Genetic and Genomic Analyses of *Drosophila melanogaster* Models of Chromatin Modification Disorders

Rebecca A. MacPherson, Vijay Shankar, Robert R. H. Anholt* and Trudy F. C. Mackay*

Center for Human Genetics and Department of Genetics and Biochemistry, Clemson University, 114 Gregor Mendel Circle, Greenwood, SC 29646, USA

*RRHA and TFCM are co-corresponding authors. Address correspondence to Robert R. H. Anholt or Trudy F. C. Mackay

e-mail: ranholt@clemson.edu, tmackay@clemson.edu

Keywords: Coffin-Siris syndrome, Nicolaides-Baraitser syndrome, SWI/SNF-related intellectual disability disorders, Cornelia de Lange syndrome, RNAi, RNA sequencing

Running Head: Fly models of chromatin modification disorders

ABSTRACT

Switch/Sucrose Non-Fermentable (SWI/SNF)-related intellectual disability disorders (SSRIDDs) and Cornelia de Lange syndrome are rare syndromic neurodevelopmental disorders with overlapping clinical phenotypes. SSRIDDs are associated with the BAF (*Brahma-Related Gene-1* Associated Factor) complex, whereas CdLS is a disorder of chromatin modification associated with the cohesin complex. Here, we used RNA interference in *Drosophila melanogaster* to reduce expression of six genes (*brm*, *osa*, *Snr1*, *SMC1*, *SMC3*, *vtd*) orthologous to human genes associated with SSRIDDs and CdLS. These fly models exhibit changes in sleep, activity, startle behavior (a proxy for sensorimotor integration) and brain morphology. Whole genome RNA sequencing identified 9,657 differentially expressed genes (FDR < 0.05), 156 of which are differentially expressed in both sexes in SSRIDD- and CdLS-specific analyses, including *Bap60,* which is orthologous to *SMARCD1*, a SSRIDD-associated BAF component. k-means clustering reveals genes co-regulated within and across SSRIDD and CdLS fly models. RNAi-mediated reduction of expression of six genes co-regulated with focal genes *brm*, *osa*, and/or *Snr1* recapitulated changes in behavior of the focal genes. Based on the assumption that fundamental biological processes are evolutionarily conserved, Drosophila models can be used to understand underlying molecular effects of variants in chromatin-modification pathways and may aid in discovery of drugs that ameliorate deleterious phenotypic effects.

INTRODUCTION

Switch/Sucrose Non-Fermenting (SWI/SNF)-related intellectual disability disorders (SSRIDDs) and Cornelia de Lange syndrome (CdLS) are syndromic neurodevelopmental Mendelian disorders of chromatin modification. SSRIDDs, including Coffin-Siris syndrome (CSS) and Nicolaides-Baraitser syndrome (NCBRS), stem from variants in genes of the *Brahma-Related Gene-1* Associated Factor (BAF) complex, also known as the mammalian SWI/SNF complex (Hoyer *et al*. 2012; Santen *et al*. 2012;

Tsurusaki *et al*. 2012; Van Houdt *et al*. 2012; Tsurusaki *et al*. 2014; Hempel *et al*. 2016; Bramswig *et al*. 2017; Bogershausen and Wollnik 2018; Vasileiou *et al*. 2018; Gazdagh *et al*. 2019; Machol *et al*. 2019; Zawerton *et al*. 2019). CdLS is associated with variants in genes that encode components of the cohesin complex (Krantz *et al*. 2004; Deardorff *et al*. 2007; Deardorff *et al*. 2012; Gil-Rodriguez *et al*. 2015; Boyle *et al*. 2017; Huisman *et al*. 2017; Olley *et al*. 2018).

SSRIDD patients exhibit neurodevelopmental delay, intellectual disability, hypotonia, seizures, and sparse hair growth, as well as cardiac, digit, and craniofacial anomalies, where the severity and spectrum of affected phenotypes are dependent upon the specific variant or affected gene product (reviewed in Bogershausen and Wollnik 2018; Schrier Vergano *et al*. 2021; Vasko *et al*. 2021). For example, many SSRIDD patients with variants in *ARID1B* tend to have milder phenotypes including normal growth, milder facial gestalt, and no central nervous system (CNS) abnormalities, whereas most variants in *SMARCB1* are associated with more severe phenotypes, including profoundly delayed developmental milestones, seizures, kidney malformations, and CNS abnormalities (Bogershausen and Wollnik, 2018; Schrier Vergano *et al*. 2021). Furthermore, variants in *ARID1B* are associated with SSRIDD, Autism Spectrum disorder, and non-syndromic intellectual disability (Hoyer *et al*. 2012; De Rubeis *et al*. 2014; Iossifov *et al*. 2014; Vissers *et al*. 2016; van der Sluijs *et al*. 2019). Brain malformations, such as agenesis of the corpus callosum, Dandy-Walker malformation, and cerebellar hypoplasia, have also been observed in 20-30% of all patients with variants in the BAF complex (Vasko *et al*. 2022), but are most commonly observed in patients with variants in *SMARCB1* (Bogershausen and Wollnik 2018).

CdLS patients also display a clinical spectrum including intellectual disability, hirsutism, synophrys, and digit, craniofacial, and CNS anomalies (reviewed in Kline *et al*. 2018; Avagliano *et al*. 2020; Selicorni *et al*. 2021). As in SSRIDDs, some phenotypes are more highly associated with a specific gene, but phenotypic

severity can vary widely across variants within the same gene. For example, most patients with variants in *SMC1A* show milder developmental delay and intellectual disability compared to their classical *NIPBL*-CdLS counterparts, but about 40% of *SMC1A* patients exhibit severe epileptic encephalopathy and intellectual disability (Jansen *et al*. 2016; Symonds *et al*. 2017; Selicorni *et al.* 2021).CdLS has also been reclassified as a spectrum of cohesinopathies (Van Allen *et al*. 1993; Kline *et al*. 2018). Patients with pathogenic variants in many genes involved in chromatin accessibility and regulation have overlapping symptoms with CdLS (Parenti et al., 2017; Aoi *et al*. 2019; Cucco *et al*. 2020).

D. melanogaster is well-suited for modeling human disorders, as large numbers of flies can be raised economically without ethical or regulatory restrictions. Additionally, SSRIDD- and CdLS-associated genes are highly conserved in flies and a wide variety of genetic tools are available to create fly models of human diseases (Hu *et al*. 2011; Perkins *et al*. 2015; Zirin *et al*. 2020). Previous groups have used *D. melanogaster* to investigate SSRIDDs and CdLS and have observed phenotypes relevant to disease presentation in humans, including changes in sleep, brain function, and brain morphology (Pauli *et al*. 2008; Schuldiner *et al*. 2008; Wu *et al*. 2015; Chubak *et al*. 2019). These studies have provided insight into potential disease pathogenesis and suggested that certain subtypes of SSRIDD and CdLS can be modeled in the fly, but they were not performed in controlled genetic backgrounds.

Here, we present behavioral and transcriptomic data on Drosophila models of SSRIDDs and CdLS in a common genetic background. RNAi-mediated knockdown of Drosophila orthologs of SSRIDD- and CdLS-associated genes show gene- and sex-specific changes in brain structure and sensorimotor integration, as well as increased locomotor activity and decreased night sleep. Transcriptomic analyses show distinct differential gene expression profiles for each focal gene.

METHODS

Drosophila Genes and Stocks

We selected SSRIDD-, and CdLS-associated genes with a strong fly ortholog (*Drosophila* RNAi Screening Center Integrative Ortholog Prediction Tool (DIOPT) score > 9) (Hu *et al*. 2011) and a corresponding *attp2* fly line available from the Transgenic RNAi Project (TRiP) (Perkins *et al*. 2015; Zirin *et al*. 2020). We excluded human genes that were orthologous to multiple fly genes to increase the likelihood of aberrant phenotypes upon knockdown of a single fly ortholog. We used *attp40* TRiP lines when assessing phenotypes associated with knockdown of co-regulated genes. We used the y^{1}, sc^{*}, v^{1}, sev^{21}; *TRiP2; TRiP3* genotype as the control *UAS* line in all experiments. With the exception of the initial viability screen, we crossed all RNAi lines to a weak ubiquitous *GAL4* driver line, *Ubi156-GAL4* (Garlapow *et al*. 2015). Table S1A lists the Drosophila stocks used.

Drosophila Culture

For all experiments, we maintained flies at a controlled density on standard cornmeal/molasses medium (Genesee Scientific, El Cajon, CA) supplemented with yeast in controlled environmental conditions (25˚C, 50% relative humidity, 12-hour light-dark cycle (lights on at 6 am)). Crosses contained five flies of each sex, with fresh food every 48 hours. After eclosion, we aged flies in mixed-sex vials at a density of 20 flies per vial until used in experiments. We performed experiments on 3-5-day old flies from 8 am to 11 am, unless otherwise noted.

Viability

For the initial viability screen of Drosophila orthologs of SSRIDD- and CdLS-associated genes, we crossed *attp2* TRiP lines and the control line to three ubiquitous *GAL4* driver lines. For the viability screen of co-regulated genes, we crossed *attp40* TRiP lines and the control line to the *Ubi156-Gal4* driver line. From

115 days 0-15, we noted the developmental stage. For stocks that contained balancers, we recorded the
116 associated phenotypic marker in eclosed progeny.
117
118 **Quantitative Real-Time PCR (qRT-PCR)**
119 For the qRT-PCR analyses of gene expression of RNAi targets of *brm*, *osa*, *SMC1*, *SMC3*, *Snr1*, and *vtd*, we
120 flash froze 3-5-day old whole flies on dry ice and then collected, sexes separately, 30 flies per sample.
121 We stored frozen flies and their extracted RNA at -80˚C. We extracted RNA using the Qiagen RNeasy Plus
122 Mini Kit (Qiagen, Hilden, Germany) by homogenizing tissue with 350µL of RLT Plus Buffer containing β-
123 mercaptoethanol (Qiagen) and DX reagent (Qiagen), using a bead mill at 5m/second for 2 minutes. We
124 quantified RNA with the Qubit RNA BR Assay Kit (ThermoFisher Scientific, Waltham, MA) on a Qubit
125 Fluorometer (ThermoFisher Scientific) according to the manufacturer's specifications. We synthesized
126 cDNA using iScript Reverse Transcription Supermix (Bio-Rad Laboratories, Inc., Hercules, CA) according
127 to the manufacturer's instructions. We quantified expression using quantitative real-time PCR with
128 SYBR™ Green PCR Master Mix (ThermoFisher Scientific), according to manufacturer specifications, but
129 with a total reaction volume of 20µL. We used three biological and three technical replicates per sample
130 and calculated percent knockdown using the ΔΔ*ct* method (Livak and Schmittgen 2001). Table S1B
131 contains primer sequences used. For the qRT-PCR analyses of gene expression for the co-regulated
132 genes *Alp10*, *CG40485*, *CG5877*, *IntS12*, *Mal-A4*, and *Odc1*, we extracted RNA using the Direct-zol RNA
133 MiniPrep Plus Kit (Zymo Research, Irvine, CA) and homogenized tissue with 350µL of Tri-Reagent, using a
134 bead mill at 5m/second for 2 minutes. We used two technical replicates in the qRT-PCR analyses of co-
135 regulated genes.
136
137 **Startle-Induced Locomotor Response**

We assessed startle response using a variation of a previously described assay (Yamamoto *et al*. 2008). In summary, 36-50 flies per sex per line were placed into individual vials to acclimate 24 hours prior to testing. To standardize the mechanical startle stimulus, we placed a vial housing a single 3-5-day old fly in a chute. Removal of a supporting dowel allows the vial to drop from a height of 42 cm, after which it comes to rest horizontally (Huggett *et al*. 2021). We measured the total time the fly spent moving during a period of 45 s immediately following the drop. We also recorded whether the fly demonstrated a tapping phenotype, a series of leg extensions without forward movement. Time spent tapping was not considered movement for startle calculations.

Sleep and Activity

We used the Drosophila Activity Monitoring System (DAM System, TriKinetics, Waltham, MA) to assess sleep and activity phenotypes. At 1-2 days of age, we placed flies into DAM tubes containing 2% agar with 5% sucrose, sealed with a rubber cap (TriKinetics) and a small piece of yarn. We collected data for 7 days on a 12-hour light-dark cycle, with sleep defined as at least 5 minutes of inactivity. We discarded data from flies that did not survive the entire testing period, leaving 18-32 flies per sex per line for analysis. We processed the raw sleep and activity data using ShinyR-DAM (Cichewicz and Hirsh 2018) and used the resulting output data for statistical analysis.

Dissection and Staining of Brains

We dissected brains from cold-anesthetized flies in cold phosphate buffered saline (PBS), before we fixed the brains with 4% paraformaldehyde (v/v in PBS) for 15 minutes, washed with PAXD buffer (1x PBS, 0.24% (v/v) Triton-X 100, 0.24% (m/v) sodium deoxycholate, and 5% (m/v) bovine serum albumin) three times for 10 minutes each, and then washed three times with PBS. We blocked fixed brains with 5% Normal Goat Serum (ThermoFisher Scientific; in PAXD) for 1 hour with gentle agitation, then stained

with 2-5 µg/mL of Mouse anti-Drosophila 1D4 anti-Fasciclin II (1:4) (Developmental Studies Hybridoma Bank; Iowa City, IA) for 16-20 hours at 4˚C. We washed brains three times with PAXD for 10 minutes and stained them with Goat anti-Mouse IgG-AlexaFluor488 (1:100) (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA) for 4 hours. Then, we washed brains with PAXD three times for 10 minutes each prior to mounting with ProLong Gold (ThermoFisher Scientific). We performed all steps at room temperature with gentle agitation during incubations.

Brain Measurements

We analyzed 17-20 brains per sex per line using a Leica TCS SPE confocal microscope. We visualized Z-stacks of each brain using Icy v. 2.2.0.0 (de Chaumont *et al*. 2012).

We measured ellipsoid body height and ellipsoid body width by measuring vertical ellipsoid body length from dorsal to ventral, and horizontal ellipsoid body length from left to right (relative to the fly). We also measured lengths of the mushroom body alpha and beta lobes by drawing a single 3D line (3DPolyLine Tool within Icy) through the center of each lobe, adjusting the position of the line while progressing through the z-stack. We measured alpha lobes from the dorsal end of the alpha lobe to the alpha/beta lobe heel (where the alpha and beta lobes overlap) and beta lobes from the median end of the beta lobe to the alpha/beta lobe heel. We normalized the measurements for each brain using the distance between the left and right heels of the mushroom body (heel-heel distance). We used the average alpha and beta lobe lengths for each brain for subsequent analyses. In the case of one missing alpha or beta lobe, we did not calculate an average and instead, used the length of the remaining lobe for analysis. If both alpha or both beta lobes were missing, we removed that brain for analysis of the missing lobes, but retained it for analysis of the other brain regions.

We also recorded gross morphological abnormalities of the mushroom body alpha and beta lobes, including missing lobe, skinny lobe, extra projections, abnormal alpha lobe outgrowth, and beta lobes crossing the midline for each brain. We selected these phenotypes based on prior studies on gross mushroom body morphology (Zwarts *et al*. 2015; Chubak *et al*. 2019).

Statistical Analyses

Unless noted below, we analyzed all behavioral data and brain morphology data in SAS v3.8 (SAS Institute, Cary, NC) using the "PROC GLM" command according to the Type III fixed effects factorial ANOVA model $Y = \mu + L + S + L{\times}S + \varepsilon$, where Y is the phenotype, μ is the true mean, L is the effect of line (*e.g.* RNAi line versus the control), S is the effect of sex (males, females), and ε is residual error. We performed comparisons between an RNAi line and its control. We also performed additional analyses for each sex separately.

We used a Fisher's Exact test (*fisher.test* in R v3.63) to analyze the proportion of flies tapping during startle experiments, the number of brains with a specific morphological abnormality, and the number of brains with any gross morphological abnormality.

We performed Levene's and Brown-Forsythe's Tests for unequal variances on the same data set used for the analysis of lobe lengths. For both tests, we used the *leveneTest* command ((*car* v3.0-11, Fox and Sanford 2019) in R v3.6.3) to run a global analysis comparing all genotypes as well as pairwise comparisons.

RNA Sequencing

We synthesized libraries from 100ng of total RNA using the Universal RNA-seq with Nuquant + UDI kit (Tecan Genomics, Inc., CA) according to manufacturer recommendations. We converted RNA into cDNA using the integrated DNase treatment and used the Covaris ME220 Focused-ultrasonicator (Covaris, Woburn, MA) to generate 350bp fragments. We performed ribosomal RNA depletion and bead selection using Drosophila AnyDeplete probes and RNAClean XP beads (Beckman Coulter, Brea, CA), respectively. We purified libraries after 17 cycles of PCR amplification. We measured library fragment sizes on the Agilent Tapestation using the Agilent High Sensitivity DNA 1000 kit (Agilent Technologies) and quantified library concentration using the Qubit 1X dsDNA High Sensitivity Assay kit (Thermo Fisher Scientific). We pooled libraries at 4nM and loaded them onto an Illumina S1 flow cell (Illumina, Inc., San Diego, CA) for paired-end sequencing on a NovaSeq6000 (Illumina, Inc., San Diego, CA). We sequenced three biological replicates of pools of 30 flies each per sex per line. We sequenced each sample to a depth of ~30 million total reads; we resequenced samples with low read depth (<8 million uniquely mapped reads).

We used the default Illumina BaseSpace NovaSeq sequencing pipeline to demultiplex the barcoded sequencing reads. We then merged S1 flow cell lanes, as well as reads from different runs. We filtered out short and low-quality reads using the *AfterQC* pipeline (v0.9.7) (Chen *et al*. 2017) and quantified remaining levels of rRNA via the bbduk command (Bushnell 2014). We aligned reads to the reference genome (*D. melanogaster* v6.13) using GMAP-GSNAP (Wu *et al*. 2016) and counted these unique alignments to Drosophila genes using the featurecounts pipeline from the Subread package (Liao *et al*. 2013). We excluded genes with a median expression across all samples of less than 3 and genes where greater than 25% of the samples had a counts value of 0. We then normalized the data based on gene length and library size using GeTMM (Smid *et al*. 2018) prior to differential expression analysis.

Differential Expression Analyses

We performed multiple analyses for differential expression in SAS (v3.8; Cary, NC) using the "PROC glm" command. We first performed a fixed effects factorial ANOVA model $Y = \mu + L + S + L{\times}S + \varepsilon$, where Line ($L$, all RNAi and control genotypes) and Sex (S) are cross-classified main effects and Line×Sex ($L{\times}S$) is the interaction term, Y is gene expression, μ is the overall mean, and ε is residual error. We then performed the same analyses only for genes associated with SSRIDDs or for CdLS; i.e., 9,657 genes that were significantly differentially expressed (FDR < 0.05 for the Line and/or Line×Sex terms) in the full model. We ran the ANOVA model for each RNAi genotype compared to the control. Finally, we ran ANOVAs ($Y = \mu + L + \varepsilon$) separately for males and females for the disease-specific and individual RNAi analyses.

Gene Ontology and k-means Clustering Analyses

We performed Gene Ontology (GO) statistical overrepresentation analyses on the top 1,000 differentially expressed genes for the Line term (GO Ontology database released 2022-03-22, Pantherdb v16.0 (Mi *et al.* 2013; Thomas *et al.* 2022)) in each disease-specific and pairwise analysis for GO Biological Process, Molecular Function, and Reactome Pathway terms. For the analyses performed on sexes separately, we used the top 600 differentially expressed genes based on the significance of the Line term. The numbers of differentially expressed genes used in GO enrichment gave maximal GO enrichment with minimal redundancy compared to other numbers of differentially expressed genes.

We performed k-means clustering (average linkage algorithm), sexes separately, on Ge-TMM normalized least squares means of 533 genes that had the highest Log2 fold change (FC) in expression. We identified the cutoff threshold value for Log2FC by first sorting genes in a descending order of maximal absolute value of Log2FC, then fitted lines to roughly linear segments of the generated distribution and designated the cutoff threshold as the Log2FC value of the index at the intersection of the two fitted lines. We used hierarchical clustering (Average Linkage algorithm, WPGMA) to determine the

approximate number of natural clusters, then performed clustering with varying values of k to determine the largest number of unique, but not redundant, expression patterns. We also performed GO statistical overrepresentation analyses on genes in each k-means cluster (GO Ontology database released 2022-07-01, Pantherdb v17.0 (Mi *et al.* 2013; Thomas *et al.* 2022)) in each disease-specific and pairwise analysis for GO Biological Process, Molecular Function, and Reactome Pathway terms.

RESULTS

Drosophila Models of SSRIDDs and CdLS

We identified Drosophila orthologs of 12 human genes associated with the SSRIDD chromatin remodeling disorders and CdLS with a DIOPT score > 9 and for which TRiP RNAi lines in a common genetic background and without predicted off-target effects were publicly available. Using these criteria, the Drosophila genes *Bap111*, *brm*, *osa*, and *Snr1* are models of SSRIDD-associated genes *ARID1A*, *ARID1B*, *SMARCA2, SMARCA4*, *SMARCB1*, and *SMARCE1*; and *Nipped-B*, *SMC1*, *SMC3*, and *vtd* are models of CdLS-associated genes *NIPBL*, *SMC1A*, *SMC3*, and *RAD21* (Table S2).

We obtained *UAS*-RNAi lines generated in the same genetic background for each of the fly orthologs and crossed these RNAi lines to each of three ubiquitous *GAL4* drivers to assess viability (Figure S1). We selected ubiquitous drivers since the human SSRIDD- and CdLS-associated genes and Drosophila orthologs are ubiquitously expressed, and SSRIDD and CdLS patients carry pathogenic variants in all cells. We initially crossed each *UAS*-RNAi line to three ubiquitous *GAL4* drivers (*Actin-GAL4*, *Ubiquitin-GAL4*, and *Ubi156-GAL4*) and assessed viability and degree of gene knockdown in the F1 progeny (Figure S1). *Ubiquitin-GAL4*-mediated gene knockdown resulted in viable progeny in only three of the eleven *UAS*-RNAi lines, with most progeny dying during the embryonic or larval stage (Figure S1). Based on

these data, we selected the weak ubiquitin driver *Ubi156-GAL4* (Garlapow *et al*. 2015) and the *UAS-RNAi* lines for *brm*, *osa*, *Snr1*, *SMC1*, *SMC3*, and *vtd* for further study (Table 1). With the exception of *Ubi156>osa* males which had ~15% gene knockdown, RNAi knockdown of all genes ranged from 40-80% (Table S3). Given that SSRIDDs and CdLS are largely autosomal dominant disorders, knockdown models that retain some degree of gene expression are reflective of the genetic landscape of SSRIDD and CdLS patients.

Effects on Startle Response

Given the neurological and musculoskeletal clinical findings in SSRIDD, and CdLS patients (Bogershausen and Wollnik 2018; Kline *et al*. 2018; Avagliano *et al*. 2020; Schrier Vergano *et al*. 2021; Selicorni *et al*. 2021; Vasko *et al*. 2022), we assessed startle-induced sensorimotor integration for RNAi of *brm*, *osa*, *Snr1*, *SMC1*, *SMC3*, and *vtd* relative to their control genotype. Almost all genotypes exhibited a decreased startle response across both sexes ($p < 0.02$ for all by-sex by-genotype comparisons to the control, Figure 1A, Table S4). Males with *osa* or *brm* knockdown did not exhibit changes in startle response ($p > 0.05$), and females with *Snr1* knockdown showed an increased startle response ($p < 0.0001$). In the lines where both sexes were affected, we observed more extreme phenotypes in males (Figure 1A).

While testing flies for startle response, we noticed that some flies exhibited a specific locomotion phenotype we termed “tapping”. Tapping is characterized by repetitive extension and retraction of individual legs as if to walk, but without progressive movement in any direction (File S1). Compared to the control (example shown in File S2), we observed an increase in the number of flies exhibiting tapping behavior in male flies with knockdown of *brm* ($p = 0.0267$), *osa* ($p = 0.0026$), *Snr1* ($p = 0.0005$) and *vtd* ($p = 0.0002$) (Figure 1B, Table S4). We also observed increases in tapping behavior in females with

304 knockdown of *Snr1* and *vtd* that fall just outside of a significance level of 0.05 (p = 0.0563 for both
305 genes); Figure 1B, Table S4). The tapping and startle phenotypes were not evident across all genes
306 associated with a specific disorder.
307
308 **Effects on Sleep and Activity**
309 We hypothesized that hypotonia and sleep disturbances observed in SSRIDD and CdLS patients (Liu and
310 Krantz 2009; Stavinoha *et al*. 2011; Rajan *et al*. 2012; Zambrelli *et al*. 2016; Bogenshausen and Wollnik
311 2018; Schrier Vergano *et al*. 2021; Vasko *et al*. 2021) may correspond to changes in activity and sleep in
312 Drosophila models. Sleep disturbances were also observed in a previous Drosophila model of *NIPBL*-
313 CdLS (Wu *et al*. 2015). Therefore, we quantified activity and sleep phenotypes for RNAi-mediated
314 knockdown of *brm*, *osa*, *Snr1*, *SMC1*, *SMC3*, and *vtd*. All RNAi genotypes showed increases in overall
315 spontaneous locomotor activity (p < 0.02 for all by-sex by-genotype comparisons to the control, Figure
316 2A, Table S4). This increase in spontaneous locomotor activity was most pronounced in males with
317 knockdown of *osa* (p < 0.0001); this was the only genotype for which males were more active than
318 females (Figure 2A, Table S4). All RNAi genotypes showed decreases in night sleep (p < 0.0001 for all by-
319 sex by-genotype comparisons to the control). Flies with knockdown of *osa* (males, p < 0.0001; females, p
320 < 0.0001) and females with knockdown of *vtd* (p < 0.0001) spent about half of the nighttime awake, the
321 least amount of sleep across all flies tested (Figure 2B, Table S4). In addition to increased activity, the
322 Drosophila models of SSRIDDs and CdLS have fragmented sleep: the number of sleep bouts at night was
323 increased for all lines and sexes compared to the control (p < 0.0001 for all by-sex by-genotype
324 comparisons to the control, except *SMC1* males, p = 0.0023, Figure 2C, Table S4).
325
326 **Effects on Brain Morphology**

To assess changes in brain structure in *brm*, *osa*, *Snr1*, *SMC1*, *SMC3*, and *vtd* RNAi genotypes, we focused on the mushroom body and the ellipsoid body, as prior studies on SSRIDDs in flies have shown changes in mushroom body structure (Chubak *et al.* 2019), and the mushroom body has been linked with regulation of sleep and activity in Drosophila (Joiner *et al.* 2006; Pitman *et al.* 2006; Guo *et al.* 2011; Sitaraman *et al.* 2015). Furthermore, SSRIDD and CdLS patients often present with intellectual disability and CNS abnormalities (Bogershausen and Wollnik 2018; Kline *et al.* 2018; Avagliano *et al.* 2020; Schrier Vergano *et al.* 2021; Selicorni *et al.* 2021; Vasko *et al.* 2022). In the Drosophila brain, the mushroom body mediates experience-dependent modulation of behavior (reviewed in Modi *et al.* 2020), making the mushroom body and the ellipsoid body, which mediates sensory integration with locomotor activity, suitable targets for examining changes in brain structure. We used confocal microscopy to quantify the lengths of both alpha and beta lobes of the mushroom body, as well as the horizontal and vertical lengths of the ellipsoid body (Figures 3A-B). The lengths of these lobes were measured in three dimensions, capturing the natural curvature of the alpha and beta lobes of the mushroom body instead of relying upon a 2D measurement of a 3D object.

We observed sex-specific changes in brain morphology (Figure 3C-D). Females, but not males, showed decreased ellipsoid body dimensions with knockdown of *Snr1* (horizontal, $p = 0.0002$; vertical, $p < 0.0444$, Table S4), while knockdown of *vtd* in females showed decreased alpha ($p = 0.0088$) and beta ($p = 0.0433$) lobe lengths. In addition to sex-specific effects, we observed sexually dimorphic effects; females with knockdown of *brm* showed decreases in alpha lobe and horizontal ellipsoid body length ($p = 0.0409$, $p = 0.0224$, respectively), while *brm* knockdown males showed increases in alpha lobe and horizontal ellipsoid body length ($p = 0.0301$, $p = 0.0305$, respectively; Figure 4, Table S4). Levene's tests for equality of variances indicate that the ellipsoid body measurements have sex-specific unequal environmental variances in some genotypes compared to the control (Figure 4, Table S4). These results

351 show that these models of SSRIDDs and CdLS show morphological changes in the mushroom body and
352 ellipsoid body.
353
354 We also recorded gross morphological abnormalities, such as missing lobes, beta lobes crossing the
355 midline, and impaired/abnormal alpha lobe outgrowth (Figure 3C-D). Although each abnormality was
356 observed across multiple genotypes, only flies with knockdown of *osa* demonstrated consistent brain
357 abnormalities. Male and female *osa* knockdown flies both exhibited an increased number of alpha lobes
358 with impaired outgrowth (males: $p < 0.0001$, females: $p < 0.0025$, Figure 4E, Table S4), and the *osa*
359 knockdown males also showed a significant number of beta lobe midline defects ($p = 0.0471$, Figure 4F,
360 Table S4). Males with knockdown of *SMC1* and *vtd* also showed increased numbers of abnormal brains
361 ($p = 0.0471$, $p = 0.0202$ respectively; Figure 4G, Table S4). Changes in brain morphology are more gene-
362 and sex-dependent than changes in sleep, activity, and startle response.
363
364 **Effects on Genome-wide Gene Expression**
365 We performed genome-wide analysis of gene expression for the *brm*, *osa*, *Snr1*, *SMC1*, *SMC3*, and *vtd*
366 RNAi genotypes and their control, separately for males and females. We first performed a factorial fixed
367 effects analysis of variance (ANOVA) for each expressed transcript, partitioning variance in gene
368 expression between sexes, lines, and the line by sex interaction for all seven genotypes. We found that
369 8,481 and 6,490 genes were differentially expressed (FDR < 0.05 for the Line and/or Line×Sex terms,
370 Table S5), for a total of 9,657 unique genes.
371
372 *brm*, *osa*, *Snr1* and their human orthologs (Tables 1, S2) are part of the same protein complex (BAF
373 complex in humans, BAP-complex in flies). Therefore, we evaluated whether other BAP complex
374 members *Bap55*, *Bap60*, and *Bap111* (which are orthologous to human BAF complex members *ACTL6A*,

375 *SMARCD1*, and *SMARCE1*, respectively), are differentially expressed in the analysis of all genes. We
376 observed differential expression of strong fly orthologs (DIOPT > 9) of additional BAF complex subunits
377 in the global model and found that *Bap55* and *Bap60* (FDR-corrected Line *p*-values: 0.0123, 0.01306,
378 respectively; Table S5), but not *Bap111*, are differentially expressed. We did not observe differential
379 expression of *Nipped-B* in the global analysis. *Nipped-B* is a member of the fly cohesin complex along
380 with *SMC1*, *SMC3*, and *vtd,* and is orthologous to the human cohesin complex member *NIPBL*.
381
382 We next performed separate pairwise analyses for SSRIDD-associated fly orthologs and CdLS-associated
383 fly orthologs against the control genotype using the subset of 9,657 unique differentially expressed
384 genes from the full ANOVA model (Tables 2, S5). We also performed these analyses on sexes separately
385 (Tables 2, S5). The number of differentially expressed genes at a given FDR threshold varies across
386 pairwise comparisons and across sexes. For example, females with knockdown of *brm* and *Snr1* have
387 583 and 3,026 differentially expressed genes (FDR < 0.05), respectively, whereas males with knockdown
388 of these genes have 2,996 and 3,376 differentially expressed genes (FDR < 0.05), respectively (Tables 2,
389 S5). We observed the largest number of differentially expressed genes in flies with knockdown of *Snr1*
390 (Tables 2, S5). At FDR < 0.0005, there were still 1,059 genes differentially expressed in *Snr1* males (Table
391 S5). A greater number of differentially expressed genes are upregulated than downregulated in flies with
392 knockdown of *brm*, *SMC1*, *SMC3*, and *vtd* (Table S5). In contrast, flies with knockdown of *osa* and *Snr1*
393 have a greater number of downregulated genes (Table S5). Flies with knockdown of *Snr1* and *SMC1* had
394 the greatest percentage of differentially expressed genes shared between males and females: 12.2%
395 (698) and 7.6% (348) respectively (Table S6). *Snr1* also had the greatest percent knockdown by RNAi.
396 Only four genes are differentially expressed in all pairwise comparisons of knockdown lines versus the
397 control line, in both males and females; all are computationally predicted genes (Table S6).
398

We performed k-means clustering to examine patterns of co-regulated expression, separately for males (k=8) and females (k=10). We identified the cutoff threshold value for Log2FC by first sorting genes in a descending order of maximal absolute value of Log2FC (Table S7). We fitted lines to roughly linear segments of the generated distribution and designated the cutoff threshold as the Log2FC value of the index at the intersection of the two fitted lines (Figure S2, Table S7). The genes in each cluster are listed in Table S8. Although many clusters reveal gene-specific expression patterns (*e.g.* Cluster F1, F9, F10, Figure 5; Clusters M1, M6, Figure 6), Clusters F7 and F8 show disease-specific patterns, where knockdown of *brm, osa,* and *Snr1* clusters separately from *SMC1*, *SMC3*, and *vtd* (Figure 5). This is not surprising, as *brm, osa,* and *Snr1* are part of the fly BAF complex and models for SSRIDDs, whereas *SMC1*, *SMC3*, and *vtd* are associated with the fly cohesin complex and are models for CdLS. We also observed patterns involving genes from both SSRIDDs and CdLS. Clusters F4 and M3 contain genes upregulated in response to knockdown of *SMC3, osa,* and *brm* and downregulated in response to knockdown of *Snr1* and *SMC1* (Figures 5-6) Clusters F5 and M5 contain genes upregulated only in flies with knockdown of *osa* and *Snr1* (Figures 5-6). Notably, many long noncoding RNAs (lncRNAs) feature prominently in many of the male and female clusters (Figures 5-6; Tables S7, S8).

To infer functions of these differentially expressed genes, we performed Gene Ontology (GO) analyses on the top approximately 600 (1000) differentially expressed genes for sexes separately (sexes pooled) (Table S9). These analyses reveal that differentially expressed genes associated with knockdown of CdLS-associated fly orthologs are involved in chromatin organization, regulation and processing of RNA, reproduction and mating behavior, peptidyl amino acid modification, and oxidoreductase activity (Table S9). We also see sex-specific effects, such as muscle cell development in males and neural projection development in females (Table S9). Differentially expressed genes associated with knockdown of SSRIDD-associated fly orthologs in males are involved in mating behavior, cilia development, and muscle

423 contraction, while we see overrepresented ontology terms involved in chromatin modification, mitotic
424 cell cycle, and serine hydrolase activity in females (Table S9). We observed more alignment of GO terms
425 across genes and sexes in the CdLS fly models (*SMC1*, *SMC3*, *vtd*) than in SSRIDD fly models (*brm*, *osa*,
426 *Snr1*). There were no overrepresented GO terms for females in the CdLS-specific analysis. However, in
427 the 156 genes shared across both sexes and both the SSRIDD and CdLS disease-level analyses, we see an
428 overrepresentation of muscle cell development and actin assembly and organization (Table S9). GO
429 enrichment on k-means clusters does not reveal over-representation of any biological processes,
430 molecular functions or pathways for Clusters F7, F8, F4, F5, and M3 (Table S10). Genes involved in alpha-
431 glucosidase activity are overrepresented in Cluster M5 (Table S10).
432
433 We generated Venn diagrams (Figure S3) to display the degree of similarity in differentially expressed
434 genes across analyses, including the 156 genes shared across SSRIDD and CdLS males and females (Table
435 S6). Interestingly, 93% (2689/2907) of genes differentially expressed in a disease-specific analysis of
436 CdLS males were also differentially expressed in CdLS females or in SSRIDD fly models (Table S6). This is
437 in contrast to CdLS females, SSRIDD males, and SSRIDD females, in which about 25% of the differentially
438 expressed genes were specific to a single analysis (Table S6). Approximately 24 and 56 percent of the
439 differentially expressed genes (FDR<0.05) in pairwise comparisons for males and females, respectively,
440 have a predicted human ortholog (DIOPT > 9) (Table S11).
441
442 **Co-Regulated Genes**
443 We selected a subset of co-regulated genes from gene expression analyses as potential modifiers of the
444 focal genes *brm*, *osa*, and/or *Snr1.* We chose genes that had a significant effect (Line FDR < 0.05) in
445 analyses pooled across sexes, a suggestive effect (Line FDR < 0.1) for each sex separately, a greater than
446 or less than two-fold-change in both sexes, a strong human ortholog (DIOPT > 9), and an available

447 *attp40* TRiP RNAi line (the same genetic background as the focal genes). We increased the FDR threshold
448 to 0.1 for the sex-specific pairwise analyses to account for the decreased power of these analyses
449 compared to those with sexes combined. This resulted in 31 genes (Table S12). We further narrowed our
450 selection by prioritizing genes for further study with potential roles in neurological tissues, metabolism,
451 chromatin, orthologs associated with disease in humans, and computationally predicted genes of
452 unknown function. The six fly genes we selected for further study are *Alp10*, *CG40485*, *CG5877*, *IntS12*,
453 *Mal-A4*, and *Odc1*, which are orthologous to human genes *ALPG*, *DHRS11*, *NRDE2*, *INTS12*, *SLC3A1*, and
454 *ODC1*, respectively (human ortholog with highest DIOPT score listed; Table S12). All six genes tested
455 were co-regulated with *Snr1*, but *CG40485* and *CG5877* were not co-regulated with *osa* and *brm* models
456 of SSRIDDs (Table S6).
457
458 For each target gene, we crossed the *UAS-RNAi* line to the *Ubi156-GAL4* driver and performed qRT-PCR
459 to assess the magnitude of reduction in gene expression. All co-regulated genes had reduced expression
460 in both sexes (Table S13). We then assessed the effects of these genes on startle response, sleep, and
461 activity. Knockdown of *Mal-A4*, *CG5877* and *Alp10* showed changes in startle response times for both
462 sexes (Figure S4A, Table S14). *Mal-A4* demonstrated sexually dimorphic changes in startle response
463 similar to flies with *Snr1* knockdown, as females showed an increase ($p = 0.0215$) and males showed a
464 decrease ($p < 0.0001$) in startle response (Figure S4A, Table S14). We also quantified tapping behavior in
465 these co-regulated genes and found that flies with knockdown of *CG5877* and *Odc1* showed an increase
466 in tapping behavior compared to the control, similar to flies with knockdown of *osa* and *Snr1* (Figure 1B),
467 although we only observed tapping in females with knockdown of *Odc1* (Figure S4B, Table S14; *CG5877*
468 females: $p = 0.0266$, *CG5877* males: $p < 0.0001$; *Odc1* females: $p = 0.0125$).
469

With the exception of *CG40485*, which showed no changes in sleep or activity for either sex, all male RNAi genotypes had increased nighttime sleep bouts ($p < 0.03$), decreased night sleep ($p < 0.03$), and, with the additional exception of *CG5877* RNAi flies, increased overall activity ($p < 0.006$) (Figure S4, Table S14). Knockdown of *Mal-A4* and *Odc1* also showed increased activity for females ($p = 0.0049$, $p = 0.0044$, respectively). Only knockdown of *CG5877* resulted in increased night sleep for females ($p = 0.0014$) (Figure S4C-D, Table S14). These changes in activity and sleep phenotypes largely parallel those observed for SSRIDD fly models (Figure 2, Table S14).

Based on effects on startle response, tapping behavior, locomotor activity, night sleep, and sleep bouts, none of the phenotypes associated with RNAi of the co-regulated genes exactly matched the phenotypes associated with RNAi of the SSRIDD focal genes in both magnitude and direction. However, three genes (*Mal-A4*, *CG5877*, *Odc1*) exhibited at least one altered phenotype in both sexes (Figure S4). These phenotypic observations provide evidence that *Mal-A4*, *CG5877*, and/or *Odc1* may be interacting with the focal genes of the SSRIDD fly models.

DISCUSSION

Variants in members of the mammalian SWI/SNF complex (BAF complex) give rise to SSRIDDs, Mendelian disorders with a wide range of phenotypic manifestations, including Coffin-Siris and Nicolaides-Baraitser syndromes (reviewed in Bogershausen and Wollnik 2018; Schrier Vergano *et al.* 2021). The diverse consequences of such variants and variation in penetrance of similar variants in different affected individuals suggest the presence of segregating genetic modifiers. Such modifiers may represent targets for ameliorating therapies or serve as indicators of disease severity, yet they cannot be easily identified in humans due to the limited sample size of individuals with rare disorders. In addition to identifying potential modifiers, Drosophila models can be used to understand underlying

494 molecular effects of variants in chromatin-modification pathways and may aid in discovery of drugs that
495 ameliorate deleterious phenotypic effects.
496
497 We used a systematic comparative genomics approach to generate Drosophila models of disorders of
498 chromatin modification, based on the assumption that fundamental elements of chromatin modification
499 are evolutionarily conserved. First, we reduced expression of BAF and cohesin complex orthologs
500 through targeted RNA interference with a *GAL4* driver that induces minimal lethality. We assessed
501 consequences of target gene knockdown on behaviors that mimic those affected in patients with
502 SSRIDDs and CdLS. We used startle behavior, a proxy for sensorimotor integration, and sleep and activity
503 phenotypes to assess the effects of variants in fly orthologues of human genes associated with similar
504 behavioral disorders. These Drosophila models show increased activity, decreased night sleep, and
505 changes in sensorimotor integration. Although we cannot readily recapitulate cognitive developmental
506 defects in Drosophila, these behavioral phenotypes along with brain morphology measurements provide
507 a representative spectrum of behaviors that correlate with human disease phenotypes. We observed
508 gene-specific effects. In addition to showing the largest changes in sleep and activity phenotypes, only
509 *osa* RNAi flies showed stunted mushroom body alpha lobes. Furthermore, only females with knockdown
510 of *Snr1* showed an increase in startle response times. Our neuroanatomical studies focused on
511 morphological changes in the ellipsoid body and mushroom bodies. We cannot exclude effects on other
512 regions in the brain.
513
514 Next, we performed whole genome transcriptional profiling to identify co-regulated genes with each
515 focal gene and used stringent filters to identify candidate modifier genes from the larger subset of co-
516 regulated genes. k-means clustering reveals co-regulated genes unique to knockdown of a single protein
517 complex member (Figures S4, S5), yet also shows genes co-regulated in response to knockdown of

518 several, but not all, members of the fly cohesin and SWI/SNF complexes. Gene-specific and cross-disease
519 effects are intriguing, since *brm*, *osa*, and *Snr1* are part of the fly SWI/SNF complex, and *SMC1*, *SMC3*,
520 and *vtd* are part of the fly cohesin complex, yet have widespread gene-specific downstream effects on
521 gene regulation. Upon knockdown of one protein complex member, we did not necessarily find changes
522 in gene expression of other members of the same complex. It is possible that a compensatory
523 mechanism exists that maintains transcript levels of other fly SWI/SNF or cohesin complex members or
524 the focal genes themselves (Dorsett 2009; Raab *et al*. 2017; Van der Vaart *et al*. 2020), such as with
525 *Nipped-B* in a CdLS fly model (Wu *et al*. 2015). Furthermore, the abundance of lncRNAs co-regulated
526 with focal genes (Figures S4, S5, Table S8) is intriguing given the association between lncRNAs,
527 chromatin modification, and changes in gene expression in both flies and humans (Li *et al*. 2019; Statello
528 *et al*. 2021).
529
530 *Snr1* is part of the Brahma complex, a core component of the BAP complex and is orthologous to
531 *SMARCB1* (Table S2). *Odc1*, which encodes ornithine decarboxylase, is orthologous to *ODC1* (Table S12),
532 which is associated with Bachmann-Bupp syndrome, a rare neurodevelopmental disorder with alopecia,
533 developmental delay, and brain abnormalities (Prokop *et al*. 2021; Bupp *et al*. 2022). Ornithine
534 decarboxylase is the rate-limiting step of polyamine synthesis, which provides critical substrates for cell
535 proliferation and differentiation (reviewed in Wallace *et al*. 2003; Pegg 2016). Polyamines interact with
536 nucleic acids and transcription factors to modulate gene expression (Watanabe *et al*. 1991; Hobbs and
537 Gilmour 2000; Miller-Fleming *et al*. 2015; Maki *et al*. 2017). *CG5877* is predicted to mediate post-
538 transcriptional gene silencing as part of the spliceosome (Herold *et al*. 2009) and is orthologous to
539 human *NRDE2* (Table S12). *Mal-A4* is predicted to be involved in carbohydrate metabolism (Inomata *et*
540 *al*. 2019) and is orthologous to *SLC3A1* (Table S12). We observed extensive sexual dimorphism in

behavioral phenotypes and transcriptional profiles upon knockdown of SSRIDD- and CdLS-associated genes.

Although we are not aware of transcriptional profiles currently available for SSRIDD patients, RNA sequencing of post-mortem neurons from CdLS patients have shown dysregulation of hundreds of neuronal genes (Weiss *et al*. 2021). RNA sequencing in a *Nipped-B*-mutation fly model of *NIPBL*-CdLS found differential expression of ~2800 genes in the imaginal disc (FDR < 0.05) (Wu *et al*. 2015). Thus, we believe the number of differentially expressed genes upon gene knockdown reported herein is comparable to previous studies.

DATA AVAILABILITY

All high throughput sequencing data are deposited in GEO GSE213763.

Raw behavioral data, qPCR data, and coding scripts are available on GitHub at https://github.com/rebeccamacpherson/Dmel_models_CSS_NCBRS_CdLS. All *UAS*-RNAi lines used in this study are available at the Bloomington Drosophila Stock Center, except the ubiquitous RNAi driver *Ubi156-GAL4* and the double RNAi lines, which are available upon request.

ACKNOWLEDGEMENTS

We thank Dr. Lakshmi Sunkara for assistance with RNA sequencing, Marion R. Campbell III, Miller Barksdale, and Rachel C. Hannah for technical assistance with behavioral assays and brain dissections. We thank Dr. Joshua Walters for helping create Figure 3 and helping dissect brains, and Dr. Richard Steet at the Greenwood Genetic Center for suggestions. We thank Katelynne Collins and Tori Gyorey for assistance with the RNAi studies. We thank the TRiP at Harvard Medical School (NIH/NIGMS R01-GM084947) for providing transgenic RNAi fly stocks used in this study.

FUNDING STATEMENT

This work was funded by NIH grants R01 GM128974 and P20 GM139769 to TFCM and RRHA, and F31 HD106719 to RAM.

DECLARATION OF INTEREST

The authors have no competing interests to report.

AUTHOR CONTRIBUTIONS

RAM performed all experiments. VS assisted with RNA sequencing analysis; TFCM conceptualized the research program and TFCM and RRHA directed the research program. TFCM, RRHA, and RAM provided resources and wrote the manuscript.

LITERATURE CITED

1. Aoi H, Mizuguchi T, Ceroni JR, Kim VEH, Furquim I, Honjo RS, Iwaki T, Suzuki T, Sekiguchi F, Uchiyama Y, et al. 2019. Comprehensive genetic analysis of 57 families with clinically suspected Cornelia de Lange syndrome. *J Hum Genet.* 64(10):967-78.

2. Avagliano L, Parenti I, Grazioli P, Di Fede E, Parodi C, Mariani M, Kaiser FJ, Selicorni A, Gervasini C, Massa V. 2020. Chromatinopathies: A focus on Cornelia de Lange syndrome. *Clin Genet.* 97(1):3-11.

3. Bögershausen N, Wollnik B. 2018. Mutational landscapes and phenotypic spectrum of SWI/SNF-related intellectual disability disorders. *Front Mol Neurosci.* 11:252.

4. Boyle MI, Jespersgaard C, Nazaryan L, Bisgaard A-, Tümer Z. 2017. A novel *RAD21* variant associated with intrafamilial phenotypic variation in Cornelia de Lange syndrome – review of the literature. *Clin Genet.* 91(4):647-9.

589 5. Bramswig NC, Caluseriu O, Lüdecke H-, Bolduc FV, Noel NCL, Wieland T, Surowy HM, Christen H-,
590 Engels H, Strom TM, et al. 2017. Heterozygosity for *ARID2* loss-of-function mutations in individuals
591 with a Coffin–Siris syndrome-like phenotype. *Hum Genet.* 136(3):297-305.
592 6. Bupp C, Michale J, VanSickle E, Rajasekaran S, Bachmann AS. 2022. Bachmann-Bupp syndrome. In:
593 GeneReviews ®. Adam MP, Everman DB, Mirzaa GM et al., editors. Seattle, WA: University of
594 Washington, Seattle.
595 7. Burns RB, Dobson CB. 1981. Standard error of the difference between means. In: *Experimental*
596 *Psychology*. Springer, Dordrecht.
597 8. Bushnell B. BBMap. SOURCEFORGE. 2014.
598 9. Chen S, Huang T, Zhou Y, Han Y, Xu M, Gu J. 2017. AfterQC: Automatic filtering, trimming, error
599 removing and quality control for fastq data. *BMC Bioinformatics.* 18(Suppl 3):80.
600 10. Chubak MC, Nixon KCJ, Stone MH, Raun N, Rice SL, Sarikahya M, Jones SG, Lyons TA, Jakub TE,
601 Mainland RLM, et al. 2019. Individual components of the SWI/SNF chromatin remodelling complex
602 have distinct roles in memory neurons of the Drosophila mushroom body. *Dis Model*
603 *Mech.* 12(3):dmm037325.
604 11. Cichewicz K, Hirsh J. 2018. ShinyR-DAM: A program analyzing Drosophila activity, sleep and circadian
605 rhythms. *Commun Biol.* 1(1):25.
606 12. Cucco F, Sarogni P, Rossato S, Alpa M, Patimo A, Latorre A, Magnani C, Puisac B, Ramos FJ, Pié J, et
607 al. 2020. Pathogenic variants in *EP300* and *ANKRD11* in patients with phenotypes overlapping
608 Cornelia de Lange syndrome. *Am J Med Genet A.* 182(7):1690-6.
609 13. de Chaumont F, Dallongeville S, Chenouard N, Hervé N, Pop S, Provoost T, Meas-Yedid V,
610 Pankajakshan P, Lecomte T, Le Montagner Y, et al. 2012. Icy: An open bioimage informatics
611 platform for extended reproducible research. *Nat Methods.* 9(7):690-6.

612 14. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, Kou Y, Liu L, Fromer M, Walker S,
613 et al. 2014. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature.*
614 515(7526):209-15.
615 15. Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, Minamino M, Saitoh K, Komata M, Katou Y,
616 Clark D, et al. 2012. *HDAC8* mutations in Cornelia de Lange syndrome affect the cohesin acetylation
617 cycle. *Nature.* 489:313-7.
618 16. Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, Gil-Rodríguez C, Arnedo M, Loeys B,
619 Kline AD, et al. 2007. Mutations in cohesin complex members *SMC3* and *SMC1A* cause a mild
620 variant of Cornelia de Lange syndrome with predominant mental retardation. *Am J Hum Genet.*
621 80(3):485-94.
622 17. Dorsett D. 2009. Cohesin, gene expression and development: Lessons from Drosophila. *Chromosome*
623 *Res.* 17(2):185-200.
624 18. Fox J, Weisberg S. 2019. *An R companion to applied regression*. Third ed. Thousand Oaks, CA: Sage.
625 19. Garlapow M, Huang W, Yarboro M, Peterson K, Mackay T. 2015. Quantitative genetics of food intake
626 in *Drosophila melanogaster*: e0138129. *PLoS One.* 10(9).
627 20. Gazdagh G, Blyth M, Scurr I, Turnpenny PD, Mehta SG, Armstrong R, McEntagart M, Newbury-Ecob
628 R, Tobias ES, Joss S. 2019. Extending the clinical and genetic spectrum of *ARID2* related intellectual
629 disability. A case series of 7 patients. *Eur J Med Genet.* 62(1):27-34.
630 21. Gil-Rodríguez MC, Deardorff MA, Ansari M, Tan CA, Parenti I, Baquero-Montoya C, Ousager LB,
631 Puisac B, Hernández-Marcos M, Teresa-Rodrigo ME, et al. 2015. De novo heterozygous mutations
632 in *SMC3* cause a range of Cornelia de Lange syndrome-overlapping phenotypes. *Hum Mutat.*
633 36(4):454-62.
634 22. Guo F, Yi W, Zhou M, Guo A. 2011. Go signaling in mushroom bodies regulates sleep in
635 Drosophila. *Sleep.* 34(3):273-81.

636 23. Hempel A, Pagnamenta AT, Blyth M, Mansour S, McConnell V, Kou I, Ikegawa S, Tsurusaki Y,
637 Matsumoto N, Lo-Castro A, et al. 2016. Deletions and de novo mutations of *SOX11* are associated
638 with a neurodevelopmental disorder with features of Coffin–Siris syndrome. *J Med Genet.*
639 53(3):152-62.
640 24. Herold N, Will CL, Wolf E, Kastner B, Urlaub H, Lührmann R. 2009. Conservation of the protein
641 composition and electron microscopy structure of drosophila melanogaster and human
642 spliceosomal complexes. *Mol Cell Biol.* 29(1):281-301.
643 25. Hobbs CA, Gilmour SK. 2000. High levels of intracellular polyamines promote histone
644 acetyltransferase activity resulting in chromatin hyperacetylation. *J Cell Biochem.* 77(3):345-60.
645 26. Hoyer J, Ekici A, Endele S, Popp B, Zweier C, Wiesener A, Wohlleber E, Dufke A, Rossier E, Petsch C, et
646 al. 2012. Haploinsufficiency of *ARID1B*, a member of the SWI/SNF-A chromatin-remodeling
647 complex, is a frequent cause of intellectual disability. *Am J Hum Genet.* 90(3):565-72.
648 27. Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N, Mohr SE. 2011. An integrative
649 approach to ortholog prediction for disease-focused and other functional studies. *BMC*
650 *Bioinformatics.* 12(1):357.
651 28. Huggett SB, Hatfield JS, Walters JD, McGeary JE, Welsh JW, Mackay TFC, Anholt RRH, Palmer RHC.
652 2021. Ibrutinib as a potential therapeutic for cocaine use disorder. *Transl Psychiatry.* 11(1):623.
653 29. Huisman S, Mulder PA, Redeker E, Bader I, Bisgaard A, Brooks A, Cereda A, Cinca C, Clark D, Cormier-
654 Daire V, et al. 2017. Phenotypes and genotypes in individuals with *SMC1A* variants. *Am J Med*
655 *Genet A.* 173A(8):2108-25.
656 30. Inomata N, Takahasi KR, Koga N. 2019. Association between duplicated maltase genes and the
657 transcriptional regulation for the carbohydrate changes in *Drosophila melanogaster. Gene.*
658 686:141-145.

659 31. Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives
660 L, Patterson KE, et al. 2014. The contribution of de novo coding mutations to autism spectrum
661 disorder. *Nature.* 515(7526):216-21.
662 32. Jansen S, Kleefstra T, Willemsen MH, de Vries P, Pfundt R, Hehir-Kwa JY, Gilissen C, Veltman JA, de
663 Vries BBA, Vissers LELM. 2016. De novo loss-of-function mutations in X-linked *SMC1A* cause severe
664 ID and therapy-resistant epilepsy in females: Expanding the phenotypic spectrum. *Clin*
665 *Genet.* 90(5):413-9.
666 33. Joiner WJ, Crocker A, White BH, Sehgal A. 2006. Sleep in Drosophila is regulated by adult mushroom
667 bodies. *Nature.* 441(7094):757-60.
668 34. Kline AD, Moss JF, Selicorni A, Bisgaard A, Deardorff MA, Gillett PM, Ishman SL, Kerr LM, Levin AV,
669 Mulder PA, et al. 2018. Diagnosis and management of Cornelia de Lange syndrome: First
670 international consensus statement. *Nat Rev Genet.* 19(10):649-66.
671 35. Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A,
672 Morris CA, et al. 2004. Cornelia de Lange syndrome is caused by mutations in *NIPBL*, the human
673 homolog of *Drosophila melanogaster Nipped-B*. *Nat Genet.* 36(6):631-5.
674 36. Li K, Tian Y, Yuan Y, Fan X, Yang M, He Z, Yang D. 2019. Insights into the functions of lncRNAs in
675 Drosophila. *Int J Mol Sci*. 20(18):4646.
676 37. Liao Y, Smyth GK, Shi W. 2013. The subread aligner: Fast, accurate and scalable read mapping by
677 seed-and-vote. *Nucleic Acids Res.* 41(10):e108.
678 38. Liu J, Krantz ID. 2009. Cornelia de Lange syndrome, cohesin, and beyond. *Clin Genet.* 76(4):303-14.
679 39. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative
680 PCR and the 2(−Δ Δ C(T)) method. *Methods.* 25(4):402-8.
681 40. Machol K, Rousseau J, Ehresmann S, Garcia T, Nguyen TTM, Spillmann RC, Sullivan JA, Shashi V, Jiang
682 Yh, Stong N, et al. 2019. Expanding the spectrum of BAF-related disorders: De novo variants in

683 *SMARCC2* cause a syndrome with intellectual disability and developmental delay. *Am J Hum Genet.*
684 104(1):164-78.
685 41. Maki K, Shibata T, Kawabata S. 2017. Transglutaminase-catalyzed incorporation of polyamines masks
686 the DNA-binding region of the transcription factor relish. *J Biol Chem.* 292(15):6369-80.
687 42. Mi H, Muruganujan A, Thomas PD. 2013. PANTHER in 2013: Modeling the evolution of gene function,
688 and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* 41(Database
689 issue):D377-86.
690 43. Miller-Fleming L, Olin-Sandoval V, Campbell K, Ralser M. 2015. Remaining mysteries of molecular
691 biology: The role of polyamines in the cell. *J Mol Biol.* 427(21):3389-406.
692 44. Modi MN, Shuai Y, Turner GC. 2020. The Drosophila mushroom body: From architecture to algorithm
693 in a learning circuit. *Annu Rev Neurosci.* 43(1):465-84.
694 45. Olley G, Ansari M, Bengani H, Grimes GR, Rhodes J, von Kriegsheim A, Blatnik A, Stewart FJ, Wakeling
695 E, Carroll N, et al. 2018. *BRD4* interacts with *NIPBL* and *BRD4* is mutated in a Cornelia de Lange-like
696 syndrome. *Nat Genet.* 50(3):329-32.
697 46. Parenti I, Teresa-Roigo ME, Pozojevic J, Gil SR, Bader I, Braunholz D, Bramswig NC, Gervasini C,
698 Larizza L, Pfeiffer L, et al. 2017. Mutations in chromatin regulators functionally link Cornelia de
699 Lange syndrome and clinically overlapping phenotypes. *Hum Genet.* 136(3):307-20.
700 47. Pauli A, Althoff F, Oliveira RA, Heidmann S, Schuldiner O, Lehner CF, Dickson BJ, Nasmyth K. 2008.
701 Cell-type-specific TEV protease cleavage reveals cohesin functions in Drosophila neurons. *Dev*
702 *Cell.* 14(2):239-51.
703 48. Pegg AE. 2016. Functions of polyamines in mammals. *J Biol Chem.* 291(29):14904-12.
704 49. Perkins LA, Holderbaum L, Tao R, Hu Y, Sopko R, McCall K, Yang-Zhou D, Flockhart I, Binari R, Shim H,
705 et al. 2015. The transgenic RNAi project at Harvard medical school: Resources and
706 validation. *Genetics.* 201(3):843-52.

707 50. Pitman J, McGill J, Keegan K, Allada R. 2006. A dynamic role for the mushroom bodies in promoting
708 sleep in Drosophila. *Nature.* 441(7094):753-6.
709 51. Prokop JW, Bupp CP, Frisch A, Bilinovich SM, Campbell DB, Vogt D, Schultz CR, Uhl KL, VanSickle E,
710 Rajasekaran S, et al. 2021. Emerging role of *ODC1* in neurodevelopmental disorders and brain
711 development. *Genes.* 12(4):470.
712 52. Raab JR, Runge JS, Spear CC, Magnuson T. 2017. Co-regulation of transcription by BRG1 and BRM,
713 two mutually exclusive SWI/SNF ATPase subunits. *Epigenetics Chromatin.* 10(1):62.
714 53. Rajan R, Benke JR, Kline AD, Levy HP, Kimball A, Mettel TL, Boss EF, Ishman SL. 2012. Insomnia in
715 Cornelia de Lange syndrome. *Int J Pediatr Otorhinolaryngol.* 76(7):972-5.
716 54. Santen G, Aten E, Sun Y, Almomani R, Gilissen C, Nielsen M, Kant SG, Snoeck IN, Peeters E, Hilhorst-
717 Hofstee Y, et al. 2012. Mutations in SWI/SNF chromatin remodeling complex gene *ARID1B* cause
718 Coffin-Siris syndrome. *Nat Genet.* 44(4):379-80.
719 55. Schrier Vergano S, Santen G, Wieczorek D, Wollnik B, Matsumo N, Deardorff MA. 2013. Coffin-Siris
720 syndrome. In: GeneReviews ®. Adam MP, Everman DB, Mirzaa GM, et al., editors. Seattle, WA:
721 University of Washington, Seattle.
722 56. Schuldiner O, Berdnik D, Levy JM, Wu JS, Luginbuhl D, Gontang AC, Luo L. 2008. piggyBac-based
723 mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon
724 pruning. *Dev Cell.* 14(2):227-38.
725 57. Selicorni A, Mariani M, Lettieri A, Massa V. 2021. Cornelia de Lange syndrome: From a disease to a
726 broader spectrum. *Genes.* 12(7):1075.
727 58. Sitaraman D, Aso Y, Rubin GM, Nitabach MN. 2015. Control of sleep by dopaminergic inputs to the
728 Drosophila mushroom body. *Front Neural Circuits.* 9:73.
729 59. Smid M, Coebergh van den Braak, Robert, van de Werken H, van Riet J, Galen A, Weerd V, Daane M,
730 Bril S, Lalmahomed Z, Kloosterman WP, et al. 2018. Gene length corrected trimmed mean of M-

values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. *BMC Bioinformatics.* 19(1):236.

60. Statello L, Guo C, Chen L, Huarte M. 2021. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol.* 22:96-118.

61. Stavinoha RC, Kline AD, Levy HP, Kimball A, Mettel TL, Ishman SL. 2010. Characterization of sleep disturbance in Cornelia de Lange syndrome. *Int J Pediatr Otorhinolaryngol.* 75(2):215-8.

62. Symonds JD, Joss S, Metcalfe KA, Somarathi S, Cruden J, Devlin AM, Donaldson A, DiDonato N, Fitzpatrick D, Kaiser FJ, et al. 2017. Heterozygous truncation mutations of the *SMC1A* gene cause a severe early onset epilepsy with cluster seizures in females: Detailed phenotyping of 10 new cases. *Epilepsia.* 58(4):565-75.

63. Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albou L, Mi H. 2022. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* 31(1):8-22.

64. Tsurusaki Y, Koshimizu E, Ohashi H, Phadke S, Kou I, Shiina M, Suzuki T, Okamoto N, Imamura S, Yamashita M, et al. 2014. De novo *SOX11* mutations cause Coffin–Siris syndrome. *Nat Commun.* 5(1):4011.

65. Tsurusaki Y, Okamoto Nobuhiko, Fukushima Y, Homma T, Kato M, Hiraki Y, Yamagata T, Yano S, Mizuno S, Sakazume S, et al. 2012. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet.* 44(4):376-8.

66. van Allen MI, Filippi G, Siegel-Bartelt J, Yong S, McGillivray B, Zuker RM, Smith CR, Magee JF, Ritchie S, Toi A, et al. 1993. Clinical variability within Brachmann-de Lange syndrome: A proposed classification system. *Am J Med Genet.* 47(7):947-58.

67. van der Sluijs, Pleuntje J., Jansen S, Vergano SA, Adachi-Fukuda M, Alanay Y, AlKindy A, Baban A, Bayat A, Beck-Wödl S, Berry K, et al. 2019. The *ARID1B* spectrum in 143 patients: From nonsyndromic intellectual disability to Coffin–Siris syndrome. *Genet Med.* 21(6):1295-307.

755 68. van der Vaart A, Godfrey M, Portegijs V, Heuvel S. 2020. Dose-dependent functions of SWI/SNF BAF
756 in permitting and inhibiting cell proliferation in vivo. *Sci Adv.* 6(21):eaay3823.
757 69. van Houdt, Jeroen K. J, Nowakowska BA, Sousa SB, van Schaik, Barbera D. C, Seuntjens E, Avonce N,
758 Sifrim A, Abdul-Rahman OA, van den Boogaard, Marie-José H, Bottani A, et al. 2012. Heterozygous
759 missense mutations in *SMARCA2* cause Nicolaides-Baraitser syndrome. *Nat Genet.* 44(4):445-9.
760 70. Vasileiou G, Vergarajauregui S, Endele S, Popp B, Büttner C, Ekici AB, Gerard M, Bramswig NC,
761 Albrecht B, Clayton-Smith J, et al. 2018. Mutations in the BAF-complex subunit *DPF2* are associated
762 with Coffin-Siris syndrome. *Am J Hum Genet.* 102(3):468-79.
763 71. Vasko A, Schrier Vergano SA. 2022. Language impairments in individuals with Coffin-Siris
764 syndrome. *Front Neurosci.* 15:802583.
765 72. Vasko A, Drivas TG, Schrier Vergano SA. 2021. Genotype-phenotype correlations in 208 individuals
766 with Coffin-Siris syndrome. *Genes.* 12(6):937.
767 73. Vissers, Lisenka E. L. M., Gilissen C, Veltman JA. 2016. Genetic studies in intellectual disability and
768 related disorders. *Nat Rev Genet.* 17(1):9-18.
769 74. Wallace HM, Fraser AV, Hughes A. 2003. A perspective of polyamine metabolism. *Biochem J.*
770 376(1):1-14.
771 75. Watanabe S, Kusama-Eguchi K, Kobayashi H, Igarashi K. 1991. Estimation of polyamine binding to
772 macromolecules and ATP in bovine lymphocytes and rat liver. *J Biol Chem.* 266(31):20803-9.
773 76. Weiss FD, Calderon L, Wang Y, Georgieva R, Guo Y, Cvetesic N, Kaur M, Dharmalingam G, Krantz ID,
774 Lenhard B, et al. 2021. Neuronal genes deregulated in Cornelia de Lange syndrome respond to
775 removal and re-expression of cohesin. *Nat Commun.* 12(1):2919.
776 77. Wu TD, Reeder J, Lawrence M, Becker G, Brauer MJ. 2016. GMAP and GSNAP for genomic sequence
777 alignment: Enhancements to speed, accuracy, and functionality. In: Statistical genomics: Methods
778 and protocols. Mathé E, Davis S, editors. New York, NY: Springer New York. 283 p.

78. Wu Y, Gause M, Xu D, Misulovin Z, Schaaf CA, Mosarla RC, Mannino E, Shannon M, Jones E, Shi M, et al. 2015. Drosophila *Nipped-B* mutants model Cornelia de Lange syndrome in growth and behavior. *PLoS Genet.* 11(11):e1005655.

79. Yamamoto A, Zwarts L, Callaerts P, Norga K, Mackay TFC, Anholt RRH. 2008. Neurogenetic networks for startle-induced locomotion in *Drosophila melanogaster*. *Proc Natl Acad Sci USA.* 105(34):12393-8.

80. Zambrelli E, Fossati C, Turner K, Taiana M, Vignoli A, Gervasini C, Russo S, Furia F, Masciadri M, Ajmone P, et al. 2016. Sleep disorders in Cornelia de Lange syndrome. *Am J Med Genet C.* 172(2):214-21.

81. Zawerton A, Yao B, Yeager JP, Pippucci T, Haseeb A, Smith JD, Wischmann L, Kühl SJ, Dean JCS, Pilz DT, et al. 2019. De novo *SOX4* variants cause a neurodevelopmental disease associated with mild dysmorphism. *Am J Hum Genet.* 104(2):246-59.

82. Zirin J, Hu Y, Liu L, Yang-Zhou D, Colbeth R, Yan D, Ewen-Campen B, Tao R, Vogt E, VanNest S, et al. 2020. Large-scale transgenic Drosophila resource collections for loss- and gain-of-function studies. *Genetics.* 214(4):755-67.

83. Zwarts L, Vanden Broeck L, Cappuyns E, Ayroles JF, Magwire MM, Vulsteke V, Clements J, Mackay TFC, Callaerts P. 2015. The genetic basis of natural variation in mushroom body size in *Drosophila melanogaster*. *Nat Commun.* 6(1):10115.

797 **Table 1. Drosophila genes used in fly models.** The table indicates fly genes used in SSRIDD and CdLS fly
798 models, as well as the respective human orthologs and MIM numbers, associated human disease and
799 respective MIM numbers, and DIOPT scores. Human orthologs are only included in the table if the DIOPT
800 score is greater than 9.

Fly Gene	Human Ortholog(s)	Human Ortholog MIM number(s)	Associated Human Disease	Phenotype MIM Number(s)	DIOPT score
brm	*SMARCA2, SMARCA4*	600014, 603254	SSRIDD (NCBRS, CSS 4)	601358, 614609	13, 12
osa	*ARID1A, ARID1B*	603024, 614556	SSRIDD (CSS 2, CSS 1)	614607, 135900	12, 12
SMC1	*SMC1A*	300040	Cornelia de Lange syndrome 2	300590	12
SMC3	*SMC3*	606062	Cornelia de Lange syndrome 3	610759	12
Snr1	*SMARCB1*	601607	SSRIDD (CSS 3)	614608	15
vtd	*RAD21, RAD21L1*	606462, 619533	Cornelia de Lange syndrome 4	614701	11, 10

801

802

803 **Table 2. Differentially expressed gene counts.** The table shows the number of differentially expressed
804 genes (FDR < 0.05) for the Line and/or Line × Sex terms for each pairwise analysis of knockdown vs
805 control, sexes together and sexes separately.
806

	Analysis			
	Both sexes		Females only	Males only
Comparison	Line	Line×Sex	Line	Line
brm vs. Control	2808	1652	583	2995
osa vs. Control	2179	1059	1135	1580
Snr1 vs. Control	4996	3632	3026	3376
SMC1 vs. Control	2714	1727	2540	2395
SMC3 vs. Control	1874	586	2711	1161
vtd vs. Control	1998	961	818	1630

807
808

FIGURE LEGENDS

Figure 1. Altered startle response phenotypes in SSRIDD and CdLS fly models. Startle phenotypes of flies with *Ubi156-GAL4*-mediated RNAi knockdown. (A) Boxplots showing the time, in seconds, spent moving after an initial startle force. Asterisks represent sex-specific pairwise comparisons with the control. (B) Bar graphs showing the percentage of flies that exhibit tapping behavior (see File S1 and S2) following an initial startle stimulus. Females and males are shown in purple and green, respectively. See Table S4 for ANOVAs (A) and Fisher's Exact Tests (B). N = 36-50 flies per sex per line. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$.

Figure 2. Altered sleep and activity phenotypes in SSRIDD and CdLS fly models. Boxplots displaying activity and sleep phenotypes of flies with *Ubi156-GAL4*-mediated RNAi knockdown. (A) total activity; (B) proportion of time spent asleep at night; (C) number of sleep bouts at night. Females and males are shown in purple and green, respectively. N = 18-32 flies per sex per line. See Table S4 for ANOVAs. Asterisks indicate pairwise comparisons of each line to the control, sexes separately. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$.

Figure 3. Examples of mushroom body abnormalities in SSRIDD and CdLS fly models. Images of a wild type mushroom body annotated with measurement descriptors for (A) mushroom body alpha and beta lobes, and heel-heel normalization measurement; and (B) ellipsoid body measurements. Images of select brains from flies with *Ubi156-GAL4*-mediated RNAi knockdown of *osa* showing (C) stunted alpha lobe outgrowth and narrowed alpha lobe head in a female *osa*-deficient fly brain; and (D) beta lobe crossing the midline/fused beta lobes, as well as a skinny alpha lobe in a male *osa*-deficient fly brain. Images shown are z-stack maximum projections from confocal imaging. Triangular arrowheads indicate the abnormalities. The scale bar represents 25 μM.

Figure 4. SSRIDD and CdLS fly models show gene-specific changes in mushroom body and ellipsoid body. Boxplots showing (A) the average alpha lobe and (B) beta lobe length for each brain; (C) ellipsoid body height (vertical direction; dorsal-ventral) and (D) width (left-right; lateral). Bar graphs showing the percentage of brains that (E) have a stunted alpha lobe(s)/narrowed alpha lobe head(s); (F) have a beta lobe(s) crossing the midline, including fused beta lobes; and (G) display one of more of the following defects: skinny alpha lobe, missing alpha lobe, skinny beta lobe, missing beta lobe, stunted alpha lobe/narrowed alpha lobe head, beta lobe crossing the midline/fused beta lobes, extra projections off of the alpha lobe, extra projections off of the beta lobe. See Figure 3. All brains were dissected from flies with *Ubi156-GAL4*-mediated RNAi knockdown. For panels A-D, brains missing only one alpha or beta lobe are represented by the length of the remaining lobe and brains missing both alpha lobes or both beta lobes were not included in the analyses. For panels E-G, data were analyzed with a Fisher's Exact test, sexes separately. Asterisks (*) and diamonds (panels A-D only; ◊) represent pairwise comparisons of the knockdown line versus the control in ANOVAs or Fisher's Exact tests, and Levene's tests for unequal variances, respectively. See Table S4 for ANOVAs, Fisher's Exact and Levene's Test results. Females and males are shown in purple and green, respectively. N = 17-20 brains per sex per line. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$. ◊: $p < 0.05$, ◊◊: $p < 0.01$.

Figure 5. k-means clusters for females.

k-means clusters (k = 10, average linkage algorithm) based on expression patterns of the 535 genes with maximal absolute value of the fold-change in expression, compared to the control. Blue and yellow indicate lower and higher expression, respectively.

Figure 6. k-means clusters for males.

k-means clusters (k = 8, average linkage algorithm) based on expression patterns of the 535 genes with maximal absolute value of the fold-change in expression, compared to the control. Blue and yellow indicate lower and higher expression, respectively.

SUPPLEMENTARY INFORMATION

Figure S1. Gross viability observations in potential CSS/NCBRS and CdLS fly models. Life stage shown is the final stage of the *Drosophila* life cycle where live individuals were observed. "X" indicates flies did not have detectable levels of gene knockdown, as quantified via qRT-PCR.

Figure S2. Selection of genes for k-means clustering. Elbow plots of maximal fold change in expression plotted against rank order (blue) across all analyses for each of 9657 genes (A) and for genes with a maximum fold change difference greater than 4 (B). See Table S7. The red and green lines were fit to roughly linear segments of the generated distribution (blue). The orange lines are drawn from the plot elbow (determined by the x coordinate of the intersection of the green and red lines) to the x and y axes.

Figure S3. Overlap of differentially expressed genes in SSRIDD and CdLS fly models. Venn diagrams displaying the number of differentially expressed genes (FDR < 0.05), in **SSRIDD** and CdLS fly models, sexes separately. Pairwise gene-specific analyses from (A) CdLS fly models and (B) **SSRIDD** fly models. Panel (C) shows overlap of disease-specific analyses, pooled across disease-associated genes.

880 **Figure S4. Altered phenotypes due to knockdown of co-regulated genes.** Bar plots displaying
881 differences in the average values of the experimental line versus the control line for (A) startle response,
882 (B) percent of flies tapping, (C) total activity, and (D) proportion of time asleep at night. All lines have
883 *Ubi156-GAL4*-mediated RNAi knockdown. Females and males are shown in purple and green,
884 respectively. See Table S14 for ANOVAs (A,B,D) and Fishers Exact Tests (C). N=29-32 per sex per line.
885 Error bars represent standard error of the difference based on error propagation (Burns and Dobson
886 1981). Asterisks represent pairwise analyses of the experimental line vs the control, sexes separately. *:
887 $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.
888
889 **Table S1. Fly reagents and primer sequences.** Drosophila reagents and primer sequences. (A)
890 Drosophila lines used. (B) Primer sequences used for qRT-PCR. BDSC: Bloomington Drosophila Stock
891 Center.
892
893 **Table S2. Ortholog prediction scores for potential focal genes.** Human-Drosophila ortholog prediction
894 scores generated using Drosophila RNAi Screening Center Integrative Ortholog Prediction Tool (DIOPT).
895 Human genes associated with SSRIDDs and Cornelia de Lange syndrome.
896
897 **Table S3. Percent knockdown of focal genes.** Average RNAi-mediated qRT-PCR knockdown of focal
898 genes.
899
900 **Table S4. Quantification of changes in behavior and brain morphology from knockdown of focal genes.**
901 Quantification of changes in behavior and brain morphology from RNAi knockdown. Statistical analyses
902 characterizing SSRIDD and CdLS fly models. (A) ANOVAs for startle response. (B) Fisher's Exact Tests for
903 tapping behavior. (C) ANOVAs for sleep and activity measurements. (D) ANOVAS for mushroom body

904 lobe lengths. (E) Levene's and Brown-Forsythe Tests for unequal variances of mushroom body lobe
905 length data. (F) Gross brain abnormalities. Line and Sex are fixed effects. df: degrees of freedom, SS:
906 Type III Sum of Squares, MS: Mean Squares.
907
908 **Table S5. ANOVA results from differential expression analyses.** Gene name, gene symbol, FlyBase ID,
909 normalized read counts (counts per million), and raw and Benjamini-Hochberg FDR adjusted *p*-values for
910 all genes for all model terms used in the ANOVA analyses. (A) Full model using all knockdown lines and
911 the control according to the model $Y = \mu + Line + Sex + Line \times Sex + \square$ for 15915 genes. (B-G) Pairwise
912 comparisons of single gene knockdown vs. the control (sexes together $Y = \mu + Line + Sex + Line \times Sex + \square$;
913 and sexes separately $Y = \mu + Line + \square$) on the 9657 genes from the full model differentially expressed
914 (FDR < 0.05) for the *Line* and/or *Line x Sex* terms. (B) *brm*. (C) *osa*. (D) *Snr1*. (E) *SMC1*. (F) *SMC3*. (G) *vtd*.
915 (H-I) Disease-specific comparisons (sexes together $Y = \mu + Line + Sex + Line \times Sex + \square$; and sexes
916 separately $Y = \mu + Line + \square$). (H) SSRIDDs. (I) Cornelia de Lange syndrome (CdLS).
917
918 **Table S6. Overlap of differentially expressed genes across analyses.** FDR-corrected *p*-values less than
919 0.05 for the *Line* term of each of the 9657 genes. (A) Pairwise analyses of each knockdown line
920 compared to the control, sexes separately. (B) Disease-specific analyses, sexes separately. NA indicates
921 FDR-corrected *P*-values for the effect of *Line* greater than 0.05.
922
923 **Table S7. k-means threshold.** (A) Average log2 fold change values for each differentially expressed gene
924 for each set of samples, as well as maximum, minimum across all samples. (B) Determination of
925 threshold by ranking, indexing and fitting lines to fold change plots. fc: log2 fold change; f: females, m:
926 males.
927

928 **Table S8. k-means clustering gene lists.** Lists of genes within each k-means cluster. (A) Females. (B)
929 Males.

930

931 **Table S9. Gene Ontology (GO) analyses for differentially expressed genes.** "Analysis" indicates the gene
932 set used in the analysis.

933

934 **Table S10. Gene Ontology (GO) analyses for k-means clusters.** "Analysis" indicates the gene set used in
935 the analysis.

936

937 **Table S11. Ortholog prediction scores for differentially expressed genes**. Drosophila-human ortholog
938 prediction scores, generated using Drosophila RNAi Screening Center Integrative Ortholog Prediction
939 Tool (DIOPT). Differentially expressed fly genes for each by-sex pairwise comparison.

940

941 **Table S12. Ortholog prediction scores and known disease associations for co-regulated genes.**
942 Drosophila-human ortholog prediction scores, generated using Drosophila RNAi Screening Center
943 Integrative Ortholog Prediction Tool (DIOPT) and Online Mendelian Inheritance of Man (OMIM)-derived
944 known disease/phenotype associations and corresponding MIM numbers. Subset of 31 Drosophila
945 genes co-regulated with *brm*, *osa*, and/or *Snr1*.

946

947 **Table S13. Percent knockdown of co-regulated genes.** Average RNAi-mediated qRT-PCR knockdown of
948 co-regulated genes.

949

950 **Table S14. Quantification of changes in behavior from knockdown of co-regulated genes.**
951 Quantification of changes in behavior from RNAi knockdown of co-regulated genes. (A) ANOVAs for

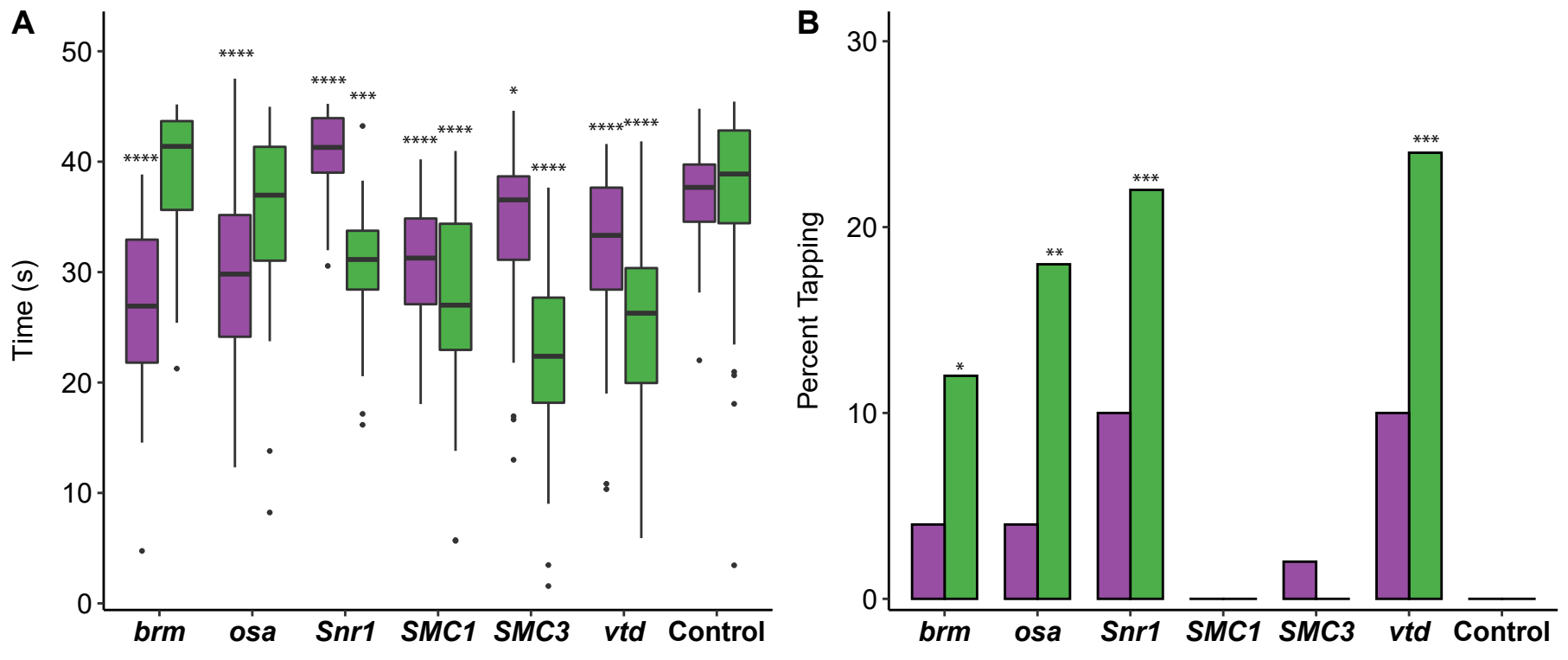
952 startle response. (B) Fisher's Exact Tests for tapping behavior. (C) ANOVAs for sleep and activity
953 measurements. Line and Sex are fixed effects. df: degrees of freedom, SS: Type III Sum of Squares, MS:
954 Mean Squares.

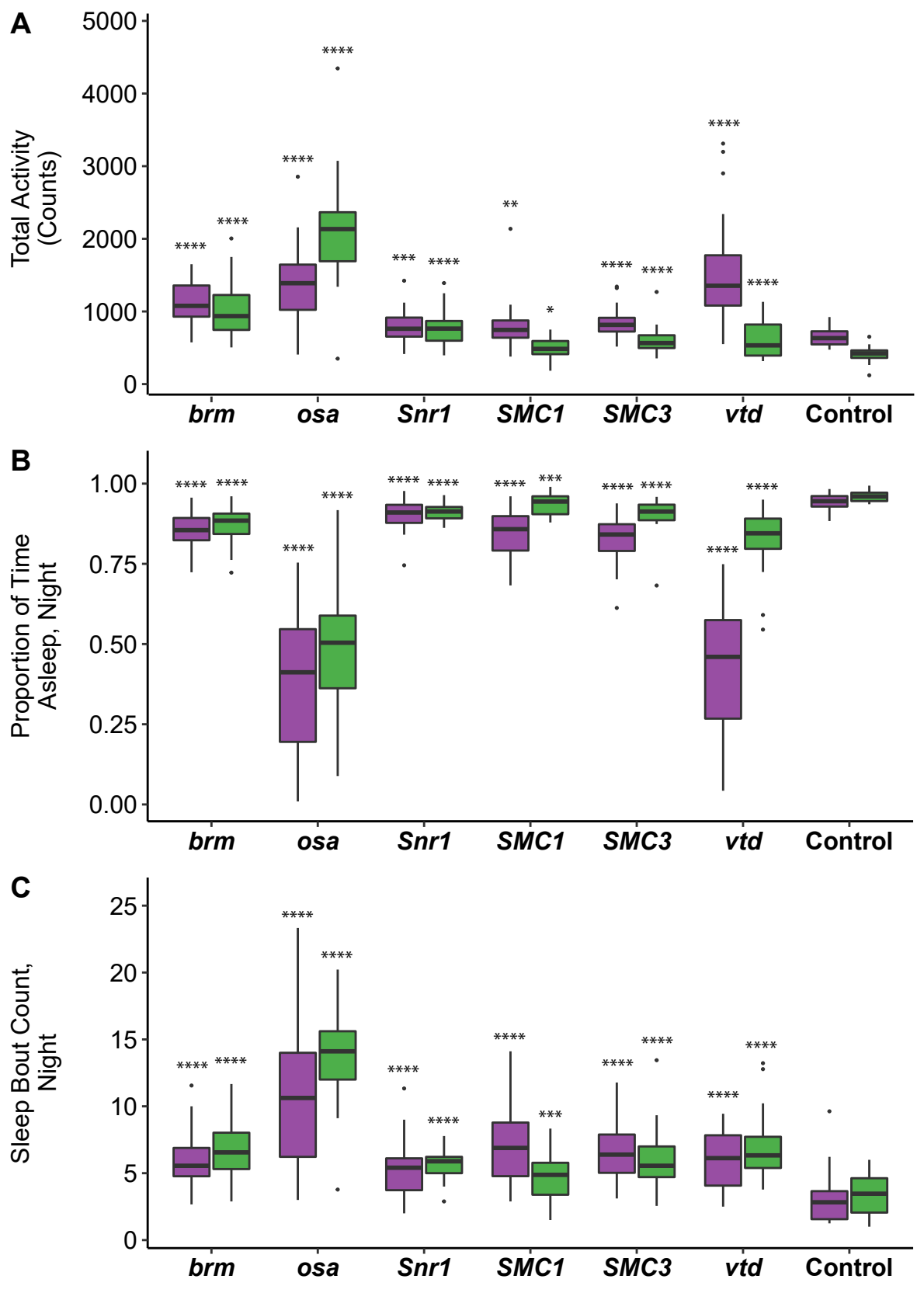
955

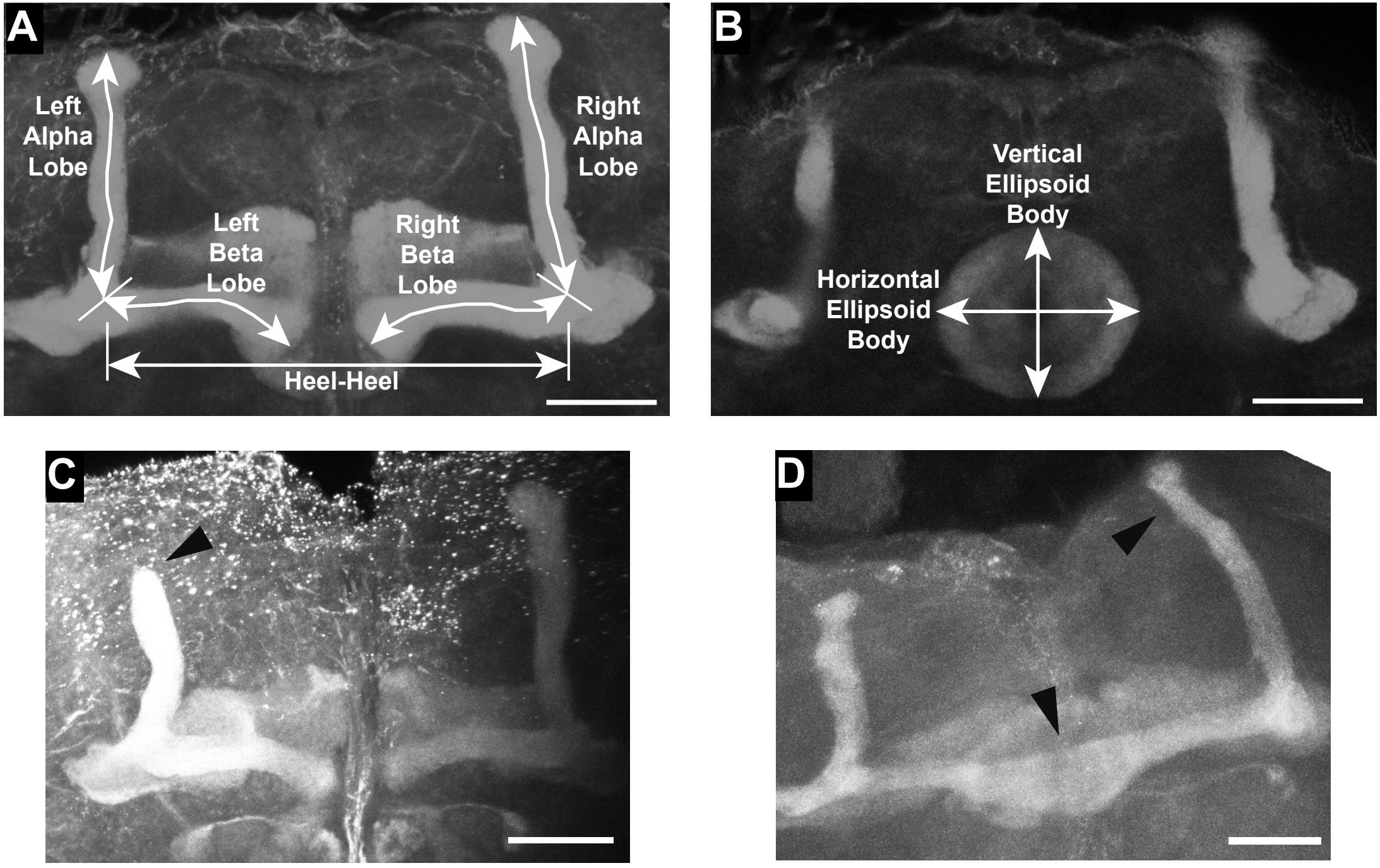
956 **File S1. Video of tapping behavior in a male fly with knockdown of *vtd* following a startle response.**

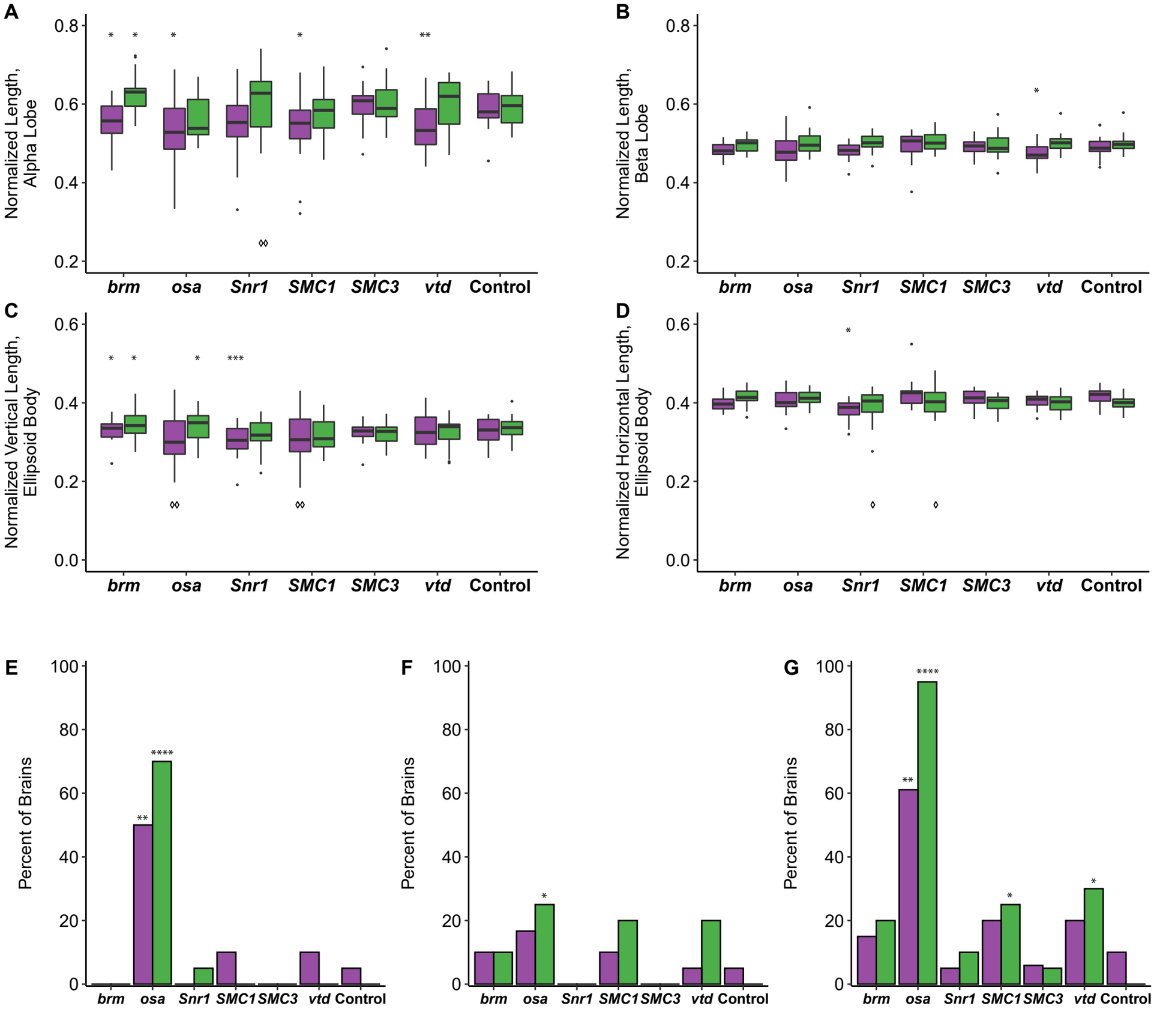
957

958 **File S2. Video of control male fly following a startle response.**


A
Time (s)
0
10
20
30
40
50
****
****
****
***
****
****
*
****
****
****
brm
osa
Snr1
SMC1
SMC3
vtd
Control
B
Percent Tapping
0
10
20
30
*
**
***
***
brm
osa
Snr1
SMC1
SMC3
vtd
Control


A
Total Activity (Counts)
5000
4000
3000
2000
1000
0
brm
osa
Snr1
SMC1
SMC3
vtd
Control
B
Proportion of Time Asleep, Night
1.00
0.75
0.50
0.25
0.00
C
Sleep Bout Count, Night
25
20
15
10
5
0


A
Left
Alpha
Lobe
Right
Alpha
Lobe
Left
Beta
Lobe
Right
Beta
Lobe
Heel-Heel
B
Vertical
Ellipsoid
Body
Horizontal
Ellipsoid
Body
C
D


A
Normalized Length, Alpha Lobe
0.8
0.6
0.4
0.2
brm
osa
Snr1
SMC1
SMC3
vtd
Control
B
Normalized Length, Beta Lobe
C
Normalized Vertical Length, Ellipsoid Body
0.6
0.4
0.2
0.0
D
Normalized Horizontal Length, Ellipsoid Body
E
Percent of Brains
100
80
60
40
20
0
F
G