AJHG The American Journal of Human Genetics

A syndromic neurodevelopmental disorder caused by mutations in SMARCD1, a core SWI/SNF subunit required for context-dependent regulation of neuronal genes in Drosophila

--Manuscript Draft--

Manuscript Number:	AJHG-D-18-00665R3
Full Title:	A syndromic neurodevelopmental disorder caused by mutations in SMARCD1, a core SWI/SNF subunit required for context-dependent regulation of neuronal genes in Drosophila
Article Type:	Article
Keywords:	Intellectual Disability Neurodevelopmental Disorder SWI/SNF Memory Drosophila Gene Regulation
Corresponding Author:	Jamie M. Kramer Western University London, CANADA
First Author:	Kevin C.J. Nixon
Order of Authors:	Kevin C.J. Nixon
	Justine Rousseau
	Max H. Stone
	Mohammed Sarikahya
	Sophie Ehresmann
	Seiji Mizuno
	Naomichi Matsumoto
	Noriko Miyake
	Diana Baralle
	Shane McKee
	Kosuke Izumi
	Alyssa Ritter
	Solveig Heide
	Delphine Héron
	Christel Depienne
	Hannah Titheradge
	Philippe M. Campeau
	Jamie M. Kramer
Abstract:	Mutations in several genes encoding components of the SWI/SNF chromatin remodeling complex cause neurodevelopmental disorders (NDDs). Here, we report on 5 individuals with mutations in SMARCD1, presenting with developmental delay, intellectual disability, hypotonia, feeding difficulties, and small hands and feet. The mutations were proven to be de novo in 4 of the 5 individuals, by trio exome sequencing. Mutations in other SWI/SNF components cause Coffin-Siris syndrome and Nicolaides-Baraitser syndrome, or other syndromic and non-syndromic NDDs. Although the individuals presented here have dysmorphisms and some clinical overlap

	with these syndromes, they lack the typical facial dysmorphisms. To gain insight into the function of SMARCD1 in neurons, we investigated the Drosophila ortholog, Bap60, in postmitotic memory-forming neurons of the adult Drosophila mushroom body (MB). Targeted knockdown of Bap60 in the MB of adult flies causes defects in long-term memory. Mushroom body specific transcriptome analysis revealed that Bap60 is required for context-dependent expression of genes involved in neuron function and development in juvenile flies when synaptic connections are actively being formed in response to experience. Taken together, we identify a NDD caused by SMARCD1 mutations and establish a role for the SMARCD1 ortholog Bap60 in regulation of neurodevelopmental genes during a critical time window of juvenile adult brain development that is essential in establishing neuronal circuits that are required for learning and memory.

A syndromic neurodevelopmental disorder caused by mutations in *SMARCD1*, a core SWI/SNF subunit required for context-dependent regulation of neuronal genes in *Drosophila*

Kevin C.J. Nixon[1,†], Justine Rousseau[2,†], Max H. Stone[3,4], Mohammed Sarikahya[3], Sophie Ehresmann[2], Seiji Mizuno[5], Naomichi Matsumoto[6], Noriko Miyake[6], DDD study, Diana Baralle[7], Shane McKee[8], Kosuke Izumi[9], Alyssa Ritter[9], Solveig Heide[10], Delphine Héron[10], Christel Depienne[11,12,13], Hannah Titheradge[14], Jamie M. Kramer[1,3,4*], and Philippe M. Campeau[2,15*]

Primary affiliations only:

[1]Department of Physiology and Pharmacology, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada. N6A 5C1

[2]CHU-Sainte Justine Research Center, University of Montreal, Montreal, QC, Canada. H3T 1C5

[3]Department of Biology, Faculty of Science, Western University, London, Ontario, Canada. N6A 5B7

[4]Division of Genetics and Development, Children's Health Research Institute, London, Ontario, Canada. N6C 2V5

[5]Department of Pediatrics, Central Hospital, Aichi Human Service Center, Kasugai 480-0392, Japan.

[6]Yokohama City University Graduate School of Medicine, Yokohama 236-0004, Japan.

[7]Faculty of Medicine, University of Southampton, Southampton SO17 1BJ, UK.

[8]Belfast City Hospital, Belfast BT9 7AB, Northern Ireland, UK.

[9]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.

[10]AP-HP, Hôpital Pitié-Salpêtrière, Département de Génétique et de Cytogénétique ; Centre de Référence Déficience Intellectuelle de Causes Rares ; GRC UPMC « Déficience Intellectuelle et Autisme », F-75013, Paris, France.

[11]INSERM, U 1127, CNRS UMR 7225, Sorbonne Universités, UPMC Univ Paris 06 UMR S 1127, Institut du Cerveau et de la Moelle épinière, ICM, F-75013, Paris, France.

[12]IGBMC, CNRS UMR 7104/INSERM U964/Université de Strasbourg, 67400 Illkirch, France.

[13]Institute of Human Genetics, University Hospital Essen, University of Duisburg-Essen, 45 239 Essen, Germany.

[14]Birmingham Women's and Children's NHS Foundation Trust Mindelsohn Way, Birmingham B15 2TG, UK.

[15]Department of Pediatrics, University of Montreal, Montreal, QC, Canada. H4A 3J1

†These authors contributed equally to this work.
*These authors contributed equally to this work; corresponding authors.

Address:
Philippe Campeau
Division of Medical Genetics, Room 6727, Sainte-Justine Hospital
3175 Cote-Sainte-Catherine
Montreal QC Canada H3T 1C5
Tel: 438-396-7899 Fax: 514 345-4766
p.campeau@umontreal.ca

Jamie Kramer
Physiology and Pharmacology
Medical Science Building, Room 266
Western University
1151 Richmond St.
London, Ontario, Canada, N6A 5C1
Tel: 519-661-3740
James.Kramer@schulich.uwo.ca

Abstract

Mutations in several genes encoding components of the SWI/SNF chromatin remodeling complex cause neurodevelopmental disorders (NDDs). Here, we report on 5 individuals with mutations in *SMARCD1*, presenting with developmental delay, intellectual disability, hypotonia, feeding difficulties, and small hands and feet. The mutations were proven to be *de novo* in 4 of the 5 individuals, by trio exome sequencing. Mutations in other SWI/SNF components cause Coffin-Siris syndrome and Nicolaides-Baraitser syndrome, or other syndromic and non-syndromic NDDs. Although the individuals presented here have dysmorphisms and some clinical overlap with these syndromes, they lack the typical facial dysmorphisms. To gain insight into the function of SMARCD1 in neurons, we investigated the *Drosophila* ortholog, Bap60, in postmitotic memory-forming neurons of the adult *Drosophila* mushroom body (MB). Targeted knockdown of Bap60 in the MB of adult flies causes defects in long-term memory. Mushroom body specific transcriptome analysis revealed that Bap60 is required for context-dependent expression of genes involved in neuron function and development in juvenile flies when synaptic connections are actively being formed in response to experience. Taken together, we identify a NDD caused by *SMARCD1* mutations and establish a role for the SMARCD1 ortholog Bap60 in regulation of neurodevelopmental genes during a critical time window of juvenile adult brain development that is essential in establishing neuronal circuits that are required for learning and memory.

Introduction

Regulation of gene expression in neurons is critical for normal brain development and for normal cognitive functioning in adults[1–4]. Chromatin structure is an important factor in modulating gene expression[5]. The SWI/SNF chromatin remodeling complex (also known as the BAF complex in mammals) is a highly conserved protein complex that utilizes energy from ATP to alter nucleosome-DNA interactions, resulting in more open chromatin for transcription factor binding[3,6–8]. In mammals, there are multiple cell-type specific conformations of the SWI/SNF complex, including npBAF, specific for neuronal progenitors, and nBAF, specific for postmitotic neurons[1,9–12]. Each form of the SWI/SNF complex contains 10-15 proteins encoded by 29 genes[11]. The SWI/SNF complex is important for regulation of gene expression programs involved in neuronal differentiation and brain-region specification in mice[2,4,9,13–17]. However, the complex is also essential in mature neurons for memory formation, synaptic plasticity, and activity responsive neurite outgrowth[3,4,18].

Disruption of genes encoding chromatin regulators is an important cause of neurodevelopmental disorders (NDDs), which are a heterogeneous group of disorders including intellectual disability and autism[19,20]. Mutations in genes encoding several different SWI/SNF subunits cause syndromic NDDs, including Nicolaides-Baraitser Syndrome (MIM: 601358) and Coffin-Siris Syndrome (MIM: 135900)[21–24]. SWI/SNF mutations are also involved in other syndromic and non-syndromic NDDs[25,26] and psychiatric disorders such as schizophrenia[27–30]. In total, mutations in 11 of the 28 genes encoding SWI/SNF components have been implicated in NDDs[21-23,25–28,30–34], emphasizing the essential role of this complex in neuron development and function. It

remains to be determined if mutations in the remaining subunits are also involved in NDDs or other brain disorders.

Here, we characterize mutations in *SMARCD1* (MIM:601735) in individuals presenting with a syndromic NDD. *SMARCD1* encodes a core component of the SWI/SNF complex that has not previously been associated with NDDs. We show that the *Drosophila* SMARCD1 ortholog Bap60 is required in the mushroom body (MB) of adult flies for normal long-term memory. The MB is the learning and memory center of the fly brain[35,36]. We find that Bap60 has a profound effect on the expression of neurodevelopmental genes in the MB during a critical time window of juvenile brain development when synaptic connections are formed in response to early life experiences.

Subjects and Methods

Participant enrolment

Individual 1 was enrolled in a study protocol that was approved by the institutional review boards of Yokohama City University School of Medicine. Individual 2 was enrolled in a study of the Groupe Hospitalier Pitié-Salpêtrière approved by the INSERM institutional review board. Individuals 3 and 4 were enrolled in the Deciphering Developmental Disorders (DDD) study[37]. Contact with the clinicians was made through the DDD website, and the individuals were enrolled in a study approved by the institutional review board of the CHU Sainte-Justine. Individual 5 had exome sequencing on a clinical basis and the family consented to the sharing of clinical

information without photos. Genematcher was used to connect with the clinicians of individuals 2 and 5.

Exome sequencing

For individual 1, genomic DNA was enriched for exons using the SureSelect All Human Exon kit (Agilent). Libraries were sequenced on the Illumina HiSeq, and analysis was performed as described[38]. For individual 2, exome sequencing was performed as described in[39] (as for family 1 in that citation). For the individuals 3 and 4, exome sequencing was done as part of the DDD project. Exomes were enriched using the Agilent SureSelect 55MB Exome Plus library followed by Illumina HiSeq sequencing, and analysis was performed as described[37]. Exome variants passing the filtering criteria were evaluated by the DDD study's internal clinical review team including two consultant clinical geneticists. For individual 5, exome sequencing was done at GeneDx. Using genomic DNA from the proband and parents, the exonic regions and flanking splice junctions of the genome were captured using the Clinical Research Exome kit (Agilent). Sequencing was done on an Illumina system with 100bp or greater paired-end reads. Reads were aligned to human genome build GRCh37/UCSC hg19 and analyzed for sequence variants using a custom-developed analysis tool. Additional sequencing technology and variant interpretation protocol (e.g. Sanger) has been previously described[40]. The general assertion criteria for variant classification are publicly available on the GeneDx ClinVar submission page.

***In silico* assessment of the variants**

Secondary structure of the SMARCD1 protein (NCBI Reference Sequence: NP_003067.3) was done using protein paint. The Prediction of the coiled coil domain was perform using the NPS@: Network Protein Sequence Analysis tool. A 3D model of the full SMARCD1 protein was predicted using the I-TASSER online suite using standard parameters. The model with the highest confidence was selected and modeled using Pymol.

Fly stocks and culture

Flies were reared on standard cornmeal-agar media with a 12h/12h light cycle in 70% humidity. The following stocks were obtained from the Bloomington *Drosophila* Stock Center (Indiana University): short hairpin Bap60 RNAi lines generated by the Transgenic RNAi Project (TRiP)[41] (*UAS-Bap60^{32503}* and *UAS-Bap60^{33954}*); the TRiP short hairpin *mCherry^{RNAi}* line (stock #35785); the mushroom body specific Gal4 driver line *R14H06-Gal4* (stock #48667) from the Janelia Flylight[42] collection, and the temperature sensitive Gal80 (*Gal80^{ts}*) driven by the tubulin promoter (stock #7019). *UAS-unc84::GFP* was a gift from G.L. Henry. Bap60 RNAi lines (*UAS-Bap60^{32503}* and *UAS-Bap60^{33954})* were assessed for knockdown efficiency in a parallel study[43]. When ubiquitously expressed with an *Act-Gal4* driver, the *UAS-Bap60^{32503}* RNAi induces 100% lethality (no adults eclose) as expected, since null mutations in Bap60 are lethal at the embryo and larval stage[44]. RT-qPCR on larvae with ubiquitous knockdown shows a significant reduction of mRNA by more than 50%[43]. *UAS-Bap60^{33954}* did not induce complete lethality under *Act-Gal4* expression as 17% of progeny did survive to

adulthood. The more potent RNAi line *UAS-Bap60^{32503}* was chosen for follow up RNA-sequencing studies.

Courtship Conditioning

Courtship conditioning assays were performed as described previously[45]. Training was performed by pairing individual males with premated wild type females. Long-term memory was induced through a seven-hour training period and tested 24 hours after training. For each fly pair, a courtship index (CI) was determined by manually observing the percentage of time spent courting during a 10-minute period. Statistically, loss of memory was identified using two complimentary methods. Reduction of the mean CI of trained ($CI_{trained}$) flies compared to naïve (CI_{naive}) of the same genotype was compared using a Kruskal-Wallis test followed by pairwise comparisons using Dunn's test. A learning index (LI) was also calculated ($LI=(CI_{naive}-CI_{trained})/CI_{naive}$). LIs were compared between genotypes using a randomization test (10,000 bootstrap replicates) performed with a custom R script[45] and the resulting p-values were corrected for multiple testing using the method of Bonferroni.

The temperature-sensitive Gal80^{ts} system[46] was used to specifically knockdown Bap60 in the adult mushroom body. Flies of the genotype *tubGal80^{ts}; R14H06-Gal4* were crossed to Bap60 RNAi lines, *UAS-Bap60^{32503}* and *UAS-Bap60^{33954}*, as well as a control *UAS-mCherry^{RNAi}* that is present in the same genetic background as the Bap60 RNAi transgenes. Crosses were raised at 18ºC, the permissive temperature which allows Gal80^{ts} to repress Gal4 mediated transcription. At eclosion, male flies of the genotypes *Gal80^{ts}/+; R14H06-Gal4/UAS-Bap60^{32503}, Gal80^{ts}/+; R14H06-Gal4/UAS-*

Bap60^{33954} and Gal80^{ts}/+; R14H06-Gal4/UAS-mCherry^{RNAi} were transferred to 29ºC, which causes Gal80 inactivation, allowing Gal4-mediated induction of UAS-RNAi transgenes. After five days, collected males were tested for long-term memory using courtship conditioning as described above.

Isolation of nuclei in tagged cell types (INTACT)

INTACT was performed as described previously[47]. Fifty juvenile (0-3h old) or mature (1-5 days old) adult male flies of the genotype UAS-unc84::GFP/+; R14H06-Gal4/UAS-Bap60^{32503} (Bap60-KD) and UAS-unc84::GFP/+;R14H06-Gal4/UAS-mCherry^{RNAi} (control) were anesthetized using CO_2 and flash frozen using liquid nitrogen. Protein G Dynabeads (Invitrogen) were adsorbed to 5μg of anti-GFP antibody (Invitrogen, G10362) in PBS 0.1% Tween-20 (PBST) for 10 minutes at room temperature. Beads were then isolated using a magnet and resuspended in PBST. The flies were then vortexed and placed in ice-cold sieves to separate and isolate their heads. Heads were then added to homogenization buffer (25mM KCl pH 7.8, 5mM $MgCl_2$, 20mM Tricine, 150nM spermine, 500nM spermidine, 10nM β-glycerophosphate, 250mM sucrose, 1X Pierce protease inhibitor tablets – EDTA-free (ThermoFisher Scientific) or 1X Halt protease inhibitor cocktail – EDTA-free (ThermoFisher Scientific)), and the suspension was homogenized for approximately 1 minute using a standard tissue homogenizer at 30,000 rpm. The suspension was then placed in a Dounce homogenizer with NP40 (ThermoFisher Scientific) added to an end concentration of 0.3% and homogenized 6 times with the tight pestle. The homogenate was filtered using a 40μm strainer and pre-cleared with non-antibody bound beads for 10 minutes.

Antibody-bound beads were added to the homogenate for 30 minutes, followed by washing in homogenization buffer. Total RNA was isolated from immunoprecipitated nuclei using the Arcturus PicoPure RNA isolation kit (ThermoFisher Scientific) with DNase digestion using the RNase-free DNase kit (Quiagen) according to the manufacturer's instructions. Isolated RNA quality was then assessed using the Bioanalyzer 2100 Pico RNA kit (Agilent), by visual examination of rRNA peak integrity.

RNA-sequencing and analysis

RNA from MB-enriched samples was used for RNA-seq library generation using the NuGEN Ovation *Drosophila* RNA-Seq System (BioLynx), according to the manufacturer's instructions. Size selection using Agencourt SPRI select beads (Beckman-Coulter) was used to select for library sizes of 200bp. Library size was assessed using the Bioanalyzer 2100 DNA high sensitivity kit (Agilent). Sequencing was performed using the Illumina NextSeq500 at the London Regional Genomics Centre (Robarts Research Institute, London, Ontario) using the high output v2 75 cycle kit to a read length of 75bp for single-end reads.

Raw sequencing reads were trimmed using Prinseq quality trimming[48] and a minimum base quality score of 30. Read quality was then assessed using FastQC. Trimmed reads were aligned to the *Drosophila melanogaster* reference genome (BDGP release 6)[49,50] using the STAR aligner[51]. An average of 54,803,904 and 46,899,292 high quality uniquely aligned reads with a maximum of four mismatches were obtained from juvenile Bap60-KD MBs (N=3) and controls (N=2) respectively and an average of 26,865,090 and 35,504,102 high quality uniquely aligned reads with a maximum of four

mismatches were obtained from mature Bap60-KD MBs (N=5) and controls (N=5) respectively (**Table S1**). The number of reads per gene was quantified using HTSeq-count[52] with the --type flag indicating 'gene' (**Table S1**). Y chromosome genes, rRNA genes, and genes with no counts across all samples were excluded, leaving 12,922 and 13,440 genes for downstream analysis for juvenile and mature samples, respectively.

Raw gene counts were normalized and differential expression analysis between Bap60-KD MBs and controls was performed using the R package DESeq2[53]. Differentially expressed genes were defined as genes with a fold-change >1.5 or > 2 and a Benjamini-Hochberg adjusted p-value <0.05. Gene ontology (GO) enrichment analysis was performed on upregulated and downregulated differentially expressed genes using Panther[54] (FDR<0.05).

Classification of tissue specific genes

To generate lists of tissue-specific genes, we used normalized gene expression values from Brown *et al.*[55] for several tissues including, adult head (9 samples), adult carcass (3 samples), and adult digestive tract (3 samples)**.** The relative enrichment in gene expression levels for each of these tissues was calculated by comparing the mean bases per kilobase per million reads (bpkm) for each specific tissue to the mean bpkm across all remaining tissues. Tissue-specific and tissue-depleted genes are defined as having an average $\log_2$ fold-change in expression outside of one standard deviation of the fold-change in expression of all genes in all tissue types (**Table S2**). Adult heads were used to represent a neuron-enriched tissue, while the adult carcass, consisting of tissues remaining after removal of the head, digestive tract, and reproductive organs,

represents a muscle-enriched tissue. A list of MB-specific genes (**Table S2**) was generated using a published MB-specific transcriptome that was obtained using the same INTACT protocol described here[47]. MB-specific genes were defined as genes that were significantly upregulated more than 2-fold in MB-enriched samples compared to biologically paired whole head input samples. Statistical significance of over or under representation of tissue-specific genes in differentially upregulated and downregulated genes in Bap60-KD MBs was determined using a hypergeometric test.

Accession Numbers

RNA-seq data are available at the NCBI Gene Expression Omnibus through the accession number GSE122864.

Results

Human Genetic and Clinical Data

In a cohort of individuals collected because of clinical overlap with Coffin-Siris syndrome, *de novo* mutations in several genes encoding members of the SWI-SNF complex were identified[21]. As an expansion of this study, an individual was identified with a *SMARCD1* variant (c.990C>G, p.Asp330Glu – individual 1), but the *de novo* status of the variant could not be confirmed since the father was not available for testing. Subsequently, as part of the Deciphering Developmental Disorders (DDD) study[37], two individuals were identified with *de novo SMARCD1* variants (c.1457G>A, p.Trp486* and c.1483T>C, p.Phe495Leu, individuals 3 and 4). Through GeneMatcher[56], two additional individuals were identified. One with a *de novo* c.1336A>G, p.Arg446Gly (individual 2) variant in a cohort of individuals with agenesis of the corpus callosum, and

another individual who had clinical exome sequencing for developmental delay and other anomalies who was identified to have a truncating variant (c.1507C>T, p.Arg503*, individual 5). For individuals 2-5 de novo occurrence of the variants was identified by analysis of exome trios (mother, father, proband).

The variants and the associated clinical features are shown in **Figure 1, Figure 2**, **Table 1, and Table S3**. There is phenotypic overlap with Coffin-Siris syndrome in that two individuals have a hypoplastic 5th toenail and all have intellectual disability or developmental delay, however, they do not have the typical facial features (wide mouth with thick everted lips, broad nose, thick eyebrows, and long eyelashes) (see **Figure 2** for pictures). The sparse hair, thin upper vermilion, thick lower vermilion in individuals 1 and 3 suggests overlapping facial features with Nicolaides-Baraitser syndrome, although the overall facial gestalt is quite different. All had feeding difficulties, and three had hypotonia. All individuals had developmental delay and none had epilepsy. Three individuals had small hands and feet. The dysmorphisms were variable and were most notable in individual 4 with the p.Phe495Leu variant, who also had the most profound neurodevelopmental disorder (at 3 years of age, he could not talk or walk).

Assessment of the identified *SMARCD1* variants

SMARCD1 is 515 amino acids long with a characteristic SWIB domain and three predicted coiled coil domains (**Figure 1A**). The identified *SMARCD1* variants, 2 nonsense and 3 missense, are all clustered either within, or in close proximity, to these domains. The two nonsense variants (p.Trp486*, p.Arg503*) are located within 50 base pairs of the last exon junction so should escape nonsense mediated decay. These

variants are predicted to produce a truncated protein lacking the last coiled coil domain, C3 (**Figure 1**). One of the missense variants (p.Asp330Glu) is located in the SWIB domain while the two other missense mutations (p.Arg446Gly, p.Phe495Leu) are located near or in the C2 and C3 coiled coil domains, respectively (**Figure 1A**). All of the missense mutations cluster close together in a 3D model of SMARCD1 (**Figure 1C**). The SMARCD2 (MIM:601736) SWIB domain and the coiled coil regions have been shown to be essential to mediate SMARCD2 specific function in granulocytic development and indeed, deletion of the coiled coil domain located at the extreme C-terminus of the protein was shown to be sufficient to disrupt the binding between SMARCD2 and the BAF complex[57].Thus all of the identified mutations have the potential to disrupt a putatively functional domain of SMARCD1.

Prior to the identification of the p.Arg446Gly and p.Arg503* mutation through GeneMatcher, we had generated expression constructs for the variants p.Asp330Glu, p.Trp486* and p.Phe495Leu. Detection of a similar level of ectopic expression of the wildtype and the mutated proteins, as visualized by western blot, suggests that the mutations do not dramatically affect protein stability (**Figure S1**). To address the question as to whether those three mutations could impair the incorporation of SMARCD1 into the SWI/SNF complex, we performed a co-immunoprecipitation (Co-IP) assay. None of the three mutations abolished the interaction between SMARCD1 and the ATPase subunit SMARCA4 (MIM:603254) nor the scaffolding subunit SMARCC1 (MIM: 601732; **Figure S1**). Although these mutations do not appear to disrupt the interactions of SMARCD1 with SMARCC1 and SMARCA4, they might affect the

function of SMARCD1 otherwise, possibly by altering interactions with other SWI/SNF proteins or transcription factors that recruit the SWI/SNF complex to specific genes.

We used several bioinformatic resources to assess the impact of the identified *SMARCD1* variants *in silico*. *SMARCD1* has an ExAC pLI score of 1 (since only one loss-of-function (LoF) variant was observed in the ExAC cohort while 26 were predicted), indicating that the gene is intolerant to heterozygous LoF variants[58]. It also had a positive ExAC Z-score of 3.95, indicating it is also relatively intolerant to missense variants (74 observed while 183 were predicted). DECIPHER gives a %HI score of 16.64%, which is close to the ranks indicating there is a high likelihood that the gene will exhibit haploinsufficiency (below 10%)[59]. DOMINO gives a 99.8% probability for the gene to be associated with an autosomal dominant condition[60].

We used wANNOVAR[61] to predict pathogenicity scores, using tools that consider conservation (polyphen[62], LRT, MutationAssessor[63], GERP[64], PhyloP[65], phastCons[66], SiPhy[67]), function (SIFT[68], PROVEAN[69]), or a combination thereof (FATHMM[70], MutationTaster[71], DANN[72], CADD[73]). We considered the following prediction cutoffs: damaging, rank scores above 0.5, and CADD phred scores above 25 (**Table S4**). Using these cutoffs, MutationTaster, FATHMM-MKL, and PROVEAN considered all assessed variants deleterious. LRT, GERP, PhyloP, phastCons, SyPhy, SIFT, DANN and CADD predicted most assessed variants to be deleterious. MutationAssessor gave medium scores, and polyphen2 predicted 2 out of 3 assessed variants to be benign. p.Arg446Gly was the variant with the most tools giving low pathogenicity predictions, mostly because of the poor conservation in animals further removed from humans (see *lamprey, C. elegans* and *Drosophila*, **Figure 1B**).

The *Drosophila* SMARCD1 ortholog Bap60 is required for memory

Given that all of the identified *SMARCD1* mutations might lead to a loss-of-function by affecting the SWIB or coiled coil domains, we assessed in *Drosophila* the consequence of loss of Bap60, the fly SMARCD1 ortholog. To do this, we generated Bap60 knockdown flies using previously characterized transgenic RNAi lines (see methods)[43]. RNAi knockdown of Bap60 was targeted to the MB, the learning and memory center of the fly brain, using *R14H06-Gal4*. Knockdown was confined to adult flies using the temperature sensitive Gal4 inhibitor, Gal80^{ts} (see methods). In a parallel study, we have characterized the role of the SWI/SNF complex in MB development[43]. Here we decided to focus on adult flies as a model for understanding the post-natal function of SMARCD1 in the brain, which could be more relevant towards the long-term goal of developing therapies.

We tested memory in MB-specific Bap60 knockdown flies using a classic paradigm known as courtship conditioning[45,74]. In this assay, male flies display a learned reduction of courtship behaviour after sexual rejection by a non-receptive pre-mated female. As expected, the control flies expressing an RNAi construct targeting *mCherry* showed a normal reduction in courtship behaviour after being sexually rejected by a mated female (**Figure 3A**). However, this reduction was not seen in flies expressing RNAi transgenes targeting Bap60 (**Figure 3A**). Bap60 knockdown flies also showed a significantly lower learning index (LI) as compared to the controls (**Figure 3B**). This assay was performed using two different Bap60 RNAi lines, which have unique target sequences. For both knockdown lines similar reduction in memory was

observed, indicating that memory defects are not due to off-target effects. These results suggest that Bap60 is required for normal memory, post development, in adult neurons.

Bap60 is required for expression of neuron-specific genes during a critical period of juvenile adult MB development

We used INTACT (see methods) to investigate the effect of Bap60 knockdown on gene expression in the specific MB cells that were targeted for RNAi knockdown in our learning and memory assays. In juvenile adult insects, the mushroom body undergoes a period of development and synaptogenesis in the first hours after eclosion[75–81]. During this time, neuronal connections are formed that are critical for normal learning and memory later in life[77,79]. We analyzed the MB-specific transcriptome in Bap60 knockdown flies compared to controls in early juvenile adults (0-3 hours after eclosion), and more mature adults (1-5 days after eclosion). We observed a greater effect of Bap60 on gene expression at the juvenile stage than in the mature adult MB (**Figure 4A and 4B**). Differential expression analysis between the juvenile Bap60-KD MBs and controls using DESeq2[53] yielded 416 differentially expressed genes (P_{adj}<0.05 and Fold-Change > 1.5), of which 156 were upregulated and 260 were downregulated (**Figure 4A**, **Table S5**). In contrast, only 68 differentially expressed genes (29 upregulated and 39 downregulated) were observed in mature Bap60-KD MBs (**Figure 3B, Table S5**). Differential expression of several genes was confirmed by RT-qPCR in independent biological replicates (**Figure S2**). We performed gene ontology (GO) enrichment analysis to investigate the functions of the differentially expressed genes. Differentially expressed genes from the mature MB showed very little GO

enrichment (**Table S5**). GO enrichment analysis of the upregulated genes from the juvenile MB revealed many terms related to muscle, such as “myofibril assembly” and “sarcomere organization” (**Table S5**). GO enrichment analysis of downregulated genes in the juvenile MB revealed neuron-related terms such as “neurotransmitter metabolic process”, “synaptic vesicle”, and “regulation of synaptic plasticity”, as well as developmental terms such as “nervous system development” and “anatomical structure development” (**Table S5**). This suggested an important role for Bap60 in the regulation of neurodevelopmental genes in the early juvenile adult MB.

Having observed a downregulation of neuron-related genes and an upregulation of muscle-related genes in juvenile Bap60-KD MBs, we reasoned that Bap60 may be required at this stage to activate expression of neuron-specific genes that contribute to cell identity. To test this, we used existing tissue specific RNA-seq data to establish a list of 904 “neuron-specific genes” that are enriched in heads compared to other tissues[55], and 118 “MB-specific genes” that are enriched in MB-specific INTACT samples compared to whole head[47] (see methods, **Table S2)**. Of the 260 genes that are downregulated in Bap60-KD MBs, 78 are neuron-specific and 31 are MB-specific, which is significantly more than expected by chance ($p<10^{-25}$, hypergeometric test) (**Figure 4C**). On average, these genes are expressed at a consistent level in controls in juvenile and mature adult MBs, but have reduced expression in Bap60-KD MBs at the juvenile stage that recovers to normal levels in mature adults (**Figure 4D**). These trends were validated by RT-qPCR for a selection of genes in an independent experiment (**Figure S2,** *prt* and *jdp*). Muscle-specific genes (enriched in carcass compared to other tissues) were not significantly over-represented among genes that were downregulated in

Bap60-KD MBs. Taken together, these results suggest that Bap60 plays a context dependent role in activating the expression of neuron-specific genes in the MB.

Bap60 is required for expression of developmental genes that are preferentially activated in the juvenile MB.

In addition to neuronal genes, our GO enrichment analysis of genes that are downregulated in Bap60 mutants revealed many terms related to development (**Table S5**). We reasoned that Bap60 might be required for activation of key genes that are involved in experience dependent MB development in juvenile adults[79,82]. To test this, we performed differential expression analysis comparing the MB-specific transcriptome in early juvenile adults to mature adults (**Table S6**). In controls, 549 genes were significantly increased by 2-fold or more in the juvenile adult MB compared to mature adults. Of these 549 juvenile enriched MB genes, 385 genes were also > 2-fold enriched in Bap60-KD flies, while 164 were not (**Figure 5A**). On average, these 164 genes show significantly lower expression in juvenile Bap60-KD MBs compared to controls, and this difference is no longer observed in mature flies (**Figure 5B**). These expression trends were validated for a selection of genes in an independent RT-qPCR experiment (**Figure S2**). Interestingly, the 164 juvenile induced genes that are not induced in Bap60-KD MBs show a strong GO enrichment for terms related to development, and immune response (**Figure 5C**). In contrast, the 385 genes that show juvenile enrichment in both Bap60-KD and control MBs show very little GO enrichment. This suggests that Bap60 is required for activation of developmental genes during a critical period when the juvenile adult MB needs to establish new experience-dependent

synaptic connections that are required for normal learning and memory throughout life[75–81].

Discussion

Characterization of a neurodevelopmental disorder caused by *SMARCD1* mutations

In this study, we describe a genetic disorder characterized by mutations in *SMARCD1*, which encodes a component of the SWI/SNF chromatin remodeling complex. Mutations in several other SWI/SNF genes are implicated in syndromic NDDs that are typically characterized by intellectual disability, abnormalities of the 5th digit, and characteristic facial features[83]. The 5 individuals described here do fit in this spectrum, with intellectual disability and a low penetrance of 5th digit abnormalities, but lack the typical facial dysmorphisms seen in other SWI/SNF related disorders.

The identified *SMARCD1* variants are all clustered in the C-terminal end of the protein, and in a 3D model of the protein, are also located in close proximity (**Figure 1**). Although these missense and protein-truncating mutations are predicted to be damaging, we do not know the precise functional effect of these mutations. One missense variant is located in the highly conserved SWIB domain, while the other 4 variants are located in or near the C-terminal coiled coil domains. We show here that those mutations do not disrupt the interaction with other SWI/SNF components SMARCA4 and SMARCC1. Recently, Mashtalir *et al.,* published an elegant study on the molecular organisation of the SWI/SNF complexes[84]. They demonstrated that the SWI/SNF complex is composed of an ATPase module (SMARCA4/2, ACTB,

ACTL6A/B), a core scaffolding module (SMARCC1/2, SMARCD1/2/3, SMARCE1, SMARCB1) and an ARID module that confers the specificity of the different SWI/SNF conformations (canonical BAF, PBAF and non-canonical BAF). The minimal core module of the SWI/SNF complex is composed of the SMARCD subunit and either a homo or heterodimer of SMARCC1 and SMARCC2. SMARCD1 can interact directly with SMARCC1/2 via the SWIB domain and the unstructured region of the protein towards the N-terminal side of the SWIB domain. There are minimal interactions between the C-terminal end of SMARCD1 and SMARCC1/2. Similarly, most interactions between SMARCA4 and SMARCD1 involve the N-terminal portion of the protein. These results are in agreement with our Co-IPs showing that the tested C-terminal mutations do not disrupt the interaction between SMARCD1 and SMARCC1/SMARCA4. Interestingly, Mashtalir *et al.* show that the C-terminal region located downstream of the SMARCD1 SWIB domain - where our mutations cluster - is implicated in the binding of ARID subunits (ARID1A/B and ARID2), and necessary for the formation of the fully assembled BAF and PBAF conformations of the mammalian SWI/SNF complex.[84]. It will be interesting to test if the mutations identified here affect the interaction with ARID subunits.

Bap60 in memory and MB-specific transcriptome regulation in juvenile *Drosophila*

Although the role of some SWI/SNF components in neuron development and function is well described[85], there was previously no work investigating the function of SMARCD1 in the nervous system. We show here that the *Drosophila* ortholog Bap60 is

required in the adult fly mushroom body for normal memory (**Figure 3**). The requirement for Bap60 in the adult fly brain is consistent with evidence from mammals suggesting the presence of a SWI/SNF complex that is only present in differentiated neurons[3,18,86]. Indeed, it has been shown that the neuron-specific subunit BAF53b is also essential in adults for normal memory[3]. Taken together, this suggests that SWI/SNF-related NDDs might result from defective gene regulation postnatally in differentiated neurons.

Using MB-specific transcriptome analysis we found that Bap60 has a greater effect on gene regulation in the MB of juvenile adult flies compared to mature flies. In particular, Bap60 seems to be important for activating expression of neuronal genes (**Figure 4**) and developmental genes that normally show increased expression in juvenile MBs (**Figure 5**). This is interesting because the MB is known to undergo structural alterations and form new synaptic connections during the early stages of juvenile adult life[76,77,87,88]. These changes are dependent on sensory input, suggesting that some of the brain's circuitry is developed in response to early life experience[75,76,88]. This early experience-dependent plasticity in the mushroom body is required for normal memory at later life stages in flies[76,77,80,81] and bees [78,88]. Although much more complex, human brains also show periods of experience dependent plasticity, especially during adolescence[89]. So called "environmental enrichment" therapy for autism is based on the idea that defects in neural circuitry might be corrected by providing increased sensorimotor experience[89,90]. It will be interesting to further investigate the mechanisms of SWI/SNF mediated gene regulation in experience-dependent brain plasticity. Understanding the role of SWI/SNF in the postnatal brain could open up possibilities for therapy, whereas prenatal developmental intervention seems unlikely.

Here we identify 5 mutations in SMARCD1, a subunit of the BAF complex that has not been previously associated with neurodevelopmental disorder. Moreover, we show that its *Drosophila* ortholog, Bap60, regulates neurodevelopmental gene expression during a critical time window of juvenile adult brain development that is essential in establishing neuronal circuits that are required for learning and memory. Altogether, our results highlight the role of SMARCD1 in establishing proper cognitive functions.

Supplemental Data

Document S1. Figures S1 and S2, Tables S3, S4, and S7, and supplementary methods.

Table S1. Alignment summary of sequencing reads to the reference genome and counts to genic features using HTSeq-Count.

Table S2. Lists of neuron-specific and mushroom body- specific genes.

Table S5. Complete results comparing juvenile Bap60-KD MB with juvenile control and comparing mature Bap60-KD MB with mature control.

Table S6. Complete results comparing mature control with juvenile control and comparing mature Bap60-KD MB with juvenile Bap60-KD MB.

Declaration of Interests

The authors declare no competing interests.

Acknowledgements

Thanks to the individuals and their families who participated in this study. We thank the Bloomington Drosophila Stock Center and the Transgenic RNAi Project at Harvard Medical School for creating and providing fly stocks used in this study. Thanks to David Carter at the London Regional Genomics Centre for help with sequencing. This work was funded by the Canadian Institute of Health Research, the Canada Research Chairs program, the Canadian Foundation for Innovation, the Fonds de Recherche Santé Québec, and by AMED under grant numbers, JP18ek0109280, JP18dm0107090, JP18ek0109301, JP18ek0109348 and JP18kk020500; JSPS KAKENHI Grant Numbers, JP17H01539 and JP16H05160; and Takeda Science Foundation (N. Miyake, N. Matsumoto). We thank Agnès Rastetter for technical assistance and confirmation of the mutation proband 2. The DDD study presents independent research commissioned by the Health Innovation Challenge Fund (grant number HICF-1009-003), a parallel funding partnership between the Wellcome Trust and the Department of Health, and the Wellcome Trust Sanger Institute (grant number WT098051). The views expressed in this publication are those of the author(s) and not necessarily those of the Wellcome Trust or the Department of Health.

Web Resources

Deciphering Developmental Disorders, https://decipher.sanger.ac.uk/

FastQC, http://www.bioinformatics.babraham.ac.uk/projects/fastqc

GeneDX ClinVar submission page,

http://www.ncbi.nlm.nih.gov/clinvar/submitters/26957/

I-TASSER Online Suite, https://zhanglab.ccmb.med.umich.edu/I-TASSER/

NPS@: Network Protein Analysis, https://npsa-prabi.ibcp.fr/

Online Mendelian Inheritance in Man (OMIM), http://www.omim.org

Protein Paint, https://pecan.stjude.cloud/proteinpaint

PyMol, https://pymol.org/2/

References

1. Ho, L. *et al.* An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 5181–6 (2009).
2. Bachmann, C. *et al.* mSWI/SNF (BAF) Complexes Are Indispensable for the Neurogenesis and Development of Embryonic Olfactory Epithelium. *PLoS Genet.* **12,** 1–29 (2016).
3. Vogel-Ciernia, A. *et al.* The neuron-specific chromatin regulatory subunit BAF53b is necessary for synaptic plasticity and memory. SUPP. *Nat. Neurosci.* **16,** 552–61 (2013).
4. Tuoc, T. *et al.* Ablation of BAF170 in Developing and Postnatal Dentate Gyrus Affects Neural Stem Cell Proliferation, Differentiation, and Learning. *Mol. Neurobiol.* **54,** 4618–4635 (2016).
5. Barrett, R. M. & Wood, M. a. Beyond transcription factors: the role of chromatin modifying enzymes in regulating transcription required for memory. *Learn. Mem.* **15,** 460–467 (2008).
6. Kassabov, S. R., Zhang, B., Persinger, J. & Bartholomew, B. SWI / SNF Unwraps , Slides , and Rewraps the Nucleosome. *Mol. Cell* **11,** 391–403 (2003).
7. Whitehouse, I. *et al.* Nucleosome mobilization catalysed by the yeast SWI/SNF complex. *Nature* **400,** 784–787 (1999).
8. van Holde, K. & Yager, T. Models for chromatin remodeling: a critical comparison.

Biochem. Cell Biol. **81,** 169–72 (2003).

9. Lessard, J. *et al.* An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* **55,** 201–215 (2007).
10. Ronan, J. L., Wu, W. & Crabtree, G. R. From neural development to cognition : unexpected roles for chromatin. *Nat. Rev. Genet.* **14,** 347–359 (2013).
11. Kadoch, C. *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* **45,** 592–601 (2013).
12. Olave, I., Wang, W., Xue, Y., Kuo, A. & Crabtree, G. R. Identification of a polymorphic, neuron-specific chromatin remodeling complex. *Genes Dev.* **16,** 2509–2517 (2002).
13. Nguyen, H. *et al.* Epigenetic regulation by BAF (mSWI/SNF) chromatin remodeling complexes is indispensable for embryonic development. *Cell Cycle* **15,** 1317–1324 (2016).
14. Narayanan, R. *et al.* Loss of BAF (mSWI/SNF) Complexes Causes Global Transcriptional and Chromatin State Changes in Forebrain Development. *Cell Rep.* **13,** 1842–1854 (2015).
15. Ninkovic, J. *et al.* The BAF complex interacts with Pax6 in adult neural progenitors to establish a neurogenic cross-regulatory transcriptional network. *Cell Stem Cell* **13,** 403–418 (2013).
16. Tuoc, T. C., Narayanan, R. & Stoykova, A. BAF chromatin remodeling complex: Cortical size regulation and beyond. *Cell Cycle* **12,** 2953–2959 (2013).
17. Matsumoto, S. *et al.* Brg1 is required for murine neural stem cell maintenance and gliogenesis. *Dev. Biol.* **289,** 372–383 (2006).
18. Wu, J. I. *et al.* Regulation of Dendritic Development by Neuron-Specific Chromatin Remodeling Complexes. *Neuron* **56,** 94–108 (2007).
19. Iwase, S. *et al.* Epigenetic Etiology of Intellectual Disability. *J. Neurosci.* **37,** 10773–

10782 (2017).

20. Kleefstra, T., Schenck, A., Kramer, J. M. & Bokhoven, H. Van. The genetics of cognitive epigenetics. *Neuropharmacology* **80,** 83–94 (2014).
21. Tsurusaki, Y. *et al.* Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat. Genet.* **44,** 376–378 (2012).
22. Bramswig, N. C. *et al.* Heterozygosity for ARID2 loss-of-function mutations in individuals with a Coffin–Siris syndrome-like phenotype. *Hum. Genet.* **136,** 297–305 (2017).
23. Wieczorek, D. *et al.* A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum. Mol. Genet.* **22,** 5121–5135 (2013).
24. Santen, G. W. E. *et al.* Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat. Genet.* **44,** 379–380 (2012).
25. Dias, C. *et al.* BCL11A Haploinsufficiency Causes an Intellectual Disability Syndrome and Dysregulates Transcription. *Am. J. Hum. Genet.* **99,** 253–274 (2016).
26. Rivière, J. *et al.* De novo mutations in the actin genes ACTB and ACTG1 cause Baraitser-Winter syndrome. *Nat. Genet.* **44,** 440–444 (2012).
27. Kleefstra, T. *et al.* Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *Am. J. Hum. Genet.* **91,** 73–82 (2012).
28. Koga, M. *et al.* Involvement of SMARCA2/BRM in the SWI/SNF chromatin-remodeling complex in schizophrenia. *Hum. Mol. Genet.* **18,** 2483–2494 (2009).
29. Loe-Mie, Y. *et al.* SMARCA2 and other genome-wide supported schizophrenia-associated genes: Regulation by REST/NRSF, network organization and primate-specific evolution. *Hum. Mol. Genet.* **19,** 2841–2857 (2010).
30. Nord, A. S. *et al.* Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. *Eur. J. Hum. Genet.* **19,** 727–731 (2011).
31. Wolff, D. *et al.* In-Frame Deletion and Missense Mutations of the C-Terminal Helicase

Domain of SMARCA2 in Three Patients with Nicolaides-Baraitser Syndrome. *Mol. Syndromol.* **2,** 237–244 (2011).

32. Van Houdt, J. K. *et al.* Heterozygous missense mutations in SMARCA2 cause Nicolaides-Baraitser syndrome. *Nat. Genet.* **44,** 445–450 (2012).
33. Marom, R. *et al.* Heterozygous variants in ACTL6A , encoding a component of the BAF complex , are associated with intellectual disability. *Hum. Mutat.* **38,** 1365–1371 (2017).
34. Machol, K. *et al.* Expanding the Spectrum of BAF-Related Disorders : De Novo Variants in SMARCC2 Cause a Syndrome with Intellectual Disability and Developmental Delay. *Am. J. Hum. Genet.* **104,** 164–178 (2019).
35. Heisenberg, M., Borst, A., Wagner, S. & Byers, D. Drosophila mushroom body mutants are deficient in olfactory learning. *J. Neurogenet.* **2,** 1–30 (1985).
36. de Belle, J. S. & Heisenberg, M. Associative odor learning in Drosophila abolished by chemical ablation of mushroom bodies. *Science* **263,** 692–5 (1994).
37. Wright, C. F. *et al.* Genetic diagnosis of developmental disorders in the DDD study: A scalable analysis of genome-wide research data. *Lancet* **385,** 1305–1314 (2015).
38. Miyatake, S. *et al.* ANKRD11 variants cause variable clinical features associated with KBG syndrome and Coffin-Siris-like syndrome. *J. Hum. Genet.* **62,** 741–746 (2017).
39. Marsh, A. P. L. *et al.* DCC mutation update: Congenital mirror movements, isolated agenesis of the corpus callosum, and developmental split brain syndrome. *Hum. Mutat.* **39,** 23–39 (2018).
40. Tanaka, A. J. *et al.* Mutations in SPATA5 Are Associated with Microcephaly, Intellectual Disability, Seizures, and Hearing Loss. *Am. J. Hum. Genet.* **97,** 457–464 (2015).
41. Perkins, L. A. *et al.* The transgenic RNAi project at Harvard medical school: Resources and validation. *Genetics* **201,** 843–852 (2015).
42. Li, H. H. *et al.* A GAL4 driver resource for developmental and behavioral studies on the larval CNS of Drosophila. *Cell Rep.* **8,** 897–908 (2014).

43. Chubak, M. C. *et al.* Systematic functional characterization of the intellectual disability-associated SWI/SNF complex reveals distinct roles for the BAP and PBAP complexes in post-mitotic memory forming neurons of the Drosophila mushroom body. *BioRxiv: 408500* (2018). doi:10.1101/408500

44. Möller, A., Avila, F. W., Erickson, J. W. & Jäckle, H. Drosophila BAP60 is an Essential Component of the Brahma Complex , Required for Gene Activation and Repression. *J. Mol. Biol.* **352,** 329–337 (2005).

45. Koemans, T. S. *et al.* Drosophila Courtship Conditioning As a Measure of Learning and Memory. *J. Vis. Exp.* 1–11 (2017). doi:10.3791/55808

46. McGuire, S. E., Mao, Z. & Davis, R. L. Spatiotemporal Gene Expression Targeting with the TARGET and Gene-Switch Systems in Drosophila. *Sci. Signal.* pl6-pl6 (2004). doi:10.1126/stke.2202004pl6

47. Jones, S. G., Nixon, K. C. J., Chubak, M. C. & Kramer, J. M. Mushroom Body Specific Transcriptome Analysis Reveals Dynamic Regulation of Learning and Memory Genes After Acquisition of Long-Term Courtship Memory in Drosophila. *G3 (Bethesda).* **8,** 3433–3446 (2018).

48. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27,** 863–864 (2011).

49. Adams, M. D. *et al.* The Genome Sequence of Drosophila melanogaster. *Science (80-.).* **287,** 2185–2195 (2000).

50. Consortium Drosophila 12 Genomes. Evolution of genes and genomes on the Drosophila phylogeny. *Nature* **450,** 203–218 (2007).

51. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).

52. Anders, S., Pyl, P. T. & Huber, W. HTSeq — a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).

53. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 1–21 (2014).

54. Mi, H. *et al.* PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* **45,** D183–D189 (2017).

55. Brown, J. B. *et al.* Diversity and dynamics of the Drosophila transcriptome. *Nature* **512,** 393–399 (2014).

56. Sobreira, N., Schiettecatte, F., Valle, D. & Hamosh, A. GeneMatcher: A Matching Tool for Connecting Investigators with an Interest in the Same Gene. *Hum. Mutat.* **36,** 928–930 (2016).

57. Priam, P. *et al.* SMARCD2 subunit of SWI / SNF chromatin-remodeling complexes mediates granulopoiesis through a CEBP ε dependent mechanism. *Nat. Genet.* **49,** 753–764 (2017).

58. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536,** 285–291 (2016).

59. Huang, N., Lee, I., Marcotte, E. M. & Hurles, M. E. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet.* **6,** 1–11 (2010).

60. Quinodoz, M. *et al.* DOMINO: Using Machine Learning to Predict Genes Associated with Dominant Disorders. *Am. J. Hum. Genet.* **101,** 623–629 (2017).

61. Wang, K., Li, M. & Hakonarson, H. ANNOVAR : functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38,** e164 (2010).

62. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations a. *Nat. Methods* **7,** 248–249 (2010).

63. Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations : application to cancer genomics. *Nucleic Acids Res.* **39,** e118 (2011).

64. Cooper, G. M. *et al.* Distribution and intensity of constraint in mammalian genomic

sequence. *Genome Res.* **15,** 901–913 (2005).

65. Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R. & Siepel, A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **20,** 110–121 (2010).

66. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15,** 1034–1050 (2005).

67. Garber, M. *et al.* Identifying novel constrained elements by exploiting biased substitution patterns. *Bioinformatics* **25,** i54–i62 (2009).

68. Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M. & Ng, P. C. SIFT missense predictions for genomes. *Nat. Protoc.* **11,** 1–9 (2016).

69. Choi, Y. & Chan, A. P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **31,** 2745–2747 (2015).

70. Shihab, H. A. *et al.* Predicting the Functional, Molecular, and Phenotypic Consequences of Amino Acid Substitutions using Hidden Markov Models. *Hum. Mutat.* **34,** 57–65 (2013).

71. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2 : mutation prediction for the deep-sequencing age. *Nat. Publ. Gr.* **11,** 361–362 (2014).

72. Quang, D., Chen, Y. & Xie, X. DANN : a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics* **31,** 761–763 (2015).

73. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Publ. Gr.* **46,** 310–315 (2014).

74. Kamyshev, N. G., Iliadi, K. G. & Bragina, J. V. Drosophila conditioned courtship: two ways of testing memory. *Learn. Mem.* **6,** 1–20 (1999).

75. Barth, M., Hirsch, H. V, Meinertzhagen, I. A. & Heisenberg, M. Experience-dependent developmental plasticity in the optic lobe of Drosophila melanogaster. *J. Neurosci.* **17,** 1493–504 (1997).

76. Barth, M. & Heisenberg, M. Vision affects mushroom bodies and central complex in Drosophila melanogaster. *Learn. Mem.* **4,** 219–229 (1997).

77. Seugnet, L., Suzuki, Y., Donlea, J. M., Gottschalk, L. & Shaw, P. J. Sleep deprivation during early-adult development results in long-lasting learning deficits in adult Drosophila. *Sleep* **34,** 137–146 (2011).

78. Jones, B. M., Leonard, A. S., Papaj, D. R. & Gronenberg, W. Plasticity of the worker bumblebee brain in relation to age and rearing environment. *Brain. Behav. Evol.* **82,** 250–261 (2013).

79. Doll, C. A. & Broadie, K. Impaired activity-dependent neural circuit assembly and refinement in autism spectrum disorder genetic models. *Front. Cell. Neurosci.* **8,** 1–26 (2014).

80. Doll, C. A. & Broadie, K. Activity-dependent FMRP requirements in development of the neural circuitry of learning and memory. *Development* **142,** 1346–1356 (2015).

81. Doll, C. A., Vita, D. J. & Broadie, K. Fragile X Mental Retardation Protein Requirements in Activity-Dependent Critical Period Neural Circuit Refinement. *Curr. Biol.* **27,** 2318–2330.e3 (2017).

82. Tessier, C. R. & Broadie, K. Drosophila fragile X mental retardation protein developmentally regulates activity-dependent axon pruning. *Development* **135,** 1547–1557 (2008).

83. Kosho, T., Miyake, N. & Carey, J. C. Coffin-Siris syndrome and related disorders involving components of the BAF (mSWI/SNF) complex: Historical review and recent advances using next generation sequencing. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **166,** 241–251 (2014).

84. Mashtalir, N. *et al.* Modular Organization and Assembly of SWI / SNF Family Chromatin Remodeling Complexes Article Modular Organization and Assembly of SWI / SNF Family Chromatin Remodeling Complexes. *Cell* **175,** 1–17 (2018).

85. Sokpor, G., Xie, Y., Rosenbusch, J. & Tuoc, T. Chromatin Remodeling BAF (SWI / SNF) Complexes in Neural Development and Disorders. *Front. Mol. Neurosci.* **10,** 1–22 (2017).

86. Olave, I., Wang, W., Xue, Y., Kuo, A. & Crabtree, G. R. Identification of a polymorphic, neuron-specific chromatin remodeling complex. *Genes Dev.* **16,** 2509–2517 (2002).

87. Heisenberg, M., Heusipp, M. & Wanke, C. Structural plasticity in the Drosophila brain. *J. Neurosci.* **15,** 1951–1960 (1995).

88. Cabirol, A., Brooks, R., Groh, C., Barron, A. B. & Devaud, J. M. Experience during early adulthood shapes the learning capacities and the number of synaptic boutons in the mushroom bodies of honey bees (Apis mellifera). *Learn. Mem.* **24,** 557–562 (2017).

89. Ismail, F. Y., Fatemi, A. & Johnston, M. V. Cerebral plasticity: Windows of opportunity in the developing brain. *Eur. J. Paediatr. Neurol.* **21,** 23–48 (2017).

90. Simpson, J. & Kelly, J. P. The impact of environmental enrichment in laboratory rats-Behavioural and neurochemical aspects. *Behav. Brain Res.* **222,** 246–264 (2011).

Figures Legends

Figure 1: Characterization of *SMARCD1* variants at the protein level. (**A**) Primary structure of SMARCD1. The SWIB domain is shown in green and the predicted coiled-coil domains (C1, C2, C3) are shown in pink. The location of the identified variants is indicated. Missense and nonsense variants are in blue and orange, respectively. (**B**) Amino acid conservation of regions around missense variants across several species. (**C**) Predicted 3D model of SMARCD1. Missense variants are show in red. Residues truncated by the W486* variant are colored in blue, and the following residues truncated by both W486* and R503* variants are colored in orange.

Figure 2: Photos of individuals identified with *SMARCD1* mutations. Individual 1, age 7 y, note ear malformation, sparse hair and hypoplastic 5th toenail. Individual 2, age 11 y, note upturned, thick earlobes, low hairline, short hands and slender fingers. Individual 3, age 6 y, note low hairline, long eyebrows, upturned earlobes, small hands and feet. Individual 4, age 2 y 11 m, note flat nasal bridge, hypertelorism, prominent philtral pillars, broad square face with temporal narrowing, wide mouth with downturned corners, myopathic facial appearance, small hands.

Figure 3: Adult-specific knockdown of Bap60 in the MB results in reduced capacity for long-term courtship memory. (**A**) Boxplots showing courtship indices (CI) for the mCherry RNAi control and Bap60 knockdown flies in long-term courtship conditioning assays (+ indicates the mean). The mean CI for naïve (N) and trained (T) flies of the same genotype was compared using a Wilcoxon test. (**B**) Learning indices (LI) for control and Bap60 knockdown flies. Adult specific Bap60 knockdown resulted in

a significantly reduced LI (randomization test, 10,000 bootstrap replicates) relative to control flies. * p<0.05, ** p<0.01, **** p<0.0001.

Figure 4: Bap60 is required for expression of neuron-specific genes in the juvenile MB. (**A** and **B**) Volcano plots showing differentially expressed genes (p_{adj}<0.05 and fold-change > 1.5) represented in blue (downregulated) or orange (upregulated) in (**A**) juvenile Bap60-KD MB and (**B**) mature Bap60-KD MB flies compared to controls of the same age. (**C**) Fold enrichment of muscle-, neuron-, and mushroom body- specific genes (Brown *et al*. and Jones *et al.* [47,55]; see methods) among downregulated genes in juvenile and mature Bap60-KD MBs (**p<0.01; ****p<0.0001; hypergeometric test). (**D**) Average normalized expression (±SEM) of neuron- and mushroom body- specific genes in control (blue) and *Bap60-KD MB* (orange) flies at the juvenile and mature stages.

Figure 5: Bap60 is required for increased expression of developmental genes in the juvenile MB. (**A**) Venn diagram showing overlap genes that are significantly increased in expression in the juvenile MB compared to the mature MB (juvenile-induced genes) in control and Bap60-KD. (**B**) Average normalized expression (±SEM) of the 164 genes that are induced in the juvenile MB in controls only. (**C**) Gene ontology enrichment for biological processes of the 164 juvenile-induced genes identified in controls only. Terms with at least 10 genes and an FDR < 0.05 are displayed.

Table 1: Summary of clinical findings (see Table S3 for additional details)

Individual	1	2	3	4	5
Mutation	NM_003076.4 c.990C>G, p.Asp330Glu	NM_003076.4 c.1336A>G, p.Arg446Gly	NM_003076.4 c.1457G>A, p.Trp486*	NM_003076.4 c.1483T>C, p.Phe495Leu	NM_003076.4 c.1507C>T, p.Arg503*
	Heterozygous in child, absent in mother, father not available	*de novo*	*de novo*	*de novo*	*de novo*
Gender	Female	Male	Male	Male	Female
Duration of gestation (weeks)	37w2d	38	40-41	40	39.4
Birth Weight, g	2598 (25th %ile)	4150 (>99th %ile)	3884 (90th %ile)	3480 (75th %ile)	2835 (20th %ile)
Birth Length, cm	44,5 (1st %ile)	52 (75th %ile)	54 (75th %ile)	NA	48.9 (5th %ile)
Birth OFC, cm	32 (3rd %ile)	37 (>99th %ile)	NA	NA	31.75 (10th %ile)
Congenital anomalies	Esophageal atresia, bronchial fistula.	Agenesis of the corpus callosum, macrosomia	Ankyloglossia (operated)	Mild hydronephrosis on antenatal USS	
Post-natal - age at last assessment	7y.o.	11y 6m	6y 3m	3y	27 months
Weight, kg	17,74 (3rd %ile)	75.5 kg (+3.3 SD)	19.4kg (21st %ile)	16.5kg (97th %ile)	9.9 kg (-2.2 SD)
Height, cm	106,2 (-2.9 SD)	161 cm (+2 SD)	105cm (-2.3 SD)	111.8cm (+3.9 SD)	84.2 cm (15th %ile)
OFC, cm	NA	59cm (+3.8 SD)	52cm (60th %ile)	49.3 cm (41 %ile)	44.3 cm (-2.4 SD)
Developmental delay	+	+	+	+	+
Delayed walking	+	+ (17 m)	+ (2 y)	Not attained	- (15 m)
Delayed speech	NA	+ (uses sentences)	Not attained, some sign language	Not attained	+ (words 18 m, now 2-3 word phrases)
Intellectual disability	+	+	+	+	NA
Hypotonia	+	-	+	+	-
Neuroradiology		MRI age 1 month : complete agenesis of the corpus callosum	No abnormality seen	MRI age 3: Prominence of the lateral and third ventricles. Cavum septum pellucidum.	MRI at age 22months: Nonspecific small right frontal subcortical white matter FLAIR

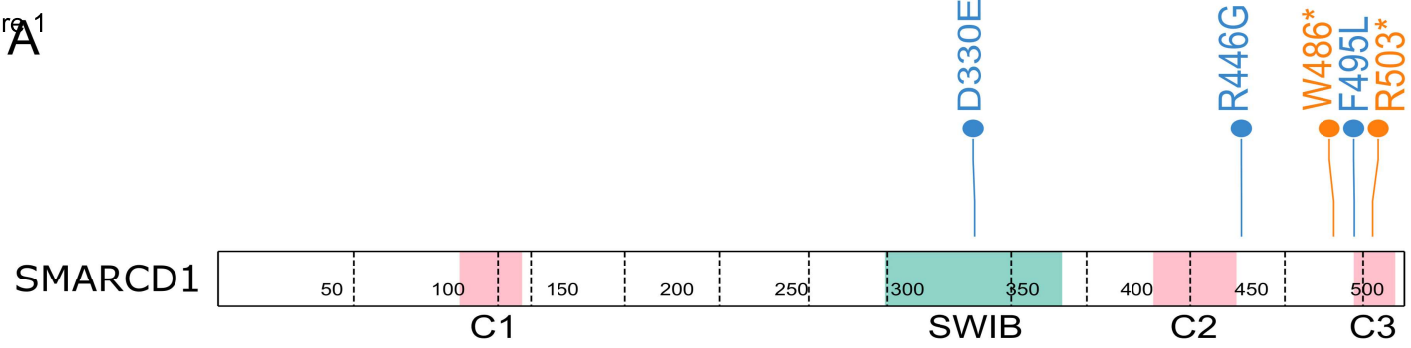
				T1 low, T2/FLAIR high in peritrigonal white matter bilaterally. Prominent perivascular spaces within the deep white matter of frontal lobes bilaterally. 4 mm rounded area of T1 low signal at pineal gland, with rim enhancement on postcontrast. Stable.	hyperintensity.
Feeding difficulties	Feeding and sucking difficulties	Tube feeding during 2 first weeks. No problem after.	Slow feeder, slow with milk and solids, gastroesophageal reflux	Ongoing, Nasogastric feeding followed by fundoplication & PEG tube, gastroesophageal reflux	Gastroesophageal reflux in infancy and difficulty with regulating self-feeding
Hair	Sparse	Low hairline	low hairline	Temporal deficiency	NI
Ears	external ear malformation	Upturned, thick ear lobes	upturned earlobes	low-set, thickened, simple shape	NI
Hearing loss	+, malformation of ossicle	-	+, mild bilateral	-	-
Hands and feet	Hypoplastic fifth toenails	Short hands, slender fingers	small hands and feet	Thick loose palmar & plantar skin, small hands & feet, short 5th fingers	NI
Teeth anomalies	-	-	small rounded teeth and pointed canines	thick, stubby teeth, delayed eruption, thick gums	-
Other dysmorphisms or relevant information	High palate, cleft soft palate		frontal bossing, long eyebrows, plagiocephaly	broad, square face, temporal narrowing flat nasal bridge,	Prominent metopic suture Bulbous nasal tip

				short nose, very prominent philtral pillars hypertelorism, divergent squint wide mouth, downturned corners, thick gums, high palate, "jowly" , myopathic appearance;, undescended testes,	Bifid uvula with normal palate Facial asymmetry when crying

Legend: +, present; -, absent; NA, not available; Nl, normal.

Figure 1

A



B

	Asp330					Arg446					Phe495				
Human	L	Q	**D**	P	H	F	A	**R**	D	P	R	Y	**F**	Y	S
Macaque	L	Q	**D**	P	H	F	A	**R**	D	P	R	Y	**F**	Y	S
Mouse	L	Q	**D**	P	H	F	A	**R**	D	P	R	Y	**F**	Y	S
Dog	L	Q	**D**	P	H	F	A	**R**	D	P	R	Y	**F**	Y	S
Elephant	L	Q	**D**	P	H	F	A	**R**	D	P	R	Y	**F**	Y	S
Chicken	L	Q	**D**	P	H	F	A	**R**	D	P	R	Y	**F**	Y	S
Xenopus	L	Q	**D**	P	H	F	A	**R**	D	P	R	Y	**F**	Y	S
Zebrafish	L	Q	**D**	P	H	F	A	**R**	D	P	R	Y	**F**	Y	S
Lamprey	L	Q	**D**	S	H	F	A	**Q**	D	P	R	Y	**F**	Y	S
C. elegans	L	Q	**D**	P	Q	F	S	**N**	E	P	R	Y	**M**	Y	Q
Drosophila	L	Q	**D**	A	H	F	A	**K**	D	P	R	Y	**F**	F	T

C

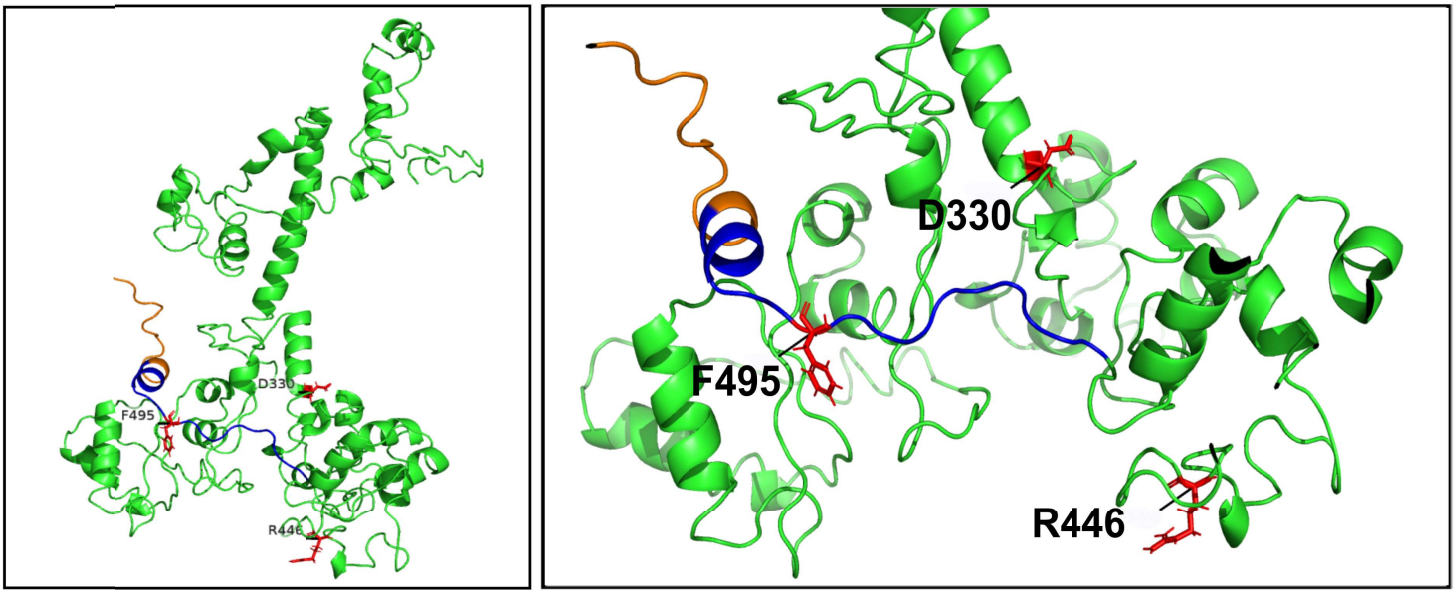


Figure 2

Click here to access/download;Figure;Figure 2.tif

Figure 3

Click here to access/download;Figure;Figure 3.tif

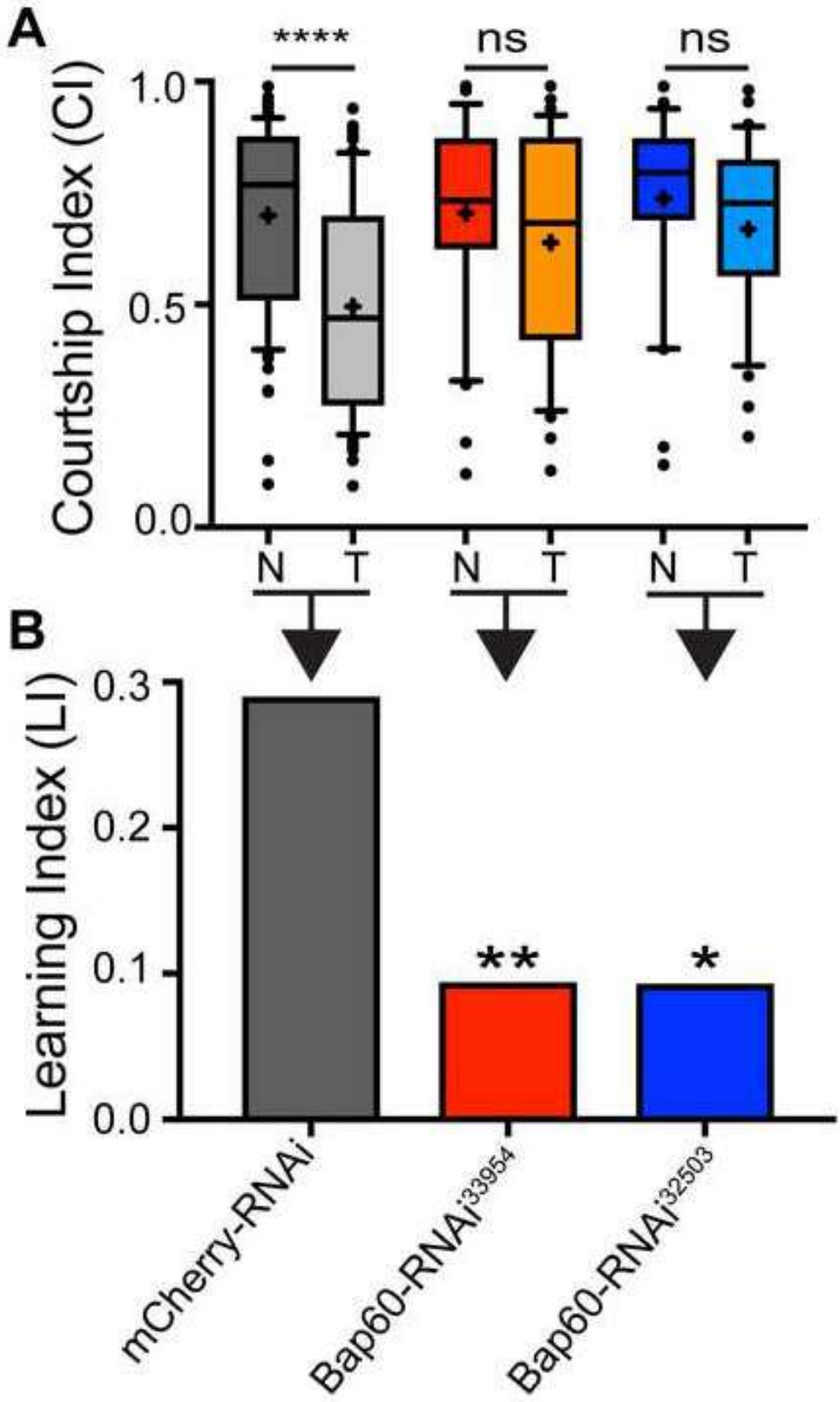


Figure 4

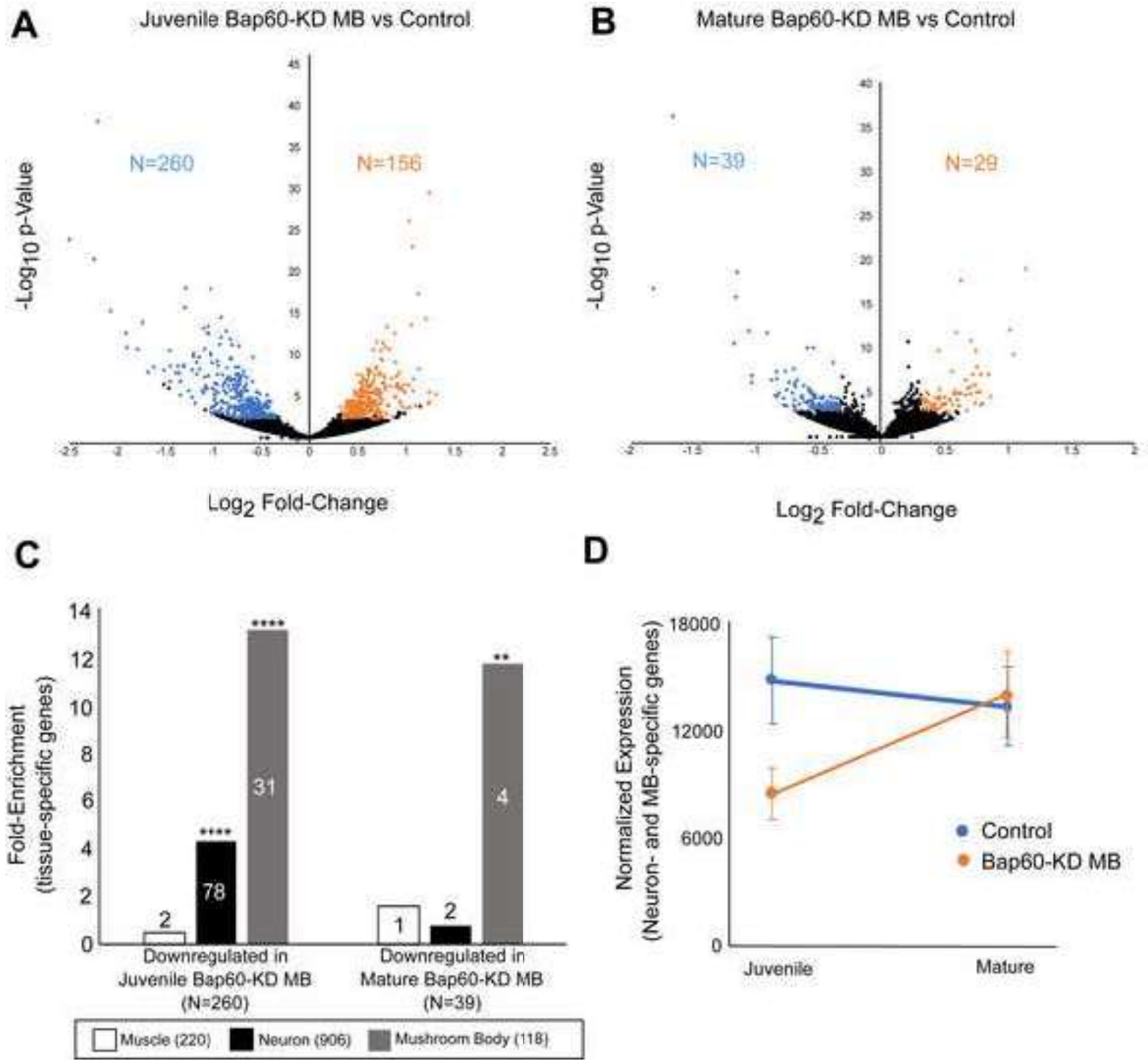


Figure 5