Supplemental Method

Srsf2$^{P95H/+}$ CRISPR knockout pooled library screen

At the start of the screen, 130x10^6 Cas9 expressing cells from each cell line (n=3 independently derived cell lines per genotype) were infected with the Brie CRISPR knockout pooled library (Virus aliquots provided by the Victorian Centre for Functional Genomics; Addgene #73633)[39]. The number of cells infected was calculated by multiplying the number of guides in the library (78,637) by the desired number of cells for each guide at the start of the experiment (aiming for 500 copies per sgRNA), divided by the multiplicity of infection required (MOI=0.3). For each cell line, the 130x10^6 cells were separated into four T75 flasks (Nunc™) and, in each spin-infection flask, 35x10^6 cells were mixed with 122µL of concentrated lentivirus containing the Brie library and 35µL of polybrene (8µg/mL, Sigma-Aldrich) in 35mL of IMDM-Cas9 media. The cells/virus mix was then spun at 1,100g for 90 minutes and the cells were re-incubated immediately after the spin-infection. Approximately 4 hours later, the mixture was transferred into two T-175 flasks (Falcon) with the addition of 65mL of IMDM-Cas9 media and cells were re-incubated. After 48 hours, cells were counted and diluted to 8x10^5 cells/mL before adding puromycin (0.5µg/mL; Merck). After 4 days of puromycin selection, the puromycin and blasticidin were washed out and the cells were cultured in IMDM with 10%FBS, 1% GM-CSF and 400nM 4-hydroxy tamoxifen (Merck Millipore). This timepoint is labelled day 0 of the screen. The tamoxifen was removed from the cells after 4 days and cells were then passaged at 2x10^5/mL every 2 days or 1x10^5/mL for a 3-day interval for a total of 18 days. A minimum of 40x10^6 cells was maintained during each passage. On day 0, day 4, day 11 and day 18 of the screen, pellets of 40x10^6 cells were collected for gDNA isolation and library preparation.

Data availability

All datasets related to this work are deposited in GEO

(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE165506).

Supplemental Table

Supp Table 1 Datasets used in comparison of *SRSF2/Srsf2* mis-spliced genes.

Supp Table 2 Mapped read counts and coverage of 24 libraries in *Srsf2* CRISPR screen.

Supplemental Figures

Supp Figure 1 Generation and characterization of Hoxb8 *R26*Cre-ERT2$^{ki/+}$ GM-CSF and Cas9-blasticidin resistant Hoxb8 *R26*Cre-ERT2$^{ki/+}$ GM-CSF cell lines. (A) A schematic diagram illustrating the generation of Hoxb8 *R26*Cre-ERT2$^{ki/+}$ GM-CSF and Cas9-blasticidin resistant Hoxb8 *R26*Cre-ERT2$^{ki/+}$ GM-CSF cell lines. (B) FACS plot of Hoxb8 GM-CSF cells stained with antibodies against CD11b and Gr-1. (C) Cas9 Hoxb8 GM-CSF *Srsf2*$^{+/+}$ and *Srsf2*$^{P95H/+}$ cell lines have similar proliferation rates (n=3 per genotype). (D) Cas9 Hoxb8 GM-CSF *Srsf2*$^{P95H/+}$ cells achieve full recombination after 4 days of 400nM tamoxifen treatment and remain stably recombined after tamoxifen withdrawal. Cas9 *Srsf2*$^{P95H/+}$ cell lines are: #243, #81 and #83. Cas9 *Srsf2*$^{+/+}$ cell lines are #281, #144 and #145.

Supp Figure 2 Heatmap of top differentially expressed genes with absolute log fold change (absLogFC) >2 and FDR < 0.05 between Cas9 *Srsf2*$^{+/+}$ and *Srsf2*$^{P95H/+}$ cell lines (n=3).

Supp Figure 3 The sample count distribution and gini index of *Srsf2* CRISPR screen. (**A**) Sample count distribution of sgRNA of all 24 libraries showing unbiased representation of sgRNA. (**B**) The Gini index of sgRNA in 24 libraries indicating fewer clones of the sgRNA pool remain as the screen progresses.

Supp Figure 4 The correlation between all six individual cell lines across all timepoints of the screen. The color code and number in each box indicate the concordance between the two samples. The color scale is shown at the top right. A red color and number close to 1 means the two samples are near identical. In contrast, a blue color and number close to 0 means the two samples are not correlated.

Supp Figure 5 Depletion of the CEG2 core essential genes at day 4 of the CRISPR screen. The top 10 depleted CEG2 essential genes are labelled. *Hoxb8* and *Csf2a* are highlighted in green and marked in blue dots. Grey dots: all the sgRNA identified; yellow dots: sgRNA targeting CEG2 essential genes.

Supp Figure 6 CDK6 is a druggable target of Srsf2^{P95H} cell lines. (**A**) The top 30 genes that are exclusively negatively selected in the *Srsf2*$^{P95H/+}$ cell lines at day 18. Genes are ranked by MaGeCK score from the lowest (the most negatively enriched) to highest. FDR<0.05 for *Srsf2*$^{+/+}$ cell lines, FDR<0.01 for *Srsf2*$^{P95H/+}$ cell lines. (**B**) The guide count for CDK2 sgRNA, and CDK4 sgRNA (**C**) in *Srsf2*$^{+/+}$ and *Srsf2*$^{P95H/+}$ cells show no depletion over the course of the screen. (**D**) *Srsf2*$^{P95H/+}$ cell lines are not sensitive to inhibition of mitochondrial OXPHOS targeting agents, Tigecycline and Mubritinib (n=2 independent lines per genotype).

Supplemental files:

Dataset S1. Gene level (a) and transcript level (b) expression of the GMCSF cells

Dataset S2. Splicing GMCSF cells (a) rMATS (b) PSI-sigma

Dataset S3. CRISPR results (a) Counts (b)

SRSF2_Mageck_MLE_Screen_Results_reps_combined (c)

Combined_gene_results_annotated_MaGeCK

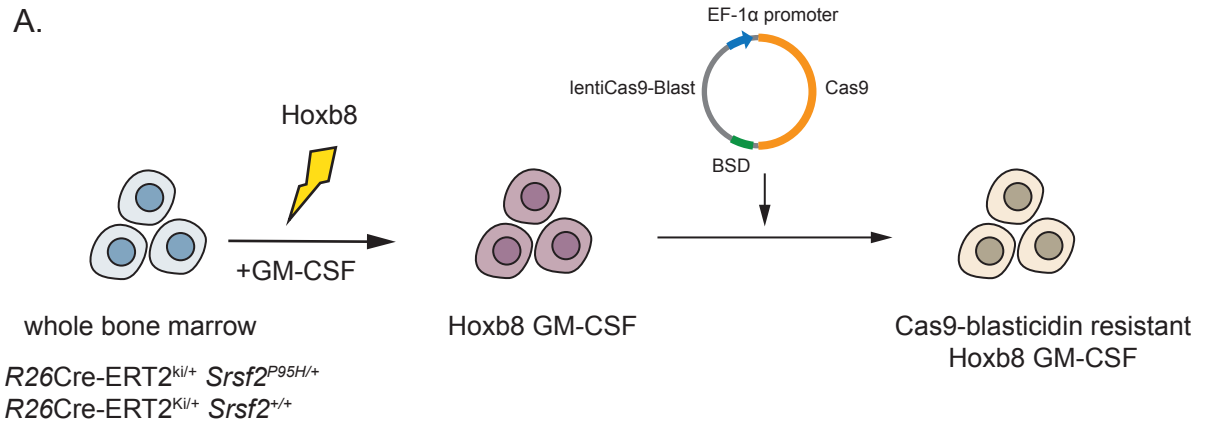
Supp Table 1 Xu et al., 2021

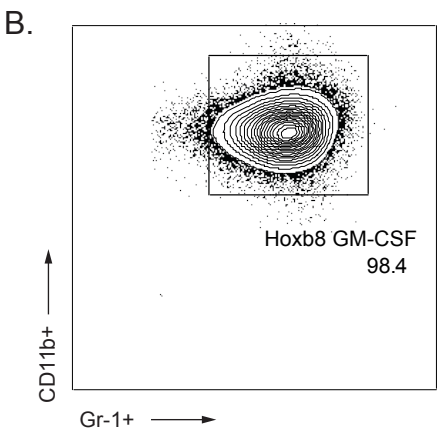
Dataset description	Dataset URL	Database, license, and accessibility information	Reference
Human datasets			
mRNA splicing data from primary CMML (n=13; 3 with *SRSF2* mutation) and AML (n=9; 5 with *SRSF2* mutation) patients*	Dataset retrieved from table S1 in the supplement	Available in the supplement information table S1	(Kim et al., 2015)
mRNA profile of MDS-L cells (in triplicate) transfected with *Srsf2* P95H shRNA construct	https://www.ncbi.nlm.nih.gov /geo/query/acc.cgi?acc=GSE61052	Publicly available at Gene Expression Omnibus (GEO): GSE61052. Mis-spliced gene list available in the supplement information table S4	(Komeno et al., 2015)
mRNA profile of K562 CRISPR cell clones (with wild-type or mutant SRSF2, n=4 per genotype)	https://www.ncbi.nlm.nih.gov /geo/query/acc.cgi?acc=GSE71299	Publicly available at Gene Expression Omnibus (GEO): GSE71299 Mis-spliced gene list available in the supplement information table S1	(Zhang et al., 2015)
mRNA profile of primary MDS bone marrow samples (n=115)	Dataset retrieved from table S7 in the supplement	Available in the supplement information tables S7	(Qiu et al., 2016)
mRNA profile of primary CMML (n=20; 7 with *SRSF2* mutation) patients	Dataset retrieved from table SX-XIV in the supplement	Available in the supplement information tables SX-XIV	(Hurtado et al., 2018)
Murine datasets			
mRNA profiles of murine model (MP/LSK cells) expressing *Srsf2* WT and mutants*	https://www.ncbi.nlm.nih.gov /geo/query/acc.cgi?acc=GSE65349	Publicly available at Gene Expression Omnibus (GEO): GSE65349 Mis-spliced gene list available in the supplement information table S1	(Kim et al., 2015)
mRNA profile of murine model (MP/LSK cells) expressing *Srsf2* WT or *Srsf2*$^{P95H/+}$ *	http://trace.ddbj.nig.ac.jp /DRASearch/submission?acc= DRA006224	Publicly available at DNA Data Bank of Japan repository: DRA006224 Mis-splicing gene list available in supplemental table 2	(Kon et al., 2018)
mRNA profile of murine model (sorted LK+ eYFP+ cells) expressing *Srsf2* WT or *Srsf2*$^{P95H/+}$	https://www.ncbi.nlm.nih.gov /geo/query/acc.cgi?acc=GSE99852	Publicly available at Gene Expression Omnibus (GEO): GSE99852 Mis-splicing gene list available in supplemental dataset 2	(Smeets et al., 2018)
mRNA profile of Hoxb8 immortalised GM-CSF cell lines expressing *Srsf2* WT or *Srsf2*$^{P95H/+}$	N/A	Available in this manuscript	Published with this manuscript

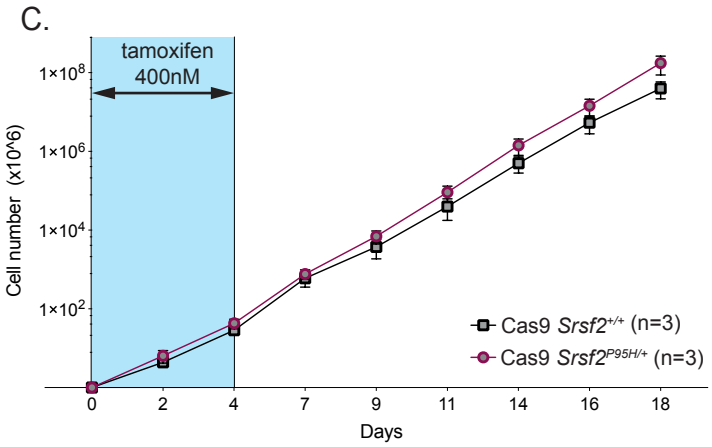
*These datasets contain mis-splicing data from two different cell/disase types

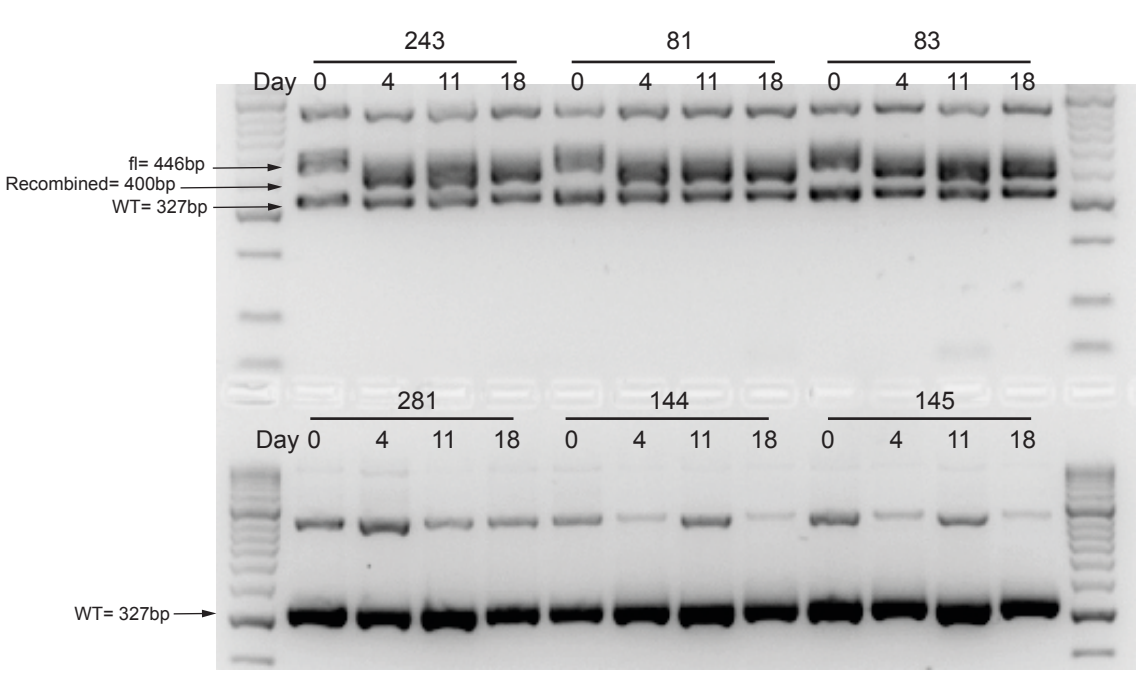
Sample_name	Mappability (%)	Total_reads	Mapped_reads	Average Coverage (Copy number /sgRNA)
WT_D0_S144	85.72347	38583293	33074936	420.6
WT_D4_S144	85.87031	34396145	29536075	375.6
WT_D11_S144	84.44593	32585779	27517364	349.9
WT_D18_S144	87.12968	36646854	31930287	406.0
WT_D0_S145	86.08064	30683490	26412546	335.9
WT_D4_S145	86.90123	34345631	29846775	379.6
WT_D11_S145	84.85053	32027437	27175449	345.6
WT_D18_S145	86.29772	30632262	26434944	336.2
WT_D0_S281	85.63385	33078080	28326034	360.2
WT_D4_S281	86.1783	32858200	28316637	360.1
WT_D11_S281	85.12764	30345497	25832405	328.5
WT_D18_S281	86.19909	37822483	32602635	414.6
P95H_D0_S81	84.96886	26147819	22217505	282.5
P95H_D4_S81	84.81972	34788027	29507107	375.2
P95H_D11_S81	84.13214	32715519	27524267	350.0
P95H_D18_S81	86.06658	32080958	27610984	351.1
P95H_D0_S83	85.048	32903333	27983628	355.9
P95H_D4_S83	85.34779	28279733	24136126	306.9
P95H_D11_S83	84.42695	30106447	25417955	323.2
P95H_D18_S83	85.38084	28122359	24011106	305.3
P95H_D0_S243	85.01022	36751029	31242132	397.3
P95H_D4_S243	85.58722	36307614	31074678	395.2
P95H_D11_S243	85.15725	32322150	27524653	350.0
P95H_D18_S243	86.3542	28848293	24911713	316.8

Supp figure 1 Xu et al., 2021

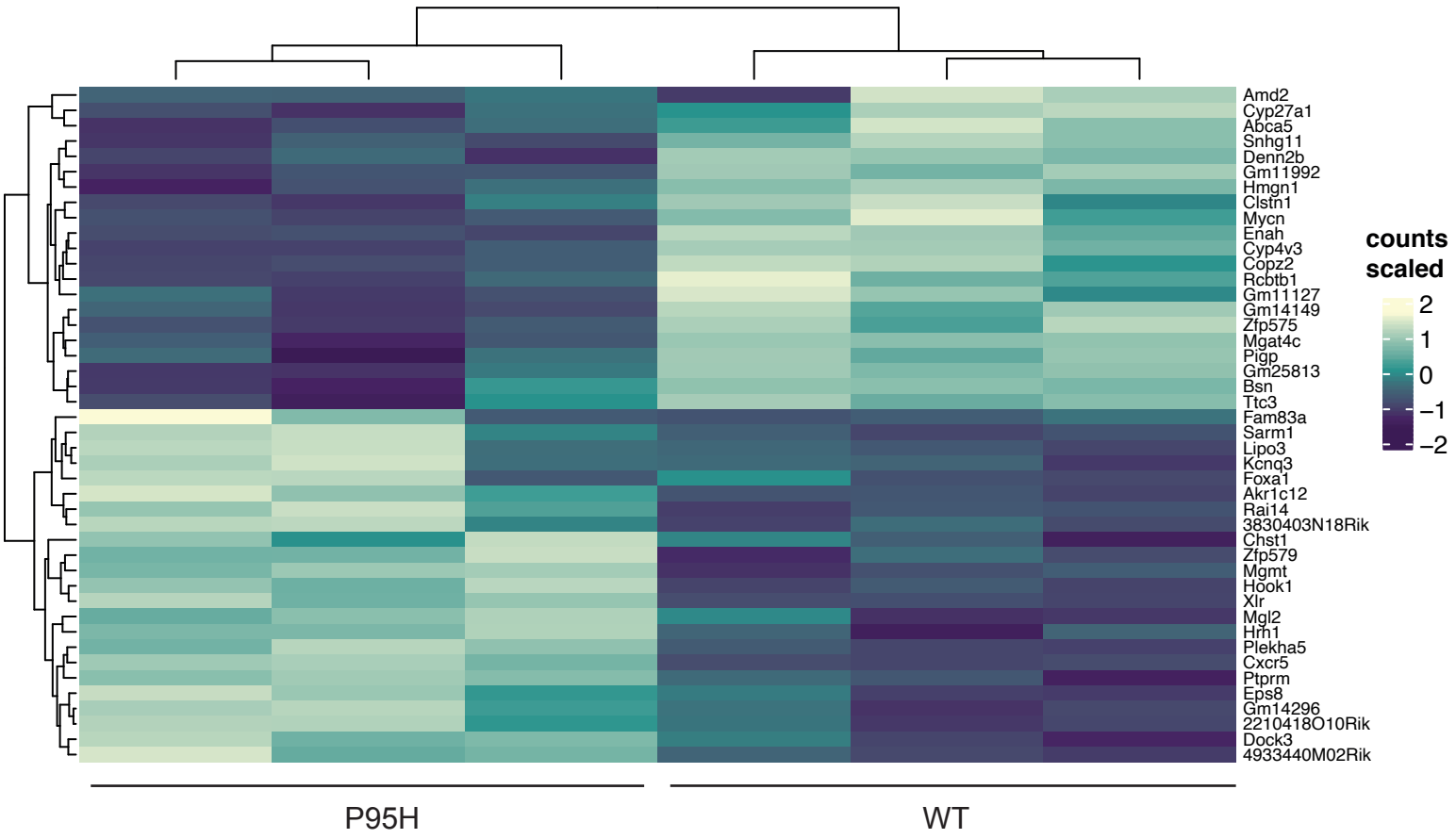






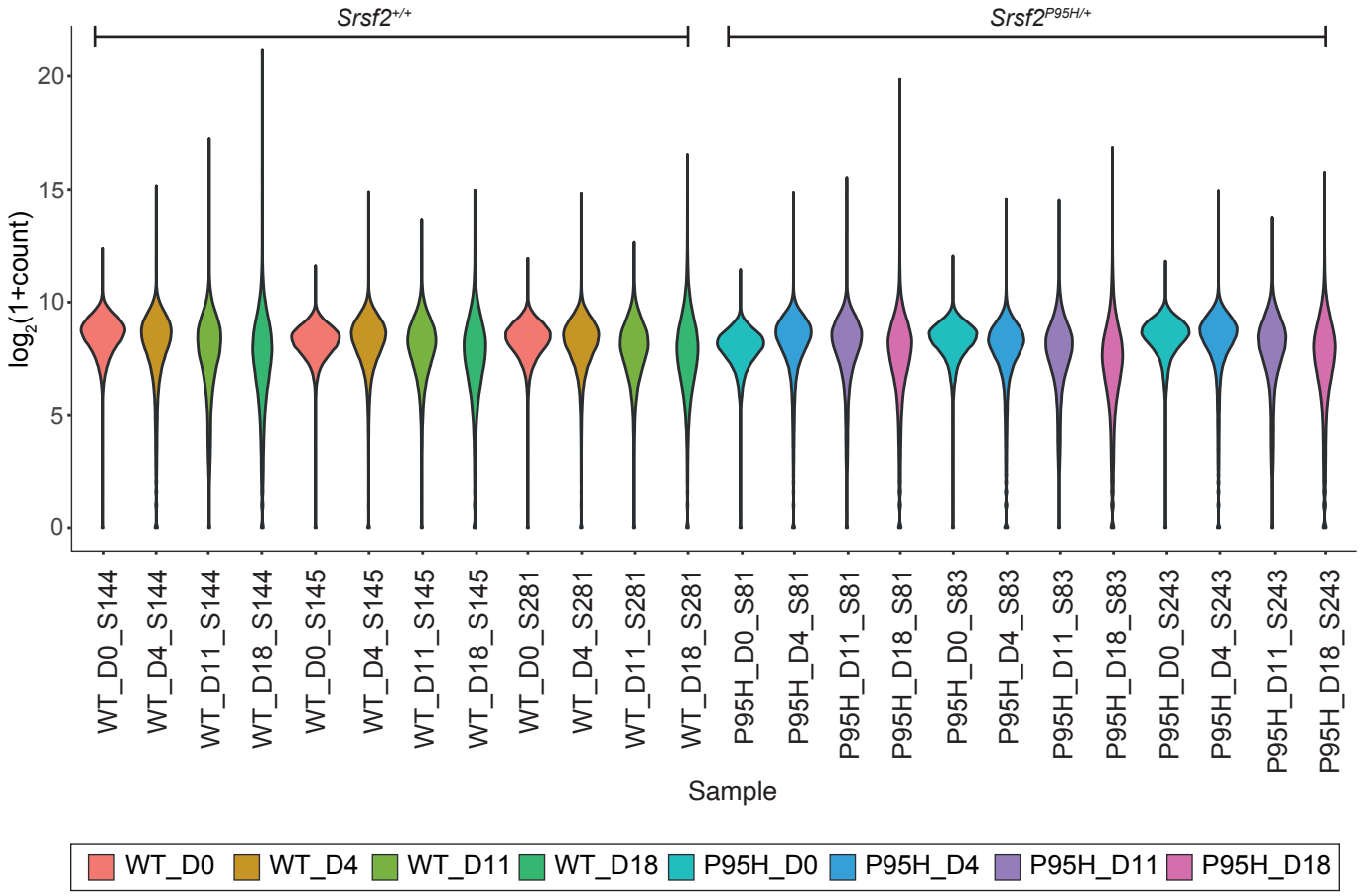


Supp figure 2 Xu et al., 2021

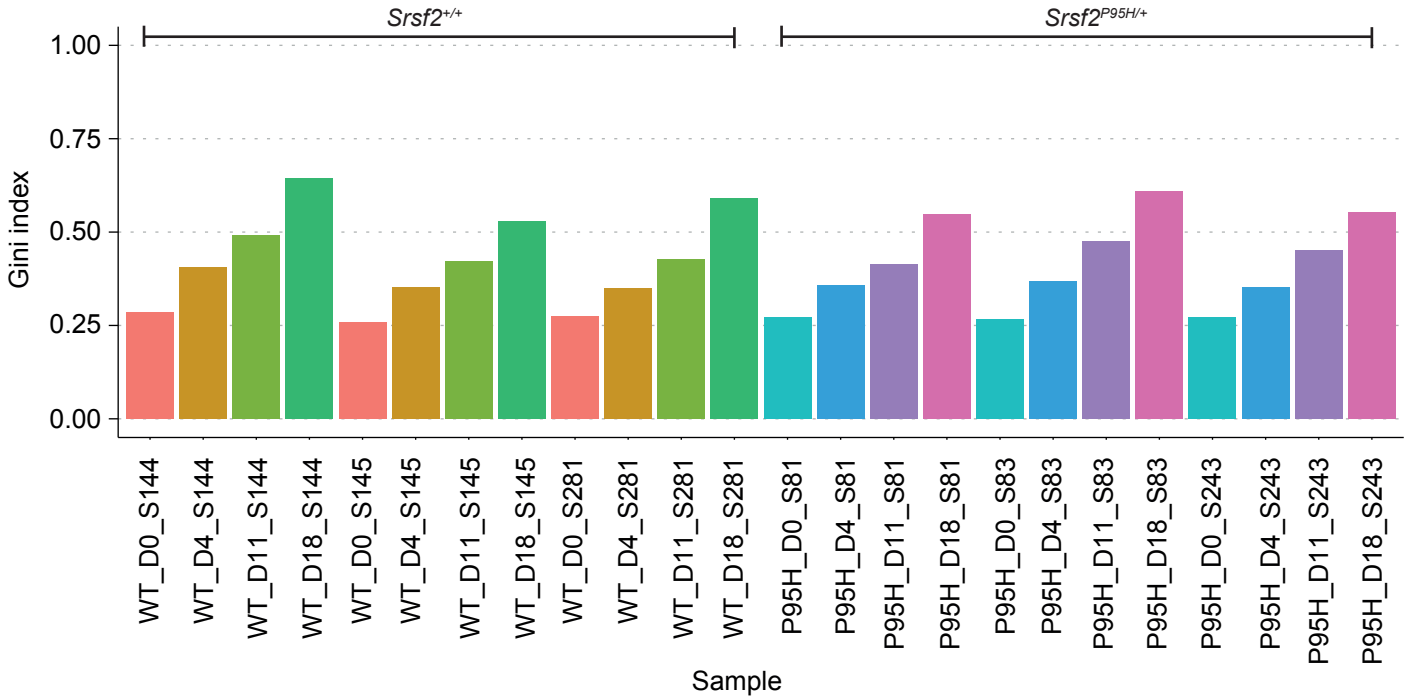


Supp figure 3 Xu et al., 2021

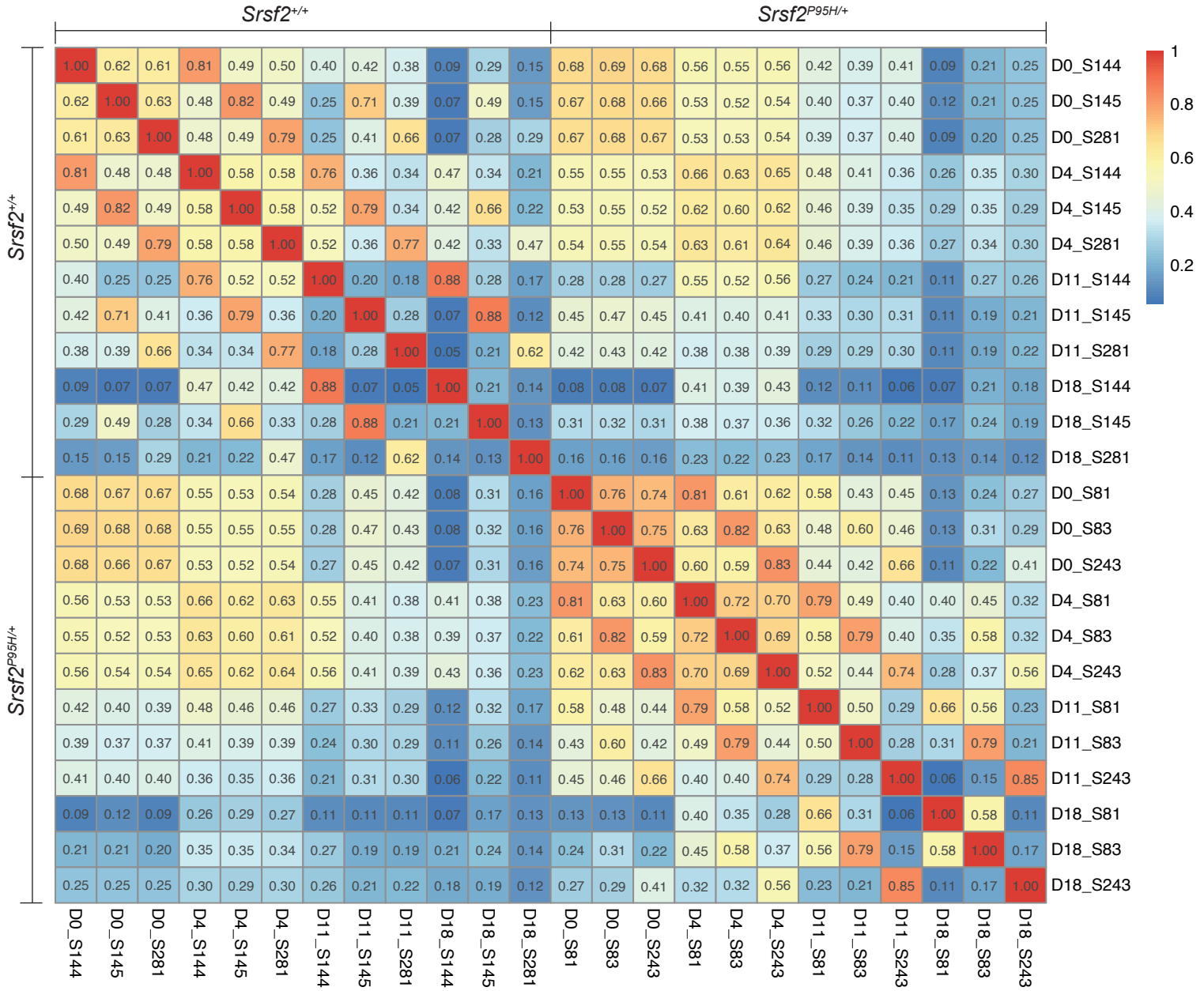
A. Sample count distribution



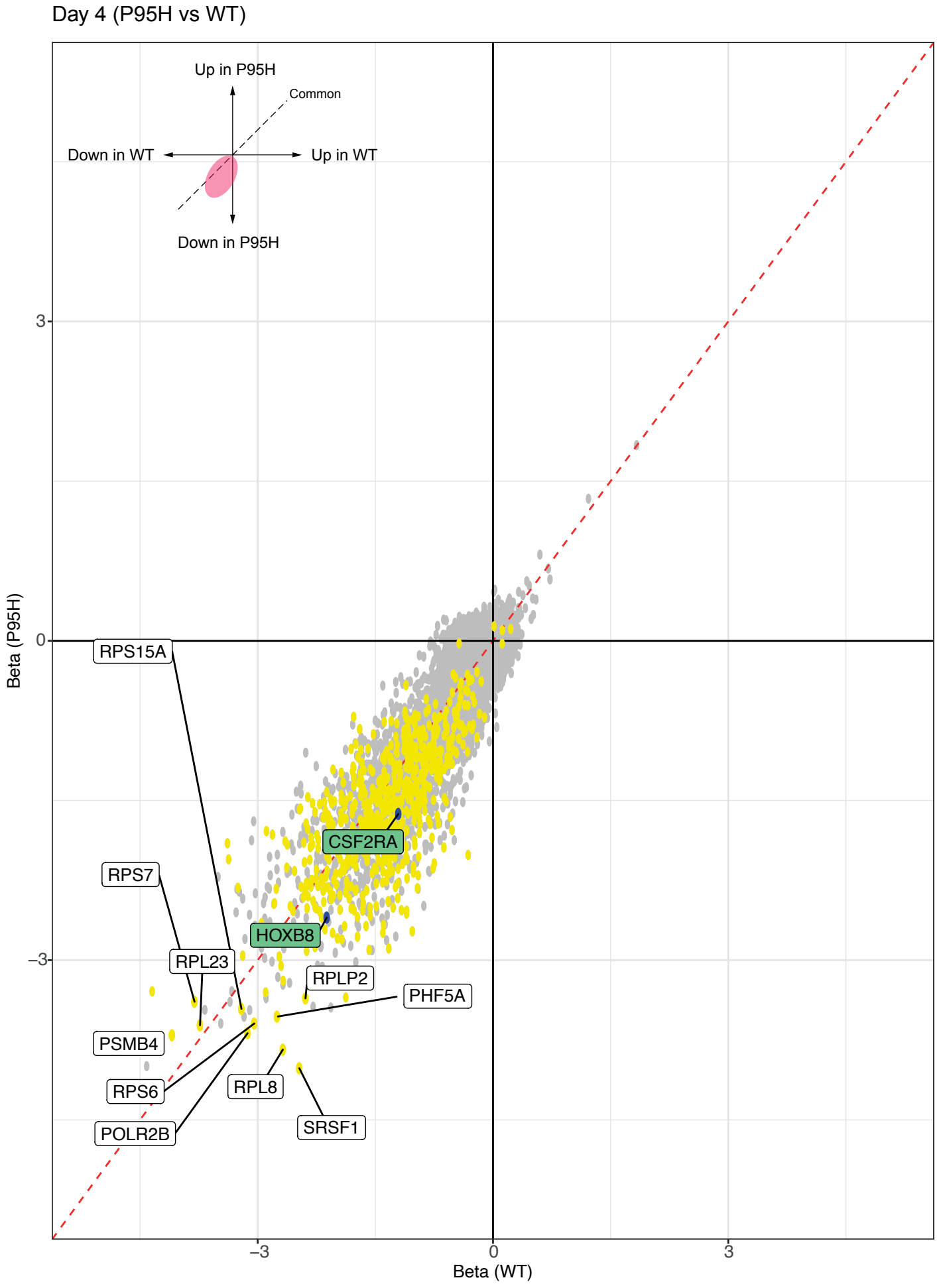
B. Gini index



Supp figure 4 Xu et al., 2021



Supp figure 5 Xu et al., 2021



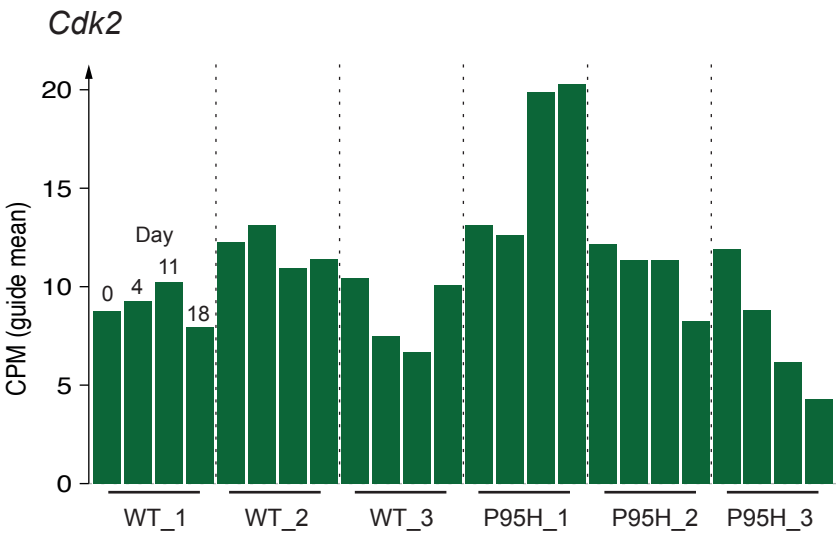
Supp figure 6 Xu et al., 2021

A.

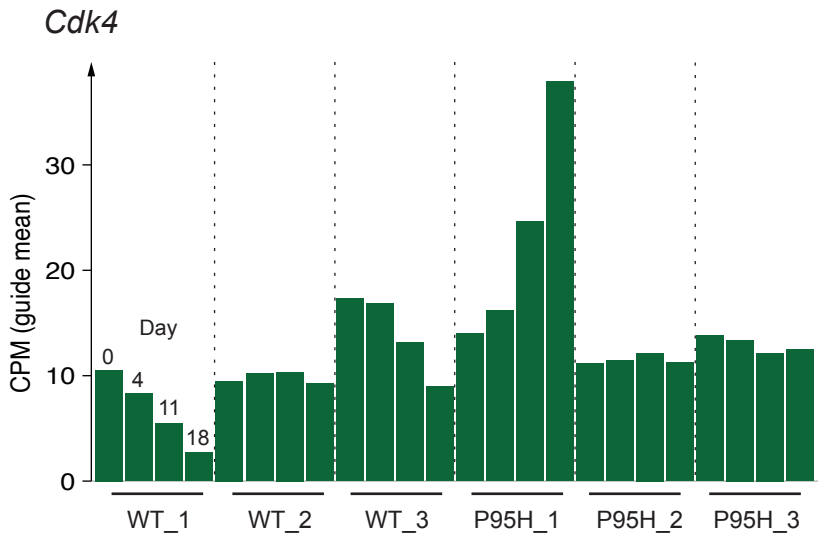
D18	MaGeCK Score
Serf2	4.60x10^-08
Mogs	5.90x10^-08
Qrsl1	1.46x10^-07
Gcn1l1	1.65x10^-07
Dap3	1.91x10^-07
Mtmr9	1.96x10^-07
Nsun4	2.56x10^-07
Vars2	3.12x10^-07
Vps51	4.72x10^-07
Pstk	5.27x10^-07
Mrps17	6.48x10^-07
Dars2	7.01x10^-07
Pkmyt1	7.14x10^-07
Mre11a	7.41x10^-07
Mrps27	1.14x10^-06
Asf1a	1.18x10^-06
Rpl21	1.39x10^-06
Rad50	1.39x10^-06
Lias	1.94x10^-06
Uqcrc1	1.94x10^-06
Cox6b1	1.97x10^-06
Atp5a1	2.20x10^-06
Wdr83	2.27x10^-06
Uqcc2	2.80x10^-06
Ptar1	2.97x10^-06
Ankrd49	3.00x10^-06
Dynll1	3.16x10^-06
Cdk6	3.44x10^-06
Ei24	3.51x10^-06
Elp5	3.63x10^06

MaGeCK score

B.



C.



D.