'usage de deux fonds génétiques différents a permis d'apporter un grand nombre de précisions nécessaires à l'interprétation des phénotypes observés. Nous avons notamment observé que le fond B6N potentialisait les altérations d'activité, alors que le fond C3B potentialisait les déficits de comportements sociaux des animaux *Del/+* et *Dup/+* pour la région *Sult1a1-Spn*.

Partie 3

Modélisation des syndromes de délétion et de duplication 17q21.31

Introduction

Le syndrome de Koolen-De Vries causé par la délétion d'un segment de 500 à 650 kb du locus 17q21.31 se caractérise sur le plan clinique par une déficience intellectuelle, un comportement amical, une hypotonie du jeune sujet, une dysmorphie faciale caractéristique et de nombreuses anomalies congénitales (Koolen et al., 2008; Koolen et al., 2006). La duplication réciproque se caractérise par une symptomatique hétérogène incluant des troubles du spectre autistique ainsi qu'un retard psychomoteur (Grisart et al., 2009; Kirchhoff et al., 2007; Kitsiou-Tzeli et al., 2012). La région critique des syndromes de délétions et duplications de la région 17q21.31 présente une taille de 424 kb et inclut 5 gènes codants : *CRHR1, SPPL2C, MAPT, STH* et *KANSL1*. Récemment, des études de génétique humaine ont identifié des mutations perte de fonction ainsi que des délétions atypiques restreintes au gène *KANSL1* chez 6 patients présentant des symptômes similaires à ceux de la délétion 17q21.31 (Koolen et al., 2012; Zollino et al., 2012). *KANSL1* qui code pour une protéine d'interaction du complexe NSL (Dias et al., 2014) impliqué dans la régulation de la transcription est donc le principal gène candidat pour les phénotypes associés au syndrome de Koolen-De Vries.

Dans cette étude, nous présentons la génération et la caractérisation des premiers modèles murins pour les réarrangements de la région de synténie 17q21.31 localisée sur le locus 11E1 ainsi que la caractérisation des animaux KO hétérozygotes pour le gène *Kansl1*. La caractérisation comportementale des animaux porteurs de la délétion (*Del/+*) et de la duplication (*Dup/+*) a révélé de nombreux phénotypes au niveau de la masse et de la taille corporelle, de l'activité circadienne, des capacités mnésiques, des comportements sociaux et de la transmission synaptique au niveau de l'hippocampe. De manière intéressante, les animaux *Kansl1*$^{+/-}$ présentent de nombreux phénotypes similaires aux animaux *Del/+*.

Ce dernier manuscrit est en cours de préparation. Une échocardiographie a été effectuée pour les animaux de la cohorte *Del-Dup* dont nous sommes en attente des résultats d'analyse. En collaboration avec le Dr. Giovanni Iacono du NCMLS (Nijmejen Centre for Molecular Lige Sciences), des expériences de séquençage à ARN (RNA-seq) à partir de prélèvements d'hippocampe sont en cours pour les animaux *Del/+* et *Kansl1*$^{+/-}$ et leurs congénères sauvages. Nous allons prochainement réaliser une étude IRM cérébrale des animaux de la cohorte *Del-Dup* en collaboration avec le Pr. Mark Henkelman, directeur du MICe (Mouse Imaging Centre, Toronto Centre for Phenogenomics).

Il est important de préciser que les résultats de caractérisation de la cohorte *Kansl1*$^{+/-}$ constituent des données préliminaires. Effectivement, la cohorte comprenait un nombre très faible d'animaux incluant dix animaux sauvages et huit animaux mutants. L'importante létalité associée à l'haplo-insuffisance du gène *Kansl1* a nécessité l'inclusion d'animaux isolés. Pour éviter tout biais lié à ces conditions d'hébergement, nous avons repoussé l'étude des comportements sociaux de ces animaux. Une deuxième cohorte *Kansl1*$^{+/-}$ a été générée et sera prochainement caractérisée. L'analyse comportementale comportera l'ensemble des tests présentés dans ce manuscrit pour confirmations des phénotypes observés ainsi que deux tests de comportements sociaux incluant le test d'interaction sociale en champ ouvert ainsi que le test de sociabilité à trois chambres. En collaboration avec le Dr. Maksym Kopanitsa, des analyses électrophysiologiques sur coupes d'hippocampe aiguës seront également réalisées pour les animaux *Kansl1*$^{+/-}$ et leurs congénères sauvages.

Mouse models of *Kansl1* haploinsuffisancy and 17q21.31 rearrangements resume major phenotypes found in patients

Thomas ARBOGAST[1], Claire CHEVALIER[1], Maksym KOPANITSA[3], Nurudeen O. AFINOWI[3], Marie-Christine BIRLING[2], Marie-France CHAMPY[2], Hamid MEZIANE[2], Yann HERAULT[1,2]

Affiliations

[1]IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), Dpt of Translational Medicine and Neurogenetics; CNRS, UMR7104; INSERM, U964; Strasbourg University; 1 rue Laurent Fries, F-67404 Illkirch-Graffenstaden, France

[2]Institut Clinique de la Souris, ICS; PHENOMIN; GIE CERBM; 1 rue Laurent Fries, F-67404 Illkirch-Graffenstaden, France

[3]Synome Ltd, Moneta Building, Babraham Research Campus, Cambridge CB22 3AT, UK

Corresponding author: Yann Hérault

Keywords: Copy number variation, neurodevelopmental disorder, intellectual disability, sociability, mouse model, learning and memory

Running title

Rearrangements of the *Arf2-Kansl1* region and Kansl1 haploinsuffisancy

ABSTRACT

Deletion and duplication syndromes represent recurrent chromosomal abnormalities associated with distinct phenotypic features. Here we interested in the 17q21.31 rearrangements associated with intellectual disability and sociability alterations. Molecular analysis of patients defined a recurrent genomic fragment of 500k-650kb which contains 5 protein-coding genes: *CRHR1*, *SPPL2C*, *MAPT*, *STH*, and *KANSL1*. Deletions of the 17q21.31 region lead to Koolen-de Vries syndrome characterized by intellectual disability, hypotonia, friendly behavior, facial dysmorphism and several anomalies affecting notably heart, kidneys and central nervous system. Reciprocal duplications, more rare, are associated with autism spectrum disorders and psychomotor delay. Atypical deletions and loss-of-function mutations restricted to *KANSL1* gene have been found in several patients presenting similar phenotypes than carriers of the 17q21.31 deletion. *KANSL1* gene, which encodes for a modifier of the histone acetyltransferase KAT8, is a strong candidate to explain phenotypes observed in Koolen-de Vries patients. To confirm this hypothesis, we generated heterozygous *Kansl1* mutant mice (*Kansl1*$^{+/-}$) and mice carrying the deletion (*Del/+*) and the reciprocal duplication (*Dup/+*) for the 17q21.31 syntenic region found on mouse chromosome 11E1. Characterization studies revealed several phenotypes affecting body mass and size, adipogenesis, activity, social behaviors, recognition and associative memories. Very interestingly, *Kansl1*$^{+/-}$ mice showed similar anatomical and behavioral phenotypes affecting *Del/+* mice. Overall, this study corroborates human genetic data about major implication of *KANSL1* in the Koolen-de Vries syndrome. Our mouse models provide new genetic tools for the identification of pathways implicated in the pathophysiology of 17q21.31 rearrangements, and will be valuable for the development of therapeutic approaches.

INTRODUCTION

During the last decade, an important number of neurodevelopmental diseases has been associated with copy number variants (CNVs) consisting in deletions and duplications of DNA segments that are 50bp or larger (Cooper et al., 2011; Grayton et al., 2012). Main mechanism of CNV generation is non-allelic homologous recombination between segmental duplications (SDs). The locus 17q21.31 occurs in humans as a direct (H1) or inverted (H2) haplotypes with differential predispositions to disease (Canu et al., 2009; Ghidoni et al., 2006; Li et al., 2014; Permuth-Wey et al., 2013). The H2 inversion polymorphism found at a frequency of ~20% in Europeans (Koolen et al., 2006; Steinberg et al., 2012; Zody et al., 2008) encompasses directly oriented low copy repeat (LCRs) subunits prone to generate CNVs. The 17q21.31 locus encompasses several genomic structural variants ranging from 200 to 800 kb including recurrent genomic fragments of 500-650kb which contains 5 protein-coding genes: *CRHR1*, *SPPL2C*, *MAPT*, *STH*, and *KANSL1*. The 17q21.31 deletion or Koolen-de Vries syndrome has a population prevalence of approximately 1/16,000 (Koolen et al., 2008). Carriers present characteristic facial dysmorphism such as long face, proeminent ears and tubular nose with bulbous nasal tip. Clinical features include intellectual disability, friendly behavior, hypotonia, and several anomalies affecting notably heart, kidneys and central nervous system (Dubourg et al., 2011; Koolen et al., 2008; Tan et al., 2009). The reciprocal duplication is much rarer. To date, only 6 patients carrying t 17q21.31 duplications have been described in the literature (Grisart et al., 2009; Kirchhoff et al., 2007; Kitsiou-Tzeli et al., 2012). Symptomatic is quite heterogeneous and includes craniofacial malformations, psychomotor delay and autism spectrum disorders.

Recently, loss-of-function mutations and atypical deletions restricted to the *KANSL1* gene have been found in 6 patients who display similar phenotypes than 17q21.31 deletion carriers (Koolen et al., 2012; Zollino et al., 2012). *KASNL1* gene encodes the 1,105 amino acid long nuclear protein KAT8 regulatory NSL complex subunit 1 (KANSL1). KANSL1 protein is a member of the evolutionarily conserved NSL (NonSpecific Lethal) complex which plays important roles in various cellular functions, including transcription regulation and stem cell identity maintenance and reprogramming (Mendjan et al., 2006; Raja et al., 2010). The NSL complex contains the histone acetyltransferase MOF (Males Absent on the First) encode by *KAT8* gene which acetylates histone H4 (H4K16) on lysine 16 and weakly on lysines 5 and 8 (H4K5 and H4K8, respectively) to facilitate activation of transcription (Cai et

al., 2010; Dou et al., 2005). Recent works on fly have shown that KANSL1 acts as a scaffold protein interacting with four NSL subunits including WDR5 which plays an important role in assembling distinct histone-modifying complexes with different epigenetic regulatory roles (Dias et al., 2014).

Genes of the 17q21.31 region have been found to be conserved on mouse chromosome 11E1. *Crhr1, Sppl2c, Mapt* and *Kansl1* orthologs have been found in the same orientation than H1 human haplotype. *Mapt* and *Chrh1* knock-out (KO) mutant mice have been generated but no phenotype has been observed for heterozygous mutant animals (Ikegami et al., 2000; Wang et al., 2011). So far, no KO mice for *Kansl1* and *Sppl2c* genes and no mouse models of 17q21.31 rearrangements has been described in the literature. Here we present the first 17q21.31 deletion (*Del*/+) and/or duplication (*Dup*/+) mouse models for the *Arf2-Kansl1* genetic interval and the first heterozygous *Kansl1* mutant mice (Kansl1$^{+/-}$). We found that *Arf2-Kansl1* rearrangements affect body size and mass, adipogenesis, activity, social behavior, recognition and associative memories. *Del*/+ mice displayed body size reduction, adipogenesis deficits, hypoactivity, higher level of social interaction and impairments for recognition and associative memories. *Dup*/+ mice showed body mass and size reduction, hypoactivity and associative memory improvements. Skull shape anomalies and alterations of synaptic transmission in the hippocampus have also been found in *Del*/+ and *Dup*/+ mice. Interestingly, *Kansl1*$^{+/-}$ mice showed phenotypes pretty similar than *Del*/+ mice including body mass and size reduction, adipogenesis deficits, motor coordination improvements and impairments for recognition and associative memories. *Del*/+ and *Kansl1*$^{+/-}$ mice resume principal features of Koolen-de Vries patients including hypersociability and cognitive deficits. Our mouse models represent optimal genetic tools to develop and evaluate first therapeutic strategies for 17q21.31 rearrangement syndromes.

RESULTS

Generation and transmission of mice carrying rearrangements of the *Arf2–Kansl1* genetic interval

To generate mouse models for 17q21.31 rearrangements, we introduced LoxP sites flanking synthenic region located on chromosome 11E1 (Fig. 1A). Mutant mice were obtained by in vivo TAMERE strategy (Brault et al., 2006; Herault et al., 1998). LoxP sites were first introduced by homologous recombination in embryonic stem cells at *Arf2* and *Kansl1* loci in the same orientation using specific vector and the corresponding mouse line were generated. Selection cassettes were excised then mice were crossed with Hprt<tm1(cre)Mnn> mouse line, expressing the Cre recombinase under the control of the X-linked hypoxanthine guanine phosphoribosyl transferase gene promoter active in oocytes (Tang et al., 2002). Females born from this mating and bearing both the Hprt<tm1(cre)Mnn> transgene and loxP sites in a trans configuration were mated with wild-type C57BL/6N males (Fig. 1B). We recovered mice carrying the deletion (*Del/+*) and the duplication (*Dup/+*) for the *Arf2-Kansl1* region with a recombination frequency of respectively 11 and 7 animals out of 140 newborns.

Del/+ and *Dup/+* mice were first amplified separately on pure B6N background. Over 338 animals obtained by crossing Dup/+ with wild-type (2n) mice, 152 were carrying the duplication (transmission rate of 45.0%; Table 1). Only 59 over 161 animals obtained by crossing Del/+ with 2n mice were carrying the deletion (36.6%). In order to perform characterization analysis with cohorts containing all genotypes, we crossed Del/+ with Dup/+ mice. Over 189 animals generated, we obtained 71 wild-type (37.6%), 14 Del/+ (7.4%), 48 Dup/+ (25.4%) and 56 Del/Dup (29.6%) animals. These data indicate an important lethality associated with deletion of the *Arf2-Kansl1* region (Chi-squared test, $p < 0.001$).

Weight evolution of adult animal revealed that all mutant mice were underweight in comparison with wild-types with significant results for Dup/+ mice (Repeated Measures ANOVA "genotype" $F_{(3,300)} = 5.529$, $p = 0.002$; *Del/+* vs wt: $p = 0.101$, *Dup/+* vs wt: $p = 0.002$, *Del/Dup* vs wt: $p = 0.109$; Fig. 1D). Similarly, body sizes of mutant animals were significantly reduced ($H_{(3, 55)} = 28.036$, $p < 0.001$; Mann-Whitney *U* test, *Del/+* vs wt: $p = 0.004$, *Dup/+* vs wt: $p < 0.001$, *Del/Dup* vs wt: $p = 0.003$; Fig. 1E). Finally, we evaluated body fat percentage of animals by qNMR and found adiposity reduction for *Del/+* mice ($H_{(3,}$

$_{55)}$ = 8.120, p = 0.044; *Del/+* vs wt: p = 0.027; Fig. 1F). Mutant animals displayed normal feeding behaviors in comparison with wild-types (Supplementary Fig. S1C-D).

Rearrangements of the *Arf2–Kansl1* region induce circadian hypoactivity and fine motor coordination alterations

Activity of animals was first evaluated with the circadian activity test. *Del/+* and *Dup/+* mice displayed a normal pattern of circadian locomotor activity (Fig. S1) but their baseline levels of locomotor and vertical activities were distinct from those of wild-types (Fig. 2A). In comparison with wild-type littermates, *Del/+* mice showed reduced locomotor activity during the dark phase ($F_{(3,46)}$ = 5.791, p = 0.002; *Del/+* vs wt: p = 0.002) and the light phase ($F_{(3,46)}$ = 4.260, p = 0.010; *Del/+* vs wt: p = 0.006) and reduced vertical activity during light phase ($H_{(3,46)}$ = 12.861, p = 0.005; *Del/+* vs wt: p = 0.002) whereas *Dup/+* and *Del/Dup* mice carrying both rearrangements behaved similarly than wild-types (p > 0.05 for all measures). When tested in novel open field arena, *Del/+* and *Dup/+* mice displayed respectively higher and reduced level of rearing behavior but phenotypes failed to reach statistical significance in comparison with wild-type littermates ($F_{(3,58)}$ = 3.923, p = 0.013; *Del/+* vs wt: p = 0.185, *Dup/+* vs wt: p = 0.266, *Del/+* vs *Dup/+*: p = 0.007; Fig. 2B). No phenotype was found for exploration locomotor activity and time spent in the arena center (Supplementary Table S1). In the elevated plus maze, *Del/+* and *Dup/+* mice explored the same number of arms and spent similar duration in open arms than controls (Table S1) which indicates no obvious anxiety in mutant animals.

Motor coordination and motor learning were evaluated in the rotarod test. Mice were placed on an accelerating rotarod and latencies to fall off the rod were recorded in 3 successive daily sessions (Fig 2C). Again, *Del/+* and *Dup/+* mice displayed respectively motor coordination improvements and deficits but phenotypes failed to reach statistical significance in comparison with wild-type littermates (Repeated Measures ANOVA "genotype" $F_{(3,116)}$ = 3.610, p = 0.018; *Del/+* vs wt: p = 0.215, *Dup/+* vs wt: p = 0.346, *Del/+* vs *Dup/+*: p = 0.010). Compared to wild-types, no alteration of motor coordination capacities was found in *Del/Dup* mice. In the grip test, no alteration of muscular strength was detected in mutant mice (Table S2). Compared to wild-types, no alteration of activity and coordination capacities was found in *Del/Dup* mice.

Effects of *Arf2–Kansl1* dosage on social behaviors

We first investigated social behaviors in the three-chamber sociability test. All genotypes explored the three chambers similarly excluding a place preference in the arena. During the social interest session, *Del/+* mice spent more displayed an interaction time with the first congener significantly higher than wild-types ($F_{(3,51)}$ = 3.447, p = 0.023; *Del/+* vs wt: p = 0.017; Fig 2D, Session 1). In the social discrimination session, *Del/+* displayed higher interaction time with familiar and novel congeners without statistical significance (Fig. S2D). In comparison with wild-type littermates, total interaction time with the two congeners during social discrimination session was again higher for *Del/+* mice ($F_{(3,51)}$ = 6.034, p = 0.001; *Del/+* vs wt: p = 0.002; Fig 2D, Session 2). No alteration of social behaviors was found for *Dup/+* and *Del/Dup* mice in comparison with wild-types (p > 0.05 for all measures). Percentage of time spent to explore the novel congener was similar between genotypes indicating normal social discrimination in mutant mice (Fig 2E). Social behaviors of animals were also evaluated in the social interaction test. In this test, 2 mice of same genotype from different home-cages were put in open field and let free to interact during 10 min (Fig. S2E-G). Latency to first contact was increased for *Dup/+* mice without statistical significance. Interestingly, we observed higher and lower level of social interaction for *Del/+* and *Dup/+* mice respectively. Nevertheless, as animal effective was divided by 2, effects failed to reach statistical significance. *Del/Dup* mice behaved similarly than wild-types.

Rearrangements of the *Arf2–Kansl1* region induce cognitive function alterations

We then investigated cognitive functions of animals. In the Y maze test, no working memory alteration was observed (Table S2). We used the Morris water maze to evaluate spatial learning and memory. Again, mutant mice displayed similar performances than wild-types (Fig. S3). Recognition memory of mice was assessed with the novel object recognition task. We investigated whether mutant mice could discriminate a novel object from a previously explored object after a retention delay of 3 hours (Fig 2F). During the acquisition session, all genotypes spent an equal amount of time exploring the sample object (Table S2). In the subsequent choice session *Del/+* mice displayed a significant memory impairment compared to wild-types ($F_{(3,43)}$ = 3.081, p = 0.037; *Del/+* vs wt: p = 0.040). *Dup/+* and *Del/Dup* mice showed no alteration of recognition memory. Finally, we evaluated aversive associative memory with the fear conditioning test (Fig 2G). No difference in baseline between was

noticed between genotypes in conditioning session (Table S2). In the context session, *Dup/+* mice displayed an increase of freezing with significant difference in the last 2 min of test ($F_{(3,56)} = 6.399$, $p < 0.001$; *Dup/+* vs wt: $p = 0.027$). Cue session was performed 5 hours after the context session. During presentation of conditioning cue, all genotypes increased their freezing. Phenotypes were observed in the second cue during which *Del/+* and *Dup/+* showed a freezing level respectively lower and higher in comparison with wild-type littermates ($H_{(3, 56)} = 20.609$, $p < 0.001$; *Del/+* vs wt: $p = 0.002$, *Dup/+* vs wt: $p = 0.036$). *Del/Dup* mice carrying both rearrangements behaved similarly than wild-types ($p > 0.05$ for all measures).

Effects of *Arf2-Kansl1* dosage on synaptic transmission and synaptic plasticity in hippocampal slices

To explore if *Arf2-Kansl1* region regulates electrophysiological parameters in mouse neurons, we assessed basal synaptic transmission and synaptic plasticity by measuring field excitatory postsynaptic potentials (fEPSPs) in acute hippocampal slices from *Del/+* and *Dup/+* mice. In Del/+ mutants, we observed decreased fEPSP slopes in mutant slices, especially in response to higher stimulus strengths (Fig. 3A). Mean slopes of fEPSPs invoked by the maximum stimulus strength (4.2 V) were significantly smaller in slices from *Del/+* mice (1.46 ± 0.09 mV/ms) compared to wt littermates (1.87 ± 0.09 mV; $F_{(1,13.34)} = 8.31$; $p = 0.025$; two-way nested ANOVA, genotype effect). Mean paired-pulse ratio of slopes of fEPSPs evoked at a 50 ms interpulse interval was also slightly, but significantly lower in mutant slices (Fig. 3C; $F_{(1, 11.04)} = 6.506$; $P = 0.027$. No significant changes in LTP elicited by theta-burst stimulation were seen in slices of *Del/+* mice (Fig. 3E).

Basal synaptic strength was slightly up-regulated in slices from Dup/+ mice and a significant interaction was observed between genotype and stimulus strength (Fig. 3B). *Post hoc* Bonferroni comparisons that treated data from each slice as independent, revealed that mean fEPSP slopes in *Dup/+* mice were significantly higher at two submaximal stimuli. However, a more stringent nested ANOVA design failed to reveal significant differences in fEPSP slopes evoked by maximum (4.2 V) stimulus ($F_{(1, 8.67)} = 3.09$; $p = 0.113$). Similarly, paired-pulse facilitation was nominally higher in slices from *Dup/+* mice (Fig. 3D), but the effect did not reach statistical significance ($F_{(1, 4.92)} = 5.402$; $p = 0.069$). Also, long-term potentiation elicited by the theta-burst stimulation was not different between the mutant and litter-matched wt mice (Fig. 3F).

Rearrangements of the *Arf2-Kansl1* region induce craniofacial features

In order to study the influence of 17q21.31 CNVs on craniofacial structure, we performed an analysis of computed tomography (CT) cranial scans of animal heads with ulterior reconstruction of 3D skull images using 39 cranial landmarks (Fig. 4A). Separate *Del/+* and *Dup/+* cohorts were used for the analysis. No alteration of skull size was found for *Del/+* (p = 0.115; Fig. 5B) and *Dup/+* (p = 0.115; Fig. 4D). Concerning skull shape, we noticed alterations very close from statistical significance in both Del/+ (p = 0.077; Fig. 4C) and Dup/+ (p = 0.052; Fig. 4E).

Hetezozygous *Kansl1* mutant mice recapitulate phenotypes observed in mice carrying deletion of the *Arf2-Kansl1* region

We repeated our analysis with a cohort of heterozygous *Kansl1* knock-out (KO) mice (Kansl1$^{+/-}$). Weight evolution of adult animal revealed that *Kansl1*$^{+/-}$ mice were underweight in comparison with wild-types (Repeated Measures ANOVA "genotype" $F_{(1,30)} = 11.729$, p = 0.004; Fig. 4A). Similarly, body sizes of mutant animals were significantly reduced ($F_{(1,15)} = 11.516$, p = 0.004; Fig. S4B). Finally, we evaluated body fat percentage of animals by qNMR and also found adiposity reduction in Kansl1$^{+/-}$ mice ($F_{(1,15)} = 6.813$, p = 0.020; Fig. S4C).

In the circadian activity test, *Kansl1*$^{+/-}$ mice displayed normal patterns of activity (Fig. S5A, B) but their baseline levels of locomotor and vertical activities were distinct from those of wild-types (Fig. 5A). In comparison with wild-type littermates, *Kansl1*$^{+/-}$ mice showed reduced locomotor activity during the dark phase ($F_{(1,16)} = 8.482$, p = 0.010) and the light phase($F_{(1,16)} = 8.573$, p = 0.010). Mutant animals also displayed trends for vertical hyperactivity during the dark phase. Feeding behaviors during circadian activity test were unchanged in mutant animals (Fig. S5C, D). When tested in novel open field arena, *Kansl1*$^{+/-}$ mice travelled the same distance and spent similar time in the arena center in comparison with wild-types but mutant mice showed a significant increase of rears during the 30 min of test ($F_{(1,15)} = 4.846$, p = 0.044; Fig. 5B, Table S3). To further investigate this vertical hyperactivity, we performed visual observations of animals in odorless home-cages (Fig. 5C). Mutant mice displayed an increased level of rearing ($F_{(1,15)} = 7.207$, p = 0.017) and decreased level of digging ($F_{(1,15)} = 12.268$, p = 0.003). These results indicate a global alteration of *Kansl1*$^{+-}$ activity characterized by locomotor hypoactivity and vertical hyperactivity.

In the rotarod test, *Kansl1*$^{+-}$ mice displayed a clear improvement of motor coordination and motor learning in comparison with wild-types (Repeated Measures ANOVA "genotype" $F_{(1,30)} = 115.867$, $p < 0.001$; Fig. 5D). No alteration of grip strength capacities was noticed in the grip test (Table S3). Recognition memory of mice was again assessed in the novel object recognition task with a delay of 3 hours between acquisition and choice sessions. During the acquisition session, mutant and wild-type mice showed similar exploration of the first object (Table S3). In the subsequent choice session, *Kansl1*$^{+-}$ mice displayed a significant memory impairment compared to wild-types ($F_{(1,15)} = 22.566$, $p < 0.001$; Fig. 5E). Finally, we evaluated associative memory with the fear conditioning test (Fig 5F). No difference in baseline and post-shock freezing levels was noticed between genotypes in conditioning session (Table S3). In the context session, *Kansl1*$^{+-}$ mice displayed a global increase of freezing with significant difference in the last 2 min of test ($H_{(1,16)} = 6.419$, $p = 0.011$). In the cue session, a decrease of freezing level was noticed for *Kansl1*$^{+-}$ mice during the second cue ($F_{(1,16)} = 16.748$, $p < 0.001$).

DISCUSSION

17q21.31 deletion and duplication syndromes are neurodevelopmental disorder associated with combination of physiological and cognitive defects. 17q21.31 deletion or Koolen-de Vries syndrome has an important prevalence and is well characterized. Patients present facial dymorphisms, mild to moderate intellectual disabilities, a friendly behavior and several abnormalities affecting heart, kidneys and central nervous system. On the opposite, 17q21.31 duplication is less pathogenic and only few carrier patients have been described. Symptomatic includes craniofacial malformations, psychomotor delay and autism spectrum disorders without striking recurrent phenotype. Interestingly, patients carrying deletions and loss-of-function mutations and deletions restricted to the *KANSL1* gene display similar phenotypes than 17q21.31 deletion carriers.

In order to dissect the molecular mechanisms underlying the 17q21.31 CNV syndromes, we generated and characterized mice carrying the deletion (*Del/+*) and the duplication (*Dup/+*) for the *Arf2-Kansl1* region and *Kansl1* heterozygous KO mice ($Kansl1^{+/-}$). We identified a set of anatomical, metabolic, behavioral and electrophysiological phenotypes. Some alterations were common to all the mice including body weight and body size reduction and skull shape changes. An important number of behavioral phenotypes were found for *Del/+* mice including circadian hypoactivity, higher level of social interaction and impairments for recognition memory and associative memory (Table 2). On the contrary, Dup/+ mice only displayed associative memory improvements. Interestingly, *Kansl1* KO mice displayed similar anatomical and behavioral phenotypes than *Del/+* mice including body size reduction, adipogenesis deficit and memory deficits which suggests that *Kansl1* gene may be the major driver of phenotypes observed in *Del/+* mice (Table 2).

Electrophysiological experiments show opposite effects of *Arf2-Kansl1* rearrangements on basal synaptic transmission and short-term plasticity of excitatory synapses in the hippocampus. It is notable that deletion and duplication of this region induce opposite changes in affected parameters: deletion significantly decreases synaptic strength and paired-pulse facilitation, while duplication led to slightly elevated basal synaptic responses and nominally (albeit not statistically) higher paired-pulse facilitation. It would be interesting to explore in future, whether these changes are explained by changes in AMPA receptor numbers and/or altered transmitter release probability within individual synapses.

Phenotypes observed in mouse models for 17q21.31 rearrangements and *Kans11* haploinsuffisancy are very close to human symptoms. Overall results presented here confirm human genetics studies about *KANSL1* gene as the major driver of symptoms of the Koolen-de Vries syndrome. By using genetic approaches to modulate the expression of *Kansl1* gene in *Del/+* and *Dup/+* mice, we think that our models will help for the identification of molecular processes associated with 17q21.31 deletion and duplication syndromes and for the development of therapeutic treatments.

MATERIALS AND METHODS

Mouse lines, genotyping and ethical statement

17p21.31Yah mice were generated in a C57BL/6N genetic background through Cre-LoxP in vivo recombination using a mouse line carrying two loxP sites inserted upstream upstream *Arf2* and downstream *Kansl1* genes in a trans configuration. Deletion of *Arf2–Kansl1* region was identified by PCR using primers Fwd1 (5'-TCCTAACCCACGGGTCAGCCTA-3') and Rev2 (5'-GCCCATACGGATGTTTCCTTCCAA-3'). Duplication of the same region was identified using primers Fwd2 (5'- CACCACGGGAGCAAGCAACTGAG-3') and Rev1 (5'-CCAGCCCACTCAGACTTTCCAG AAT-3'). Wild-type allele was identified using Fwd1 and Rev1 primers. PCR reactions gave deletion, duplication and wild-type products of 448 bp, 653 bp and 341 bp long respectively (Figure 1C). All mice were genotyped by PCR reaction following this program: 95°C /5min; 35x (95°C/30sec, 65°C/30sec, 70°C/1min), 70°C/5min.

Kansl1-het mutant mice ($Kansl1^{+/-}$) were generated in a C57BL/6N genetic background from the unique IKMP ES cell clone HEPD0766_8_G02. $Kansl1^{tm1b(EUCOMM)Hmgu}$ (Skarnes et al., 2011) were obtained by breeding $Kansl1^{tm1a(EUCOMM)Hmgu}$ mice with a Cre delete developed at the "Institut Clinique de la Souris" (ICS) (Birling et al., 2012). Deletion of the targeted exon was identified using primers Ri1Fwd (5'-GCACATGGCTGAATATCGACGGT-3') and LxRwd (5'-ACTGATGGCGAGCTCAGACCATAAC-3'). PCR reactions gave a product of 471 bp long with the program following: 95°C /4min; 35x (94°C/30sec, 62°C/30sec, 72°C/1min), 72°C/7min. The experimental procedures were approved by the local ethical committee Com'Eth under accreditation number (2012-069) with YH, as the principal investigator in this study, (accreditation 67-369).

Behavioral analysis

Only male mice were used for behavioral studies. Cohorts of experimental animals were formed by selecting mice coming from litters containing a minimum of two male pups were selected. After weaning, animals were gathered by litters in 39 x 20 x 16 cm cage (Green Line, Techniplast, Italy) where they had free access to water and food (D04 chow diet, Safe). Temperature was maintained at 23±1°C and the light cycle was controlled as 12 hour light and 12 hour dark (lights on at 7 am). Mice were transferred from the animal facility to the phenotyping area at the age of 10 weeks. In the testing days, animals were transferred to

experimental room antechambers 30 min before the start of the experiment. All the experiments were done between 8:00 AM and 2:00 PM. Behavioral experiments were conducted between 12 and 20 weeks of age.

We crossed *Del/+* with *Dup/+* animals and generated two *Del/Dup* cohorts of males (n total = 24 wt, 18 *Del/+*, 11 *Dup/+*, 11 *Del/Dup*). To produce experimental groups, only animals coming from litters containing a minimum of two male pups were selected. Were passed animals at the same age under the same behavioral pipeline and we pooled the data together. Tests were administered in the following order: circadian activity, elevated plus maze, open field, novel object recognition with 3h delay, novel object recognition, three-chamber test, Morris water maze, social interaction, rotarod, grip test, and fear conditioning. Resting period of 2 days to 1 week was used between two consecutives tests. Body weights of animals were recorded once a week (same day at the same time) from 13 to 21 weeks of age.

Kansl1$^{+/-}$ cohort including 10 wt and 8 *Kansl1*$^{+/-}$ males was generated by in vitro fertilization. Due to a high lethality, some animals were isolated in home cages including 3 mutant mice and 2 wt animals. Due to a small number of mutant mice including 3 isolated animals over 8, we did not performed social behavioral tests. Tests were administered in the following order: circadian activity, elevated plus maze, open field, novel object recognition, repetitive behaviors, rotarod, grip test, and fear conditioning. Body weights of animals were recorded at 15, 17 and 19 weeks of age. Resting period of 2 days to 1 week was used between two consecutives tests.

Circadian activity (AC): The test allows the evaluation of endogenous activity and feeding behavior over a complete light/dark cycle. Testing was done in individual cages (11 x 21 x 18 cm^3) fitted with infrared captors linked to an electronic interface (Imetronic, France) which provide automated measures of position and locomotor activity were used for the experiment. Mice were put in cages at 11 am the first day and removed the second day at 7 pm. The light cycle was controlled as 12 h light and 12 h dark (lights on at 7 am). The 32 hours of test consist in three different phases: the habituation phase (from 11am to 7pm on the first day), the night/dark phase (from 7pm on the first day to 7am on the second day), and the day/light phase (from 7am to 7pm on the second day). Feeding behaviors were evaluated using an automated lickometer and a 20mg pellet feeder (Test Diet, Hoffman La-Roche).

Open-field (OF): The test is used to evaluate exploration behavior. Mice were tested in automated open fields (44.3 x 44.3 x 16.8 cm) made of PVC with transparent walls and a black floor, and covered with translucent PVC (Panlab, Barcelona, Spain). The open field arena is divided into central and peripheral regions and is homogeneously illuminated at 150 Lux. Each mouse was placed in the periphery of the open field and allowed to explore freely the apparatus for 30 min. The distance travelled, the number of rears and time spent in the central and peripheral parts of the arena were recorded over the test session.

Repetitive behavior: Mice were put individually in clean home-cages dimly lighted (60 lux) without the pellets and water bottle. The occurrence of repetitive behaviors (rearing, jumping, climbing, digging, grooming) was observed for 10 min and scored using an ethological keyboard (Viewpoint, Labwatcher, France).

Elevated plus maze: The apparatus consists of two opposed open arms (30 x 5 cm) crossed by two enclosed arms (30 x 5 x 15 cm), and elevated 66 cm from the floor. The light intensity at the extremity of the open arms was kept at 50 Lux. Each mouse was tested for 5 min after being placed in the central platform and allowed to explore freely the apparatus. The number of entries and time spent in the open arms were used as an anxiety index. Closed arm entries and rears in the closed arms were used as measures of general motor activity.

Y-maze: This test which is used to evaluate short-term working memory is based on the innate preference of animals to explore arms that has not been previously explored, a behavior that, if occurring with a frequency greater than 50%, is called spontaneous alternation. The apparatus consists in a Y-shaped maze with three white, opaque plexiglas arms of equivalent length forming a 120° angle with each other. The arms have walls with specific motifs allowing distinction from each other. After introduction at the center of the maze at 60 Lux, animal were allowed to freely explore the maze for 6 minutes. The number of arm entries and the number of triads were recorded in order to calculate the percentage of alternation.

Novel object recognition task (NOR): This test evaluates object recognition memory and is based on the innate preference of rodents to explore novelty. The test was carried in an open field arena as previously described (Goeldner et al., 2008). On the first day, mice were habituated to the arena for 30 min at 60 Lux. The following day, animals were submitted to a first 10-min acquisition trial during which they were individually placed in the presence of the object A (marble or dice) placed at 10cm of one corner of the box. The exploration time of

object A (when the animal's snout was directed toward the object at a distance ≤1 cm) was recorded. A 10-min retention trial (second trial) was conducted 3 hours later. The familiar object (object A) and novel object (object B) were placed at 10 cm of two open field corner (the distance between the two objects was about 20 cm) and the exploration time of the two objects was recorded. A discrimination index was defined as $(t_B/(t_A + t_B)) \times 100$. All mice which did not explore the first object more than 3 seconds during the acquisition trial were excluded from the analysis.

Morris water maze paradigm (MWM): The test is a paradigm for spatial learning and memory. The protocol is adapted from the already established protocol (Duchon et al., 2011). The apparatus consists in a circular pool (150-cm diameter, 60-cm height) filled to a depth of 40 cm with water maintained at 20°C–22°C and made opaque using a white aqueous emulsion (Acusol OP 301 opacifier). An escape platform, made of 6 cm diameter rough plastic, is submerged 1 cm below the water surface. The test began with 6 days of acquisition, 4 trials per day, at 120 Lux. Each trial started with the mice facing the interior wall of the pool and ended when animals climb on the platform or after a maximum searching time of 90 sec. The platform was at the same position for all the four trials but starting positions changed randomly between each trial with departures from each cardinal point. Travelled distances to find the platform and swimming speeds were analyzed each day. On 7th day, mice were given a single trial of 60 seconds trial during the probe test or removal session in which the platform has been removed. The distance traveled and duration spent in each quadrant (NW, NE, SW, SE) were recorded. Annulus crossing index was calculated as the number of times that animals crossed the exact platform position. On 12th and 13th days, mice were given reversal sessions with 4 trials of 90 seconds per day. The platform was made visible by a small dark ball placed 12cm on top of the platform, while the external cues were hidden by surrounding the pool with a black curtain. In order to be sure that the mouse used the platform cue, starting position and platform position were changed for each trial.

Fear conditioning: The test is used to assess aversive learning and memory. Polymodal operant chambers (Coulbourn Instruments, USA) consisting of steel rod floor and plexiglas walls of 18.5 x 18 x 21.5 cm are used for this test. On the first day, during the conditioning session, mice were allowed to acclimate in the chambers for 4 min, then a light/tone (10 kHz, 80-dB) conditioned stimulus (CS) was presented for 20 sec and coterminated by a mild (0.4 mA, 1 sec) footshock unconditioned stimulus (US). Mice were left in the chambers for

another 2 min then were put back to their homecages. Context session was performed 24 hours after the conditioning session. Mice were placed into the same chambers and allowed to explore for 6 min without presentation of light/auditory CS. Cue session was performed 5 hours after the context session. Animals were tested in distinct appearance chamber (grey plastic ground, black opaque plastic walls, change of light orientation) and were allowed to habituate for 2 min then presented to light/auditory CS. This sequence was repeated once again. During each session, animal movement was monitored to detect freezing behavior.

Social interaction task: The test allows the evaluation of social behaviors. In each session, 2 mice of same genotype and similar body weight housed in different cages were put in open field area during 10 min. The duration of sniffing and following social behaviors were recorded during the session.

Three-chamber sociability: The test allows the evaluation of social preference and discrimination using a specific apparatus (Stoelting, Dublin) with three successive and identical chambers (20cm x 40cm x 22(height) cm with 5cm x 8cm openings allowing access between the chambers). The testing protocol is performed as previously described (Moy et al., 2004). In the habituation, mice were allowed to explore the three chambers freely for 10 min. In the second phase, the test mouse was places in the central box, while an unfamiliar mouse (stranger 1) was put in one of the wire cages in a random and balanced manner. The doors were reopened and the test mouse was allowed to explore the chambers for 10 min. Time spent to explore empty cage and cage with stranger 1 was recorded. In the third phase, the social discrimination was evaluated with a new mouse placed into the empty wire cage and the test mouse was allowed to explore again the entire arena for 10 min, having the choice between the familiar mouse (stranger 1 or 1) and the novel mouse (stranger 2 or 2). A social discrimination index or social preference was defined as the percentage of time spent exploring the novel mouse as $(t_2/(t_1 + t_2)) \times 100$.

Rotarod: The test assesses sensorimotor coordination and balance. The apparatus (Biosed, France) is a rotating bar of 5 cm diameter (hard plastic materiel covered by grey rubber foam) on which mice are placed facing the direction of rotation. Animals were first habituated to stay on the rod for 30 seconds at a constant speed of 4 rotations per minute (rpm). This was followed by 3 training days with 4 trials per day. Mice were placed on an accelerating rod increasing from 4 rpm to 40 rpm in 5 min of time. The test was stopped when mice fell down

from the rod or when they made more than one passive rotation. The latency to fall and the maximal speed before falling was recorded.

Grip test: The test is used to evaluate muscular strength. Mice were first weighted and tested with a handy force gauge (Bioseb, France). Animals were placed on the instrument grid and pulled by the tail until letting go. The force (g) was related to animal weight (g).

Fat percentage and body size evaluation

Body composition by qNMR: Analysis was performed on of 20-weeks old animals. This procedure was used to give precise analysis of the body composition for fat content by nuclear magnetic resonance apparatus and Minispec+ analyser (Bruker, Germany). The test was conducted during light period on conscious fed mice.

Body size: Analysis was performed on of 20-weeks old animals. Animals were anesthetized under isoflurane and distances from snout to tail basis were measured with a FST ruler.

Hippocampal slices electrophysiology

Acute hippocampal slices were used to record field excitatory post synaptic potentials (fEPSPs), by the MEA60 electrophysiological suite (Multi Channel Systems, Reutlingen, FRG) as previously described (Coba et al., 2012; Kopanitsa et al., 2006). Eight set-ups consisting of a MEA1060-BC pre-amplifier and a filter amplifier (gain 550x) were run simultaneously by data acquisition units operated by MC_Rack software. Raw electrode data were digitized at 10 kHz and stored on a PC hard disk for subsequent analysis. To record fEPSPs, a hippocampal slice was placed into the well of 5x13 3D multi electrode array (MEA) biochip (Qwane Biosciences, Lausanne, Switzerland). The slice was guided to a desired position with a fine paint brush and gently fixed over MEA electrodes by a silver ring with attached nylon mesh lowered vertically by a one-dimensional U-1C micromanipulator (You Ltd, Tokyo, Japan). MEA biochips were fitted into the pre-amplifier case and fresh ACSF was delivered to the MEA well through a temperature-controlled perfusion cannula that warmed perfused media to 32°C. Monopolar stimulation of Schäffer collateral/commissural fibers through array electrodes was performed by STG4008 stimulus generator (Multi Channel Systems, Reutlingen, FRG). Biphasic (positive/negative, 100 µs/a

phase) voltage pulses were used. Amplitude, duration and frequency of stimulation were controlled by MC_Stimulus II software. All experiments were performed using two-pathway stimulation of Schäffer collateral/commissural fibers. Our previous experiments that utilized MEAs, demonstrated that largest LTP was recorded in proximal part of apical dendrites of CA1 pyramidal neurons (Kopanitsa et al., 2006). We have therefore picked a single principal recording electrode in the middle of proximal part of CA1 and assigned two electrodes for stimulation of the control and test pathways on the subicular side and on the CA3 side of *stratum radiatum* respectively. The distance from the recording electrode to the test stimulation electrode was 400-510 µm and to the control stimulation electrode 316-447 µm. To evoke orthodromic fEPSPs, test and control pathways were activated in succession at a frequency of 0.02 Hz. Baseline stimulation strength was adjusted to evoke a response that corresponded to 40% of the maximal attainable fEPSP at the recording electrode located in proximal *stratum radiatum*. Slope of the negative part of fEPSPs was used as a measure of the synaptic strength. Paired stimulation with an interpulse interval of 50 ms was used to observe paired-pulse facilitation (PPF) in baseline conditions in the test pathway before LTP induction. PPF was calculated by dividing the negative slope of fEPSP obtained in response to the second pulse by the amplitude of fEPSP amplitude evoked by the preceding pulse. To induce LTP, 10 bursts of baseline strength stimuli were administered at 5 Hz to test pathway with 4 pulses given at 100 Hz per burst (total 40 stimuli). LTP plots were scaled to the average of the first five baseline points. Normalisation of LTP values was performed by dividing the fEPSP amplitude in the tetanised pathway by the amplitude of the control fEPSP at corresponding time points. Normalised LTP values averaged across the period of 61-65 min after theta-burst stimulation were used for statistical comparison.

Craniofacial analysis

Craniums of 13 weeks old females obtained from *Del/+* cohort (n = 10 wt and 10 Del/+) and *Dup/+* cohort (n = 9 wt and 9 Dup/+) were stored in 100% ethanol. 3D coordinates of 38 cranial relevant landmarks were recorded using Landmark software and posterior comparisons were performed using the Euclidean distance matrix analysis (EDMA) with the software WinEDMA (version 1.0.1 beta). 3D data were converted into linear distances compiling into a matrix. Both the form (size of the skull) difference matrix (FDM) and the shape difference matrix (SDM) were analyzed. A ratio different from 1 (FDM) or 0 (SDM) for any linear distance indicates that the two samples are not similar for that measure.

Confidence intervals were estimated using a non-parametric bootstrapping algorithm. For each linear distance, the null hypothesis was rejected if the 90% confidence interval does not include 1 (FDM) or 0 (SDM): rejection of the null hypothesis enables localization of differences to specific landmarks and linear distances. Bootstrap Distributions of T (FDM) and Z (SDM) were calculated as follow: for each FDM is calculated a T value. The location of the T observed from the FDM, allows to calculate the probability (*p*) in this distribution.

Statistical analyses

Results were processed for statistical analysis using the Sigma Plot software (Sigma). All acquired behavioral data were analyzed using a one-way or a two-way ANOVA with a repeated measures followed by Student's t-test or Tukey's post-hoc test whenever it was appropriate. Otherwise, the non-parametric Kruskal-Wallis analysis and Mann-Whitney *U* test was done. One group Student t-tests were used to compare recognition index values to the chance level (50%). A Pearson's chi-squared test was used to evaluate mutant allele transmission. Data are represented as the mean ± SEM and the significant threshold was $p < 0.05$ or otherwise indicated.

In electrophysiological experiments input-output relationships were initially compared by mixed model repeated-measures ANOVA and Bonferroni *post hoc* test implemented in Prism 5 (GraphPad Software, Inc., San Diego, CA, USA) using individual slice data as independent observations. Since several slices were routinely recorded from every mouse, fEPSPmax, PPF and LTP values between wild-type and mutant mice were then compared using two-way nested ANOVA design with genotype (group) and mice (sub-group) as fixed and random factors correspondingly (STATISTICA v. 10, StatSoft, Inc., Tulsa, OK, USA). DF error was computed using Satterthwaite method and main genotype effect was considered significant if $P < 0.05$. Graph plots and normalisation were performed using OriginPro 8.5 (OriginLab, Northampton, MA, USA). Throughout the text, electrophysiological data are presented as mean ± s.e.m. with *n* and *N* indicating number of slices and mice respectively.

REFERENCES

Birling, M.C., Dierich, A., Jacquot, S., Hérault, Y., and Pavlovic, G. (2012). Highly-efficient, fluorescent, locus directed cre and FlpO deleter mice on a pure C57BL/6N genetic background. Genesis *50*, 482-489.

Brault, V., Pereira, P., Duchon, A., and Herault, Y. (2006). Modeling chromosomes in mouse to explore the function of genes, genomic disorders, and chromosomal organization. Plos Genetics *2*, 911-919.

Cai, Y., Jin, J., Swanson, S.K., Cole, M.D., Choi, S.H., Florens, L., Washburn, M.P., Conaway, J.W., and Conaway, R.C. (2010). Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex. J Biol Chem *285*, 4268-4272.

Canu, E., Boccardi, M., Ghidoni, R., Benussi, L., Testa, C., Pievani, M., Bonetti, M., Binetti, G., and Frisoni, G.B. (2009). H1 haplotype of the MAPT gene is associated with lower regional gray matter volume in healthy carriers. Eur J Hum Genet *17*, 287-294.

Coba, M.P., Komiyama, N.H., Nithianantharajah, J., Kopanitsa, M.V., Indersmitten, T., Skene, N.G., Tuck, E.J., Fricker, D.G., Elsegood, K.A., Stanford, L.E.*, et al.* (2012). TNiK is required for postsynaptic and nuclear signaling pathways and cognitive function. J Neurosci *32*, 13987-13999.

Cooper, G.M., Coe, B.P., Girirajan, S., Rosenfeld, J.A., Vu, T.H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V.*, et al.* (2011). A copy number variation morbidity map of developmental delay. Nature Genetics *43*, 838-U844.

Dias, J., Van Nguyen, N., Georgiev, P., Gaub, A., Brettschneider, J., Cusack, S., Kadlec, J., and Akhtar, A. (2014). Structural analysis of the KANSL1/WDR5/KANSL2 complex reveals that WDR5 is required for efficient assembly and chromatin targeting of the NSL complex. Genes Dev *28*, 929-942.

Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L., and Roeder, R.G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell *121*, 873-885.

Dubourg, C., Sanlaville, D., Doco-Fenzy, M., Le Caignec, C., Missirian, C., Jaillard, S., Schluth-Bolard, C., Landais, E., Boute, O., Philip, N.*, et al.* (2011). Clinical and molecular characterization of 17q21.31 microdeletion syndrome in 14 French patients with mental retardation. Eur J Med Genet *54*, 144-151.

Duchon, A., Pothion, S., Brault, V., Sharp, A.J., Tybulewicz, V.L.J., Fisher, E.M.C., and Herault, Y. (2011). The telomeric part of the human chromosome 21 from Cstb to Prmt2 is not necessary for the locomotor and short-term memory deficits observed in the Tc1 mouse model of Down syndrome. Behavioural Brain Research *217*, 271-281.

Ghidoni, R., Signorini, S., Barbiero, L., Sina, E., Cominelli, P., Villa, A., Benussi, L., and Binetti, G. (2006). The H2 MAPT haplotype is associated with familial frontotemporal dementia. Neurobiol Dis *22*, 357-362.

Goeldner, C., Reiss, D., Wichmann, J., Meziane, H., Kieffer, B.L., and Ouagazzal, A.M. (2008). Nociceptin receptor impairs recognition memory via interaction with NMDA receptor-dependent mitogen-activated protein kinase/extracellular signal-regulated kinase signaling in the hippocampus. J Neurosci *28*, 2190-2198.

Grayton, H.M., Fernandes, C., Rujescu, D., and Collier, D.A. (2012). Copy number variations in neurodevelopmental disorders. Progress in Neurobiology *99*, 81-91.

Grisart, B., Willatt, L., Destrée, A., Fryns, J.P., Rack, K., de Ravel, T., Rosenfeld, J., Vermeesch, J.R., Verellen-Dumoulin, C., and Sandford, R. (2009). 17q21.31 microduplication patients are characterised by behavioural problems and poor social interaction. J Med Genet *46*, 524-530.

Herault, Y., Rassoulzadegan, M., Cuzin, F., and Duboule, D. (1998). Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). Nature Genetics *20*, 381-384.

Ikegami, S., Harada, A., and Hirokawa, N. (2000). Muscle weakness, hyperactivity, and impairment in fear conditioning in tau-deficient mice. Neurosci Lett *279*, 129-132.

Kirchhoff, M., Bisgaard, A.M., Duno, M., Hansen, F.J., and Schwartz, M. (2007). A 17q21.31 microduplication, reciprocal to the newly described 17q21.31 microdeletion, in a girl with severe psychomotor developmental delay and dysmorphic craniofacial features. Eur J Med Genet *50*, 256-263.

Kitsiou-Tzeli, S., Frysira, H., Giannikou, K., Syrmou, A., Kosma, K., Kakourou, G., Leze, E., Sofocleous, C., Kanavakis, E., and Tzetis, M. (2012). Microdeletion and microduplication 17q21.31 plus an additional CNV, in patients with intellectual disability, identified by array-CGH. Gene *492*, 319-324.

Koolen, D.A., Kramer, J.M., Neveling, K., Nillesen, W.M., Moore-Barton, H.L., Elmslie, F.V., Toutain, A., Amiel, J., Malan, V., Tsai, A.C., *et al.* (2012). Mutations in the chromatin modifier gene KANSL1 cause the 17q21.31 microdeletion syndrome. Nat Genet *44*, 639-641.

Koolen, D.A., Sharp, A.J., Hurst, J.A., Firth, H.V., Knight, S.J., Goldenberg, A., Saugier-Veber, P., Pfundt, R., Vissers, L.E., Destrée, A., *et al.* (2008). Clinical and molecular delineation of the 17q21.31 microdeletion syndrome. J Med Genet *45*, 710-720.

Koolen, D.A., Vissers, L., Pfundt, R., de Leeuw, N., Knight, S.J.L., Regan, R., Kooy, R.F., Reyniers, E., Romano, C., Fichera, M., *et al.* (2006). A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. Nature Genetics *38*, 999-1001.

Kopanitsa, M.V., Afinowi, N.O., and Grant, S.G. (2006). Recording long-term potentiation of synaptic transmission by three-dimensional multi-electrode arrays. BMC Neurosci *7*, 61.

Li, Y., Chen, J.A., Sears, R.L., Gao, F., Klein, E.D., Karydas, A., Geschwind, M.D., Rosen, H.J., Boxer, A.L., Guo, W., *et al.* (2014). An epigenetic signature in peripheral blood associated with the haplotype on 17q21.31, a risk factor for neurodegenerative tauopathy. PLoS Genet *10*, e1004211.

Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J.*, et al.* (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila. Mol Cell *21*, 811-823.

Moy, S.S., Nadler, J.J., Perez, A., Barbaro, R.P., Johns, J.M., Magnuson, T.R., Piven, J., and Crawley, J.N. (2004). Sociability and preference for social novelty in five inbred strains: an approach to assess autistic-like behavior in mice. Genes Brain Behav *3*, 287-302.

Permuth-Wey, J., Lawrenson, K., Shen, H.C., Velkova, A., Tyrer, J.P., Chen, Z., Lin, H.Y., Chen, Y.A., Tsai, Y.Y., Qu, X.*, et al.* (2013). Identification and molecular characterization of a new ovarian cancer susceptibility locus at 17q21.31. Nat Commun *4*, 1627.

Raja, S.J., Charapitsa, I., Conrad, T., Vaquerizas, J.M., Gebhardt, P., Holz, H., Kadlec, J., Fraterman, S., Luscombe, N.M., and Akhtar, A. (2010). The nonspecific lethal complex is a transcriptional regulator in Drosophila. Mol Cell *38*, 827-841.

Skarnes, W.C., Rosen, B., West, A.P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A.O., Thomas, M., Harrow, J., Cox, T.*, et al.* (2011). A conditional knockout resource for the genome-wide study of mouse gene function. Nature *474*, 337-342.

Steinberg, K.M., Antonacci, F., Sudmant, P.H., Kidd, J.M., Campbell, C.D., Vives, L., Malig, M., Scheinfeldt, L., Beggs, W., Ibrahim, M.*, et al.* (2012). Structural diversity and African origin of the 17q21.31 inversion polymorphism. Nat Genet *44*, 872-880.

Tan, T.Y., Aftimos, S., Worgan, L., Susman, R., Wilson, M., Ghedia, S., Kirk, E.P., Love, D., Ronan, A., Darmanian, A.*, et al.* (2009). Phenotypic expansion and further characterisation of the 17q21.31 microdeletion syndrome. J Med Genet *46*, 480-489.

Tang, S.H.E., Silva, F.J., Tsark, W.M.K., and Mann, J.R. (2002). A Cre/loxP-deleter transgenic line in mouse strain 129S1/SvImJ. Genesis *32*, 199-202.

Wang, X.D., Chen, Y., Wolf, M., Wagner, K.V., Liebl, C., Scharf, S.H., Harbich, D., Mayer, B., Wurst, W., Holsboer, F.*, et al.* (2011). Forebrain CRHR1 deficiency attenuates chronic stress-induced cognitive deficits and dendritic remodeling. Neurobiol Dis *42*, 300-310.

Zody, M.C., Jiang, Z., Fung, H.C., Antonacci, F., Hillier, L.W., Cardone, M.F., Graves, T.A., Kidd, J.M., Cheng, Z., Abouelleil, A.*, et al.* (2008). Evolutionary toggling of the MAPT 17q21.31 inversion region. Nat Genet *40*, 1076-1083.

Zollino, M., Orteschi, D., Murdolo, M., Lattante, S., Battaglia, D., Stefanini, C., Mercuri, E., Chiurazzi, P., Neri, G., and Marangi, G. (2012). Mutations in KANSL1 cause the 17q21.31 microdeletion syndrome phenotype. Nat Genet *44*, 636-638.

LEGENDS TO FIGURES

Figure 1. Mouse models for 17q21.31 rearrangements. **(A)** Top: haplotype H1 of the human 17q21.31 region. All genomic positions are given according to UCSC human genome browser GRCh38/hg38. Bottom: 17q21.31 syntenic region on mouse chromosome 11E1. All genomic positions are given according to UCSC mouse genome browser GRCm38/mm10. **(B)** Strategy for in vivo cre-mediated recombination and crossing strategy. LoxP sites were inserted upstream of *Arf2* and downstream of *Kanl1*. Breeding were done in order to have in one trans-loxer female expressing the Hprt<tm1(cre)Mnn> transgene and carrying the two loxP sites in a *trans* configuration. Last step consisted in mating trans-loxer females with wild-type males to generate, in the progeny, individuals carrying the deletion or the duplication of the *Arf2-Kansl1* region. *Del/+* and *Dup/+* animals were crossed together to generate *Del/Dup* cohorts. **(C)** Molecular validation. PCR products specific for the *Del* and *Dup* alleles are 448-bp and 653-bp long respectively. **(D)** Evolution of body weight (g). **(E)** Body fat percentage of 20-weeks old animals measured by qNMR. **(F)** Body length (distance from snout to tail basis) of 20-weeks animals. Data are mean ± SEM. **(A)** Repeated Measures ANOVA "genotype" analysis. **(B, C)** Kruskal-Wallis analysis, Mann-Whitney *U* test. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{\#}$significantly different from all other groups.

Figure 2. Behavioral characterization of the *Del-Dup* cohort. **(A)** Circadian activity test. The light-light cycle was controlled as 12-hours light and 12-hours dark (lights on at 7 am). Graphs plot spontaneous locomotor activity (count) and rearing behavior (count) during dark and light phases. **(B)** Open field test. Exploratory locomotor activity (distance travelled in m) and rearing behavior (count) over 30 min of test. **(C)** Motor coordination evaluation. The graph plots latency (s) that mice stayed on the rod before falling under acceleration speed (4-40 rpm in 5 minutes) over 3 consecutive days of test. **(D-E)** Three-chamber sociability test. **(D)** Sniffing time (s) of the first congener in the social interest session (session 1) and sniffing time of familiar and novel congener in the social discrimination session (session 2). No delay was used between the two sessions. **(E)** Preference index was calculated as the percentage of time spent interacting with the stranger 2 in the second session of test. **(F)** Novel object recognition task. Recognition index reflects the ability of mice to distinguish the novel object from the familiar object after a retention delay of 3 hours. The dashed line materializes a chance level of 50%. All genotypes performed significantly above the chance level. **(G)** Fear conditioning test. Graphs plot the percentage of freezing during sessions of test. The 6-min

context session was run 24-hours after conditioning. The 8-min cue session was performed 5-hours after the context session. A sequence of 2-min with no cue and 2 min with light/auditory conditioning stimulus was repeated two times. Data are mean ± SEM. One-way ANOVA analysis, Tukey's test otherwise Kruskal-Wallis analysis, Mann-Whitney *U* test. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{\#}$significantly different from all other groups.

Figure 3. Effect of *Arf2-Kansl1* dosage on electrophysiological parameters measured in the CA1 area of hippocampal slices. **(A-B)** Baseline synaptic transmission. **(A)** Transmission was significantly decreased in *Del/+* mice ($F_{(9, 531)} = 7.59$; $P < 0.0001$). Input-output relationships illustrate averaged fEPSP slopes in slices from *Del/+* ($n = 28$; $N = 9$) and wt mice ($n = 33$; $N = 8$) in response to stimulation of Schäffer collaterals by biphasic voltage pulses of 0.1 – 4.2 V. Representative families of fEPSP traces are given on the right side. **(B)** Transmission was slightly enhanced in *Dup/+* mice ($F_{(9, 468)} = 3.54$; $P = 0.0003$). Input-output relationships illustrate averaged fEPSP slopes in slices from *Dup/+* ($n = 27$; $N = 9$) and wt mice ($n = 27$; $N = 10$). **(C-D)** Paired-pulse facilitation. **(C)** Paired-pulse facilitation was slightly but significantly decreased in *Del/+* animals ($n = 28$; $N = 9$) as compared to their wt littermates ($n = 33$; $N = 8$; $F_{(1, 11.04)} = 6.506$; $P = 0.027$). Representative fEPSP sweeps are presented on the right side. **(D)** Paired-pulse facilitation was nominally higher in *Dup/+* animals ($n = 27$; $N = 9$) as compared to their wt littermates ($n = 27$; $N = 10$), but the effect did not reach statistical significance ($F_{(1, 4.92)} = 5.402$; $P = 0.069$). **(E-F)** Theta-burst stimulation elicited pathway-specific long-term potentiation of synaptic transmission in hippocampal CA1 area. **(E)** Normalised magnitude of this potentiation 60-65 min after LTP induction did not differ in *Del/+* mice (163 ± 4%; $n = 27$; $N = 9$; $P = 0.88$) relatively to their wt counterparts (165 ± 8%; $n = 33$; $N = 8$). Examples of test pathway fEPSP traces immediately before and 1 h after theta-burst stimulation are given on the right side. **(F)** Normalised magnitude of this potentiation 60-65 min after LTP induction did not differ in *Dup/+* mice (166 ± 7%; $n = 27$; $N = 9$; $P = 0.29$) relatively to their wt counterparts (154 ± 6%; $n = 26$; $N = 10$)

Figure 4. Effect of *Arf2-Kansl1* dosage on craniofacial parameters. **(A)** Representative reconstructed 3D skull images and landmark distribution. Euclidian distances between the different landmaks allowed calculation of both the form (or size) difference (FD) and the shape difference (SD). Analysis revealed no size alteration of Del/+ **(B)** and Dup/+ **(D)** skulls. Alteration phenotypes near from statistical significance were found for the skull shape of *Del/+* **(C)** and *Dup/+* **(E)** animals.

Figure 5. Behavioral characterization of the *Kansl1*$^{+/-}$ cohort. **(A)** Circadian activity test. The light-light cycle was controlled as 12-hours light and 12-hours dark (lights on at 7 am). Graphs plot spontaneous locomotor activity (count) and rearing behavior (count) during dark and light phases. **(B)** Open field test. Exploratory locomotor activity (distance travelled in m) and rearing behavior (count) over 30 min of test. **(C)** Repetitive behaviors. The graph plots the occurrences of rearing, jumping, climbing and digging behaviors during 10 min of observation in a novel home-cage. **(D)** Motor coordination evaluation. The graph plots latency (s) that mice stayed on the rod before falling under acceleration speed (4-40 rpm in 5 minutes) over 3 consecutive days of test. **(E)** Novel object recognition task. Recognition index reflects the ability of mice to distinguish the novel object from the familiar object after retention delay of 3 hours. The dashed line materializes a chance level of 50%. All genotypes performed significantly above the chance level. **(F)** Fear conditioning test. Graphs plot the percentage of freezing during sessions of test. The 6-min context session was run 24-hours after conditioning. The 8-min cue session was performed 5-hours after the context session. A sequence of 2-min with no cue and 2 min with light/auditory conditioning stimulus was repeated two times. Data are mean ± SEM. One-way ANOVA analysis, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Supplementary Figure 1. Spontaneous locomotor activity and feeding behavior of the *Del-Dup* cohort during the circadian activity test. Patterns of locomotor activity **(A)** and rearing behavior **(B)** profiles during 32-hours of test. **(C-D)** Feeding behaviors. Water **(C)** and pellet **(D)** consumption during the 32-hours of test. **(E)** Pellets loss by animals which passed through the holed ground. All graphs depict mean ± SEM.

Supplementary Figure 2. Social behaviors of the *Del-Dup* cohort. **(A-D)** Three-chamber sociability test. Time (s) spent in the different chambers **(A, B)** and time spent sniffing cages **(C, D)**. In **A** and **B**, chamber 1 contained a mouse (familiar) and chamber 2 is empty. In **C** and **D**, a new mouse (novel) was placed in chamber 2. **(E-G)** Social interaction test. 2 mice of same genotype from different home-cages were put in open field area during 10 min. Graph plots the latency to first contact **(E)**, and the duration of sniffing **(F)** and following **(G)** social behaviors. All graphs depict mean ± SEM. Sudent's t-test, $^{*}p < 0.05$.

Supplementary Figure 3. Spatial learning and memory performances of the *Del-Dup* cohort in the Morris water maze test. **(A)** Distance travelled to find the platform along acquisition (A1 – A6) and reversal (Cued1) sessions. **(B)** Corresponding latencies to find the platform.

(C-D) Removal session. Mice are scored for the percentage of time spent in the different quadrant of the MWM **(C)** and the annulus crossing **(D)** during one single trial where the platform was removed. All graphs depict mean ± SEM.

Supplementary Figure 4. Body mass and size and adiposity of *Kansl1*$^{+/-}$ cohort. **(A)** Evolution of body weight (g). **(B)** Body fat percentage of 20-weeks old animals measured by qNMR. **(C)** Body length (distance from snout to tail basis) of 20-weeks animals. Data are mean ± SEM. **(A)** Repeated Measures ANOVA "genotype" analysis. **(B, C)** Sudent's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$.

Supplementary Figure 5. Spontaneous locomotor activity and feeding behavior of *Kansl1*$^{+/-}$ cohort during circadian activity test. Patterns of locomotor activity **(A)** and rearing behavior **(B)** profiles during 32-hours of test. **(C-D)** Feeding behaviors. Water **(C)** and pellet **(D)** consumption during the 32-hours of test. **(E)** Pellets loss by animals which passed through the holed ground. All graphs depict mean ± SEM.

FIGURES

Figure 1. Mouse models for 17q21.31 rearrangements

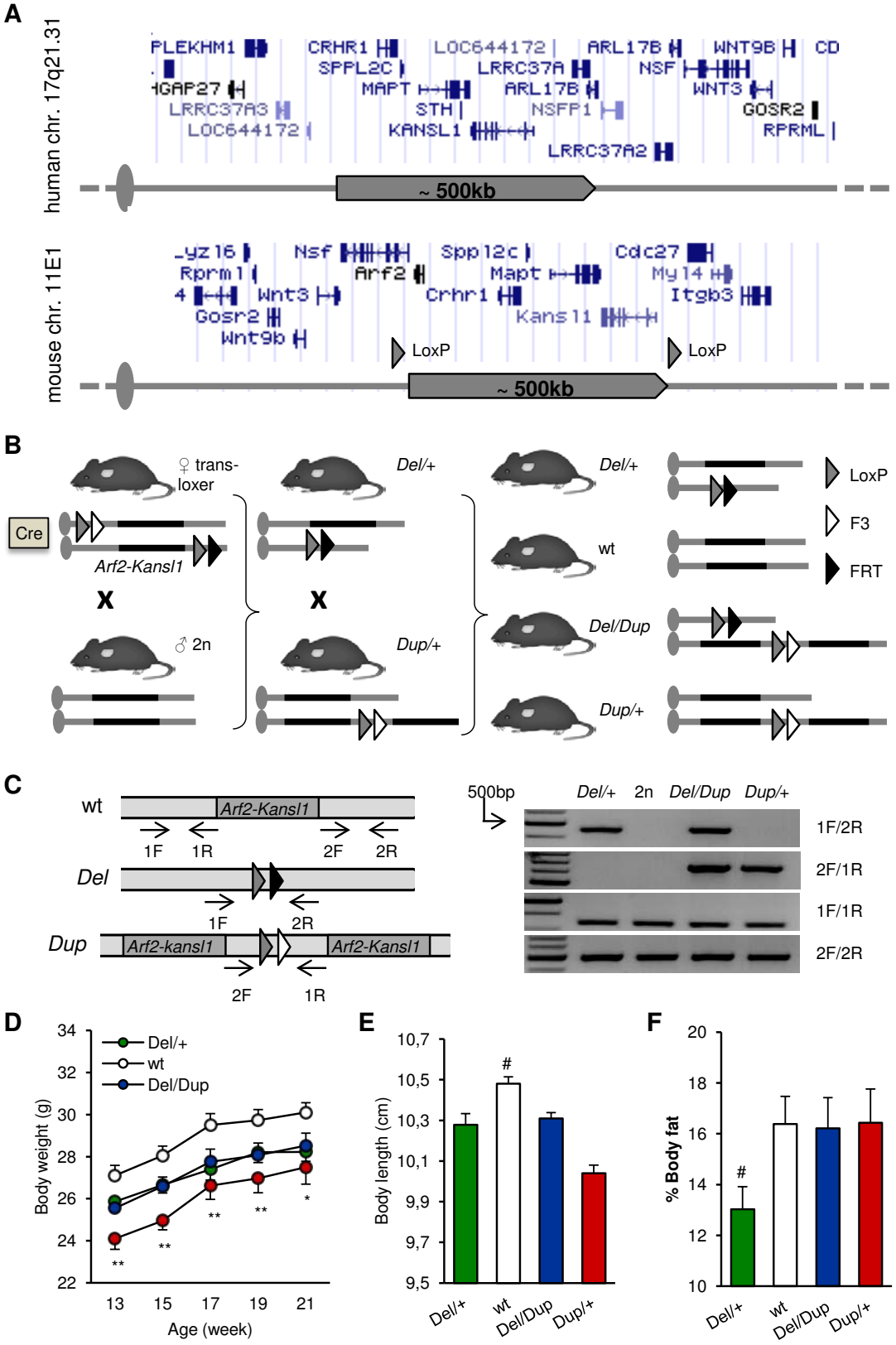


Figure 2. Behavioral characterization of the *Del-Dup* cohort

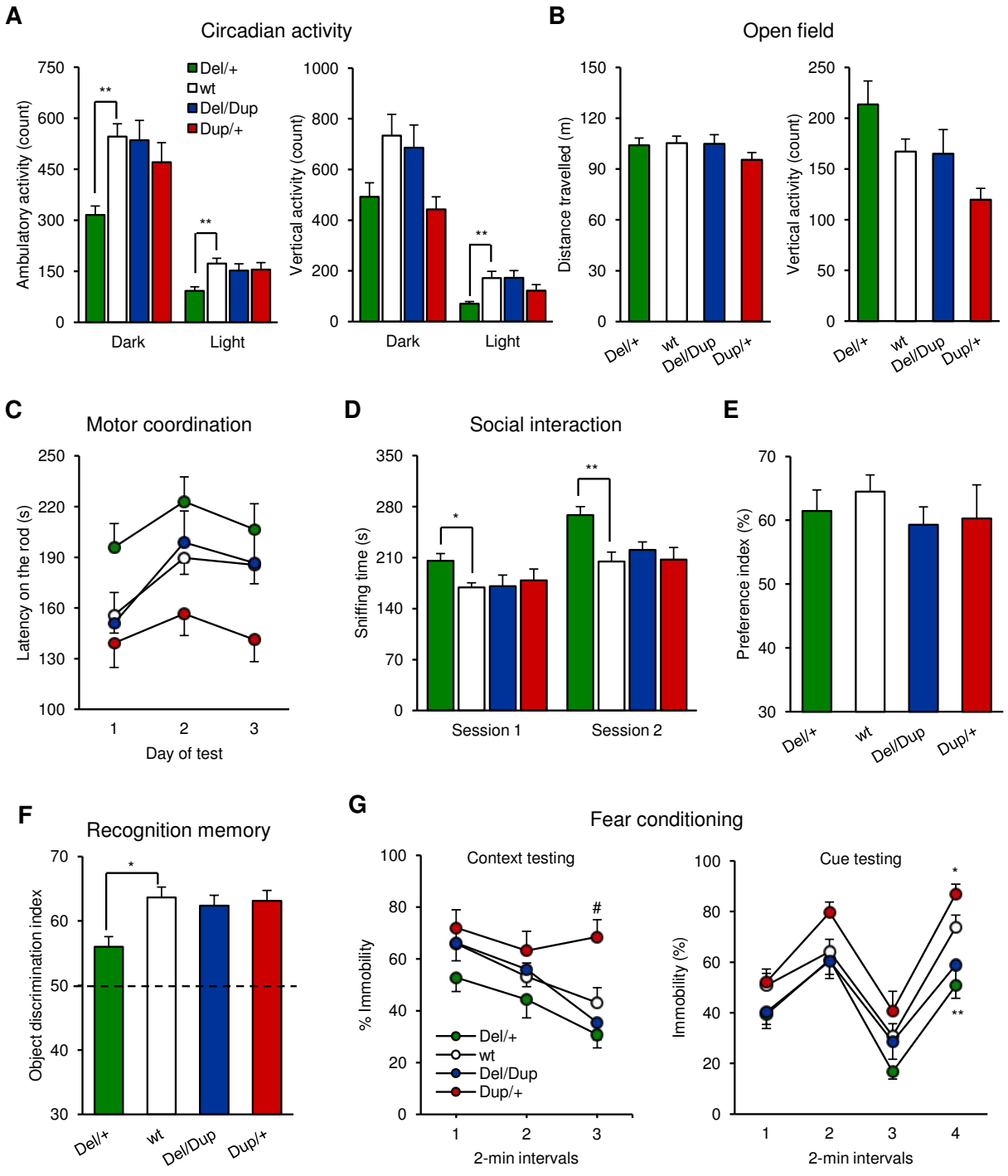


Figure 3. Effect of *Arf2-Kansl1* dosage on electrophysiological parameters measured in the CA1 area of hippocampal slices

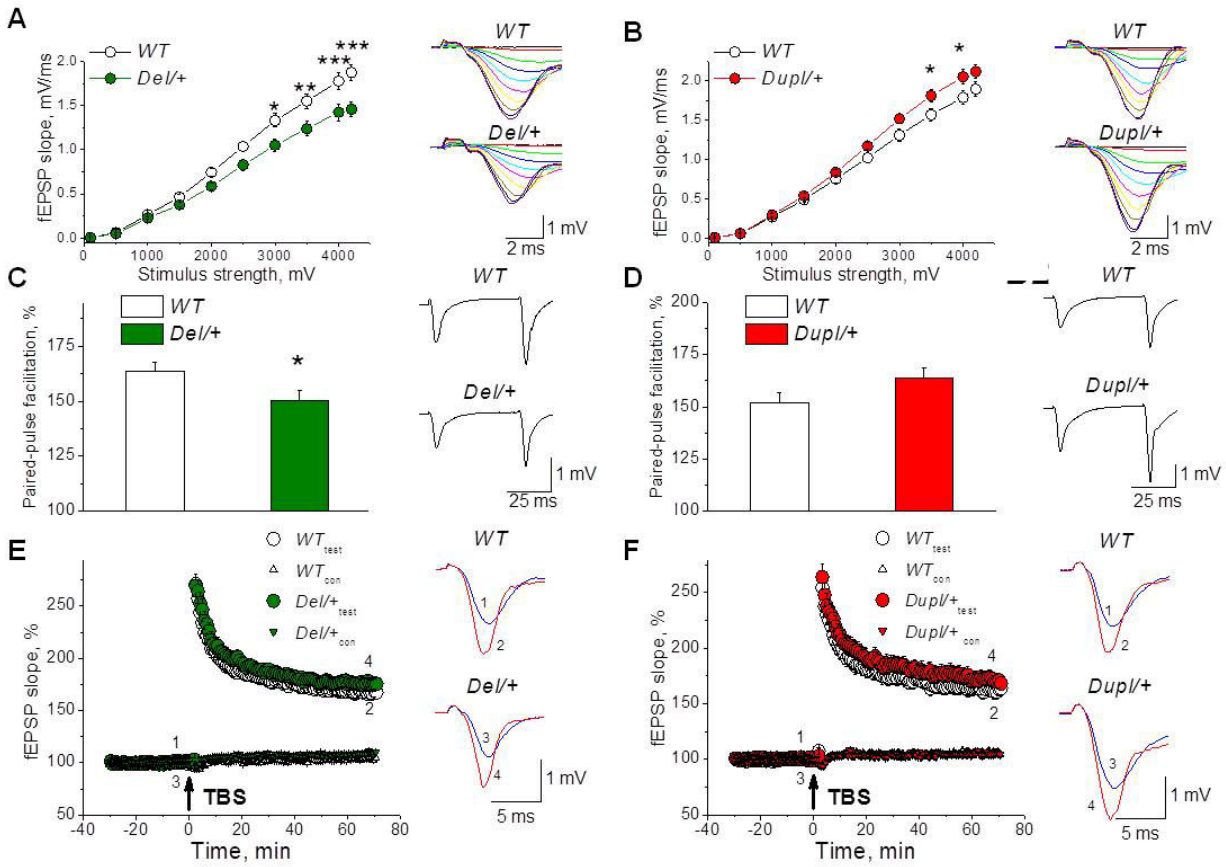


Figure 4. Effect of *Arf2-Kansl1* dosage on craniofacial parameters

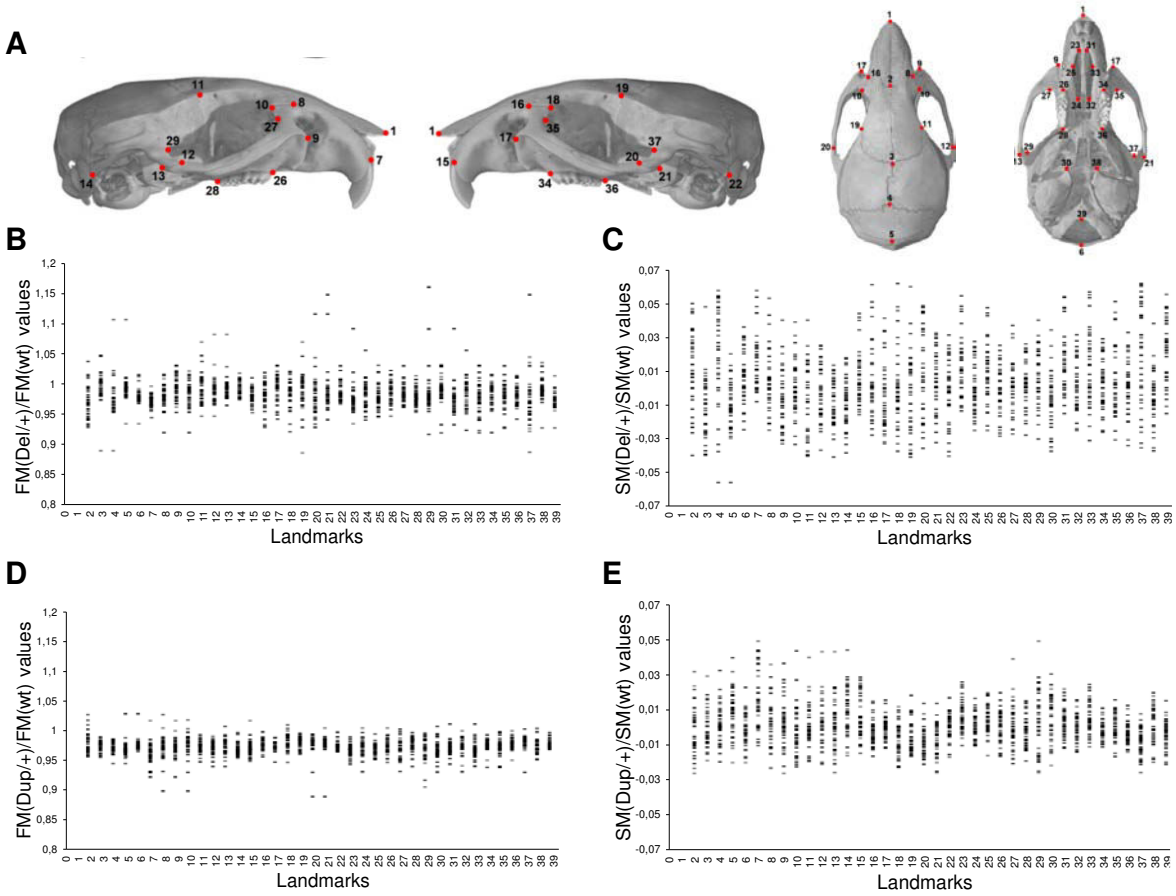
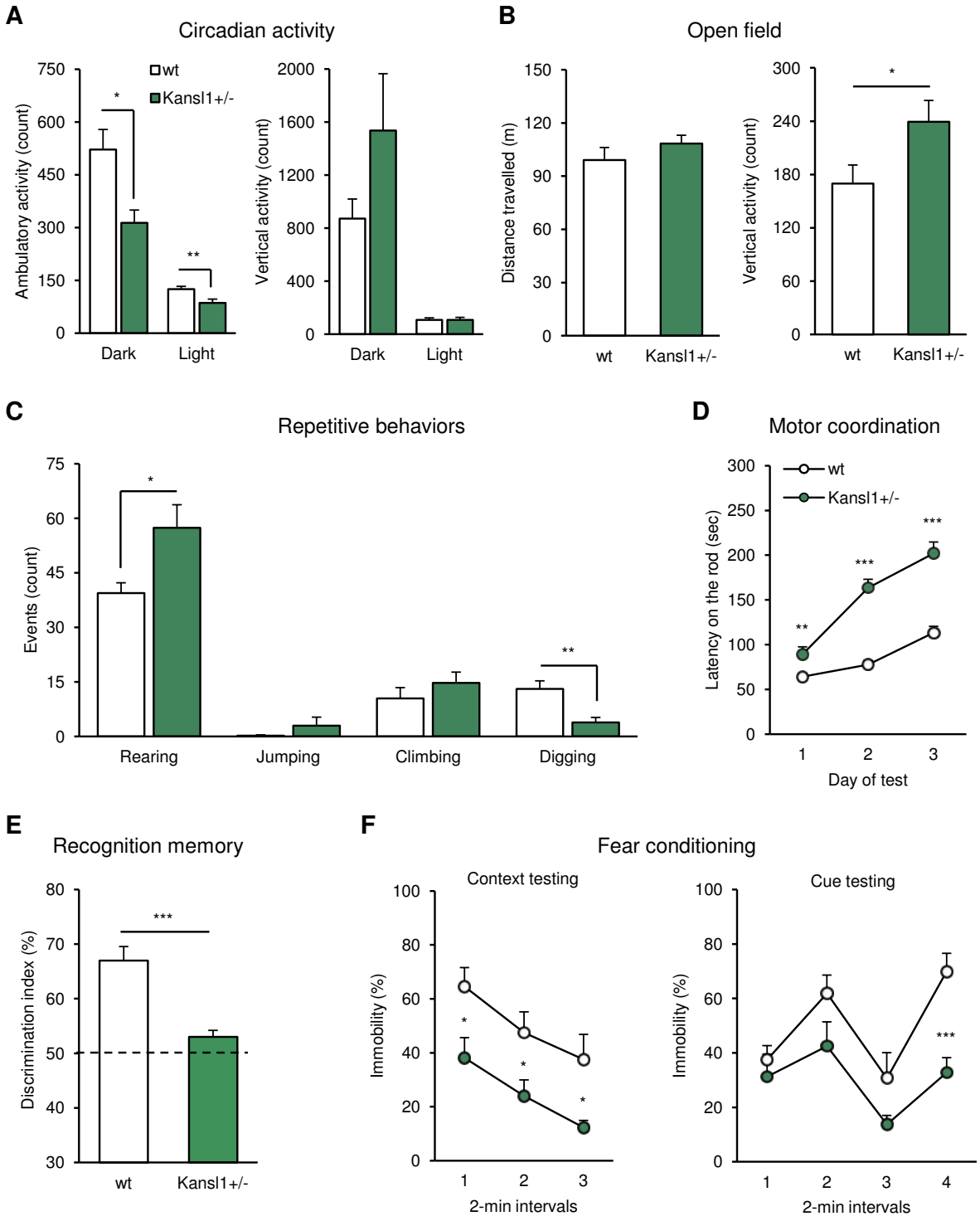


Figure 5. Behavioral characterization of the *Kansl1*$^{+/-}$ cohort



TABLES

Crossing	Genotype	Observed number	Observed ratio	χ^2	*p*
Del/+ x wt	wt	102	63.4%	5.74	0.02
	Del/+	59	36.6%		
Dup/+ x wt	2n	186	55.0%	1.71	0.20
	Dup/+	152	45.0%		
Del/+ x *Dup*/+	wt	71	37.6%	11.9	$5.5x10^{-04}$
	Del/+	14	7.4%	23.4	$1.3x10^{-06}$
	Dup/+	48	25.4%	0.01	0.91
	Del/Dup	56	29.6%	1.62	0.20

Table 1. Transmission rates of the *Arf2-Kansl1* deletion (Del) and duplication (Dup) alleles observed at weaning. A reduced transmission rate is observed for the Del allele. Combination of both Del and Dup alleles leads to rescue for this lethality.

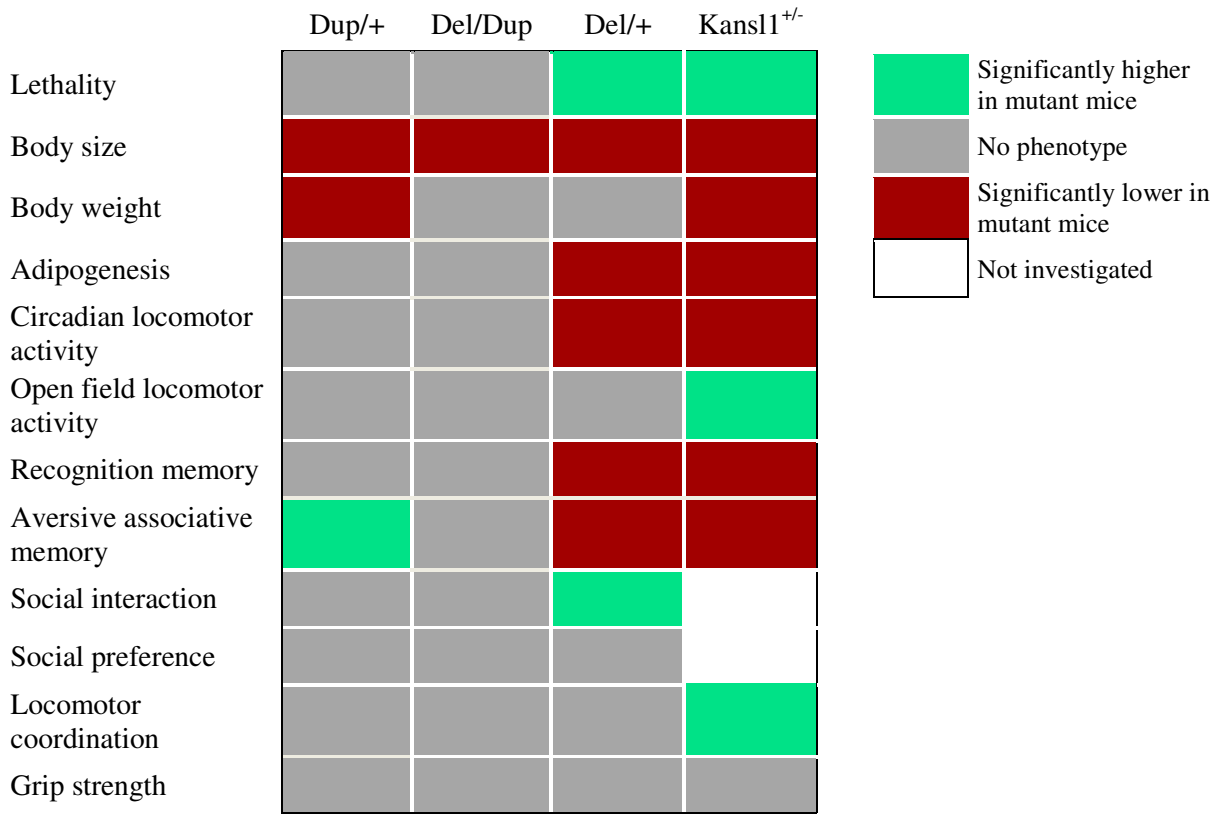
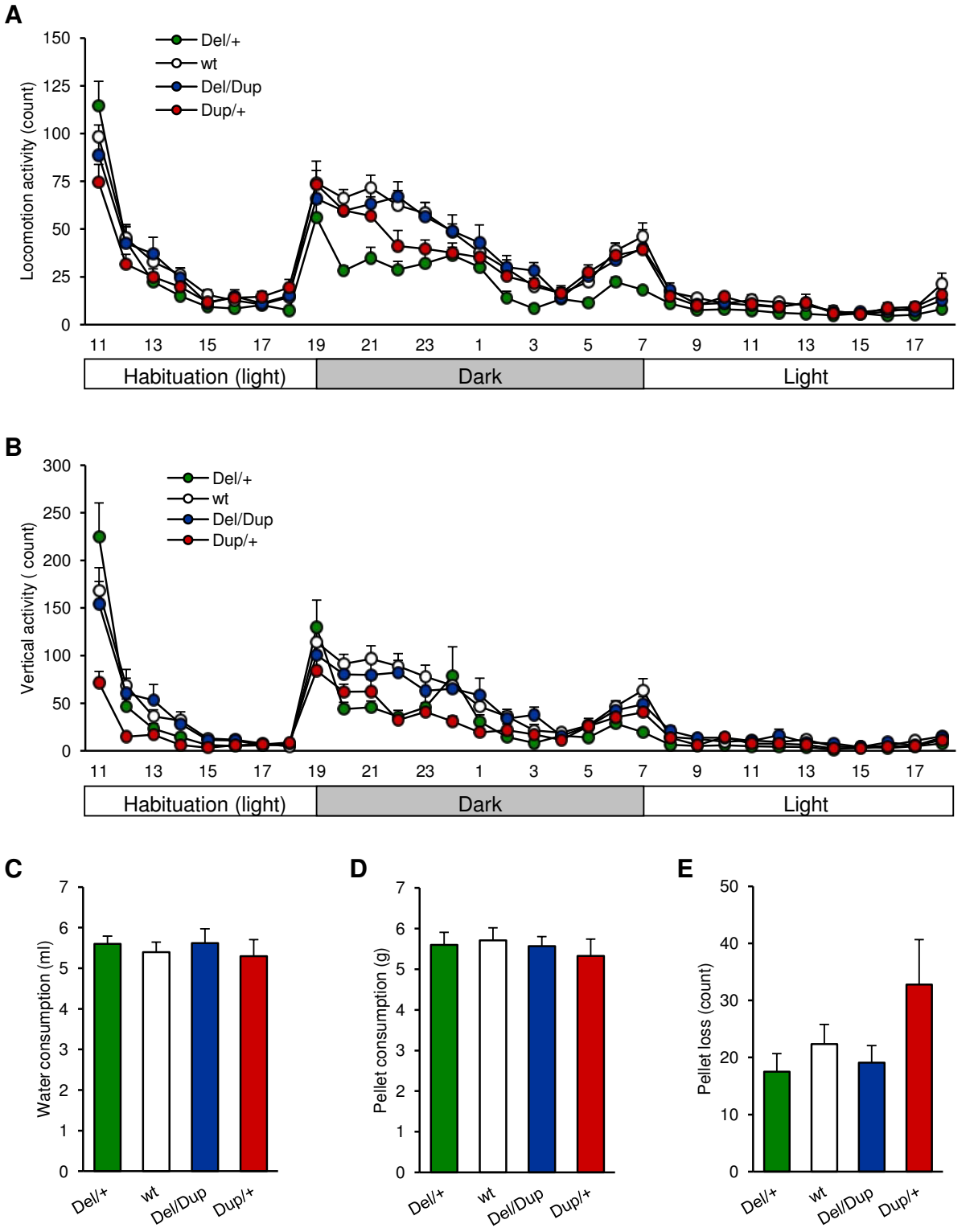


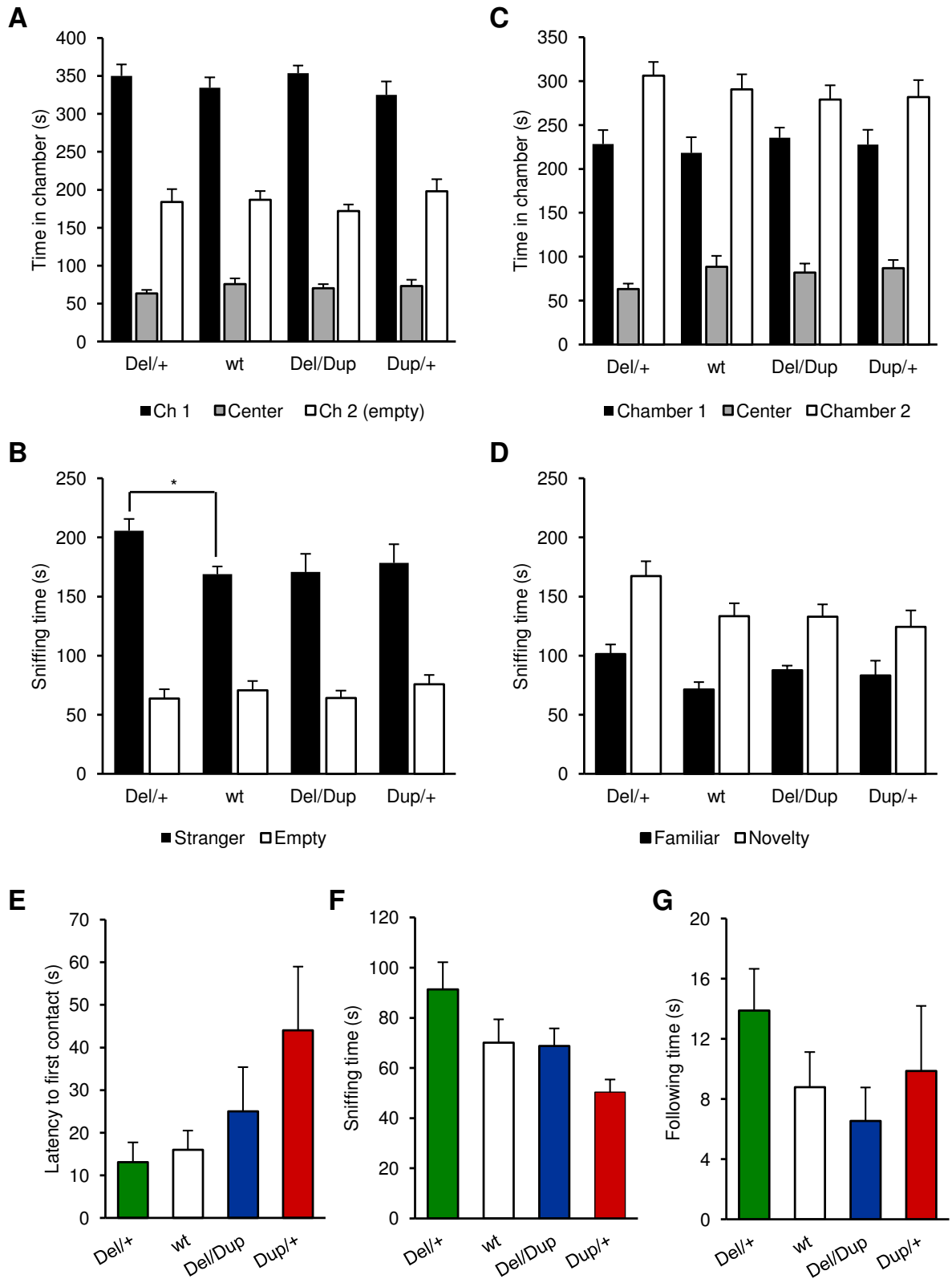
Table 2. Map of phenotypes observed in mice carrying deletion and/or duplication for the *Arf2-Kansl1* genetic interval and heterozygous *Kansl1* mutant mice in comparison with wild-type littermates. Few phenotypes were found in *Dup/+* mice showing body weight and size reduction and associative memory improvements. Only one phenotype was observed for *Del/Dup* consisting in body size reduction. Interestingly, *Del/+* mice and *Kansl1+/-* mice displayed an important number of similar phenotypes including a high lethality, a reduction of body size, adipogenesis deficits, spontaneous hypoactivity and impairments for recognition memory and aversive associative memory. *Kansl1*$^{+/-}$ animals showed body weight reduction and locomotor coordination improvements while *Del/+* animals showed trends for these phenotypes. *Kanl1+/-* mice also displayed vertical hyperactivity in open field and in odorless home-cages, a phenotype which was not observed for *Del/+* mice. Finally, *Del/+* mice also showed an increased level for social interaction. Due to a small number of mutant animals including some isolated mice, social behaviors have not been investigated yet for the *Kansl1+/-* cohort.

SUPPLEMENTARY INFORMATIONS

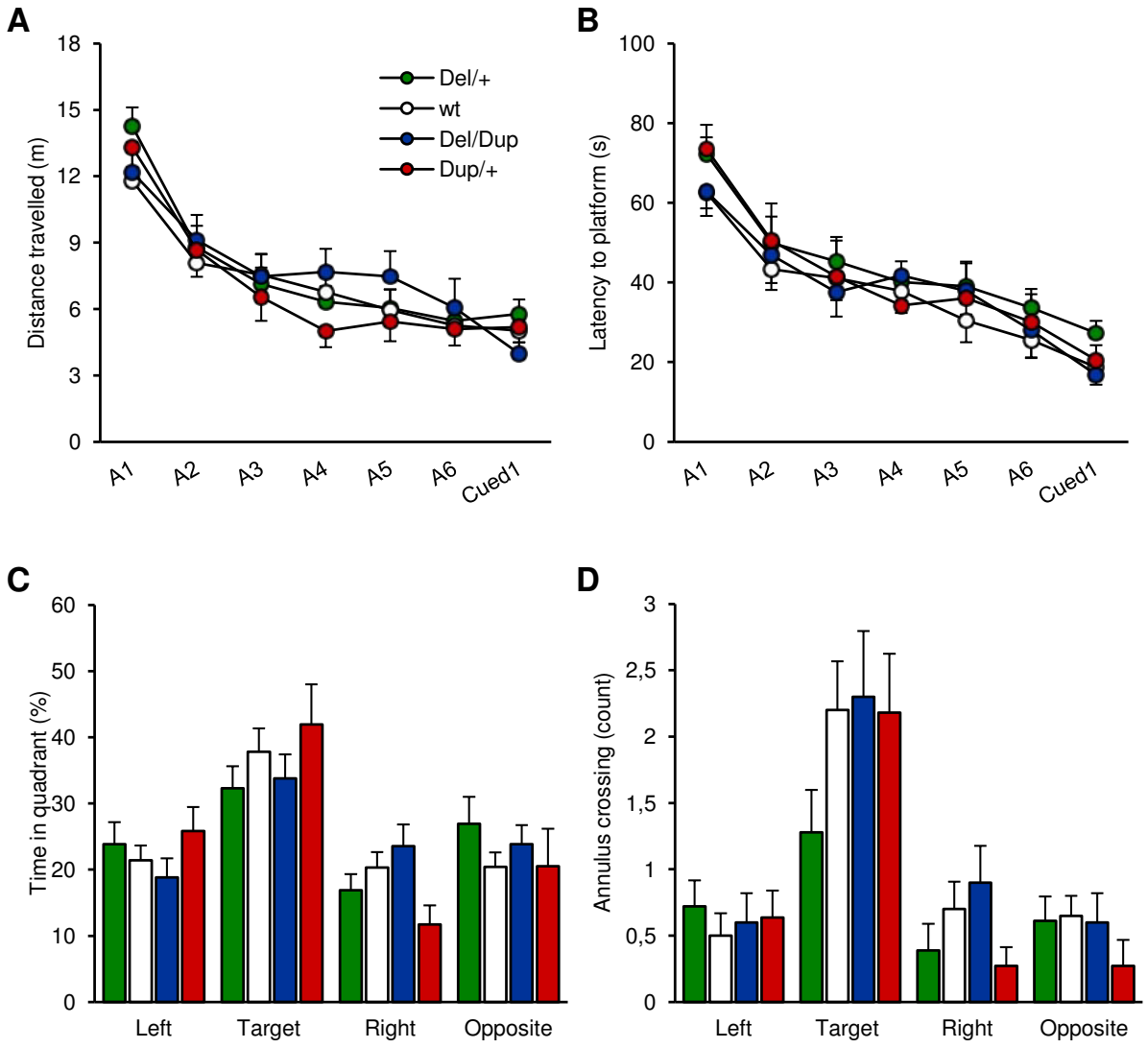
Supplementary Figure 1. Spontaneous locomotor activity and feeding behavior of the *Del-Dup* cohort in the circadian activity test



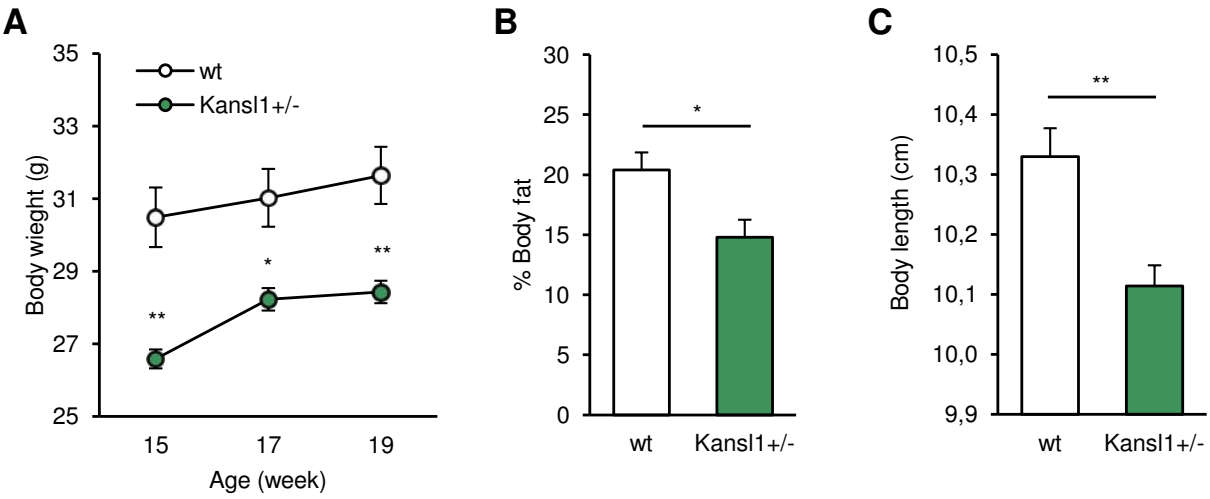
Supplementary Figure 2. Social behaviors of the *Del-Dup* cohort



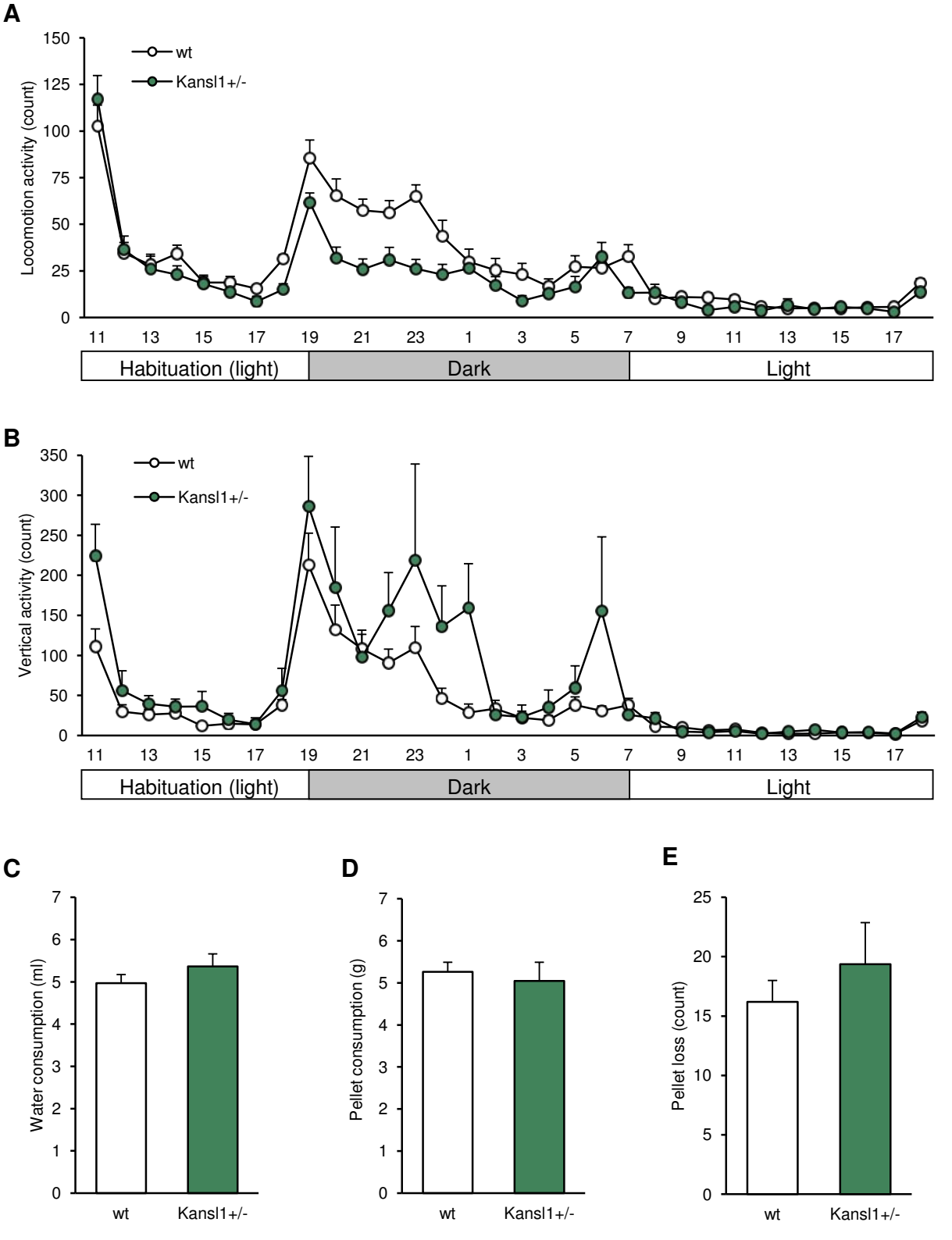
Supplementary Figure 3. Spatial learning and memory performances of the *Del-Dup* cohort in the Morris water maze test.



Supplementary Figure 4. Body mass and size and adiposity of *Kansl1*$^{+/-}$ cohort



Supplementary Figure 5. Spontaneous locomotor activity and feeding behavior of *Kansl1*$^{+/-}$ cohort during circadian activity test.



Test	Parameter	Genotype			
		Del/+	**wt**	***Del/Dup***	***Dup/+***
Circadian Activity	Hab ambulatory activity (count)	223 ± 28	241± 17	230 ± 35	191 ± 23
	Hab vertical activity (count)	326 ± 41	337 ± 58	324 ± 64	125 ± 21
	Dark ambulatory activity (count)	316 ± 26**	546 ± 37	535 ± 58	470 ± 58
	Dark vertical activity (count)	492 ± 56	734 ± 84	686 ± 90	443 ± 50
	Light ambulatory activity (count)	92.4 ± 12**	172 ± 16	152 ± 20	155 ± 21
	Light vertical activity (count)	70.6 ± 8.6**	172 ± 27	173 ± 29	122 ± 23
	Total ambulatory activity (count)	639 ± 56***	974 ± 48	932 ± 104	836 ± 94
	Total vertical activity (count)	895 ± 78	1248 ± 145	1190 ± 165	699 ± 76*
	Total food consumption (g)	5.6 ± 0.3	5.7 ± 0.3	5.6 ± 0.2	5.3 ± 0.4
	Total water consumption (ml)	5.6 ± 0.2	5.4 ± 0.3	5.6 ± 0.4	5.3 ± 0.4
Open Field	Locomotor exploration (m)	104 ± 4	105 ± 4	105 ± 5	95.4 ± 4.3
	Rears (count)	214 ± 23	167 ± 12	165 ± 24	120 ± 11
	Time in center (%)	7.0 ± 0.7	8.7 ± 1.4	8.4 ± 1.2	10.2 ± 1.2
Elevated Plus Maze	Arm entries (count)	11.8 ± 1.2	14.2 ± 0.9	15.2 ± 1.1	10.9 ± 1.2
	Open arm time (%)	2.7 ± 0.8	5.7 ± 0.9	7.9 ± 1.7	4.7 ± 1.1

Supplementary Table 1. Behavioral characterization of the *Del-Dup* cohort. Circadian activity test revealed vertical hypoactivity for both mutant mice. *Del/+* mice showed ambulatory hypoactivity during dark and light phases and vertical hypoactivity during light phase. *Dup/+* mice showed vertical hypoactivity during habituation phase. Globally, *Del/+* and *Dup/+* mice showed respectively ambulatory and vertical hypoactivity. No alteration of feeding behavior was noticed during the test. In the open field test, *Del/+* and *Dup/+* mice showed respectively trends for rearing beharior increase and decrease. No phenotype was observed in the elevated plus maze test. For all these tests, *Del/Dup* mice showed similar activity and behaviors than wild-type mice. Data are mean ± SEM. One-way ANOVA analysis, Tuckey's test otherwise Kruskal-Wallis analysis, Mann-Whitney *U* test. Stars indicate significant difference with wild-type littermates. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Test	Parameter	Genotype			
		Del/+	**wt**	***Del/Dup***	***Dup/+***
Y Maze	Arm entries (count)	25.1 ± 2.1	22.5 ± 2.6	20.0 ± 2.1	15.8 ± 2.1
	Alternation (%)	64.9 ± 5.9	64.6 ± 3.7	63.8 ± 4.9	63.8 ± 5.9
New Object Recognition 3 hour delay	S1 object A exploration (s)	9.0 ± 0.6	9.3 ± 0.8	8.1 ± 1.0	6.2 ± 0.9
	S2 object A exploration (s)	4.6 ± 0.6	3.4 ± 0.7	2.7 ± 0.5	2.3 ± 0.7
	S2 object B exploration (s)	6.3 ± 1.2	6.0 ± 1.2	4.5 ± 0.7	3.4 ± 0.5
	Discrimination index (%)	56.0 ± 1.6*	63.7 ± 1.9	62.4 ± 2.8	63.1 ± 3.3
Morris water maze	D6 distance to platform (m)	5.5 ± 0.6	5.3 ± 0.7	6.1 ± 1.3	5.1 ± 0.7
	D6 latency to platform (s)	33.7 ± 4.7	25.6 ± 4.4	28.0 ± 7.0	30.1 ± 7.0
	Probe - time in target quad. (%)	32.3 ± 3.3	37.8 ± 3.5	33.8 ± 3.6	41.9 ± 6.0
Fear Conditioning	Baseline freezing (%)	11.4 ± 2.1	15.5 ± 3.4	12.7 ± 3.4	25.1 ± 6.9
	Post-choc freezing (%)	19.9 ± 3.3	30.7 ± 5.6	15.6 ± 3.1	31.7 ± 2.8
	Contextual freezing (%)	42.7 ± 5.5	54.1 ± 5.0	52.6 ± 5.9	67.9 ± 5.9
	First Cue freezing (%)	60.9 ± 5.7	64.2 ± 4.8	60.5 ± 6.9	79.7 ± 4.1
	Second cue freezing (%)	50.8 ± 5.1**	73.8 ± 4.7	59.0 ± 7.1	87.0 ± 3.9*
Three-Chamber Sociability	S1 First stranger exploration (s)	206 ± 10*	169 ± 6	171 ± 15	179 ± 16
	S2 First stranger exploration (s)	101 ± 8	71.4 ± 6.3	87.6 ± 4.0	83.1 ± 12.6
	S2 Second stranger exploration (s)	167 ± 13	133 ± 11	133 ± 11	124 ± 14
	Preference index (%)	61.4 ± 3.3	64.5 ± 2.6	59.3 ± 2.8	60.3 ± 5.3
Social Interaction	First contact latency (s)	13.1 ± 4.6	16.0 ± 4.5	25.0 ± 10.4	44.0 ± 14.9
	Sniffing time (s)	91.4 ± 10.9	70.1 ± 9.4	68.8 ± 7.0	50.4 ± 5.1
	Following time (s)	13.9 ± 2.8	8.8 ± 2.3	6.5 ± 2.2	9.9 ± 4.3
Rotarod	D1 Time on the rod (s)	196 ± 14	156 ± 11	151 ± 18	139 ± 15
	D2 Time on the rod (s)	223 ± 15	190 ± 10	199 ± 19	157 ± 13
	D3 Time on the rod (s)	207 ± 15	186 ± 11	187 ± 19	141 ± 13
Grip Test	Grip stength (g/g body weight)	8.0 ± 0.2	7.8 ± 0.2	8.2 ± 0.3	8.2 ± 0.4

Supplementary Table 2. Behavioral characterization (following table) of the *Del-Dup* cohort. No phenotype was seen in Y maze test. Recognition memory was evaluated with the novel object recognition test. In the first session (S1) of test, no difference in object exploration was noticed. After a 3-hour retention delay, *Del/+* mice showed a deficit for novel object discrimination. In the Morris water maze, no change of spatial learning and memory was observed. Mutant mice and wild-type littermates travelled the same distance and needed the same time to find the platform from first day to sixth day (D6) of acquisition. In the probe test (PT) on day 7, when the platform is removed, all mice spent the same percentage of time in the target quadrant. In the fear conditioning test, all mice displayed similar level of activity in the conditioning session before footshock. Post-choc freezing was decreased for *Del/+* and *Del/Dup* without significance. In the 6-min contextual session, *Dup/+* showed a global higher

freezing without significance. In cue sessions, *Del/+* and *Dup/+* displayed respectively lower and higher level of freezing in comparison with wild-type littermates. In the three-chamber sociability test, Del/+ mice showed increased social interaction with the first stranger in the first session 1 (S1). In the second session (S2), *Del/+* mice presented trends for higher level of social interaction with the first stranger and the second stranger. No alteration of social preference was found between genotypes. In the social interaction test, in comparison with wild-types, *Del/+* displayed trends for shorter first contact latency and higher level of social interaction whereas *Dup/+* mice displayed trends for longer first contact latency and lower level of social interaction. In the rotarod test, *Del/+* and *Dup/+* mice showed respectively improvements and deficits for motor coordination in comparison with wild-types without statistical significance from first day (D1) to third day (D3) of test. No phenotype was observed in the grip test. For all these tests, *Del/Dup* mice showed similar performances than wild-type mice. Data are mean ± SEM. One-way ANOVA analysis, Tuckey's test otherwise Kruskal-Wallis analysis, Mann-Whitney *U* test. Stars indicate significant difference with wild-type littermates, $^{*}p < 0.05$, $^{**}p < 0.01$.

Test	Parameter	Genotype	
		wt	*Kansl1*$^{+/-}$
Circadian Activity	Hab ambulatory activity (count)	253 ± 21	243± 33
	Hab vertical activity (count)	235 ± 47	426 ± 89
	Dark ambulatory activity (count)	522 ± 57	314 ± 36 *
	Dark vertical activity (count)	872 ± 146	1536 ± 427
	Light ambulatory activity (count)	125 ± 8	86.3 ± 10.3 **
	Light vertical activity (count)	108 ± 15	108 ± 18
	Total ambulatory activity (count)	930 ± 77	658 ± 71 *
	Total vertical activity (count)	1253 ± 194	2126 ± 522
	Total food consumption (g)	5.3 ± 0.2	5.0 ± 0.5
	Total water consumption (ml)	5.0 ± 0.2	5.4 ± 0.3
Open Field	Locomotor exploration (m)	99.1 ± 7.1	108 ± 5
	Rears (count)	170 ± 21	240 ± 24 *
	Time in center (%)	7.1 ± 1.2	10.0 ± 1.3
Repetitive Behaviors	Rearing (count)	39.4 ± 2.8	57.4 ± 6.3 *
	Jumping (count)	0.22 ± 0.22	3.0 ± 2.3
	Climbing (count)	10.4 ± 3.0	14.8 ± 2.9
	Digging (count)	13.1 ± 2.2	3.9 ± 1.4 **
Elevated Plus Maze	Arm entries (count)	14.3 ± 1.8	14.5± 2.2
	Open arm time (%)	4.0 ± 2.3	1.3 ± 0.6
New Object Recognition 3 hour delay	S1 object A exploration (s)	7.2 ± 0.5	5.4 ± 0.7
	S2 object A exploration (s)	2.2 ± 0.4	2.5 ± 0.5
	S2 object B exploration (s)	4.3 ± 0.6	2.8 ± 0.5
	Discrimination index (%)	67.0 ± 2.6	53.0 ± 1.2 ***
Fear Conditioning	Baseline freezing (%)	9.9 ± 1.7	13.7 ± 4.8
	Post-choc freezing (%)	25.7 ± 5.7	14.2 ± 3.3
	Contextual freezing (%)	49.8 ± 7.1	24.8 ± 4.7 *
	First Cue freezing (%)	61.8 ± 6.8	42.5 ± 8.9
	Second cue freezing (%)	69.8 ± 6.8	32.7 ± 5.6 ***
Rotarod	D1 Time on the rod (s)	64.2 ± 3.3	89.3 ± 8.1 **
	D2 Time on the rod (s)	77.7 ± 4.5	164 ± 9 ***
	D3 Time on the rod (s)	114 ± 7	202 ± 13 ***
Grip Test	Grip stength (g/g body weight)	9.7 ± 0.4	10.0 ± 0.3

Supplementary Table 3. Behavioral characterization of the *Kansl1*$^{+/-}$ cohort. Circadian activity test revealed ambulatory hypoactivity in both dark and light phases for mutant mice. No alteration of feeding behavior was noticed during the test. In the open field test, *Kansl1*$^{+/-}$ mice increased level of rears. Observation of repetitive behaviors in home-cages during 10 min showed less digging and more rearing in *Kansl1*$^{+/-}$ mice in comparison with wild-types. No phenotype was observed in the elevated plus maze test. In the first session (S1) of NOR

test, mutant and wild-type mice showed similar exploration duration of the first object. In the second session (S2), *Kansl1*$^{+/-}$ mice displayed an important deficit of novel object discrimination after a retention delay of 3 hours. In the fear conditioning test, mice displayed similar level of activity in the conditioning session before footshock. Post-choc freezing was decreased for *Kansl1*$^{+/-}$ without reaching statistical significance. In the 6-min contextual session, *Kansl1*$^{+/-}$ showed a lower level of freezing. In cue sessions, mutant mice displayed again lower level of freezing with significant difference with wild-types during the second cue. In the rotarod test, *Kansl1*$^{+/-}$ mice displayed an important improvement of motor coordination from first day (D1) to third day (D3) of test. No phenotype was observed in the grip test. Data are mean ± SEM. Student's t-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Discussion et conclusion

La caractérisation des animaux porteurs des réarrangements pour la région *Arf2-Kansl1* a révélé un grand nombre de phénotypes majoritairement associés à la délétion de la région. Une importante létalité a été observée pour les animaux *Del/+* alors que la duplication réciproque n'a aucune incidence sur la viabilité des animaux. Les animaux adultes *Dup/+* présentent néanmoins une diminution de la taille et de la masse corporelle et les animaux adultes *Del/+* présentent des tendances de diminution de la masse corporelle, une diminution de la taille ainsi qu'un déficit d'adipogenèse. En plus de ces phénotypes, l'analyse cranio-faciale a également révélé des altérations proches de la significativité pour la forme crânienne des animaux *Del/+* et *Dup/+*.

La caractérisation comportementale a révélé un grand nombre de phénotypes pour les animaux *Del/+* incluant une hypoactivité locomotrice en test d'activité circadienne, un niveau augmenté de l'interaction sociale, un déficit de mémoire de reconnaissance ainsi qu'un déficit de mémoire associative en test de conditionnement à la peur. Au contraire, les animaux *Dup/+* ne présentent qu'une amélioration de la mémoire associative aversive. Des tendances de déficits de la coordination motrice sont également observées pour les animaux *Dup/+*.

Des analyses électrophysiologiques sur coupes d'hippocampe aiguës ont été réalisées. Pour les animaux *Del/+*, une diminution de la facilitation synaptique à court-terme (PPF pour Paired Pulse Facilitation) ainsi qu'une diminution des potentiels excitateurs post-synaptiques (fEPSP field Excitatory PostSynaptic Potential) pour les hautes stimulations ont été observées. Ces résultats sont consistants avec les déficits de mémoire de reconnaissance observés. A l'opposé, les analyses effectuées sur les animaux *Dup/+* ont révélé des tendances d'augmentation de la PPF et une augmentation des fEPSP. Aucune altération de la LTP n'a été observée.

L'étude préliminaire des animaux mutants hétérozygotes *Kansl1*$^{+/-}$ a révélé un grand nombre de phénotypes similaires aux animaux *Del/+* incluant une importante létalité, une diminution de la taille corporelle, une réduction de l'adipogenèse ainsi que des déficits pour la mémoire de reconnaissance et pour la mémoire associative aversive. Une diminution de la masse corporelle ainsi qu'une amélioration de la coordination motrice ont également été observés pour les animaux *Kansl1*$^{+/-}$.

Notre modélisation des syndromes de délétion 17q21.31 et de l'haplo-insuffisance pour le gène *Kansl1* confirme les études de génétique humaine indiquant que l'haplo-insuffisance de *KANSL1* suffit à générer les phénotypes associés au syndrome de Koolen-De Vries.

Les déficits mnésiques et l'augmentation du niveau d'interaction sociale des animaux *Del/+* sont proches de la déficience intellectuelle et du comportement affectifs des patients porteurs de la délétion 17q21.31 et de mutations perte de fonction pour le gène *KANSL1*. De prochaines études seront dédiées à la caractérisation des comportements sociaux ainsi qu'à l'analyse électrophysiologique sur coupes d'hippocampe aiguës des animaux *Kansl1*$^{+/-}$.

Alors que très peu de patients porteurs de la duplication réciproque 17q21.31 ont identifiés, un nombre limités de phénotypes a été observé pour les animaux *Dup/+* qui présentent une réduction de la taille et de la masse corporelles, des altérations de la forme crânienne ainsi qu'une amélioration de la mémoire associative aversive. Nos résultats suggèrent une faible pathogénicité de la duplication 17q21.31 qui pourrait expliquer le faible nombre de patients identifiés ainsi que l'hétérogénéité de la symptomatique associée.

Conclusion et perspectives

Modélisation de la monosomie 21 partielle

La lignée Ms5Yah n'étant pas viable sur un fond consanguin, nous avons maintenu les animaux sur un fond génétique B6NC3B. Un taux de transmission extrêmement faible de 22.4% a été observé pour les animaux porteurs de la délétion pour la région *App-Runx1*. Des tests de viabilité effectués au stade embryonnaire E18.5 n'ont révélé aucune létalité prénatale. Néanmoins, les animaux Ms5Yah présentaient une réduction de la masse corporelle associée et plus de la moitié des animaux mutants présentaient une incapacité respiratoire. L'haplo-insuffisance des gènes de la région *App-Runx1* est donc associée à une importante mortalité post-natale. Dans le but de déterminer les causes de cette létalité, nous avons réalisé une étude anatomique des fœtus présentant une incapacité respiratoire. Néanmoins, hormis l'absence d'ouverture des alvéoles pulmonaires, aucune anomalie probante n'a été observée.

La région de synténie *Kcne1-Runx1* est comprise dans la région *App-Runx1*. Cet intervalle de 560 kb critique pour la viabilité comprenant les gènes *KCNE1, DSCR1, CLIC6* et *RUNX1* a été associé aux malformations cardiaques de la PM21 (Lindstrand et al., 2010). Néanmoins, aucun défaut de l'anatomie et de la fonction cardiaque n'a été observé chez les animaux Ms5Yah (Raveau et al., 2012). Chez l'homme, l'haplo-insuffisance du gène *RUNX1* est également associée à une thrombopénie (Jalagadugula et al., 2010; Liew and Owen, 2011; Shinawi et al., 2008). De manière intéressante, l'analyse hématologique révèle une importante réduction du nombre de plaquettes sanguines pour les animaux Ms5Yah. La durée de coagulation des animaux mutants est également deux fois plus longue en comparaison aux congénères sauvages. Ainsi, l'haplo-insuffisance en *Runx1* pourrait expliquer en partie la fragilité des animaux Ms5Yah.

Les animaux Ms5Yah survivants présentent un important déficit de la taille et de la masse corporelle ainsi que d'importantes difficultés de coordination. A l'âge adulte, les animaux mutants présentent une démarche caractéristique lors de l'observation des animaux en cage d'hébergement. La caractérisation comportementale des animaux Ms5Yah a révélé d'importants déficits d'apprentissage et de mémoire spatiale dans le test de piscine de Morris (MWM) ainsi que des déficits de la coordination et de l'apprentissage locomoteur dans les tests de rotarod. De manière intéressante, ces phénotypes murins sont extrêmement proches de la petite taille, du retard psychomoteur et de la déficience intellectuelle associés aux patients porteurs de délétion pour la région *APP-SOD1* du locus 21q21.3 (Lindstrand et al., 2010; Lyle et al., 2009; Roberson et al., 2011). Une hypertonie et de nombreuses malformations cranio-

faciales sont également observées chez ces patients. Néanmoins, l'analyse cranio-faciale n'a indiqué aucune altération de la forme et de la taille du crâne des animaux Ms5Yah. Aucune anomalie de la force d'agrippement de n'a été observée pour les animaux mutants. Il serait intéressant de réaliser un électromyogramme (EMG) pour compléter l'analyse pour déterminer si les défauts de coordination des animaux mutants sont liés au fonctionnement musculaire. Des analyses électrophysiologiques sur coupes d'hippocampe aiguës seront également prochainement réalisées pour compléter notre caractérisation du modèle Ms5Yah.

Nous nous sommes particulièrement intéressés aux déficits observés lors du test MWM. Les déficits d'apprentissage et de mémoire des animaux modèles sont très intéressants car ils peuvent être associés à la déficience intellectuelle des patients. Les régions cérébrales impliquées dans la mémoire de reconnaissance incluent principalement l'hippocampe auquel s'ajoutent des afférences des cortex entorhinal, périrhinal et parahippocampique (Broadbent et al., 2010; Squire et al., 2007). C'est donc la région hippocampique que nous avons ciblée pour notre étude transcriptomique sur puce Affymetrix. L'analyse a révélé que la majorité des gènes de la région *App-Runx1* est sensible aux effets de dose et que la délétion a un impact important sur la régulation du génome. Effectivement, nous avons identifié 96 gènes surexprimés et 96 gènes sous-régulés incluant 33 gènes localisés dans la région *App-Runx1*. L'analyse GO a révélé une perturbation des gènes impliqués dans l'adhésion cellulaire incluant le gène *App*. Ce gène est un candidat intéressant pour expliquer les phénotypes cognitifs observés chez le modèle Ms5Yah. Chez la souris adulte, les données de l'Allen Brain Atlas indiquent une forte expression d'*App* au niveau de l'hippocampe. La protéine précurseur de l'amyloïde est impliquée dans de nombreux processus neuronaux incluant la synaptogenèse (Wang et al., 2009), la plasticité synaptique et la mémoire (Turner et al., 2003). Les modèles murins mutants homozygote $App^{-/-}$ présentent une réduction de la masse corporelle et une activité locomotrice diminuée (Zheng et al., 1995). Les souris $App^{-/-}$ âgées présentent également des déficits d'apprentissage et de mémoire et des altérations de la LTP hippocampique (Ring et al., 2007). Nous avons également observé une surexpression du gène *Klk6* qui code pour une sérine protéase. De manière intéressante, KLK6 est un biomarqueur de la maladie d'Alzheimer (Diamandis et al., 2000) ; de plus, l'un des substrats de la protéase KLK6 est la protéine APP (Ashby et al., 2010). Néanmoins, nous ne pouvons sous-estimer la contribution d'autres gènes comme *Cbs, Olig1* et *Olig2* dans les déficits d'apprentissage observés chez les animaux Ms5Yah.

L'ensemble de nos résultats indique que les animaux Ms5Yah présentent de nombreux phénotypes pouvant être associés à la symptomatique de la monosomie 21 partielle. Notre caractérisation confirme les effets délétères des délétions de la région 21q21.3q22.11 observés lors des études de génétique humaine. Contrairement aux autres modèles générés pour les différentes régions de synténie de l'HSA21 qui montrent des phénotypes relativement légers, les animaux porteurs de la délétion *App-Runx1* présentent une importante létalité et de sévères déficits de coordination locomotrice, d'apprentissage et de mémoire. Le modèle Ms5Yah représente un nouvel outil génétique permettant d'identifier les voies de signalisation et les mécanismes moléculaires impliqués dans la physiopathologie de la PM21 associé aux délétions de la région 21q21.3q22.11. De prochaines études pourraient être dédiées à la recherche de molécules permettant de restaurer les capacités d'apprentissage des animaux Ms5Yah. Ces études pourraient potentiellement aboutir au développement des premières stratégies thérapeutiques permettant d'améliorer les capacités intellectuelles des patients.

Modélisation des délétions et duplications 16p11.2

Chez l'homme, les délétions de la région 16p11.2 BP4-BP5 apparaissent majoritairement *de novo* alors que les duplications réciproques sont majoritairement héritées de l'un des deux parents (Girirajan et al., 2012). Pour les deux réarrangements, la transmission se fait principalement du côté maternel. Lors de l'amplification des animaux mutants porteurs de la délétion *(Del/+)* ou de la duplication *(Dup/+)* de la région *Sult1a1-Spn*, aucune différence de transmission n'a été observée en fonction du sexe du parent porteur. Sur un fond génétique consanguin B6N, une importante létalité a été observée pour les animaux *Del/+* alors qu'aucune incidence n'a été observée pour les animaux *Dup/+*. En test de viabilité, une importante diminution du poids ainsi que des difficultés respiratoires ont été observées pour les animaux *Del/+* indiquant un retard développemental alors qu'aucune anomalie n'a été observée pour les animaux *Dup/+*. Ces résultats montrent un effet particulièrement délétère de l'haplo-insuffisance des gènes de la région 16p11.2 par rapport à leur surexpression. Ceci est cohérent avec l'importante proportion de délétions *de novo* retrouvée chez l'homme alors que la duplication réciproque peut être transmise par des parents asymptomatiques ou présentant des symptômes modérés. De plus, le nombre plus élevé de patients porteurs de délétion 16p11.2 BP4-BP5 identifié dans les cohortes d'ID est conséquent à une pathogénicité augmentée de la délétion par rapport à la duplication réciproque (Cooper et al., 2011).

Les patients porteurs des réarrangements présentent des phénotypes opposés pour le volume crânien et le BMI. Alors que la délétion est associée à la macrocéphalie et à l'obésité et à l'hyperphagie, la duplication réciproque est associée à la microcéphalie et à l'insuffisance pondérale et à un comportement alimentaire sélectif et restrictif (Jacquemont et al., 2011; McCarthy et al., 2009; Shinawi et al., 2010; Walters et al., 2010; Weiss et al., 2008; Zufferey et al., 2012). De manière intéressante, la caractérisation des animaux mutants a également révélé de nombreux phénotypes opposés. Les animaux *Del/+* présentaient un poids anormalement bas et une faible quantité de tissue adipeux, une hyperactivité ainsi qu'un déficit de la mémoire de reconnaissance. Au contraire, les animaux *Dup/+* montraient une augmentation du poids et de la quantité de tissue adipeux, une hypoactivité ainsi qu'une amélioration de la mémoire de reconnaissance. Chez les animaux, les phénotypes de masse corporelle et d'adipogenèse ne sont pas associés à des altérations du comportement alimentaire. Ainsi, les phénotypes de masse corporelle et d'adipogenèse observés chez les patients et les animaux mutants sont diamétralement contraires.

Lors de notre caractérisation, nous avons également étudié les animaux porteurs des deux réarrangements *Del/Dup*. Ces animaux sont donc porteurs des deux régions *Sult1a1-Spn* sur un seul chromosome. Un unique phénotype a été observé pour le test de reconnaissance du nouvel objet. Comme pour les animaux *Dup/+*, les animaux *Del/Dup* montraient une l'amélioration de la mémoire de reconnaissance. Ces résultats suggèrent que la structure de l'ADN de la duplication *Sult1a1-Spn* pourrait impacter sur l'expression génique par des effets de position ou des perturbations de la chromatine pouvant potentiellement aboutir à l'amélioration de la mémoire de reconnaissance des animaux.

La récurrence de phénotypes opposés observés chez les patients et les animaux mutants suggère l'implication de gènes sensibles aux effets de dose. Nos études transcriptomiques sur puce Affymetrix ont indiqué que la majorité des gènes de la région *Sult1a1-Spn* est sensible au dosage génique. L'analyse en composantes principales a indiqué un effet plus fort de la délétion sur l'expression des gènes de la région ce qui pourrait expliquer les phénotypes des animaux *Del/+*. Un impact très limité a été observé pour la régulation du génome. Ces résultats suggèrent l'implication de gènes localisés au sein de la région *Sult1a1-Spn* dans l'apparition de phénotypes de métabolisme énergétique, d'activité et de mémoire de reconnaissance observés chez les animaux *Del/+* et *Dup/+*. De futures analyses transcriptomiques à partir de prélèvements d'hippocampe d'animaux *Del/Dup* nous permettront d'étudier l'impact des deux réarrangements sur l'expression génique.

Concernant les altérations de BMI, la régulation de gènes sensibles au dosage génique situés en dehors de la région 16p11.2 pourrait diverger entre l'homme et la souris. Le gène *SH2B1* situé à 800 kb en amont de la région 16p11.2 BP4-BP5 est impliqué dans la signalisation de la leptine et le régulation de la prise alimentaire (Allison and Myers, 2014). Ce gène est localisé dans la région 16p11.2 BP2-BP3 dont la délétion est également associée à l'obésité et à l'ID (Bachmann-Gagescu et al., 2010). *SH2B1* pourrait être dérégulé par des effets de position des réarrangements BP4-BP5. Aucune régulation de ce gène n'a été observée au niveau des lignées lymphoblastoïdes dérivées des cellules sanguines des patients (Jacquemont et al., 2011; Walters et al., 2010) mais cela ne veut pas dire que le gène *SH2B1* ne soit pas dérégulé au niveau du système nerveux central des patients. Chez les animaux mutants, nous n'avons observé aucune dérégulation du gène *Sh2b1* dans les régions cérébrales étudiées ce qui est consistant avec l'absence d'altération du comportement alimentaire.

Nous nous sommes également intéressés au volume crânien. Nous avons réalisé une analyse cranio-faciale permettant d'étudier les altérations de forme et de taille du crâne des animaux. L'analyse a révélé une altération de la forme du crâne pour les animaux *Del/+* et *Dup/+* ainsi qu'une diminution de la taille des crânes *Del/+*. Néanmoins, une diminution de la taille fémorale a également été observée pour les animaux *Del/+* indiquant un retard développemental qui affecte l'ensemble du squelette. Ici encore, nos résultats sont différents de la symptomatique humaine. Nos résultats divergent également des études réalisées sur le poisson zèbre pour lesquelles la surexpression de *KCTD13* induit une diminution de la taille de la tête de l'animal, alors que l'injection de morpholino dirigé contre l'orthologue du poisson induit une augmentation de la taille de la tête (Golzio et al., 2012) . Chez la souris, l'haplo-insuffisance des gènes de la région de synténie 16p11.2 incluant une sous-expression du gène *Kctd13* engendre une diminution du volume crânien, alors que la surexpression des gènes n'a pas de conséquence sur le volume crânien. Néanmoins, nous n'excluons pas la potentielle implication du gène *Kctd13* dans les altérations de la forme du crâne des animaux mutants. Nous sommes actuellement en attente des résultats IRM de la structure cérébrale des animaux *Del/+* et *Dup/+*. Ces analyses réalisées en collaboration avec le Dr. Nathalie Just du CIB (Centre d'Imagerie BioMédicale de Lausanne) nous permettront d'évaluer les altérations neuroanatomiques de nos modèles.

Chez l'homme, les réarrangements de la région 16p11.2 BP4-BP5 sont associés à un grand nombre de troubles neuropsychiatriques dont les plus fréquents incluent l'ID et l'ASD (Cooper et al., 2011; Weiss et al., 2008). Les deux réarrangements sont également associés à

l'épilepsie et à l'ADHD alors que seule la duplication est associée à la schizophrénie, au trouble bipolaire et à la dépression (Reinthaler et al., 2014; Shinawi et al., 2010; Steinberg et al., 2014; Zufferey et al., 2012). Sur un fond consanguin B6N, les animaux *Del/+* présentaient des stéréotypies de redressement vertical, une hyperactivité, ainsi qu'un déficit de mémoire de reconnaissance. Ces phénotypes sont à associer à l'ID et aux troubles d'activité des patients. Les stéréotypies sont également associés aux modèles murins d'ASD (Crawley, 2007). Néanmoins, aucune altération des comportements d'interaction sociale et de préférence sociale pour le nouvel individu n'a été observée. Les animaux *Dup/+* présentaient une hypoactivité ainsi qu'une amélioration de la mémoire de reconnaissance ce qui est éloigné de la symptomatique humaine.

L'importante variabilité et la faible pénétrance des troubles neuropsychiatriques suggèrent l'implication de polymorphismes secondaires dans la modulation de la pathogénicité des réarrangements 16p11.2. Une étude récente a notamment montré l'enrichissement de seconds larges CNVs chez les patients porteurs de la duplication (Girirajan et al., 2012). Dans le but d'étudier l'influence du fond génétique, nous avons décidé de répliquer notre analyse avec des animaux *Del/+* et *Dup/+* maintenus sur un fond hybride B6NC3B. Les animaux de fond C3B présentent de bonnes capacités cognitives et une activité plus faible en comparaison à la souche B6N (Mandillo et al., 2008). Une importante atténuation des phénotypes d'activité a donc été observée pour les animaux mutants. Néanmoins, les phénotypes de mémoire de reconnaissance sont conservés. De manière très intéressante, une diminution des comportements sociaux a été observée chez les animaux *Del/+* et *Dup/+* dans le test d'interaction sociale en champ ouvert. Dans le test de sociabilité à trois chambres, aucun phénotype d'interaction sociale n'a été observé, mais les animaux *Del/+* ont présenté un important déficit de préférence sociale pour le nouvel individu. Nos résultats corroborent donc les études humaines et suggèrent l'implication de polymorphismes secondaires dans la modulation des phénotypes neuropsychiatriques associés aux réarrangements de la région 16p11.2 BP4-BP5. Chez la souris, les allèles de fond génétique B6N potentialisent les altérations d'activité alors que les allèles du fond génétique C3B potentialisent les déficits d'interaction sociale. Le choix du fond génétique est donc primordial pour la modélisation murine des troubles neuropsychiatriques (Kerr et al., 2013; Pietropaolo et al., 2011).

Encore une fois, nous nous somme intéressés plus particulièrement aux phénotypes de mémoire de reconnaissance. Les études électrophysiologiques ont révélé une altération de la transmission synaptique au niveau de l'hippocampe. De manière étonnante, une diminution de la LTP a été observée pour les animaux *Del/+* présentant un déficit des capacités mnésiques mais aussi pour les animaux *Dup/+* présentant une amélioration des capacités mnésiques. La région *Sult1a1-Spn* contient une trentaine de gènes codants dont la majorité est exprimée au niveau central et sensible aux effets de dose. L'identification des gènes candidats pour les phénotypes mnésiques présente donc une grande complexité. Plusieurs gènes présents dans la région *Sult1a1-Spn* sont des candidats très intéressants pour les phénotypes mnésiques observés chez la souris. Ces gènes incluent *Doc2a*, *Kctd13*, *Prrt2*, *Sez6l2* et *Taok2* dont les fonctions sont décrites dans l'introduction. Les travaux de Golzio *et al.* (2012) ont montré que la sous-expression de *Kctd13* induisait une augmentation de la prolifération des neurones corticaux de l'embryon murin alors que sa surexpression induisait un ralentissement de la prolifération cellulaire couplé à une augmentation de l'apoptose. Les études de De Anda *et al.* (2012) ont également montré que la modulation de l'expression du gène *Taok2* dans des cultures de neurones corticaux de souris impactait sur la formation dendritique. Nous allons prochainement réaliser une analyse morphologique de l'arborisation dendritique sur coupes d'hippocampe colorées par la méthode Golgi Cox pour les animaux des quatre génotypes *Del/+*, wt, *Del/Dup* et *Dup/+*. Chez la souris adulte, les données de l'Allen Brain Atlas indiquent une forte expression de *Kctd13* et de *Taok2* au niveau de l'hippocampe et du cervelet. Nous avons donc choisi ces deux gènes pour réaliser des études de restauration par injection bilatérale de vecteurs AAV au niveau de l'hippocampe des animaux *Del/+*. Ces expériences seront réalisées dans les prochains mois. Les animaux seront évalués pour des tests comportementaux incluant l'activité circadienne, le champ ouvert, le labyrinthe en Y et la reconnaissance du nouvel objet avant et après l'injection stéréotaxique qui sera effectuée chez des animaux âgés de dix semaines. L'évaluation des capacités de discrimination du nouvel objet avant et après injection permettra d'étudier l'implication potentielle des gènes *Kctd13* et *Taok2* dans les déficits cognitifs des animaux *Del/+*. Dans le cas de la restauration de l'expression du gène *Taok2*, nous prévoyons également de réaliser une analyse morphologique de l'arborisation dendritique hippocampique des animaux *Del/+* et leur congénères sauvages. La comparaison aux résultats préalables sans injection AAV nous permettra d'étudier la conséquence de l'expression du gène *Taok2* sur la formation dendritique.

En parallèle, la caractérisation comportementale de lignés murines KO hétérozygotes et homozygotes pour les gènes de la région *Sult1a1-Spn* est en cours dans notre équipe. Aucun phénotype mnésique n'a encore été observé pour les lignées caractérisées. Dans le but d'étudier l'effet de dose d'un potentiel gène candidat *X*, il serait très intéressant de croiser les animaux *Dup/+* avec les animaux mutants hétérozygotes $X^{+/-}$. Cette stratégie permettrait d'étudier les animaux $X^{+/-}$ mais aussi les animaux *X/Dup* et d'étudier si le retour à deux copies du gène candidat *X* chez les animaux porteurs de la duplication abaisse les capacités mnésiques au niveau des animaux sauvages. Cette stratégie élégante a précédemment été utilisée dans notre équipe et à permis de démontrer l'implication du gè*ne Cbs* dans les déficits mnésique d'un modèle de trisomie 21 partielle.

Modélisation des délétions et duplications 17q21.31

Chez l'homme, les réarrangements de la région 17q21.31 apparaissent *de novo*. L'haplotype H2 qui favorise les NAHR est transmis de manière équilibrée par les deux parents (Dubourg et al., 2011; Koolen et al., 2008). De manière similaire, lors de l'amplification des animaux mutants porteurs de la délétion (*Del/+*) ou de la duplication (*Dup/+*) de la région *Arf2-Kansl1*, aucune différence de transmission n'a été observée en fonction du sexe du parent porteur. Sur un fond génétique consanguin B6N, une importante létalité a été observée pour les animaux porteurs de la délétion de la région de synténie 17q21.31 (*Del/+*) alors qu'aucune réduction de la transmission n'a été observée pour les animaux porteurs de la duplication réciproque (*Dup/+*). Des tests de viabilité seront prochainement réalisés pour étudier la létalité et la de masse corporelle des animaux mutants au stade embryonnaire E18.5.

La caractérisation des animaux montre des phénotypes très proche de la symptomatique humaine. Les patients porteurs de la délétion de la région 17q21.31 présentent une déficience intellectuelle ainsi qu'un comportement de type hypersociable. De manière similaire, les animaux *Del/+* présentent des déficits de mémoire de reconnaissance et de mémoire associative aversive ainsi qu'une augmentation du niveau d'interaction sociale. Depuis 2007, seulement cinq patients porteurs de la duplication réciproques ont été caractérisés. La symptomatique, largement hétérogène, comprend des traits autistiques et un retard psychomoteur. Les animaux *Dup/+* présentent très peu de phénotypes incluant une diminution de la taille et de la masse corporelle ainsi qu'une amélioration de la mémoire associative aversive. Les études électrophysiologiques ont révélé une altération de la transmission

synaptique au niveau de l'hippocampe. Pour les animaux *Del/+*, une diminution de la facilitation synaptique à court-terme (PPF) ainsi qu'une diminution des potentiels excitateurs post-synaptiques (fEPSP) ont été observées ce qui est consistant avec les déficits de mémoire de reconnaissance observés. A l'opposé, les analyses effectuées sur les animaux *Dup/+* ont révélé des tendances d'augmentation de la PPF et une augmentation des fEPSP.

Lors de notre caractérisation, nous avons également étudié les animaux porteurs des deux réarrangements *Del/Dup*. Ces animaux sont donc porteurs des deux régions *Arf2-Kansl1* sur un seul chromosome. Comme pour les animaux *Del/+* et *Dup/+*, les animaux *Del/Dup* montraient une diminution de la taille corporelle. Néanmoins aucun phénotype comportemental n'a été observé pour les animaux *Del/Dup*.

En complément de l'étude des réarrangements de la région de synténie 17q21.31, nous avons entrepris la caractérisation des animaux mutants hétérozygotes pour le gène *Kansl1* (*Kansl1*$^{+/-}$). De nombreux phénotypes similaires aux animaux 17q21.31 *Del/+* ont été observés. Effectivement, une importante létalité est associée aux animaux mutants. Les animaux *Kansl1*$^{+/-}$ survivants présentent des réductions de masse et de taille corporelles ainsi qu'un déficit d'adipogenèse. Une amélioration de la coordination motrice ainsi que des déficits de mémoire de reconnaissance et de mémoire associative aversive ont également été observés. Les comportements sociaux des animaux *Kansl1*$^{+/-}$ seront prochainement évalués avec une seconde cohorte. Nous sommes en attente des résultats d'analyses de RNA-seq réalisées à partir de prélèvements d'hippocampe pour les animaux des cohortes *Del/+*, *Kansl1*$^{+/-}$ et de leurs congénères sauvages. La comparaison des profils d'expression pour les animaux *Del/+* et *Kansl1*$^{+/-}$ permettra de déterminer si l'haplo-insuffisance du gène *Kansl1* suffit à induire des régulations génomiques similaires à la délétion de la région de synthénie 17q21.31. Le gène *Kansl1* interagit avec le complexe NSL (Dias et al., 2014) impliqué dans la régulation de la transcription par acétylation des lysines K5, K8 et K16 de l'histone H4 (Cai et al., 2010; Dou et al., 2005). Dans le but d'étudier ces modifications épigénétiques, des expériences d'immunoprécipitation de la chromatine (ChIP-seq) seront prochainement réalisées à partir de prélèvements d'hippocampe et de cortex d'animaux *Del/+*, *Kansl1*$^{+/-}$ et de leurs congénères sauvages.

Nos résultats suggèrent une faible pathogénicité de la duplication 17q21.31 ce qui pourrait expliquer le faible nombre de patients identifiés ainsi que l'hétérogénéité de la symptomatique associée. Alors que deux études indépendantes ont associé le gène *KANSL1* aux symptômes la délétion 17q21.31 (Koolen et al., 2012; Zollino et al., 2012), il paraît moins

évident que *KANSL1* soit impliqué dans les symptômes de la duplication. Les troubles de la sociabilité observés chez certains patients pourraient potentiellement être associés à la surexpression du gène *CRHR1* dont certains SNPs sont associés à l'anxiété, à la dépression ainsi qu'à la dépendance alcoolique (Chen et al., 2010; Rogers et al., 2013; Treutlein et al., 2006). Disposant de la lignée *Kansl1*$^{+/-}$, il serait intéressant de tester la contre-hypothèse chez la souris et de croiser les animaux *Dup/+* avec les animaux *Kansl1*$^{+/-}$. La caractérisation des animaux permettrait de savoir si le retour à deux copies du gène *Kansl1* chez les animaux porteurs de la duplication *Arf2-Kansl1* a des conséquences anatomiques et comportementales.

Nos résultats confirment l'effet délétère de la délétion 17q21.31 associée à de nombreuses altérations anatomiques et cognitives aboutissant à des déficits de l'apprentissage et de la mémoire ainsi qu'à un intérêt accru pour les comportements sociaux. De plus, des résultats préliminaires indiquent que l'haplo-insuffisance du gène *Kansl1*suffit à engendrer les phénotypes observés chez les animaux *Del/+*. Des études de restauration de l'expression du gène *Kansl1* par injection bilatérale de vecteurs AAV dans l'hippocampe vont être prochainement réalisée pour les animaux *Del/+*. Derechef, les souris seront évaluées pour des tests comportementaux incluant l'activité circadienne, le champ ouvert, le labyrinthe en Y et la reconnaissance du nouvel objet avant et après l'injection stéréotaxique qui sera effectuée chez des animaux âgés de dix semaines. S'il y a restauration de la discrimination du nouvel objet après injection, cela confirmera l'implication du gène *Kansl1* dans les déficits de mémoire de reconnaissance observés pour les animaux 17q21.31 *Del/+*.

Conclusion générale

Nos résultats mettent en lumière l'effet particulièrement délétère des délétions par rapport aux duplications réciproques pour les trois régions étudiées. Les déficits d'apprentissage et de mémoire des animaux pouvant être associés à la déficience intellectuelle des patients, nos résultats corroborent l'important enrichissement des larges délétions incluant des régions géniques par rapport aux duplications réciproques dans les groupes de patients atteints de retards développementaux (Conrad et al., 2010b; Cooper et al., 2011; Grayton et al., 2012).

La délétion de la région *App-Runx1* est associée à une importante létalité post-natale et de sévères déficits de la coordination locomotrice, de la mémoire et de l'apprentissage. Au contraire, des études menées au sein de l'équipe indiquent que les animaux Ts5Yah porteurs de la duplication réciproque *App-Runx1* ne présentent aucun déficit de transmission. De plus,

les études préliminaires de caractérisation comportementale n'ont révélé aucun phénotype pour les animaux Ts5Yah. La délétion *App-Runx1* constitue un exemple de CNV pour lequel, bien que des phénotypes sévères soient associés à la délétion, aucune anomalie ne résulte de la duplication réciproque. Chez l'homme, alors que les délétions de la région 21q21.3q22.11 sont également associées à des phénotypes extrêmement délétères, aucune trisomie partielle restreinte à la région 21q21.3q22.11 n'a été identifiée chez des patients présentant une déficience intellectuelle (Lyle et al., 2009). C'est également le cas pour le syndrome de Sotos associé à la délétion de la région 5q35 (Fickie et al., 2011) ainsi que le syndrome de délétion 2q37 (Leroy et al., 2013). Cette situation facilite le développement de thérapies géniques destinées à rétablir l'expression de gènes candidats situés dans la région de délétion. Effectivement, la dose n'étant pas toujours totalement maîtrisée, l'injection de vecteurs viraux destinés à rétablir l'expression d'un gène peut engendrer sa surexpression associée à des effets délétères rédhibitoires à la thérapeutique.

Les réarrangements de la région 17q21.31 constituent un exemple intermédiaire. Notre modélisation murine a révélé de nombreux phénotypes pour la délétion incluant une létalité post-natale, des déficits mnésiques et une augmentation du niveau d'interaction sociale. Au contraire, la duplication réciproque est associée à une amélioration de la mémoire associative aversive et n'engendre aucune létalité. Chez l'homme, la délétion est associée à une déficience intellectuelle, un comportement de type hypersociable ainsi qu'une dysmorphie faciale caractéristique. La duplication réciproque est extrêmement rare et la symptomatique associée est largement hétérogène ce qui suggère l'implication de polymorphismes secondaires dans la modulation de la pathogénicité de la duplication 17q21.31. Les études de génétique humaines ont identifié *KANSL1* comme gène causatif du syndrome de délétion 17q21.31 ou syndrome de Koolen-de Vries. Chez la souris, nos études confirment que l'haplo-insuffisance du gène *Kansl1* engendre des déficits cognitifs similaires aux animaux *Del/+*. Ces résultats sont extrêmement prometteurs pour le développement de futures stratégies thérapeutiques pouvant potentiellement aboutir à l'amélioration des capacités cognitives des patients atteints du syndrome de Koolen-de Vries.

En ce qui concerne les réarrangements de la région 16p11.2 BP4-BP5, des phénotypes opposés sont observés chez l'homme ce qui suggère l'implication de gène(s) à effet de dose linéaire. Les deux réarrangements sont associés à la déficience intellectuelle, aux troubles du spectre autistique et à l'épilepsie. Des phénotypes opposés sont observés pour le volume crânien et l'indice de masse corporelle. Chez la souris, nous avons également observé de

nombreux phénotypes opposés au niveau de la masse corporelle, de l'adipogenèse, de l'activité locomotrice et de la mémoire de reconnaissance. Si nos résultats confirment l'implication de gène(s) sensible(s) aux effets de dose dans les phénotypes associées aux réarrangements 16p11.2, de nombreuses disparités sont néanmoins observées entre l'homme et l'animal, notamment au niveau de la masse corporelle, de l'adipogenèse et du volume crânien. Ces résultats suggèrent la possible implication de gène(s) sensible(s) à la dose situés en dehors de la région 16p11.2 BP4-BP5 qui pourraient diverger entre l'homme et la souris. En addition aux réarrangements de la région 16p11.2, la majeure partie des syndromes de déficience intellectuelle associés aux CNVs inclut des délétions et des duplications réciproques pathogènes (Cooper et al., 2011). On peut citer les syndromes de délétion et de duplication des régions 1p36, 2q13, 4p16.3, 7q11.23, 15q11.2, 15q11q13, 17p11.2 et 22q11.2. Ces conditions pathologiques compliquent considérablement la possible prise en charge thérapeutique.

Le symptôme commun aux maladies syndromiques étudiées dans nos travaux est la déficience intellectuelle. Les premiers essais thérapeutiques sont aujourd'hui dédiés à l'amélioration des capacités cognitives de patients atteints de maladies neurodéveloppementales incluant le syndrome de Down (De la Torre et al., 2014). Nos modèles murins représentent de nouveaux outils génétiques permettant la compréhension des mécanismes moléculaires aboutissant à la diminution des capacités intellectuelles pour les patients porteurs de réarrangements pour les régions 21q21.3q22.11, 16p11.2 BP4-BP5 et 17q21.31. Effectivement, les déficits de mémoire de reconnaissance observés chez les modèles 16p11.2 *Del/+* et 17q21.31 *Del/+* ainsi que les déficits d'apprentissage locomoteur et de mémoire spatiale observés chez le modèle Ms5Yah peuvent être directement associés à la diminution des capacités cognitives des patients. L'identification des gènes candidats responsables de ces déficits permettra la compréhension des mécanismes moléculaires associés. De futures études pourront être dédiées à la recherche de molécules restaurant les déficits d'apprentissage observés chez les animaux 16p11.2 *Del/+*, 17q21.31 *Del/+* et Ms5Yah pouvant potentiellement aboutir au développement des premières stratégies thérapeutiques.

Références bibliographiques

(1989). Proposal for revised classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy. Epilepsia *30*, 389-399.

Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., McVean, G.A., and Consortium, G.P. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061-1073.

Abrahams, B.S., and Geschwind, D.H. (2008). Advances in autism genetics: on the threshold of a new neurobiology. Nat Rev Genet *9*, 341-355.

Albertson, D.G., and Pinkel, D. (2003). Genomic microarrays in human genetic disease and cancer. Hum Mol Genet *12 Spec No 2*, R145-152.

Alfi, O.S., Donnell, G.N., Allderdice, P.W., and Derencsenyi, A. (1976). The 9p- syndrome. Ann Genet *19*, 11-16.

Allison, M.B., and Myers, M.G. (2014). 20 YEARS OF LEPTIN: Connecting leptin signaling to biological function. J Endocrinol *223*, T25-T35.

Arlt, M.F., Wilson, T.E., and Glover, T.W. (2012). Replication stress and mechanisms of CNV formation. Curr Opin Genet Dev *22*, 204-210.

Ashby, E.L., Kehoe, P.G., and Love, S. (2010). Kallikrein-related peptidase 6 in Alzheimer's disease and vascular dementia. Brain Research *1363*, 1-10.

Bachmann-Gagescu, R., Mefford, H.C., Cowan, C., Glew, G.M., Hing, A.V., Wallace, S., Bader, P.I., Hamati, A., Reitnauer, P.J., Smith, R.*, et al.* (2010). Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. Genet Med *12*, 641-647.

Bah, J., Quach, H., Ebstein, R.P., Segman, R.H., Melke, J., Jamain, S., Rietschel, M., Modai, I., Kanas, K., Karni, O.*, et al.* (2004). Maternal transmission disequilibrium of the glutamate receptor GRIK2 in schizophrenia. Neuroreport *15*, 1987-1991.

Bailey, J.A., Gu, Z., Clark, R.A., Reinert, K., Samonte, R.V., Schwartz, S., Adams, M.D., Myers, E.W., Li, P.W., and Eichler, E.E. (2002). Recent segmental duplications in the human genome. Science *297*, 1003-1007.

Ballif, B.C., Hornor, S.A., Jenkins, E., Madan-Khetarpal, S., Surti, U., Jackson, K.E., Asamoah, A., Brock, P.L., Gowans, G.C., Conway, R.L.*, et al.* (2007). Discovery of a previously unrecognized microdeletion syndrome of 16p11.2-p12.2. Nat Genet *39*, 1071-1073.

Barbouti, A., Stankiewicz, P., Nusbaum, C., Cuomo, C., Cook, A., Höglund, M., Johansson, B., Hagemeijer, A., Park, S.S., Mitelman, F.*, et al.* (2004). The breakpoint region of the most common isochromosome, i(17q), in human neoplasia is characterized by a complex genomic architecture with large, palindromic, low-copy repeats. Am J Hum Genet *74*, 1-10.

Bedoyan, J.K., Kumar, R.A., Sudi, J., Silverstein, F., Ackley, T., Iyer, R.K., Christian, S.L., and Martin, D.M. (2010). Duplication 16p11.2 in a child with infantile seizure disorder. Am J Med Genet A *152A*, 1567-1574.

Bi, W., Sapir, T., Shchelochkov, O.A., Zhang, F., Withers, M.A., Hunter, J.V., Levy, T., Shinder, V., Peiffer, D.A., Gunderson, K.L.*, et al.* (2009). Increased LIS1 expression affects human and mouse brain development. Nat Genet *41*, 168-177.

Blumenfeld, H. (2005). Cellular and network mechanisms of spike-wave seizures. Epilepsia *46 Suppl 9*, 21-33.

Bonaglia, M.C., Giorda, R., Beri, S., De Agostini, C., Novara, F., Fichera, M., Grillo, L., Galesi, O., Vetro, A., Ciccone, R.*, et al.* (2011). Molecular mechanisms generating and stabilizing terminal 22q13 deletions in 44 subjects with Phelan/McDermid syndrome. PLoS Genet *7*, e1002173.

Bornstein, S.R., Webster, E.L., Torpy, D.J., Richman, S.J., Mitsiades, N., Igel, M., Lewis, D.B., Rice, K.C., Joost, H.G., Tsokos, M.*, et al.* (1998). Chronic effects of a nonpeptide corticotropin-releasing hormone type I receptor antagonist on pituitary-adrenal function, body weight, and metabolic regulation. Endocrinology *139*, 1546-1555.

Bosia, M., Buonocore, M., Guglielmino, C., Pirovano, A., Lorenzi, C., Marcone, A., Bramanti, P., Cappa, S.F., Aguglia, E., Smeraldi, E.*, et al.* (2012). Saitohin polymorphism and executive dysfunction in schizophrenia. Neurol Sci *33*, 1051-1056.

Boyden, E.D., Campos-Xavier, A.B., Kalamajski, S., Cameron, T.L., Suarez, P., Tanackovic, G., Tanackovich, G., Andria, G., Ballhausen, D., Briggs, M.D., *et al.* (2011). Recurrent dominant mutations affecting two adjacent residues in the motor domain of the monomeric kinesin KIF22 result in skeletal dysplasia and joint laxity. Am J Hum Genet *89*, 767-772.

Brault, V., Pereira, P., Duchon, A., and Herault, Y. (2006). Modeling chromosomes in mouse to explore the function of genes, genomic disorders, and chromosomal organization. Plos Genetics *2*, 911-919.

Broadbent, N.J., Gaskin, S., Squire, L.R., and Clark, R.E. (2010). Object recognition memory and the rodent hippocampus. Learn Mem *17*, 5-11.

Burgess, T., Downie, L., Pertile, M.D., Francis, D., Glass, M., Nouri, S., and Pszczola, R. (2014). Monosomy 21 seen in live born is unlikely to represent true monosomy 21: a case report and review of the literature. Case Rep Genet *2014*, 965401.

Bzymek, M., and Lovett, S.T. (2001). Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. Proc Natl Acad Sci U S A *98*, 8319-8325.

Cai, Y., Jin, J., Swanson, S.K., Cole, M.D., Choi, S.H., Florens, L., Washburn, M.P., Conaway, J.W., and Conaway, R.C. (2010). Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex. J Biol Chem *285*, 4268-4272.

Chakrabarti, S., and Fombonne, E. (2005). Pervasive developmental disorders in preschool children: confirmation of high prevalence. Am J Psychiatry *162*, 1133-1141.

Chang, J.J., Liu, C.J., Liu, J.H., Chiou, T.J., Tzeng, C.H., and Chen, P.M. (1992). Monosomy 21 in two patients with acute nonlymphocytic leukemia. Cancer Genet Cytogenet *61*, 122-125.

Chapman, J., Rees, E., Harold, D., Ivanov, D., Gerrish, A., Sims, R., Hollingworth, P., Stretton, A., Holmans, P., Owen, M.J., *et al.* (2013). A genome-wide study shows a limited contribution of rare copy number variants to Alzheimer's disease risk. Hum Mol Genet *22*, 816-824.

Chelly, J., Khelfaoui, M., Francis, F., Chérif, B., and Bienvenu, T. (2006). Genetics and pathophysiology of mental retardation. Eur J Hum Genet *14*, 701-713.

Chen, A.C., Manz, N., Tang, Y., Rangaswamy, M., Almasy, L., Kuperman, S., Nurnberger, J., O'Connor, S.J., Edenberg, H.J., Schuckit, M.A., *et al.* (2010). Single-nucleotide polymorphisms in corticotropin releasing hormone receptor 1 gene (CRHR1) are associated with quantitative trait of event-related potential and alcohol dependence. Alcohol Clin Exp Res *34*, 988-996.

Chen, J., Yu, S., Fu, Y., and Li, X. (2014). Synaptic proteins and receptors defects in autism spectrum disorders. Front Cell Neurosci *8*, 276.

Chen, W.J., Lin, Y., Xiong, Z.Q., Wei, W., Ni, W., Tan, G.H., Guo, S.L., He, J., Chen, Y.F., Zhang, Q.J., *et al.* (2011). Exome sequencing identifies truncating mutations in PRRT2 that cause paroxysmal kinesigenic dyskinesia. Nat Genet *43*, 1252-1255.

Chettouh, Z., Croquette, M.F., Delobel, B., Gilgenkrants, S., Leonard, C., Maunoury, C., Prieur, M., Rethore, M.O., Sinet, P.M., Chery, M., *et al.* (1995). MOLECULAR MAPPING OF 21 FEATURES ASSOCIATED WITH PARTIAL MONOSOMY-21 - INVOLVEMENT OF THE APP-SOD1 REGION. American Journal of Human Genetics *57*, 62-71.

Chubb, J.E., Bradshaw, N.J., Soares, D.C., Porteous, D.J., and Millar, J.K. (2008). The DISC locus in psychiatric illness. Mol Psychiatry *13*, 36-64.

Coe, B.P., Witherspoon, K., Rosenfeld, J.A., van Bon, B.W., Vulto-van Silfhout, A.T., Bosco, P., Friend, K.L., Baker, C., Buono, S., Vissers, L.E., *et al.* (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay. Nat Genet *46*, 1063-1071.

Conrad, C., Vianna, C., Schultz, C., Thal, D.R., Ghebremedhin, E., Lenz, J., Braak, H., and Davies, P. (2004). Molecular evolution and genetics of the Saitohin gene and tau haplotype in Alzheimer's disease and argyrophilic grain disease. J Neurochem *89*, 179-188.

Conrad, D.F., Andrews, T.D., Carter, N.P., Hurles, M.E., and Pritchard, J.K. (2006). A high-resolution survey of deletion polymorphism in the human genome. Nat Genet *38*, 75-81.

Conrad, D.F., Bird, C., Blackburne, B., Lindsay, S., Mamanova, L., Lee, C., Turner, D.J., and Hurles, M.E. (2010a). Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. Nat Genet *42*, 385-391.

Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y.J., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., *et al.* (2010b). Origins and functional impact of copy number variation in the human genome. Nature *464*, 704-712.

Consortium, C.-D.G.o.t.P.G. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet *381*, 1371-1379.

Consortium, I.H. (2003). The International HapMap Project. Nature *426*, 789-796.

Consortium, I.H. (2005). A haplotype map of the human genome. Nature *437*, 1299-1320.

Consortium, I.S. (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature *455*, 237-241.

Cooper, G.M., Coe, B.P., Girirajan, S., Rosenfeld, J.A., Vu, T.H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V., *et al.* (2011). A copy number variation morbidity map of developmental delay. Nature Genetics *43*, 838-U844.

Crawley, J.N. (2007). Mouse behavioral assays relevant to the symptoms of autism. Brain Pathology *17*, 448-459.

Crepel, A., Steyaert, J., De la Marche, W., De Wolf, V., Fryns, J.P., Noens, I., Devriendt, K., and Peeters, H. (2011). Narrowing the Critical Deletion Region for Autism Spectrum Disorders on 16p11.2. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics *156B*, 243-245.

Davisson, M.T., Schmidt, C., Reeves, R.H., Irving, N.G., Akeson, E.C., Harris, B.S., and Bronson, R.T. (1993). Segmental trisomy as a mouse model for Down syndrome. Prog Clin Biol Res *384*, 117-133.

de Anda, F.C., Rosario, A.L., Durak, O., Tran, T., Gräff, J., Meletis, K., Rei, D., Soda, T., Madabhushi, R., Ginty, D.D., *et al.* (2012). Autism spectrum disorder susceptibility gene TAOK2 affects basal dendrite formation in the neocortex. Nat Neurosci *15*, 1022-1031.

de Cid, R., Riveira-Munoz, E., Zeeuwen, P.L., Robarge, J., Liao, W., Dannhauser, E.N., Giardina, E., Stuart, P.E., Nair, R., Helms, C., *et al.* (2009). Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis. Nat Genet *41*, 211-215.

de Kovel, C.G., Trucks, H., Helbig, I., Mefford, H.C., Baker, C., Leu, C., Kluck, C., Muhle, H., von Spiczak, S., Ostertag, P., *et al.* (2010). Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. Brain *133*, 23-32.

de la Chapelle, A., Herva, R., Koivisto, M., and Aula, P. (1981). A deletion in chromosome 22 can cause DiGeorge syndrome. Hum Genet *57*, 253-256.

De la Torre, R., De Sola, S., Pons, M., Duchon, A., de Lagran, M.M., Farré, M., Fitó, M., Benejam, B., Langohr, K., Rodriguez, J., *et al.* (2014). Epigallocatechin-3-gallate, a DYRK1A inhibitor, rescues cognitive deficits in Down syndrome mouse models and in humans. Mol Nutr Food Res *58*, 278-288.

de Smith, A.J., Walters, R.G., Coin, L.J., Steinfeld, I., Yakhini, Z., Sladek, R., Froguel, P., and Blakemore, A.I. (2008). Small deletion variants have stable breakpoints commonly associated with alu elements. PLoS One *3*, e3104.

de Vries, B., Callenbach, P.M., Kamphorst, J.T., Weller, C.M., Koelewijn, S.C., ten Houten, R., de Coo, I.F., Brouwer, O.F., and van den Maagdenberg, A.M. (2012). PRRT2 mutation causes benign familial infantile convulsions. Neurology *79*, 2154-2155.

Deans, B., Griffin, C.S., Maconochie, M., and Thacker, J. (2000). Xrcc2 is required for genetic stability, embryonic neurogenesis and viability in mice. EMBO J *19*, 6675-6685.

DeScipio, C., Conlin, L., Rosenfeld, J., Tepperberg, J., Pasion, R., Patel, A., McDonald, M.T., Aradhya, S., Ho, D., Goldstein, J., *et al.* (2012). Subtelomeric deletion of chromosome 10p15.3: clinical findings and molecular cytogenetic characterization. Am J Med Genet A *158A*, 2152-2161.

Diamandis, E.P., Yousef, G.M., Petraki, C., and Soosaipillai, A.R. (2000). Human kallikrein 6 as a biomarker of Alzheimer's disease. Clinical Biochemistry *33*, 663-667.

Dias, J., Van Nguyen, N., Georgiev, P., Gaub, A., Brettschneider, J., Cusack, S., Kadlec, J., and Akhtar, A. (2014). Structural analysis of the KANSL1/WDR5/KANSL2 complex reveals that WDR5 is required for efficient assembly and chromatin targeting of the NSL complex. Genes Dev *28*, 929-942.

Dibbens, L.M., Mullen, S., Helbig, I., Mefford, H.C., Bayly, M.A., Bellows, S., Leu, C., Trucks, H., Obermeier, T., Wittig, M.*, et al.* (2009). Familial and sporadic 15q13.3 microdeletions in idiopathic generalized epilepsy: precedent for disorders with complex inheritance. Hum Mol Genet *18*, 3626-3631.

Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L., and Roeder, R.G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell *121*, 873-885.

Dubourg, C., Sanlaville, D., Doco-Fenzy, M., Le Caignec, C., Missirian, C., Jaillard, S., Schluth-Bolard, C., Landais, E., Boute, O., Philip, N.*, et al.* (2011). Clinical and molecular characterization of 17q21.31 microdeletion syndrome in 14 French patients with mental retardation. Eur J Med Genet *54*, 144-151.

Dziuba, P., Dziekanowska, D., and Hübner, H. (1976). A female infant with monosomy 21. Hum Genet *31*, 351-353.

Eichler, E.E. (2001). Recent duplication, domain accretion and the dynamic mutation of the human genome. Trends Genet *17*, 661-669.

Elenkov, I.J., Webster, E.L., Torpy, D.J., and Chrousos, G.P. (1999). Stress, corticotropin-releasing hormone, glucocorticoids, and the immune/inflammatory response: acute and chronic effects. Ann N Y Acad Sci *876*, 1-11; discussion 11-13.

Elliott, B., Richardson, C., and Jasin, M. (2005). Chromosomal translocation mechanisms at intronic alu elements in mammalian cells. Mol Cell *17*, 885-894.

Fei, Q., Wu, Z., Wang, H., Zhou, X., Wang, N., Ding, Y., Wang, Y., and Qiu, G. (2010). The association analysis of TBX6 polymorphism with susceptibility to congenital scoliosis in a Chinese Han population. Spine (Phila Pa 1976) *35*, 983-988.

Fernandez, B.A., Roberts, W., Chung, B., Weksberg, R., Meyn, S., Szatmari, P., Joseph-George, A.M., MacKay, S., Whitten, K., Noble, B.*, et al.* (2010). Phenotypic spectrum associated with de novo and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder. Journal of Medical Genetics *47*, 195-203.

Fickie, M.R., Lapunzina, P., Gentile, J.K., Tolkoff-Rubin, N., Kroshinsky, D., Galan, E., Gean, E., Martorell, L., Romanelli, V., Toral, J.F.*, et al.* (2011). Adults with Sotos syndrome: review of 21 adults with molecularly confirmed NSD1 alterations, including a detailed case report of the oldest person. Am J Med Genet A *155A*, 2105-2111.

Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet *84*, 524-533.

Fombonne, E. (2005). Epidemiology of autistic disorder and other pervasive developmental disorders. J Clin Psychiatry *66 Suppl 10*, 3-8.

Foust, K.D., Wang, X., McGovern, V.L., Braun, L., Bevan, A.K., Haidet, A.M., Le, T.T., Morales, P.R., Rich, M.M., Burghes, A.H.*, et al.* (2010). Rescue of the spinal muscular atrophy phenotype in a mouse model by early postnatal delivery of SMN. Nat Biotechnol *28*, 271-274.

Friedman, J.I., Vrijenhoek, T., Markx, S., Janssen, I.M., van der Vliet, W.A., Faas, B.H., Knoers, N.V., Cahn, W., Kahn, R.S., Edelmann, L.*, et al.* (2008). CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy. Mol Psychiatry *13*, 261-266.

Friedman, J.M., Baross, A., Delaney, A.D., Ally, A., Arbour, L., Armstrong, L., Asano, J., Bailey, D.K., Barber, S., Birch, P.*, et al.* (2006). Oligonucleotide microarray analysis of genomic imbalance in children with mental retardation. Am J Hum Genet *79*, 500-513.

Fromer, M., Moran, J.L., Chambert, K., Banks, E., Bergen, S.E., Ruderfer, D.M., Handsaker, R.E., McCarroll, S.A., O'Donovan, M.C., Owen, M.J.*, et al.* (2012). Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. Am J Hum Genet *91*, 597-607.

Fujimoto, A., Towner, J.W., Ebbin, A.J., Kahlstrom, E.J., and Wilson, M.G. (1974). Inherited partial duplication of chromosome No. 15. J Med Genet *11*, 287-291.

Gao, L., Tse, S.W., Conrad, C., and Andreadis, A. (2005). Saitohin, which is nested in the tau locus and confers allele-specific susceptibility to several neurodegenerative diseases, interacts with peroxiredoxin 6. J Biol Chem *280*, 39268-39272.

Garzicić, B., Guć-Sćekić, M., Pilić-Radivojević, G., Ignjatović, M., and Vilhar, N. (1988). A case of monosomy 21. Ann Genet *31*, 247-249.

Ghebranious, N., Giampietro, P.F., Wesbrook, F.P., and Rezkana, S.H. (2007). A novel microdeletion at 16p11.2 harbors candidate genes for aortic valve development, seizure disorder, and mild mental retardation. American Journal of Medical Genetics Part A *143A*, 1462-1471.

Gilissen, C., Hehir-Kwa, J.Y., Thung, D.T., van de Vorst, M., van Bon, B.W., Willemsen, M.H., Kwint, M., Janssen, I.M., Hoischen, A., Schenck, A.*, et al.* (2014). Genome sequencing identifies major causes of severe intellectual disability. Nature *511*, 344-347.

Girirajan, S., and Eichler, E.E. (2010). Phenotypic variability and genetic susceptibility to genomic disorders. Hum Mol Genet *19*, R176-187.

Girirajan, S., Rosenfeld, J.A., Coe, B.P., Parikh, S., Friedman, N., Goldstein, A., Filipink, R.A., McConnell, J.S., Angle, B., Meschino, W.S.*, et al.* (2012). Phenotypic heterogeneity of genomic disorders and rare copy-number variants. N Engl J Med *367*, 1321-1331.

Golzio, C., Willer, J., Talkowski, M.E., Oh, E.C., Taniguchi, Y., Jacquemont, S., Reymond, A., Sun, M., Sawa, A., Gusella, J.F.*, et al.* (2012). KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. Nature *485*, 363-U111.

Grayton, H.M., Fernandes, C., Rujescu, D., and Collier, D.A. (2012). Copy number variations in neurodevelopmental disorders. Progress in Neurobiology *99*, 81-91.

Greenberg, F. (1993). DiGeorge syndrome: an historical review of clinical and cytogenetic features. J Med Genet *30*, 803-806.

Grisart, B., Willatt, L., Destrée, A., Fryns, J.P., Rack, K., de Ravel, T., Rosenfeld, J., Vermeesch, J.R., Verellen-Dumoulin, C., and Sandford, R. (2009). 17q21.31 microduplication patients are characterised by behavioural problems and poor social interaction. J Med Genet *46*, 524-530.

Griswold, A.J., Ma, D., Cukier, H.N., Nations, L.D., Schmidt, M.A., Chung, R.H., Jaworski, J.M., Salyakina, D., Konidari, I., Whitehead, P.L.*, et al.* (2012). Evaluation of copy number variations reveals novel candidate genes in autism spectrum disorder-associated pathways. Hum Mol Genet *21*, 3513-3523.

Gu, W., Zhang, F., and Lupski, J.R. (2008). Mechanisms for human genomic rearrangements. Pathogenetics *1*, 4.

Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A., and McKusick, V.A. (2005). Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res *33*, D514-517.

Handsaker, R.E., Korn, J.M., Nemesh, J., and McCarroll, S.A. (2011). Discovery and genotyping of genome structural polymorphism by sequencing on a population scale. Nat Genet *43*, 269-276.

Hart, H. (2008). 'Puppet' children. A report on three cases (1965). Dev Med Child Neurol *50*, 564.

Hastings, P.J., Ira, G., and Lupski, J.R. (2009a). A microhomology-mediated break-induced replication model for the origin of human copy number variation. PLoS Genet *5*, e1000327.

Hastings, P.J., Lupski, J.R., Rosenberg, S.M., and Ira, G. (2009b). Mechanisms of change in gene copy number. Nature Reviews Genetics *10*, 551-564.

Hauser, W.A., Annegers, J.F., and Kurland, L.T. (1993). Incidence of epilepsy and unprovoked seizures in Rochester, Minnesota: 1935-1984. Epilepsia *34*, 453-468.

Hauser, W.A., Annegers, J.F., and Rocca, W.A. (1996). Descriptive epidemiology of epilepsy: contributions of population-based studies from Rochester, Minnesota. Mayo Clin Proc *71*, 576-586.

Haviv-Chesner, A., Kobayashi, Y., Gabriel, A., and Kupiec, M. (2007). Capture of linear fragments at a double-strand break in yeast. Nucleic Acids Res *35*, 5192-5202.

He, H., Tan, C.K., Downey, K.M., and So, A.G. (2001). A tumor necrosis factor alpha- and interleukin 6-inducible protein that interacts with the small subunit of DNA polymerase delta and proliferating cell nuclear antigen. Proc Natl Acad Sci U S A *98*, 11979-11984.

Helbig, I., Mefford, H.C., Sharp, A.J., Guipponi, M., Fichera, M., Franke, A., Muhle, H., de Kovel, C., Baker, C., von Spiczak, S.*, et al.* (2009). 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. Nat Genet *41*, 160-162.

Helbig, I., Scheffer, I.E., Mulley, J.C., and Berkovic, S.F. (2008). Navigating the channels and beyond: unravelling the genetics of the epilepsies. Lancet Neurol *7*, 231-245.

Herault, Y., Rassoulzadegan, M., Cuzin, F., and Duboule, D. (1998). Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). Nature Genetics *20*, 381-384.

Hildebrand, M.S., Dahl, H.H., Damiano, J.A., Smith, R.J., Scheffer, I.E., and Berkovic, S.F. (2013). Recent advances in the molecular genetics of epilepsy. J Med Genet *50*, 271-279.

Hopman, S., Merks, J., Eussen, H., Douben, H., Snijder, S., Hennekam, R., de Klein, A., and Caron, H. (2013). Structural genome variations in individuals with childhood cancer and tumour predisposition syndromes. Eur J Cancer *49*, 2170-2178.

Horev, G., Ellegood, J., Lerch, J.P., Son, Y.E.E., Muthuswamy, L., Vogel, H., Krieger, A.M., Buja, A., Henkelman, R.M., Wigler, M.*, et al.* (2011). Dosage-dependent phenotypes in models of 16p11.2 lesions found in autism. Proceedings of the National Academy of Sciences of the United States of America *108*, 17076-17081.

Iafrate, A.J., Feuk, L., Rivera, M.N., Listewnik, M.L., Donahoe, P.K., Qi, Y., Scherer, S.W., and Lee, C. (2004). Detection of large-scale variation in the human genome. Nat Genet *36*, 949-951.

Izquierdo, M.A., Shoemaker, R.H., Flens, M.J., Scheffer, G.L., Wu, L., Prather, T.R., and Scheper, R.J. (1996). Overlapping phenotypes of multidrug resistance among panels of human cancer-cell lines. Int J Cancer *65*, 230-237.

Jacobsen, P., Hauge, M., Henningsen, K., Hobolth, N., Mikkelsen, M., and Philip, J. (1973). An (11;21) translocation in four generations with chromosome 11 abnormalities in the offspring. A clinical, cytogenetical, and gene marker study. Hum Hered *23*, 568-585.

Jacquemont, S., Reymond, A., Zufferey, F., Harewood, L., Walters, R.G., Kutalik, Z., Martinet, D., Shen, Y.P., Valsesia, A., Beckmann, N.D.*, et al.* (2011). Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. Nature *478*, 97-U111.

Jalagadugula, G., Mao, G.F., Kaur, G., Goldfinger, L.E., Dhanasekaran, D.N., and Rao, A.K. (2010). Regulation of platelet myosin light chain (MYL9) by RUNX1: implications for thrombocytopenia and platelet dysfunction in RUNX1 haplodeficiency. Blood *116*, 6037-6045.

Jallon, P., Loiseau, P., and Loiseau, J. (2001). Newly diagnosed unprovoked epileptic seizures: presentation at diagnosis in CAROLE study. Coordination Active du Réseau Observatoire Longitudinal de l' Epilepsie. Epilepsia *42*, 464-475.

Jarvis, S.E., and Zamponi, G.W. (2005). Masters or slaves? Vesicle release machinery and the regulation of presynaptic calcium channels. Cell Calcium *37*, 483-488.

Kallioniemi, A., Kallioniemi, O.P., Sudar, D., Rutovitz, D., Gray, J.W., Waldman, F., and Pinkel, D. (1992). Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science *258*, 818-821.

Kamnasaran, D., Muir, W.J., Ferguson-Smith, M.A., and Cox, D.W. (2003). Disruption of the neuronal PAS3 gene in a family affected with schizophrenia. J Med Genet *40*, 325-332.

Kaneko, Y., Ikeuchi, T., Sasaki, M., Stakae, Y., and Kuwajima, S. (1975). A male infant with monosomy 21. Humangenetik *29*, 1-7.

Karayiorgou, M., Morris, M.A., Morrow, B., Shprintzen, R.J., Goldberg, R., Borrow, J., Gos, A., Nestadt, G., Wolyniec, P.S., and Lasseter, V.K. (1995). Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. Proc Natl Acad Sci U S A *92*, 7612-7616.

Katz, G., and Lazcano-Ponce, E. (2008). Intellectual disability: definition, etiological factors, classification, diagnosis, treatment and prognosis. Salud Publica Mex *50 Suppl 2*, s132-141.

Kelley, R.I., Zackai, E.H., Emanuel, B.S., Kistenmacher, M., Greenberg, F., and Punnett, H.H. (1982). The association of the DiGeorge anomalad with partial monosomy of chromosome 22. J Pediatr *101*, 197-200.

Kempf, M., Clement, A., Faissner, A., Lee, G., and Brandt, R. (1996). Tau binds to the distal axon early in development of polarity in a microtubule- and microfilament-dependent manner. J Neurosci *16*, 5583-5592.

Kerr, T.M., Muller, C.L., Miah, M., Jetter, C.S., Pfeiffer, R., Shah, C., Baganz, N., Anderson, G.M., Crawley, J.N., Sutcliffe, J.S.*, et al.* (2013). Genetic background modulates phenotypes of serotonin transporter Ala56 knock-in mice. Mol Autism *4*, 35.

Kirchhoff, M., Bisgaard, A.M., Duno, M., Hansen, F.J., and Schwartz, M. (2007). A 17q21.31 microduplication, reciprocal to the newly described 17q21.31 microdeletion, in a girl with severe psychomotor developmental delay and dysmorphic craniofacial features. Eur J Med Genet *50*, 256-263.

Kirov, G., Gumus, D., Chen, W., Norton, N., Georgieva, L., Sari, M., O'Donovan, M.C., Erdogan, F., Owen, M.J., Ropers, H.H.*, et al.* (2008). Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. Hum Mol Genet *17*, 458-465.

Kitsiou-Tzeli, S., Frysira, H., Giannikou, K., Syrmou, A., Kosma, K., Kakourou, G., Leze, E., Sofocleous, C., Kanavakis, E., and Tzetis, M. (2012). Microdeletion and microduplication 17q21.31 plus an additional CNV, in patients with intellectual disability, identified by array-CGH. Gene *492*, 319-324.

Kjeldsen, M.J., Corey, L.A., Christensen, K., and Friis, M.L. (2003). Epileptic seizures and syndromes in twins: the importance of genetic factors. Epilepsy Res *55*, 137-146.

Kleefstra, T., Schenck, A., Kramer, J.M., and van Bokhoven, H. (2014). The genetics of cognitive epigenetics. Neuropharmacology *80*, 83-94.

Kleinjan, D.A., and van Heyningen, V. (2005). Long-range control of gene expression: emerging mechanisms and disruption in disease. Am J Hum Genet *76*, 8-32.

Knoll, J.H., Nicholls, R.D., Magenis, R.E., Graham, J.M., Lalande, M., and Latt, S.A. (1989). Angelman and Prader-Willi syndromes share a common chromosome 15 deletion but differ in parental origin of the deletion. Am J Med Genet *32*, 285-290.

Kojima, T., Fukuda, M., Aruga, J., and Mikoshiba, K. (1996). Calcium-dependent phospholipid binding to the C2A domain of a ubiquitous form of double C2 protein (Doc2 beta). J Biochem *120*, 671-676.

Koolen, D.A., Kramer, J.M., Neveling, K., Nillesen, W.M., Moore-Barton, H.L., Elmslie, F.V., Toutain, A., Amiel, J., Malan, V., Tsai, A.C.*, et al.* (2012). Mutations in the chromatin modifier gene KANSL1 cause the 17q21.31 microdeletion syndrome. Nat Genet *44*, 639-641.

Koolen, D.A., Pfundt, R., de Leeuw, N., Hehir-Kwa, J.Y., Nillesen, W.M., Neefs, I., Scheltinga, I., Sistermans, E., Smeets, D., Brunner, H.G.*, et al.* (2009). Genomic microarrays in mental retardation: a practical workflow for diagnostic applications. Hum Mutat *30*, 283-292.

Koolen, D.A., Sharp, A.J., Hurst, J.A., Firth, H.V., Knight, S.J., Goldenberg, A., Saugier-Veber, P., Pfundt, R., Vissers, L.E., Destrée, A.*, et al.* (2008). Clinical and molecular delineation of the 17q21.31 microdeletion syndrome. J Med Genet *45*, 710-720.

Koolen, D.A., Vissers, L., Pfundt, R., de Leeuw, N., Knight, S.J.L., Regan, R., Kooy, R.F., Reyniers, E., Romano, C., Fichera, M.*, et al.* (2006). A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. Nature Genetics *38*, 999-1001.

Korbel, J.O., Urban, A.E., Affourtit, J.P., Godwin, B., Grubert, F., Simons, J.F., Kim, P.M., Palejev, D., Carriero, N.J., Du, L.*, et al.* (2007). Paired-end mapping reveals extensive structural variation in the human genome. Science *318*, 420-426.

Koszul, R., and Fischer, G. (2009). A prominent role for segmental duplications in modeling eukaryotic genomes. C R Biol *332*, 254-266.

Kumar, R.A., KaraMohamed, S., Sudi, J., Conrad, D.F., Brune, C., Badner, J.A., Gilliam, T.C., Nowak, N.J., Cook, E.H., Dobyns, W.B.*, et al.* (2008). Recurrent 16p11.2 microdeletions in autism. Human Molecular Genetics *17*, 628-638.

Kumar, R.A., Marshall, C.R., Badner, J.A., Babatz, T.D., Mukamel, Z., Aldinger, K.A., Sudi, J., Brune, C.W., Goh, G., Karamohamed, S., *et al.* (2009). Association and mutation analyses of 16p11.2 autism candidate genes. PLoS One *4*, e4582.

Kurotaki, N., Shen, J.J., Touyama, M., Kondoh, T., Visser, R., Ozaki, T., Nishimoto, J., Shiihara, T., Uetake, K., Makita, Y., *et al.* (2005). Phenotypic consequences of genetic variation at hemizygous alleles: Sotos syndrome is a contiguous gene syndrome incorporating coagulation factor twelve (FXII) deficiency. Genet Med *7*, 479-483.

Lai, M.C., Lombardo, M.V., and Baron-Cohen, S. (2014). Autism. Lancet *383*, 896-910.

Lara, P.C., Pruschy, M., Zimmermann, M., and Henríquez-Hernández, L.A. (2011). MVP and vaults: a role in the radiation response. Radiat Oncol *6*, 148.

Laurençot, C.M., Scheffer, G.L., Scheper, R.J., and Shoemaker, R.H. (1997). Increased LRP mRNA expression is associated with the MDR phenotype in intrinsically resistant human cancer cell lines. Int J Cancer *72*, 1021-1026.

Lee, S.H., Ripke, S., Neale, B.M., Faraone, S.V., Purcell, S.M., Perlis, R.H., Mowry, B.J., Thapar, A., Goddard, M.E., Witte, J.S., *et al.* (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat Genet *45*, 984-994.

Lejeune, J., Berger, R., Rethore, M.O., Archamba.L, Jerome, H., Thieffry, S., Aicardi, J., Broyer, M., Lafourca.J, Cruveill.J, *et al.* (1964). MONOSOMIE PARTIELLE POUR UN PETIT ACROCENTRIQUE. Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences *259*, 4187-&.

Leonard, H., and Wen, X. (2002). The epidemiology of mental retardation: challenges and opportunities in the new millennium. Ment Retard Dev Disabil Res Rev *8*, 117-134.

Leroy, C., Landais, E., Briault, S., David, A., Tassy, O., Gruchy, N., Delobel, B., Grégoire, M.J., Leheup, B., Taine, L., *et al.* (2013). The 2q37-deletion syndrome: an update of the clinical spectrum including overweight, brachydactyly and behavioural features in 14 new patients. Eur J Hum Genet *21*, 602-612.

Lieber, M.R. (2008). The mechanism of human nonhomologous DNA end joining. J Biol Chem *283*, 1-5.

Liew, E., and Owen, C. (2011). Familial myelodysplastic syndromes: a review of the literature. Haematologica-the Hematology Journal *96*, 1536-1542.

Lin, F.L., Sperle, K., and Sternberg, N. (1984). Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. Mol Cell Biol *4*, 1020-1034.

Lindstrand, A., Malmgren, H., Sahlen, S., Schoumans, J., Nordgren, A., Ergander, U., Holm, E., Anderlid, B.M., and Blennow, E. (2010). Detailed molecular and clinical characterization of three patients with 21q deletions. Clinical Genetics *77*, 145-154.

Liu, B., Wang, T., Wang, L., Wang, C., Zhang, H., and Gao, G.D. (2011a). Up-regulation of major vault protein in the frontal cortex of patients with intractable frontal lobe epilepsy. J Neurol Sci *308*, 88-93.

Liu, P., Carvalho, C.M., Hastings, P.J., and Lupski, J.R. (2012). Mechanisms for recurrent and complex human genomic rearrangements. Curr Opin Genet Dev *22*, 211-220.

Liu, P., Erez, A., Nagamani, S.C., Dhar, S.U., Kołodziejska, K.E., Dharmadhikari, A.V., Cooper, M.L., Wiszniewska, J., Zhang, F., Withers, M.A., *et al.* (2011b). Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. Cell *146*, 889-903.

Lohmueller, K.E., Pearce, C.L., Pike, M., Lander, E.S., and Hirschhorn, J.N. (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet *33*, 177-182.

Lowery, M.C., Morris, C.A., Ewart, A., Brothman, L.J., Zhu, X.L., Leonard, C.O., Carey, J.C., Keating, M., and Brothman, A.R. (1995). Strong correlation of elastin deletions, detected by FISH, with Williams syndrome: evaluation of 235 patients. Am J Hum Genet *57*, 49-53.

Lowery, R.L., and Majewska, A.K. (2010). Intracranial injection of adeno-associated viral vectors. J Vis Exp.

Lu, X., Ye, K., Zou, K., and Chen, J. (2014). Identification of copy number variation-driven genes for liver cancer via bioinformatics analysis. Oncol Rep.

Lupski, J.R. (1998). Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. Trends Genet *14*, 417-422.

Lupski, J.R., and Stankiewicz, P. (2005). Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet *1*, e49.

Lupski, J.R., Wise, C.A., Kuwano, A., Pentao, L., Parke, J.T., Glaze, D.G., Ledbetter, D.H., Greenberg, F., and Patel, P.I. (1992). Gene dosage is a mechanism for Charcot-Marie-Tooth disease type 1A. Nat Genet *1*, 29-33.

Lyle, R., Bena, F., Gagos, S., Gehrig, C., Lopez, G., Schinzel, A., Lespinasse, J., Bottani, A., Dahoun, S., Taine, L.*, et al.* (2009). Genotype-phenotype correlations in Down syndrome identified by array CGH in 30 cases of partial trisomy and partial monosomy chromosome 21. European Journal of Human Genetics *17*, 454-466.

MacDonald, J.R., Ziman, R., Yuen, R.K., Feuk, L., and Scherer, S.W. (2014). The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic Acids Res *42*, D986-992.

MacIntyre, D.J., Blackwood, D.H., Porteous, D.J., Pickard, B.S., and Muir, W.J. (2003). Chromosomal abnormalities and mental illness. Mol Psychiatry *8*, 275-287.

Mandillo, S., Tucci, V., Hölter, S.M., Meziane, H., Banchaabouchi, M.A., Kallnik, M., Lad, H.V., Nolan, P.M., Ouagazzal, A.M., Coghill, E.L.*, et al.* (2008). Reliability, robustness, and reproducibility in mouse behavioral phenotyping: a cross-laboratory study. Physiol Genomics *34*, 243-255.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A.*, et al.* (2009). Finding the missing heritability of complex diseases. Nature *461*, 747-753.

Marini, C., Conti, V., Mei, D., Battaglia, D., Lettori, D., Losito, E., Bruccini, G., Tortorella, G., and Guerrini, R. (2012). PRRT2 mutations in familial infantile seizures, paroxysmal dyskinesia, and hemiplegic migraine. Neurology *79*, 2109-2114.

Maroun, C., Schmerler, S., and Hutcheon, R.G. (1994). Child with Sotos phenotype and a 5:15 translocation. Am J Med Genet *50*, 291-293.

Marseglia, G., Scordo, M.R., Pescucci, C., Nannetti, G., Biagini, E., Scandurra, V., Gerundino, F., Magi, A., Benelli, M., and Torricelli, F. (2012). 372 kb microdeletion in 18q12.3 causing SETBP1 haploinsufficiency associated with mild mental retardation and expressive speech impairment. Eur J Med Genet *55*, 216-221.

Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y.*, et al.* (2008). Structural variation of chromosomes in autism spectrum disorder. American Journal of Human Genetics *82*, 477-488.

Maulik, P.K., Mascarenhas, M.N., Mathers, C.D., Dua, T., and Saxena, S. (2011). Prevalence of intellectual disability: a meta-analysis of population-based studies. Res Dev Disabil *32*, 419-436.

McCarroll, S.A., Huett, A., Kuballa, P., Chilewski, S.D., Landry, A., Goyette, P., Zody, M.C., Hall, J.L., Brant, S.R., Cho, J.H.*, et al.* (2008). Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease. Nat Genet *40*, 1107-1112.

McCarthy, S.E., Makarov, V., Kirov, G., Addington, A.M., McClellan, J., Yoon, S., Perkins, D.O., Dickel, D.E., Kusenda, M., Krastoshevsky, O.*, et al.* (2009). Microduplications of 16p11.2 are associated with schizophrenia. Nature Genetics *41*, 1223-U1285.

McVey, M., and Lee, S.E. (2008). MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet *24*, 529-538.

Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J.*, et al.* (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila. Mol Cell *21*, 811-823.

Migdalska, A.M., van der Weyden, L., Ismail, O., White, J.K., Sanchez-Andrade, G., Logan, D.W., Arends, M.J., Adams, D.J., and Sanger Mouse Genetics, P. (2012). Modeling Partial Monosomy for Human Chromosome 21q11.2-q21.1 Reveals Haploinsufficient Genes Influencing Behavior and Fat Deposition. Plos One *7*.

Miller, D.T., Adam, M.P., Aradhya, S., Biesecker, L.G., Brothman, A.R., Carter, N.P., Church, D.M., Crolla, J.A., Eichler, E.E., Epstein, C.J.*, et al.* (2010). Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet *86*, 749-764.

Mills, R.E., Walter, K., Stewart, C., Handsaker, R.E., Chen, K., Alkan, C., Abyzov, A., Yoon, S.C., Ye, K., Cheetham, R.K., *et al.* (2011). Mapping copy number variation by population-scale genome sequencing. Nature *470*, 59-65.

Min, B.J., Kim, N., Chung, T., Kim, O.H., Nishimura, G., Chung, C.Y., Song, H.R., Kim, H.W., Lee, H.R., Kim, J., *et al.* (2011). Whole-exome sequencing identifies mutations of KIF22 in spondyloepimetaphyseal dysplasia with joint laxity, leptodactylic type. Am J Hum Genet *89*, 760-766.

Mochida, S., Orita, S., Sakaguchi, G., Sasaki, T., and Takai, Y. (1998). Role of the Doc2 alpha-Munc13-1 interaction in the neurotransmitter release process. Proc Natl Acad Sci U S A *95*, 11418-11422.

Moon, H.J., Yim, S.V., Lee, W.K., Jeon, Y.W., Kim, Y.H., Ko, Y.J., Lee, K.S., Lee, K.H., Han, S.I., and Rha, H.K. (2006). Identification of DNA copy-number aberrations by array-comparative genomic hybridization in patients with schizophrenia. Biochem Biophys Res Commun *344*, 531-539.

Morris, R.G., Garrud, P., Rawlins, J.N., and O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. Nature *297*, 681-683.

Mossink, M.H., van Zon, A., Scheper, R.J., Sonneveld, P., and Wiemer, E.A. (2003). Vaults: a ribonucleoprotein particle involved in drug resistance? Oncogene *22*, 7458-7467.

Mullen, S.A., Carvill, G.L., Bellows, S., Bayly, M.A., Trucks, H., Lal, D., Sander, T., Berkovic, S.F., Dibbens, L.M., Scheffer, I.E., *et al.* (2013). Copy number variants are frequent in genetic generalized epilepsy with intellectual disability. Neurology *81*, 1507-1514.

Nickerson, E., Greenberg, F., Keating, M.T., McCaskill, C., and Shaffer, L.G. (1995). Deletions of the elastin gene at 7q11.23 occur in approximately 90% of patients with Williams syndrome. Am J Hum Genet *56*, 1156-1161.

Nobile, C., and Striano, P. (2014). PRRT2: A major cause of infantile epilepsy and other paroxysmal disorders of childhood. Prog Brain Res *213*, 141-158.

Nomura, J., and Takumi, T. (2012). Animal models of psychiatric disorders that reflect human copy number variation. Neural Plast *2012*, 589524.

Olson, L.E., Roper, R.J., Sengstaken, C.L., Peterson, E.A., Aquino, V., Galdzicki, Z., Siarey, R., Pletnikov, M., Moran, T.H., and Reeves, R.H. (2007). Trisomy for the Down syndrome 'critical region' is necessary but not sufficient for brain phenotypes of trisomic mice. Human Molecular Genetics *16*, 774-782.

Ou, Z., Stankiewicz, P., Xia, Z., Breman, A.M., Dawson, B., Wiszniewska, J., Szafranski, P., Cooper, M.L., Rao, M., Shao, L., *et al.* (2011). Observation and prediction of recurrent human translocations mediated by NAHR between nonhomologous chromosomes. Genome Res *21*, 33-46.

Pankratz, N., Dumitriu, A., Hetrick, K.N., Sun, M., Latourelle, J.C., Wilk, J.B., Halter, C., Doheny, K.F., Gusella, J.F., Nichols, W.C., *et al.* (2011). Copy number variation in familial Parkinson disease. PLoS One *6*, e20988.

Papapetrou, C., Putt, W., Fox, M., and Edwards, Y.H. (1999). The human TBX6 gene: cloning and assignment to chromosome 16p11.2. Genomics *55*, 238-241.

Pelzer, N., de Vries, B., Kamphorst, J.T., Vijfhuizen, L.S., Ferrari, M.D., Haan, J., van den Maagdenberg, A.M., and Terwindt, G.M. (2014). PRRT2 and hemiplegic migraine: a complex association. Neurology *83*, 288-290.

Pietropaolo, S., Guilleminot, A., Martin, B., D'Amato, F.R., and Crusio, W.E. (2011). Genetic-background modulation of core and variable autistic-like symptoms in Fmr1 knock-out mice. PLoS One *6*, e17073.

Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segraves, R., Lucas, J., and Gray, J. (1988). Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc Natl Acad Sci U S A *85*, 9138-9142.

Portmann, T., Yang, M., Mao, R., Panagiotakos, G., Ellegood, J., Dolen, G., Bader, P.L., Grueter, B.A., Goold, C., Fisher, E., *et al.* (2014). the Basal Ganglia of a Mouse Model of 16p11.2 Deletion Syndrome. Cell Reports *7*, 1077-1092.

Potocki, L., Chen, K.S., Park, S.S., Osterholm, D.E., Withers, M.A., Kimonis, V., Summers, A.M., Meschino, W.S., Anyane-Yeboa, K., Kashork, C.D., *et al.* (2000). Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. Nat Genet *24*, 84-87.

Prokopenko, I., Langenberg, C., Florez, J.C., Saxena, R., Soranzo, N., Thorleifsson, G., Loos, R.J., Manning, A.K., Jackson, A.U., Aulchenko, Y.*, et al.* (2009). Variants in MTNR1B influence fasting glucose levels. Nat Genet *41*, 77-81.

Raja, S.J., Charapitsa, I., Conrad, T., Vaquerizas, J.M., Gebhardt, P., Holz, H., Kadlec, J., Fraterman, S., Luscombe, N.M., and Akhtar, A. (2010). The nonspecific lethal complex is a transcriptional regulator in Drosophila. Mol Cell *38*, 827-841.

Raveau, M., Lignon, J.M., Nalesso, V., Duchon, A., Groner, Y., Sharp, A.J., Dembele, D., Brault, V., and Herault, Y. (2012). The App-Runx1 Region Is Critical for Birth Defects and Electrocardiographic Dysfunctions Observed in a Down Syndrome Mouse Model. Plos Genetics *8*.

Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W.W.*, et al.* (2006). Global variation in copy number in the human genome. Nature *444*, 444-454.

Reeves, R.H., Irving, N.G., Moran, T.H., Wohn, A., Kitt, C., Sisodia, S.S., Schmidt, C., Bronson, R.T., and Davisson, M.T. (1995). A mouse model for Down syndrome exhibits learning and behaviour deficits. Nat Genet *11*, 177-184.

Reid, C.A., Berkovic, S.F., and Petrou, S. (2009). Mechanisms of human inherited epilepsies. Prog Neurobiol *87*, 41-57.

Reinthaler, E.M., Lal, D., Lebon, S., Hildebrand, M.S., Dahl, H.H., Regan, B.M., Feucht, M., Steinböck, H., Neophytou, B., Ronen, G.M.*, et al.* (2014). 16p11.2 600 kb Duplications confer risk for typical and atypical Rolandic epilepsy. Hum Mol Genet.

Ricard, G., Molina, J., Chrast, J., Gu, W., Gheldof, N., Pradervand, S., Schütz, F., Young, J.I., Lupski, J.R., Reymond, A.*, et al.* (2010). Phenotypic consequences of copy number variation: insights from Smith-Magenis and Potocki-Lupski syndrome mouse models. PLoS Biol *8*, e1000543.

Richtsmeier, J.T., Baxter, L.L., and Reeves, R.H. (2000). Parallels of craniofacial maldevelopment in Down syndrome and Ts65Dn mice. Dev Dyn *217*, 137-145.

Riedel, G., Micheau, J., Lam, A.G., Roloff, E.L., Martin, S.J., Bridge, H., de Hoz, L., Poeschel, B., McCulloch, J., and Morris, R.G. (1999). Reversible neural inactivation reveals hippocampal participation in several memory processes. Nat Neurosci *2*, 898-905.

Ring, S., Weyer, S.W., Kilian, S.B., Waldron, E., Pietrzik, C.U., Filippov, M.A., Herms, J., Buchholz, C., Eckman, C.B., Korte, M.*, et al.* (2007). The secreted beta-amyloid precursor protein ectodomain APPs alpha is sufficient to rescue the anatomical, behavioral, and electrophysiological abnormalities of APP-deficient mice. Journal of Neuroscience *27*, 7817-7826.

Risch, N., and Merikangas, K. (1996). The future of genetic studies of complex human diseases. Science *273*, 1516-1517.

Roberson, E.D.O., Wohler, E.S., Hoover-Fong, J.E., Lisi, E., Stevens, E.L., Thomas, G.H., Leonard, J., Hamosh, A., and Pevsner, J. (2011). Genomic analysis of partial 21q monosomies with variable phenotypes. European Journal of Human Genetics *19*, 235-238.

Rodríguez-Caballero, A., Torres-Lagares, D., Rodríguez-Pérez, A., Serrera-Figallo, M.A., Hernández-Guisado, J.M., and Machuca-Portillo, G. (2010). Cri du chat syndrome: a critical review. Med Oral Patol Oral Cir Bucal *15*, e473-478.

Rogers, J., Raveendran, M., Fawcett, G.L., Fox, A.S., Shelton, S.E., Oler, J.A., Cheverud, J., Muzny, D.M., Gibbs, R.A., Davidson, R.J.*, et al.* (2013). CRHR1 genotypes, neural circuits and the diathesis for anxiety and depression. Mol Psychiatry *18*, 700-707.

Roll, P., Rudolf, G., Pereira, S., Royer, B., Scheffer, I.E., Massacrier, A., Valenti, M.P., Roeckel-Trevisiol, N., Jamali, S., Beclin, C.*, et al.* (2006). SRPX2 mutations in disorders of language cortex and cognition. Hum Mol Genet *15*, 1195-1207.

Ropers, H.H. (2010). Genetics of early onset cognitive impairment. Annu Rev Genomics Hum Genet *11*, 161-187.

Saha, S., Chant, D., Welham, J., and McGrath, J. (2005). A systematic review of the prevalence of schizophrenia. PLoS Med *2*, e141.

Sahún, I., Marechal, D., Lopes Pereira, P., Nalesso, V., Gruart, A., Delgado Garcia, J.M., Antonarakis, S.E., Dierssen, M., and Herault, Y. (2014). Cognition and Hippocampal Plasticity in the Mouse Is Altered by Monosomy of a Genomic Region Implicated in Down Syndrome. Genetics.

Schaaf, C.P., Goin-Kochel, R.P., Nowell, K.P., Hunter, J.V., Aleck, K.A., Cox, S., Patel, A., Bacino, C.A., and Shinawi, M. (2011). Expanding the clinical spectrum of the 16p11.2 chromosomal rearrangements: three patients with syringomyelia. Eur J Hum Genet *19*, 152-156.

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J.*, et al.* (2007). Strong association of de novo copy number mutations with autism. Science *316*, 445-449.

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Månér, S., Massa, H., Walker, M., Chi, M.*, et al.* (2004). Large-scale copy number polymorphism in the human genome. Science *305*, 525-528.

Sharp, A.J., Locke, D.P., McGrath, S.D., Cheng, Z., Bailey, J.A., Vallente, R.U., Pertz, L.M., Clark, R.A., Schwartz, S., Segraves, R.*, et al.* (2005). Segmental duplications and copy-number variation in the human genome. Am J Hum Genet *77*, 78-88.

Sharp, A.J., Mefford, H.C., Li, K., Baker, C., Skinner, C., Stevenson, R.E., Schroer, R.J., Novara, F., De Gregori, M., Ciccone, R.*, et al.* (2008). A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. Nat Genet *40*, 322-328.

Sharp, A.J., Selzer, R.R., Veltman, J.A., Gimelli, S., Gimelli, G., Striano, P., Coppola, A., Regan, R., Price, S.M., Knoers, N.V.*, et al.* (2007). Characterization of a recurrent 15q24 microdeletion syndrome. Hum Mol Genet *16*, 567-572.

She, X.W., Horvath, J.E., Jiang, Z.S., Liu, G., Furey, T.S., Christ, L., Clark, R., Graves, T., Gulden, C.L., Alkan, C.*, et al.* (2004). The structure and evolution of centromeric transition regions within the human genome. Nature *430*, 857-864.

Sheen, C.R., Jewell, U.R., Morris, C.M., Brennan, S.O., Férec, C., George, P.M., Smith, M.P., and Chen, J.M. (2007). Double complex mutations involving F8 and FUNDC2 caused by distinct break-induced replication. Hum Mutat *28*, 1198-1206.

Shen, Y., Chen, X., Wang, L., Guo, J., Shen, J., An, Y., Zhu, H., Zhu, Y., Xin, R., Bao, Y.*, et al.* (2011). Intra-family phenotypic heterogeneity of 16p11.2 deletion carriers in a three-generation Chinese family. Am J Med Genet B Neuropsychiatr Genet *156*, 225-232.

Shibata, H., Shibata, A., Ninomiya, H., Tashiro, N., and Fukumaki, Y. (2002). Association study of polymorphisms in the GluR6 kainate receptor gene (GRIK2) with schizophrenia. Psychiatry Res *113*, 59-67.

Shimojima, K., Inoue, T., Fujii, Y., Ohno, K., and Yamamoto, T. (2009). A familial 593-kb microdeletion of 16p11.2 associated with mental retardation and hemivertebrae. European Journal of Medical Genetics *52*, 433-435.

Shinawi, M., Erez, A., Shardy, D.L., Lee, B., Naeem, R., Weissenberger, G., Chinault, A.C., Cheung, S.W., and Plon, S.E. (2008). Syndromic thrombocytopenia and predisposition to acute myelogenous leukemia caused by constitutional microdeletions on chromosome 21q. Blood *112*, 1042-1047.

Shinawi, M., Liu, P.F., Kang, S.H.L., Shen, J., Belmont, J.W., Scott, D.A., Probst, F.J., Craigen, W.J., Graham, B.H., Pursley, A.*, et al.* (2010). Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. Journal of Medical Genetics *47*, 332-341.

Sisodiya, S.M., Martinian, L., Scheffer, G.L., van der Valk, P., Cross, J.H., Scheper, R.J., Harding, B.N., and Thom, M. (2003). Major vault protein, a marker of drug resistance, is upregulated in refractory epilepsy. Epilepsia *44*, 1388-1396.

Slack, A., Thornton, P.C., Magner, D.B., Rosenberg, S.M., and Hastings, P.J. (2006). On the mechanism of gene amplification induced under stress in Escherichia coli. PLoS Genet *2*, e48.

Smith, A.C., McGavran, L., Robinson, J., Waldstein, G., Macfarlane, J., Zonona, J., Reiss, J., Lahr, M., Allen, L., and Magenis, E. (1986). Interstitial deletion of (17)(p11.2p11.2) in nine patients. Am J Med Genet *24*, 393-414.

SOTOS, J.F., DODGE, P.R., MUIRHEAD, D., CRAWFORD, J.D., and TALBOT, N.B. (1964). CEREBRAL GIGANTISM IN CHILDHOOD. A SYNDROME OF EXCESSIVELY RAPID GROWTH AND ACROMEGALIC FEATURES AND A NONPROGRESSIVE NEUROLOGIC DISORDER. N Engl J Med *271*, 109-116.

Sparrow, D.B., McInerney-Leo, A., Gucev, Z.S., Gardiner, B., Marshall, M., Leo, P.J., Chapman, D.L., Tasic, V., Shishko, A., Brown, M.A.*, et al.* (2013). Autosomal dominant spondylocostal dysostosis is caused by mutation in TBX6. Hum Mol Genet *22*, 1625-1631.

Spillantini, M.G., and Goedert, M. (2013). Tau pathology and neurodegeneration. Lancet Neurol *12*, 609-622.

Squire, L.R., Wixted, J.T., and Clark, R.E. (2007). Recognition memory and the medial temporal lobe: a new perspective. Nat Rev Neurosci *8*, 872-883.

Stankiewicz, P., and Lupski, J.R. (2002). Genome architecture, rearrangements and genomic disorders. Trends Genet *18*, 74-82.

Stankiewicz, P., and Lupski, J.R. (2010). Structural variation in the human genome and its role in disease. Annu Rev Med *61*, 437-455.

Stefansson, H., Rujescu, D., Cichon, S., Pietiläinen, O.P., Ingason, A., Steinberg, S., Fossdal, R., Sigurdsson, E., Sigmundsson, T., Buizer-Voskamp, J.E*., et al.* (2008). Large recurrent microdeletions associated with schizophrenia. Nature *455*, 232-236.

Steinberg, K.M., Antonacci, F., Sudmant, P.H., Kidd, J.M., Campbell, C.D., Vives, L., Malig, M., Scheinfeldt, L., Beggs, W., Ibrahim, M*., et al.* (2012). Structural diversity and African origin of the 17q21.31 inversion polymorphism. Nat Genet *44*, 872-880.

Steinberg, S., de Jong, S., Mattheisen, M., Costas, J., Demontis, D., Jamain, S., Pietilainen, O.P.H., Lin, K., Papiol, S., Huttenlocher, J*., et al.* (2014). Common variant at 16p11.2 conferring risk of psychosis. Molecular Psychiatry *19*, 108-114.

Stratton, R.F., Dobyns, W.B., Airhart, S.D., and Ledbetter, D.H. (1984). New chromosomal syndrome: Miller-Dieker syndrome and monosomy 17p13. Hum Genet *67*, 193-200.

Sullivan, P.F., Kendler, K.S., and Neale, M.C. (2003). Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch Gen Psychiatry *60*, 1187-1192.

Szatmari, P., Paterson, A.D., Zwaigenbaum, L., Roberts, W., Brian, J., Liu, X.Q., Vincent, J.B., Skaug, J.L., Thompson, A.P., Senman, L*., et al.* (2007). Mapping autism risk loci using genetic linkage and chromosomal rearrangements. Nat Genet *39*, 319-328.

Takemoto, T., Uchikawa, M., Yoshida, M., Bell, D.M., Lovell-Badge, R., Papaioannou, V.E., and Kondoh, H. (2011). Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. Nature *470*, 394-398.

Tan, T.Y., Aftimos, S., Worgan, L., Susman, R., Wilson, M., Ghedia, S., Kirk, E.P., Love, D., Ronan, A., Darmanian, A*., et al.* (2009). Phenotypic expansion and further characterisation of the 17q21.31 microdeletion syndrome. J Med Genet *46*, 480-489.

Thomas, P.D., Kejariwal, A., Campbell, M.J., Mi, H., Diemer, K., Guo, N., Ladunga, I., Ulitsky-Lazareva, B., Muruganujan, A., Rabkin, S*., et al.* (2003). PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. Nucleic Acids Res *31*, 334-341.

Toral-Lopez, J., Gonzalez-Huerta, L.M., and Cuevas-Covarrubias, S.A. (2012). Complete monosomy mosaic of chromosome 21: Case report and review of literature. Gene *510*, 175-179.

Treutlein, J., Kissling, C., Frank, J., Wiemann, S., Dong, L., Depner, M., Saam, C., Lascorz, J., Soyka, M., Preuss, U.W*., et al.* (2006). Genetic association of the human corticotropin releasing hormone receptor 1 (CRHR1) with binge drinking and alcohol intake patterns in two independent samples. Mol Psychiatry *11*, 594-602.

Treutlein, J., Mühleisen, T.W., Frank, J., Mattheisen, M., Herms, S., Ludwig, K.U., Treutlein, T., Schmael, C., Strohmaier, J., Bösshenz, K.V*., et al.* (2009). Dissection of phenotype reveals possible association between schizophrenia and Glutamate Receptor Delta 1 (GRID1) gene promoter. Schizophr Res *111*, 123-130.

Turner, D.J., Miretti, M., Rajan, D., Fiegler, H., Carter, N.P., Blayney, M.L., Beck, S., and Hurles, M.E. (2008). Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. Nat Genet *40*, 90-95.

Turner, P.R., O'Connor, K., Tate, W.P., and Abraham, W.C. (2003). Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory. Progress in Neurobiology *70*, 1-32.

van Bokhoven, H. (2011). Genetic and epigenetic networks in intellectual disabilities. Annu Rev Genet *45*, 81-104.

van Os, J., and Kapur, S. (2009). Schizophrenia. Lancet *374*, 635-645.

van Vliet, E.A., Aronica, E., Redeker, S., and Gorter, J.A. (2004). Expression and cellular distribution of major vault protein: a putative marker for pharmacoresistance in a rat model for temporal lobe epilepsy. Epilepsia *45*, 1506-1516.

Visscher, P.M. (2008). Sizing up human height variation. Nat Genet *40*, 489-490.

Voss, M., Schröder, B., and Fluhrer, R. (2013). Mechanism, specificity, and physiology of signal peptide peptidase (SPP) and SPP-like proteases. Biochim Biophys Acta *1828*, 2828-2839.

Walsh, T., McClellan, J.M., McCarthy, S.E., Addington, A.M., Pierce, S.B., Cooper, G.M., Nord, A.S., Kusenda, M., Malhotra, D., Bhandari, A.*, et al.* (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science *320*, 539-543.

Walters, R.G., Jacquemont, S., Valsesia, A., de Smith, A.J., Martinet, D., Andersson, J., Falchi, M., Chen, F., Andrieux, J., Lobbens, S.*, et al.* (2010). A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. Nature *463*, 671-U104.

Wang, Y., Gao, L., Conrad, C.G., and Andreadis, A. (2011). Saitohin, which is nested within the tau gene, interacts with tau and Abl and its human-specific allele influences Abl phosphorylation. J Cell Biochem *112*, 3482-3488.

Wang, Z.L., Wang, B.P., Yang, L., Guo, Q.X., Aithmitti, N., Zhou, S.Y., and Zheng, H. (2009). Presynaptic and Postsynaptic Interaction of the Amyloid Precursor Protein Promotes Peripheral and Central Synaptogenesis. Journal of Neuroscience *29*, 10788-10801.

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P.*, et al.* (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520-562.

Weiss, L.A., Shen, Y.P., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M.A.R., Green, T.*, et al.* (2008). Association between microdeletion and microduplication at 16p11.2 and autism. New England Journal of Medicine *358*, 667-675.

Wheeler, E., Huang, N., Bochukova, E.G., Keogh, J.M., Lindsay, S., Garg, S., Henning, E., Blackburn, H., Loos, R.J.F., Wareham, N.J.*, et al.* (2013). Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. Nature Genetics *45*, 513-U576.

Willer, C.J., Speliotes, E.K., Loos, R.J., Li, S., Lindgren, C.M., Heid, I.M., Berndt, S.I., Elliott, A.L., Jackson, A.U., Lamina, C.*, et al.* (2009). Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet *41*, 25-34.

WILLIAMS, J.C., BARRATT-BOYES, B.G., and LOWE, J.B. (1961). Supravalvular aortic stenosis. Circulation *24*, 1311-1318.

Williams, N.M., O'Donovan, M.C., and Owen, M.J. (2006). Chromosome 22 deletion syndrome and schizophrenia. Int Rev Neurobiol *73*, 1-27.

Wu, Y.Q., Heilstedt, H.A., Bedell, J.A., May, K.M., Starkey, D.E., McPherson, J.D., Shapira, S.K., and Shaffer, L.G. (1999). Molecular refinement of the 1p36 deletion syndrome reveals size diversity and a preponderance of maternally derived deletions. Hum Mol Genet *8*, 313-321.

Yasuhiko, Y., Kitajima, S., Takahashi, Y., Oginuma, M., Kagiwada, H., Kanno, J., and Saga, Y. (2008). Functional importance of evolutionally conserved Tbx6 binding sites in the presomitic mesoderm-specific enhancer of Mesp2. Development *135*, 3511-3519.

Yu, T., Clapcote, S.J., Li, Z.Y., Liu, C.H., Pao, A., Bechard, A.R., Carattini-Rivera, S., Matsui, S.I., Roder, J.C., Baldini, A.*, et al.* (2010). Deficiencies in the region syntenic to human 21q22.3 cause cognitive deficits in mice. Mammalian Genome *21*, 258-267.

Yunis, J.J., and Sanchez, O. (1974). A new syndrome resulting from partial trisomy for the distal third of the long arm of chromosome 10. J Pediatr *84*, 567-570.

Zheng, H., Jiang, M.H., Trumbauer, M.E., Sirinathsinghji, D.J.S., Hopkins, R., Smith, D.W., Heavens, R.P., Dawson, G.R., Boyce, S., Conner, M.W.*, et al.* (1995). BETA-AMYLOID PRECURSOR PROTEIN-DEFICIENT MICE SHOW REACTIVE GLIOSIS AND DECREASED LOCOMOTOR-ACTIVITY. Cell *81*, 525-531.

Zhu, X., Need, A.C., Petrovski, S., and Goldstein, D.B. (2014). One gene, many neuropsychiatric disorders: lessons from Mendelian diseases. Nat Neurosci *17*, 773-781.

Zody, M.C., Jiang, Z., Fung, H.C., Antonacci, F., Hillier, L.W., Cardone, M.F., Graves, T.A., Kidd, J.M., Cheng, Z., Abouelleil, A.*, et al.* (2008). Evolutionary toggling of the MAPT 17q21.31 inversion region. Nat Genet *40*, 1076-1083.

Zollino, M., Orteschi, D., Murdolo, M., Lattante, S., Battaglia, D., Stefanini, C., Mercuri, E., Chiurazzi, P., Neri, G., and Marangi, G. (2012). Mutations in KANSL1 cause the 17q21.31 microdeletion syndrome phenotype. Nat Genet *44*, 636-638.

Zufferey, F., Sherr, E.H., Beckmann, N.D., Hanson, E., Maillard, A.M., Hippolyte, L., Mace, A., Ferrari, C., Kutalik, Z., Andrieux, J.*, et al.* (2012). A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. Journal of Medical Genetics *49*, 660-668.

Thomas ARBOGAST

Apport des modèles murins à la compréhension des maladies associées aux variations du nombre de copies : monosomie 21 partielle et délétions et duplications des régions 16p11.2 et 17q21.31

Les variations du nombre de copies (CNVs) incluent les délétions et les duplications de régions chromosomiques d'une taille variant de 50 pb à plusieurs Mb. Depuis 2005, les études d'association pangénomiques (GWAS) ont permis d'associer certains larges CNVs à des maladies syndromiques associées à la déficience intellectuelle incluant les syndromes de DiGeorge, Williams, Angelman, etc. En fonction de la densité génique de la région d'intérêt et de la variabilité des phénotypes associés, l'étude de la physiopathologie des syndromes peut être extrêmement complexe. La modélisation murine offre de nombreux avantages pour l'identification des gènes candidats et la compréhension des mécanismes moléculaires associés à ces pathologies.

Les travaux présentés dans ce manuscrit consistent en la caractérisation des modèles murins pour cinq maladies syndromiques associées aux CNVs : la monosomie 21 partielle ainsi que les réarrangements des régions 16p11.1 et 17q21.31. Les caractérisations anatomiques, métaboliques et comportementales des animaux nous ont permis d'évaluer un grand nombre de paramètres associés à la symptomatique humaine. Nous avons également réalisé des analyses électrophysiologiques et transcriptomiques en ciblant nos investigations sur l'hippocampe, structure cérébrale qui joue un rôle central dans les processus de mémoire et d'apprentissage. Ce projet de recherche s'inscrit dans une perspective plus large qui est l'identification des gènes candidats pour les phénotypes observés et le développement de premières stratégies thérapeutiques pouvant potentiellement aboutir à l'amélioration des capacités cognitives des patients.

Mots clés : variation du nombre de copies, déficience intellectuelle, modèles murins, apprentissage et mémoire

Copy number variations (CNVs) include deletions and duplications of chromosomal regions ranging in size from 50bp to several Mb. Since 2005, genome-wide association studies (GWAS) have associated some large CNVs to syndromic diseases linked to intellectual disability including DiGeorge, Williams, Angelman syndroms, etc. Depending on the gene density of the region of interest and the variability of symptoms, the study of the pathophysiology of syndromes can be extremely complex. Mouse modeling show many advantages for the identification of candidate genes and the understanding of molecular mechanisms associated with these diseases.

The work presented in this manuscript consists of the characterization of mouse models of five syndromic diseases associated with CNVs: partial monosomy 21 and rearrangements of 16p11.2 and 17q21.31 regions. Anatomical, metabolic and behavioral characterizations of animals allowed us to evaluate a broad number of parameters associated with human phenotypes. We also performed electrophysiological and transcriptomic analysis focusing our investigation on the hippocampus which has a major role in learning and memory processes. This project is part of a wider perspective which is the identification of candidate genes for the different phenotypes we observe in the mouse and the development of first treatment strategies which can potentially lead to the improvement of cognitive capacity of patients.

Key words: Copy number variation, intellectual disability, mouse models, learning and memory