D1.4

Model development data including genetic perturbation screens and gene-drug synergies

Project number	826121
Project acronym	iPC
Project title	individualizedPaediatricCure: Cloud-based virtual-patient models for precision paediatric oncology
Start date of the project	1st January, 2019
Duration	53 months
Programme	H2020-SC1-DTH-2018-1

Deliverable type	Report
Deliverable reference number	SC1-DTH-07-826121 / D1.4 / V1.0
Work package contributing to the deliverable	WP1
Due date	July 2022 – M43
Actual submission date	29th July, 2022

Responsible organisation	MPG
Editor	Mathurin Dorel
Dissemination level	PU
Revision	V1.0

Abstract	We report on the generation of CROPseq and drug screening data for two Ewing Sarcoma cell lines, one Hepatoblastoma cell line and one B-cell Acute Lymphoblastic Leukemia cell line.
Keywords	CROPseq, CRISPR, drug screen, perturbation data, molecular data, Ewing Sarcoma, Hepatoblastoma

The project iPC has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 826121.

Editor

Mathurin Dorel (MPG)

Contributors (ordered according to beneficiary numbers)

Cornelia Eckert (CHARITE)

Oliver Delattre (CURIE)

Cornelia Armengol (IGTP)

Roland Kappler (LMU)

Disclaimer

Executive Summary

CRISPR technology enables knocking out of genes with little off-target effect and has become the technology of choice to study genetic regulation. In this deliverable, we report on the establishment of the CROPseq pipeline for iPC; and on the generation of CROPseq and CRISPR screen data, with or without drug treatment, in four paediatric cancer cell lines. We provide an overview of those complex datasets, which will be analysed in more details in collaboration with the individual tumor working groups of iPC. Those data can be used to train mechanistic model, and are already being integrated in the Modcell model for iPC.

Table of Content

List of Figures

List of Tables

Chapter 1 Introduction

Perturbation of biological systems is the key experiment to understand how biological components interact with one another. Only by stimulating or removing a component is it possible to establish unambiguously the directions and components involved in signalling and regulation pathways. Many tools exist to perturb a biological system. Small molecules are of high interest because they can potentially be given to patients as therapy. As a complement, genetic perturbation methods such as CRISPR, overexpression or siRNA give the possibility to perturb any gene and help direct the drug discovery efforts. The use of targeted therapy in cancer has also highlighted the importance of understanding the topology and regulation mechanisms of signalling networks in order to exploit vulnerabilities of cancer cells and overcome resistance. In depth understanding of signalling networks can help design therapies targeting non-druggable mutations indirectly by exploiting synthetic lethality (Aly & Ganesan, 2011).

Chapter 2 Establishment of the CROPseq workflow

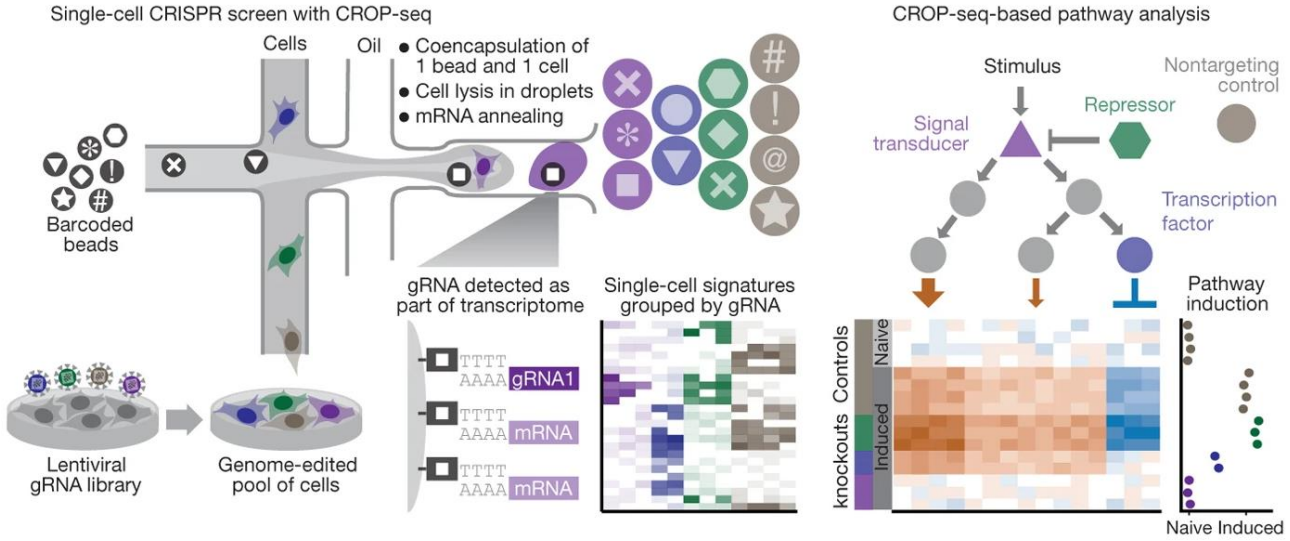

Figure 1: Principle of CROPseq (Datlinger et al., 2017).

Cells are transfected with a gRNA plasmid library with plasmid designed such that the spacer sequence of the guide is transcribed into a poly-adenylated RNA that can be captured with mRNA by poly-dT capture beads typically used in single cell sequencing. Bioinformatics analysis can then assign a guide to each single cell barcode to group cells by sgRNA. Pathway analysis methods can then be used to reconstruct the signalling pathways from the overlap in perturbations observed for the various guides.

Classical perturbation experiments are expensive and offer limited resolution. The CROPseq method was developed to address this shortcoming. Using a plasmid designed to render the sgRNA readable using standard single cell technology (Figure 1), it measures the effect of many perturbations in parallel at a single cell resolution (Datlinger et al., 2017). To perform CROPseq, we first inserted Staphylococcus Aureus Cas9 (saCas9) into cell lines of interest. To date, we generated A673cas9 (ES), TC71cas9 (ES), HepG2cas9 (HB) and 697cas9 (ALL). All cell lines with Cas9 were sequenced with RNAseq and Cas9 activity was confirmed using a T7 endonuclease assay in adherent cells or the CRISPRtest essential knock-out assay (Cellecta) in suspension cells.

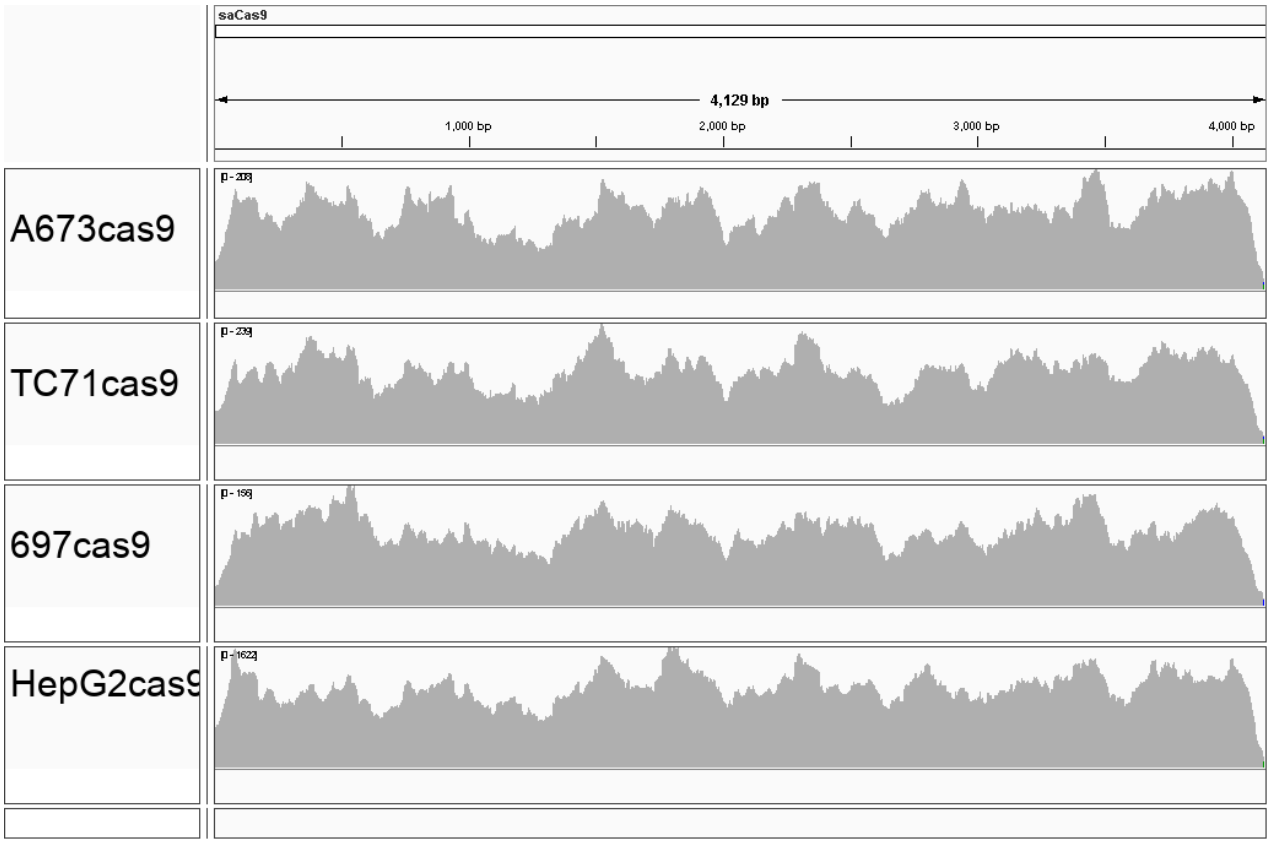

Figure 2: RNAseq coverage of the saCas9 ORF for iPC cas9 cell lines visualized using igv

In parallel, we generated lentiviral libraries containing the CROPseq-Puro plasmid (AddGene #86708) with sgRNA spacer sequences targeting about 60 genes each with three different guides, 6 positive controls targeting TUBB, MVD and HCFC1 each with two different guides, and 16 non-targeting negative controls. We designed the libraries to be lineage specific: the 60 targeted genes consisted of 40 shared signalling genes and about 20 lineage-specific genes (See Table 1). The libraries used guides from the Brunello library (Sanson et al., 2018) when available and were balanced to optimize the information obtained from the CROPseq experiment (Figure 3, Figure 4 and Figure 5)

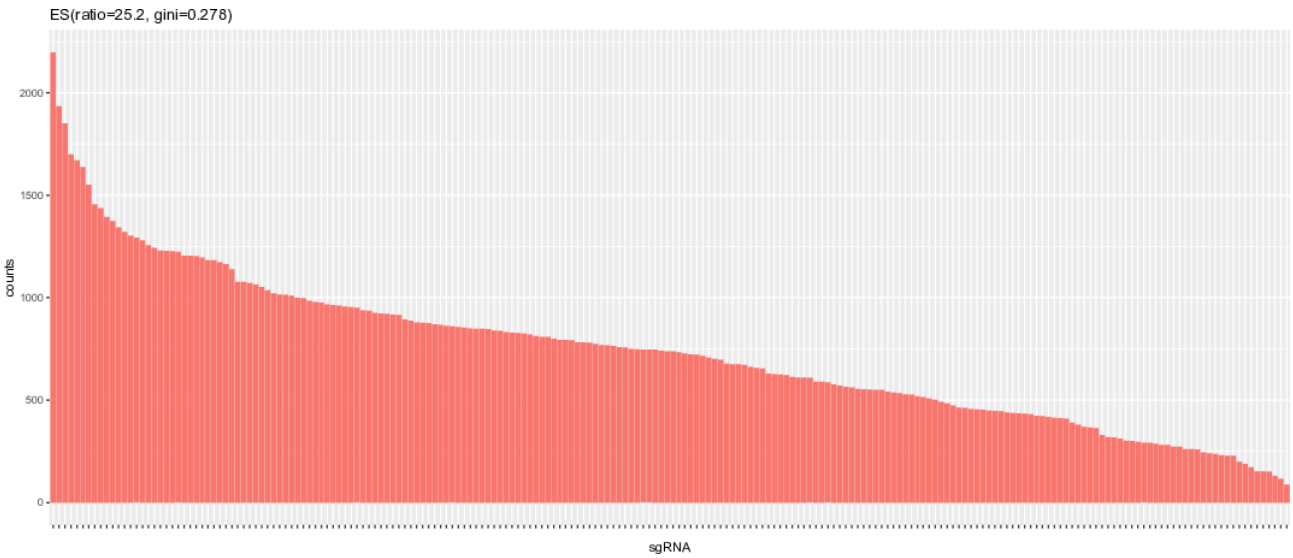

Figure 3: Coverage of the ES_CROP1 sgRNA library

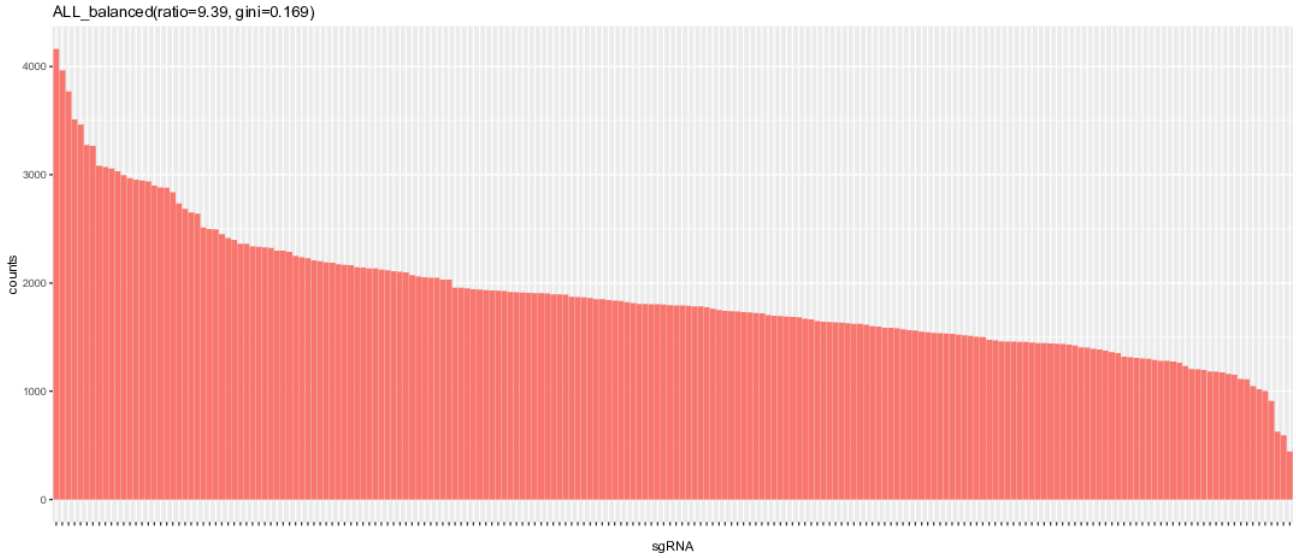

Figure 4: Coverage of ALL_CROP1 sgRNA library

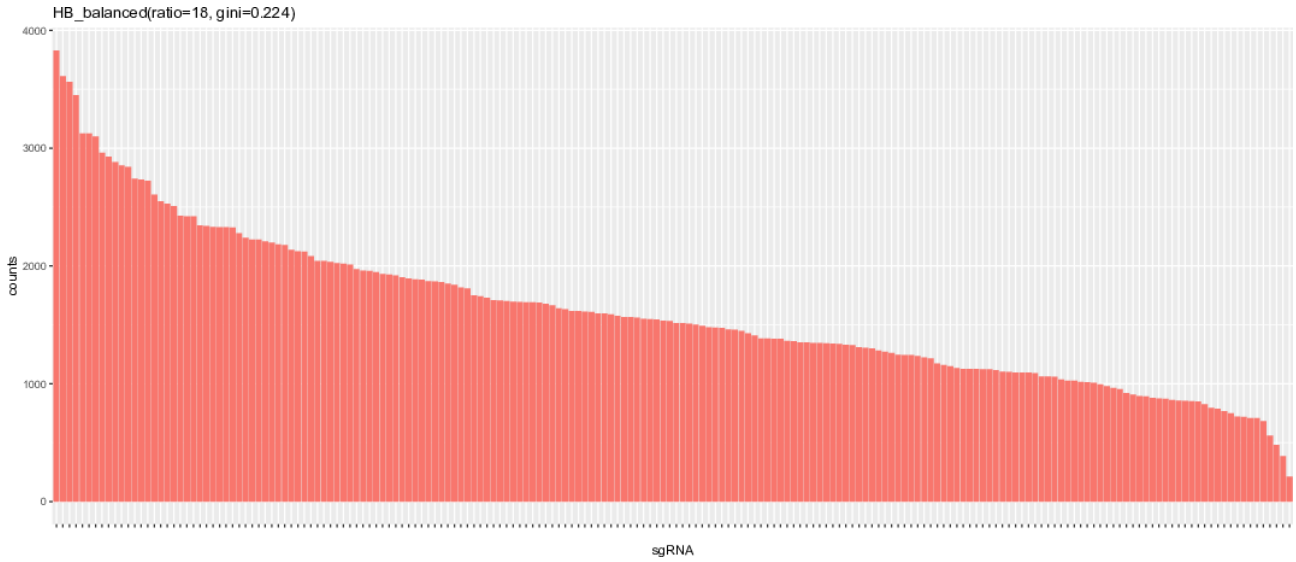

Figure 5: Coverage of the HB_CROP1 sgRNA libary

Chapter 3 Analysis of the Ewing Sarcoma CROPseq datasets

For iPC, we generated CROPSeq datasets for the Ewing Sarcoma cell lines A673 and TC-71. Staphylococcus Aureus Cas9 was stably expressed in each cell line, which were then exposed to a CROPseq library containing 206 guides targeting 62 genes. This library targets 40 genes involved in common signalling pathways with three guides each and 22 genes with specific importance in Ewing Sarcoma such as EWSR1 with three guides as well (Table 3). Similarly, we generated CROPseq dataset for the Acute Lymphocytic Leukemia cell line 697 treated with 20 genes specific for ALL and the hepatoblastoma cell line HepG2 with 21 genes specific for hepatoblastoma. As expected from dropout screens, cells with guides inactivating essential genes are depleted in all CROPseq datasets. More surprisingly, there is a negative correlation between the number of cells and the number of differentially expressed genes, suggesting that knock outs which affect the cells the most tend to also affect proliferation and survival (Figure 6, Figure 7, Figure 8, Figure 9). One particular exception to this trend is STAG2 in the Ewing Sarcoma cell line A673 (Figure 7), where many genes are affected but the cells are enriched compared to the initial library. This can be explained by the fact that STAG2 loss moderates the strong oncogenic signal of the EWS-FLI1 fusion in Ewing Sarcoma (Adane et al., 2021; Surdez et al., 2021). Interestingly however, FLI1 knock out does not alter the transcriptome. This result also reproduces findings that STAG2 inactivation enhances growth in A673 but decreases it in TC71 (Adane et al., 2021).

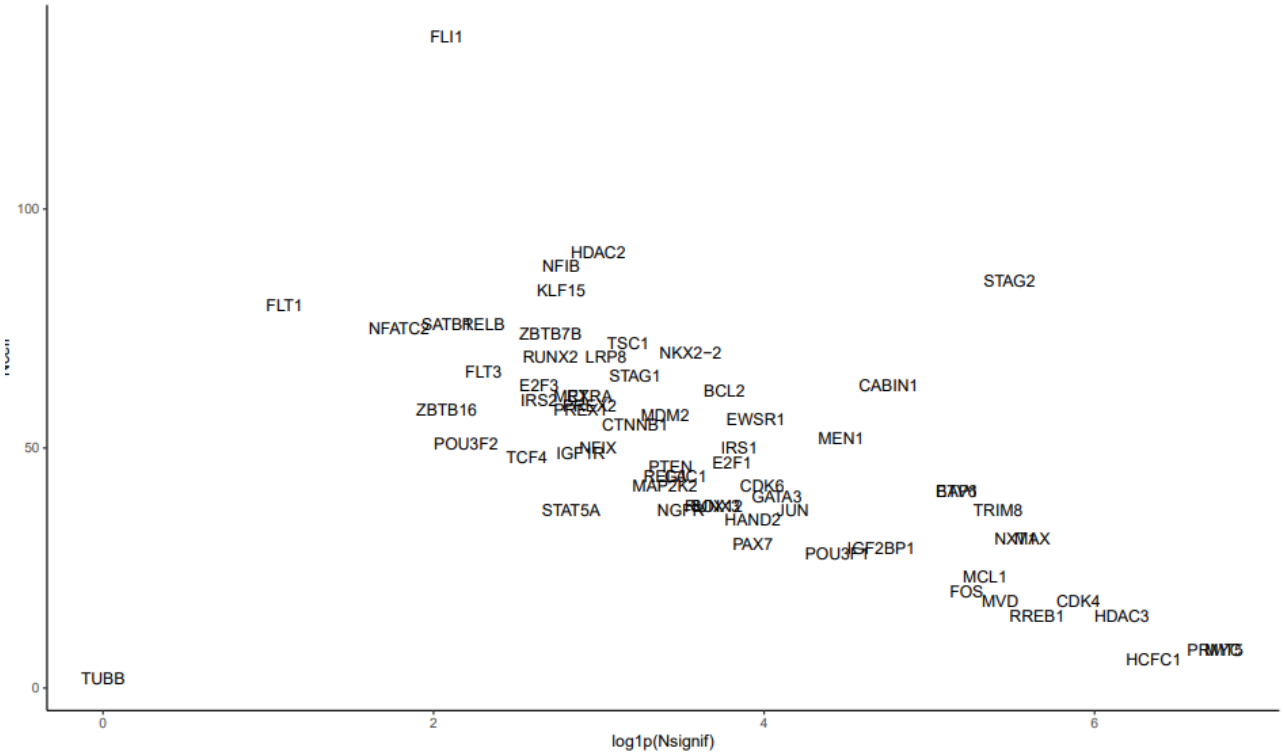

Figure 6: Number of cells versus number of differentially expressed genes in A679cas9 treated with ES_CROP1

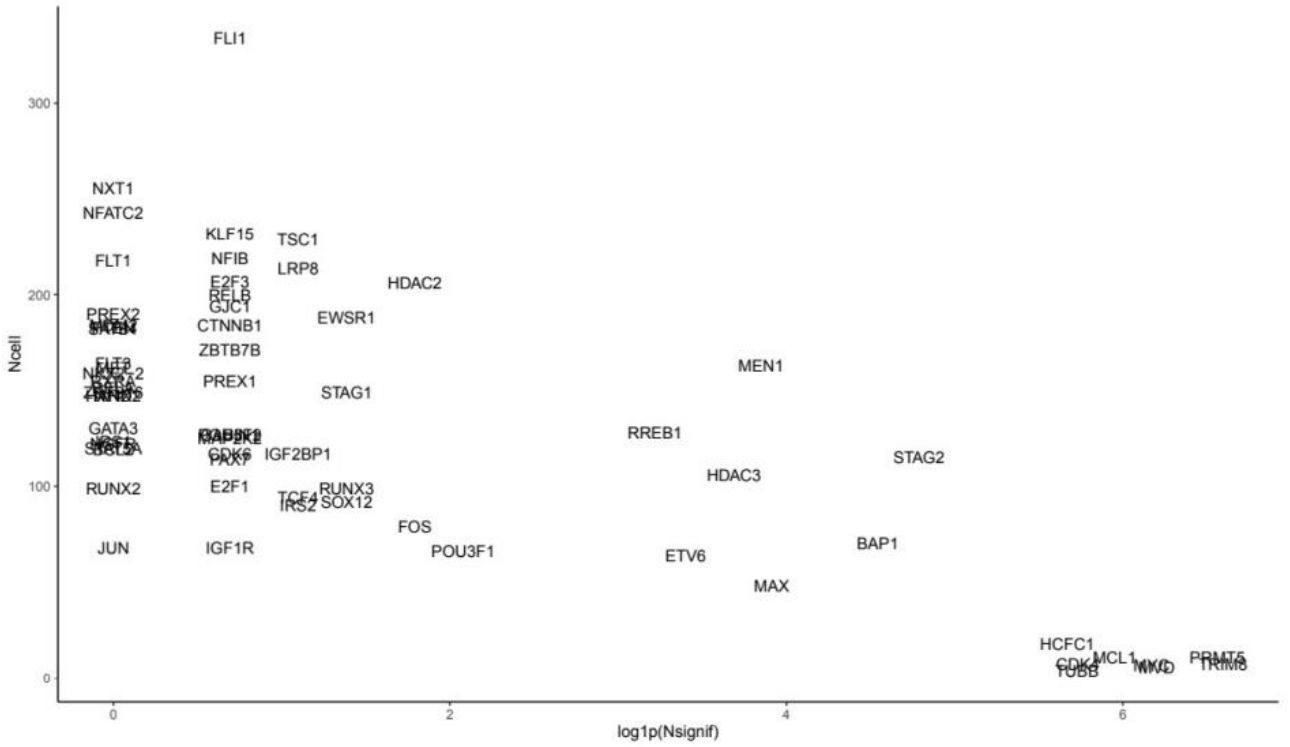

Figure 7: Number of cells versus number of differentially expressed genes in TC-71cas9 treated with ES_CROP1

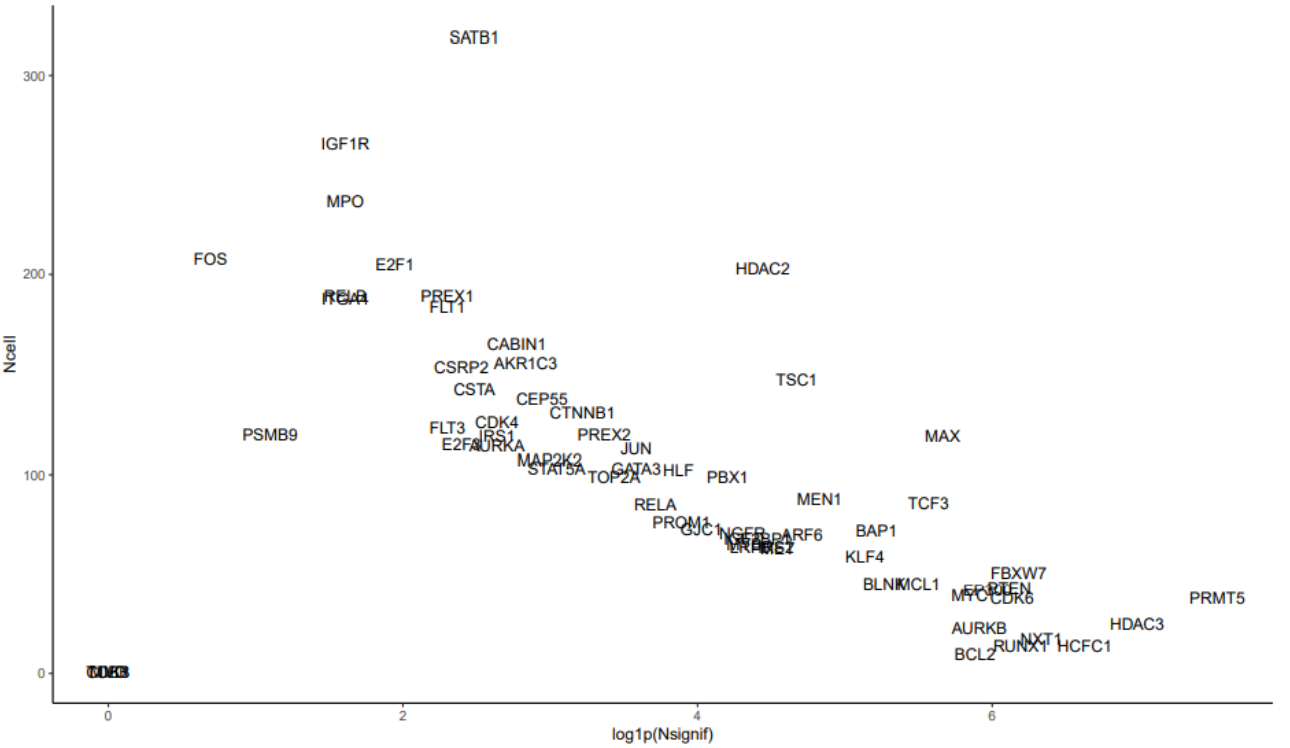

Figure 8: Number of cells versus numbers of differentially expressed genes in 697cas9 treated with ALL_CROP1

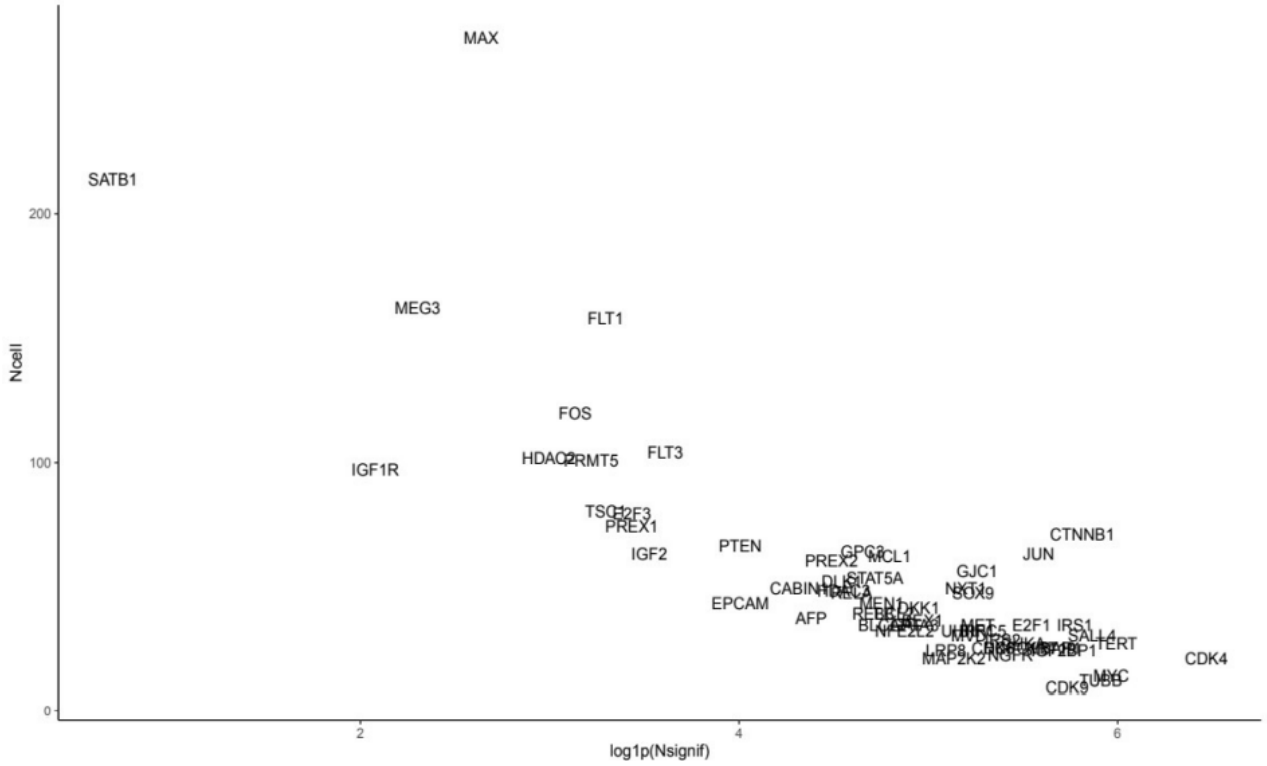

Figure 9: Number of cells versus number of differentially expressed genes in HepG2cas9 treated with HB_CROP1

The many dropouts arising from the use of single cell sequencing make calling differentially expressed genes from CROPseq data difficult. However single cell data provide new ways of analysing the data, such as cell cycle phase analysis and the possibility to correlate gene expression, which we used to find the effect of the gene knock-out. Looking at cell cycle phase distribution shows that most knock-out do not alter the cycling of the cells. TUBB KO is a notable exception in A673, with all 6 cells being stuck in G1. In other cell lines, the remaining cells from knock-outs with low counts also tend to be stuck in G1.

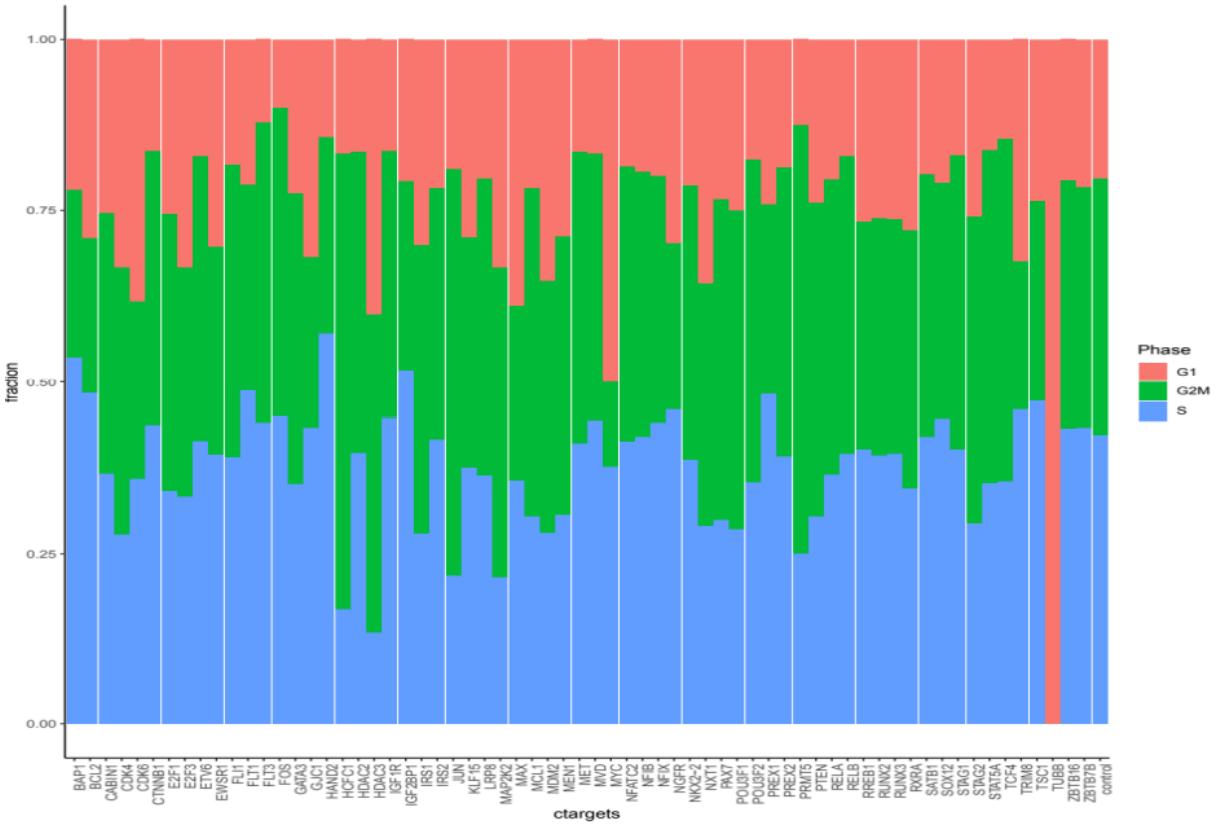

Figure 10: Distribution of cell cycle phase per target gene in A673cas9 treated with ES_CROP1

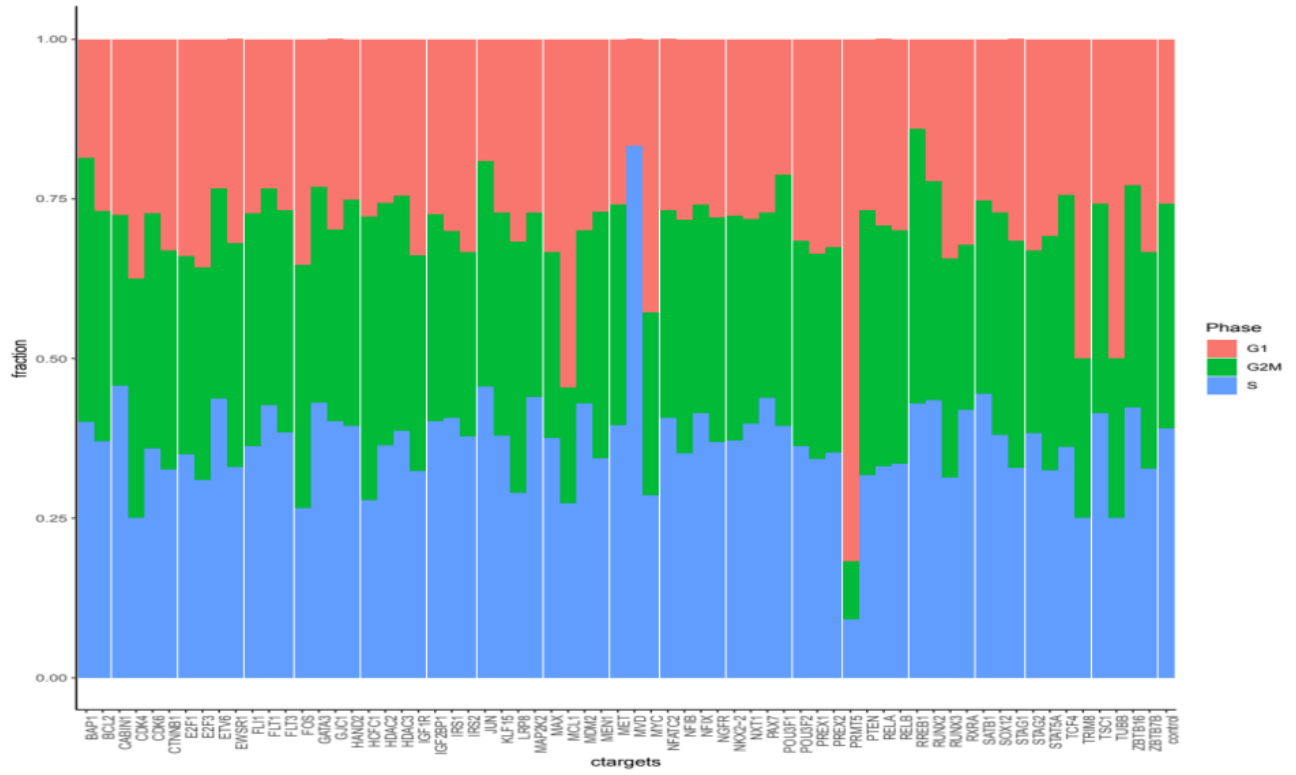

Figure 11: Distribution of cell cycle phase per target gene in TC-71cas9 treated with ES_CROP1

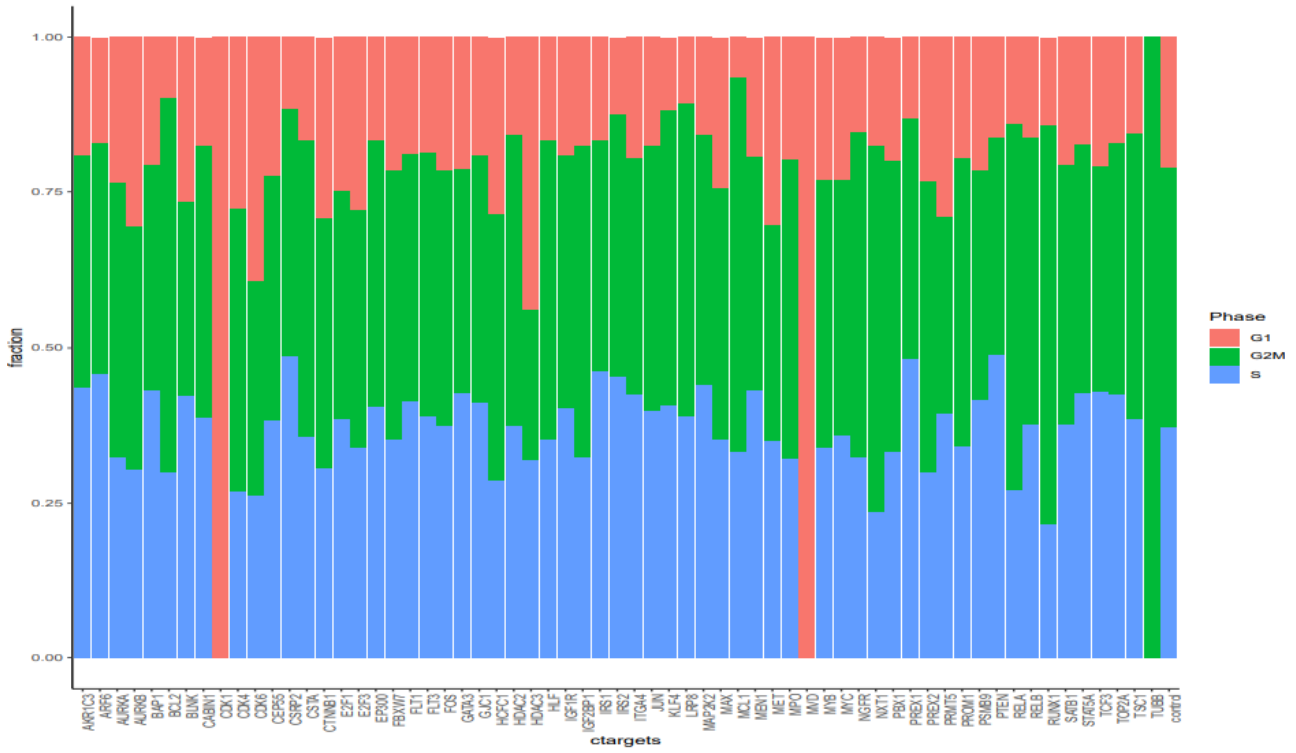

Figure 12: Distribution of cell cycle phase per target gene in 697cas9 treated with ALL_CROP1

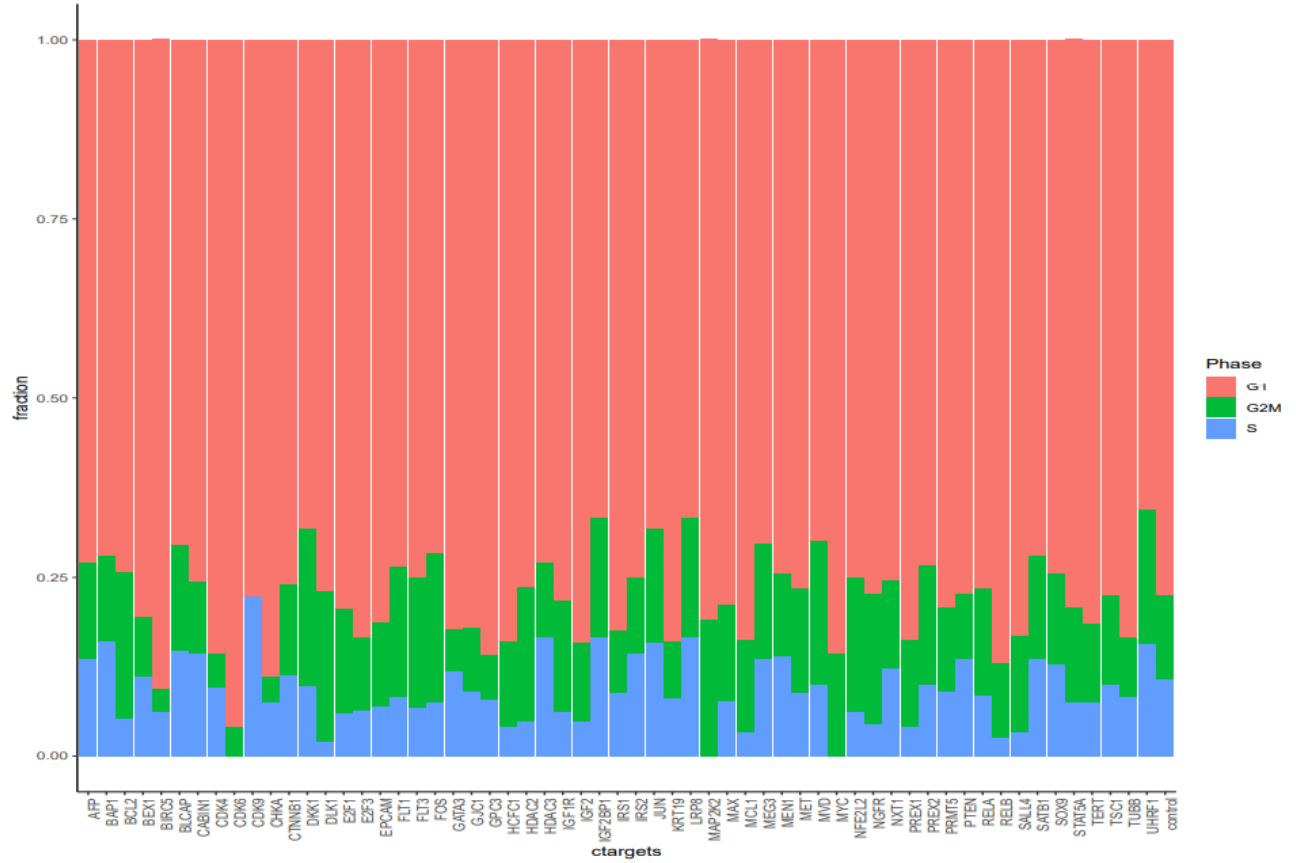

Figure 13: Distribution of cell cycle phase per target gene in HepG2cas9 treated with HB_CROP1

Finally, we used Mixscape workflow (Papalexi 2021), designed for CROPseq, to identify cells with guides that differ from the control and identify the genes differentially expressed in those cells (e.g. Figure 14 and Figure 15). Shortly, Mixscape annotates perturbed cells as knocked-out (KO) or non-perturbed (NP) depending on a perturbation score computed as a fold-change to the closest controls. KO cells are then used to call genes differentially expressed after the perturbation. The set of efficient knock-outs and the differentially expressed genes vary between lineages, but show high similarity for both Ewing Sarcoma cell lines, with a common regulations by STAG2, TRIM8, RREB1, BAP1 and ETV6. The CROPseq experiments are summarized in figures 16, 17, 18 and 19 which show the normalized expression of the top 20 differentially expressed genes for each effective knockouts. Only cells identified as proper KO by mixscape are plotted, with a downsampling to 50 for each target gene. The lists of genes will be used to refine the model by associating transcription factors with their targets in a lineage specific context and fitting their specific interaction kinetics. For non transcription factor perturbations, the data will be used to refine the model fit.

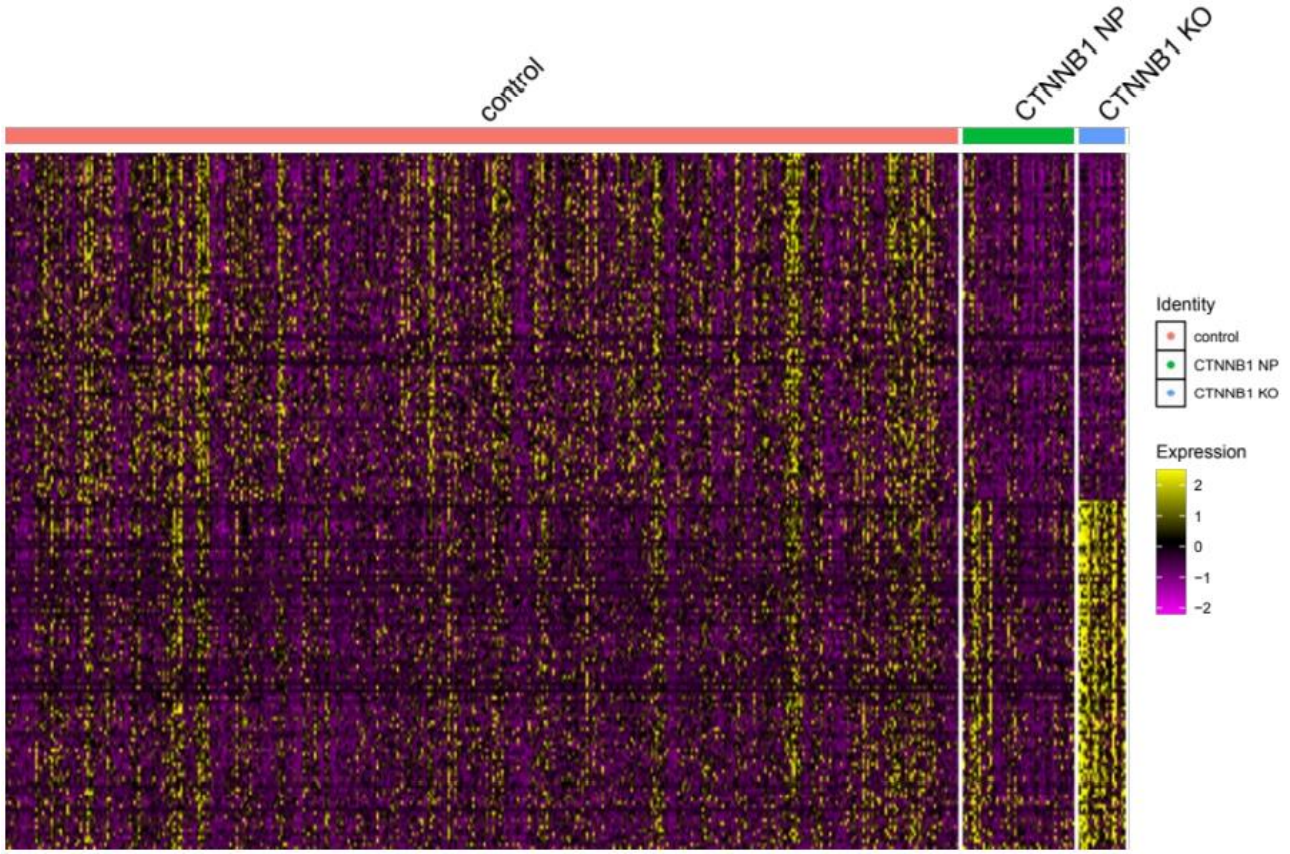

Figure 14: Expression in cells with a guide against CTNNB1 or a control guide of genes differentially expressed in CTNNB1 KO HepG2cas9 cells

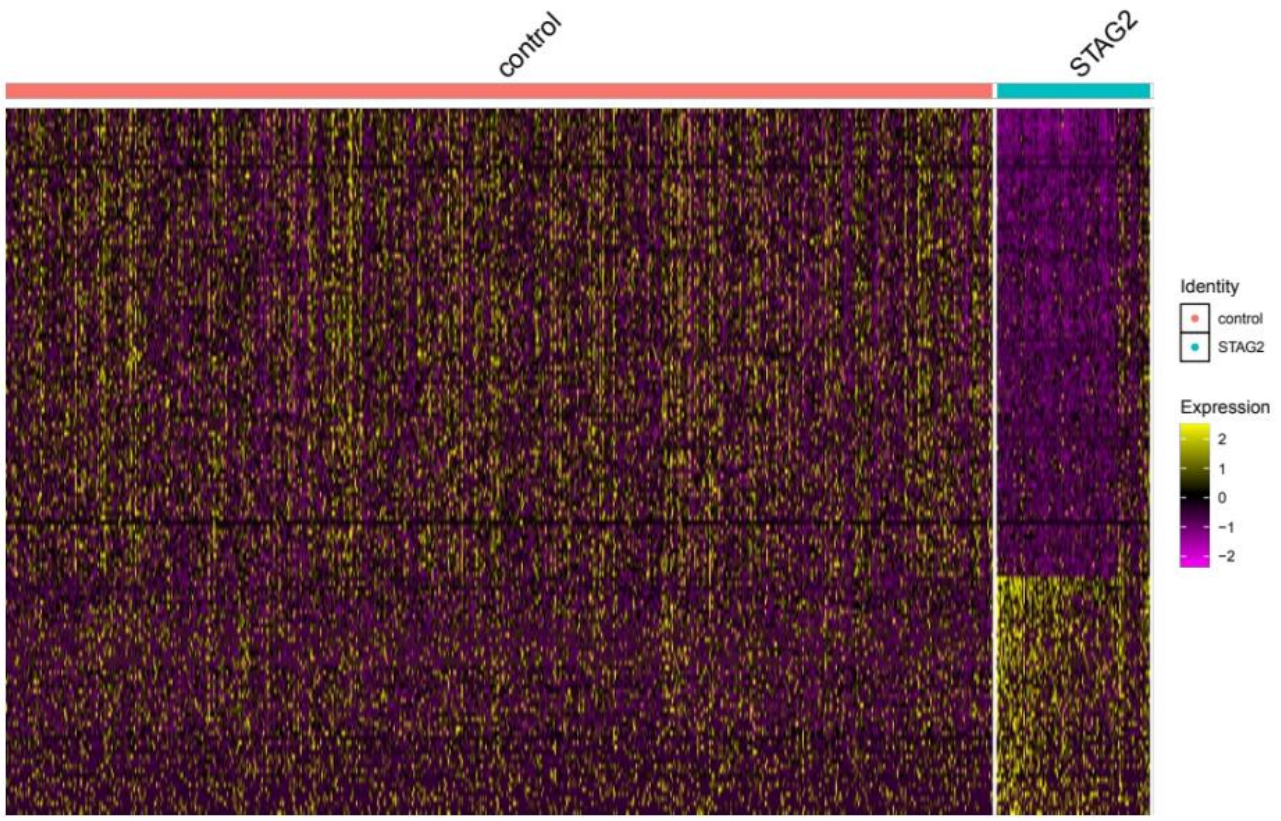

Figure 15: Expression in cells with a guide against STAG2 or a control guide of genes differentially expressed in STAG2 KO TC-71cas9 cells

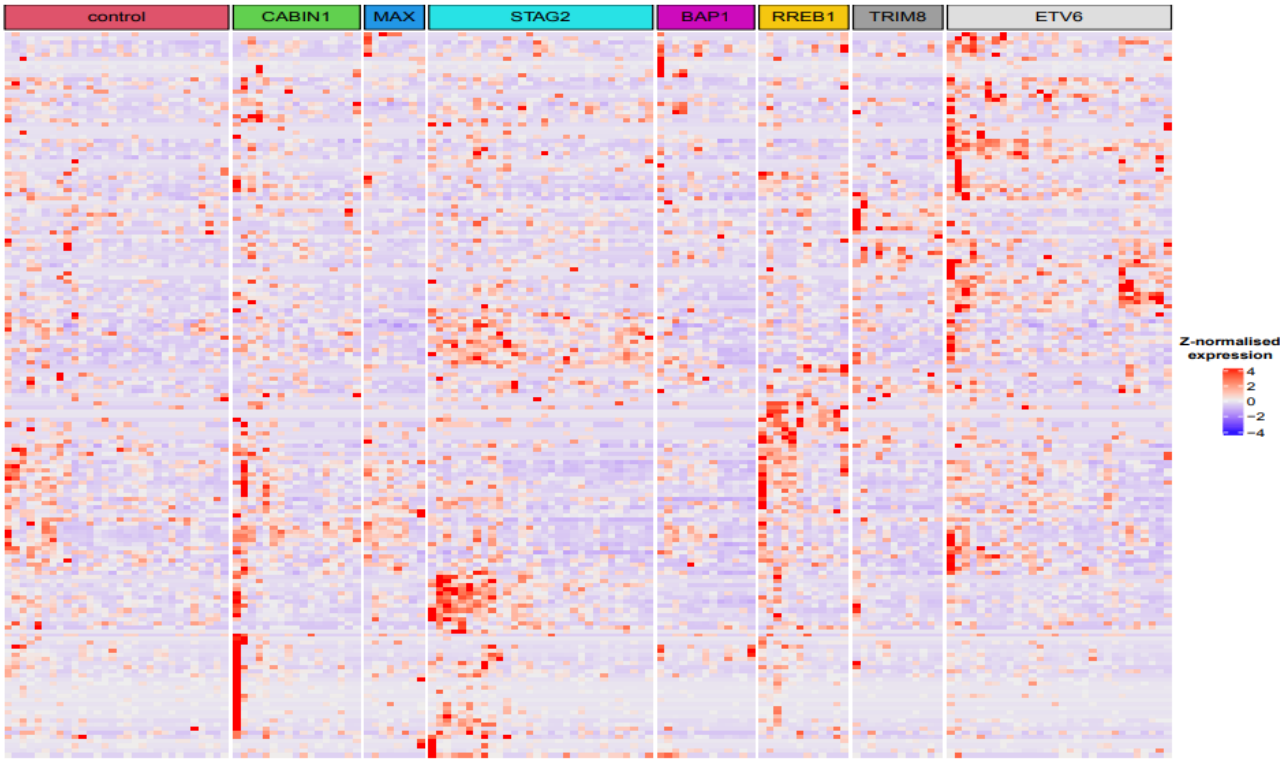

Figure 16: Summary of the A673 ES_CROP1 dataset

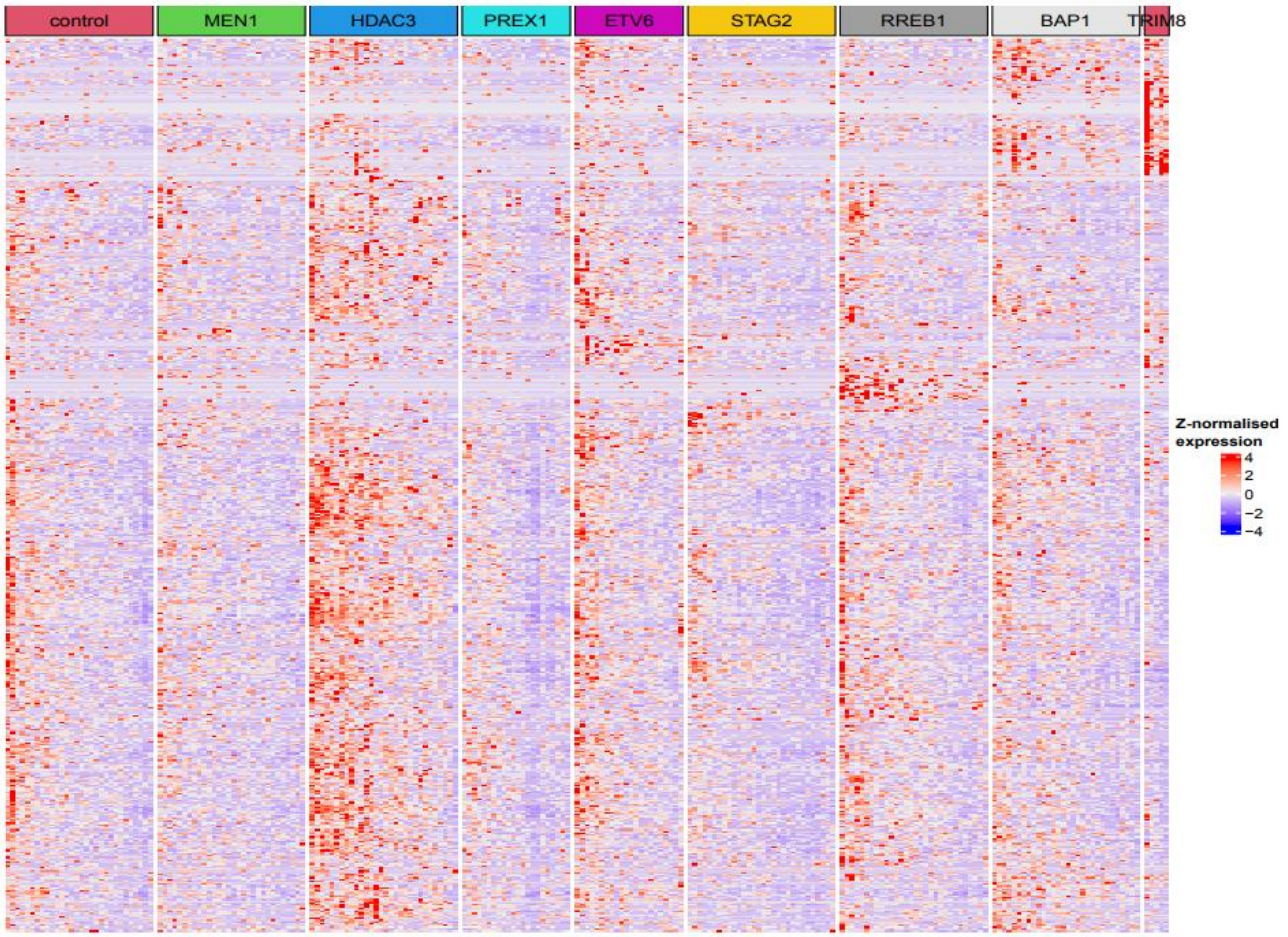

Figure 17: Summary of the TC71 ES_CROP1 dataset

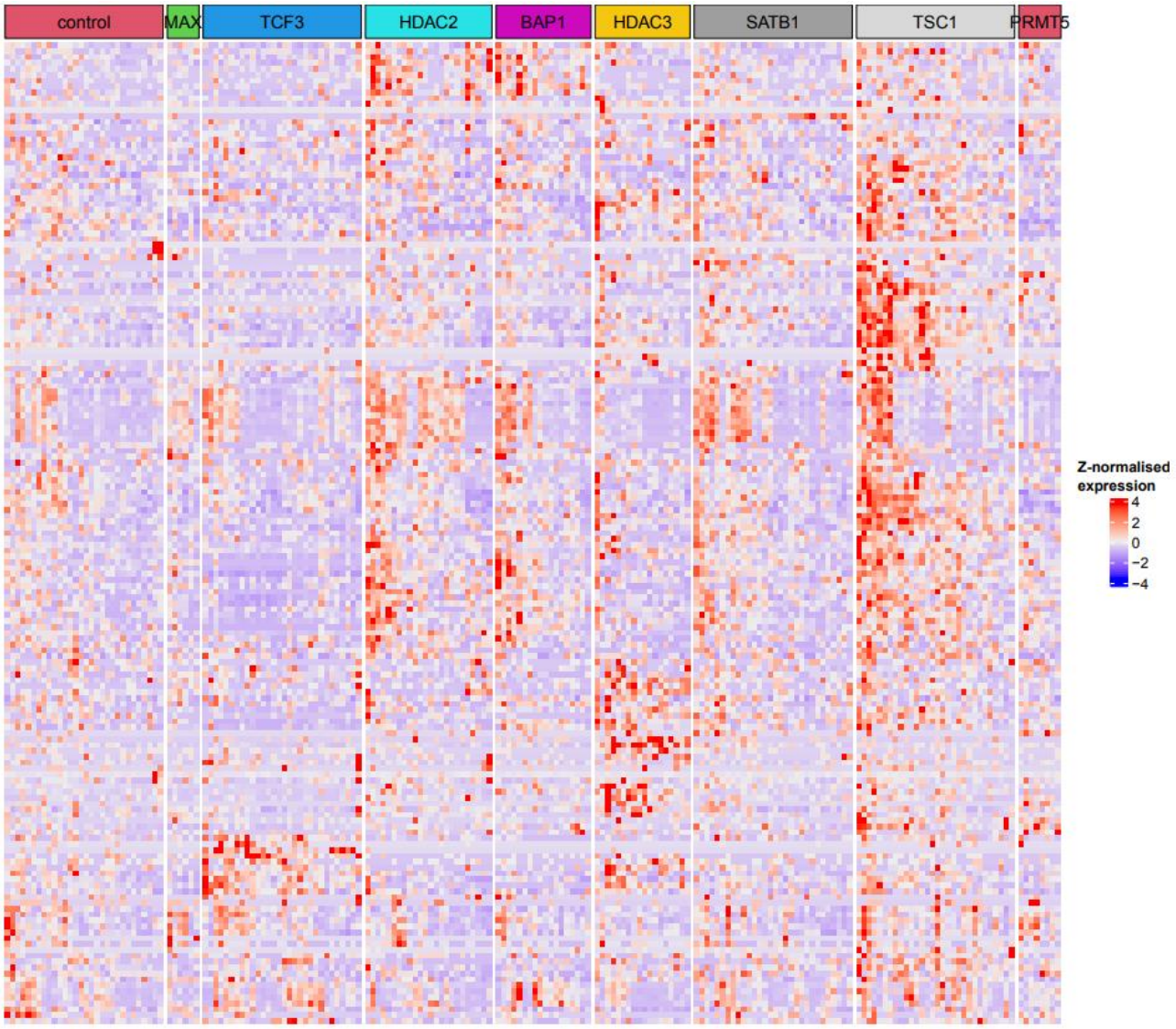

Figure 18: Summary of the 697 ALL_CROP1 dataset

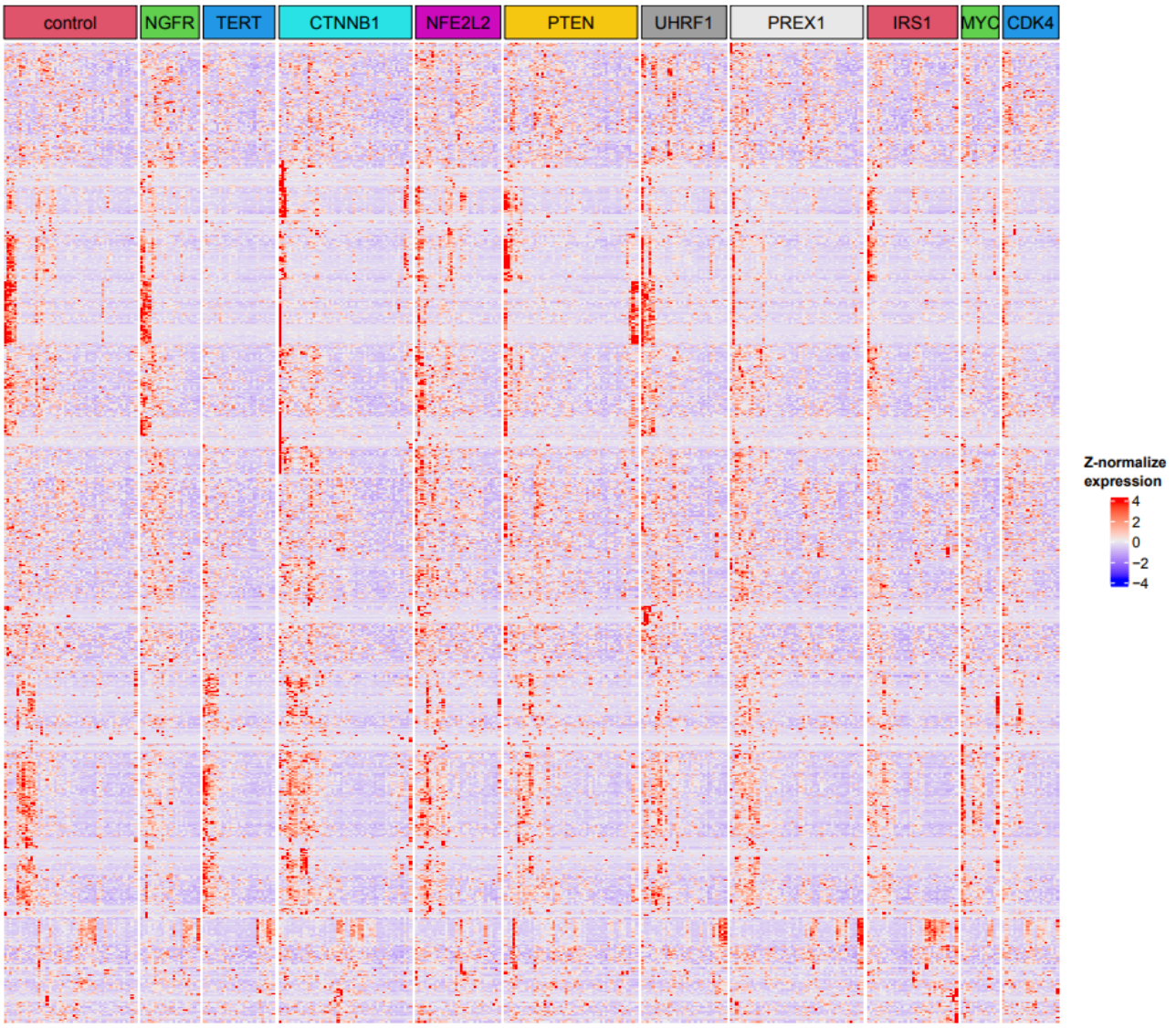

Figure 19: Summary of the HepG2 HB_CROP1 experiment

Chapter 4 CRISPR drop-out screen

In order to expand the knowledge on drug resistance and the model predictions, we performed CRISPR screens with various drugs used for cancer treatment. We assembled a larger library sgRNA library containing 1790 guides targeting genes from the model at Alacris to perform CRISPR screens in drug contexts of the cell lines.

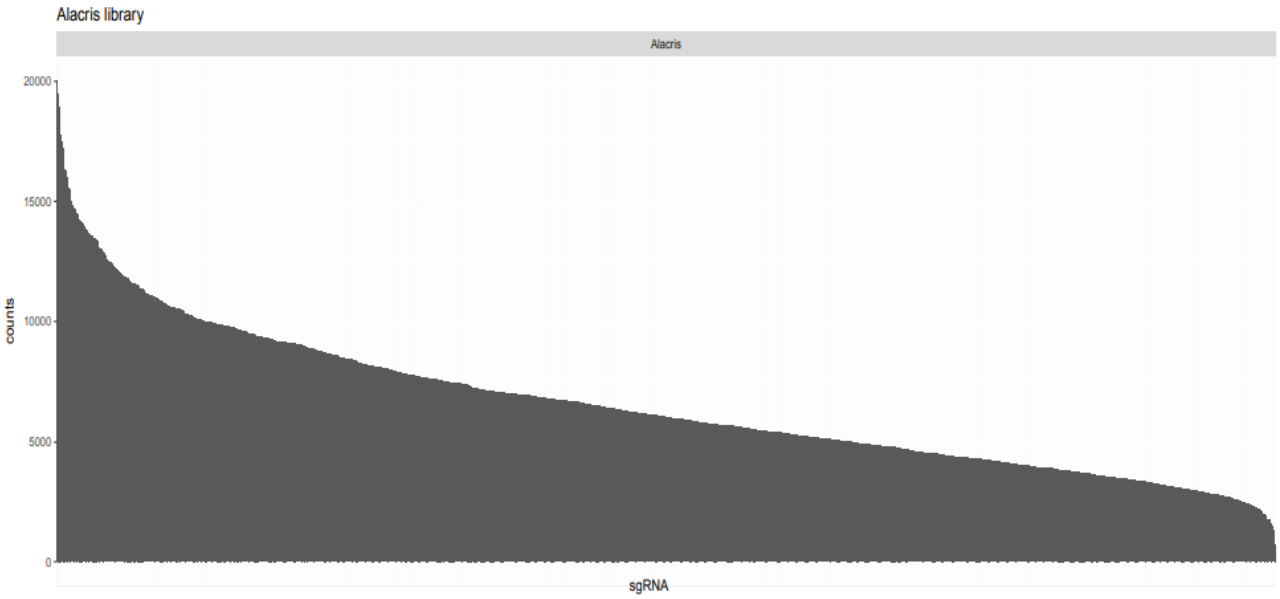

Figure 20: Coverage of the Alacris library, guides are ordered by number of reads

The first batch of inhibitors, consisting of Venetoclax (Bcl-2 inhibitor used in leukemia) and Olaparib (PARP inhibitor used in breast cancer) was used on 697 (ALL), HepG2 (HB) and TC71 (ES). The main source of variation in the resulting data set is the cell lines, and the second source of variation the time point (Figure 21). This indicates that the cell line specific dependencies dominate the variation in sgRNA abundance. This is expected from such data, with different cell lines having different dependencies that are only marginally affected by the drugs, and the distribution of guides changing with time as selection changes their relative abundance (Figure 22).

Figure 21: PCA of the drug CRISPR-screen data

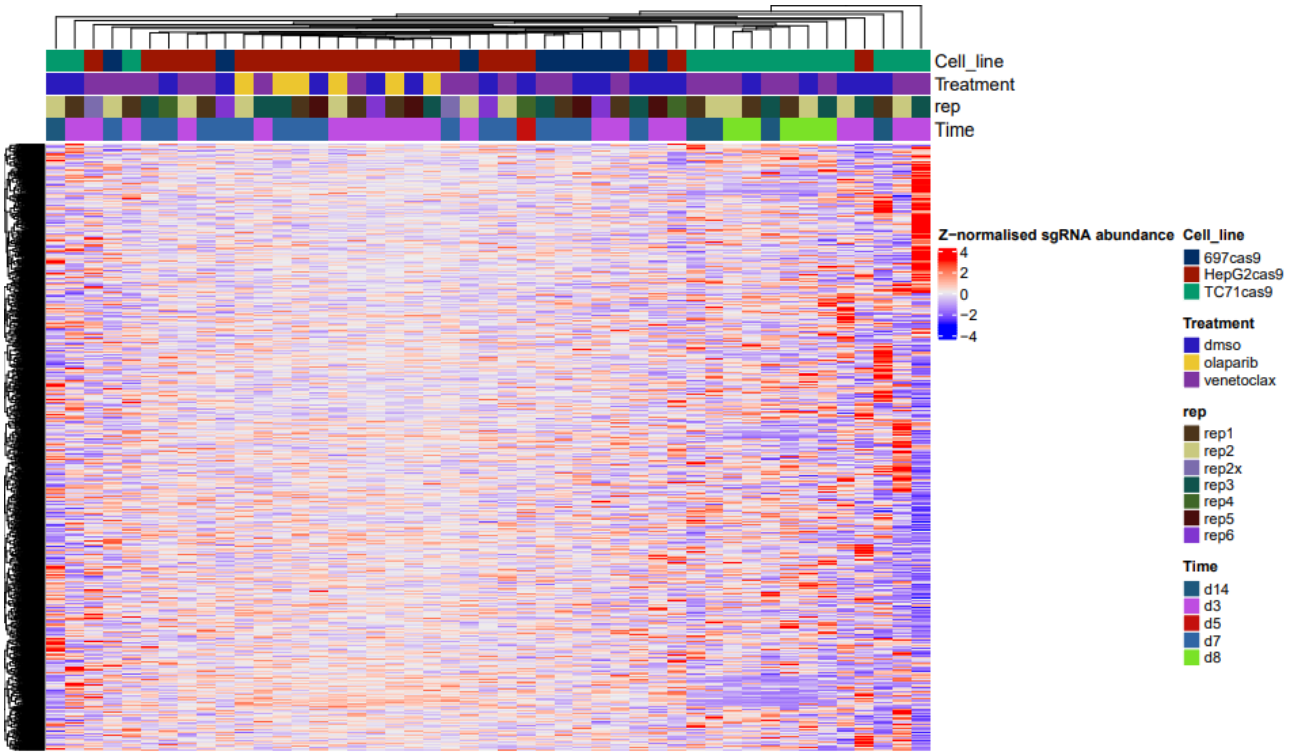

Figure 22: sgRNA abundance Z-normalised across samples

Chapter 5 Summary and Conclusion

In collaboration under the umbrella of the iPC project, we established the CROPseq method and designed libraries dedicated to paediatric tumors. CRISPR is a recent tool added to perturbation biology that allows precise knock-out of any genomic location. CROPseq, by coupling CRISPR knock-out with single cell sequencing technology, gives unprecedented insights into the response of paediatric cancer cell lines to various perturbation that will of great value to improve signalling models.

In addition, we used the expertise acquired with CRISPR libraries to perform gene-drug interaction screens that will provide important phenotypic readouts to train our models. Indeed, the relative abundance of guides after a dropout screen is directly related to relative growth rates of the cells with those perturbations, which can be readily integrated in the model as a growth rate after knock-out provided that a basal growth rate is available for the cell line.

Finally, we are continuing to use our pipeline to expand this dataset with CROPseq and CRISPR screen data using more cell lines representing other cancer lineages or subtypes and various drugs with different mechanism of actions.

Chapter 6 List of Abbreviations

Abbreviation	Translation
ALL	Acute Lymphoblastic Leukemia
ES	Ewing Sarcoma
HB	Hepatoblastoma
ORF	Open Reading Frame

Chapter 7 Bibliography

[1] Adane, B., Alexe, G., Seong, B. K. A., Lu, D., Hwang, E. E., Hnisz, D., Lareau, C. A., Ross, L., Lin, S., Dela Cruz, F. S., Richardson, M., Weintraub, A. S., Wang, S., Iniguez, A. B., Dharia, N. V., Conway, A. S., Robichaud, A. L., Tanenbaum, B., Krill-Burger, J. M., … Stegmaier, K. (2021). STAG2 loss rewires oncogenic and developmental programs to promote metastasis in Ewing sarcoma. Cancer Cell, 39(6), 827-844.e10. https://doi.org/10.1016/J.CCELL.2021.05.007

[2] Aly, A., & Ganesan, S. (2011). BRCA1, PARP, and 53BP1: conditional synthetic lethality and synthetic viability. Journal of Molecular Cell Biology, 3(1), 66–74. https://doi.org/10.1093/jmcb/mjq055

[3] Datlinger, P., Rendeiro, A. F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., Schuster, L. C., Kuchler, A., Alpar, D., & Bock, C. (2017). Pooled CRISPR screening with single-cell transcriptome readout. Nature Methods, 14(3), 297–301. https://doi.org/10.1038/nmeth.4177

[4] Surdez, D., Zaidi, S., Grossetête, S., Laud-Duval, K., Ferre, A. S., Mous, L., Vourc'h, T., Tirode, F., Pierron, G., Raynal, V., Baulande, S., Brunet, E., Hill, V., & Delattre, O. (2021). STAG2 mutations alter CTCF-anchored loop extrusion, reduce cis-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma. Cancer Cell, 39(6), 810-826.e9. https://doi.org/10.1016/J.CCELL.2021.04.001

[5] Sanson, K. R., Hanna, R. E., Hegde, M., Donovan, K. F., Strand, C., Sullender, M. E., Vaimberg, E. W., Goodale, A., Root, D. E., Piccioni, F., & Doench, J. G. (2018). Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. Nature Communications, 9(1), 5416. https://doi.org/10.1038/s41467-018-07901-8

Chapter 8 Data Tables

Common genes	E2F3,IRS1,MCL1,BAP1,RELB,NXT1,CABIN1,RELA, IGF1R,MAX,HDAC2,JUN,SATB1,NGFR,BCL2,MYC,PTEN, MET,CTNNB1,IRS2,FLT3,MAP2K2,STAT5A,HDAC3,E2F1, PRMT5,GATA3,PREX2,TSC1,IGF2BP1,PREX1,FLT1, LRP8,CDK6,FOS,CDK4,MEN1,GJC1
Positive controls	MVD,TUBB,HCFC1
HB_CROP1 specific genes	KRT19,IGF2,SOX9,BLCAP,DLK1,CDK9,BIRC5,NFE2L2, UHRF1,SALL4,TERT,EPCAM,DKK1,AFP,GPC3,BEX1, MEG3,CHKA
ALL_CROP1 specific genes	TOP2A,CSTA,ARF6,HLF,PBX1,PROM1,CDK1,CEP55, KLF4,AKR1C3,MYB,AURKA,PSMB9,EP300,CSRP2,TCF3, BLNK,MPO,RUNX1,AURKB,ITGA4,FBXW7
ES_CROP1 specific genes	TRIM8,POU3F1,RXRA,NFATC2,RUNX2,TCF4,STAG2, PAX7,POU3F2,EWSR1,HAND2,RUNX3,FLI1,RREB1, NFIB,KLF15,ETV6,SOX12,STAG1,NFIX,MDM2,ZBTB7B, ZBTB16,NKX2-2

Table 1: Gene categorisation for the CROPseq libraries

sgRNA_name	target_name	spacer_sequence
sgTOP2A_1	TOP2A	TCCCGTCAGAACATGGACCC
sgTOP2A_2	TOP2A	AGCATTGTAAAGATGTATCG
sgTOP2A_3	TOP2A	TGTACGCTTATCCTGACTGA
sgCSTA_1	CSTA	GATACCTGGAGGCTTATCTG
sgCSTA_2	CSTA	ACCCGCCACTCCAGAAATCC
sgCSTA_3	CSTA	AATCCAGGAGATTGTTGATA
sgARF6_1	ARF6	CCTGGCGAGCCTCATCGATG
sgARF6_2	ARF6	GTGTAGTAATGCCGCCAGAG
sgARF6_3	ARF6	CCAAGGTCTCATCTTCGTAG
sgHLF_1	HLF	TGATGAGAGTAACAGCCCGA
sgHLF_2	HLF	GGTCTGGCTCATAACCCACT
sgHLF_3	HLF	TGGAAAGTATCTCCGTCATA

sgPBX1_1	PBX1	CAGCAACCCTTACCCCAGTG
sgPBX1_2	PBX1	CAGGAGATTCATCACGTGGG
sgPBX1_3	PBX1	TTCACACAACACATTAAACA
sgPROM1_1	PROM1	TGTCGTTGCTGTAACAAATG
sgPROM1_2	PROM1	CTCACCTGCTACGACAGTCG
sgPROM1_3	PROM1	CTGTGAACCTTACACGAGCA
sgCDK1_1	CDK1	GACAAAACACAATCCCCTGT
sgCDK1_2	CDK1	GTATTCCAAAAGCTCTGGCA
sgCDK1_3	CDK1	ACCCTTATACACAACTCCAT
sgCEP55_1	CEP55	CTAGAAAATTCGAGTCCTTG
sgCEP55_2	CEP55	CAATACACTCCGTTTATCAC
sgCEP55_3	CEP55	GCGACTGAGAGACCAACTGA
sgKLF4_1	KLF4	AGCGATACTCACGTTATTCG
sgKLF4_2	KLF4	CCTGCGGCCAGAATTGGACC
sgKLF4_3	KLF4	GCCAGGTTGAAGGGAGCCGT
sgAKR1C3_1	AKR1C3	AGAAATCTAGCAATTTACTC
sgAKR1C3_2	AKR1C3	AATGAGCAGAATCTATATGG
sgAKR1C3_3	AKR1C3	GGGTGTCAAACTTCAACCGC
sgMYB_1	MYB	AGTCTGGAAAGCGTCACTTG
sgMYB_2	MYB	ACCAGGCACACAAGAGACTG
sgMYB_3	MYB	TATTTACATGTAACGCTACA
sgAURKA_1	AURKA	CTTCGAATGACAGTAAGACA
sgAURKA_2	AURKA	CCATATAGAAAATAATCCTG
sgAURKA_3	AURKA	CCTGAAAACTCACCGAAGGT
sgPSMB9_1	PSMB9	GTAGATGCGCTCGTGCAGCG
sgPSMB9_2	PSMB9	CCAGACCCATTACCCCGGTG
sgPSMB9_3	PSMB9	TATCAGCTATAAATATCGAG
sgEP300_1	EP300	GGTACGACTAGGTACAGGCG

sgEP300_2	EP300	ATGGTGAACCATAAGGATTG
sgEP300_3	EP300	GTGGCACGAAGATATTACTC
sgCSRP2_1	CSRP2	GGCACGCTTAACATGGACCG
sgCSRP2_2	CSRP2	GGATTTGTTGTAGGCCTGTG
sgCSRP2_3	CSRP2	GTGGTACACGGTCCTCCCAC
sgTCF3_1	TCF3	GTTATTGCTTGAGTGATCCG
sgTCF3_2	TCF3	GCTGGGCGATAAGGCACCGG
sgTCF3_3	TCF3	CACCAGCACGAGCGTATGGT
sgBLNK_1	BLNK	CAGAGGCGAGTATATAGGTG
sgBLNK_2	BLNK	CAAGTCCCACCCAAACCCAA
sgBLNK_3	BLNK	GTCTGTGACTTGACCCTCGG
sgMPO_1	MPO	TCATTGTAGGAACGGTACGT
sgMPO_2	MPO	AGTAGGATAGGAGTTCCATG
sgMPO_3	MPO	AAGTAAGAGGGTGTGCATGG
sgRUNX1_1	RUNX1	CTGATCGTAGGACCACGGTG
sgRUNX1_2	RUNX1	CACTTCGACCGACAAACCTG
sgRUNX1_3	RUNX1	TAGATGATCAGACCAAGCCC
sgAURKB_1	AURKB	ATTCTAGAGTATGCCCCCCG
sgAURKB_2	AURKB	TCTTTCCGGAGGACTCGCTG
sgAURKB_3	AURKB	CATCAACCCATACTGCAGGT
sgITGA4_1	ITGA4	AGTTCCAATACCTACCACGA
sgITGA4_2	ITGA4	CATATTTGTCACTTCCAACG
sgITGA4_3	ITGA4	CTCACCATCGGTTCGCCCCG
sgFBXW7_1	FBXW7	ACAGAATTGATACTAACTGG
sgFBXW7_2	FBXW7	AAGAGCGGACCTCAGAACCA
sgFBXW7_3	FBXW7	GTTGGAGTAGAACCTAGACC
NTC_1	CTRL00018	GGTGATCCTAGTCGACTGGC
NTC_2	CTRL00022	GCTATTGTCGGCTGGATGGA

NTC_3	CTRL00080	GCGCATCAATATGCCCGCAC
NTC_4	CTRL00087	GAGGTCCCGCCTCCGCCCAA
NTC_5	CTRL00096	GGCTAACGCACGCCGAGGTG
NTC_6	CTRL00196	GGGCCTTAATAGTTCAACGC
NTC_7	CTRL00275	GCAGGACGATAGTAACGTCC
NTC_8	CTRL00320	GAGAACGTGATAAGACTCGG
NTC_9	CTRL00405	GGACTGACGTATACGCTTGC
NTC_10	CTRL00484	GTGAATACCCGTGACGGACA
NTC_11	CTRL00493	GGCAATCGGCACGGCAAGGT
NTC_12	CTRL00545	GGGCTGTTCTCACTCGTAGC
NTC_13	CTRL00616	GGTTAAAGAGATTGTACGCC
NTC_14	CTRL00637	GGGTTGCACGTAGGATATTC
NTC_15	CTRL00640	GCGTATTGTATCCGCCACCG
NTC_16	CTRL00673	GAGTAAAGCCGATTATATCG
sgMVD_1	MVD	AAGCTGACAGGCAGTACCGT
sgMVD_2	MVD	CTGGTGCAGAGTGACGCTCA
sgTUBB_1	TUBB	GCTGACCACACCAACCTACG
sgTUBB_2	TUBB	CCCCACCGGCACCTACCACG
sgHCFC1_1	HCFC1	GTGGAAGTGTACCAACACGC
sgHCFC1_2	HCFC1	GGTGCCTTTCACAACCAACG
sgE2F3_1	E2F3	GGACCTCAAACTGTTAACCG
sgE2F3_2	E2F3	AGATGGTTTAAAAACCCCCA
sgE2F3_3	E2F3	AATCTCCCTCAGAAAAAACG
sgIRS1_1	IRS1	CCGAAGCACTAGATCGCCGT
sgIRS1_2	IRS1	TCTTGCTGGTCAGGCAAAGG
sgIRS1_3	IRS1	GCCCCCCCCGACGCTCCAAG
sgMCL1_1	MCL1	AGGCGCTGGAGACCTTACGA
sgMCL1_2	MCL1	GTAATAACACCAGTACGGAC

sgMCL1_3	MCL1	AGTCGCTGGAGATTATCTCT
sgBAP1_1	BAP1	CACGGACGTATCATCCACCA
sgBAP1_2	BAP1	GAACCGTCAGACAGTACTAG
sgBAP1_3	BAP1	TCTACCCCATTGACCATGGT
sgRELB_1	RELB	ATTGAGCGGAAGATTCAACT
sgRELB_2	RELB	GCCTCATATCGGGACCAGCA
sgRELB_3	RELB	CGGTGCAGTCTTTCCCCACG
sgNXT1_1	NXT1	ACTACACCACCATGGATAAG
sgNXT1_2	NXT1	AGCATTGCCATTCCAGACCA
sgNXT1_3	NXT1	ACAGATGACAACAAGGACCG
sgCABIN1_1	CABIN1	GTAATCGTGGTCAATCGGAG
sgCABIN1_2	CABIN1	GTAGTGCAGCAAGTAAACGG
sgCABIN1_3	CABIN1	CTGGAGAACCTAACCAACGG
sgRELA_1	RELA	TCAATGGCTACACAGGACCA
sgRELA_2	RELA	GCTTCCGCTACAAGTGCGAG
sgRELA_3	RELA	GGAAGATCTCATCCCCACCG
sgIGF1R_1	IGF1R	GGAGAACGACCATATCCGTG
sgIGF1R_2	IGF1R	TTCCGAAATTTACCGCATGG
sgIGF1R_3	IGF1R	GGTACAATGTGAAAGGCCGA
sgMAX_1	MAX	TATTCCAGGAAGAGCAACCG
sgMAX_2	MAX	GAAGAGCATTCTGCCGCTTG
sgMAX_3	MAX	AATATATCCAGTATATGCGA
sgHDAC2_1	HDAC2	GATGTATCAACCTAGTGCTG
sgHDAC2_2	HDAC2	TACAACAGATCGTGTAATGA
sgHDAC2_3	HDAC2	CCTCCTCCAAGCATCAGTAA
sgJUN_1	JUN	GGCGGCGCAGCCGGTCAACG
sgJUN_2	JUN	GCTCTCGGACGGGAGGAACG
sgJUN_3	JUN	TGAACCTGGCCGACCCAGTG

sgSATB1_1	SATB1	TAGGTGTTGATACGAGCCCA
sgSATB1_2	SATB1	TATTCATAGATCTACTGACA
sgSATB1_3	SATB1	CATTGAATATGATTGCAAGG
sgNGFR_1	NGFR	ACAGGCCTGTACACACACAG
sgNGFR_2	NGFR	CGACGGCACGTATTCCGACG
sgNGFR_3	NGFR	GTGTGGACCGTGTAATCCAA
sgBCL2_1	BCL2	TGTCGCAGAGGGGCTACGAG
sgBCL2_2	BCL2	CTGACGCCCTTCACCGCGCG
sgBCL2_3	BCL2	GGCCTTCTTTGAGTTCGGTG
sgMYC_1	MYC	CTTCGGGGAGACAACGACGG
sgMYC_2	MYC	AGAGTGCATCGACCCCTCGG
sgMYC_3	MYC	CTGCGGGGAGGACTCCGTCG
sgPTEN_1	PTEN	AGAGCGTGCAGATAATGACA
sgPTEN_2	PTEN	CCAATTCAGGACCCACACGA
sgPTEN_3	PTEN	AGCTGGCAGACCACAAACTG
sgMET_1	MET	CCGATCGCACACATTTGTCG
sgMET_2	MET	AGCTGTGGCAGCGTCAACAG
sgMET_3	MET	CTCACTGATATCGAATGCAA
sgCTNNB1_1	CTNNB1	AAGGTTATGCAAGGTCCCAG
sgCTNNB1_2	CTNNB1	CAACTGGTAGTCCATAGTGA
sgCTNNB1_3	CTNNB1	ATGCAATGACTCGAGCTCAG
sgIRS2_1	IRS2	GCAGTAAGAGCCAATCGTCG
sgIRS2_2	IRS2	CAGGGTGTATTCATCCAGCG
sgIRS2_3	IRS2	ACAGTGGGTACATGCGCATG
sgFLT3_1	FLT3	AAAGCTGTTCATGTGAACCA
sgFLT3_2	FLT3	GGTGCTTTGCGATTCACAGG
sgFLT3_3	FLT3	GTAACCAAAGCTGATTGACT
sgMAP2K2_1	MAP2K2	AAGCACCAGATCATGCACCG

sgMAP2K2_2	MAP2K2	ACGGCGAGTTGCATTCGTGC
sgMAP2K2_3	MAP2K2	GGCCCATCCCCTACCAGCGA
sgSTAT5A_1	STAT5A	ACATTCTGTACAATGAACAG
sgSTAT5A_2	STAT5A	ATCAAGCGTGCTGACCGGCG
sgSTAT5A_3	STAT5A	CGTGCACATGAATCCCCCCC
sgHDAC3_1	HDAC3	TCATCAATGCCATCCCGCAG
sgHDAC3_2	HDAC3	ACCTGGAGCACAATGCACGT
sgHDAC3_3	HDAC3	TGGGTCAATGCCAGGCGATG
sgE2F1_1	E2F1	AAGGTCCTGACACGTCACGT
sgE2F1_2	E2F1	CACAGGTGTGAAATCCCCGG
sgE2F1_3	E2F1	CTGGATGCCCTCAAGGACGT
sgPRMT5_1	PRMT5	GGAGAAAAACCCAAATGCCG
sgPRMT5_2	PRMT5	GGTACTGAGAGTATTTGATG
sgPRMT5_3	PRMT5	GAAGATTCGCAGGAACTCCG
sgGATA3_1	GATA3	TCCAAGACGTCCATCCACCA
sgGATA3_2	GATA3	AGGTACCCTCCGACCCACCA
sgGATA3_3	GATA3	GGAGCTGTACTCGGGCACGT
sgPREX2_1	PREX2	TGTCTGTTCCAACATAAACG
sgPREX2_2	PREX2	GGAGCATGTGAGTCTGACAG
sgPREX2_3	PREX2	TGGGTCATGATCTCTGAACA
sgTSC1_1	TSC1	GTGACATCGGCTGAACGATG
sgTSC1_2	TSC1	TGACATGACCCAGTAACGAG
sgTSC1_3	TSC1	TCAAAAAGATGATCATACGG
sgIGF2BP1_1	IGF2BP1	ATTCGGTGCCCAAAAAACAA
sgIGF2BP1_2	IGF2BP1	CAAGATCATCTTACAAGCGG
sgIGF2BP1_3	IGF2BP1	CTCGTCCGGGCAGTCCACGA
sgPREX1_1	PREX1	GTGGATCTACCACACCCATG
sgPREX1_2	PREX1	GTGTAAACAATACTCCAAGG

sgPREX1_3	PREX1	ATGTTGCGGAACTCAAAGTG
sgFLT1_1	FLT1	ACAGCCACAGTCCGGCACGT
sgFLT1_2	FLT1	AGGTTGAGGGATACCATATG
sgFLT1_3	FLT1	CTTACCATATATATGCACTG
sgLRP8_1	LRP8	GCCGTCGCACACGAACACGG
sgLRP8_2	LRP8	CTCGTCGCATCTCCACACAG
sgLRP8_3	LRP8	GAGTGTCTGCACAACAATGG
sgCDK6_1	CDK6	GCCCGCGACTTGAAGAACGG
sgCDK6_2	CDK6	AACACTCCAGAGATCCACGG
sgCDK6_3	CDK6	TGGCTCACCTGACCACGTTG
sgFOS_1	FOS	GGAAAAACTAGAGTTCATCC
sgFOS_2	FOS	GTCGAGATGGCAGTGACCGT
sgFOS_3	FOS	GTAGTAAGAGAGGCTATCCC
sgCDK4_1	CDK4	CCAGATGGCACTTACACCCG
sgCDK4_2	CDK4	AGTGTGAGAGTCCCCAATGG
sgCDK4_3	CDK4	GTCCACATATGCAACACCTG
sgMEN1_1	MEN1	GAACGTTGGTAGGGATGACG
sgMEN1_2	MEN1	CCAGGCATGATCCTCAGACA
sgMEN1_3	MEN1	CCAGACAGTCAATGCCGGTG
sgGJC1_1	GJC1	GAACACCCAGAAGCGTACAT
sgGJC1_2	GJC1	CATCTTCCCGAATCCGTCGT
sgGJC1_3	GJC1	AATGCGCTGGAAACAACACC
sgPREX1_1	PREX1	GTGGATCTACCACACCCATG
sgPREX1_2	PREX1	GTGTAAACAATACTCCAAGG
sgPREX1_3	PREX1	ATGTTGCGGAACTCAAAGTG
sgFLT1_1	FLT1	ACAGCCACAGTCCGGCACGT
sgFLT1_2	FLT1	AGGTTGAGGGATACCATATG
sgFLT1_3	FLT1	CTTACCATATATATGCACTG

sgLRP8_1	LRP8	GCCGTCGCACACGAACACGG
sgLRP8_2	LRP8	CTCGTCGCATCTCCACACAG
sgLRP8_3	LRP8	GAGTGTCTGCACAACAATGG
sgCDK6_1	CDK6	GCCCGCGACTTGAAGAACGG
sgCDK6_2	CDK6	AACACTCCAGAGATCCACGG
sgCDK6_3	CDK6	TGGCTCACCTGACCACGTTG
sgFOS_1	FOS	GGAAAAACTAGAGTTCATCC
sgFOS_2	FOS	GTCGAGATGGCAGTGACCGT
sgFOS_3	FOS	GTAGTAAGAGAGGCTATCCC
sgCDK4_1	CDK4	CCAGATGGCACTTACACCCG
sgCDK4_2	CDK4	AGTGTGAGAGTCCCCAATGG
sgCDK4_3	CDK4	GTCCACATATGCAACACCTG
sgMEN1_1	MEN1	GAACGTTGGTAGGGATGACG
sgMEN1_2	MEN1	CCAGGCATGATCCTCAGACA
sgMEN1_3	MEN1	CCAGACAGTCAATGCCGGTG
sgGJC1_1	GJC1	GAACACCCAGAAGCGTACAT
sgGJC1_2	GJC1	CATCTTCCCGAATCCGTCGT
sgGJC1_3	GJC1	AATGCGCTGGAAACAACACC

Table 2: ALL_CROP1 sgRNA spacers

shRNA_name	target_gene	spacer_sequence
sgTRIM8_1	TRIM8	GACAGTGGAGGTCCTAGACA
sgTRIM8_2	TRIM8	TCAGGTGGCCGATCTTAGTG
sgTRIM8_3	TRIM8	GGGACACTCGGTGTGCGACG
sgPOU3F1_1	POU3F1	CTTCTGCACTTCGCGGTACG
sgPOU3F1_2	POU3F1	GCCTGCGCGTACAGCCCGAG
sgPOU3F1_3	POU3F1	ATGTCCGTGTGCGTGTCCGT
sgRXRA_1	RXRA	CCTACGTGGAGGCAAACATG
sgRXRA_2	RXRA	AGGACTGCCTGATTGACAAG

sgRXRA_3	RXRA	AGGAAGCCATGTTTCCTGAG
sgNFATC2_1	NFATC2	ATGTAAAGTTCTGCCCCGTG
sgNFATC2_2	NFATC2	ACATTGGAAGAAAGAACACG
sgNFATC2_3	NFATC2	GCCGCAGCCCTCATCTCACG
sgRUNX2_1	RUNX2	GAATGCGCCCTAAATCACTG
sgRUNX2_2	RUNX2	TGAAATGCGCCTAGGCACAT
sgRUNX2_3	RUNX2	GACTGGTCATAGGACCACGG
sgTCF4_1	TCF4	GACTACAATAGGGACTCGCC
sgTCF4_2	TCF4	CGATGGAAAGTGGACATCGG
sgTCF4_3	TCF4	CTAGCAATAATCCCCGAAGG
sgSTAG2_1	STAG2	TGGAGATTATCCACTTACCA
sgSTAG2_2	STAG2	ATACCTTGTGGATAGCATGT
sgSTAG2_3	STAG2	AATACTAACCTTGAACCGAC
sgPAX7_1	PAX7	CGAACTTGATTCTGAGCACG
sgPAX7_2	PAX7	AGCGCAAGCAGCGACGCAGT
sgPAX7_3	PAX7	GCTTACGCCAACGGGCGCGG
sgPOU3F2_1	POU3F2	CAAACTGGGATTTACCCAAG
sgPOU3F2_2	POU3F2	TCCATGGGAGCGTCCAACGG
sgPOU3F2_3	POU3F2	CACGCCGCTAACCACCACCC
sgEWSR1_1	EWSR1	TGGGTCTTCATAGGACACTG
sgEWSR1_2	EWSR1	ATGAGTGGCCCTGATAACCG
sgEWSR1_3	EWSR1	GGTCTGCCCATAGGTTGCAG
sgHAND2_1	HAND2	TGGACCACTCCCATTACGGG
sgHAND2_2	HAND2	CACCCGGTGGTGCACCACGA
sgHAND2_3	HAND2	TGGTGTCGGCGGGTACGTTG
sgRUNX3_1	RUNX3	CCCCAGGATGCATTATCCCG
sgRUNX3_2	RUNX3	CACTGCGGCCCACGAAGCGA
sgRUNX3_3	RUNX3	CCGTGCCGTACCTTGGATTG

sgFLI1_1	FLI1	CCAACATGACCACCAACGAG
sgFLI1_2	FLI1	ACTCAATCGTGAGGATTGGT
sgFLI1_3	FLI1	ACTGTGTAAAATGAACAAGG
sgRREB1_1	RREB1	CCTGGTGCGAAACAAACCTG
sgRREB1_2	RREB1	CTTCTCGATATCCTTGCGGG
sgRREB1_3	RREB1	CAGCACAACACAGACACTGG
sgNFIB_1	NFIB	CTATCATGACATGAACTCGG
sgNFIB_2	NFIB	GATTGGATAAGACACAGCAC
sgNFIB_3	NFIB	TGGAAAGTACCGATGGAGAG
sgKLF15_1	KLF15	ACGTCATCAGGATCACCCAA
sgKLF15_2	KLF15	GGCAGGTTCAAGTTGGAGGA
sgKLF15_3	KLF15	GGCAACCTTGACATTCTCTG
sgETV6_1	ETV6	TGGTGCACATTATCCACGGA
sgETV6_2	ETV6	AATGGTGAAAAAAGAATCCG
sgETV6_3	ETV6	TGGAACATGAAGTGGCGTCG
sgSOX12_1	SOX12	TGTGGTCGCAGCACGAACGG
sgSOX12_2	SOX12	GCGCTCCGCTTGTCCCCGAG
sgSOX12_3	SOX12	GGAGAAGATCCCGTTCGTGC
sgSTAG1_1	STAG1	TGGACACCCTCAACAGAATG
sgSTAG1_2	STAG1	AGATCGATTCAATCATTCTG
sgSTAG1_3	STAG1	TGGCTGGACTCTTCATGACA
sgNFIX_1	NFIX	TCAGATAGTTCAAACCAGCA
sgNFIX_2	NFIX	CATCCGGCCCGAGTTCCGCG
sgNFIX_3	NFIX	CAGCTACTACAACATCAACC
sgMDM2_1	MDM2	GAGAACATTACCGGATTCGA
sgMDM2_2	MDM2	TACCATGATCTACAGGAACT
sgMDM2_3	MDM2	AGACACTTATACTATGAAAG
sgZBTB7B_1	ZBTB7B	TCCGGATGGTGAGGTCACAT

sgZBTB7B_2	ZBTB7B	AGCAAACCACCTAGTCCCTG
sgZBTB7B_3	ZBTB7B	TGTATAGGCAAATTCAAGGA
sgZBTB16_1	ZBTB16	CATCTCGAAGCATTCCAGCG
sgZBTB16_2	ZBTB16	GGTCGAGCTTCCTGATAACG
sgZBTB16_3	ZBTB16	GGAGCTACACTATGGGCGAG
sgNKX2-2_1	NKX2-2	AATGACAAGGAGACCCCGGG
sgNKX2-2_2	NKX2-2	GCTACTTACGGGAGTACTGA
sgNKX2-2_3	NKX2-2	CCTGCCGGACACCAACGATG
NTC_1	CTRL00018	GGTGATCCTAGTCGACTGGC
NTC_2	CTRL00022	GCTATTGTCGGCTGGATGGA
NTC_3	CTRL00080	GCGCATCAATATGCCCGCAC
NTC_4	CTRL00087	GAGGTCCCGCCTCCGCCCAA
NTC_5	CTRL00096	GGCTAACGCACGCCGAGGTG
NTC_6	CTRL00196	GGGCCTTAATAGTTCAACGC
NTC_7	CTRL00275	GCAGGACGATAGTAACGTCC
NTC_8	CTRL00320	GAGAACGTGATAAGACTCGG
NTC_9	CTRL00405	GGACTGACGTATACGCTTGC
NTC_10	CTRL00484	GTGAATACCCGTGACGGACA
NTC_11	CTRL00493	GGCAATCGGCACGGCAAGGT
NTC_12	CTRL00545	GGGCTGTTCTCACTCGTAGC
NTC_13	CTRL00616	GGTTAAAGAGATTGTACGCC
NTC_14	CTRL00637	GGGTTGCACGTAGGATATTC
NTC_15	CTRL00640	GCGTATTGTATCCGCCACCG
NTC_16	CTRL00673	GAGTAAAGCCGATTATATCG
sgMVD_1	MVD	AAGCTGACAGGCAGTACCGT
sgMVD_2	MVD	CTGGTGCAGAGTGACGCTCA
sgTUBB_1	TUBB	GCTGACCACACCAACCTACG
sgTUBB_2	TUBB	CCCCACCGGCACCTACCACG

sgHCFC1_1	HCFC1	GTGGAAGTGTACCAACACGC
sgHCFC1_2	HCFC1	GGTGCCTTTCACAACCAACG
sgE2F3_1	E2F3	GGACCTCAAACTGTTAACCG
sgE2F3_2	E2F3	AGATGGTTTAAAAACCCCCA
sgE2F3_3	E2F3	AATCTCCCTCAGAAAAAACG
sgIRS1_1	IRS1	CCGAAGCACTAGATCGCCGT
sgIRS1_2	IRS1	TCTTGCTGGTCAGGCAAAGG
sgIRS1_3	IRS1	GCCCCCCCCGACGCTCCAAG
sgMCL1_1	MCL1	AGGCGCTGGAGACCTTACGA
sgMCL1_2	MCL1	GTAATAACACCAGTACGGAC
sgMCL1_3	MCL1	AGTCGCTGGAGATTATCTCT
sgBAP1_1	BAP1	CACGGACGTATCATCCACCA
sgBAP1_2	BAP1	GAACCGTCAGACAGTACTAG
sgBAP1_3	BAP1	TCTACCCCATTGACCATGGT
sgRELB_1	RELB	ATTGAGCGGAAGATTCAACT
sgRELB_2	RELB	GCCTCATATCGGGACCAGCA
sgRELB_3	RELB	CGGTGCAGTCTTTCCCCACG
sgNXT1_1	NXT1	ACTACACCACCATGGATAAG
sgNXT1_2	NXT1	AGCATTGCCATTCCAGACCA
sgNXT1_3	NXT1	ACAGATGACAACAAGGACCG
sgCABIN1_1	CABIN1	GTAATCGTGGTCAATCGGAG
sgCABIN1_2	CABIN1	GTAGTGCAGCAAGTAAACGG
sgCABIN1_3	CABIN1	CTGGAGAACCTAACCAACGG
sgRELA_1	RELA	TCAATGGCTACACAGGACCA
sgRELA_2	RELA	GCTTCCGCTACAAGTGCGAG
sgRELA_3	RELA	GGAAGATCTCATCCCCACCG
sgIGF1R_1	IGF1R	GGAGAACGACCATATCCGTG
sgIGF1R_2	IGF1R	TTCCGAAATTTACCGCATGG

sgIGF1R_3	IGF1R	GGTACAATGTGAAAGGCCGA
sgMAX_1	MAX	TATTCCAGGAAGAGCAACCG
sgMAX_2	MAX	GAAGAGCATTCTGCCGCTTG
sgMAX_3	MAX	AATATATCCAGTATATGCGA
sgHDAC2_1	HDAC2	GATGTATCAACCTAGTGCTG
sgHDAC2_2	HDAC2	TACAACAGATCGTGTAATGA
sgHDAC2_3	HDAC2	CCTCCTCCAAGCATCAGTAA
sgJUN_1	JUN	GGCGGCGCAGCCGGTCAACG
sgJUN_2	JUN	GCTCTCGGACGGGAGGAACG
sgJUN_3	JUN	TGAACCTGGCCGACCCAGTG
sgSATB1_1	SATB1	TAGGTGTTGATACGAGCCCA
sgSATB1_2	SATB1	TATTCATAGATCTACTGACA
sgSATB1_3	SATB1	CATTGAATATGATTGCAAGG
sgNGFR_1	NGFR	ACAGGCCTGTACACACACAG
sgNGFR_2	NGFR	CGACGGCACGTATTCCGACG
sgNGFR_3	NGFR	GTGTGGACCGTGTAATCCAA
sgBCL2_1	BCL2	TGTCGCAGAGGGGCTACGAG
sgBCL2_2	BCL2	CTGACGCCCTTCACCGCGCG
sgBCL2_3	BCL2	GGCCTTCTTTGAGTTCGGTG
sgMYC_1	MYC	CTTCGGGGAGACAACGACGG
sgMYC_2	MYC	AGAGTGCATCGACCCCTCGG
sgMYC_3	MYC	CTGCGGGGAGGACTCCGTCG
sgPTEN_1	PTEN	AGAGCGTGCAGATAATGACA
sgPTEN_2	PTEN	CCAATTCAGGACCCACACGA
sgPTEN_3	PTEN	AGCTGGCAGACCACAAACTG
sgMET_1	MET	CCGATCGCACACATTTGTCG
sgMET_2	MET	AGCTGTGGCAGCGTCAACAG
sgMET_3	MET	CTCACTGATATCGAATGCAA

sgCTNNB1_1	CTNNB1	AAGGTTATGCAAGGTCCCAG
sgCTNNB1_2	CTNNB1	CAACTGGTAGTCCATAGTGA
sgCTNNB1_3	CTNNB1	ATGCAATGACTCGAGCTCAG
sgIRS2_1	IRS2	GCAGTAAGAGCCAATCGTCG
sgIRS2_2	IRS2	CAGGGTGTATTCATCCAGCG
sgIRS2_3	IRS2	ACAGTGGGTACATGCGCATG
sgFLT3_1	FLT3	AAAGCTGTTCATGTGAACCA
sgFLT3_2	FLT3	GGTGCTTTGCGATTCACAGG
sgFLT3_3	FLT3	GTAACCAAAGCTGATTGACT
sgMAP2K2_1	MAP2K2	AAGCACCAGATCATGCACCG
sgMAP2K2_2	MAP2K2	ACGGCGAGTTGCATTCGTGC
sgMAP2K2_3	MAP2K2	GGCCCATCCCCTACCAGCGA
sgSTAT5A_1	STAT5A	ACATTCTGTACAATGAACAG
sgSTAT5A_2	STAT5A	ATCAAGCGTGCTGACCGGCG
sgSTAT5A_3	STAT5A	CGTGCACATGAATCCCCCCC
sgHDAC3_1	HDAC3	TCATCAATGCCATCCCGCAG
sgHDAC3_2	HDAC3	ACCTGGAGCACAATGCACGT
sgHDAC3_3	HDAC3	TGGGTCAATGCCAGGCGATG
sgE2F1_1	E2F1	AAGGTCCTGACACGTCACGT
sgE2F1_2	E2F1	CACAGGTGTGAAATCCCCGG
sgE2F1_3	E2F1	CTGGATGCCCTCAAGGACGT
sgPRMT5_1	PRMT5	GGAGAAAAACCCAAATGCCG
sgPRMT5_2	PRMT5	GGTACTGAGAGTATTTGATG
sgPRMT5_3	PRMT5	GAAGATTCGCAGGAACTCCG
sgGATA3_1	GATA3	TCCAAGACGTCCATCCACCA
sgGATA3_2	GATA3	AGGTACCCTCCGACCCACCA
sgGATA3_3	GATA3	GGAGCTGTACTCGGGCACGT
sgPREX2_1	PREX2	TGTCTGTTCCAACATAAACG

sgPREX2_2	PREX2	GGAGCATGTGAGTCTGACAG
sgPREX2_3	PREX2	TGGGTCATGATCTCTGAACA
sgTSC1_1	TSC1	GTGACATCGGCTGAACGATG
sgTSC1_2	TSC1	TGACATGACCCAGTAACGAG
sgTSC1_3	TSC1	TCAAAAAGATGATCATACGG
sgIGF2BP1_1	IGF2BP1	ATTCGGTGCCCAAAAAACAA
sgIGF2BP1_2	IGF2BP1	CAAGATCATCTTACAAGCGG
sgIGF2BP1_3	IGF2BP1	CTCGTCCGGGCAGTCCACGA
sgPREX1_1	PREX1	GTGGATCTACCACACCCATG
sgPREX1_2	PREX1	GTGTAAACAATACTCCAAGG
sgPREX1_3	PREX1	ATGTTGCGGAACTCAAAGTG
sgFLT1_1	FLT1	ACAGCCACAGTCCGGCACGT
sgFLT1_2	FLT1	AGGTTGAGGGATACCATATG
sgFLT1_3	FLT1	CTTACCATATATATGCACTG
sgLRP8_1	LRP8	GCCGTCGCACACGAACACGG
sgLRP8_2	LRP8	CTCGTCGCATCTCCACACAG
sgLRP8_3	LRP8	GAGTGTCTGCACAACAATGG
sgCDK6_1	CDK6	GCCCGCGACTTGAAGAACGG
sgCDK6_2	CDK6	AACACTCCAGAGATCCACGG
sgCDK6_3	CDK6	TGGCTCACCTGACCACGTTG
sgFOS_1	FOS	GGAAAAACTAGAGTTCATCC
sgFOS_2	FOS	GTCGAGATGGCAGTGACCGT
sgFOS_3	FOS	GTAGTAAGAGAGGCTATCCC
sgCDK4_1	CDK4	CCAGATGGCACTTACACCCG
sgCDK4_2	CDK4	AGTGTGAGAGTCCCCAATGG
sgCDK4_3	CDK4	GTCCACATATGCAACACCTG
sgMEN1_1	MEN1	GAACGTTGGTAGGGATGACG
sgMEN1_2	MEN1	CCAGGCATGATCCTCAGACA

sgMEN1_3	MEN1	CCAGACAGTCAATGCCGGTG
sgGJC1_1	GJC1	GAACACCCAGAAGCGTACAT
sgGJC1_2	GJC1	CATCTTCCCGAATCCGTCGT
sgGJC1_3	GJC1	AATGCGCTGGAAACAACACC

Table 3: ES_CROP1 sgRNA list

shRNA_name	target_gene	spacer_sequence
sgKRT19_1	KRT19	CAGGAAATCAGTACGCTGAG
sgKRT19_2	KRT19	CGCGCGCCCAGCATTCACGG
sgKRT19_3	KRT19	TAGTGGCTGTAGTCGCGGGA
sgIGF2_1	IGF2	AGTTCTTCCAATATGACACC
sgIGF2_2	IGF2	CTCACCGGAAGCACGGTCGG
sgIGF2_3	IGF2	CAGTGAGACCCTGTGCGGCG
sgSOX9_1	SOX9	TTCACCGACTTCCTCCGCCG
sgSOX9_2	SOX9	ACCATGTCCGAGGACTCCGC
sgSOX9_3	SOX9	ACTCACCCGAGTGCTCGCCG
sgBLCAP_1	BLCAP	GCTATAGCAGATAAGGAACA
sgBLCAP_2	BLCAP	ACCAAGGCACAAATTGTGCA
sgBLCAP_3	BLCAP	TGCAAGGCTTCCGTTCCAGG
sgDLK1_1	DLK1	CCCCCAAAATGGATTCTGCG
sgDLK1_2	DLK1	AGGCAATTTCTGCGAGATCG
sgDLK1_3	DLK1	CAGACACTCGTAGCTCACCT
sgCDK9_1	CDK9	CCAGAGTGTCACCACACGGT
sgCDK9_2	CDK9	TCTCCCGCAAGGCTGTAATG
sgCDK9_3	CDK9	GCGGTTATAGGGGGAAGCTG
sgBIRC5_1	BIRC5	ATGCGGTGGTCCTTGAGAAA
sgBIRC5_2	BIRC5	GAACATAAAAAGCATTCGTC
sgBIRC5_3	BIRC5	GCTGCGCCTGCACCCCGGAG
sgNFE2L2_1	NFE2L2	CTTCCACTTCAGAATCACTG

sgNFE2L2_2	NFE2L2	CACATCCAGTCAGAAACCAG
sgNFE2L2_3	NFE2L2	CATACCGTCTAAATCAACAG
sgUHRF1_1	UHRF1	TGCTCGGGACACGAACATGG
sgUHRF1_2	UHRF1	CTACAACCCCGACAACCCCA
sgUHRF1_3	UHRF1	ACACCCGACTCGCTGACCTG
sgSALL4_1	SALL4	GTGACCTTGCAGGCACTACG
sgSALL4_2	SALL4	GCCACCGAACATCTCCGCGG
sgSALL4_3	SALL4	TCTGACAGCTTAAGACTCGG
sgTERT_1	TERT	CACACGCTAGTGGACCCCGA
sgTERT_2	TERT	GTGACACCACACAGAAACCA
sgTERT_3	TERT	CTCACGCAGACGGTGCTCTG
sgEPCAM_1	EPCAM	GTGCACCAACTGAAGTACAC
sgEPCAM_2	EPCAM	AGTGTTCACACACCAGCACA
sgEPCAM_3	EPCAM	GGGCCCTCCAGAACAATGAT
sgDKK1_1	DKK1	TGGTAATGATCATAGCACCT
sgDKK1_2	DKK1	TACTGCGCTAGTCCCACCCG
sgDKK1_3	DKK1	TACCCGGGTCTTTGTCGCGA
sgAFP_1	AFP	ACATTGACCACGTTCCAGCG
sgAFP_2	AFP	AGTGGCTTCTTGAACAAACT
sgAFP_3	AFP	ATTGTAGGTGCATACAGGAA
sgGPC3_1	GPC3	CTGTGGCGGTTACTGCAATG
sgGPC3_2	GPC3	GACATCAATGAGTGCCTCCG
sgGPC3_3	GPC3	CATGTACAGAATCTATGACA
sgBEX1_1	BEX1	GTGTGCCTAGAGGAAATCGT
sgBEX1_2	BEX1	AGCAGTAAACAGTCTCAGCA
sgBEX1_3	BEX1	ATTCACCAGCATCCAAAGGG
sgMEG3_1	MEG3	GCACTAGGAGCACGGTTTCC
sgMEG3_2	MEG3	TTTGTATGTTGGTGGGATCC

sgMEG3_3	MEG3	GTTTAAGTCTTTAGGTAAGA
sgCHKA_1	CHKA	CATAACGCTCTCCAGAACCA
sgCHKA_2	CHKA	CCGGGATGAACTGCTCCAGT
sgCHKA_3	CHKA	CTTGGACTCGAGGTCGCTGG
NTC_1	CTRL00018	GGTGATCCTAGTCGACTGGC
NTC_2	CTRL00022	GCTATTGTCGGCTGGATGGA
NTC_3	CTRL00080	GCGCATCAATATGCCCGCAC
NTC_4	CTRL00087	GAGGTCCCGCCTCCGCCCAA
NTC_5	CTRL00096	GGCTAACGCACGCCGAGGTG
NTC_6	CTRL00196	GGGCCTTAATAGTTCAACGC
NTC_7	CTRL00275	GCAGGACGATAGTAACGTCC
NTC_8	CTRL00320	GAGAACGTGATAAGACTCGG
NTC_9	CTRL00405	GGACTGACGTATACGCTTGC
NTC_10	CTRL00484	GTGAATACCCGTGACGGACA
NTC_11	CTRL00493	GGCAATCGGCACGGCAAGGT
NTC_12	CTRL00545	GGGCTGTTCTCACTCGTAGC
NTC_13	CTRL00616	GGTTAAAGAGATTGTACGCC
NTC_14	CTRL00637	GGGTTGCACGTAGGATATTC
NTC_15	CTRL00640	GCGTATTGTATCCGCCACCG
NTC_16	CTRL00673	GAGTAAAGCCGATTATATCG
sgMVD_1	MVD	AAGCTGACAGGCAGTACCGT
sgMVD_2	MVD	CTGGTGCAGAGTGACGCTCA
sgTUBB_1	TUBB	GCTGACCACACCAACCTACG
sgTUBB_2	TUBB	CCCCACCGGCACCTACCACG
sgHCFC1_1	HCFC1	GTGGAAGTGTACCAACACGC
sgHCFC1_2	HCFC1	GGTGCCTTTCACAACCAACG
sgE2F3_1	E2F3	GGACCTCAAACTGTTAACCG
sgE2F3_2	E2F3	AGATGGTTTAAAAACCCCCA

sgE2F3_3	E2F3	AATCTCCCTCAGAAAAAACG
sgIRS1_1	IRS1	CCGAAGCACTAGATCGCCGT
sgIRS1_2	IRS1	TCTTGCTGGTCAGGCAAAGG
sgIRS1_3	IRS1	GCCCCCCCCGACGCTCCAAG
sgMCL1_1	MCL1	AGGCGCTGGAGACCTTACGA
sgMCL1_2	MCL1	GTAATAACACCAGTACGGAC
sgMCL1_3	MCL1	AGTCGCTGGAGATTATCTCT
sgBAP1_1	BAP1	CACGGACGTATCATCCACCA
sgBAP1_2	BAP1	GAACCGTCAGACAGTACTAG
sgBAP1_3	BAP1	TCTACCCCATTGACCATGGT
sgRELB_1	RELB	ATTGAGCGGAAGATTCAACT
sgRELB_2	RELB	GCCTCATATCGGGACCAGCA
sgRELB_3	RELB	CGGTGCAGTCTTTCCCCACG
sgNXT1_1	NXT1	ACTACACCACCATGGATAAG
sgNXT1_2	NXT1	AGCATTGCCATTCCAGACCA
sgNXT1_3	NXT1	ACAGATGACAACAAGGACCG
sgCABIN1_1	CABIN1	GTAATCGTGGTCAATCGGAG
sgCABIN1_2	CABIN1	GTAGTGCAGCAAGTAAACGG
sgCABIN1_3	CABIN1	CTGGAGAACCTAACCAACGG
sgRELA_1	RELA	TCAATGGCTACACAGGACCA
sgRELA_2	RELA	GCTTCCGCTACAAGTGCGAG
sgRELA_3	RELA	GGAAGATCTCATCCCCACCG
sgIGF1R_1	IGF1R	GGAGAACGACCATATCCGTG
sgIGF1R_2	IGF1R	TTCCGAAATTTACCGCATGG
sgIGF1R_3	IGF1R	GGTACAATGTGAAAGGCCGA
sgMAX_1	MAX	TATTCCAGGAAGAGCAACCG
sgMAX_2	MAX	GAAGAGCATTCTGCCGCTTG
sgMAX_3	MAX	AATATATCCAGTATATGCGA

sgHDAC2_1	HDAC2	GATGTATCAACCTAGTGCTG
sgHDAC2_2	HDAC2	TACAACAGATCGTGTAATGA
sgHDAC2_3	HDAC2	CCTCCTCCAAGCATCAGTAA
sgJUN_1	JUN	GGCGGCGCAGCCGGTCAACG
sgJUN_2	JUN	GCTCTCGGACGGGAGGAACG
sgJUN_3	JUN	TGAACCTGGCCGACCCAGTG
sgSATB1_1	SATB1	TAGGTGTTGATACGAGCCCA
sgSATB1_2	SATB1	TATTCATAGATCTACTGACA
sgSATB1_3	SATB1	CATTGAATATGATTGCAAGG
sgNGFR_1	NGFR	ACAGGCCTGTACACACACAG
sgNGFR_2	NGFR	CGACGGCACGTATTCCGACG
sgNGFR_3	NGFR	GTGTGGACCGTGTAATCCAA
sgBCL2_1	BCL2	TGTCGCAGAGGGGCTACGAG
sgBCL2_2	BCL2	CTGACGCCCTTCACCGCGCG
sgBCL2_3	BCL2	GGCCTTCTTTGAGTTCGGTG
sgMYC_1	MYC	CTTCGGGGAGACAACGACGG
sgMYC_2	MYC	AGAGTGCATCGACCCCTCGG
sgMYC_3	MYC	CTGCGGGGAGGACTCCGTCG
sgPTEN_1	PTEN	AGAGCGTGCAGATAATGACA
sgPTEN_2	PTEN	CCAATTCAGGACCCACACGA
sgPTEN_3	PTEN	AGCTGGCAGACCACAAACTG
sgMET_1	MET	CCGATCGCACACATTTGTCG
sgMET_2	MET	AGCTGTGGCAGCGTCAACAG
sgMET_3	MET	CTCACTGATATCGAATGCAA
sgCTNNB1_1	CTNNB1	AAGGTTATGCAAGGTCCCAG
sgCTNNB1_2	CTNNB1	CAACTGGTAGTCCATAGTGA
sgCTNNB1_3	CTNNB1	ATGCAATGACTCGAGCTCAG
sgIRS2_1	IRS2	GCAGTAAGAGCCAATCGTCG

sgIRS2_2	IRS2	CAGGGTGTATTCATCCAGCG
sgIRS2_3	IRS2	ACAGTGGGTACATGCGCATG
sgFLT3_1	FLT3	AAAGCTGTTCATGTGAACCA
sgFLT3_2	FLT3	GGTGCTTTGCGATTCACAGG
sgFLT3_3	FLT3	GTAACCAAAGCTGATTGACT
sgMAP2K2_1	MAP2K2	AAGCACCAGATCATGCACCG
sgMAP2K2_2	MAP2K2	ACGGCGAGTTGCATTCGTGC
sgMAP2K2_3	MAP2K2	GGCCCATCCCCTACCAGCGA
sgSTAT5A_1	STAT5A	ACATTCTGTACAATGAACAG
sgSTAT5A_2	STAT5A	ATCAAGCGTGCTGACCGGCG
sgSTAT5A_3	STAT5A	CGTGCACATGAATCCCCCCC
sgHDAC3_1	HDAC3	TCATCAATGCCATCCCGCAG
sgHDAC3_2	HDAC3	ACCTGGAGCACAATGCACGT
sgHDAC3_3	HDAC3	TGGGTCAATGCCAGGCGATG
sgE2F1_1	E2F1	AAGGTCCTGACACGTCACGT
sgE2F1_2	E2F1	CACAGGTGTGAAATCCCCGG
sgE2F1_3	E2F1	CTGGATGCCCTCAAGGACGT
sgPRMT5_1	PRMT5	GGAGAAAAACCCAAATGCCG
sgPRMT5_2	PRMT5	GGTACTGAGAGTATTTGATG
sgPRMT5_3	PRMT5	GAAGATTCGCAGGAACTCCG
sgGATA3_1	GATA3	TCCAAGACGTCCATCCACCA
sgGATA3_2	GATA3	AGGTACCCTCCGACCCACCA
sgGATA3_3	GATA3	GGAGCTGTACTCGGGCACGT
sgPREX2_1	PREX2	TGTCTGTTCCAACATAAACG
sgPREX2_2	PREX2	GGAGCATGTGAGTCTGACAG
sgPREX2_3	PREX2	TGGGTCATGATCTCTGAACA
sgTSC1_1	TSC1	GTGACATCGGCTGAACGATG
sgTSC1_2	TSC1	TGACATGACCCAGTAACGAG

sgTSC1_3	TSC1	TCAAAAAGATGATCATACGG
sgIGF2BP1_1	IGF2BP1	ATTCGGTGCCCAAAAAACAA
sgIGF2BP1_2	IGF2BP1	CAAGATCATCTTACAAGCGG
sgIGF2BP1_3	IGF2BP1	CTCGTCCGGGCAGTCCACGA
sgPREX1_1	PREX1	GTGGATCTACCACACCCATG
sgPREX1_2	PREX1	GTGTAAACAATACTCCAAGG
sgPREX1_3	PREX1	ATGTTGCGGAACTCAAAGTG
sgFLT1_1	FLT1	ACAGCCACAGTCCGGCACGT
sgFLT1_2	FLT1	AGGTTGAGGGATACCATATG
sgFLT1_3	FLT1	CTTACCATATATATGCACTG
sgLRP8_1	LRP8	GCCGTCGCACACGAACACGG
sgLRP8_2	LRP8	CTCGTCGCATCTCCACACAG
sgLRP8_3	LRP8	GAGTGTCTGCACAACAATGG
sgCDK6_1	CDK6	GCCCGCGACTTGAAGAACGG
sgCDK6_2	CDK6	AACACTCCAGAGATCCACGG
sgCDK6_3	CDK6	TGGCTCACCTGACCACGTTG
sgFOS_1	FOS	GGAAAAACTAGAGTTCATCC
sgFOS_2	FOS	GTCGAGATGGCAGTGACCGT
sgFOS_3	FOS	GTAGTAAGAGAGGCTATCCC
sgCDK4_1	CDK4	CCAGATGGCACTTACACCCG
sgCDK4_2	CDK4	AGTGTGAGAGTCCCCAATGG
sgCDK4_3	CDK4	GTCCACATATGCAACACCTG
sgMEN1_1	MEN1	GAACGTTGGTAGGGATGACG
sgMEN1_2	MEN1	CCAGGCATGATCCTCAGACA
sgMEN1_3	MEN1	CCAGACAGTCAATGCCGGTG
sgGJC1_1	GJC1	GAACACCCAGAAGCGTACAT
sgGJC1_2	GJC1	CATCTTCCCGAATCCGTCGT
sgGJC1_3	GJC1	AATGCGCTGGAAACAACACC

Table 4: HB_CROP1 sgRNA list