Science Advances
AAAS

Supplementary Materials for

Mechanism of KMT5B haploinsufficiency in neurodevelopment in humans and mice

Sarah E. Sheppard *et al.*

Corresponding author: Holly A. F. Stessman, hollystessman@creighton.edu; Elizabeth J. Bhoj, bhoje@chop.edu

Sci. Adv. **9**, eade1463 (2023)
DOI: 10.1126/sciadv.ade1463

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Supplementary Text

Supplementary methods

RT-qPCR

From mouse brain RNA samples, cDNA was synthesized using the iScript™ Advanced cDNA Synthesis Kit for RT-qPCR (Bio-Rad, Hercules, CA USA). Gene expression was quantified using the SsoAdvanced™ Universal SYBR® Green Supermix Protocol (Bio-Rad, Hercules, CA USA) with gene-specific PrimeTime® primers (Coralville, IA, USA) or *Gapdh* primers (qMmuCID0020579; Bio-Rad). Reactions were set up using 10 µl SsoAdvanced™ Universal SYBR® Green Supermix, 1 µl forward and reverse primers, 1 µl cDNA template, and 18 µl nuclease free water. Samples were incubated on a Bio-Rad CFX Connect for 30 seconds at 95 °C followed by 40 cycles of 5 seconds at 95 °C and 15 seconds at 60 °C. Three biological replicates of each genotype were run in technical quadruplicate across all primer sets.

PrimeTime® assays

Gene target	**Assay name**
Grin2b	Mm.PT.58.42676841
Reln	Mm.PT.58.10165516
Plcb1	Mm.PT.56a.12599823
Wdfy3	Mm.PT.58.31111193

Total T4 enzyme-linked immunosorbent assay (ELISA)

Mouse blood serum and plasma were collected from adult (P53) animals via superficial temporal vein cheek bleed. Blood was collected in either PST tubes with lithium heparin (for plasma; ref# 365985; BD, Franklin Lakes, NJ, USA) or Vacutainer serum glass tubes (for serum; ref#366430, BD), centrifuged (@2000 x g for 8 minutes for plasma or @7000 x g for 90 seconds for serum), transferred into 1.5 mL tubes, and stored in a -80°C freezer. Total thyroxine (T4) concentration was determined using a commercially available enzyme-linked immunosorbent assay (ELISA) specific for T4 (Mouse TT4 ELISA Kit, MBS2600004, MyBioSource, San Diego, CA, USA). Briefly, 100 µL of serum or plasma was added to wells and incubated for 1.5 hours. The plate was washed, and a biotinylated antibody was added to each well and incubated for one hour. After another wash, avidin-peroxidase solution was added into each well and incubated for 30 minutes. The plate was washed and TMB substrate reagent added to each well followed by an incubation for 30 minutes before adding an acid stop solution. Absorbance was read using a Synergy LX (BioTek, Winooski, VT, USA) plate reader at 450 nm.

Species	Gene ID	Peptide ID	% identity (Protein)
Hu (Homo sapiens)	ENSG00000110066	ENSP00000305899 (885 aa)	52 %
Zf (Danio rerio)	ENSDARG00000041081	ENSDARP00000060190 (808 aa)	57 %

```
CLUSTAL W (1.81) multiple sequence alignment

Hu MKWLGESKNMVVNGRRNGGKLSNDHQQNQSKLQHTGKDTLKAGKNAVERRSNRCN---GN
Zf ---MGESKNMVLNGRRHGRKFSSNQPVSKSRLQNTQRSHLRQNKGSP--SVRRCSRRCGG
      :*******:****:* *:*.::  .:*:**:* :. *: .*.:     .**.   *.
Hu SGFEGQSRYVPSSGMSAKELCENDDLATSLVLDPYLGFQTHKMNTSAFPSRSSRHFSKSD
Zf APPEAERRHVPSSGMTAKELCEYDDLSTSLILDPYLGFQTHKMNT--------------
   :  *.: *:******:****** ***:***:**************
Hu SFSHNNPVRFRPIKGRQEELKEVIERFKKDEHLEKAFKCLTSGEWARHYFLNKNKMQEKL
Zf --------RFRPIKGRQRELREIIELFKKHDNLEKAFQALTSGDWTRHHFLNKTKSQEKL
           *********.**:*:** ***.::*****:.****:*:**:****.* ****
Hu FKEHVFIYLRMFATDSGFEILPCNRYSSEQNGAKIVATKEWKRNDKIELLVGCIAELSEI
Zf FKAHVFVYLRMFASDSGFEILSCNRYSSEQNGAKIVATKDWKRNDKIEHLVGCIAELSPS
   ** ***:******:*******.*****************:******** *********
Hu EENMLLRHGENDFSVMYSTRKNCAQLWLGPAAFINHDCRPNCKFVSTGRDTACVKALRDI
Zf EERMLLRHGENDFSVMYSTRKNCAQLWLGPAAFINHDCRPNCKFVSTGRDTACVKVLRDI
   **.**************************************************.****
Hu EPGEEISCYYGDGFFGENNEFCECYTCERRGTGAFKSRVGLPAPAPVINSKYGLRETDKR
Zf EPGEEISCYYGDGFFGENNEFCECYTCERRGTGAFKSKPGLPVEAPVINSKYGLRETDKR
   ***********************************: ***. ***************
Hu LNRLKKLGDSSKNSDSQSVSSNTDADTTQEKN--NATSNRKS-SVGVKKNSKSRTLTRQS
Zf LNRLKKLGESCRNSDSQSVSSNAEAD-SQEPTTVQTSLRKRTSQSCVKKHGEAKAVTRQT
   ********:*.:**********::** :** .  ::: .::: .  ***:.::::::***:
Hu MSRIPASSNST-SSKLTHINNSRVPKKLK-KPAKPLLSKIKLRNHCKRLEQKNASRKLEM
Zf LSSTPS---STSSSKRSQANISSLPKRLKSKPTQTLS---KGRRRCRGLWTKGSSRV-SA
```

Fig. S1.

KMT5B homology between human and zebrafish. Homology data exported from Ensembl. Boxed region indicates the SET protein domain.

Species	Gene ID	Peptide ID	% identity (Protein)
Hu (Homo sapiens)	ENSG00000110066	ENSP00000305899 (885 aa)	87 %
Ms (Mus musculus)	ENSMUSG00000045098	ENSMUSP00000109606 (883 aa)	87 %

```
CLUSTAL W (1.81) multiple sequence alignment

Hu MKWLGESKNMVVNGRRNGGKLSNDHQQNQSKLQ-HTGKDTLKAGKNAVERRSNRCNGNSG
Ms MKWLGDSKNMVVNGRRNGGKLSNDHQQNQSKLQQHSGKDTLKTGRNAVERRSSRCHGNSG
   *****:*************************** *:******:*:*******.**:****
Hu FEGQSRYVPSSGMSAKELCENDDLATSLVLDPYLGFQTHKMNTSAFPSRSSRHFSKSDSF
Ms FEGQSRYVPSSGMSAKELCENDDLATSLVLDPYLGFQTHKMNTSAFPSRSSRHISKADSF
   ******************************************************:**:***
Hu SHNNPVRFRPIKGRQEELKEVIERFKKDEHLEKAFKCLTSGEWARHYFLNKNKMQEKLFK
Ms SHNNPVRFRPIKGRQEELKEVIERFKKDEHLEKAFKCLTSGEWARHYFLNKNKMQEKLFK
   ************************************************************
Hu EHVFIYLRMFATDSGFEILPCNRYSSEQNGAKIVATKEWKRNDKIELLVGCIAELSEIEE
Ms EHVFIYLRMFATDSGFEILPCNRYSSEQNGAKIVATKEWKRNDKIELLVGCIAELSEIEE
   ************************************************************
Hu NMLLRHGENDFSVMYSTRKNCAQLWLGPAAFINHDCRPNCKFVSTGRDTACVKALRDIEP
Ms NMLLRHGENDFSVMYSTRKNCAQLWLGPAAFINHDCRPNCKFVSTGRDTACVKALRDIEP
   ************************************************************
Hu GEEISCYYGDGFFGENNEFCECYTCERRGTGAFKSRVGLPAPAPVINSKYGLRETDKRLN
Ms GEEISCYYGDGFFGENNEFCECYTCERRGTGAFKSRVGLPAPAPVINSKYGLRETDKRLN
   ************************************************************
Hu RLKKLGDSSKNSDSQSVSSNTDADTTQEKNNATSNRKSSVGVKKNSKSRTLTRQSMSRIP
Ms RLKKLGDSSKNSDSQSVSSNTDADTTQEKDNATSNRKSSVGVKKSSKSRALTRPSMPRVP
   ****************************:**************.****:*** **.*:*
Hu ASSNSTSSKLTHINNSRVPKKLKKPAKPLLSKIKLRNHCKRLEQKNASRKLEMGNLVLKE
Ms AASNSTSPKLVHTNNPRVPKKLRKPAKPLLSKIRLRNHCKRLDQKSASRKLEMGSLVLKE
```

```
Hu PKVVLYKNLPIKKDKEPEGPAQAAVASGCLTRHAAREHRQNPVRGAHSQGESSPCTYITR
Ms PKVVLYKNLPIKKEREPEGPAHAAVGSGCLTRHAAREHRQNHGRGAHSQGDSLPCTYTTR
   *************::******:***.**************  *******:* **** **
Hu RSVRTRTNLKEASDIKLEPNTLNGYKSSVTEPCPDSGEQLQPAPVLQEEELAHETAQKGE
Ms RSLRTRTGLKETTDIKLEPSPLDGYKNGILEPCPDSGQ--QPTPEVLEE-LAPETAHREE
   **:****.***::******..*:***..: *******:  **:* : ** ** ***:: *
Hu A--KCHKSDTGMSKKKSRQGKLVKQFAKIEESTPVHDSPGKDDAVPDLMGPHSDQGEHSG
Ms ASQECPKNDSCLSRKKFRQVKPVKHLAKTEDCSPEHSFPGKDGL-PDLPGSHPDQGEPSG
   *  :* *.*: :*:** ** * **::** *:.:* *. ****.  *** *.*.**** **
Hu TVGVPVSYTDCAPSPVGCSVVTSDSFKTKDSFRTAKSKKKRRITRYDAQLILENNSGIPK
Ms TVRVPVSHTDSAPSPVGCSVVAPDSF-TKDSFRTAQSKKKRRVTRYDAQLILENSSGIPK
   ** ****:**.**********:.*** ********:******:***********.*****
Hu LTLRRRHDSSSKTNDQENDGMNSSKISIKLSKDHDNDNNLYVAKLNNGFNSGSGSSSTKL
Ms LTLRRRHDSSSKTNDHESDGVNSSKISIKLSKDHDSDSNLYVAKLSNGVSAGPGSSSTKL
   ***************:*.**:**************.*.*******.**..:*.*******
Hu KIQLKRDEENRGSYTEGLHENGVCCSDPLSLLESRMEVDDYSQYEEESTDDSSSSEGDEE
Ms KIQLKRDEESRGPCAEGLHENGVCCSDPLSLLESQMEVDDYSQYEEDSTDESSSSEGEEE
   *********.**. :*******************:***********:***:******:**
Hu EDDYDDDFEDDFIPLPPAKRLRLIVGKDSIDIDISSRRREDQSLRLNA
Ms EEDCEDDFDDDFIPLPPAKRLRLIVGKDSIDIDISSRRREDQSLRLNA
   *:* :***:************************************
```

Fig. S2.

KMT5B homology between human and mouse. Homology data exported from Ensembl. Boxed region indicates the SET protein domain.

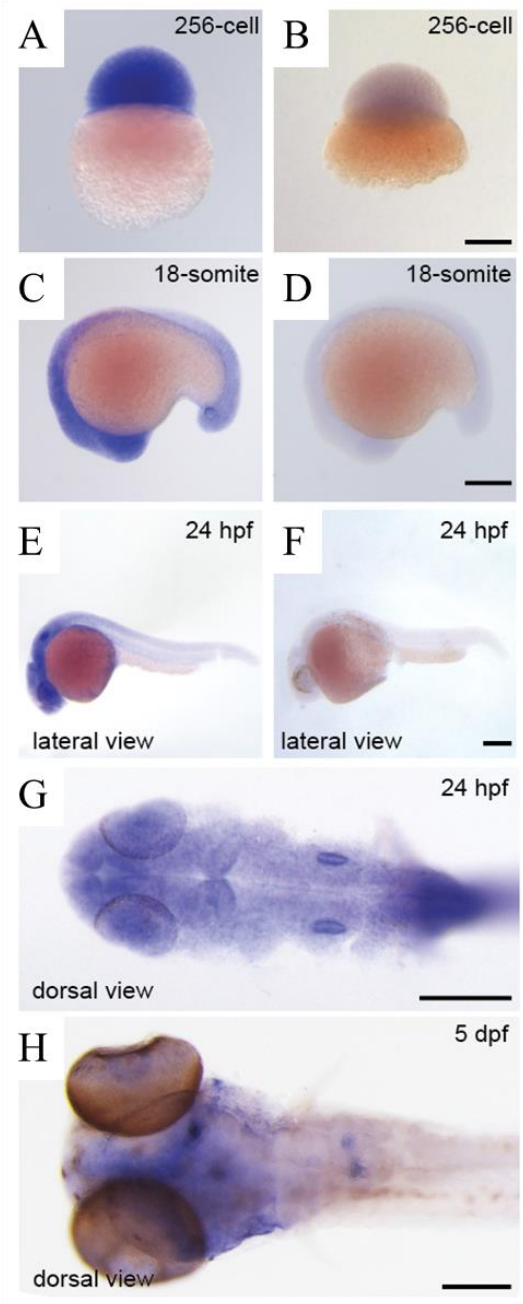

Fig. S3.

Expression of *zKmt5b* during zebrafish embryonic development. Whole mount *in situ* hybridization of *kmt5b* in zebrafish embryos and larvae using antisense (**A**, **C**, **E**, **G**) and sense (**B**, **D**, **F**, **H**) probes. Scale bar = 0.2mm. dpf = days post fertilization.

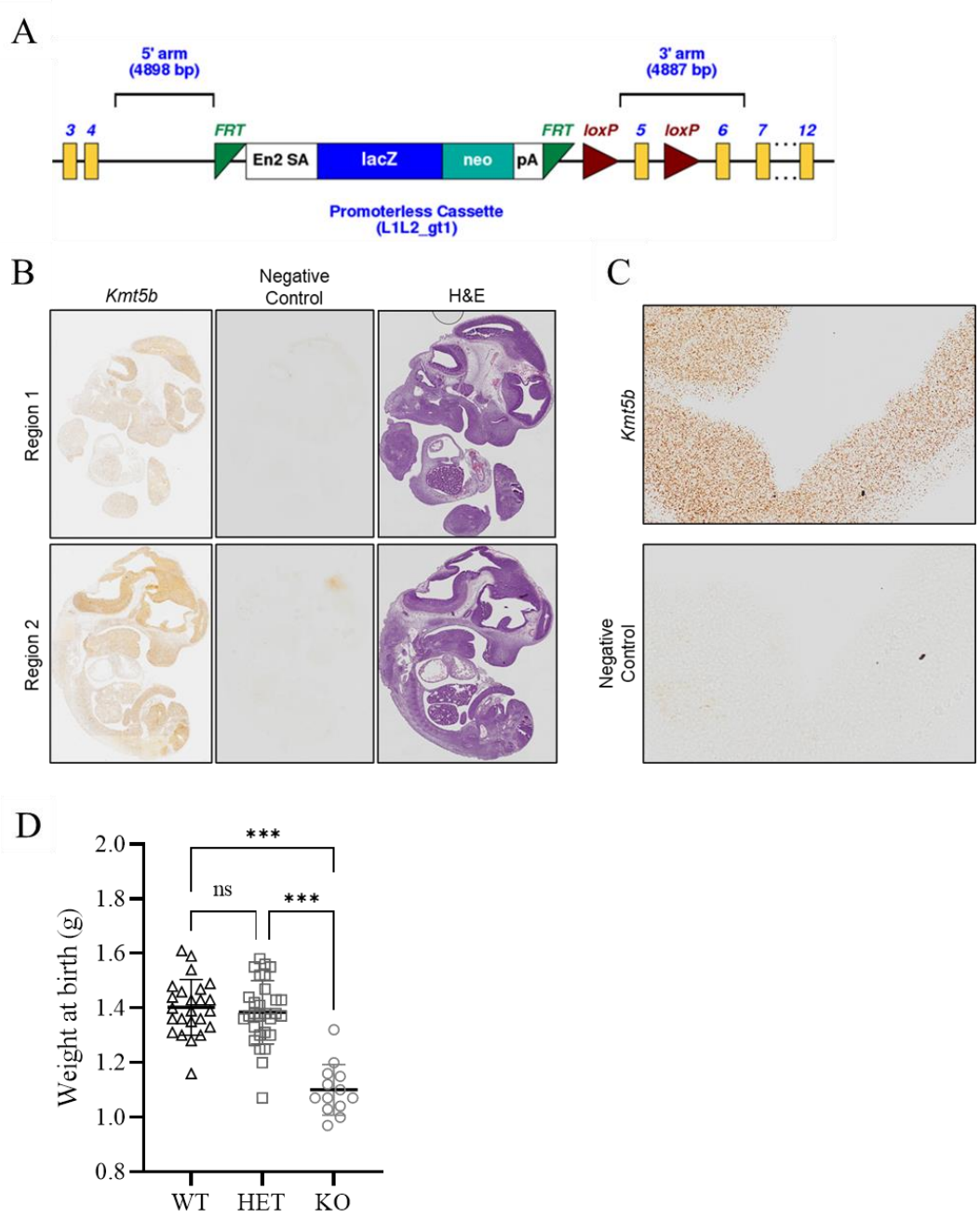

Fig. S4.

***Kmt5b* mouse model and RNAscope controls**. (**A**) Mouse model gene trap allele. (**B**) Two different regions of a fixed and serially sectioned E12.5 mouse embryo are shown stained with either a *Kmt5b* probe (left), negative control probe (middle), or hematoxylin and eosin (H&E) stain (right). (**C**) *Kmt5b* versus negative control probes shown at higher resolution (20X). (**D**) C57BL/6JInv strain background weight at birth data when carrying the gene trap allele in either one (HET) or two (KO) copies compared to wild-type (WT) littermates. ns = not significant; ***p<0.001.

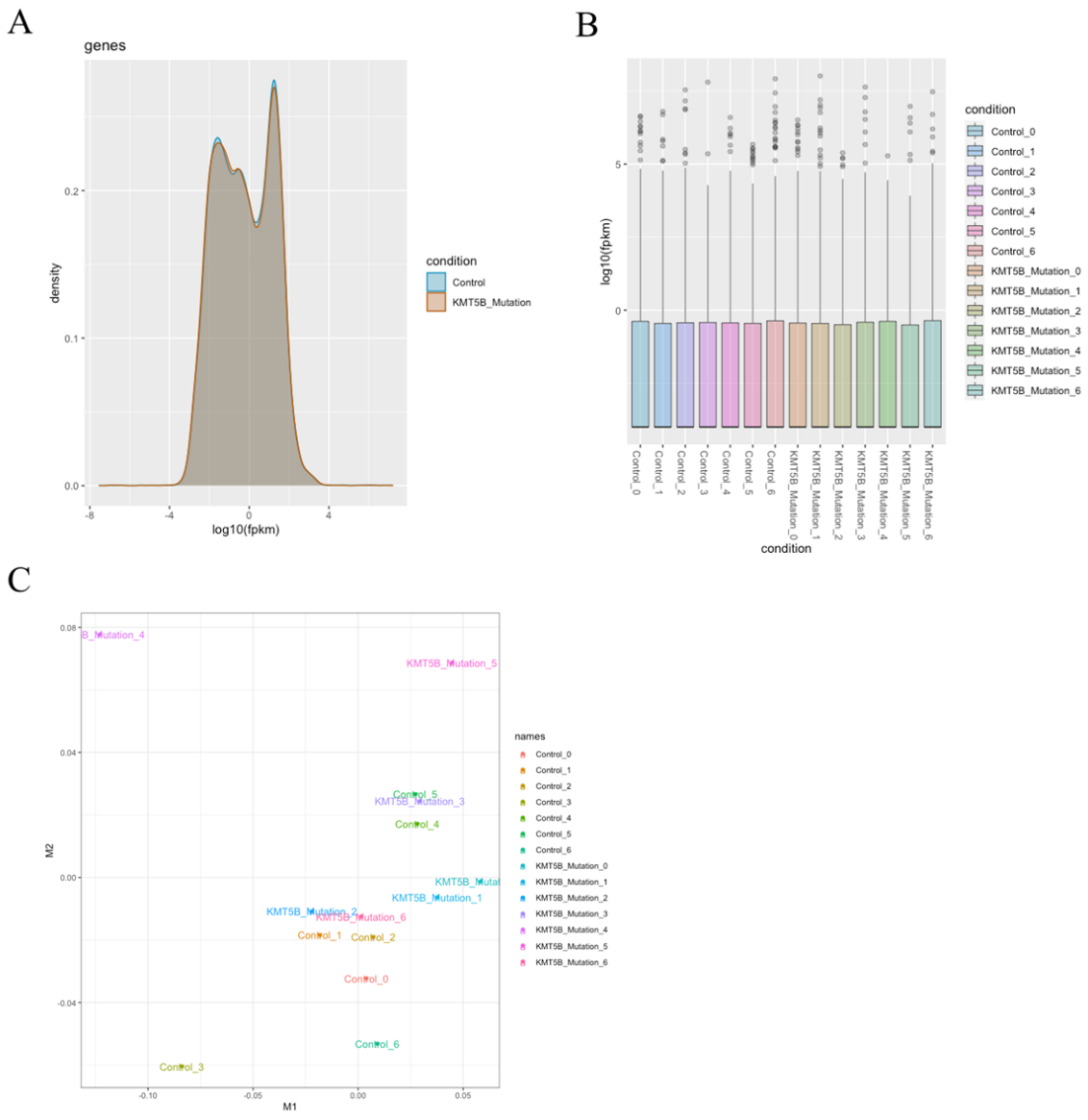

Fig. S5.

Quality control of human RNA-seq data. (**A**) Smoothed density plot of $\log_{10}$ fragments per kilobase of transcript per million fragments mapped (FPKM) per genotype (**B**) further broken down by individual sample as boxplots. (**C**) Weighted principle component analysis (i.e., MSA) plot by sample.

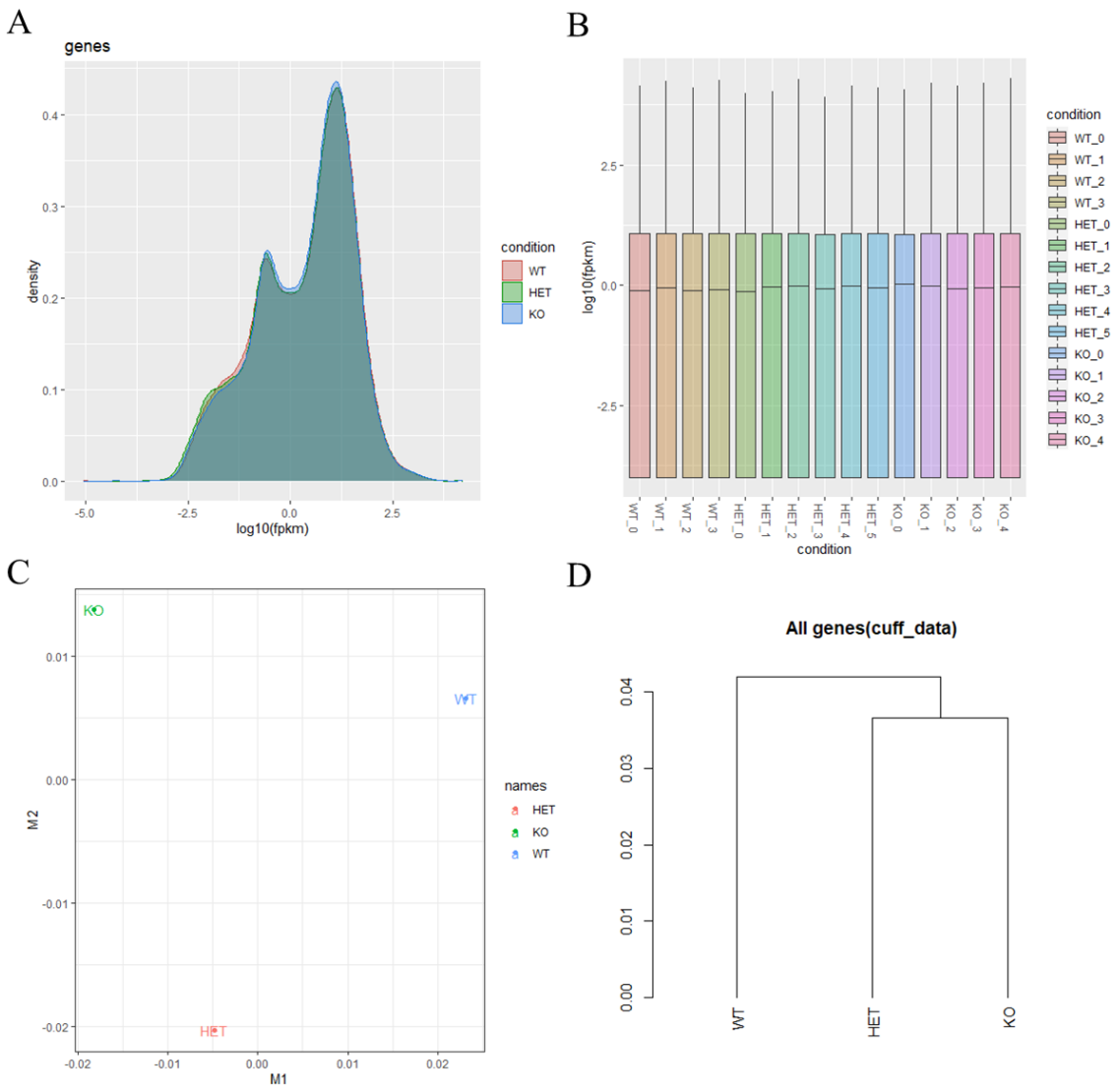

Fig. S6.

Quality control of mouse RNA-seq data. (**A**) Smoothed density plot of $\log_{10}$ fragments per kilobase of transcript per million fragments mapped (FPKM) per genotype (**B**) further broken down by individual sample as boxplots. (**C**) Weighted principle component analysis (i.e., MSA) plot and (**D**) dendrogram by genotype. WT: wild-type; HET: heterozygous; KO: knockout.

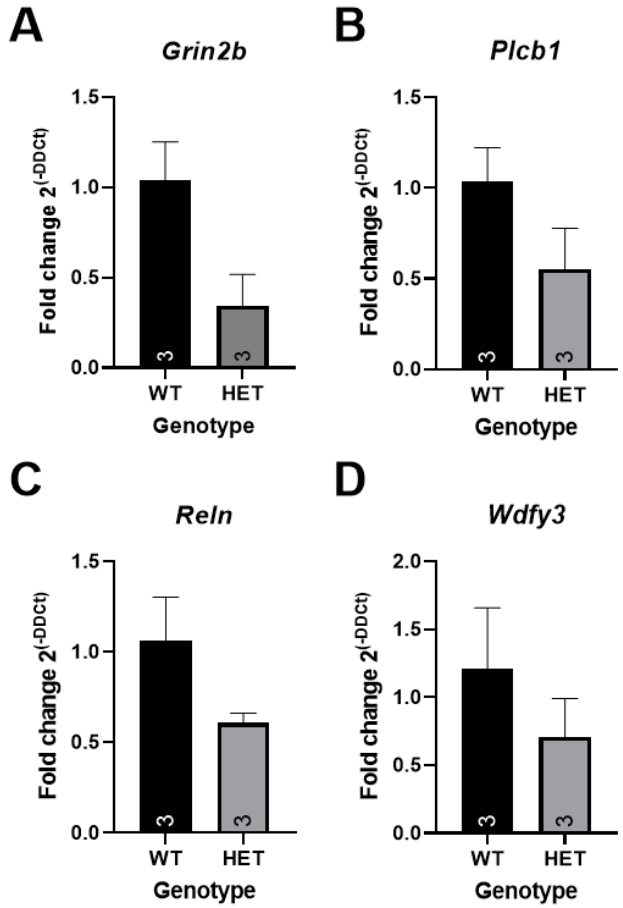

Fig. S7.

Validation of select RNA-seq results. Quantitative RT-PCR data shown for (**A**) *Grin2b*, (**B**) *Plcb1*, (**C**) *Reln*, and (**D**) *Wdfy3* in mouse WT and HET brains (E14.5). WT: wild-type; HET: heterozygous. Error bars show ± SEM; number of biological replicates are shown within each bar.

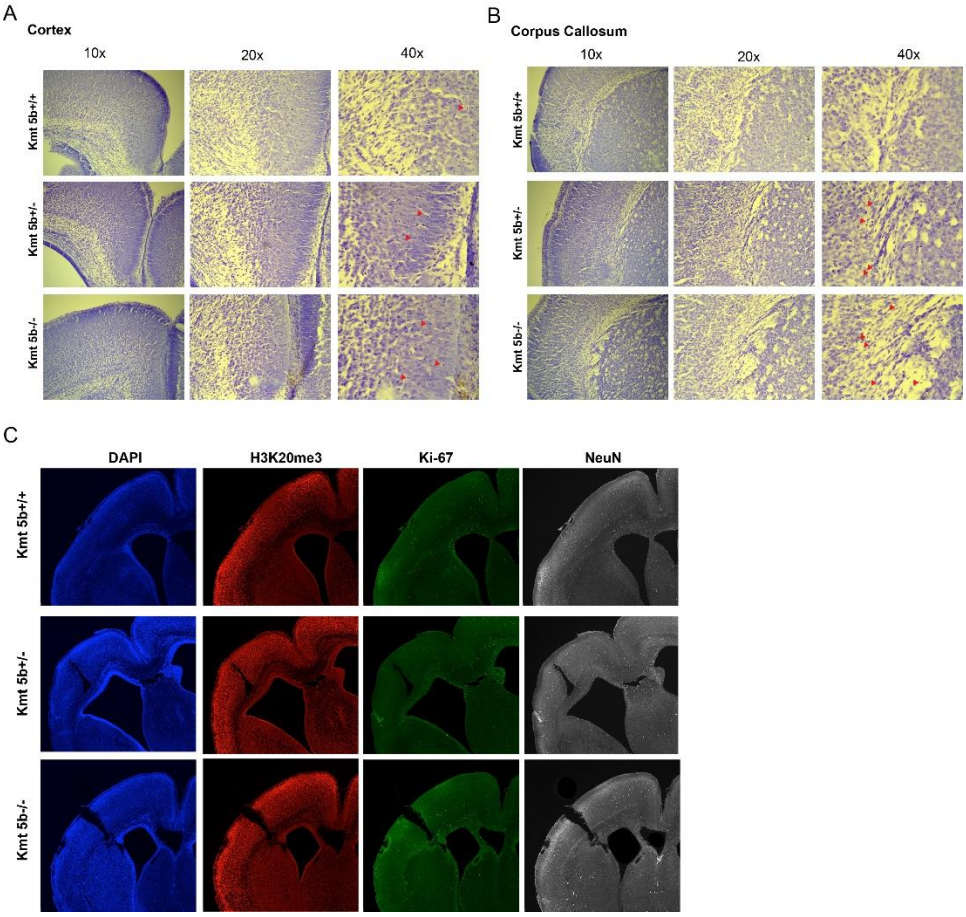

Fig. S8.

Histology in mouse brain. (A-B) Nissl staining was performed on the cortex and corpus collosum from P0 mouse pups of the three different genotypes. Increased pyknotic cells were observed in both the heterozygous and knockout mice compared to the wild-type mice. (**C**)There is no significant difference in H3K20me3, cell proliferation or mature neurons in P0 mouse brains. Brain sections were stained with DAPI, H3K20me3, Ki-67 and NeuN from Kmt5b$^{+/+}$ (WT), Kmt5b$^{+/-}$ (HET), and Kmt5b$^{-/-}$ (KO) mouse pups and imaged on a confocal microscope.

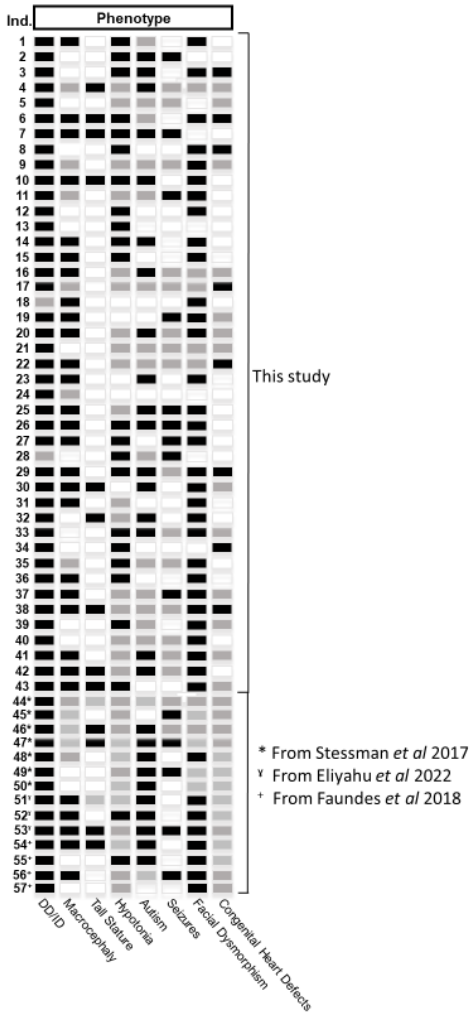

Fig. S9.

Comparison of *KMT5B* patient phenotypes across studies. Ten additional individuals have been described in detail in two other papers (Idv 44-57). Black indicates the feature is present, white indicates the feature is absent, and grey indicates that the status of the feature is unknown or not applicable. Previous studies did not look at congenital heart defects and did not evaluate for hypotonia. All individuals described had developmental delay and/or intellectual disability except for two where it is unknown. Individual 18 was not evaluated until she was an adult. She does not have intellectual disability, but no developmental information is available. Individual 28 died as an infant and could not be evaluated for developmental delay. 62% (29/47) of the combined cohort had macrocephaly and 24% (13/55) had tall stature. 83% of previously reported patients had autism but only 68% (28/41) of the combined cohort have an autism diagnosis. This could be because previous studies targeted autism cohorts while ours did not use an autism diagnosis as a selection criterion. 34% (14/41) had seizures. 83% (39/47) had facial dysmorphism.

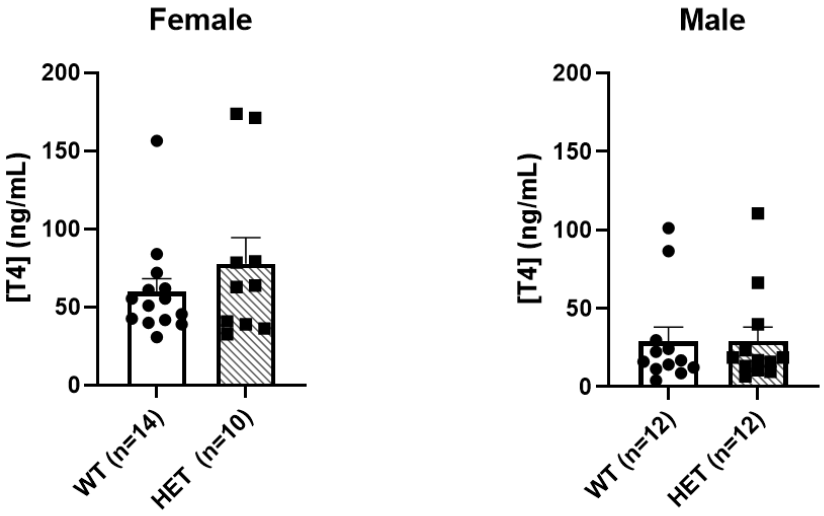

Fig. S10.

Circulating T4 levels do not differ between genotypes in adults. Plots show the level of total circulating T4 in adult animals as measured by enzyme-linked immunosorbent assay (ELISA) in (**left**) females and (**right**) males. Bars represent the average of biological replicates (n) ± SEM. WT: wild-type; HET: heterozygous.

Table S1.
***KMT5B* patient phenotypes.**
See separate spreadsheet.

Variant	ΔΔG	All contacts	SAM contacts	ZN contacts
S73F	-0.853	3		
A74T	-1.173	5		
T85I	-0.143	10		
G194E	-0.252	9		
C200R	-1.404	43	37	
D222Y	0.141	6		
E233K	-0.045	5		
N248D	-0.896	113		
I271M	-0.897	74	64	
H273R	-0.92	192	161	12
I297V	-1.601	13		
E302K	-0.971	7	3	

Table S2.
***In silico* stability computations for missense variants.**

Main Tissues	Marker #	E11.5*	E12.5*	E13.5*	E14.5*	E15.5-E16*
Forebrain	1	++	+++	+++	+++	+++
Neopallial cortex						+++
Intermediate zone						+
Midbrain	2	++	+++	+++	+++	+++
Hindbrain	3	++	+++	+++	+++	+++
Mandibular component of first branchial arch	4	+				
Second branchial arch	5	+				
Trabeculated wall of common ventricular chamber of heart (future left ventricle)	6	+				
Wall of left atrial chamber of heart (heart tissue)	7	+		+		++
Hepatic primordium (liver)	8	+	+	+	+	++
Gut, intestinal tissue	9	+				
Hindlimb	10	+			++	
Optic tissue	11	+				
Spinal cord	12			+++		++
Vertebral column	13			+		++
Muscles of the tongue	14			+		
Lower jaw	15			+		
Tail	16			+		++
Lung	17				++	
Primordium of follicle of vibrissa	18				+++	
Lung tissue	19					++
Tongue/lower jaw	20					++
Intestinal system	21					++

Table S3.
Embryonic expression of *Kmt5b*.

Brain Regions	P1-2 Brain**	P10 Brain**	P56 Brain Adult
Cortex	++	++	
Retrosplenial area			++
Retrosplenial area, ventral part, layer 2			+++
Secondary, primary motor areas			++
Somatosensory areas			++
Auditory cortex			++
Temporal cortex			++
Ectorhinal area			++
Perirhinal area			++
Entorhinal area			++
Piriform area	++	++	++
Piriform area, pyramidal layer	+++	+++	+++
Corpus callosum	+	+	+
Medial habenula			+++
Choroid plexus			+++
Expressed in all regions of the following structures			
Amygdala nuclei	++	++	++
Caudate/Putamen	++	++	++
Hippocampus	++	++	++
Dentate gyrus, granule cell layer	+++	+++	+++
Dentate gyrus, molecular layer	+	+	+
Field CA1, CA2 and CA3 pyramidal layer	+++	+++	+++
Stratum oriens			+
Stratum radiatum			+
Stratum lacunosum-moleculare			+
Thalamus	++	++	++
Hypothalamus	++	++	++
Brainstem	++	++	++
Cerebellum	++	++	
Molecular layer			+
Granule layer			+++

Table S4.
Postnatal expression of *Kmt5b*.

Table S5.

Human RNA-seq dataset.

See separate spreadsheet.

Table S6.

Human enriched brain diseases and functions.

See separate spreadsheet.

Table S7.

Human predicted upstream regulators.

See separate spreadsheet.

Table S8.

Mouse RNA-seq dataset.

See separate spreadsheet.

Table S9.

Mouse enriched brain diseases and functions.

See separate spreadsheet.

Table S10.

Mouse predicted upstream regulators.

See separate spreadsheet.