Supplemental Material

Detailed Methods

Analysis of SHROOM3 Regulation and Transcripts

Roadmap epigenomics 15 core data was accessed from Wash U browser (http://egg2.wustl.edu/roadmap/web_portal/chr_state_learning.html#core_15state). SwitchGear Genomics Transcription Start Sites were identified using the human genome browser in hg19 build with lower stringency setting (https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=647211331_6qX1NPkqJmLzEbxdLX2UaLwwrruH&c=chr4&g=switchDbTss). The LD block for the GWAS near *SHROOM3* was identified using the SNAP tool[1] (http://archive.broadinstitute.org/mpg/snap/) with a 0.8 correlation. The 11 variants identified with SNAP were then analyzed using RegulomeDB[2] for integration analysis of multiple ENCODE and roadmap datasets at each variant.

To identify cell usage of TSS1 and TSS2, Cap Analysis of Gene Expression (CAGE) data was analyzed on the datasets of FANTOM[3,4]. The robust promoter data of Phase1 and 2 FANTOM5 was extracted from the Zenbu browser for sites near either TSS1 (hg19: chr4:77,352,033-77,362,449) or TSS2 (hg19: chr4:77,504,889-77,509,140). In order to build transcript maps using multiple RNAseq datasets, reads were extracted using NCBI SRA BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=SRA&LINK_LOC=blasttab) for *SHROOM3*/*Shroom3,* aligned to transcripts using BOWTIE 2 (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml), and visualized using Unipro UGENE (http://ugene.net/). Reads mapped per million reads sequenced (RPM) were normalized between each RNAseq run by taking the number of mapped reads to each transcript divided by the total number of million reads sequenced in the dataset. Transcripts used were: mouse *Shroom3* (NM_015756), human *SHROOM3* (NM_020859), mouse *Nphs2* (XM_006496683), and human *NPHS2* (NM_014625). RNAseq datasets used were mouse podocyte single cell: SRX2248950, SRX2248951, SRX2248952, SRX2248953, SRX2248954, SRX2248955, SRX2248956, SRX2248957, SRX2248958, SRX2248959, SRX2248960, SRX2248961, SRX2248962, SRX2248963, SRX2248964, SRX2248965, SRX2248966, SRX2248967, SRX2248968, SRX2248969; mouse isolated glomeruli: SRX1637541, SRX1637540, SRX1637539, SRX1637538, SRX1637537, SRX1637536, SRX1637535, SRX1637534, SRX1637533, SRX1637532, SRX1637531; mouse single cell proximal tubule cells: SRX2884964, SRX2884963, SRX2884962, SRX2884961, SRX2884960, SRX2884959, SRX2884958, SRX2884957, SRX2884956, SRX2884955, SRX2884954, SRX2884953, SRX2884952, SRX2884951, SRX2884950, SRX2884949, SRX2884948, SRX2884947, SRX2884946; Human Protein Atlas kidney: ERX288476, ERX288607, ERX288628, ERX288622; Human Protein Atlas brain: ERX288614, ERX288639, ERX288561; HEK293: SRX3015634, SRX3015633, SRX3015632.

HiChIP analysis was performed using the Yue labs integrated viewer: (http://promoter.bx.psu.edu/hi-c/chiapet.php?method=hichip&species=human&assembly=hg19&tissue=GM12878&target=cohesin&lab=Chang&gene=shroom3&window=250&sessionID=&browser=ucsc). HiChIP data was from the cohesin datasets previously published[5]. For all of ENCODE Transcription Factor (TF) binding sites as determined by ChIP-Seq to overlap with either rs17319721 (hg19: chr4:77,361,347-77,376,346) or TSS2 (hg19: chr4:77487991-77513325) we pulled peaks from the Human Genome Browser (https://genome.ucsc.edu/) using the Transcription Factor ChIP-Seq (161 factors). For HEK293 cells we downloaded ChIP-Seq maps of 190 unique TFs from ENCODE portal (https://www.encodeproject.org) to retrieve only IDR (Irreproducible Discovery Rate) passed peaks or binding sites for each factors. Peaks were analyzed for binding overlap

of all 190 TFs with rs17319721 or TSS2 loci using a - 50 bp up-and- downstream from summit of each peak using Pybedtools[6,7]. GO enrichment analysis of all TFs groupings was performed using the Gene Ontology Enrichment tools (http://geneontology.org/). STRING analysis[8] was performed on the list of TFs found bound to show connections of TFs bound.

Electrophoresis Mobility Shift Assays

Nuclear extracts were purchased for K562 (#36015, Active Motif), HEK293 (#36033, Active Motif), and A-431 (#36004, Active Motif) and prepared from Human Primary Kidney Endothelial Cells (#H-6014, Cell Biologics), Human Primary Tubule cells (#PCKDH01, Primecells), and Primary Monkey Podocytes (#PCKDP01, Primecells) grown according to manufactures and extracted with the NE-PER kit (#78833, ThermoFisher). Recombinant proteins were purchased for FOXO1 (#TP300477, Origene) and TCF7L2 (#513437, Novoprolabs).

Biotin conjugated DNA probes (5' tagged) were ordered (IDT) for each SNP (Table S1), annealed (90°C-10min, 80°C-10min, 70°C-10min, 60°C-10min, 50°C-10min, 40°C-10min, 30°C-10min, 20°C-hold) with the reverse complement strand (nonbiotin tagged) in annealing buffer (10mM Tris, 1mM EDTA, 100mM NaCl, ph 8), and diluted to a final volume of 0.25uM. EMSAs were performed using the LightShift Chemiluminescent EMSA kit (ThermoFisher). Binding reactions consisted of 2uL of binding buffer, 1uL glycerol, 1uL NP40, 1uL Poly-dIdC, 1uL of DNA probe annealed and diluted, 2uL of nuclear extract (at 2 ug/uL), and brought to 20uL with water. Outcompetition binding experiments were performed with non-biotin tagged probe at 40X concentration in addition to the labeled probe. Binding reactions were incubated at room temp for 20 min followed by the addition of 5uL of 5X loading buffer. Reactions were run on a 6% DNA Retardation Gel (ThermoFisher, #EC6365BOX) at 100V for 1hr with 0.5X TBE and transferred to Biodyne B membrane (ThermoFisher, # 77016). Membranes were dried and crosslinked. Membranes were probed with streptavidin-HRP exactly as recommended in the LightShift Chemiluminescent EMSA kit.

CRISPR/Cas9 Modification and Expression Analysis

For generation of the single variant, CRISPR gRNAs were cloned into U6-Chimeric_BB-CBh-hSpCas9 (PX330) which was a kind gift from Feng Zhang (Addgene plasmid # 42230). Oligos for a gRNA (GAGTAGCAGGGCAAAAACA) near the SNPs were identified and ordered through Integrated DNA technologies (IDT) and cloned downstream of a U6 promoter element as previously described[9]. Single strand donor DNA was designed to include the SNP rs17319721 which also removed the PAM sequence, and ordered as ultramers through IDT. HEK293T cells were grown under recommended growth conditions. Nucleofection (Lonza) was done using SF kit with the Nucleofetor 4D in the 20ul total volume with 2 ug of PX330 and 0.1 nmoles of ssDNA donor. Cells were expanded and diluted to single cell colonies. These colonies were expanded and tested for genetic engineering at the Cas9 target, and a colony homozygous for rs17319721 was confirmed by Sanger sequencing.

For generation of a cell line without the LD block, single guide (sgRNA) oligos targeting intron 1 of *SHROOM3* were designed and cloned into pX330 plasmids as previously described[10]. pX330 plasmids, containing spCas9 and the designed sgRNAs, were transfected into HEK 293T using lipofectamine 2000 according to manufacturer's protocol (Thermo Fisher Scientific). All cell lines were cultured and maintained in Dulbecco modified Eagle medium (DMEM) with 10% fetal bovine serum (FBS), 1% L-glutamine, penicillin streptomycin (1X), and sodium pyruvate (1X). Single cell clones were isolated and genomic DNA of each clone was extracted using DNeasy Blood & Tissue Kit (Qiagen). Genotypes of each cell clones were checked by PCR amplification using primers flanking the two CRISPR target sites as cell clones harboring the desired deletion but not the unmodified cells were expected to produce PCR amplicons of ~200bp. PCR product was ran on a gel, purified and sequenced to verify deletions.

Target (F:forward, R:reverse)	Primer sequence (5' to 3')
sgRNA1_F	CACCGGTATGCTGCAGGATGACTA
sgRNA1_R	AAACTAGTCATCCTGCAGCATACC
sgRNA2_F	CACCGGGAGGATGTATCGGACTTT
sgRNA2_R	AAACAAAGTCCGATACATCCTCCC
PCR_F	TTCCCCACACTCAGAAAGGA
PCR_R	CTTATCTGCCGCTCCACCAT

Cells were grown to 70% confluency and then lysed in 350uL of RLT Plus buffer supplemented with 2-Mercaptoethanol and RNA was purified using RNaesay Plus Mini kit (QIAGEN) according to manufacturer's instructions. For all samples, cDNA was generated with RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific) using 1ug of total RNA. Quantitative PCR was performed using AmpliTaq DNA polymerase (Applied Biosystems), and the products were separated on 2% agarose gel. Quantitative real-time PCR was performed on the QuantStudio 6 Flex Real-Time PCR system (Applied Biosystems) using primers listed in Table S1 and QuantiTect SYBR Green PCR Kit (QIAGEN) with cycling conditions following: 95C- 15 min, 40x (94°C- 20° sec, 60°C- 20 sec, 70°C- 20 sec). The reactions were conducted in triplicate. Relative gene expression levels were analyzed using the Comparative CT Method (DDCT). The *GAPDH* and *B2M* gene was used as a reference.

Zebrafish Experiments to Test the Short Isoform of SHROOM3.

Human *SHROOM3* cDNA ORF (NM_020859) was purchased from OriGene Technologies (Rockville, MD). To test the function of *SHROOM3* protein domains, human *SHROOM3* mutant alleles lacking either the ASD2 (ΔASD2) or PDZ (ΔPDZ) site were generated using Phusion site-directed mutagenesis kit according to manufacturer's protocol (Thermo Fisher Scientific) using phosphorylated primers listed in the table below. All plasmids were sequenced to verify the constructs were correct. For rescue experiments in zebrafish, plasmids were linearized by *PmeI* and *in vitro* transcribed into mRNA using mMESSAGE mMACHINE® kit (Ambion) according to the manufacturer's instructions. 400 pg of full-length or mutant *SHROOM3* mRNA was co-injected along with the *shroom3* morpholino (MO) into 1- to 4- cell stage zebrafish embryos. A modified dextran clearance assay was developed from that of previous study[11]. In short, 55-h post-fertilization zebrafish were anesthetized with tricane, embedded in 1% agarose, and the cardinal vein injected with 9.2 nl 70-kDa FITC-labeled dextran (1 mg/mL). Following injection, the dorsal aorta was imaged by confocal microscopy at 1, 24, and 48 h post-injection (hpi). FITC fluorescent intensities were quantified using ImageJ software (National Institutes of Health).

Target (F:forward, R:reverse)	Primer sequence (5' to 3')
ΔPDZ _F	TGCCAGCATGATGCAGATATCTCAGGG
ΔPDZ _R	GGCGGCAGATCTCCTCGGTA
ΔASD2 _F	CCAACATTAACCTCTCCACTT
ΔASD2 _R	TAAACCCTGAGTTCCACTGAC

Immunofluorescence and Immunoprecipitation of SHROOM3 in Podocytes

Human (Celprogen, #36036-08) primary podocytes were grown on podocyte extra-cellular expansion flasks (Celprogen, #E36036-08-T75) with complete human podocyte media (Celprogen, #M36036-08S) at 5%CO_2/20%O_2. K3 (gift from Duncan lab) inducible pluripotent stem cells (iPSCs) were cultured on StemAdhere with mTeSR1 media (StemCell Technology, #85870) at 5%CO_2/5%O_2. Cells were plated at 90% confluency onto iMatrix Laminin-511-E8 (Iwai, #N-892012) coated 6-well plates (VWR, #82050-846) and grown using a previously published podocyte differentiation protocol (PMID: 29038743). The iPSC derived podocytes were plated onto 6-well plates coated with iMatrix Laminin-511-E8 and the human podocytes onto podocyte treated 6-well plates (Celprogen, E36036-08-6W).

Immunoprecipitations were performed using 100 ug of SHROOM3 antibody from either Sigma (#HPA047784) or ThermoFisher (#PA5-61589) crosslinked to resin using Pierce Co-Immunoprecipitation kit (#26149) according to recommended protocol. Primary human podocytes were lysed using the IP lysis buffer from the kit, precleared with control agarose resin, and then incubated with SHROOM3 antibody resin. A control resin without SHROOM3 antibody was used. Proteins were removed from the resin as recommended in the kit and run on 4-12% SDS-PAGE followed by either silver staining (ThermoFisher, #24612) or transfer to PDVF and western blotted with a third anti-SHROOM3 antibody (ThermoFisher, #PA5-34482) or the Sigma anti-SHROOM3 antibody.

LATS2 Kinase Assay

All phosorylation sites shown in Figure 4A were identified from UniProt (http://www.uniprot.org/) and the LATS1/2 kinase prediction done using NetPhosK[12]. Peptides for the kinase assay were ordered from GenScript at >90% purity as confirmed by HPLC and mass spec followed by suspension in water: SHROOM3 (AGPVHVRSRSSPATAD), SHROOM3 H1237A (AGPVAVRSRSSPATAD), SHROOM3 SS1241AA (AGPVHVRSRAAPATAD), SHROOM3 P1244L (AGPVHVRSRSSLATAD). LATS2 active enzyme (#L02-11G, SignalChem) was setup to run reactions at 400ng, 300ng, 200ng, 100ng, 50ng, 25ng, and a 0 control in two separate experiments. Reactions were setup to a total volume of 20uL using 5ug of each peptide, LATS2 enzyme, and 5uL Kinase Reaction Buffer (40mM Tris pH 7.5, 20mM $MgCl_2$, 0.1mg/mL BSA). Kinase reactions were performed using the ADP-Glo Kinase Assay (#V6930, Promega) such that reactions were initiated with the addition of 5uL UltraPure ATP to a final concentration of 1mM and incubated at 23°C for 30min. 25uL of ADP-Glo reagent was added, reactions incubated at 23°C for 40min, 50uL kinase detection reagent added, incubated at 23°C for 30min, and luminescence read using a BioTek Synergy H4 multimode plate reader.

Peptide Synthesis for Crystal Structures

The SHROOM3 linear motif at 1244 was identified to be a potential 14-3-3 binding site based on ELM analysis[13]. The SHROOM3-derived peptides, wt (H_2N-AGPVHVRSRpSSPATA-NH2, pS represents a phospho Ser) and P1244L mutant (H_2N-AGPVHVRSRpSSLATA-NH2), were synthesized via Fmoc solid phase peptide synthesis strategy[14] using an Intavis MultiPep RSi peptide synthesizer. The phosphorylated peptides were synthesized using Fmoc-protected amino acid building blocks (4 eq., Novabiochem) and a HBTU (Biosolve b.v., 4 eq.) /*N,N*-diisopropylethylamine (8 eq., Biosolve b.v.) coupling strategy at 50 µmol scale on a Tentagel®R RAM resin (Rapp Polymere; 0.19 mmol/g loading). The phosphoserine amino acid was specifically introduced via Fmoc-Ser(PO(OBzl)OH)-OH (Novabiochem). The peptides were cleaved using TFA/triisopropylsilane/H_2O (95/2.5/2.5 v/v), precipitated into ice-cold ether, and the resultant crude isolated as a solid pellet by centrifiguation. Peptide purification was performed using a preparative LC-MS system, which comprised of a LCQ Deca XP Max (Thermo Finnigan) ion-trap mass spectrometer equipped with a Surveyor autosampler and

Surveyor photodiode detector array (PDA) detector (Thermo Finnigan). Solvents were pumped using a high-pressure gradient system using two LC-8A pumps (Shimadzu) for the preparative system (20 mL min^{-1} flow rate) and two LC-20AD pumps (Shimadzu) for the analytical system (0.2 mL min^{-1} flow rate). The crude mixture was purified on a reverse-phase C18 column (Atlantis T3 prep OBD, 5 µm, 150 x 19 mm, Waters) using a linear acetonitrile gradient in water with 0.1% v/v trifluoracetic acid (TFA). Fractions with the correct mass were collected using a PrepFC fraction collector (Gilson Inc). Molecular characterization of SHROOM3-derived peptides (LC-MS) are as follows: wt sequence, $C_{62}H_{107}N_{25}O_{22}P_1$, Exact Mass: 1570.7724, found: 1571.7 $[M+H]^+$; P1244L mutant sequence, $C_{63}H_{111}N_{25}O_{22}P_1$, Exact Mass: 1586.8037, found: 1587.8 $[M+H]^+$.

14-3-3 Expression

His_6-tagged 14-3-3 proteins (full-length and ΔC-terminus) were expressed in NiCo21(DE3) competent cells with a pPROEX HTb plasmid, and purified using Ni^{2+}-affinity chromatography. The ΔC variant meant for crystallization was treated with TEV-protease to cleave off the His_6-tag, followed by a second Ni^{2+}-affinity column and size exclusion chromatography. The proteins were dialyzed against ITC- or crystallization-buffers before usage (described below).

Isothermal Titration Calorimetry (ITC)

The ITC measurements were performed with the Malvern MicroCal iTC_{200}. The protein and peptides were dissolved in ITC-buffer (25 mM HEPES pH 7.4, 100 mM NaCl, 10 mM $MgCl_2$, 0.5 mM TCEP). Eighteen titrations of 2 µL were performed at 37 °C (reference power: 5 µCal/sec., initial delay: 60 sec., stirring speed: 750 rpm, spacing: 180 sec.). In case of two titration series the data was merged with ConCat32 software.

Crystallography

The 14-3-3σ protein was C-terminally truncated after T231 to enhance crystallization (14-3-3σΔC). The 14-3-3 protein and peptide were dissolved in crystallization buffer (25 mM HEPES, 0.1 M NaCl, 2 mM DTT, pH7.4) and mixed in a 1:1 stoichiometry to a final protein concentration of 10 mg/mL. This peptide-protein mixture was set up for sitting-drop crystallization in crystallization liquor (0.095 M HEPES, 0.19 M $CaCl_2$, 26% (v/v) PEG 400, 5% (v/v) glycerol, pH7.3). Crystals were fished after 10 days of incubation at 4°C and flash-cooled in liquid nitrogen. Diffraction data was collected at 100 K at the X06SA/PXI beamline (Swiss Light Source, Villigen, Switzerland). The data were indexed and integrated using XDS[15] and scaled using Aimless[16]. The structure was phased by molecular replacement using PHASER[17] with PDB 3LW1[18] as search model. Coot[19] and phenix.refine[20] were used in alternating cycles of manual and automatic refinement to complete the atomic model.

SHROOM3 and 14-3-3 affinity capture

Affinity capture was performed taking 400ug of GenScript synthesized (>95% purity) N-terminally biotin tagged phospho peptides for SHROOM3 (AGPVHVRSRpSSPATAD) or SHROOM3 P1244L (AGPVHVRSRpSSLATAD) binding to 80uL Pierce™ Avidin Agarose (#20219, ThermoFisher). Beads were washed three times with PBS and then 5ug of 14-3-3 beta (LF-P0040, ThermoFisher) was captured for 1 hour at 23°C followed by two washes with PBS+500mM NaCl and two washes with PBS. Beads were boiled with LDS sample buffer and samples were run on a NuPAGE 4-12% Bis-Tris Protein Gels (#NP0321BOX, ThermoFisher). Proteins were imaged using Pierce™ Silver Stain (#24600, ThermoFisher) on a myECL system.

Molecular Dynamic Simulations

Each of the PDB files were loaded into YASARA and setup for molecular dynamic simulations. Simulations were setup using pH of 7.4 for pKa predictions, 0.997g/mL of explicit water, and

0.9% NaCl in a simulation square that extends 10 angstroms from all atoms. Simulations were run for 125 nanoseconds with the AMBER03 force field[21] and periodic boundaries capturing the structure coordinates every 24 picoseconds. This recorded a total of 5000 trajectory files for each protein that were then assessed for movement in Root-mean squared deviation (RMSD in angstroms) averaging every atom of each residue that is connected to a hydrogen throughout all 5000 trajectory files.

References in Supplemental File

1. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PIW: SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinforma. Oxf. Engl.* 24: 2938–2939, 2008

2. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M: Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* 22: 1790–1797, 2012

3. Abugessaisa I, Noguchi S, Carninci P, Kasukawa T: The FANTOM5 Computation Ecosystem: Genomic Information Hub for Promoters and Active Enhancers. *Methods Mol. Biol. Clifton NJ* 1611: 199–217, 2017

4. A promoter-level mammalian expression atlas. - PubMed - NCBI [Internet]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/24670764 [cited 2017 Dec 12]

5. Mumbach MR, Rubin AJ, Flynn RA, Dai C, Khavari PA, Greenleaf WJ, Chang HY: HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* 13: 919–922, 2016

6. Dale RK, Pedersen BS, Quinlan AR: Pybedtools: a flexible Python library for manipulating genomic datasets and annotations. *Bioinforma. Oxf. Engl.* 27: 3423–3424, 2011

7. Quinlan AR, Hall IM: BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma. Oxf. Engl.* 26: 841–842, 2010

8. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ: STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 41: D808-815, 2013

9. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F: Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8: 2281–2308, 2013

10. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F: Multiplex genome engineering using CRISPR/Cas systems. *Science* 339: 819–823, 2013

11. Hentschel DM, Mengel M, Boehme L, Liebsch F, Albertin C, Bonventre JV, Haller H, Schiffer M: Rapid screening of glomerular slit diaphragm integrity in larval zebrafish. *Am. J. Physiol. Renal Physiol.* 293: F1746-1750, 2007

12. Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S: Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4: 1633–1649, 2004

13. Dinkel H, Michael S, Weatheritt RJ, Davey NE, Van Roey K, Altenberg B, Toedt G, Uyar B, Seiler M, Budd A, Jödicke L, Dammert MA, Schroeter C, Hammer M, Schmidt T, Jehl P, McGuigan C, Dymecka M, Chica C, Luck K, Via A, Chatr-Aryamontri A, Haslam N, Grebnev G, Edwards RJ, Steinmetz MO, Meiselbach H, Diella F, Gibson TJ: ELM--the database of eukaryotic linear motifs. *Nucleic Acids Res.* 40: D242-251, 2012

14. Fmoc Solid Phase Peptide Synthesis: A Practical Approach. Oxford, New York: Oxford University Press;

15. Kabsch W: XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66: 125–132, 2010

16. Evans PR, Murshudov GN: How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* 69: 1204–1214, 2013

17. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ: Phaser crystallographic software. *J. Appl. Crystallogr.* 40: 658–674, 2007

18. Schumacher B, Mondry J, Thiel P, Weyand M, Ottmann C: Structure of the p53 C-terminus bound to 14-3-3: implications for stabilization of the p53 tetramer. *FEBS Lett.* 584: 1443–1448, 2010

19. Emsley P, Cowtan K: Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60: 2126–2132, 2004

20. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH: PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66: 213–221, 2010

21. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P: A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* 24: 1999–2012, 2003

Supplemental Figures

Table S1 Oligos used

Name	Use	Sequence (5'->3')	Assay
rs17319721	Probe	5' Biotin-CTAGTGAGTAGCAGGGCAAAAACAA[A/G]GCAGCCATTGATCTAAGTGTAAGG	EMSA
rs17319721 Scramble	Probe	5' Biotin-CTAGTGAGTAGCAGGGCAAAGGTGGAATAGCCATTGATCTAAGTGTAAGG	EMSA
ΔPDZ	F	5'-Phospho-TGCCAGCATGATGCAGATATCTCAGGG	Mut Construct
ΔPDZ	R	5'-Phospho-GGCGGCAGATCTCCTCGGTA	Mut Construct
ΔASD2	F	5'-Phospho-CCAACATTAACCTCTCCACTT	Mut Construct
ΔASD2	R	5'-Phospho-TAAACCCTGAGTTCCACTGAC	Mut Construct
rs17319721 CRISPR	gRNA	GAGTAGCAGGGCAAAAACA	CRISPR / Cas9
rs17319721 CRISPR	Donor	GAAAATAGACTCTTTTCTCTCATGCTTTGCTTTTGC TTTTGTTTATTTTTTATATTTTTACTTTTGAGTATCA GCTAGTGAGTAGCAGGGCAAAAACAA[**A**]GCAGC CATTGATCTAAGTGTAAGGCAGCAGGAGGAAGG TTAAGAGAATGAGTGTTATTCTTTAGGACTACATT CATAGTTGCATGGTGTTGCTCTTGGCTC	CRISPR / Cas9
rs17319721 Sanger	F	TGTGATGAACATAGGGTGCAA	Sanger
rs17319721 Sanger	R	CAGAGGAAAGGGGACACTGA	Sanger
SHROOM3 E2E4	F	GGATCCTACAAGACCCTCAG	qPCR
	R	TCCCCAGATTTAGTTGTGTG	
SHROOM3 E8E9	F	TGTCCACCAAGACAAATCTC	qPCR, qRT-PCR
	R	GCTTCCTTGTCTTCATTCCT	
SHROOM3 E4E5	F	CTCACTCGTGGCACACAACTA	qPCR
	R	GACCATGCTTATCTAAGGCGG	
11-Sep	F	AAGCCAGAAGTTATGAGCTTCAG	qRT-PCR
	R	GCCTCGAACTGGGCATCAAT	
FAM47E	F	AGACCTTCACGGAACAGTTGC	qRT-PCR
	R	CACGTTTACATTCCTTCCCGA	
GAPDH	F	CGGAGTCAACGGATTTGGTCGTAT	qRT-PCR
	R	AGCCTTCTCCATGGTGGTGAAGAC	
B2M	F	TGAGTATGCCTGCCGTGTGAAC	qRT-PCR
	R	ATGCTGCTTACATGTCTCGATCCC	

Table S2 Expression of transcripts for *SHROOM3* and TCF7L2 from the Human Protein Atlas. Kidney is marked in red and the top four tissues for each gene are colored in cyan.

Tissue	*SHROOM3* FPKM±SEM	*SHROOM3* Isoform	SHROOM3 Staining	*TCF7L2* FPKM ± SEM	TCF7L2 Staining
Colon	30.6 ± 3.1	2/3	Glandular	23.6 ± 2.3	Glandular, Endothelial, Ganglion
Salivary Gland	25.1 ± 1.3	2/3	Glandular	12.2 ± 1.4	Glandular
Rectum	24.1 ± 4.2	2/3	Glandular	25.7 ± 1.4 *	Glandular
Gallbladder	21.3 ± 2.5	2/3	Glandular	12.3 ± 1.6	Glandular
Stomach	21.0 ± 3.9	2/3	Glandular	25.3 ± 8.5 *	Glandular
Small Intestine	18.6 ± 0.7	2/3	Glandular	18.9 ± 0.2	Glandular
Fallopian Tube	18.3 ± 5.0	2/3	Glandular	23.4 ± 1.0	Glandular
Duodenum	17.5 ± 0.9	2/3	Glandular	20.3 ± 0.6	Glandular
Lung	16.8 ± 2.5	2/3	-	14.5 ± 0.2	Macrophages, Pneumocytes
Esophagus	16.4 ± 0.4	2/3	Squamous Epithelial	11.9 ± 1.1	Squamous epithelial cells
Prostate	11.5 ± 1.7	2/3	Glandular	13.8 ± 1.0	Glandular
Kidney	11.0 ± 0.9	2/3	Glomeruli and Tubules	9.7 ± 0.9	Glomeruli and Tubules
Thyroid Gland	9.7 ± 2.1	1	-	14.4 ± 3.4	Glandular
Smooth Muscle	7.8 ± 1.1	2/3	-	21.2 ± 4.9	Smooth muscle
Liver	6.5 ± 1.1	1	-	14.6 ± 2.5	Bile duct, Hepatocytes
Heart Muscle	6.1 ± 1.7	1	-	8.7 ± 1.5	Myocytes
Appendix	5.2 ± 0.5	2/3	Glandular	9.3 ± 1.2	Glandular, Lymphoid
Pancreas	5.1 ± 0.4	2/3	Glandular	8.2 ± 1.6	Glandular, Islets
Spleen	4.6 ± 0.9	1	-	23.8 ± 0.6 *	Red Pulp, White Pulp
Endometrium	4.0 ± 1.8	2/3	Glandular	22.1 ± 1.3	Glandular, endometrial stroma
Urinary Bladder	3.7 ± 1.5	2/3	Glandular	12.7 ± 4.1	Urothelial
Placenta	3.7 ± 1.1	2/3	-	19.4 ± 2.4	Decidual, Trophoblastic
Tonsil	3.0 ± 0.5	1	-	4.9 ± 0.2	Squamous Epithelial
Adrenal Gland	2.4 ± 0.3	2/3	-	9.1 ± 0.3	Glandular
Cerebral Cortex	2.2 ± 0.5	1	-	8.5 ± 2.8	Neuronal, Glial, Endothelial
Testis	1.7 ± 0.3	1	-	10.3 ± 1.2	Seminiferous Ducts, Leydig
Skin	1.6 ± 0.2	2/3	-	7 ± 0.5	Fibroblasts, Langerhans, Melanocytes
Adipose Tissue	1.0 ± 0.4	1	-	23.3 ± 2.7	Adipocytes, Fibroblasts
Ovary	0.5 ± 0.1	-	-	27.8 ± 1.6 *	Ovarian Stroma
Lymph Node	0.4 ± 0.1	-	-	5.9 ± 0.8	Germinal Center, Non-germinal Center
Skeletal Muscle	0.1 ± 0.1	-	-	4.0 ± 0.4	Myocytes
Bone Marrow	0.0 ± 0.0	-	-	4.7 ± 0.9	Hematopoietic

Table S3 gnomAD human variants within *SHROOM3*. Transcript ENST0000029604 annotated variants were grouped and all missense variants (red) analyzed with PolyPhen2 as either benign, possibly damaging or probably damaging followed by impact scoring for those predicted probably damaging (score>100).

Annotation	# variants
3' UTR	278
frameshift	21
inframe deletion	13
intron	298
Non-coding transcript exon	0
splice acceptor	1
splice donor	4
splice region	37
start lost	0
stop gained	11
synonymous	447
missense	1043
benign	518
possibly damaging	195
probably damaging	330
High Impact Variants (>100 Score)	35

Table S4 Top 35 high impact gnomAD human variants within *SHROOM3*.

RSID	Position	Protein Consequence	Triplet	Normalized dN-dS	Conservation Score	21 codon window	PolyPhen2	Allele Count	Impact
rs139272770	chr4:77659945 C/T	His207Tyr	CAT	-0.890867	1.25	14.75	probably damaging	846	15598.13
rs145112769	chr4:77652057 G/C	Gly186Arg	GGC	-0.476875	1	9.25	probably damaging	1611	14901.75
rs147732653	chr4:77662361 C/A	Thr1012Asn	ACT	-2.27092	0.5	15.5	probably damaging	1630	12632.5
rs144435434	chr4:77476772 G/T	Gly60Val	GGG	-1.31398	0.25	13.25	probably damaging	538	1782.125
rs201389959	chr4:77660493 T/A	Ser389Arg	AGT	-3.04693	2	15.5	probably damaging	56	1736
rs371533598	chr4:77662358 T/C	Leu1011Pro	CTG	-0.312924	1	16	probably damaging	68	1088
rs199558629	chr4:77680831 G/C	Asp1778His	GAT	-2.4882	1.5	13.25	probably damaging	38	755.25
rs141646361	chr4:77676155 G/A	Glu1507Lys	GAA	-1.69351	0.25	6.75	probably damaging	424	715.5
rs141239704	chr4:77476831 A/T	Asn80Tyr	AAT	-2.15082	1.5	14	probably damaging	34	714
rs76406459	chr4:77659958 G/A	Arg211Lys	AGG	-0.874031	1.25	13.5	probably damaging	41	691.875
rs376718092	chr4:77691999 G/A	Arg1857His	CGT	-2.19996	1.5	21.5	probably damaging	20	645
rs199533144	chr4:77357309 G/T	Gly35Val	GGA	-1.9075	1.5	15.25	probably damaging	28	640.5
rs745611618	chr4:77700166 A/C	Lys1943Gln	AAG	-2.42386	1.5	20.75	probably damaging	19	591.375
rs181194611	chr4:77663057 C/T	Pro1244Leu	CCC	-0.238437	1	10.5	probably damaging	50	525
rs367782393	chr4:77662214 C/T	Ser963Leu	TCG	-2.86125	2	10.75	probably damaging	23	494.5
rs181584053	chr4:77691937 G/A	Met1836Ile	ATG	0	1	18.5	probably damaging	24	444
rs370967544	chr4:77660063 T/C	Ile246Thr	ATT	-1.8975	1.5	16.5	probably damaging	17	420.75
rs144913986	chr4:77476909 G/C	Val106Leu	GTG	-1.32553	0.25	7.75	probably damaging	183	354.5625
rs775459440	chr4:77691866 G/A	Ala1813Thr	GCC	-1.43062	1.25	18.5	probably damaging	14	323.75
rs761396961	chr4:77476813 G/A	Asp74Asn	GAT	-3.96299	2	10.75	probably damaging	14	301
rs777929615	chr4:77659915 T/C	Ser197Pro	TCC	-0.95375	1.25	15	probably damaging	13	243.75
rs533735013	chr4:77660489 G/A	Arg388Gln	CGG	-1.28941	1.25	15.25	probably damaging	10	190.625
rs769359406	chr4:77660732 C/T	Pro469Leu	CCG	-2.38437	0.5	15.75	probably damaging	21	165.375
rs146652221	chr4:77660381 G/A	Arg352Gln	CGG	-0.932178	0.25	2.75	probably damaging	238	163.625
rs371504856	chr4:77660201 G/A	Arg292Gln	CGA	-1.43898	1.25	8	probably damaging	15	150
rs79007254	chr4:77660567 G/A	Arg414Gln	CGG	-2.05531	1.5	6.5	probably damaging	15	146.25
rs780349816	chr4:77660717 C/T	Pro464Leu	CCG	-2.74088	1	15.75	probably damaging	8	126
rs752388322	chr4:77677916 T/C	Ile1675Thr	ATT	-1.97434	1.5	16.5	probably damaging	5	123.75
rs144435434	chr4:77476772 G/C	Gly60Ala	GGG	-1.31398	0.25	13.25	probably damaging	36	119.25
rs147919147	chr4:77676154 C/G	Asp1506Glu	GAC	-5.24274	1	7	probably damaging	17	119
rs767412535	chr4:77357350 G/A	Gly49Arg	GGA	-1.19219	1.25	18.75	probably damaging	5	117.1875
rs377199268	chr4:77660023 C/A	Pro233Thr	CCA	-1.09899	0.25	11.5	probably damaging	40	115
rs770837589	chr4:77700037 C/T	Arg1900Cys	CGC	-1.77718	1.5	19	probably damaging	4	114
rs757040989	chr4:77651977 G/A	Arg159Gln	CGA	-1.03479	1.25	8.25	probably damaging	10	103.125
rs751264874	chr4:77662571 A/G	Gln1082Arg	CAG	-0.407049	1	14.5	probably damaging	7	101.5

Table S5 Top 35 high impact gnomAD human variants within *SHROOM3* for allele frequencies.

RSID	Protein Consequence	African	Ashkenazi Jewish	European (Finnish)	European (Non-Finnish)	Latino	East Asian	South Asian	Other
rs139272770	His207Tyr	0.054	0.000	0.826	0.464	0.029	0.000	0.000	0.356
rs145112769	Gly186Arg	0.012	0.000	0.000	0.016	0.003	0.011	5.087	0.294
rs147732653	Thr1012Asn	0.193	0.981	0.310	1.020	0.586	0.000	0.201	0.854
rs144435434	Gly60Val	0.042	0.345	0.054	0.279	0.218	0.005	0.101	0.294
rs201389959	Ser389Arg	0.013	0.000	0.000	0.017	0.079	0.000	0.000	0.062
rs371533598	Leu1011Pro	0.000	0.000	0.000	0.000	0.000	0.000	0.227	0.034
rs199558629	Asp1778His	0.000	0.000	0.004	0.011	0.012	0.000	0.068	0.000
rs141646361	Glu1507Lys	0.017	0.000	0.307	0.247	0.026	0.005	0.023	0.202
rs141239704	Asn80Tyr	0.004	0.000	0.000	0.026	0.000	0.000	0.000	0.000
rs76406459	Arg211Lys	0.000	0.000	0.043	0.024	0.000	0.000	0.000	0.000
rs376718092	Arg1857His	0.004	0.000	0.012	0.006	0.020	0.005	0.003	0.000
rs199533144	Gly35Val	0.000	0.000	0.000	0.025	0.000	0.000	0.000	0.000
rs745611618	Lys1943Gln	0.000	0.000	0.000	0.002	0.036	0.000	0.000	0.091
rs181194611	Pro1244Leu	0.000	0.000	0.000	0.000	0.000	0.268	0.000	0.000
rs367782393	Ser963Leu	0.014	0.022	0.000	0.000	0.000	0.191	0.000	0.028
rs181584053	Met1836Ile	0.000	0.000	0.004	0.017	0.003	0.000	0.003	0.000
rs370967544	Ile246Thr	0.013	0.000	0.000	0.002	0.018	0.017	0.010	0.018
rs144913986	Val106Leu	0.200	0.246	0.000	0.045	0.122	0.000	0.003	0.155
rs775459440	Ala1813Thr	0.004	0.000	0.000	0.002	0.000	0.058	0.000	0.000
rs761396961	Asp74Asn	0.000	0.000	0.054	0.001	0.000	0.000	0.000	0.018
rs777929615	Ser197Pro	0.000	0.000	0.000	0.000	0.039	0.000	0.000	0.000
rs533735013	Arg388Gln	0.000	0.010	0.000	0.003	0.003	0.021	0.000	0.000
rs769359406	Pro469Leu	0.004	0.000	0.000	0.009	0.000	0.000	0.029	0.000
rs146652221	Arg352Gln	0.054	0.355	0.066	0.132	0.009	0.000	0.000	0.031
rs371504856	Arg292Gln	0.025	0.000	0.000	0.006	0.000	0.000	0.000	0.031
rs79007254	Arg414Gln	0.000	0.000	0.000	0.005	0.003	0.042	0.000	0.000
rs780349816	Pro464Leu	0.012	0.000	0.000	0.002	0.000	0.005	0.003	0.000
rs752388322	Ile1675Thr	0.004	0.000	0.000	0.003	0.000	0.000	0.000	0.000
rs144435434	Gly60Ala	0.000	0.000	0.132	0.000	0.003	0.005	0.000	0.000
rs147919147	Asp1506Glu	0.000	0.000	0.000	0.013	0.003	0.000	0.000	0.000
rs767412535	Gly49Arg	0.007	0.000	0.000	0.000	0.000	0.012	0.006	0.000
rs377199268	Pro233Thr	0.000	0.000	0.000	0.000	0.006	0.201	0.000	0.000
rs770837589	Arg1900Cys	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000
rs757040989	Arg159Gln	0.000	0.000	0.000	0.001	0.009	0.006	0.016	0.000
rs751264874	Gln1082Arg	0.036	0.000	0.000	0.000	0.000	0.000	0.000	0.000

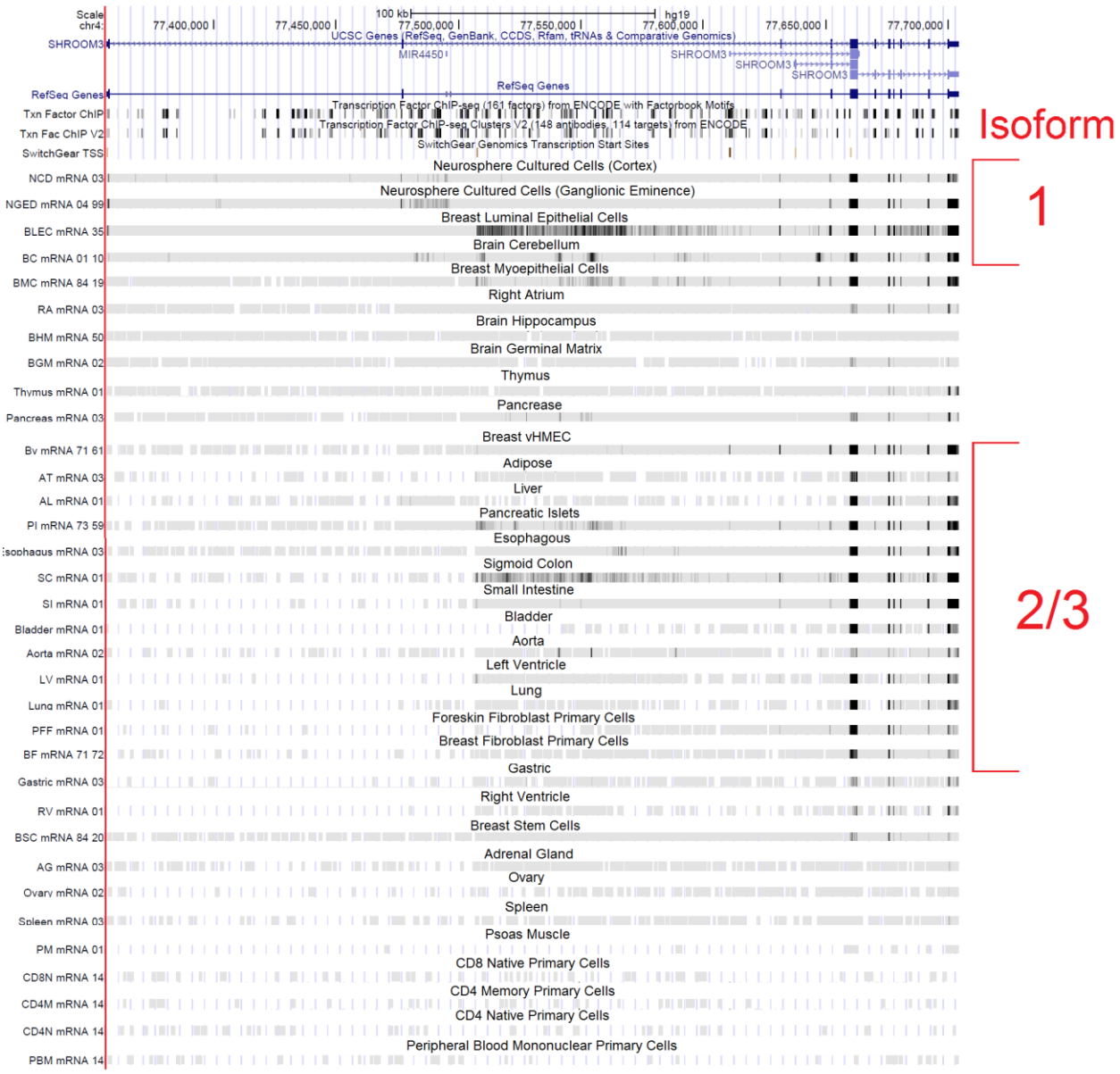
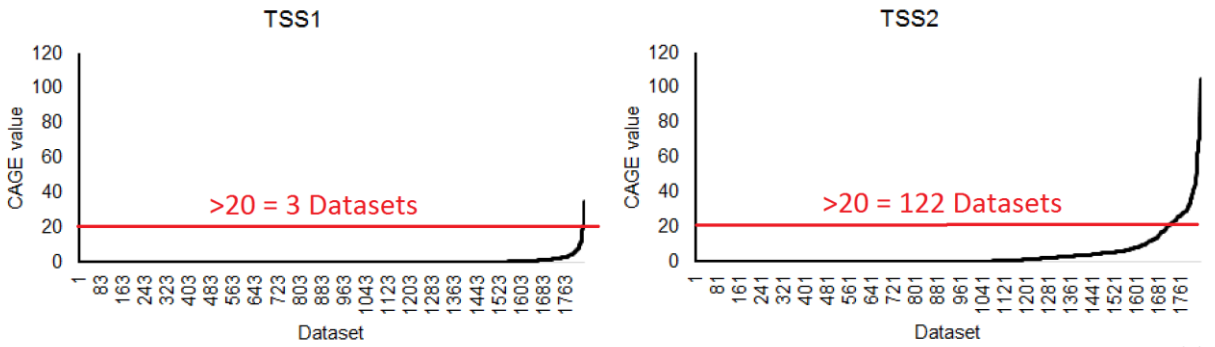


Figure S1 RNAseq of adult tissues from Roadmap Epigenomics. Darker black represents a higher density of reads mapped to the genome through RNAseq. Isoform 1 is identified by containing expression of exon 1 found on the left most side. This figure shows that isoform 1 is only used in a few cells such as neural while transcripts not including exon 1 are found in more tissues types.



experiment name	pos_strand_value
Astrocyte - cerebral cortex, donor1 : CNhs10864 ctss	35.207
small cell lung carcinoma cell line:LK-2 : CNhs11285 ctss	21.008
Ciliary Epithelial Cells, donor1 : CNhs10871 ctss	20.892
colon carcinoma cell line:COLO-320 : CNhs10737 ctss	19.908
signet ring carcinoma cell line:NUGC-4 : CNhs11270 ctss	19.507
Lens Epithelial Cells, donor2 : CNhs12568 ctss	19.476
iPS differentiation to neuron, down-syndrome donor C18-CCL54, day06, rep2 : CNhs14061 ctss	18.783
Astrocyte - cerebral cortex, donor3 : CNhs12005 ctss	15.818
Astrocyte - cerebral cortex, donor2 : CNhs11960 ctss	15.056
Neural stem cells, donor1 : CNhs11063 ctss	13.975
anaplastic squamous cell carcinoma cell line:RPMI 2650 : CNhs11889 ctss	12.245
Ciliary Epithelial Cells, donor3 : CNhs12009 ctss	12.033
gastric cancer cell line:AZ521 : CNhs11286 ctss	11.912
serous adenocarcinoma cell line:JHOS-2 : CNhs11746 ctss	11.455
Retinal Pigment Epithelial Cells, donor0 : CNhs10842 ctss	11.284
iPS differentiation to neuron, down-syndrome donor C18-CCL54, day06, rep3 : CNhs14064 ctss	11.009
testicular germ cell embryonal carcinoma cell line:NEC15 : CNhs12362 ctss	10.844
adipose, donor1 : CNhs13972 ctss	10.754
iPS differentiation to neuron, control donor C32-CRL1502, day06, rep1 : CNhs13827 ctss	8.935
iPS differentiation to neuron, down-syndrome donor C18-CCL54, day06, rep1 : CNhs13836 ctss	8.617

experiment name	pos_strand_value
small intestine, adult, pool1 : CNhs10630 ctss	104.571
lung, fetal, donor1 : CNhs11680 ctss	83.475
signet ring carcinoma cell line:NUGC-4 : CNhs11270 ctss	78.149
colon, adult, pool1 : CNhs10619 ctss	75.72
Alveolar Epithelial Cells, donor2 : CNhs12084 ctss	70.478
serous cystadenocarcinoma cell line:HTOA : CNhs11827 ctss	69.239
gall bladder, adult : CNhs12848 ctss	66.811
colon, adult, donor1 : CNhs11794 ctss	65.687
salivary gland, adult, pool1 : CNhs11677 ctss	63.674
seminal vesicle, adult : CNhs12851 ctss	63.508
aorta, adult, pool1 : CNhs11760 ctss	62.962
duodenum, fetal, donor1, tech rep1 : CNhs11781 ctss	55.656
bile duct carcinoma cell line:HuCCT1 : CNhs10750 ctss	54.55
papillotubular adenocarcinoma cell line:TGBC18TKB : CNhs10734 ctss	53.672
bronchioalveolar carcinoma cell line:NCI-H358 : CNhs11840 ctss	51.359
pituitary gland, adult, donor10252 : CNhs12229 ctss	48.492
submaxillary gland, adult : CNhs12852 ctss	47.249
lung, adult, pool1 : CNhs10625 ctss	45.598
Oligodendrocyte - precursors, donor1 : CNhs12586 ctss	45.108
lung, right lower lobe, adult, donor1 : CNhs11786 ctss	43.495

Figure S2 FANTOM CAGE data for either TSS1 or TSS2 of SHROOM3 in 1,830 datasets. More datasets use TSS2 as a start site than do TSS1 with higher positive strand values. Below each plot are the top ten tissues with noticeable elevation of neuron and astrocyte cells using TSS1 and a list of more diverse organs that use TSS2.

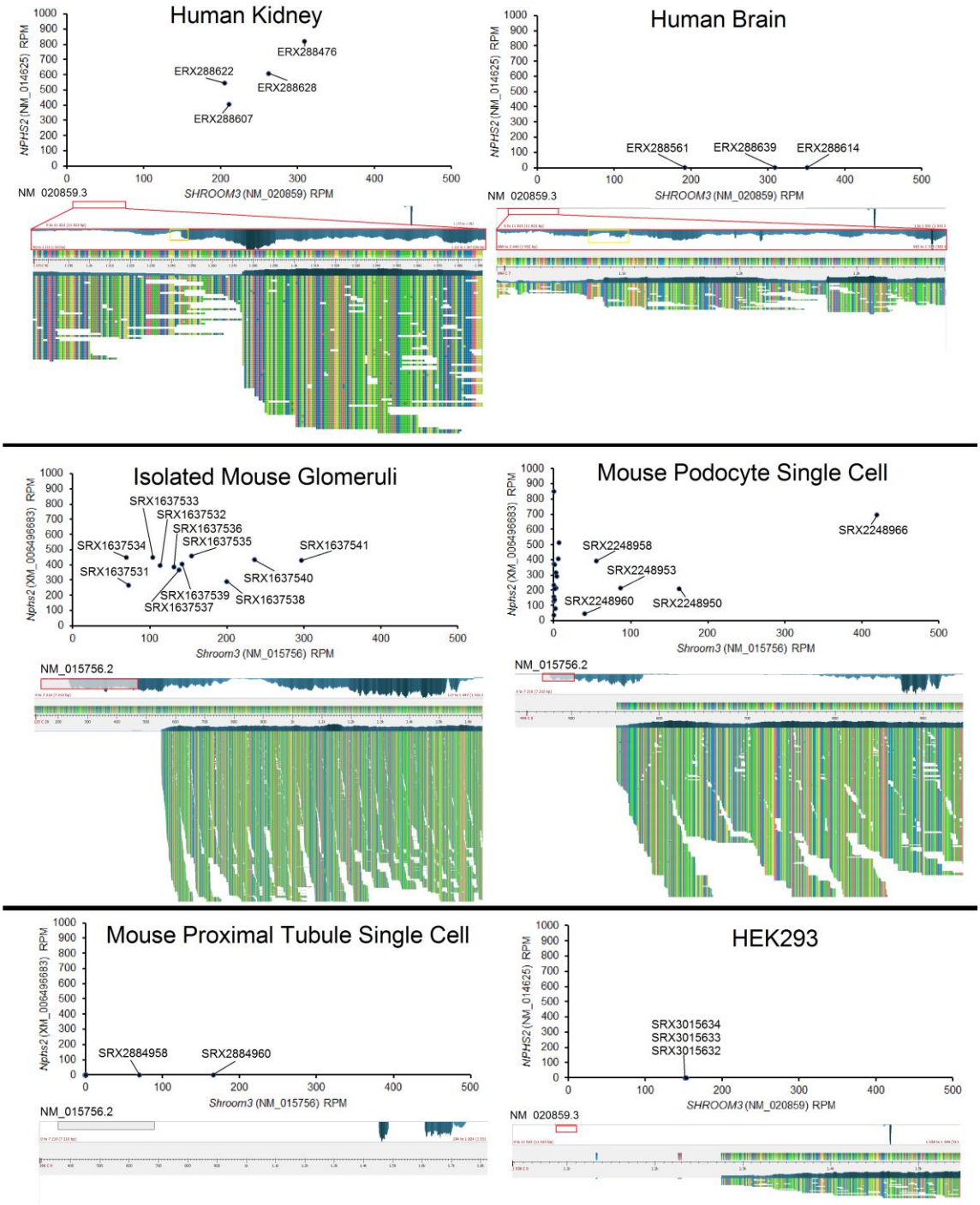


Figure S3 Transcript mapping of *SHROOM3* using publicly available RNAseq datasets. BLAST analysis was performed using *SHROOM3*/*Shroom3* transcripts (NM_015756/NM_020859) or *NPHS2*/*Nphs2* and plotted as reads mapped per million reads sequenced (RPM) for six different clusters of RNAseq datasets. Only whole kidney, isolated glomeruli, and single cell podocytes have expression of *NPHS2*/*Nphs2*. Below the RPM plots are the Bowtie2 alignments of transcripts with a zoomed in view of the secondary transcriptional start site in either mouse or human.

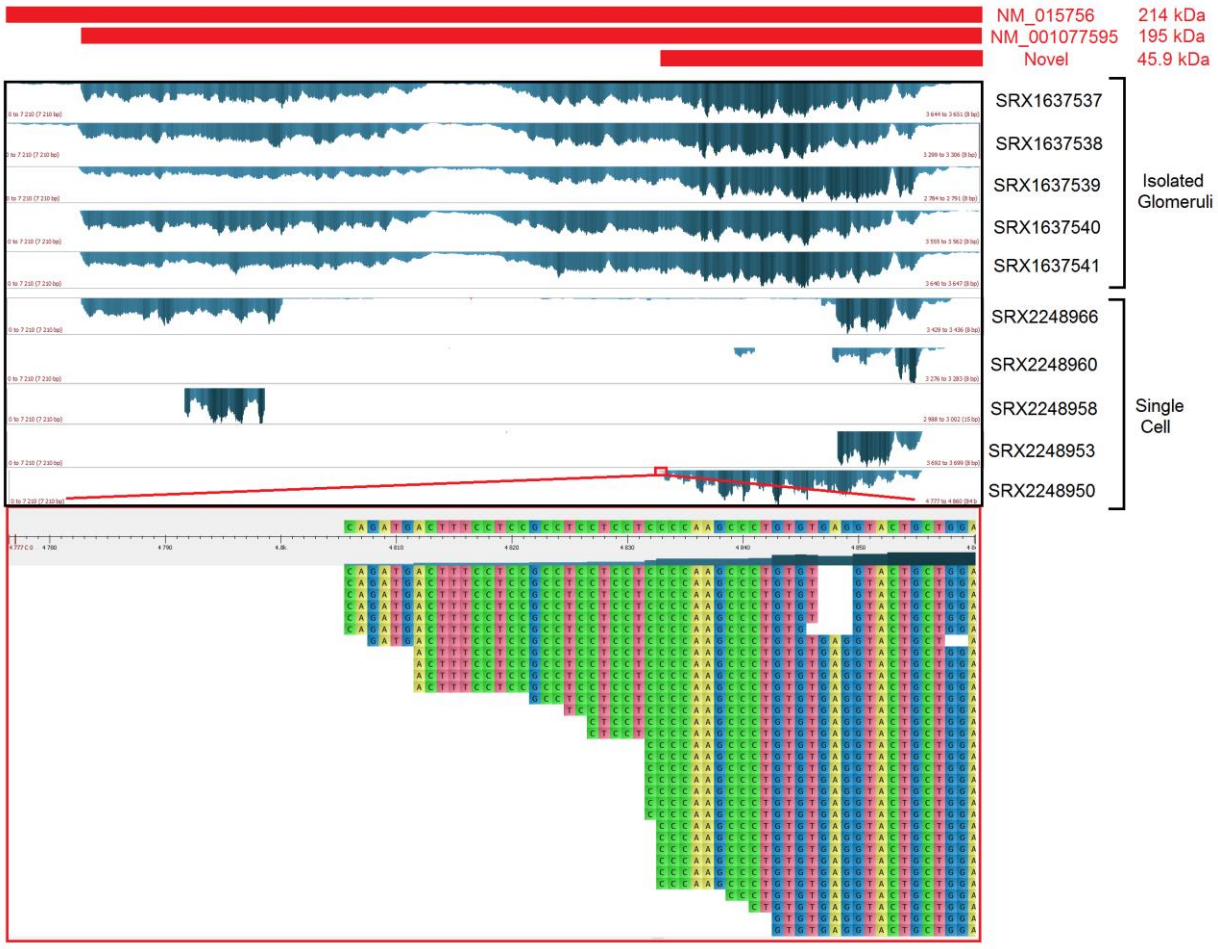


Figure S4 Transcripts in individual podocyte RNAseq datasets. Three transcripts of the mouse are shown above in red. Reads from five datasets for both the isolated glomeruli and for individual podocytes were aligned and shown for read depth. A third and novel transcript is shown in one of the isolated podocyte RNAseq datasets zoomed in in the red box. This transcript is predicted to generate a protein of 45.9 kDa that contains only the ASD2 domain.

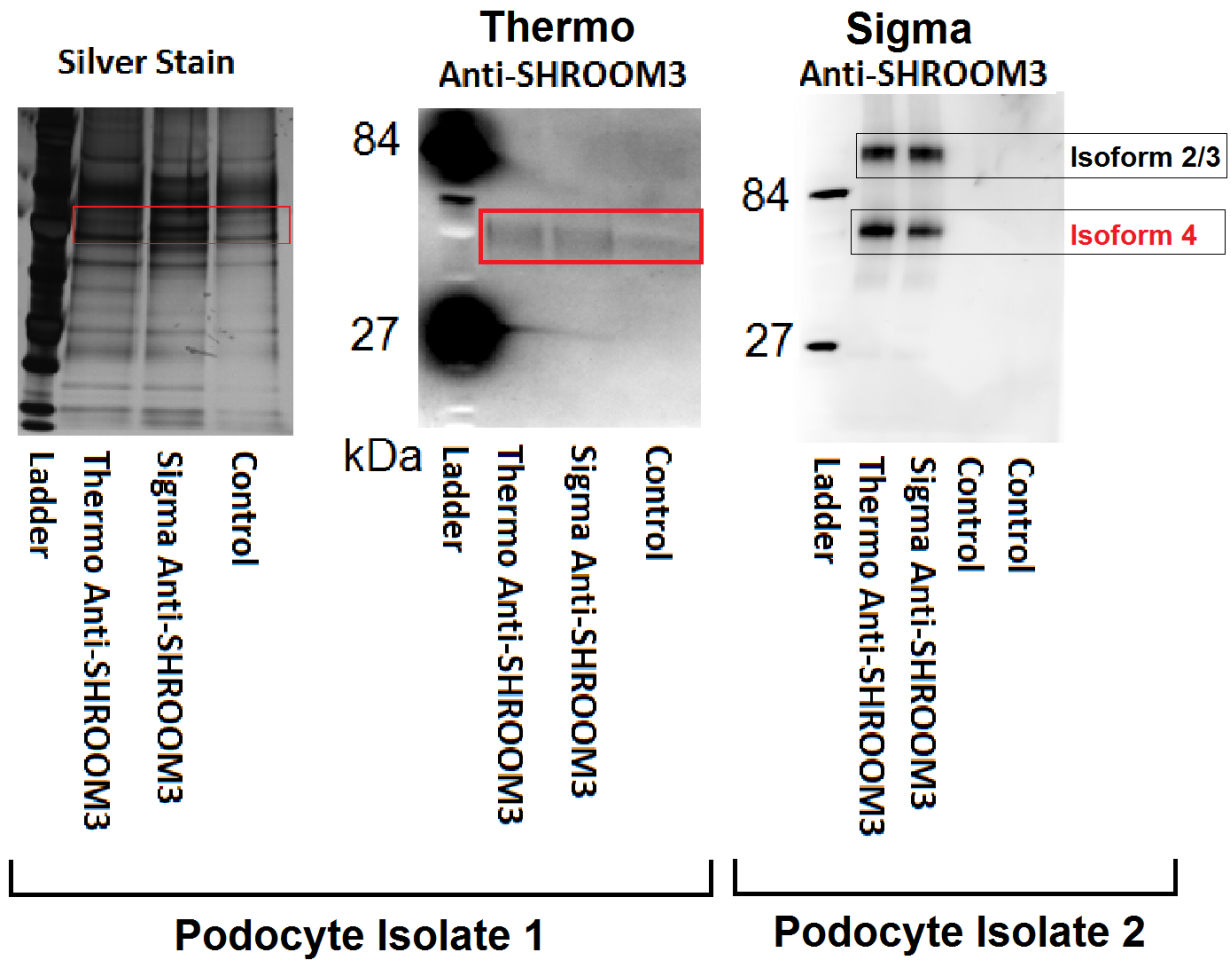


Figure S5 Immunoprecipitation of SHROOM3 from primary human podocytes. Two SHROOM3 antibodies and a negative antibody control (bottom) were used to capture proteins from cellular extracts made from two separate isolates of primary human podocytes. Captured proteins were then run on SDS-PAGE and proteins determined using either silver staining (left) or a western blot using a SHROOM3 antibody (right). Protein bands were identified at ~45kDa, the predicted size of the new SHROOM3 Isoform 4 and ~190kDa for the isoform2/3 predicted from Figure S4.

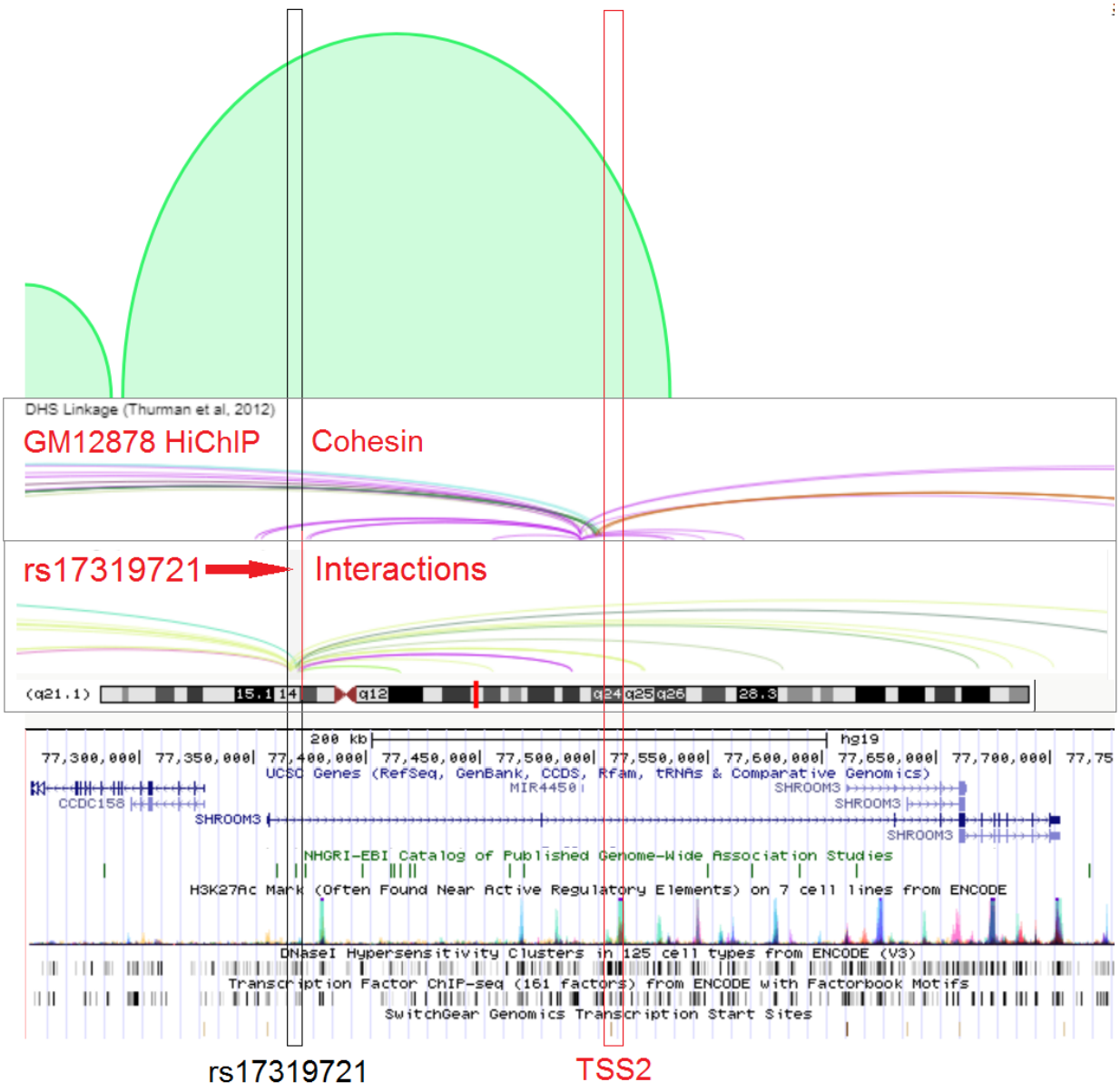


Figure S6 HiChIP data for GM12878, a cell line without SHROOM3 expression. On top is shown the DHS-Linkage, a compiled dataset for DNaseI-hypersensitive sites (DHS) of ~100 cell types of ENCODE correlating linkage in gene regulation areas. This suggests that regulation is correlated from before the first TSS through the second TSS. The GM12878 cell line does not express SHROOM3. The HiChIP data for Cohesin is shown in the middle and shows where cohesion, a factor involved in chromosome organization, links multiple regions of the genome together. Shown below the DHS data is either the global interaction sites for HiChIP, and below that is interaction sites directly around the rs17319721 variant. On the bottom is the human genome browser showing the location of SHROOM3, GWAS SNPs (green), H3K27Ac marks of ENCODE, DNaseI hypersensitive sites, transcription factor binding sites, and SwitchGear TSS sites. Boxes in black is the location of rs17319721 and boxed in red is the TSS2 site.

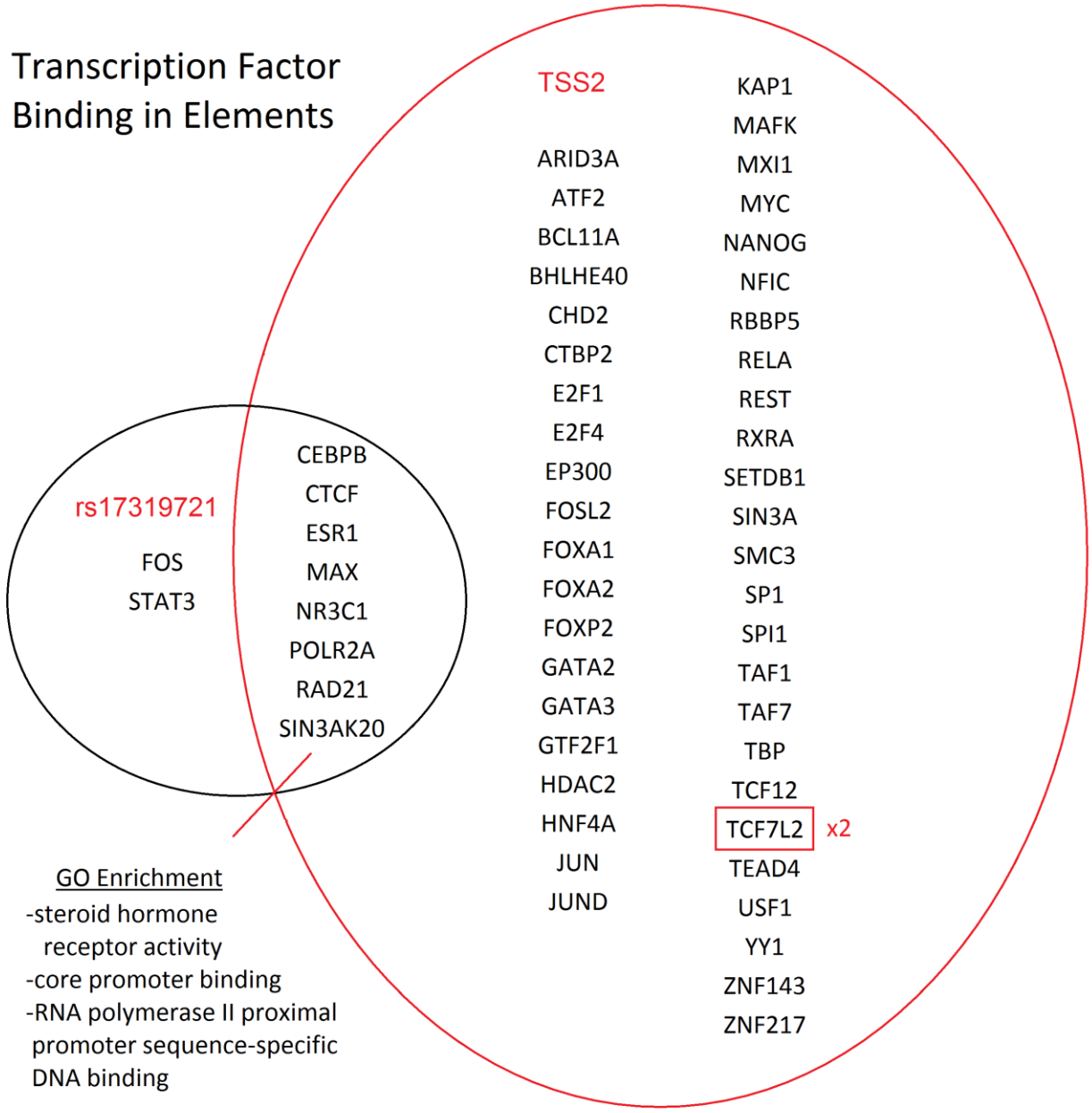


Figure S7 Transcription Factors (TFs) from ENCODE near rs17319721 and TSS2. The binding of TFs from Phase 1 transcription factor ChIP-Seq data, as determined by a threshold of the ENCODE consortium, for the region around rs17319721 (hg19: chr4:77,361,347-77,376,346) or the TSS2 (hg19: chr4:77487991-77513325). A total of 53 factors are bound to the TSS2 location and 10 near rs17319721 with 8 TFs overlapping between the two datasets including the well-known linker CTCF.

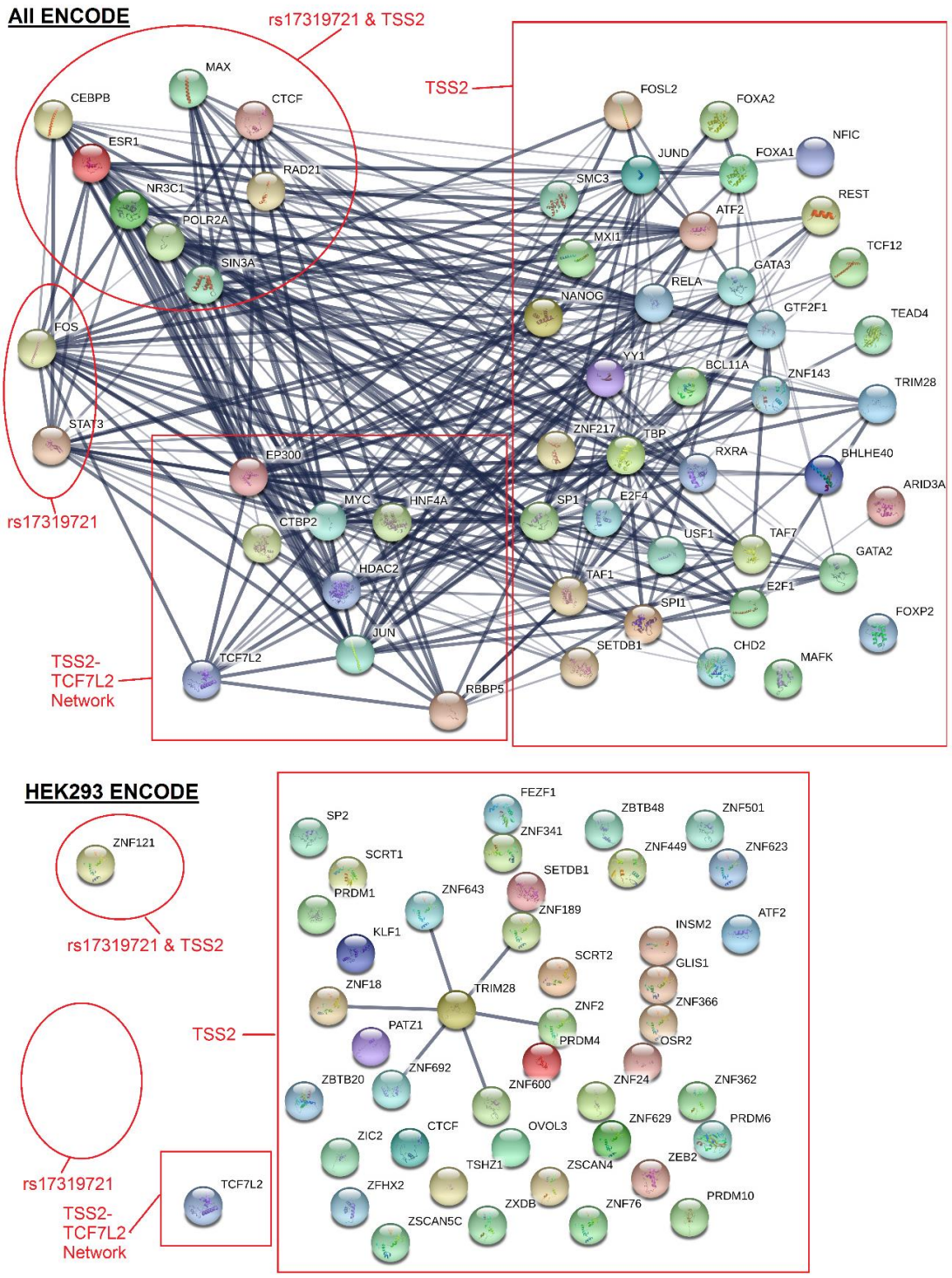


Figure S8 Pathway enrichment of TFs from all ENCODE cell ChIPseq and those of HEK293 for rs17319721 and TSS2 sites. STRING analysis was performed on the list of TFs from Figure S7 (top) or those TFs found in the HEK293 ChIP-Seq (bottom) for either rs17319721 alone, TSS2 alone, or found in both. It should be noted that TCF7L2 is found in the TSS2 site for both panels.

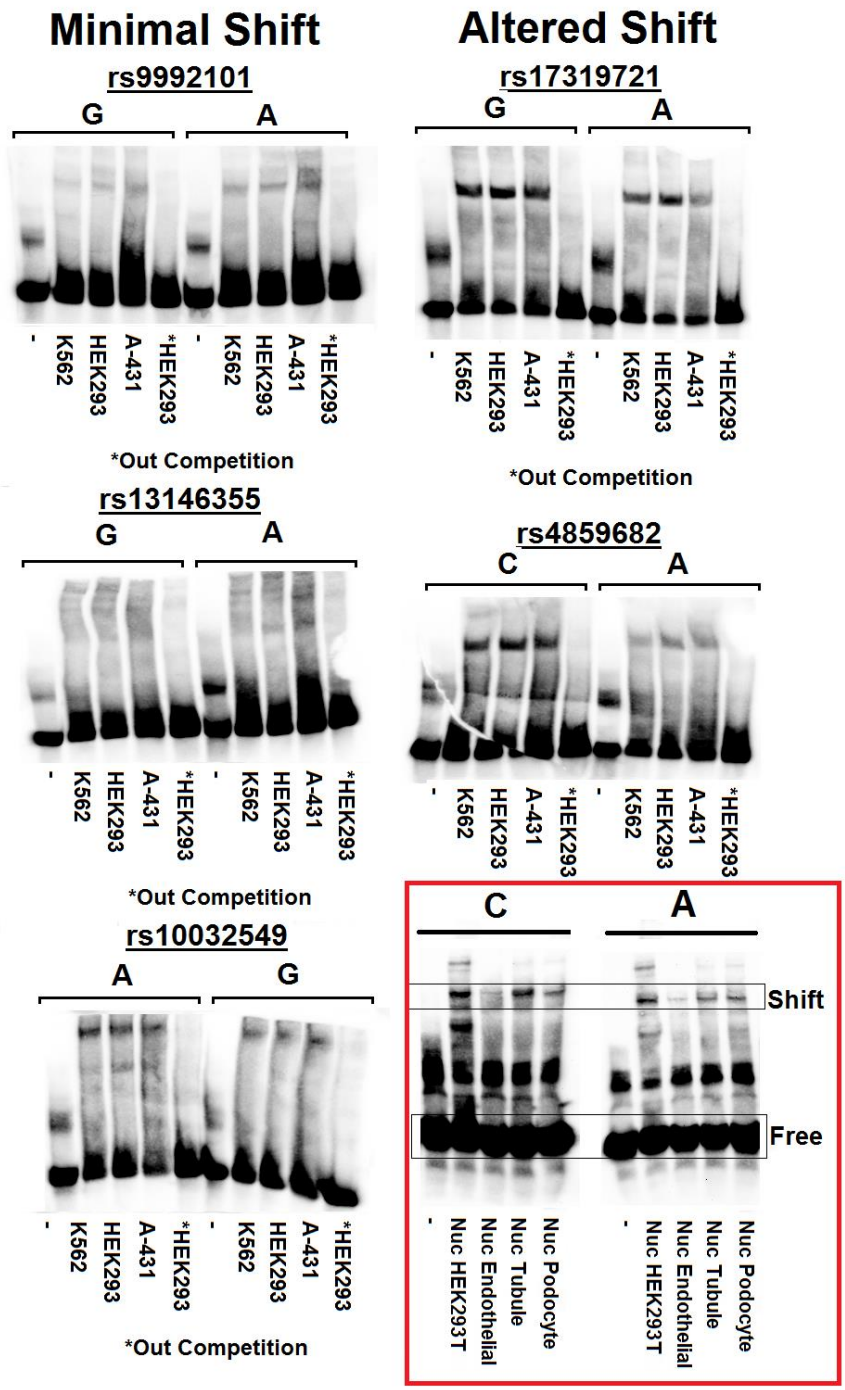


Figure S9 Testing of top five SNPs from RegulomeDB with EMSA. Each of the SNPs in LD were processed through RegulomeDB with scores shown: rs17319721 (2b: http://www.regulomedb.org/snp/chr4/77368846); rs1398016 (5); rs17253722 (5); rs28418670 (6); rs11724003 (5); rs1398018 (6); rs2870238 (No Data); rs4859682 (5); rs13146355 (5); rs9992101 (5); rs10023335 (No Data). Shifting of the major and minor allele in three cell lines that have variable expression of SHROOM3 for five probes. Cold outcompetition was performed on the HEK293 binding using 40X non-biotin tagged oligo probe in addition to the normal biotin probe. Two probes show promising shifting; however, rs4859682 does not have changes in binding to the primary kidney nuclear extracts (red box).

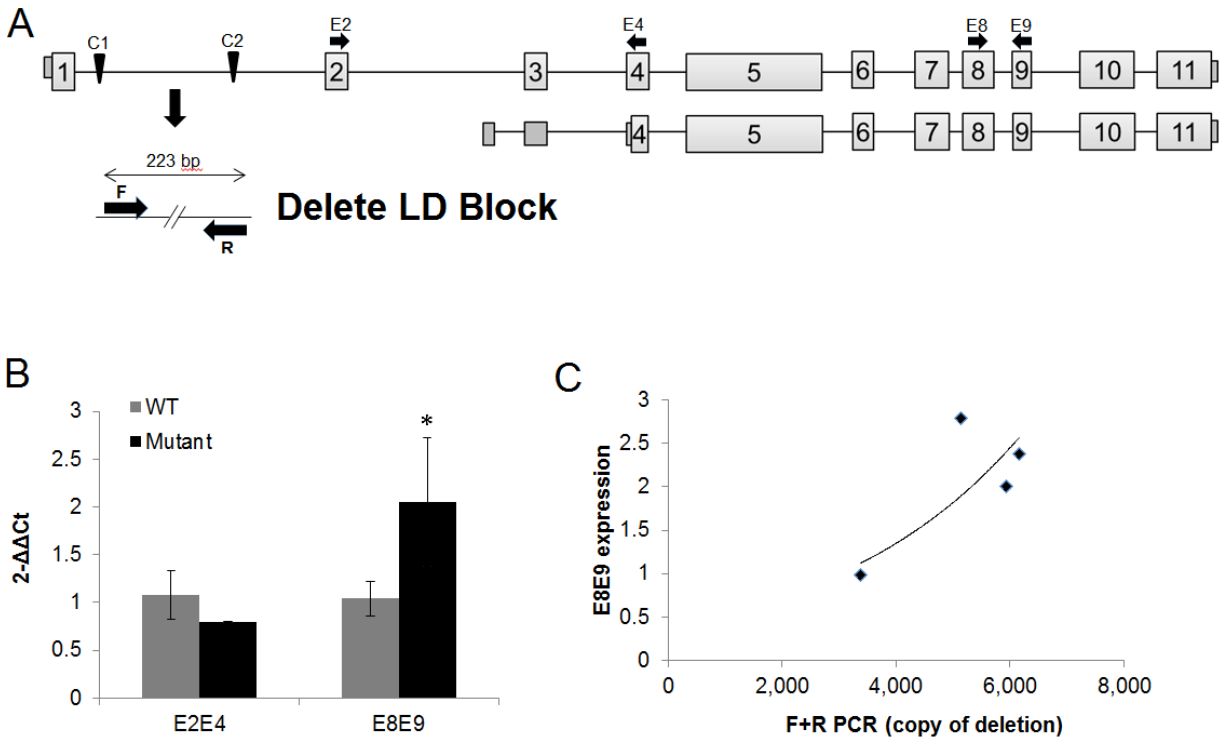


Figure S10 Removal of the entire SHROOM3 GWAS LD block changes regulation similar to rs17319721. A) A CRISPR/Cas9 system was designed to remove the entire LD block from C1 to C2, resulting in a 223 bp fragement that the F and R primers could detect successful CRISPR within single cell isolates. B) Change in expression of the E2E4 and E8E9 primer sets following the LD block removal as determined by q-RT-PCR. C) Number of modified alleles (x-axis) was determined by quantifying the F+R PCR for four separate single cell isolates and plotted relative to the E8E9 expression levels (y-axis) showing a correlation in expression change to the number of modified alleles within the cell.

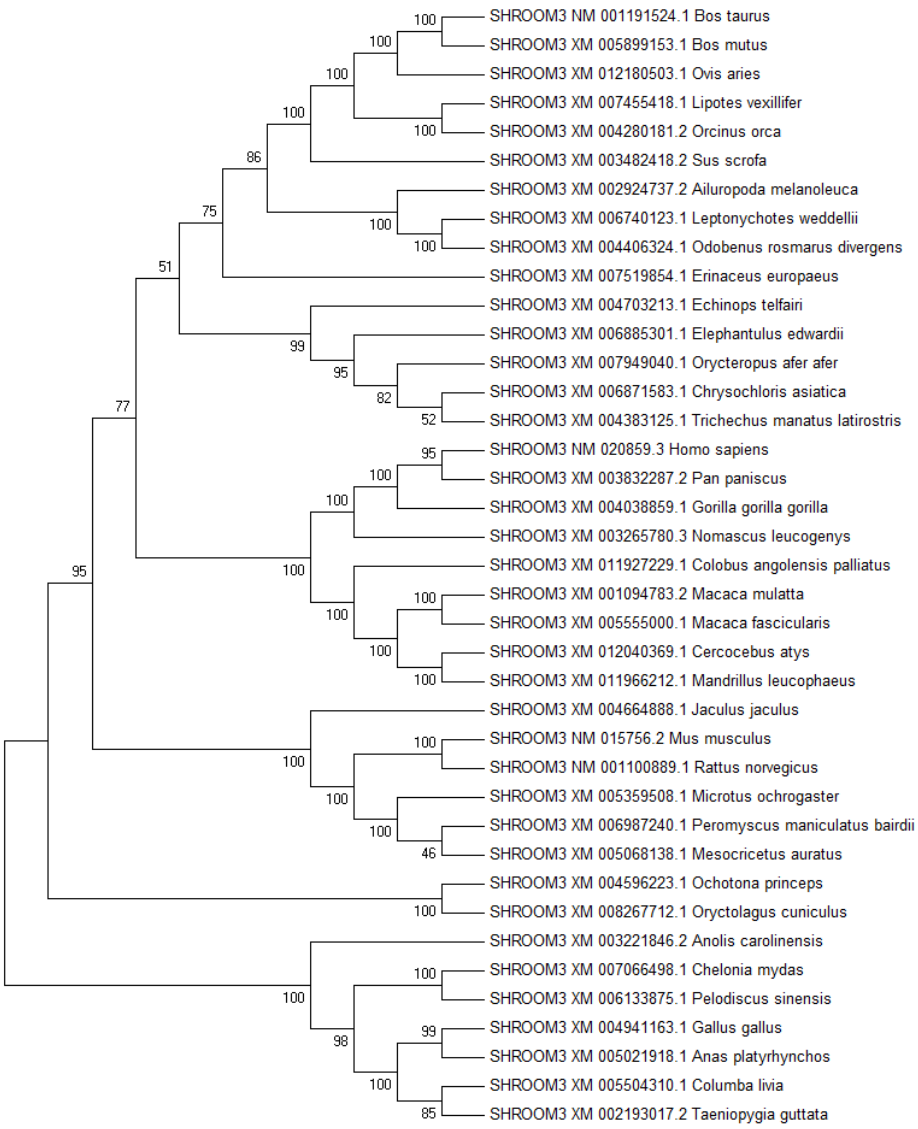


Figure S11 Molecular Phylogenetic analysis by Maximum Likelihood for full-length *SHROOM3* sequences. It can be noted that species included in our evolution range from mammals to reptiles (such as *Anolis*) and birds (such as *Gallus*). Methods for tree generation: The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The analysis involved 39 nucleotide sequences. All positions with less than 80% site coverage were eliminated. That is, fewer than 20% alignment gaps, missing data, and ambiguous bases were allowed at any position. There was a total of 5,856 positions in the final dataset.

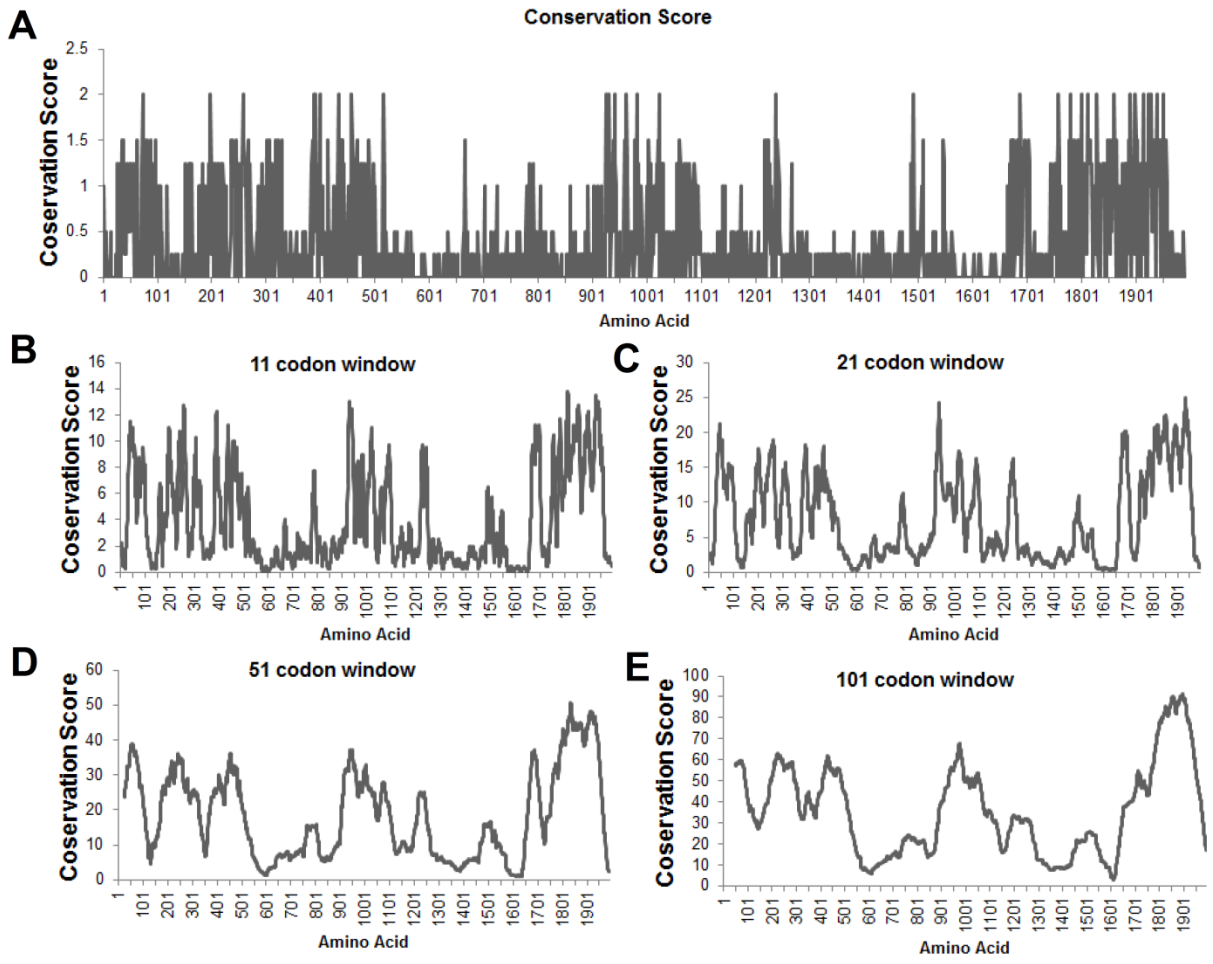


Figure S12 Sliding window for motif discovery of *Shroom3* open reading frame. A) The conservation of each amino acid was determined using the 39 sequences from Figure S2. **B-E)** The scores for each position were then put on a sliding window, ie summing the values for position X with the number of amino acids Y upstream and downstream to equal Z codons in the window. This was performed for a window of 11 (**B**), 21 (**C**), 51 (**D**), and 101 (**E**) codons. For example the 21-codon window position 101 is the combined values of evolutionary conservation for amino acid 90-111 and for 102 is 91-112.

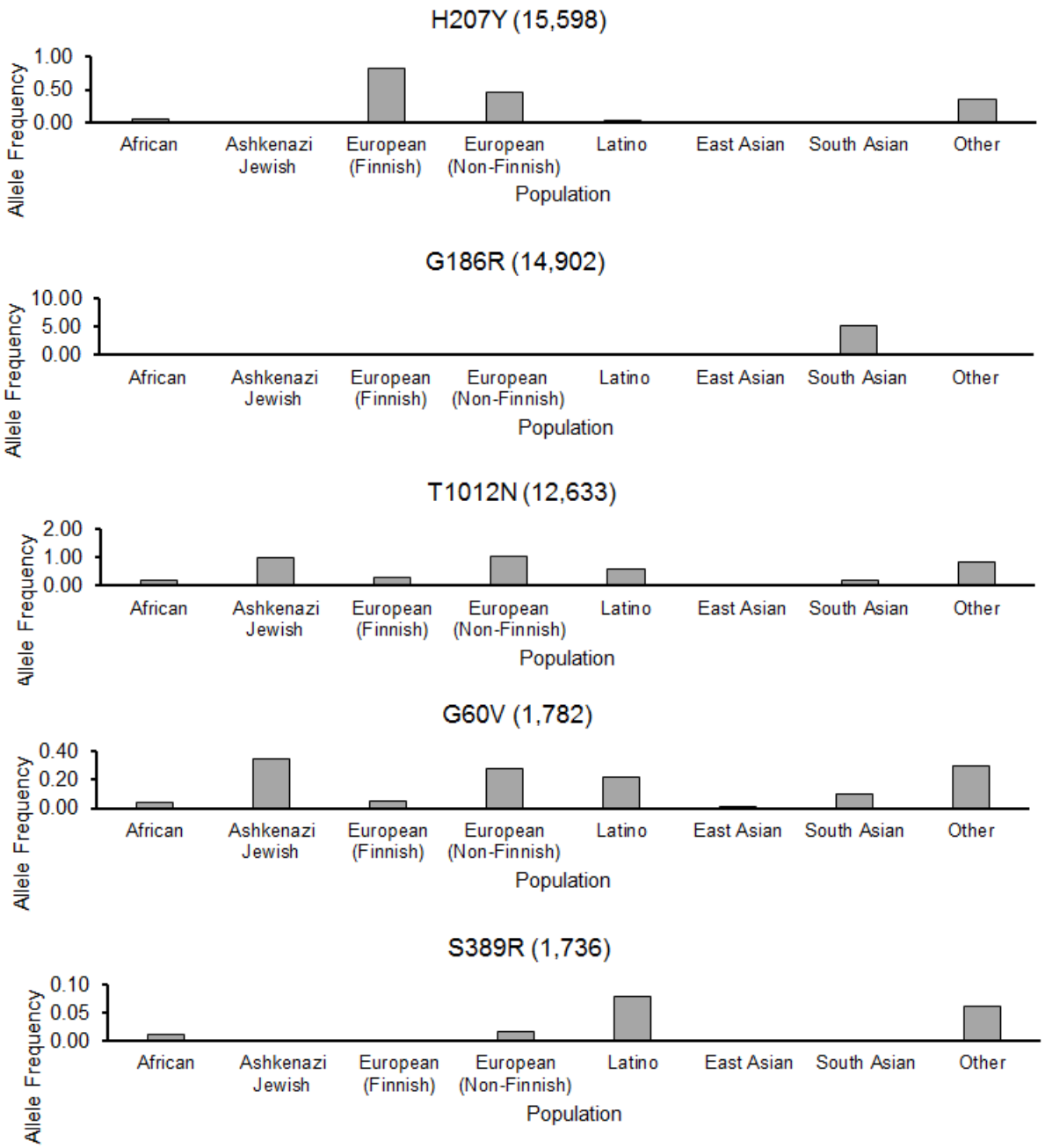


Figure S13 Analysis of variants from gnomAD. Allele frequency percentages from five variants within SHROOM3 with the top impact scores (given in parentheses) in the 8 populations of gnomAD.

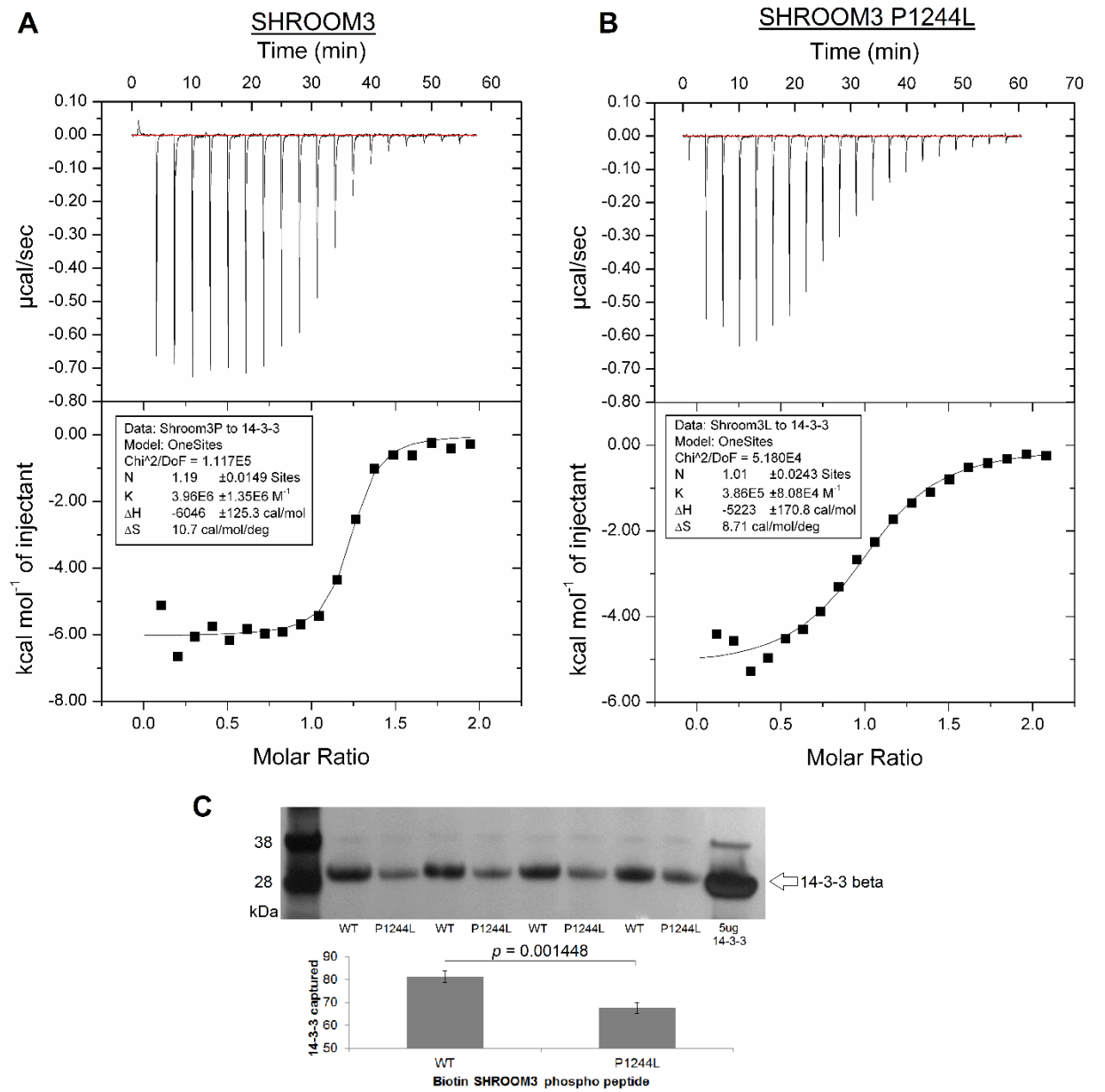


Figure S14 Biophysical characterization of the P1244L variant. **A-B)** Isothermal titration calorimetry (ITC) results of binding of the SHROOM3 (**A**) and SHROOM3 P1244L (**B**) peptides to 14-3-3. **C)** Affinity capture of recombinant 14-3-3 by a SHROOM3 phosphorylated peptides fixed to avidin in 4 separate capture experiments. The quantification of 6 independent capture experiments is shown below, with the error bars representing the SEM. Significance was determined with a Student's T-test.