Supplementary Information

Inducers of the endothelial cell barrier identified through chemogenomic screening in genome-edited hPSC-endothelial cells

Filip Roudnicky[a,1], Jitao David Zhang[c,1,3], Bo Kyoung Kim[b], Nikhil J. Pandya[c], Yanjun Lan[a], Lisa Sach-Peltason[e], Heloise Ragelle[b], Pamela Strassburger[b], Sabine Gruener[b], Mirjana Lazendic[b], Sabine Uhles[b], Franco Revelant[b], Oliv Eidam[e], Gregor Sturm[c], Verena Kueppers[b], Klaus Christensen[a], Leonard D Goldstein[f], Manuel Tzouros[c], Balasz Banfai[c], Zora Modrusan[f], Martin Graf[a], Christoph Patsch[a], Mark Burcin[a], Claas A. Meyer[a,2], Peter D. Westenskow[b,2,3], Chad A Cowan[g,h,i,2,3]

[3]To whom correspondence may be addressed.

Chad A Cowan, Jitao David Zhang, and Peter D. Westenskow

ccowan@bidmc.harvard.edu, jitao_david.zhang@roche.com, and peter.westenskow@roche.com.

www.pnas.org/cgi/doi/10.1073/pnas.1911532117

Supplementary Materials and Methods:

Integration of a GFP reporter gene into the CLDN5 locus using CRISPR/Cas9 genome editing and PiggyBac excision. A Cas9 targeting site close to the stop codon of the CLDN5 gene (GCGAGGCGTTGGATAAGCCT) was chosen, and complementary sgRNA was produced by *in vitro* transcription [Thermo Fisher(1)]. The vector construct (*SI Appendix*, Table S1 and Fig. S1B) for GFP integration through homologous recombination repair after CRISPR/Cas9 DNA double-stranded breakage was designed with homology arms flanking the left (694 bp) and right (518 bp) side of the stop codon of the human CLDN5 gene. The ATAA site, 61 nucleotides downstream of the stop codon of CLDN5, was changed to TTAA in the right homologous recombination arm to allow PiggyBac excision of the vector resistance cassette (for puromycin) and a truncated thymidine kinase gene (tTK) under the control of an EF1A promoter. Inverted terminal repeat (ITR) sequences allowing PiggyBac excision, and LoxP sites allowing Cre recombinase excision, mediated the removal of the resistance cassette. The hPSCs were pretreated with 10 μM of Y-27632 (Calbiochem) for 4 h, after which 200,000 cells were nucleofected using an Amaxa 4D-Nucleofector (Lonza) with a P3 Primary Cell 4D-Nucleofector kit (Lonza) using the CM130 program with 10.8 μg of sgRNA, 8 μg of Cas9, and 2.4 μg of plasmid vector donor. After nucleofection, the cells were treated with 10 μM of Y-27632 for 24 h. Cells were left to recover for 5 days and then expanded under selection with puromycin (200 ng/mL). After selection, cells were nucleofected with excision-only PiggyBac mRNA transposase (1.75 μg). Cells were seeded in serial dilution (1-300 cells/cm^2) on several tissue culture plates. Well-separated single cell colonies were picked after reaching 200 μm in diameter. Cells were washed and held in PBS. Colonies were detached by scratching with a sterile pipette tip, pipetting the colony, and re-plating it on a Matrigel-coated 48-well plate containing mTeSR1 medium. After 4 h, the medium was replaced with fresh mTeSR1 medium and further treated with 10 μM Y-27632 for 24 h. Cells were expanded in mTeSR1 until confluency. DNA was isolated using a BioSprint 96 DNA Blood Kit (Qiagen) according to the manufacturer protocol. Excision of the resistance cassette was evaluated by qPCR using primers that bind in the TK coding sequence (fwr- GTACCCGAGCCGATGACTTAC, rev- CCCGGCCGATATCTCA, probe- CTTCCGAGACAATCGCGAACATCTACACC) and performed in a multiplex reaction with primers for the reference gene RPP30 (fwr- GATTTGGACCTGCGA, rev- GCGGCTGTCTCCACA, probe- CTGACCTGAAGGCTCT). Reactions were performed in a Light Cyler 480 (Roche) with the light cycler Kit (Roche) according to the manufacturer's instructions. Clones with

the lowest expression of tTK, as determined by qPCR, were validated by PCR using a FastStart kit (Roche) with primers that bind to GFP or tTK or a site flanking the insert (R1- GGCTGGACAGAGAACAGGAC, F2- GCCCCCGAACCTTCAAAGA, R2- CTGCACGCCGTAGGTCAG, F3- GGAGATGGGGGAGGC TAACT). PCR products were separated on a 1 % agarose gel (Sigma) in TBE buffer (Life Technologies) for analysis and purified using the QIAquick PCR purification kit (Qiagen) prior to Sanger sequencing (Mycrosynth).

RNA isolation. RNA isolation from FACS-sorted or cultured cells was performed using a RNeasy micro or mini kit (both Qiagen) or automated Maxwell Total RNA purification kit (Promega). All procedures included DNAse I digestion.

RNA-sequencing and analysis. Total RNA from FACS-sorted or cell culture-treated samples was subjected to oligo (dT) capture and enrichment, and the resulting mRNA fraction was used to construct complementary DNA libraries. Transcriptome sequencing (RNA-seq.) was performed on an Illumina HiSeq platform using the standard protocol (TruSeq Stranded Total RNA Library, Illumina) that generates approximately 30 million reads of 50 base-pairs per sample. Experiments on FACS-sorted CLDN5-GFP+ and CLDN5-GFP- hPSC-ECs were performed using 6 replicates each from 2 different clones. The experiments with TGF-β treated samples included 3 replicates (each replicate experiment was performed as individual separated experiment). The RNA-seq. reads were then mapped to the human genome (NCBI build 37) using GSNAP (2). A comparison of 12 samples each of CLDN5-GFP+ and CLDN5-GFP- from two different clones was used. Inhibitor-treated hPSC-EC FASTQ samples were mapped to the human genome (hg19/Refseq) using STAR (3) and counting was performed using the union mode of HtSeq (4). Differential expression was performed using Deseq2 (3). Gene set enrichment analysis (GSEA) (5) was performed using the hallmark gene set at the MsigDb database (6) using weighted p2 analysis and following the default conditions with gene sets smaller than 15 and larger than 5000 genes were ignored. RNA-seq. data is available at Gene Expression Omnibus (GEO) under accession numbers GSE142321 and GSE142322.

Sample preparation and labeling for TMT-MS3 analysis. FACS-sorted CLDN5-GFP+ and CLDN5-GFP- hPSC-ECs (>1 million) were pelleted and subjected to TMT-MS3 analysis (7). In the follow-up

experiment with TGF-β inhibitors (RepSox and SB431542), cell pellets of about 2 million cells were processed in the same way. The samples were digested and prepared using the TMT labeling kit (PreOmics) with minor modifications. Briefly, frozen cell pellets were lysed in NHS-lysis buffer (100 µL) by pipetting up and down ten times. Next, the samples were boiled at 95ºC for 10 min and sonicated using a Bioruptor (10 min, 20 cycles of ON and OFF, Diagenode). Protein estimation was performed using BCA, and equal amounts (50 µg) of each sample were digested separately according to the PreOmics protocol at 37ºC for 2 h. The protein digests were then cleaned up using the PreOmics kit and labeled using the TMT10plex reagent (Thermo Fisher) using the following scheme: CLDN5-GFP+ hPSC-ECs replicates: 01, 02, 03, 04, 05 as 126, 127N, 128N, 129N, 130N; and CLDN5-GFP- hPSC-ECs replicates: 01, 02, 03, 04, 05 as 131, 127C, 128C, 129C, 130C at a ratio of 1:4 (protein to labeling reagent). In the follow up TMT11 plex reagent (Thermo Fisher) was used using the following labeling scheme: DMSO replicates 01, 02, 03 as 127C, 128N, 129C, RepSox replicates 01, 02, 03, 04 as 129N, 131C, 126, 127N and SB431542 treated replicates 01, 02, 03, 04 as 128C, 131, 130N, 130C respectively, Samples were then pooled into aliquots of 100 µg, loaded onto the PreOmics cartridges, cleaned according to the kit protocol, and lyophilized overnight. The labeling efficiency, as determined through data dependent acquisition (DDA) analysis and TMT10 as a dynamic modification on Proteome Discoverer 2.1, was determined to be 97.1 % for the FACS sorted cells and > 99 % for the follow-up experiment (TMT modified vs. unmodified PSMs).

Offline high pH basic reverse-phase fractionation for Mass Spectrometry Spectrometry and TMT-MS3 data acquisition and analysis. Pooled 100 µg labeled peptides were dissolved in 2 % Buffer B (Buffer B: 1 mM ammonium formate pH 10.0 in 100 % acetonitrile) and subjected to basic reverse phase fractionation using a YMC-Triart C18 Column (0.5 mm x 250 mm, S-3 µm particle size, 12 nm pore size) and an Agilent 1260 infinity series HPLC (Agilent Technologies) with the following gradient: 2-23 % Buffer B for 5 minutes, 23-33 % Buffer B for 25 minutes, 33-53 % Buffer B for 30 minutes, 53-100 % Buffer B for 5 minutes and 100 % Buffer B for 5 minutes. The column was equilibrated by 100-2 % buffer B for 1 minute followed by 2 % Buffer B for 14 minutes. (Buffer A: 10 mM ammonium formate pH 10.0, Buffer B 10 % Buffer A in acetonitrile). The 36 fractions were pooled into 12 fractions and lyophilized. 1 µg of each sample was then injected into the Orbitrap Fusion Lumos (Thermo fisher) instrument for TMT-MS3 analysis as in (8),

Endothelial cell expression signature. To further characterize the hPSC-derived ECs that we generated, we collected cell-identity signatures from the FANTOM 5 database (9) and used *BioQC* method (10) in either CLDN5-GFP+ or CLDN5-GFP- RNA expression data. *BioQC* performs the Wilcoxon-Mann-Whitney test to test whether genes in a gene-set have the same median expression as genes that are not in the gene-set (background). We report the results of *BioQC* with the *Enrichment Score*, which is the common logarithm transformed *P* values of one-sided (greater) Wilcoxon-Mann-Whitney test. An enrichment score of zero indicate that there is no difference between median expression of genes in the gene-set and that of background genes. A large positive enrichment score, in contrast, suggests that genes in the gene-set are positively enriched in the tested sample.

Competitive protein kinase binding assays. The scanMAX (DiscoveRx) kinase panel (including 468 kinases) was used to perform a kinome scan at 1 µM. The KINOMEscan procedure was carried out as described (11). K_d determination for ACVR1B (ALK4), TGFRB1 (ALK5), TGFBR2, ACVR2A, ACVR2B, BMPR1B, JNK2, p38α, and VEGFR-2 was done in duplicate with 11-point 3-fold serial dilutions starting from 40 µM (DiscoveRx). All of the compounds were dissolved in DMSO. Binding constants (K_d) were calculated with a standard dose-response curve using the Hill equation and the Hill Slope was set to -1. Curves were fitted using a non-linear least square fit using the Levenberg-Marquardt algorithm.

EC sprouting assays in neonatal mice. Newborn mice were injected (intraperitoneal) with 0, 0.3, 1, 3 and 10 mg/kg RepSox diluted in PBS/NaCl. The superficial layer of blood vessels in treated and control mice C57BL6 mice were investigated at P5 (postnatal day five). On day five, mice were sacrificed and the retinal cup was separated from the enucleated eye, clipped to unfold into a clover leaf-like structure obtaining four wings, stained for endothelial cells (Isolectin B4, biotin conjugated, Sigma #L2140) and with secondary antibody (DyLight 488 Streptavidin (VECTOR #SA-5488) and mounted onto a cover slip for image acquisition. Whole retinal flat mount images were acquired using an Olympus VS120 scanner (Olympus AG, Switzerland) at 20x at wavelength 488 nm. An automated analysis was performed with the Definiens Developer XD64 image analysis software (Definiens AG, Germany) applying a custom-tailored ruleset for segmentation of the vascular network allowing

quantitation of the superficial plexus and the radial expansion of the vascular front into the periphery with distance to the center and its tortuosity. All animals were kept at standard housing conditions (22±2 °C, 50-60 % humidity with a range of 40-80 % and 12/12 light/dark cycle). All the procedures were in strict accordance with the Swiss legislation on the use and care of laboratory animals.

Choroidal Neovascularization (CNV) *in vivo* model. 20 male Brown Norway (BN) rats from Charles River Germany were used for the study. The average weight of animals were 230-250 g when the experiments were executed. Six photocoagulation spots were induced in both dilated eyes with a Micron IV (600mW, 100ms) (Phoenix Research Labs, Pleasanton, CA), and rats were immediately treated with anesthetic eye drops (Fentanyl 0.005 mg/kg; Medetomidine 0.15 mg/kg; Midazolam 2 mg/kg). RepSox and vehicle control (ethanol n=6) was administered by gavage, and two doses were tested (1 mg/kg n=5; or 10 mg/kg n=5) starting two days before laser induction. The surface of the eyes were covered with Viscotears© to avoid corneal damage. The CNV-induced BN rat models were sacrificed seven days after photocoagulation. Effects were monitored in vivo using fluorescence angiography (described below).

Fundus fluorescence angiography. The fluorescence angiography (FA) was employed to image the photocoagulated lesions in the CNV-induced BN rats in both eyes. The lesions area was evaluated with Heidelberg Spectralis scanner HRA-OCT (Heidelberg Spectralis) right after the intravenous injection of Fluorescein (40 mg/kg, 20 mg/mL in NaCl 0.9 %) into lateral tail vein. All fundus images were manually analyzed to assess the level of vascularization within the lesion area by measuring hyper-fluorescence. The quantification was reported as either individual CNV area or mean per eye.

The barrier on-a-chip permeability assay (3D model). Endothelial cells were seeded in two-lane OrganoPlates (Mimetas BV) to form a tube and permeability of the endothelial barrier was assessed using fluorescent tracers. Briefly, 1.9 µL of Collagen I gel (final concentration 4 mg/mL, Collagen I from rat tail, Cultrex) was dispensed in the gel inlet and polymerized at 37 °C for 45 min. After polymerization, 2 µL of cell suspension (primary hRMECs at a final concentration of 60,000 cells/channel was added to the cell inlet using an automatic dispenser (MultiPette E3, Eppendorf). The plate was placed on the side to allow cells to attach to the gel for 60 min. Growth medium and

HBSS$^{+/+}$ were then added to the cell inlet and gel inlet respectively and the plate was incubated on the side for an additional 3-4 h at 37 °C to complete cell attachment. Medium was added to the cell outlet and plates were placed under bi-directional flow (+2° and −2° inclination every 8 min, Mimetas Rocker Mini). Endothelial tube formation was observed the day after seeding and medium was refreshed then. Cells were starved for 24 h prior treatment with 10 µM RepSox or 10 µM SB431542 for 24 h. To assess barrier permeability, starvation medium containing 0.1 mg/mL FITC-dextran (40 kDa, Sigma; 78331) and TRITC-dextran (4.4 kDa, Sigma; T1037) was perfused in the endothelium channel while starvation medium without fluorescent dye was added to the gel channel. The fluorescence intensity in the endothelium channel (IPERF) and in the receiver gel channel (IGEL) were monitored every 5 min for 20 min using Opera Phenix and Harmony software (Perkin Elmer) and the permeability coefficient (P_a) was calculated as follow:

$$P_a(cm.s^{-1}) = \frac{d}{dt}\left(\frac{I_{GEL}}{I_{PERF}}\right) \cdot V_{GEL} \cdot \frac{1}{A}$$

Where VGEL is the volume of the gel (4.54x10^{-4} cm^3) and A the surface area of the microvessel in contact with the gel (1.21x10^{-2} cm^2). P_a values were normalized to DMSO control.

Confocal imaging. Images were prepared using ZEISS LSM 710 confocal microscope (Carl Zeiss Microscopy GmbH) with 40x objective. The software, ZEN 2.1 version 14.0.7.201 was employed for imaging acquisition. The max intensity projected confocal Z-stack images were generated with ImageJ for further quantification.

Image-based quantification. Data analysis for fluorescence stained tight junction proteins in hPSC-ECs were performed using the open-source software Cell Profiler version 3.1.5 (12). The custom-analysis pipeline quantified the fluorescent intensity of each protein of interest (Claudin-5, ZO-1, and Occludin) co-localized with VE-Cadherin as the area of interest. For the transcytosis assessment, the number of Cholera Toxin Subunit B (CtxB) conglomerates in the cytosol were counted per cell using Cell Profiler. The fluorescent intensity of Caveolin-1 was extracted per cell and calculated in its average.

Fluorescence activated cell sorting (FACS) and analysis. The hPSC-ECs were dissociated from plates using Accutase (StemCell Technologies) and filtered before sorting with 30 µm filters (Miltenyi

Biotec). Dissociated cells were kept in full media during sorting and sorted in cooled collection tubes with complete EC media supplemented with 20 % FBS and 25 mM HEPES. Sorting was performed with BD FACSAria III (BD Biosciences) using 4-way purity precision mode. FACS plots were generated using the Flowjo_V10 software. For RNA-seq., a minimum of 100,000 cells were sorted. After sorting, cells were centrifuged (610 g, 10 min) and lysed in 650 µL RLT lysis buffer (Qiagen) + β-mercaptoethanol (Sigma-Aldrich, 1 %) and subsequently vortexed for 1 min at room temperature before being snap frozen. For mass spectrometry analysis, a minimum of 1,000,000 cells were sorted, washed with $PBS^{-/-}$ (Life Technologies), centrifuged (110 g, 3 min), $PBS^{-/-}$ was removed, and cell pellets were snap frozen.

Immunocytochemistry. The hPSC-ECs were plated on fibronectin (Corning)-coated cover glass (VWR) at 63,000 cells/cm^2 in 24 well culture plate. After the cells formed monolayers, the cells were treated with DMSO 0.1 % as a vehicle control, 10 µM SB431542, and 10 µM RepSox. Then after 48 hours incubation, the hPSC-ECs were fixed in 4 % PFA for 20 minutes at room temperature after PBS washing including Mg^{2+} /Ca^{2+}. Then, cover slips were washed two times with PBS again, then blocked in SuperBlock™ (Thermo) with 0.3 % TritonX for three hours at 4 °C on the shaker. Cells on the cover slip were stained with the primary antibody overnight at 4 °C in 1/10 diluted SuperBlock™ solution with 0.3 % TritonX-100. A day after, cells were washed three times for 10 minutes using PBS with 0.05 % TWEEN 20, then incubated them with secondary antibodies [Alexa Fluor™ 546 donkey anti-Goat (Invitrogen; A11056; 1:200); Alexa Fluor™ 647 donkey anti-Rabbit (Invitrogen; A31573; 1:200)] for three hours at 4 °C on the shaker. After three times of washing with PBS, Coverslips were mounted on glass-slide with Fluorescence Mounting Medium (DAKO) including DAPI. Primary antibodies were used: Human VE-Cadherin (R&D systems; AF938; 1:100), Claudin-5 Polyclonal Antibody (Invitrogen; 341600; 1:100), ZO-1 Polyclonal (Invitrogen; 402200; 1:100) and Occludin Polyclonal (Invitrogen; 711500; 1:100).

FITC-dextran permeability assay. Endothelial cells were seeded on fibronectin-coated Transwell®-96 well plates (Corning) in complete media. The bottom chamber contained 235 µL EC media and the top contained 75 µL. Cells were grown for 2 days to attach and generate confluent monolayers. Cells in the upper chamber were treated with compounds either with or without VEGFA (50 ng/mL) and

analyzed 2 days later by adding 10 µL of 4 and 40 kDa FITC-dextran (Sigma) to the top chamber and incubating for 30 min at 37 °C under 5 % CO_2. After 30 min, top wells were removed and the bottom plate, containing media and leaked FITC-dextran was measured with a fluorescent reader (Molecular Devices, excitation 485 nm, emission 535 nm).

Supplementary Figure legends:

Fig. S1. Construction of the CLDN5 transcriptional reporter in hPSCs. (A) A schematic of the Cas9 gene editing strategy for generating a CLDN5-P2A-GFP reporter is shown. A targeting vector was constructed with a promoter-less P2A-GFP sequence flanked by homology arms (LHA-left homology arm, RHA- right homology arm) and with transposase inverted terminal repeat sequences (ITR) flanking the vector resistance cassette. Targeting was performed in two steps. First, Cas9 and a sgRNA (containing sequence near the stop codon of CLDN5) catalyzed a double-stranded break in the genomic locus which was then repaired through homologous recombination between the genomic locus and the donor vector. Second, the vector resistance cassette was removed using an excision-only PiggyBac transposase. **(B)** A map of the donor vector (tTK= truncated thymidine kinase, E1Fα= elongation factor 1 alpha promoter). **(C)** An image of the PCR and gel electrophoresis used to assess clones after the first step of the editing and puromycin selection using primers as indicated (CPGP = cell pool-genome editing-puromycin selected). **(D)** PCR and gel electrophoresis again used to assess clones after excision of the resistance cassette using primers as indicated (CPE = cell pool-excised). **(E)** Clones were evaluated by qPCR using primers for tTK (as indicated in part c) to identify those that had lost tTK ("positive" clones). RPP30 was used as a reference gene. Those with delta Ct values higher than CPE (as indicated by the red dashed line) were selected for further consideration. **(F)** Selected clones were validated by PCR (using primers indicated in parts c and d) and gel electrophoresis. **(G)** Sanger sequencing of clones to confirm the insertion is correct. **(H)** Karyotyping of two clones with correct integration of the CLDN5-GFP reporter.

Fig. S2. Gene expression of EC identity and barrier related genes in CLDN5-GFP- and CLDN5-GFP+ sorted hPSC-ECs. The relative mRNA expression level from RNA-seq. data and protein (mass spectrometry) expression level of the CLDN5-GFP+hPSC-ECs in compared to CLDN5-GFP- population. The absolute mRNA expression level in RPKM of **(A)** (left) the endothelial cell marker, *CDH5* (VE-Cadherin) and (right) the epithelial cell marker, *CDH1* (E-Cadherin). **(B)** CLDN5-GFP- and CLDN5-GFP+ hPSC-EC samples were compared to gene expression of primary cells in the FANTOM 5 database. Individual dots denote average enrichment score of primary cell signatures derived from FANTOM 5 database among CLDN5-GFP+ and CLDN5-GFP- hPSC-ECs. **(C)** The heat map for visualized cell-identity signatures from FANTOM 5 database compared to CLDN5-GFP hPSC

populations. **(D)** Primary cell types that have the most similar gene expression to CLDN5-GFP- and CLDN5-GFP+ hPSC-EC samples. The RNA-seq. data for**,** the connexin gene family as components of gap junctions, and for **(E)** tight junction proteins. The mRNA expression and protein expression level of **(F)** the ATP binding cassette families. The relative mRNA expression level of the proteins related to the function of BRB/BBB. **(G)** Insulin receptor (*INSR)*, lipoprotein receptor protein (*LRP1*), and lipid transporter (*MFSD2A*). Columns are means ±SD. *=*P*<0.05, ***=*P*<0.001.

Fig. S3. Gene expression of hPSC-EC CLDN5-/+ in relation to signaling pathways involved in EC barrier integrity. Relative mRNA (RNA-seq.) and protein (mass spectrometry) expression for the following relevant pathway proteins was measured: **(A)** angiogenesis, **(B)** TGF-β signaling, **(C)** E2Fα-proliferation, and **(D)** Wnt signaling. Columns are means ±SD. ***=FDR<0.001, **=FDR<0.01, *=FDR<0.05

Fig. S4. Characterization of compounds in the SPARK library. A representation of the variability of activity of compounds created by relating standard deviation of *pAct* to average *pAct* in the CAT database. A representation of the selectivity of compounds created by relating Gini index to number of targets per gene in the CAT database. Pathway clusters generated using definitions from the Reactome database, the KEGG database, and the Gene Ontology (GO) Biological Process (BP).

Fig. S5. Dose-response treatment of selected compounds of hPSC-ECs. The CLDN5-GFP hPSC-ECs were treated with compounds, as indicated, ranging from 160 nM to 10 μM in 2-fold increments. After 2 days of treatment, the percentage of CLDN5-GFP+ hPSC-ECs was measured by flow cytometry.

Fig. S6. Molecular characterization of TGF-β inhibition by RepSox and SB431542. Compounds are shown based on capability to induce barrier, top to bottom. **(A)** Each of the compounds listed was subjected to a competitive binding assay against 468 kinases. For clarity, only those with > 65 % inhibition are shown [white = <65 % inhibition; red = >65 % inhibition, intensity of red denotes potency (from 65-100 %)]. **(B)** Dose-response kinase binding assays for a subset of kinases in the presence of each of the four compounds. K_ds higher than 40 μM were not plotted in the graph. GSEA plot of

pathway hallmarks after treatment with **(C)** RepSox or **(D)** SB431542. Pathways are sorted from left to right based on the normalized enrichment score (NES). Pathways depicted above the x-axis are upregulated while pathways below the x-axis were downregulated. Red dots and blue dots indicate pathways relevant to barrier endothelial cell biology, respectively. Pathways that were differentially regulated by the two compounds are labeled in green. **(E** and **F)** RNA-seq. was used to measure expression of **(E)** CLDN5 and PLVAP and **(F)** ESM1 at 8 hours and 48 hours following treatment with the TGF-β pathway inhibitors (SB431542 and RepSox) and VEGFA (50 ng/mL). ***=*P* or FDR<0.001, **=*P* or FDR<0.01, *=*P* or FDR<0.05

Fig. S7. The effects of RepSox treatment on the laser-induced choroidal neovascularization *in vivo* model. Quantification of average lesion size in laser-induced CNV model with/without compound treatment (DMSO control, 1 mg/kg and 10 mg/kg RepSox). **(A)** Fluorescence intensity of each lesion was measured. **(B)** The average area of all lesions per eye (6 lesions/eye) was determined.

Fig. S8. The Image intensity of VE-Cadherin in the different treatment condition. The intensity quantification of VE-Cadherin images in the different treatment conditions and without or with VEGFA. The arbitrary unit was normalized to DMSO treatment condition. Columns show mean ± SD.

Supplementary datasets legends:

Datasets S1. Genomic sequence of the donor vector for CLDN5-GFP reporter generation.

Datasets S2. RNA-seq. differential expression analysis of GFP+ vs GFP- hPSC-ECs.

Table S3. *BioQC* enrichment score of GFP- and GFP+ sorted ECs from the expression profiles in the FANTOM 5 database.

Datasets S4. Mass spectometry TMT differential expression of GFP+ vs. GFP- hPSC-ECs. Abundance Ratio: (GFPplus) / (GFPminus), Abundance Ratio Adj. P-Value: (GFPplus) / (GFPminus), Abundance Ratio Variability [%]: (GFPplus) / (GFPminus), Abundances (Grouped): GFPplus, Abundances (Grouped): GFPminus, Abundances (Grouped) coefficient of variation (CV [%]: GFPplus), Abundances (Grouped) CV [%]: GFPminus, Abundances of all samples (Scaled), Abundances of all samples (Normalized), confidence interval left (CI.L), confidence interval right (CI.R).

Datasets S5. Gene set enrichment analysis (GSEA) using the hallmark gene set of the MSigDB to compare GFP+ and GFP- hPSC-ECs. Enrichment score (ES), normalized enrichment score (NES), nominal p-value (NOM p-value). Pathways represented by genes (A) upregulated and (B) downregulated.

Datasets S6. Publicly available SPARK compounds. Only compounds that are publicly available are listed. A unique, integer index of SPARK compounds (library version 1.01; IndexSpark1.01). ChEMBL ID of the compound (ChEMBLID). Human gene symbol(s) of the target(s) (GeneSymbol). Mean pAct values in case of multiple measurements (pActMean). In case only one value is available, it is reported. Synonym according to the ChEMBL database, if available (ChEMBLPrefName).

Datasets S7. Gene targets of the SPARK library. Only compounds that are publicly available are listed. A unique, integer index of SPARK compounds (library version 1.01; IndexSpark1.01). Human entrez gene ID (EntrezGeneID). Human gene symbol(s) of the target(s) (GeneSymbol). Official gene

name (GeneName).

Datasets S8. Target clusters of the SPARK library. Target classes uniquely defined by a set of UniProt keywords (Traget Class Cluster and Target Class Name).UniProt cellular localization (UniProt Keyword). Human entrez gene ID (EntrezGeneID). Human gene symbol(s) of the target(s) (GeneSymbol). Official gene name (GeneName).

Datasets S9. Target pathways of the SPARK library. Genes were mapped to pathways from different databases (Category). Pathway description (Term). Enrichment score/Enrichment P-value/ Enrichment FDR of mapping genes to pathways terms. Human entrez gene ID (EntrezGeneID). Human gene symbol(s) of the target(s) (GeneSymbol). Official gene name (GeneName).

Datasets S10. Compounds from SPARK library (all at 10 µM) that induced a 2-fold or greater increase in the percent of GFP+ cells compared to DMSO after 2 days of treatment. Average/Standard deviation.

Datasets S11. RNA-seq. differential expression analysis after treatment of hPSC-EC with RepSox and SB431542 (both at 10 µM) with or without VEGFA. Human gene symbol (GeneSymbol). RPKM values of the expression.

Datasets S12. Gene set enrichment analysis (GSEA) using the hallmark gene set of the MSigDB after 2 days of treatment with RepSox and SB431542 (both at 10 µM). Abbreviations: enrichment score (ES), normalized enrichment score (NES). Pathways represented by genes (A) upregulated and (B) downregulated after RepSox treatment. Pathways represented by genes (C) upregulated or (D) downregulated by SB431542 treatment.

Datasets S13. Mass spectometry TMT differential expression of RepSox, SB431541 (both 10 µM) and DMSO treated hPSC-ECs. Abundance Ratio: (RepSox) / (DMSO), Abundance Ratio Adj. P-Value: (RepSox) / (DMSO), Abundance Ratio Variability [%]: (RepSox) / (DMSO), Abundance Ratio: (SB431542) / (DMSO), Abundance Ratio Adj. P-Value: (SB431542) / (DMSO), Abundance Ratio

Variability [%]: (SB431542) / (DMSO), Abundances (Grouped): RepSox, Abundances (Grouped): DMSO, Abundances (Grouped): SB431542, Abundances (Grouped) coefficient of variation (CV [%]: RepSox), Abundances (Grouped) CV [%]: DMSO, Abundances (Grouped) CV [%]: SB431542, Abundances of all samples (Scaled), Abundances of all samples (Normalized), confidence interval left (CI.L), confidence interval right (CI.R).

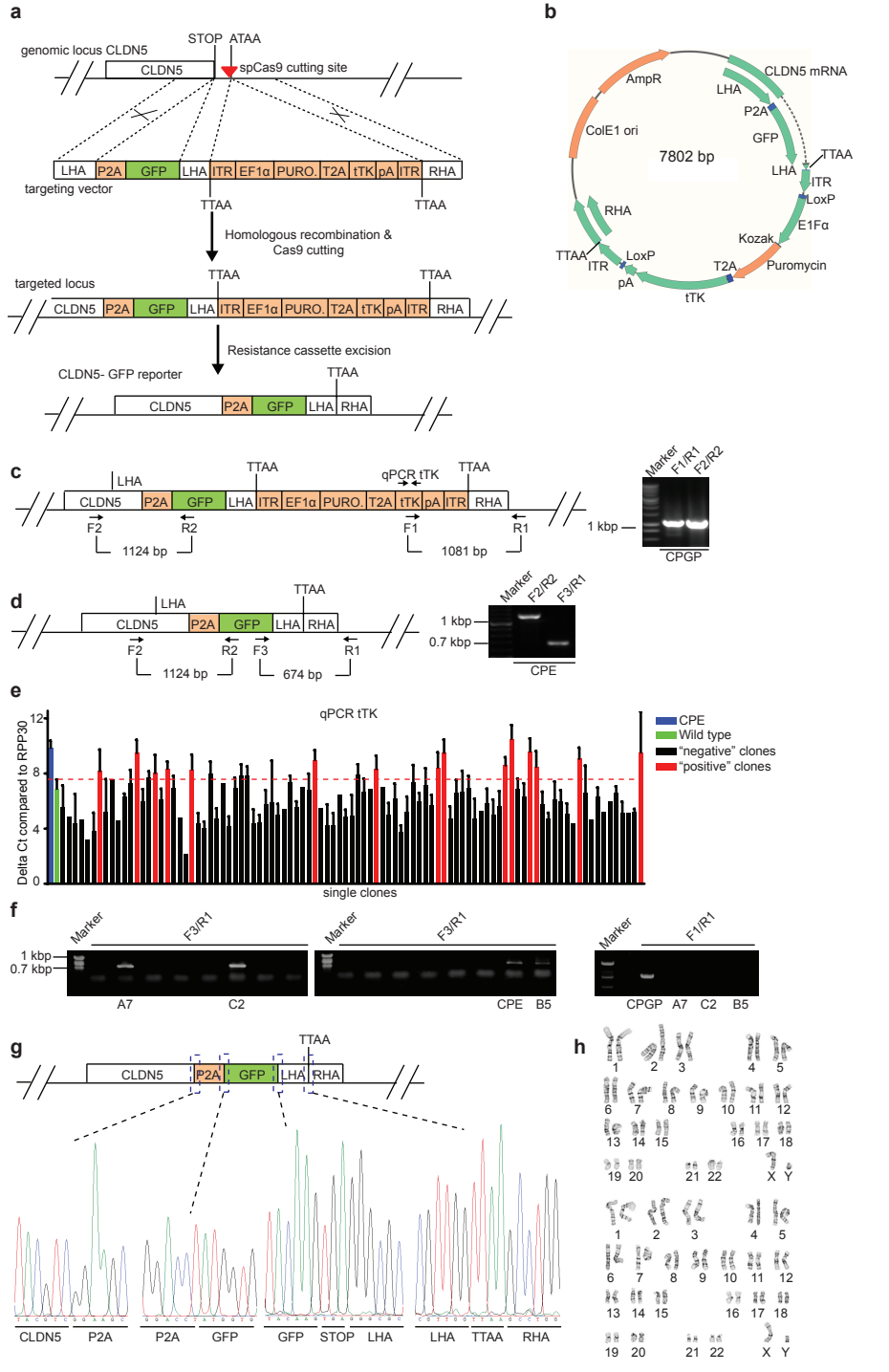
Datasets S14. Mass spectrometry TMT differential expression of RepSox, SB431541 (both 10 μM) and DMSO treated hPSC-ECs with statistical analysis of RepSox vs DMSO and SB431542 vs DMSO. Abundances of all samples (Normalized), confidence interval left (CI.L), confidence interval right (CI.R).

Dataset S15. The low average RPKM values of the expression. The listed genes, *MARVELD3*, *ABCB1*, *SLC6A8*, *NFATC2*, *TNFRSF19*, and *LRG1* that have low average RPKM values below 1.

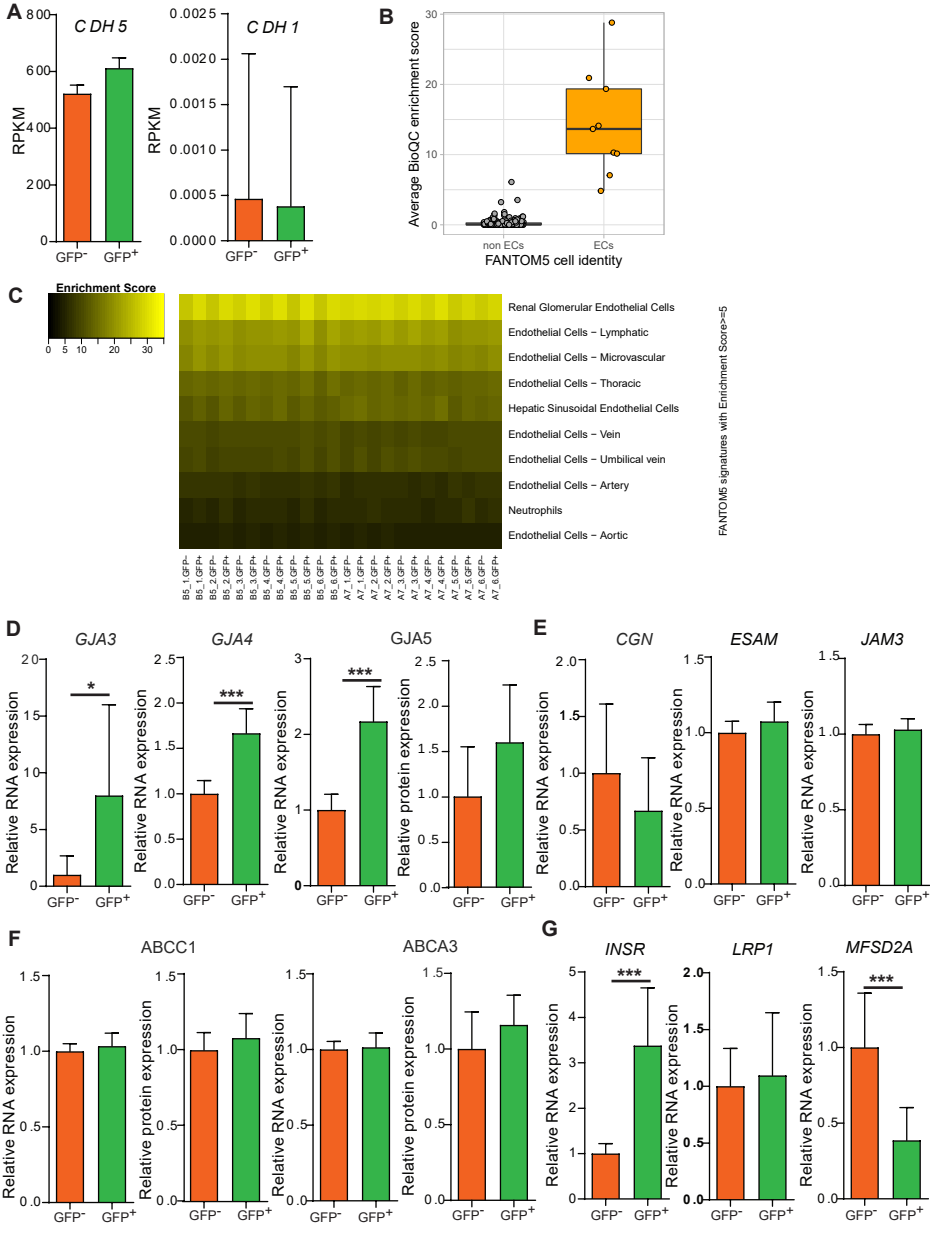
References for SI appendix:

1. Liang X, *et al.* (2015) Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J Biotechnol* 208:44-53.
2. Wu TD & Nacu S (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* 26(7):873-881.
3. Dobin A, *et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1):15-21.
4. Anders S, Pyl PT, & Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31(2):166-169.
5. Subramanian A, *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci* U S A 102(43):15545-15550.
6. Liberzon A, *et al.* (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 1(6):417-425.
7. McAlister GC, *et al.* (2014) MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal chem* 86(14):7150-7158.
8. Pandya NJ, Avila D, Dunkley T, Jagasia R, & Tzouros M (2019) TMT-MS3-Enabled Proteomic Quantification of Human IPSC-Derived Neurons. *Neuroproteomics*, ed Li KW (Springer New York, New York, NY), pp 103-117.
9. Forrest ARR, *et al.* (2014) A promoter-level mammalian expression atlas. *Nature* 507(7493):462-470.
10. Zhang JD, *et al.* (2017) Detect tissue heterogeneity in gene expression data with BioQC. *BMC Genomics* 18(1):277.
11. Fabian MA, *et al.* (2005) A small molecule–kinase interaction map for clinical kinase inhibitors. *Nat biotechnol* 23(3):329-336.
12. Carpenter AE, *et al.* (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* 7(10):R100.

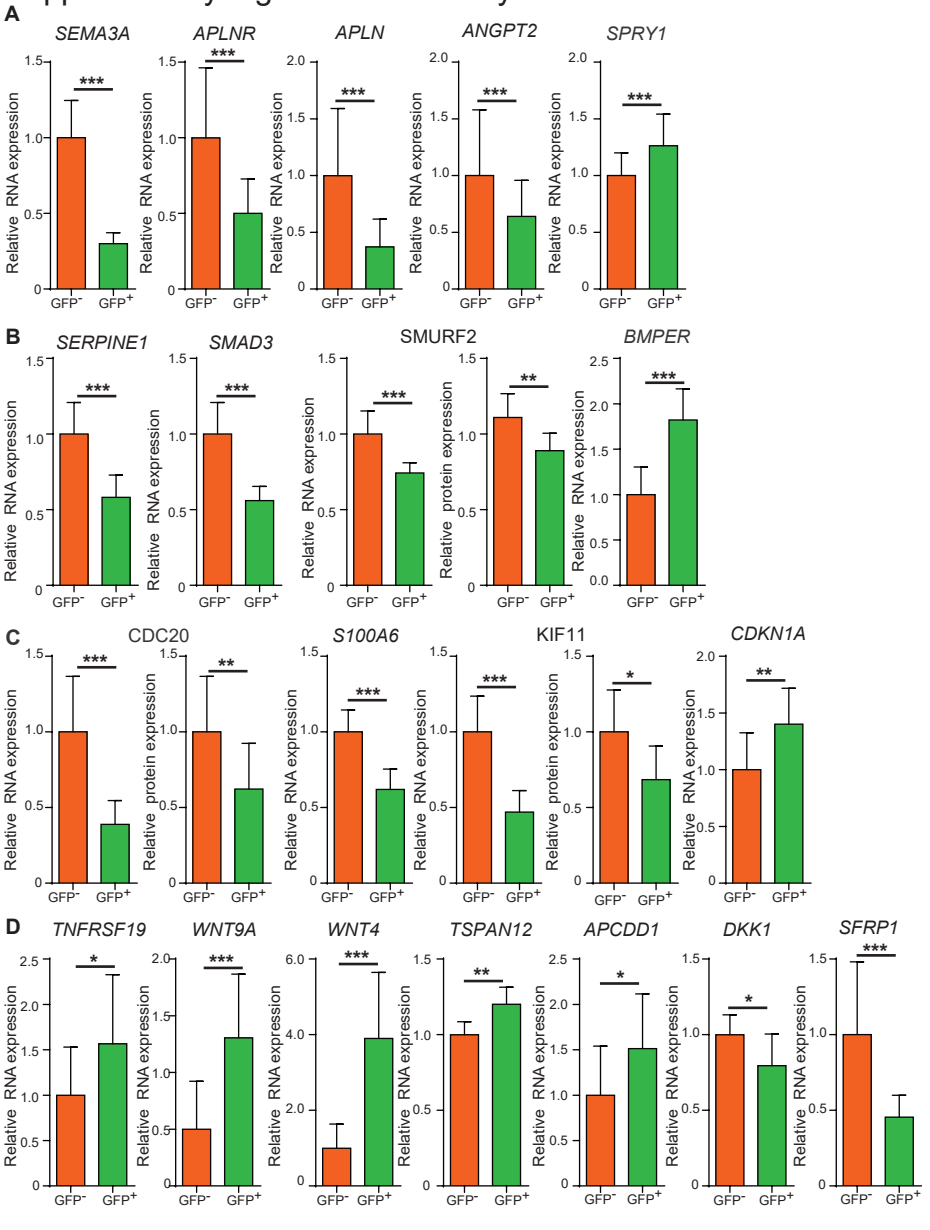
Supplementary Figure 1. Roudnicky et al.



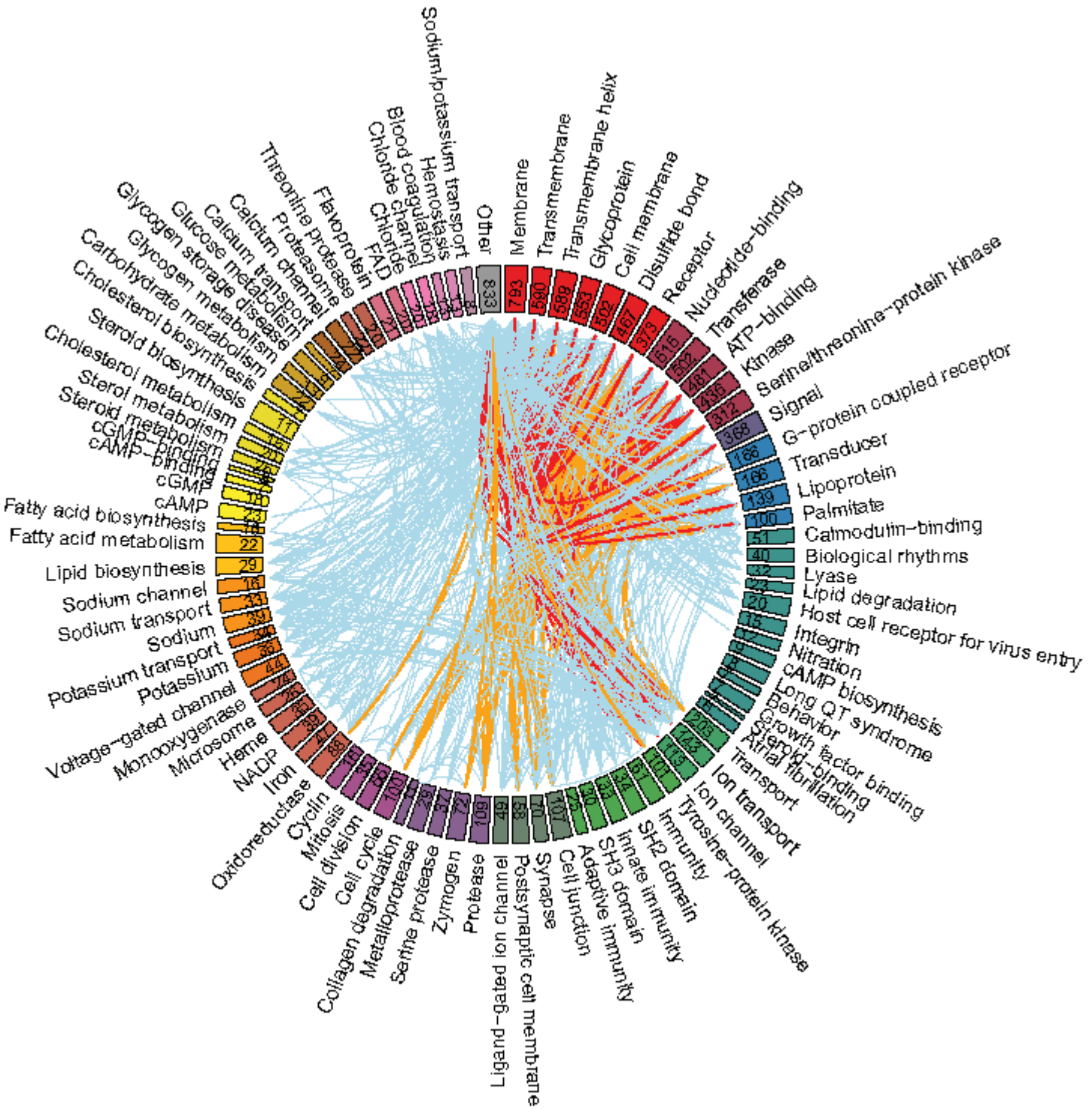
Supplementary Figure 2. Roudnicky et al.



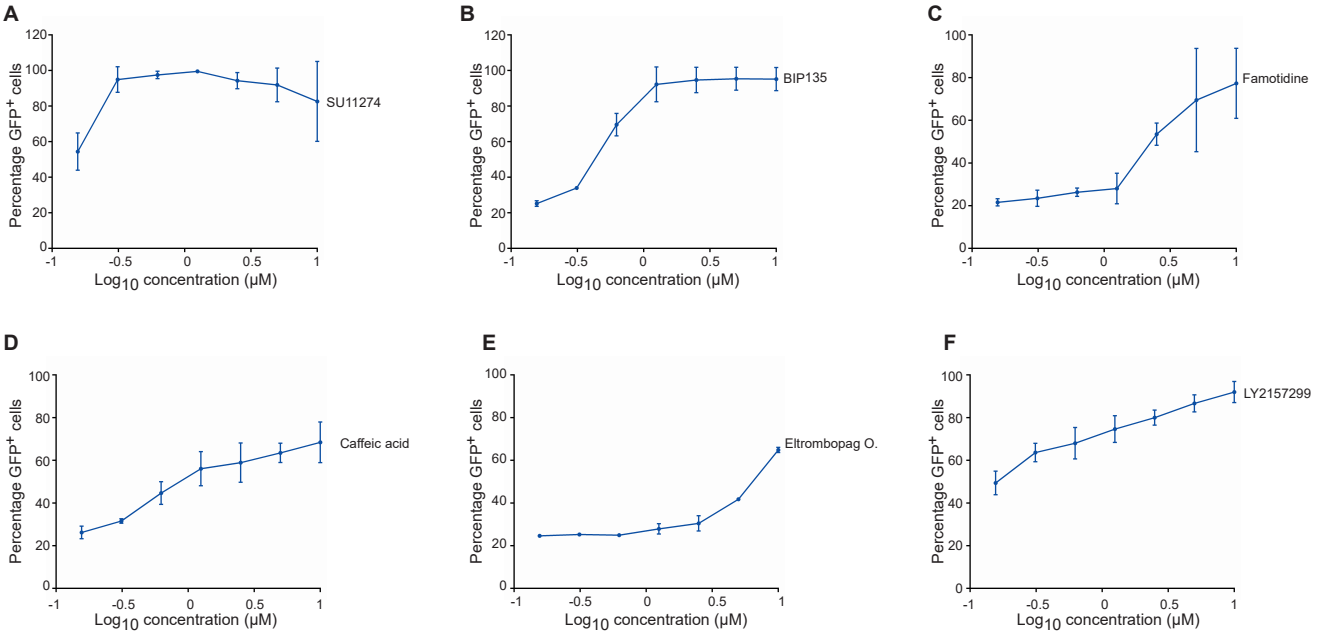
Supplementary Figure 3. Roudnicky et al.



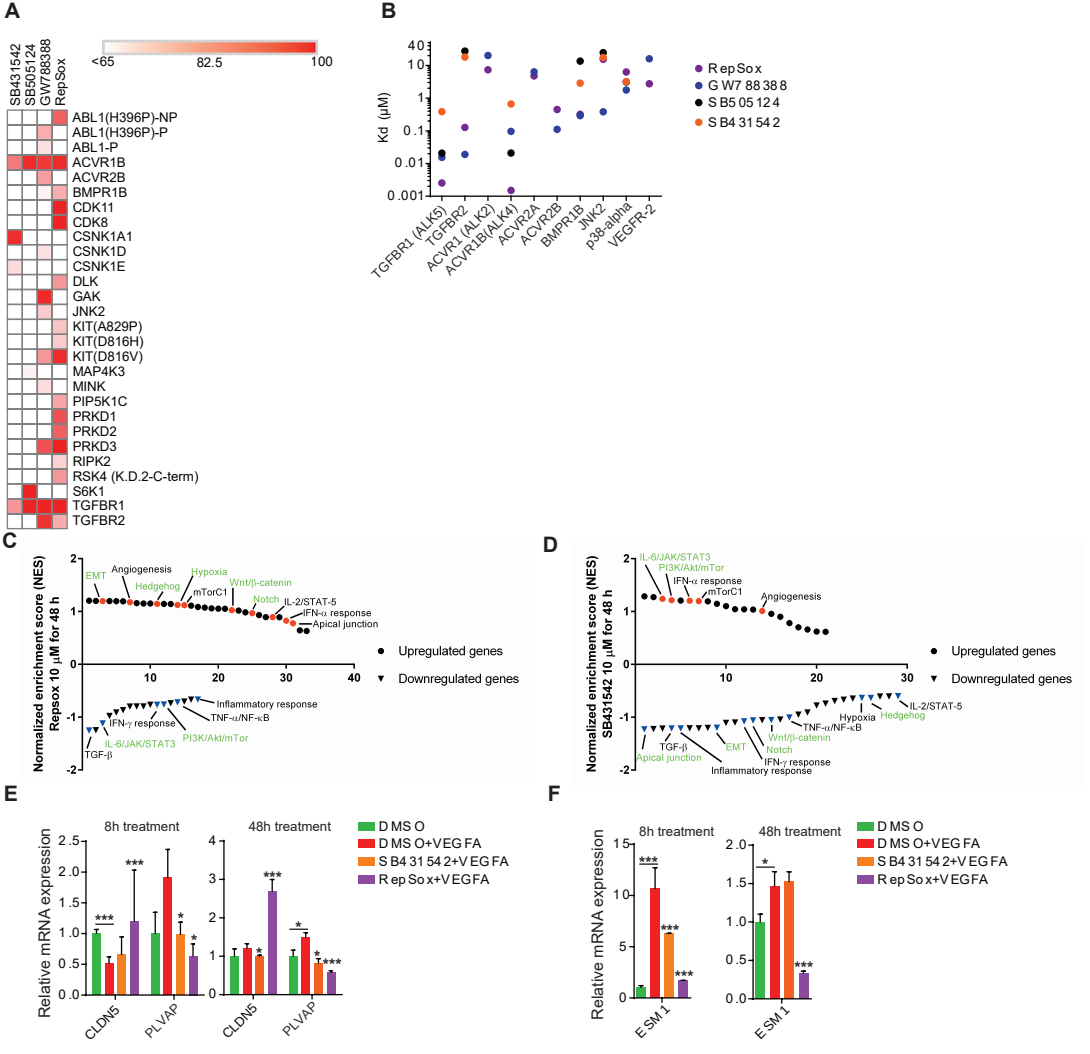
Supplementary Figure 4. Roudnicky et al.



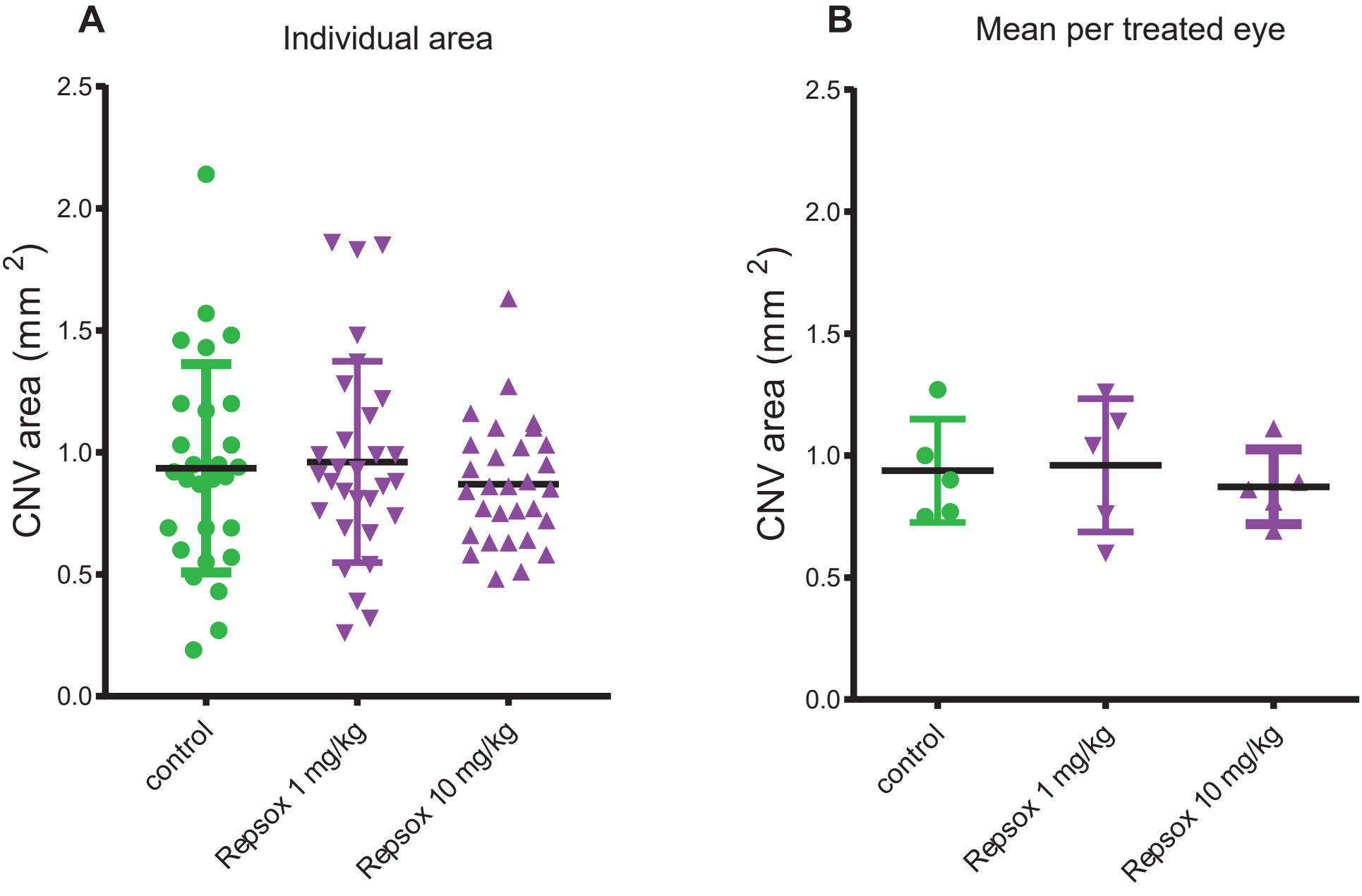
Supplementary Figure 5. Roudnicky et al.



Supplementary Figure 6. Roudnicky et al.



Supplementary Figure 7. Roudnicky et al.



Supplementary Figure 8. Roudnicky et al.