Identification of *SMARCD1* as a syndromic intellectual disability gene that is required for memory and context-dependent regulation of neuronal genes in *Drosophila*

Kevin C.J. Nixon^{1,†}, Justine Rousseau^{2,†}, Max H. Stone^{3,4}, Mohammed Sarikahya³, Sophie Ehresmann², Seiji Mizuno⁵, Naomichi Matsumoto⁶, Noriko Miyake⁶, DDD study, Diana Baralle⁷, Shane McKee⁸, Kosuke Izumi⁹, Alyssa Ritter⁹, Solveig Heide¹⁰, Delphine Héron¹⁰, Christel Depienne^{11,12,13}, Hannah Titheradge¹⁴, Jamie M. Kramer^{1,3,4*}, and Philippe M. Campeau^{2,15*}

Primary affiliations only:

¹Department of Physiology and Pharmacology, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada.

²CHU-Sainte Justine Research Center, University of Montreal, Montreal, QC, Canada.

³Department of Biology, Faculty of Science, Western University, London, Ontario, Canada.

⁴Division of Genetics and Development, Children's Health Research Institute, London, Ontario, Canada.

⁵Department of Pediatrics, Central Hospital, Aichi Human Service Center, Aichi, Japan.

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

⁷Faculty of Medicine, University of Southampton, Southampton, UK.

⁸Belfast City Hospital, Belfast, Northern Ireland, UK.

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

¹⁰AP-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.

¹¹INSERM, 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.

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

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

¹⁴Birmingham Women's and Children's NHS Foundation Trust Mindelsohn Way,
Birmingham, UK.

¹⁵Department of Pediatrics, University of Montreal, Montreal, QC, Canada.

†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: 519-661-374

James.Kramer@schulich.uwo.ca

Abstract

Mutations in several genes encoding components of the SWI/SNF chromatin remodeling complex cause syndromic intellectual disability (ID). Here, we report on 5 individuals with mutations in the *SMARCD1* gene, presenting with ID, developmental delay, hypotonia, feeding difficulties, and small extremities. The mutations were proven to be *de novo* in 4 of the 5 individuals. Mutations in other SWI/SNF components cause Coffin-Siris, Nicolaides-Baraitser, or other syndromic ID disorders. Although the individuals presented here have some clinical overlap with these disorders, 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 *SMARCD1* mutations as a novel cause of ID, 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 adult animals^{1–4}. Chromatin structure is an important factor in modulating gene expression⁵. Disruption of genes encoding chromatin regulators can result in intellectual disability (ID)^{6,7}, a prevalent neurodevelopmental disorder characterized by limitations in adaptive behaviour and intellectual function. Mutations in several subunits of the SWI/SNF (SWItch/Sucrose Non-Fermenting) ATP-dependent chromatin remodeling complex (also known as the BAF complex in mammals) cause syndromic ID disorders, such as Nicolaides-Baraitser Syndrome (OMIM: 601358) and Coffin-Siris Syndrome (OMIM: 135900)^{8–11}. SWI/SNF mutations are also potentially involved in other neurodevelopmental disorders such as schizophrenia, and autism spectrum disorder^{12–15}, emphasizing the importance of this complex in neural development and function. In total, 9 of the 28 genes encoding SWI/SNF components have been implicated in ID, yet it remains to be determined if the remaining subunits are also involved in ID or other neuronal disorders.

The SWI/SNF complex is a highly conserved chromatin remodeling complex that hydrolyzes ATP to generate the energy required to alter nucleosome-DNA interactions resulting in more open chromatin for transcription factor binding and subsequent gene expression activation^{3,16–18}. 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^{1,19–22}. The SWI/SNF complex is important for regulation of gene expression programs involved in neuronal differentiation and brain-region specification in mice^{2,4,19,23–27}. However, the complex is also essential

in mature neurons for memory formation, synaptic plasticity, and activity responsive neurite outgrowth^{3,4,28}.

Here, we characterize novel mutations in the *SMARCD1* gene in patients presenting with syndromic ID. *SMARCD1* encodes a core component of the SWI/SNF complex that has not previously been associated with ID, or any other neurodevelopmental condition. Furthermore, 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^{29,30}. 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.

Methods

Patient 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 DDD study and 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 described³¹. For individual 2, exome sequencing was performed as described in³² (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 described³³. 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 described³⁴. The general assertion criteria for variant classification are publicly available on the GeneDx ClinVar submission page (<http://www.ncbi.nlm.nih.gov/clinvar/submitters/26957/>).

Co-Immunoprecipitation

HEK293T and SK-N-AS were cultured in DMEM, 10% FBS (Wisent), 1X AA and 1X Glutamax all from ThermoFisher Scientific. SK-N-AS media was supplemented with 0.1mM NEAA (ThermoFisher Scientific) and 25mM HEPES (Wisent). pcDNA.3.1-Flag-SMARCD1 wild type (NM_003076.4) or SMARCD1-c.990C>G (p.Asp330Glu),

c.1457G>A (p.Trp486*) and c.1483T>C (p.Phe495Leu) were transfected using jetPRIME (Polyplus-transfection) according to the manufacturer's instructions. After 48h, cells were washed once with ice-cold phosphate-buffered saline, lysed in a hypotonic buffer (25 mM HEPES pH 7.9, 25 mM KCL, 50mM EDTA, 5 mM MgCl₂, 10% glycerol, 0.1% NP40, 1 mM DTT, complete mini protease inhibitors (Sigma-Aldrich)), and centrifuge at 3000 rpm for 3 minutes. The pelleted nuclei fraction (NE) was resuspended in nuclear lysis buffer (25 mM Tris, pH8, 150 mM NaCl, 1 mM EDTA, 1% triton, 5% glycerol, complete mini protease inhibitors), incubated 30 min at 4 degrees and centrifuge at 13 000 rpm for 20 min. For immunoprecipitation, nuclear extracts (100 to 300ug) were incubated with either pre-coupled M2 anti-FLAG (M8823, Sigma-Aldrich) or anti-Brg1 (H10) coupled to Dynabeads (ThermoFisher Scientific) magnetic beads. After 1.5 and 4 hours respectively, beads were washed 4 times and proteins eluted directly in 1X laemmli loading buffer supplemented with 50 mM DTT. Western blot was performed using the following antibodies: anti-SMARCC1 (PCRP-SMARCC1-1F1, DSHB), BRG-1 (H10) (sc-374197, Santa-Cruz), anti-FLAG M2 (368791, Sigma-Alrich), anti-BRG/BRM (J1) (a generous gift from J.Lessard, IRIC).

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)³⁵ (*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 Flylight³⁶ 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.

Courtship Conditioning

Courtship conditioning assays were performed as described previously³⁷. Training was performed by pairing individual males with predated 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 ($CI_{\text{naïve}}$) 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 ($LI = (CI_{\text{naïve}} - CI_{\text{trained}}) / CI_{\text{naïve}}$). LIs were compared between genotypes using a randomization test⁵² (10,000 bootstrap replicates) performed with a custom R script³⁷ and the resulting p-values were corrected for multiple testing using the method of Bonferroni.

The temperature-sensitive Gal80^{ts} system³⁸ 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-Bap60³²⁵⁰³* and *UAS-Bap60³³⁹⁵⁴*, as well as a control *UAS-mCherry^{RNAi}* that is present in the same genetic background as the RNAi transgenes. Crosses were raised at 18°C, the permissive temperature which allows Gal80^{ts} to repress Gal4 mediated transcription. At eclosion, male flies of the genotypes

Gal80^{ts/+}; R14H06-Gal4/UAS-Bap60³²⁵⁰³, *Gal80^{ts/+}; R14H06-Gal4/UAS-Bap60³³⁹⁵⁴* and *Gal80^{ts/+}; R14H06-Gal4/UAS-mCherry^{RNAi}* were transferred to 29°C, 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)

Fifty juvenile (0-3h old) or mature (1-5 days old) adult male flies of the genotype UAS-unc84::GFP/+; R14H06-Gal4/UAS-Bap60³²⁵⁰³ (Bap60-KD) and UAS-unc84::GFP/+;R14H06-Gal4/UAS-mCherry^{RNAi} (control) were anesthetized using CO₂ 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 MgCl₂, 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 trimming³⁹ and a minimum base quality score of 30. Read quality was then assessed using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>). Trimmed reads were aligned to the *Drosophila melanogaster* reference genome (BDGP release 6)^{40,41} using the STAR aligner⁴². 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 and controls 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 and controls respectively (**Table S1**). The number of reads per gene was quantified by inputting resulting reads into HTSeq-count⁴³ with the --type flag indicating 'gene' (**Table S1**). Genes encoding rRNA and genes with no counts across all samples were removed from analysis, 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 R⁴⁴ package DESeq2⁴⁵. 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 Panther⁴⁶ (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.*⁴⁷ 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 log₂ 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 published MB-specific gene expression that were obtained using the same INTACT protocol described here⁴⁸. 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 MB flies was determined using a hypergeometric test.

Results

Human Genetics 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 identified⁸. As an expansion of this study, an individual was identified with a *SMARCD1* variant (p.Asp330Glu), but the *de novo* status of the variant could not be confirmed since the father was not available for testing. As part of the DDD study, two individuals were identified with *SMARCD1* variants (p.Trp486* and p.Phe495Leu). Through Genematcher, two additional individuals were identified, a patient with a *de novo* p.Arg446Gly 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 (p.Arg503*).

Figure 1: Protein structure of *SMARCD1* and associated domains with respective variants. SWIB and coiled-coil domains are shown in red. The location of the identified variants are indicated with black arrows. Conserved regions of *SMARCD1* across species are shown below.

The variants and the associated clinical features are shown in **Figure 1, Table 1, and Table S3**. There is phenotypic overlap with Coffin-Siris syndrome in that two patients have a hypoplastic 5th toenail and all have ID, however, most 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). 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).

c.990C>G, p.Asp330Glu

c.1336A>G, p.Arg446Gly

c.1457G>A, p.Trp486*

c.1483T>C, p.Phe495Leu

Figure 2: Photos of individuals identified with *SMARCD1* mutations.

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

We performed some bioinformatic analysis of the observed variants. *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 variants⁴⁹. 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%)⁵⁰. DOMINO gives at 99.8% probability for the gene to be associated with an autosomal dominant condition⁵¹.

We used wANNOVAR to predict pathogenicity scores, those mostly considering conservation (polyphen, LRT, MutationAssessor, GERP, PhyloP, phastCons, SiPhy), function (SIFT, PROVEAN), or combination thereof (FATHMM, MutationTaster, DANN, CADD). 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 1**).

***In vitro* assessment of the variants**

We generated expression constructs with three different variants (p.Asp330Glu, p.Trp486*, p.Phe495Leu). None of the mutations affect protein stability, as all were detectable at normal levels by Western blot. Upon co-immunoprecipitation, we noted that all three mutants bind well to BRG1 and SMARCC1 (**Figure S1**). Therefore, the mutations might not affect integration in the SWI/SNF complex, but they might affect the function of SMARCD1 otherwise, possibly by changing the activity of the complex or interactions with other proteins.

Figure 3: Adult-specific knockdown of Bap60 in the MB results in reduced capacity for both short-term and long-term courtship memory. (A) Boxplots showing courtship indices (CI) for the mCherry RNAi control and Bap60 knock down 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.

The *Drosophila* SMARCD1 ortholog Bap60 is required for memory

Given that some of the mutations, such as the truncating mutations and the SWIB domain mutation might lead to a loss-of-function, we assessed in *Drosophila* the consequence of loss of Bap60, the *Drosophila* ortholog of SMARCD1. To do this, we generated Bap60 knockdown flies using previously characterized transgenic RNAi lines⁵². 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). We tested memory in MB-specific Bap60 knockdown flies using a classic paradigm known as courtship conditioning^{37,53}. In this assay, male flies display a learned reduction of courtship behavior after sexual rejection by a non-receptive pre-mated female. Control flies expressing an RNAi construct targeting mCherry showed a normal reduction in courtship behavior after being sexually rejected by a mated female (**Figure 3A**). This reduction in courtship was not seen in flies expressing RNAi transgenes targeting Bap60 (**Figure 3A**) and Bap60 knockdown flies had a significant reduction in the learning index (LI) compared to the controls (**Figure 3B**). Both Bap60 RNAi lines, which have unique target sequences, showed a similar reduction in memory, which suggests 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.

Figure 4: Mushroom body knockdown of Bap60 expression leads to

dysregulation of tissue-specific genes. (A and B) Volcano plots of gene expression

and p-Value. Differentially expressed genes ($P_{adj} < 0.05$ and fold-change > 1.5) are

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.* ^{47,48}; see methods) in downregulated genes in juvenile and mature Bap60-

KD MBs compared to the entire gene set analyzed (** $P < 10^{-3}$; **** $P < 10^{-25}$;

Hypergeometric test). **(D)** Average normalized expression ($\pm$ SEM) of neuron- and

mushroom body- specific genes in control (blue) and *Bap60-KD MB* (orange) flies at the

juvenile and mature stages.

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^{54–60}. During this time, neuronal connections are formed that are critical for normal learning and memory later in life^{56,58}. We compared the MB-specific transcriptome in Bap60 knockdown flies compared to controls in early juvenile adults (0–3 h 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 control using DESeq2⁴⁵ yielded 416 differentially expressed genes ($P_{\text{adj}} < 0.05$ and Fold-Change > 1.5), of which 156 were upregulated and 260 were downregulated in the Bap60-KD MBs (**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 differentially expressed genes. DE 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

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 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^{47,48} to establish a list of 904 “neuron-specific genes” that are enriched in heads compared to other tissues, and 118 “MB-specific genes” that are enriched in MB-specific INTACT samples compared to whole head (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 would be expected by random ($p < 10^{-25}$, 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 represented among genes that were downregulated in Bap60-KD MBs. Taken together, this data suggests Bap60 plays a context dependent role in activating the expression of neuron-specific genes in the MB.

Figure 5: Mushroom body knockdown of Bap60 expression leads to dysregulation of developmental genes at the juvenile stage. (A) Venn diagram showing overlap of upregulated genes (P_{adj} < 0.05 and fold-change > 2) between control and Bap60-KD MB files at the juvenile stage compared to a mature stage (juvenile-induced genes). (B) Average normalized expression (±SEM) of the 164 control juvenile-induced genes from (A) in both control (blue) and Bap60-KD MB (orange) flies at the juvenile and mature stages. (C) Gene ontology enrichment for biological processes of juvenile-induced genes control flies. Terms with at least 10 genes and an FDR < 0.05 are displayed.

Bap60 is required for expression of developmental genes that are preferentially activated during juvenile MB development.

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^{58,61}. To test this, we performed DE 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 165 were not (**Figure 5A**). On average, these 165 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, these 165 genes 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^{54–60}.

Discussion

Characterization of a novel neurodevelopmental disorder

In this study, we describe a novel genetic disorder characterized by mutations in *SMARCD1*, which encodes a component of the SWI/SNF chromatin remodeling complex. Several other SWI/SNF genes are implicated in syndromic neurodevelopmental disorders that are typically characterized by ID, abnormalities of the 5th digit, and characteristic facial features. The 5 individuals described here do fit in this spectrum, with ID and a low penetrance of 5th digit abnormalities, but lack the typical facial dysmorphisms seen in other SWI/SNF related disorders (**Figure 2, Table 1**).

The identified *SMARCD1* variants, 2 nonsense and 3 missense, are all clustered in the C-terminal end of the protein (**Figure 1**). The two nonsense variants are located in the last exon of the transcript or within 50 base pairs of the last exon junction, so should escape nonsense mediated decay and produce a truncated protein. One missense variant is located in the highly conserved SWIB domain, while the other two are not located in any known domain. Although these missense and protein-truncating mutations are predicted to be damaging, we do not know the precise functional effect of these mutations. We show here that the proteins are stable and do not disrupt interaction with other members of the SWI/SNF complex, BRG1 and SMARCC1. These mutations might affect binding to other proteins such as transcription factors or other SWI/SNF components that we did not test. It is possible that these mutated proteins have a loss of function or dominant negative effect by integrating into the complex and compromising its normal activity, however, this remains to be tested. Notably, missense

mutations are quite common in other SWI/SNF related disorders. While nonsense mutations are common in some SWI/SNF genes (*ARID1A*, *ARID1B*, and *ARID2*), *SMARCA4*, *SMARCA2*, *SMARCB1*, and *SMARCE1* are primarily affected by missense mutations with a predicted dominant negative effect⁶².

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 described⁶³, 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^{3,28,64}. Indeed, it has been shown that the neuron-specific subunit BAF53b is also essential in adults for normal memory³. Taken together, this suggests that SWI/SNF-related ID 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^{55,56,65,66}. These changes are dependent on sensory input, suggesting

that some of the brain's circuitry is developed in response to early life experience^{54,55,66}. This early experience dependent plasticity in the mushroom body is required for normal memory at later life stages in flies^{55,56,59,60} and bees^{57,66}. Although much more complex, human brains also show periods of experience dependent plasticity, especially during adolescence⁶⁷. 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^{67,68}. It will be interesting to see if SMARCD1 is required for experience-dependent brain plasticity in mammals.

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 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.

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. *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. Iwase, S. *et al.* Epigenetic Etiology of Intellectual Disability. *J. Neurosci.* **37**, 10773–10782 (2017).
7. Kleefstra, T., Schenck, A., Kramer, J. M. & Bokhoven, H. Van. The genetics of cognitive epigenetics. *Neuropharmacology* **80**, 83–94 (2014).
8. Tsurusaki, Y. *et al.* Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat. Genet.* **44**, 376–378 (2012).
9. Bramswig, N. C. *et al.* Heterozygosity for ARID2 loss-of-function mutations in individuals with a Coffin–Siris syndrome-like phenotype. *Hum. Genet.* (2017). doi:10.1007/s00439-017-1757-z
10. 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).
11. 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).
 12. Kleefstra, T. *et al.* Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *Am. J. Hum. Genet.* **91**, 73–82 (2012).
 13. Koga, M. *et al.* Involvement of SMARCA2/BRM in the SWI/SNF chromatin-remodeling complex in schizophrenia. *Hum. Mol. Genet.* **18**, 2483–2494 (2009).
 14. 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).
 15. 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).
 16. Kassabov, S. R., Zhang, B., Persinger, J. & Bartholomew, B. SWI / SNF Unwraps , Slides , and Rewraps the Nucleosome. *Mol. Cell* **11**, 391–403 (2003).
 17. Whitehouse, I. *et al.* Nucleosome mobilization catalysed by the yeast SWI/SNF complex. *Nature* **400**, 784–787 (1999).
 18. van Holde, K. & Yager, T. Models for chromatin remodeling: a critical comparison. *Biochem. Cell Biol.* **81**, 169–72 (2003).
 19. Lessard, J. *et al.* An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* **55**, 201–215 (2007).
 20. Ronan, J. L., Wu, W. & Crabtree, G. R. From neural development to cognition : unexpected roles for chromatin. (2013). doi:10.1038/nrg3413
 21. 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).

22. 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).
23. Nguyen, H. *et al.* Epigenetic regulation by BAF (mSWI/SNF) chromatin remodeling complexes is indispensable for embryonic development. *Cell Cycle* **15**, 1317–1324 (2016).
24. 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).
25. 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).
26. Tuoc, T. C., Narayanan, R. & Stoykova, A. BAF chromatin remodeling complex: Cortical size regulation and beyond. *Cell Cycle* **12**, 2953–2959 (2013).
27. Matsumoto, S. *et al.* Brg1 is required for murine neural stem cell maintenance and gliogenesis. *Dev. Biol.* **289**, 372–383 (2006).
28. Wu, J. I. *et al.* Regulation of Dendritic Development by Neuron-Specific Chromatin Remodeling Complexes. *Neuron* **56**, 94–108 (2007).
29. Heisenberg, M., Borst, A., Wagner, S. & Byers, D. Drosophila mushroom body mutants are deficient in olfactory learning. *J. Neurogenet.* **2**, 1–30 (1985).
30. de Belle, J. S. & Heisenberg, M. Associative odor learning in Drosophila abolished by chemical ablation of mushroom bodies. *Science* **263**, 692–5 (1994).
31. 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).
32. 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).

33. 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).
34. 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).
35. Perkins, L. A. *et al.* The transgenic RNAi project at Harvard medical school: Resources and validation. *Genetics* **201**, 843–852 (2015).
36. 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).
37. 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
38. McGuire, S. E., Mao, Z. & Davis, R. L. Spatiotemporal Gene Expression Targeting with the TARGET and Gene-Switch Systems in *Drosophila*. *Sci. Signal.* p16-p16 (2004). doi:10.1126/stke.2202004p16
39. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864 (2011).
40. Adams, M. D. *et al.* The Genome Sequence of *Drosophila melanogaster*. *Science* (80-.). **287**, 2185–2195 (2000).
41. Consortium *Drosophila* 12 Genomes. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **450**, 203–218 (2007).
42. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
43. Anders, S., Pyl, P. T. & Huber, W. HTSeq — a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
44. R Core Team. R: A language and environment for statistical computing. (2018). at <<https://www.r-project.org/>>
45. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion

- p>for RNA-seq data with DESeq2. 1–21 (2014). doi:10.1186/s13059-014-0550-8
46. 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).
 47. Brown, J. B. *et al.* Diversity and dynamics of the *Drosophila* transcriptome. *Nature* **512**, 393–399 (2014).
 48. 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)*. (2018). doi:10.1534/g3.118.200560
 49. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
 50. Huang, N., Lee, I., Marcotte, E. M. & Hurles, M. E. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet.* **6**, 1–11 (2010).
 51. Quinodoz, M. *et al.* DOMINO: Using Machine Learning to Predict Genes Associated with Dominant Disorders. *Am. J. Hum. Genet.* **101**, 623–629 (2017).
 52. 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
 53. Kamyshev, N. G., Iliadi, K. G. & Bragina, J. V. *Drosophila* conditioned courtship: two ways of testing memory. *Learn. Mem.* **6**, 1–20 (1999).
 54. 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).
 55. Barth, M. & Heisenberg, M. Vision affects mushroom bodies and central complex in

- Drosophila melanogaster. *Learn. Mem.* **4**, 219–229 (1997).
56. 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).
57. 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).
58. 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).
59. Doll, C. A. & Broadie, K. Activity-dependent FMRP requirements in development of the neural circuitry of learning and memory. *Development* **142**, 1346–1356 (2015).
60. 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).
61. Tessier, C. R. & Broadie, K. Drosophila fragile X mental retardation protein developmentally regulates activity-dependent axon pruning. *Development* **135**, 1547–1557 (2008).
62. Bögershausen, N. & Wollnik, B. Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders. *Front. Mol. Neurosci.* **11**, 1–18 (2018).
63. 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).
64. 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).
65. Heisenberg, M., Heusipp, M. & Wanke, C. Structural plasticity in the Drosophila brain. *J. Neurosci.* **15**, 1951–1960 (1995).

66. 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).
67. 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).
68. 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

Figure S1: Mutations in SMARCD1 do not impair its binding to other SWI/SNF components. Co-Immunoprecipitation of SMARCA4 (BRG1) and SMARCC1 (**A**) and reciprocally of SMARCD1 (**B**) following immunoprecipitation of FLAG-SMARCD1 or endogenous SMARCA4 (BRG1) respectively.

Figure S2: qPCR validation of RNAseq gene expression changes. Average

normalized gene expression $\pm$ SD of genes as measured through RT-qPCR (left panel) and RNA-Seq (right panel) in Bap60-KD MB and control mushroom body (MB) samples. RT-qPCR gene expression is normalized to the housekeeping genes *Eif2 β* and *β cop* and RNA-Seq gene expression is normalized from DESeq2. **(A)** Validation of overall gene expression trends for *Ldh*, *MFS3*, *sls*, *dysf*, *tmod*, *Obp56d*, *prt*, *jdp*, *sNPF*, and *Est-6*. Comparisons were made between juvenile control and juvenile Bap60-KD MB and between mature control and mature Bap60-KD MB. Genes validated by RT-qPCR were: *Ldh*, *sls*, *dysf*, *tmod*, *prt*, and *jdp*. Genes that were showing similar trends between RT-qPCR and RNA-Seq were: *MFS3*, *Obp56d*, and *Est-6*. *sNPF* could not be validated by RT-qPCR due to variation between biological replicates. **(B)** Validations of gene expression trends of developmental genes: *Salm*, *Sosie*, *kn*, and *Pvf1* that were induced in juvenile control flies, but not juvenile Bap60-KD MB. Comparisons were made between mature control and juvenile control and between mature Bap60-KD MB and juvenile Bap60-KD MB. Genes validated were: *Kn* and *Pvf1*. *Sosie* is validated despite variability of the data. Not validated is: *Salm*. Significance for RT-qPCR is determined by Student's t-test; significance for RNA-Seq is determined by binomial Wald test (DESeq2); *P<0.05, **P<0.01, ***P<0.001.

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 S3: Detailed phenotypes.

Table S4. *In silico* prediction tools for the variants.

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.

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	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>
Gender	Female	Male	Male	Male	Female
Post-natal - age at last measurements (yr)	7y.o.	11y 6m	6y 3m	3y	27 months
Developmental delay	+	+	+	+	+
Delayed walking	+	+	+	Not attained	-
Delayed speech	NA	+	+	Not attained	+
Intellectual disability	+	+	+	+	NA
Hypotonia	+	-	+	+	-
Feeding difficulties	+	+	+	+	+
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	+	-	+	-	-
Extremities	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	High palate, cleft soft palate		frontal bossing, long eyebrows	broad, square face, temporal narrowing	Prominent metopic suture Bulbous nasal tip

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

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

Table S3: Detailed phenotypes.

Individual	1	2	3	4	5
Mutation	NM_003076.4:c.990C>G, p.Asp330Glu	NM_003076.4 :c.11336A>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*
Microarray (aCGH)	NA	normal	normal	normal	normal
	Heterozygous in child, not mother, father not available	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>
Country of Origin and ethnicity	Japan, asian	France	UK, white	Northern Ireland, white	USA / European, German
Gender	Female	Male	Male	Male	Female
Pregnancy issues		Antenatal diagnosis of agenesis of the corpus callosum	Reduced fetal movements, scans suggested large head	Mild hydronephrosis on antenatal USS. Delivery: Emergency CS for failure to progress	IUGR and small head size noted in 3 rd trimester
Perinatal complications		Macrosomia at birth. Feeding difficulties with feeding tube during 15 days	very quiet baby, poor bowel movements with loose stools	Didn't cry at birth; stridor for several weeks	Mild jaundice at birth; gastroesophageal reflux as infant

duration gestation (weeks)	37w2d	38	40-41	40	39.4
-Birth Weight, g	2598	4150	3884	3480	2835 g
-Birth Length, cm	44,5	52	54cm	NA	48.9 cm
-Birth OFC, cm	32	37	NA	NA	31.75 cm
Post-natal - age at last measurements (yr)	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 (15 th %ile)
- OFC, cm	NA	59cm (+3.8 SD)	52cm (60th %ile)	49.3 cm (41 %ile)	44.3 cm (-2.4 SD)
Development					
Walking independently (months)	delay	17 months	2y	No	15 months
Speech - first words (months)		Delay, sentences	None, some sign language	No	18-20 months Still delayed but making progress including 2-3 word phrases
Feeding problems (no / yes;please specify age, duration, tube feeding needed or not)	Yes, feeding and sucking problems, congenital esophageal atresia, bronchial fistula.	Yes, 2 first weeks; tube feeding during 2 weeks. No problem after.	slow feeder, slow with milk and solids	yes - ongoing. NG-> fundoplication & PEG tube	Yes – hx of reflux in infancy and difficulty with regulating self-feeding; no need to tube feeding

Hypotonia (no / yes, congenital / yes, present for .. yrs)	Yes	no	yes, still evident but milder	yes - congenital, severe	No
Intellectual disability	Yes	yes	Yes	Severe	Too young to determine
details on development and IQ,		4 years 9 m : WPPSI Performance IQ 75, verbal IQ 50 Has special educational needs	Has a statement of special educational needs and attends a special needs school	Profound disability	Receiving special instruction, physical and speech therapy No formal IQ testing at this time
- other diagnosis such as ADHD or Autism		Behavioral disorders, frustration intolerance	No formal diagnosis of autism but has obsessive traits.	No	No
- unusual anxieties			Anxiety of new things and new situations and large places eg shopping centres	NA	Very anxious with health care providers
- medications		No	Domperidone, lazoparole, antibiotics	Omeprazole	none
Seizures (if yes, nature, frequency)	No	no	none	No	none

Neuroradiology (not performed/ yes CT +age/ yes MRI +age)		MRI age 1 month : complete agenesis of the corpus callosum	yes, no abnormality seen	MRI age 3: Prominence of the lateral and third ventricles. Cavum septum pellucidum. 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.	MRI at age 22months: Nonspecific small right frontal subcortical white matter FLAIR hyperintensity. Otherwise unremarkable brain MR without contrast.
Hearing loss (n; yes unilat; yes, bilat)	Yes, malformation of ossicle	no	mild bilateral	N	N
Vision impaired (no; yes unilat; yes, bilat)	No	no	Yes, plus nystagmus	alternating divergent squint	No – mild astigmatism not requiring

					glasses/intervention
Hair	Sparse	Low hairline	low hairline	temporal deficiency	
Forehead			frontal bossing	broad, square face, temporal narrowing	Prominent metopic suture
Nose				flat nasal bridge, sl short nose, very prominent philtral pillars	Bulbous nasal tip
Eyes			long eyebrows	hypertelorism, divergent squint	
Mouth including gums (eg hypertrophy), palate	High palate, cleft soft palate			wide mouth, downturned corners, thick gums, high palate, very "jowly" appearance; myopathic	Bifid uvula, normal palate
Ears	external ear malformation	Upturned, thick ear lobes	upturned earlobes	low-set, thickened, simple shape	normal
<u>Trunk and limbs</u>	Hypoplastic fifth toenails	Short hands, slender fingers	small hands and feet	Thick loose palmar & plantar skin, small hands & feet, short 5th fingers	none

Gastrointestinal abnormalities		no	gastroesophageal reflux, needing medication. 'disturbed bowels' very frequent stools, getting worse	reflux+++; fundoplication & PEG	Gastroesophageal reflux, now improved
Teeth anomalies		no	small rounded teeth and pointed canines	thick, stubby teeth, delayed eruption, thick gums	None
Frequent infections		no	frequent ear infections, tonsillitis. Necessitating hospitalisations	Yes; respiratory as baby	Frequent ear infection s/sp BMT
Surgeries		no	ENT- adenoids, tonsils removed and grommets inserted.	PEG, surgery for undesc testes	BMT for frequent ear infections
Further information			oral dyspraxia, ankyloglossia (operated), plagiocephaly	elevated T3 as baby; normalised frequent UTIs; undescended testes, mottled skin on legs & arms	Facial asymmetry when crying, Prominent metopic suture, Gastroesophageal reflux, now improved

Table S4. *In silico* prediction tools for the variants.

	Conservation scores										Functional		General			
	Polyphen2 HDIV	Polyphen2 HVAR	LRT	MutationAssessor	GERP++ RS rankscore	phyloP100way vertebrate rankscore	phyloP20way mammalian rankscore	phastCons100way vertebrate rankscore	phastCons20way mammalian rankscore	SiPhy_29way logOdds rankscore	SIFT	PROVEAN	FATHMM-MKL coding	Mutation Taster	DANN rankscore	CADD phred
p.Asp330Glu	D	D	D	M	0.393	0.26	0.165	0.424	0.75	0.537	D	D	D	D	0.816	26.8
p.Arg446Gly	B	B	N	M	0.173	0.389	0.044	0.715	0.625	0.36	D	D	D	D	0.698	20.9
p.Trp486*	.	.	D	.	0.693	0.997	0.605	0.715	0.697	0.934	.	.	D	D	0.559	39
p.Phe495Leu	B	B	D	M	0.693	0.886	0.626	0.715	0.888	0.744	T	D	D	D	0.182	14.24
p.Arg503*	.	.	D	.	0.6	0.502	0.391	0.715	0.75	0.684	.	.	D	D	0.732	41

D, deleterious or damaging; B, benign; M, medium; N, neutral; T, tolerated.

Shaded are the damaging predictions, those with rank scores above 0.5, and CADD phred scores above 25.