Supplemental Material

Females on standard chow
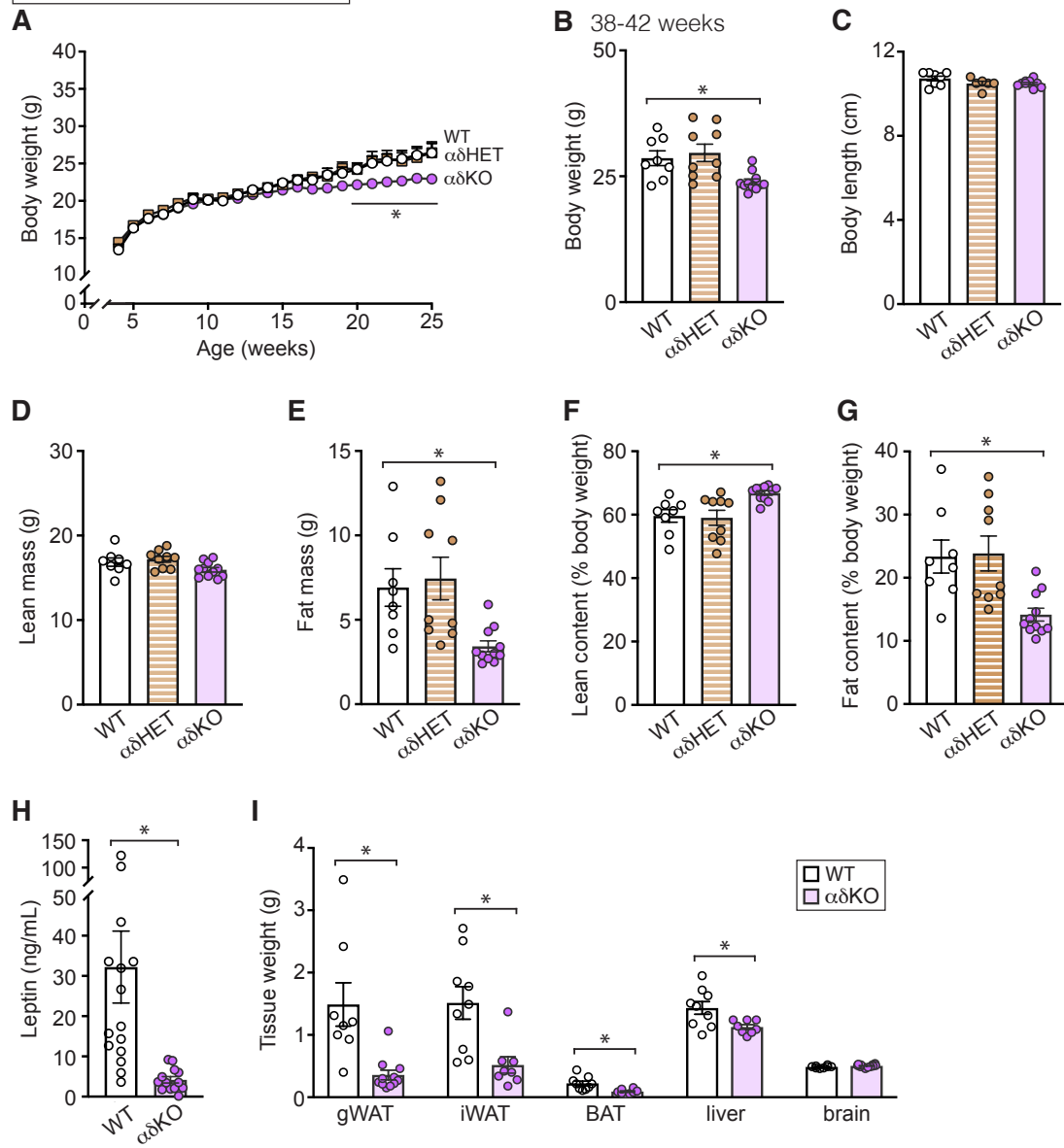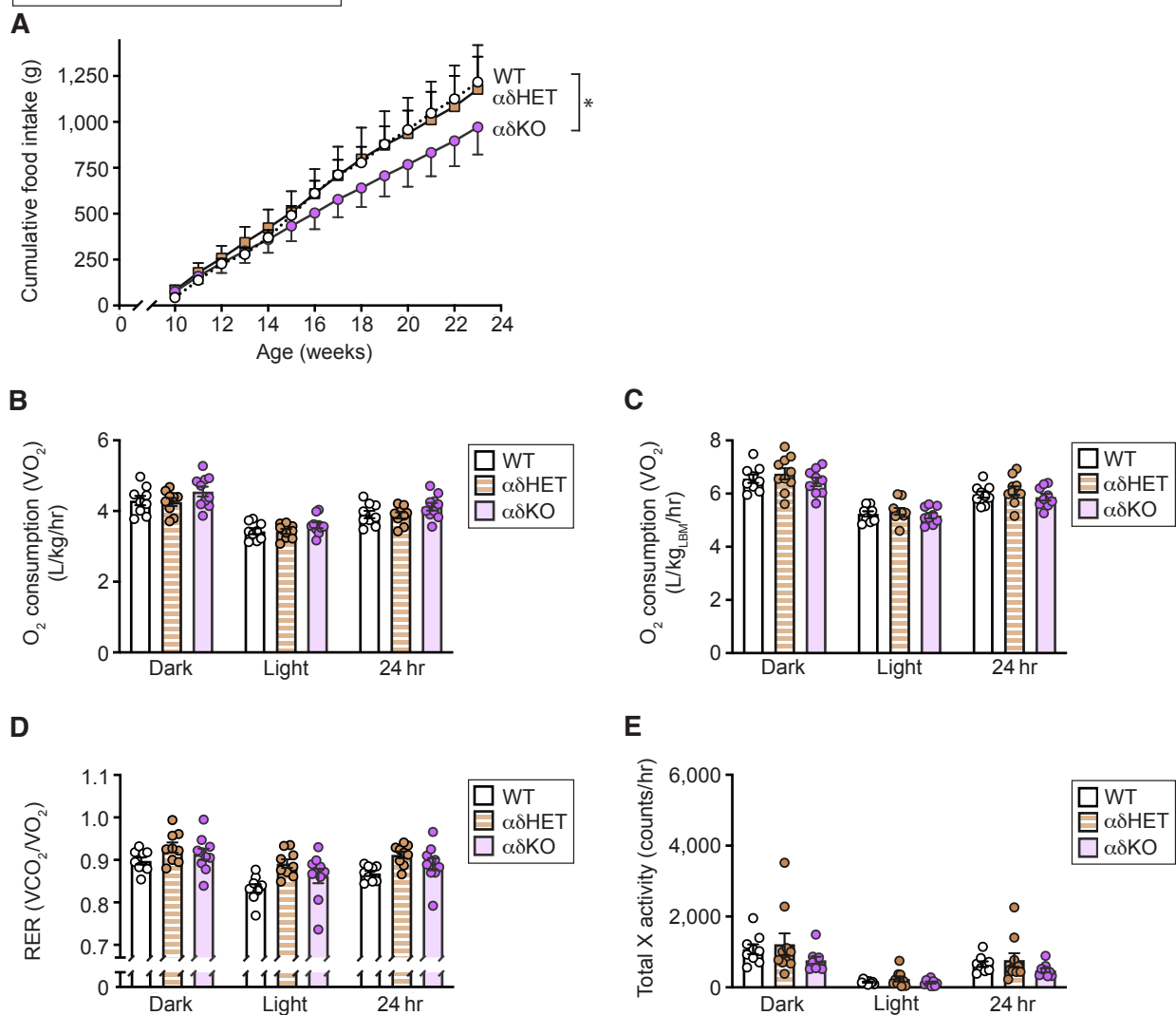
Supplementary Figure 1. Female αδKO mice fed standard chow exhibit reduced body weight, fat content, circulating leptin levels, adipose tissue weight, and liver weight. **(A)** Body weight was measured weekly from 4 to 25 weeks (n: WT = 6-8; αδHET = 7-9; αδKO = 11-13). **(B)** Body weight was measured between 38-42 weeks of age (n: WT = 8; αδHET = 9; αδKO = 11). **(C)** Body length was measured from nose to anus in 38-to-42-week-old mice (n: WT = 8; αδHET = 6; αδKO = 10). **(D-G)** Body composition of 38-to-42-week-old mice was analyzed by NMR between 9_{AM}-11_{AM} (n: WT = 8; αδHET = 9; αδKO = 11). **(H)** Blood was collected from 38-to-42-week-old mice fed *ad libitum*, and serum leptin levels were measured by ELISA (n: WT = 15; αδKO = 13). **(I)** Tissues of 38-to-42-week-old mice were dissected and weighed (n: WT = 8-10; αδKO = 8-12). gWAT, gonadal white adipose tissue; iWAT, inguinal white adipose tissue; BAT, brown adipose tissue. **Statistics: A**, two-way repeated measures ANOVA (weeks 7 to 25); **B-G**, one-way ANOVA; **H-I**, unpaired, two-tailed Student's t-test. * $P < 0.05$. Data are means ± SEM.
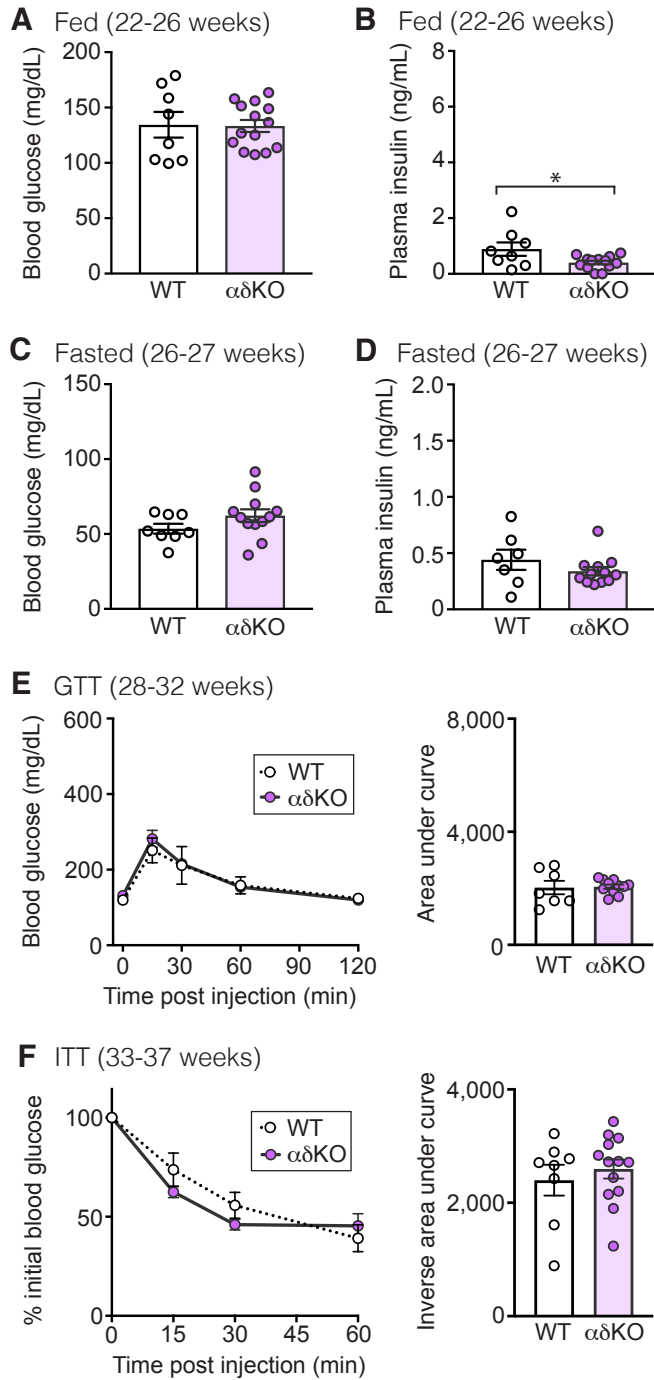
Females on standard chow

A

B

C

D

E

Supplementary Figure 2. Female αδKO mice fed standard chow exhibit decreased food consumption but normal energy expenditure. **(A)** Food intake was measured weekly from 10 to 23 weeks (n: WT = 7; αδHET = 8; αδKO = 11). **(B-E)** Energy expenditure was measured in 10-to-12-week-old mice by CLAMS (n: WT = 9; αδHET = 10; αδKO = 9). Parameters included: **(B)** oxygen consumption corrected for total body weight or **(C)** lean body mass; **(D)** respiratory exchange ratio (RER); and **(E)** total motor activity in the X dimension. RER was calculated as VCO_2/VO_2. The final 24 hours of recordings are presented. **Statistics: A**, linear regression analysis (weeks 10 to 23); **B-E**, one-way ANOVA. * $P < 0.05$. Data are means ± SEM.
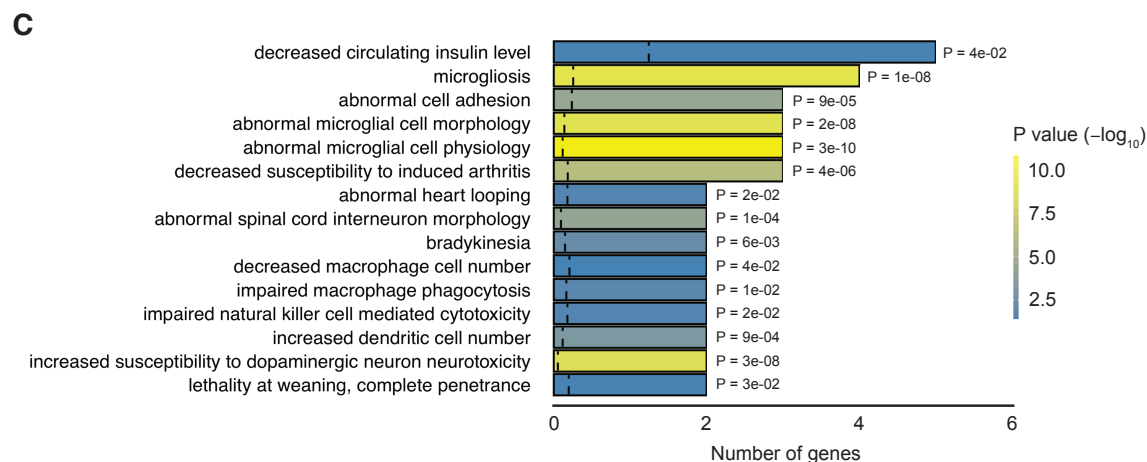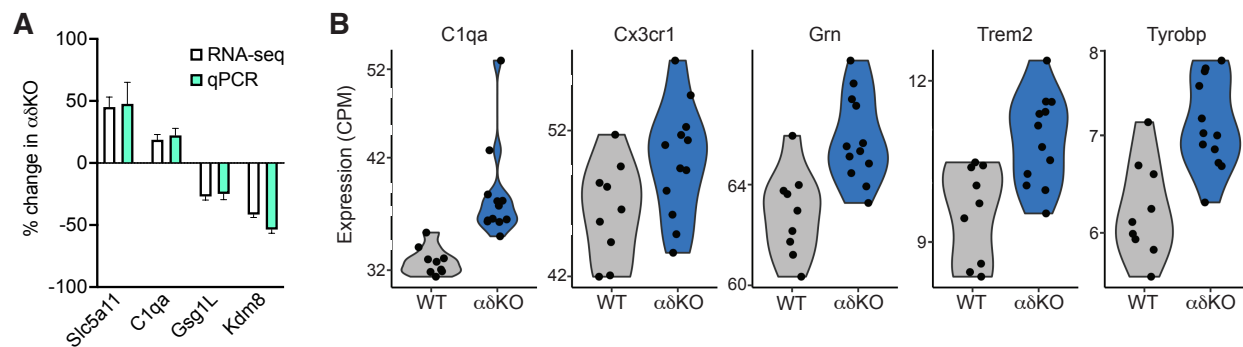
Females on standard chow

A Fed (22-26 weeks)

B Fed (22-26 weeks)

C Fasted (26-27 weeks)

D Fasted (26-27 weeks)

E GTT (28-32 weeks)

F ITT (33-37 weeks)

Supplementary Figure 3. Older female αδKO mice exhibit modestly improved glucose homeostasis. Blood was collected from 22-to-26-week-old mice fed *ad libitum* and **(A)** blood glucose levels were measured by glucometer (n: WT = 8; αδKO = 14) and **(B)** plasma insulin levels were measured by ELISA (n: WT = 8; αδKO = 13). Blood was collected from 26-to-27-week-old mice fasted overnight (16 hours) and **(C)** blood glucose levels were measured by glucometer (n: WT = 8; αδKO = 12) and **(D)** plasma insulin levels were measured by ELISA (n: WT = 8; αδKO = 12). **(E)** Glucose tolerance tests (GTT) were performed on 28-to-32-week-old-mice (i.p. glucose, 2 g/kg) (n: WT = 7; αδKO = 11). Area under curve for each animal was calculated using a baseline of y = 0. **(F)** Insulin tolerance tests (ITT) were performed on 33-to-37-week-old mice (i.p. insulin, 1 unit/kg) (n: WT = 8; αδKO = 13). Data are reported as % initial blood glucose values. Inverse area under curve for each animal was calculated using a baseline of y = 100. **Statistics**: **A-F**, unpaired, two-tailed Student's t-test. * P < 0.05. Data are means ± SEM.

Supplemental Material

A

B C1qa Cx3cr1 Grn Trem2 Tyrobp

C

decreased circulating insulin level — P = 4e-02
microgliosis — P = 1e-08
abnormal cell adhesion — P = 9e-05
abnormal microglial cell morphology — P = 2e-08
abnormal microglial cell physiology — P = 3e-10
decreased susceptibility to induced arthritis — P = 4e-06
abnormal heart looping — P = 2e-02
abnormal spinal cord interneuron morphology — P = 1e-04
bradykinesia — P = 6e-03
decreased macrophage cell number — P = 4e-02
impaired macrophage phagocytosis — P = 1e-02
impaired natural killer cell mediated cytotoxicity — P = 2e-02
increased dendritic cell number — P = 9e-04
increased susceptibility to dopaminergic neuron neurotoxicity — P = 3e-08
lethality at weaning, complete penetrance — P = 3e-02

D

Supplementary Figure 4. Transcriptional changes in αδKO hypothalami are significantly associated with terms linked to microglial function. **(A)** mRNA analyzed by RNA-seq was assessed by qPCR for expression levels of indicated genes (n = 4 animals per genotype). Graph depicts the % change in gene expression in αδKO compared to WT mice. **(B)** Violin plots depict expression of genes identified by Gene Ontology analysis. CPM, counts per million. **(C)** Bar graph displays mouse phenotypes that significantly associate with genes that have statistically significant differential expression in αδKO mice. The dashed vertical line on each bar corresponds to the frequency of finding that phenotype associated with all genes. **(D)** Heat map depicts cell type-specific expression of genes with statistically significant differential regulation in αδKO mice. POPC, proliferating oligodendrocyte precursor cell; Astro, astrocyte; MO, myelinating oligodendrocyte; OPC, oligodendrocyte precursor cell; IMO, immature myelinating oligodendrocyte; Epith, epithelial cell; SCO, subcommissural organ cells; Ependy, ependymocyte; Tany, tanycyte; Micro, microglia; Macro, macrophages. **Statistics**: **B-D**, see Supplementary Materials. Data are means.

Supplemental Material

A

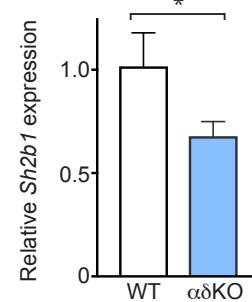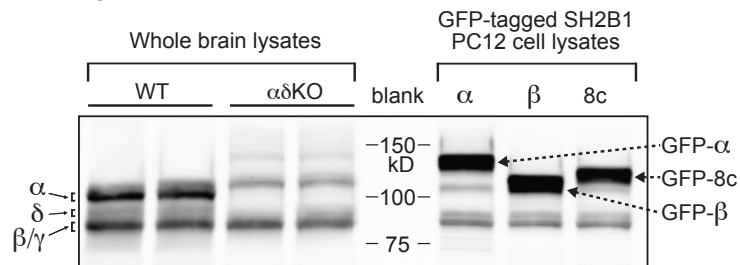| Ensembl transcript	NCBI variant	SH2B1 isoform
201	2	γ
202	1	β
203	4	β
204	5	δ
205	7	α
206	6	δ
207	–	lncRNA
208	–	73aa
209	8	NMD
–	–	α
–	–	β
–	3‡	β
–	–	γ
–	–	8c
–	–	8c
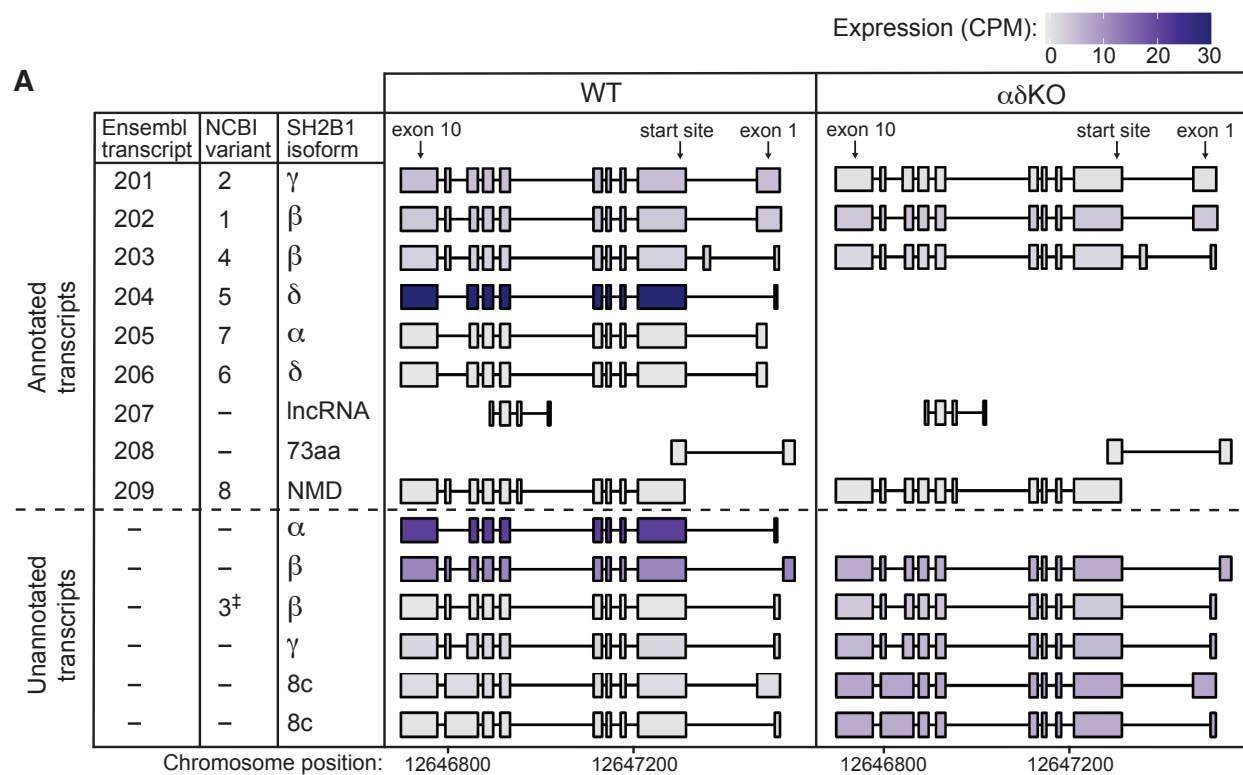
B Annotated transcripts only
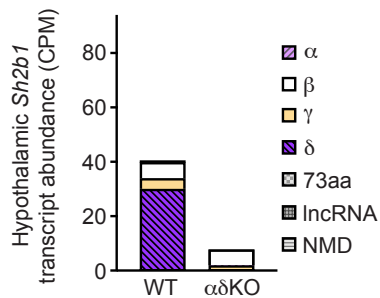
C Annotated & unannotated transcripts
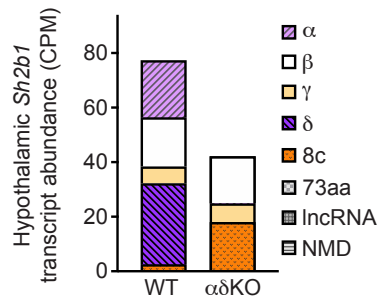
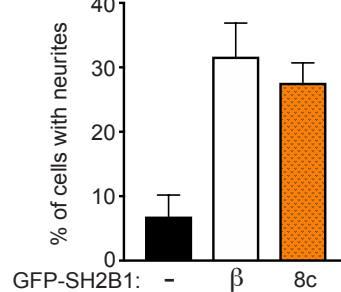**D** qPCR

E IB: α SH2B1

F

Supplementary Figure 5. RNA-seq data analysis reveals potential novel transcripts of *Sh2b1*.

(A) Schematic of *Sh2b1* transcripts and their expression levels in hypothalami of 10-to-12-week-old male mice as identified by RNA-seq (n: WT = 9; αδKO = 12). The dotted horizontal line separates annotated (top) versus unannotated (bottom) Ensembl transcripts. *Sh2b1*8c transcripts have an extended 8th exon. Validated NCBI variant numbers are listed. ‡ SH2B1 isoform assignment based on transcript schematic. CPM, counts per million; lncRNA, long non-coding RNA; aa, amino acids; NMD, nonsense-mediated decay. **(B)** *Sh2b1* transcript abundance was calculated by RNA-seq data analysis using annotated Ensembl transcripts only or **(C)** both annotated and unannotated Ensembl transcripts. Data for α, β, γ, and δ transcripts in **C** are the same as in Figure 1F. **(D)** Levels of total *Sh2b1* mRNA were measured by qPCR using an aliquot of the same hypothalamic RNA as in Fig. 8A (n: WT = 9; αδKO = 12). **(E)** Proteins in brain tissue homogenates from WT and αδKO adult male mice and lysates of PC12 cells expressing GFP-tagged SH2B1α, β, or 8c were immunoblotted with antibody to SH2B1 (α SH2B1). The migration of molecular weight standards is shown in the center of the blot. The expected migration of the different isoforms is indicated on the left (endogenous) and right (GFP-tagged). IB, immunoblot. **(F)** PC12 cells transiently expressing GFP (-), GFP-SH2B1β (β), or GFP-SH2B18c (8c) were treated with 25 ng/mL NGF for 2 days, at which point neurite outgrowth was assessed. GFP-positive cells were scored for the presence of neurites at least 2 times the length of the cell body (n = 3 experiments, 200-300 cells/condition/experiment). **Statistics**: **D**, unpaired, two-tailed Student's t-test. * $P < 0.05$. Data are means ± SEM.
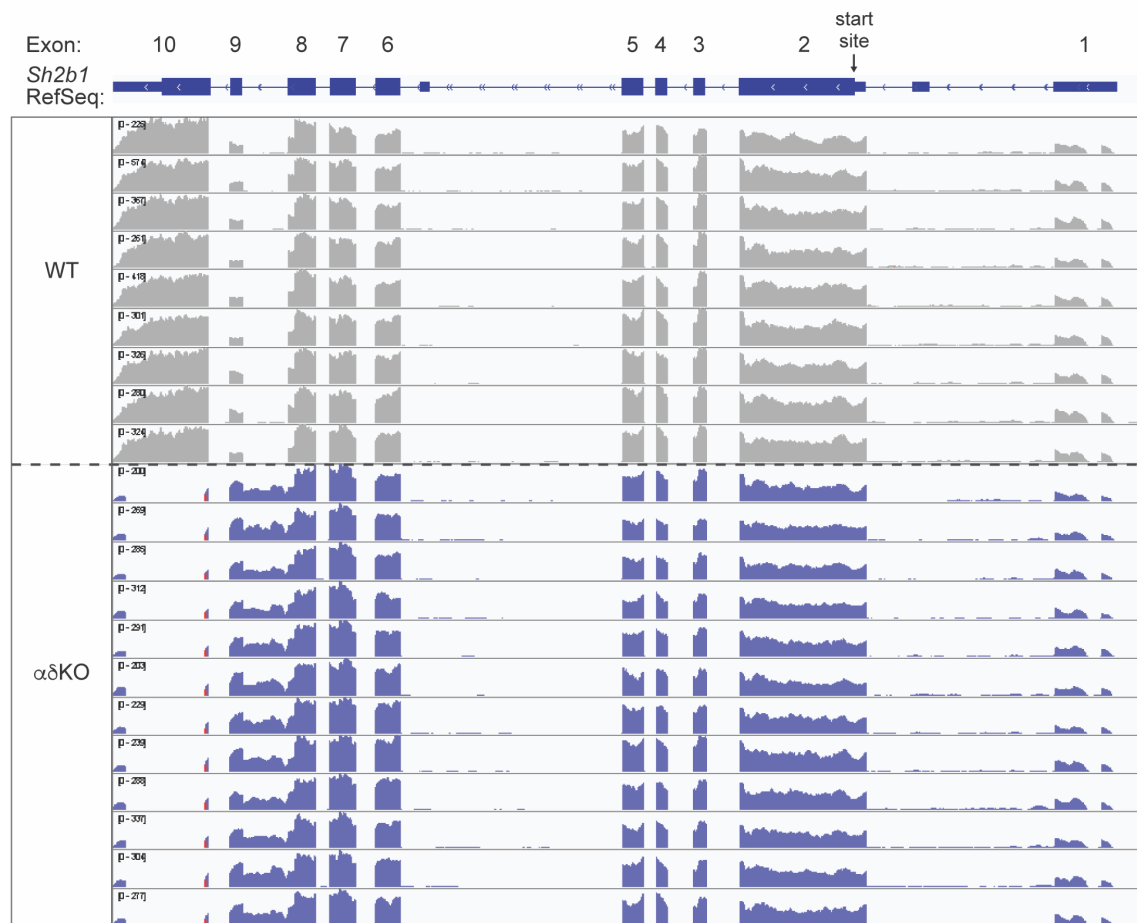
Supplemental Material

Supplementary Figure 6. RNA-seq reads from WT and αδKO hypothalami mapped to mouse *Sh2b1* Reference Sequence (RefSeq). In the *Sh2b1* RefSeq schematic at the top of the figure, blue rectangles represent exons; non-coding regions have less height than coding regions. The height of the reads tracks from individual WT (light gray) and αδKO mice (dark blue) depicts the level of mRNA expression at each nucleotide of the RefSeq for each mouse. Note that αδKO reads tracks are lacking expression for most of exon 10 as expected. However, additional expression is present between exons 8 and 9, which was unexpected and corresponds to the predicted *Sh2b1*8c transcripts (Supplementary Fig. 5). Numbers in brackets on the left side of each track (e.g. [0-225]) are the range of expression values for that track.

Supplemental Material

Supplementary Table 1. Primers for genotyping

Primer	Direction	Sequence (5' → 3')
WT	Forward	CAAAGGGGAGGTCACCATAAGAACTCAC
WT	Reverse	GTGGGCAGGTATCTCACACAAATGAGTA
Edit	Forward	GGGAATGTGCAGAACTGGACCCA
Edit	Reverse	GCGAGTGACTGTGTAACGGAGCA

Supplementary Table 2. Primers for *Sh2b1δ* and β-*actin* cDNA

Primer	Direction	Sequence (5' → 3')
Sh2b1δ	Forward	GTGCACCCAAGAAGTGAGAAC
Sh2b1δ	Reverse	CTTGTACCAACATACACACCCTTG
β-*actin*	Forward	CGTCTGGACCTGGCTGGCCGGGACC
β-*actin*	Reverse	CTAGAAGCATTTGCGGTGGACGATG

Supplementary Table 3. TaqMan Gene Expression Assays from Applied Biosystems

Gene	TaqMan Gene Expression Assay
Agrp	Mm00475829_g1
Asb4	Mm00480830_m1
C1qa	Mm00432142_m1
Cartpt	Mm04210469_m1
Ghrh	Mm00439100_m1
Gsg1l	Mm01278519_m1
Irf9	Mm00492679_m1
Kdm8	Mm00513079_m1
Pomc	Mm00435874_m1
Npy	Mm00445771_m1
Serpina3n	Mm00776439_m1
Sh2b1α	Mm01275190_m1
Sh2b1β	Mm01163373_g1
Sh2b1γ	Mm00488153_m1
Sh2b1δ	Mm01163374_m1
Slc5a11	Mm00461434_m1
36b4 (reference gene)	Mm00725448_s1
Gapdh (reference gene)	Mm99999915_g1
Tbp (reference gene)	Mm01277042_m1

Supplementary Table 4. Genes with significantly increased expression in αδKO mice

(sorted by % increase, largest to smallest)

Gene name	Gene product	Average CPM		% increase in αδKO	P value
		WT	αδKO		
Slc5a11	Solute carrier family 5 (sodium/glucose cotransporter), member 11 (a.k.a. KST1, SMIT2)	1.663	2.413	45.087	0.0318
AI467606	Expressed sequence AI467606	3.974	5.599	40.882	0.0012
Myo1f	Myosin IF (unconventional)	2.428	3.185	31.195	0.0320
C1qa	Complement component 1, q subcomponent, α polypeptide	32.498	38.561	18.658	0.0127
Itgam	Integrin α M (a.k.a. CD11b/CD18, CR3, CR3α, Mac-1, Mac-1α)	4.279	5.033	17.613	0.0407
Ctss	Cathepsin S	49.920	58.084	16.355	0.0117
Fcgr3	Fc receptor, IgG, low affinity III (a.k.a. CD16)	7.314	8.459	15.648	0.0168
Tyrobp	TYRO protein tyrosine kinase binding protein (a.k.a. DAP12, KARAP)	6.190	7.141	15.370	0.0288
C1qb	Complement component 1, q subcomponent, β polypeptide	39.367	45.255	14.957	0.0048
Trem2	Triggering receptor expressed on myeloid cells 2 (a.k.a. TREM2a, TREM2b, TREM2c)	9.441	10.832	14.732	0.0166
Tmem41a	Transmembrane protein 41a	19.207	21.883	13.936	0.0155
Kcne1l	Potassium voltage-gated channel, Isk-related family, member 1-like, pseudogene (a.k.a. MINK, KCNE5)	10.386	11.626	11.943	0.0337
Laptm5	Lysosomal-associated protein transmembrane 5 (a.k.a. E3)	27.053	30.221	11.708	0.0273
Tmem119	Transmembrane protein 119 (a.k.a. OBIF)	19.186	21.245	10.732	0.0413
Pgp	Phosphoglycolate phosphatase (a.k.a AUM, G3PP)	64.349	70.578	9.680	2.53E-04
Cx3cr1	Chemokine (C-X3-C motif) receptor 1	46.008	49.929	8.521	0.0127
Hexb	Hexosaminidase B	98.714	106.838	8.230	0.0062
Ajap1	Adherens junction associated protein 1	44.643	47.881	7.252	0.0443
Cox6a1	Cytochrome c oxidase subunit 6A1	166.605	178.519	7.151	0.0145
Ppp2r5d	Protein phosphatase 2, regulatory subunit B', δ (a.k.a. B'δ, TEG-271)	131.952	140.590	6.546	0.0407
Ctsd	Cathepsin D (a.k.a. CD, CatD)	143.229	152.248	6.297	0.0381
Cdk5r2	Cyclin-dependent kinase 5, regulatory subunit 2 (p39)	199.699	212.086	6.203	0.0159
Rplp2	Ribosomal protein, large P2	97.732	103.728	6.135	0.0159
Abhd17a	αβ hydrolase domain containing 17A (a.k.a. FAM108A)	96.430	102.187	5.971	0.0121
Grn	Progranulin (a.k.a. GP88, PEPI, PCDGF)	62.386	65.965	5.736	0.0336
Gpat4	Glycerol-3-phosphate acyltransferase 4 (a.k.a. AGPAT6, TSARG7)	107.142	113.061	5.525	0.0166
Cotl1	Coactosin-like 1 (Dictyostelium) (a.k.a. CLP)	73.903	77.760	5.218	0.0119
Rhob	Ras homolog family member B (a.k.a. ARH6, ARHB)	390.237	409.486	4.933	0.0449
Inpp5f	Inositol polyphosphate-5-phosphatase F (a.ka. SAC2)	181.284	189.614	4.595	0.0145
Fbxw8	F-box and WD-40 domain protein 8 (a.k.a. FBW6, FBW8, FBX29, FBXO29)	66.920	69.887	4.435	0.0354
Map7d1	MAP7 domain containing 1 (a.k.a. RPRC1, PARCC1, MTAP7D1)	341.982	354.815	3.752	0.0127

CPM, counts per million

Supplemental Material

Supplementary Table 5. Genes with significantly decreased expression in αδKO mice

(sorted by % decrease, largest to smallest)

		Average CPM			
				% decrease	
Gene name	Gene product	WT	αδKO	in αδKO	P value
Kdm8	Lysine (K)-specific demethylase 8 (a.k.a. JMJD5)	5.606	3.277	-41.546	2.23E-20
Epx	Eosinophil peroxidase (a.k.a. EPO)	1.271	0.842	-33.792	0.0414
Slc26a10	Solute carrier family 26, member 10	2.566	1.789	-30.264	0.0145
Fbxw10	F-box and WD-40 domain protein 10 (a.k.a FBW10, HREP, SM25H2, SM2SH2)	1.268	0.900	-29.027	0.0407
Gsg1l	Germ cell-specific gene 1-like	44.831	32.760	-26.926	5.52E-07
Slc2a4rg-ps	Slc2a4 regulator, pseudogene	6.401	5.206	-18.662	0.0119
Gm15446		18.392	15.394	-16.305	0.0336
C78859	Expressed sequence C78859, long non-coding RNA	8.953	7.497	-16.259	0.0145
Bub3	BUB3 mitotic checkpoint protein	85.535	71.881	-15.963	9.00E-07
Sfi1	SFI1 centrin binding protein	15.313	13.211	-13.727	0.0145
Leng8	Leukocyte receptor cluster (LRC) member 8	135.504	117.415	-13.349	0.0434
Mbd6	Methyl-CpG binding domain protein 6	34.378	29.797	-13.327	0.0127
Intu	Inturned planar cell polarity protein (a.k.a. PDZD6, PDZK6)	9.416	8.162	-13.316	0.0240
Sirt4	Sirtuin 4	13.946	12.167	-12.762	0.0407
Mir124a-1hg	Mir124-1 host gene (non-protein coding)	46.598	40.709	-12.639	0.0152
Uvssa	UV stimulated scaffold protein A	10.853	9.535	-12.141	0.0263
Rsrp1	Arginine/serine rich protein 1	221.905	198.823	-10.402	0.0166
Abhd14b	αβ hydrolase domain containing 14b	16.624	14.957	-10.026	0.0151
Tarbp2	TARBP2, RISC loading complex RNA binding subunit (a.k.a. PrBP)	15.433	13.893	-9.975	0.0337
Clasrp	CLK4-associating serine/arginine rich protein (a.k.a. CLASP, SFRS16, SRSF16, SWAP2)	54.974	49.735	-9.531	0.0337
Zfp512b	Zinc finger protein 512B (a.k.a. ZNF512b)	52.156	47.248	-9.410	0.0337
Mtf2	Metal response element binding transcription factor 2 (a.k.a M96, PCL2)	28.200	25.661	-9.002	0.0288
Mphosph9	M-phase phosphoprotein 9 (a.k.a. MPP9)	17.364	15.892	-8.476	0.0155
Hmgn3	High mobility group nucleosomal binding domain 3 (a.k.a. TRIP7)	28.740	26.527	-7.700	0.0293
Iqce	IQ motif containing E	32.067	29.635	-7.582	0.0218
Kmt2b	Lysine (K)-specific methyltransferase 2B (a.k.a. MLL2, WBP7)	64.157	59.958	-6.544	0.0489
Tnrc6a	Trinucleotide repeat containing 6a (a.k.a. GW182)	121.310	113.487	-6.449	0.0166
Srrm1	Serine/arginine repetitive matrix 1 (a.k.a. SRM160)	108.441	101.786	-6.137	0.0192

CPM, counts per million

Supplementary Table 6. Oligos for guide RNAs

Guide	Nucleotide #s	Strand	Sequence (5' → 3')
C	7203-7222	1	AAACccgcccatgattcatcttcc
C	7203-7222	2	CACCggaagatgaatcatgggcgg
D	8024-8043	1	AAACaacccaaccaagggtgagggC
D	8024-8043	2	CACCGccctcacccttggttgggtt

Supplementary Table 7. Primers used to generate a vector encoding GFP-SH2B18c

Primer	Direction	Sequence (5' → 3')
AarI	Forward	GCTGACGGATCCCACCTGCGCTTG TCACTAAATGAG
HindIII (reverse complement)	Reverse	GGATCCGTCAGCAAGCTTTCAGGC CATGAATCCCCCAAAAGG

Animal care

Mice were weaned at 3 weeks of age. Mice were fed standard chow or HFD starting at 3 or 4

weeks of age, respectively. Mice included in standard chow experiments were individually

housed at 5 weeks of age (leptin sensitivity experiments, Fig. 7B-D) or 9 weeks of age (all other

experiments). Mice included in HFD experiments were individually housed at 4 weeks of age.

Breeders were fed a standard chow containing 6.5% fat (#5008; LabDiet) or 9% fat (#5058;

LabDiet). When fed *ad libitum*, mice were housed in cages with corncob bedding (Bed-o'Cobs

¼", #4B; The Andersons). When fasted, mice were temporarily housed in cages with corn-free

natural soft cellulose bedding (Comfort Bedding; BioFresh) to prevent them from eating the

bedding and thus disrupting their fasting.

Generation of αδKO mice

The reverse complement of the genomic *Sh2b1* sequence in C57BL/6J mice (GenBank accession

#NC_000073, GRCm38) was used when designing CRISPR reagents. The selected guides were

designed to direct Cas9 to cut close to the ends of the desired deleted regions (Fig. 1B, top) and

have a low likelihood of off-target binding. Two sets of RNA guides, C and D, were selected. Guide C was used alone or Guides C and D were used together. The guides were expressed using the pX330 vector [1, 2], which contains a chimeric guide RNA expression cassette and a hSpCas9 expression cassette. Sense and anti-sense oligos (Supplementary Table 6) corresponding to either Guide C or Guides C and D were annealed and subcloned between the two Bbs1 sites in the pX330 vector. The resulting construct was expressed in DH5α cells. Colonies were selected, DNA isolated, and the sequence verified in the region of the inserted guide sequence. The pX330 plasmid containing the guide sequence and the Zero blunt TOPO vector containing the donor template were purified as previously described [3].

The guide(s) and donor template were tested in blastocysts of C57BL/6J x SJL F2 mice. Sequencing of DNA from blastocysts showed that 6 of 20 blastocysts (30%) injected with Guide C and 16 of 30 blastocysts (53%) injected with both Guides C and D contained the correctly edited *Sh2b1* sequence. Of the 17 pups born from mice implanted with oocytes injected with Guide C, 3 (18%) contained the correctly edited *Sh2b1* sequence and 14 were WT. Of the 58 pups born from mice implanted with oocytes injected with both Guides C and D, 7 (12%) contained the correctly edited sequence, 2 had insertions/deletions, and 52 were WT. DNA sequencing revealed germline transmission from all founders. All experiments were performed using progeny of one founder generated from Guide C.

Body weight and food intake

Body weight was assessed starting at 4 weeks of age. Food intake was determined weekly starting at 10 weeks of age for standard chow experiments or twice weekly starting at 5 weeks of age for the HFD experiment. Food intake was determined by weighing the food remaining in the

cage and subtracting the value from the weight of the food initially added to the hopper.

Following determination of food intake, old food was removed and fresh food was added. Mice

were excluded from food intake data analysis if one or more weekly measurement was missing

and/or they were identified as outliers by the following rule: their final cumulative food intake

value was greater than two standard deviations from the mean. This quantitative identification of

outliers matched with qualitative notes from individuals blinded to mouse genotypes that

identified "extreme nibblers"—mice that ground food pellets into powder and spread it around

their cage, making it difficult to accurately measure their food intake.

Glucose tolerance tests

When HFD-fed mice were subjected to glucose tolerance tests, six WT mice reached the

maximum value of the glucometer range (600 mg/dL) at one or more time points. As a result, the

average WT curve depicted (Fig. 5G) is lower than it would be if the meter had a higher

maximum value.

qPCR

Before reverse-transcription, RNA quality was confirmed using a Nanodrop spectrophotometer.

qPCR was performed with an Eppendorf Realplex2 using Mastercycler software and PCR

parameters recommended by Applied Biosystems. Three reference genes (*36b4, Gapdh, Tbp*)

were used for all assays. All cDNA samples were analyzed in triplicate and a non-template

control was included for each gene of interest. Cycle threshold (Ct) values were normalized to

the geometric mean of the Ct values of the reference genes, as in Vandesompele et al. [4]. The

expression of the reference genes did not differ between the control and experimental samples (data not shown).

RNA-seq and data analysis

RNA samples were treated with DNase I. RNA was assessed for quality using Agilent 4200 TapeStation. RNA samples were enriched for mRNA transcripts using NEBNext Poly(A) mRNA Magnetic Isolation Module (#E7490; New England BioLabs). cDNA libraries were prepared from mRNA using NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (#E7760; New England BioLabs). Libraries were assessed for quality and quantity using Agilent TapeStation and qPCR by KAPA Library Quantification Kit for Illumina Platforms (#KK4835; Kapa Biosystems). Libraries were sequenced on the Illumina NovaSeq 6000 S4 Flow Cell for 200 cycles with 100 base pair paired-end reads.

FASTQ files were inspected using FastQC 0.11.7 [5] to ensure quality reads. Fastq_quality_filter (from the fastx_toolkit 0.0.14) [6] was used to remove data with Phred score < 20. Reads that made it through filtering were then aligned to the genome using STAR 2.5.3a_modified [7] and mouse genome GRCm38 (version 92). Count tables were generated using STAR with flag --quantMode GeneCounts. Count tables were analyzed in R 3.5.2 [8] and another round of quality control was performed. Samples were inspected for library size (all samples had > 20 million mapped reads), downregulation of *Sh2b1*, and dimension reduction for detection of batch effects. RNA samples were prepared in multiple batches, which were accounted for in the analysis. Differential expression of genes was determined using DESeq2 1.22.2 [9]. GO analysis was performed using PANTHER DB [10]. *Sh2b1* transcripts were assembled using Ensembl annotation (version 92) [11]. Novel *Sh2b1* transcripts were identified

using StringTie 1.3.6 [12]. WT and αδKO samples were merged during StringTie analysis. Read sequences misaligned by StringTie were excluded from further analysis. Misalignments were identified by their robust outlier quality (e.g. all samples of the same genotype measured 0 counts per million except one sample, which measured 11 counts per million). Statistics for differential gene expression and GO analysis were calculated using the default parameters of DESeq2 and PANTHER DB, respectively. $P < 0.05$ was considered significant. Differentially expressed genes were visualized using ggplot2 3.3.0 [13] in RStudio [14].

To identify phenotype enrichment of differentially expressed genes, the differentially expressed genes were used to query the MouseMine database through the python Application Programming Interface (http://www.mousemine.org/mousemine/). The associated phenotype corresponds to the MouseMine category "ontologyTerm.name". To determine significant enrichment, all protein-coding genes for which a P value was measured by DESeq2 were fed into the same MouseMine query. Enrichment significance of phenotypes in differentially expressed genes versus all genes was determined using prop.test() in R 3.6.3. P values were corrected for multiple comparisons using the Benjamini-Hochberg method.

To determine cell type enrichment of differentially expressed genes, data from GSE87544 was downloaded from GEO as a count matrix (GSE87544_Merged_17samples_14437cells_count.txt.gz) and metadata (GSE87544_1443737Cells.SVM.cluster.identity.renamed.csv.gz). Data were imported into R and analyzed with Seurat 3.1.5 [15, 16]. Count data was scaled with NormalizeData() followed by ScaleData(), then averaged by cell type as defined by the GSE87544 metadata. Heatmap and dendrogram were plotted with ggplot2 3.3.0 and ggdendro 0.1-20.

Supplemental Material

Plasmid expressing GFP-SH2B18c

The genomic sequence of *Sh2b1* between the end of exon 8 and the start of exon 9 (nt 6658-7181) was isolated from a pCR-Blunt II-TOPO vector (Invitrogen) containing a larger segment of genomic *Sh2b1* sequence (nt 4579-8807) using primers containing the the AarI or the HindIII restriction sites (primers listed in Supplementary Table 7) and PfuUltra High-Fidelity DNA Polymerase (#600380; Agilent). cDNA encoding mouse SH2B1γ [17] was subcloned into pEGFPC1 (Clontech). The vector encoding GFP-SH2B1γ was then cleaved with the AarI and HindII restriction sites and the novel PCR product (*Sh2b1* nt 6658-7181) was inserted into this vector. The sequence was confirmed by the University of Michigan DNA Sequencing Core.

PC12 cell neurite outgrowth assay

PC12 cells were plated in tissue culture dishes coated with rat tail type I collagen (#354236; Corning) in RPMI 1640 (#A10491-01; Gibco) supplemented with 10% horse serum (HS) (#16050114; Gibco) and 5% fetal bovine serum (FBS) (#S11150; Atlanta Biologicals). For the neurite outgrowth assay, PC12 cells were plated in 6-well collagen-coated plates and transiently transfected with the indicated construct for 24 hours using Lipofectamine LTX (#15338030; Invitrogen). Cells were treated with 25 ng/mL mouse NGF 2.5S (BT.5025; Envigo) in RPMI 1640 medium containing 2% HS and 1% FBS. After 2 days, GFP-positive cells were visualized by fluorescence microscopy (20X or 40X objective, Nikon Eclipse TE200)
and scored for the presence of neurites ≥ 2 times the length of the cell body. N=3 experiments (300 cells per condition counted in 1st replicate; 200 cells per condition counted in 2nd and 3rd replicates).

Immunoblotting

Blots were incubated overnight at 4°C with mouse monoclonal antibody to SH2B1 (1:1,000) (#sc-136065; Santa Cruz) or β-tubulin (1:1,000) (#sc-55529; Santa Cruz) or ERK1/2 (1:1,000) (#4695S; Cell Signaling) in 10 mM Tris, 150 mM NaCl, pH 7.4, 0.1% Tween 20, and 3% ovalbumin from chicken egg white, followed by IRDye-800CW goat anti-mouse IgG secondary antibody (1:20,000) (#926-32210; Li-Cor) for 1 hour at room temperature.

Supplemental Material References

1. Ran, F.A., et al., *Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity.* Cell, 2013. **154**(6): p. 1380-1389.
2. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat. Protoc., 2013. **8**(11): p. 2281-2308.
3. Flores, A., et al., *Crucial Role of the SH2B1 PH Domain for the Control of Energy Balance.* Diabetes, 2019. **68**(11): p. 2049-2062.
4. Vandesompele, J., et al., *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.* Genome Biol, 2002. **3**(7): p. research0034.1-0034.11.
5. Andrews, S., *FastQC: A Quality Control Tool for High Throughput Sequence Data.* 2010.
6. Hannon, G.J., *FASTX-Toolkit.* 2010.
7. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
8. Team, R.C., *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, 2017.
9. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.
10. Thomas, P.D., et al., *PANTHER: a library of protein families and subfamilies indexed by function.* Genome Res, 2003. **13**(9): p. 2129-2141.
11. Cunningham, F., et al., *Ensembl 2019.* Nucleic Acids Res, 2019. **47**(D1): p. D745-D751.
12. Pertea, M., et al., *StringTie enables improved reconstruction of a transcriptome from RNA-seq reads.* Nat Biotechnol, 2015. **33**(3): p. 290-295.
13. Wickham, H., *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York (book), 2009.
14. Team, R., *RStudio: Integrated Development for R. RStudio, Inc., Boston, MA.* 2015.
15. Stuart, T., et al., *Comprehensive Integration of Single-Cell Data.* Cell, 2019. **177**(7): p. 1888-1902 e21.
16. Butler, A., et al., *Integrating single-cell transcriptomic data across different conditions, technologies, and species.* Nat Biotechnol, 2018. **36**(5): p. 411-420.

Supplemental Material

17.   Yousaf, N., et al., *Four PSM/SH2-B alternative splice variants and their differential roles in mitogenesis.* J. Biol. Chem., 2001. **276**(44): p. 40940-40948.