

Functional translation of malignant peripheral nerve
sheath tumour driver gene variants using CRISPR screens

Paul Metz



This thesis is submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Hertford College

Sir William Dunn School of Pathology

University of Oxford

Trinity 2024

*"Little drops of water,
Little grains of sand,
Make the mighty ocean
And the pleasant land."*

Julia Abigail Fletcher Carney, Little Things

Declaration

I declare that this thesis has never been submitted, either in the same or in a different form, to this or another university for a degree or other qualification. The entire experimental work and data analysis in this thesis was performed by me, with the exception of the specific work listed below. The work was conducted during my time as a DPhil student at the Sir William Dunn School of Pathology, under the supervision of Professor Bass Hassan. My DPhil studies were funded by the CRUK Oxford Centre.

The work presented in this thesis includes the following contributions.

- whole-genome sequencing analysis, including somatic variant and copy number variation calling of cell lines 26T, 90-8, ST-8814, s462TY and iHSC by Dr Mirvat Surakhya and Dr Stuart Brown.
- Design and synthesis of UNC7700, an EED-targeting PROTAC, was performed by Lindsey I. James, at North Carolina Chapel Hill [1].

The approximated word count is around 48,000.

Paul Metz, 12-10-2024

Acknowledgements

No achievement is ever accomplished by one person alone. During my years at Oxford, I have been fortunate to be surrounded by many incredible people who have made my DPhil a rewarding, interesting, and enjoyable journey, equipping me with the tools to face new challenges in life.

First and foremost, I would like to express my gratitude to my supervisor, Professor Bass Hassan, for giving me the opportunity to join his lab, guiding my project, and contributing to my development as a researcher. I feel privileged to have been able to present my work at scientific conferences, contribute to publications, and collaborate with other research groups. Bass has always been kind and understanding, particularly during challenging moments and when I wanted to pursue new directions in my work. I would also like to thank my co-supervisor, Joey Riepsaame, for his insights into biotechnology, DNA-editing, and CRISPR techniques. Both Bass and Joey's sense of humour during our meetings has made our work together all the more enjoyable.

I would also like to extend my thanks to Professor Lindsey I. James for providing the UNC7700 compound and to John Doench for allowing us to use his newly designed CRISPR libraries, Humagne C and D.

During my time at the Dunn School, I've been supported by many people, including those in my lab, especially Mirvat Surakhy, Richard Wallbank, Georgina Fisher, Zuzana Shah, Qinan Qian, Rachael Wilkinson, Tanya Ren, Stuart Brown, Sofia Vasconcelos, Isabel Karia, and Hannah Mallon. Mirvat and Richard, in particular, have been instrumental in teaching me lab techniques that were vital to my research. I am also grateful to Tanya, Sofia, and Rachael for helping out with my cell culture when I was away, and to Qinan, my "baby-sister", for helping me to succeed in the exciting field of Buzzing science.

There are also many wonderful people outside my lab, both in and outside the Dunn School, to whom I owe my thanks. This includes Joanna Ferreira, Nik Sergejevs, Asmaa Hachim, Yoel Klug, Zhe Ji, (Wingyan) Skyla Siu, Loges Krshnan, (the entire Carvalho lab), Simon Martina-Perez, Andrew Bassett, John Doench, Joshua Dempster, Adam Grieve, Sofia van Leur, and Eniola Awodiya. In particular, I want to thank Joana and Nik for teaching me the superior way of doing Westerns, and Asmaa for showing me how to improve FACS readouts. I've also been lucky to receive invaluable advice, suggestions, and free reagents from colleagues throughout the Dunn School, and have felt at home wherever I went in the department.

I must also thank the amazing friends I have made during my time in Oxford, whom I miss dearly while writing this from The Netherlands. You all have been essential to my DPhil journey and life in Oxford. Thank you Asmaa, Peny, Joana, Sophie van Leur, Kira, Qinan, Teresa, Sophie

Hughes, Alexandra, Rachel, Josephine, Nina, Rebecca, Dorian, Jody, Laurel, Michelle, Nik, Banashree, Simon, Andrew, Nikki, Ty, Seth, Dominique, Johanna Friege, Johanna Hettinga, Sofia Vasconcelos, Laura, Isabel, Juli, Ana Gali, Anne Sofie, Georgie, Phillipe, Lachlan, and many more.

Thank you, Alex, for everything: your support, positivity, and care have meant a lot to me, especially in the final stretch. A special thanks as well to Alex's family, Kevin, lovely Sara, Katie, Teddy, and Annie, for making me feel at home immediately.

Of course, I must also thank you all, my friends, especially mijn meissies" Lia and PJ, my dear friends Nicole and Darragh (for basically parenting me during August and September), Nina, Kim and Laura. Also, many thanks to all of my friends that visited me in Oxford and thus providing entertainment and support.

My final thanks goes to my family: Mom, Dad, Emma, Joep, and Hoetie. Your unwavering support during my life has been invaluable, especially you, Mom and Dad. You have always been there when I needed you. You have played a huge role in bringing me to where I am today.

Table of Contents

Declaration	3
Acknowledgements	4
Abstract.....	10
List of Figures	11
List of Tables.....	15
List of abbreviations	17
1 Chapter 1: Introduction.....	24
1.1 Challenges in cancer treatment.....	25
1.1.1 The cancer evolutionary cycle.....	26
1.1.2 Cancer cells do not rely on one mechanism for growth	27
1.1.3 Plasticity and intra-tumour heterogeneity in cancer evolution.....	28
1.1.4 Clonal hierarchy	29
1.2 Overcoming the cancer evolutionary cycle	31
1.2.1 Immune checkpoint inhibition.....	31
1.2.2 T-cell therapy in cancer.....	32
1.2.3 Cancer vaccines	33
1.2.4 Synthetic lethality	34
1.3 Malignant Peripheral Nerve Sheath Tumours.....	36
1.3.1 MPNST-driver gene variants.....	37
1.3.2 Polycomb repressive complex 2 loss of function in MPNST	37
1.3.3 Current therapeutic landscape and emerging targets in MPNST.....	41
1.3.4 Synthetic lethality in MPNST	43
1.4 CRISPR	45
1.4.1 Genome-wide CRISPR knockout screens	46
1.4.2 Genome-wide CRISPR screens in MPNST cell line models.....	49
1.5 Aim of the study.....	49
2 Chapter 2: Materials and Methods.....	52
2.1 Reagents and resources.....	52
2.2 Molecular biology methods	57
2.2.1 Bacterial transformation.....	57
2.2.2 Cloning pRDA174-enAsCas12a-EGFP	57
2.2.3 Cloning pRDA052_EGFP_targeting_gRNAs	58
2.2.4 Replacing mCherry with BSD in SUZ12 variant constructs	59
2.2.5 CRISPR library amplification.....	60
2.2.6 CRISPR library validation	60

2.3	Mammalian cell culture methods.....	62
2.3.1	Cell lines and cell culture methods.....	62
2.3.2	Cell line authentication	63
2.3.3	Puromycin and blasticidin kill curves.....	64
2.3.4	UNC7700 treatment of iHSC cells.....	65
2.3.5	Lentivirus Production via PEI-Mediated Transfection	65
2.3.6	Lentiviral Transduction.....	66
2.3.7	CRISPR system validation	66
2.3.8	Determination of transduction efficiency with CRISPR library lentivirus	67
2.3.9	Genome-wide CRISPR knockout screens	67
2.4	Biochemical methods	69
2.4.1	Preparation of Linear Polyethyleneimine (PEI) Solution.....	69
2.4.2	Genomic DNA extraction.....	69
2.4.3	Whole cell extract lysis.....	70
2.4.4	Protein quantification.....	70
2.4.5	SDS-PAGE and Western blotting.....	71
2.4.6	Preparation of 4% paraformaldehyde (PFA) solution.....	72
2.4.7	Immunolabelling for Flow cytometry.....	72
2.4.8	Preparation of Linear Polyethyleneimine (PEI) Solution.....	73
2.4.9	PCR amplification of gRNA cassette region, isolation and NGS-sequencing.	73
2.4.10	Synthesis of UNC7700.....	75
2.5	Bioinformatics and data analysis.....	76
2.5.1	Whole-genome sequencing analysis.....	76
2.5.2	RNA sequencing analysis	76
2.5.3	Standard CRISPR screen data analysis using Log ₂ (fold change)	77
2.5.4	CRISPR screen data analysis using Chronos.....	77
2.5.5	Common essential gene recall	78
2.5.6	Use of DepMap in this study.....	78
2.5.7	pDNA representation and gene effect correlation	79
2.5.8	Prioritisation of context-dependent hit genes	79
2.5.9	Protein structure and AlphaFold structural prediction.....	80
2.5.10	Statistical analysis.....	80
3	Chapter 3: CRISPR screen technical design variables impacting on the selection of hits	81
3.1	Introduction	82
3.2	Results.....	84
3.2.1	Developing and validating enAsCas12a expressing neural crest derived cell line models 84	
3.2.2	Stability of enAsCas12a expression using GFP in cell lines	88
3.2.3	CRISPR library amplification and validation	91
3.2.4	Summary of conducted CRISPR screens in enAsCas12a cell lines.....	93
3.2.5	CRISPR screen duration affects gRNA dropouts	96
3.2.6	Global effect size is a determinant for quality of CRISPR screen data.....	99
3.2.7	Recall of essential (lethality) genes can be applied as a quality control measure	101
3.2.8	Chronos improves CRISPR screen data analysis	103
3.2.9	Copy number variation impact and correction	110

3.2.10	Comparison of this study with DepMap screens.....	114
3.2.11	CRISPR library choice is major determinant for screen results.	116
3.2.12	Low initial gRNA representation leads to stronger gene effects during CRISPR screens.	120
3.2.13	Correction for the effect of initial gRNA representation	128
3.3	Discussion	130
4	Chapter 4: Context dependent gene essentiality and hit prioritisation	133
4.1	Introduction	134
4.2	Results.....	137
4.2.1	Cell line-specific transcriptomic responses to HPLM culture	137
4.2.2	Medium-context dependent cell culture versus physiological transcriptomics.....	145
4.2.3	HPLM culture population doubling rate alters the CRISPR screen workflow.....	149
4.2.4	Medium-specific differential essentiality of gene sets linked to transcriptomic analysis.....	154
4.2.5	Inverse associations between medium-specific gene essentiality and differential gene expression	160
4.2.6	Hit prioritisation of medium-dependent essential genes	164
4.2.7	Hit prioritisation of MPNST-specific medium-independent essential genes.....	170
4.2.8	Prediction of synthetic lethality associates with lower gene effect scores	177
4.2.9	Integration with external MPNST CRISPR screens suggests synthetic lethality and druggable gene targets	179
4.3	Discussion	184
5	Chapter 5: Polycomb Repressive Complex 2-context (PRC2) dependent synthetic lethality	190
5.1	Introduction	191
5.2	Results.....	195
5.2.1	EED-targeting PROTAC UNC7700 degrades PRC2 components and derepresses target genes.....	195
5.2.2	CRISPR screens reveal candidate PRC2 context-specific essential genes.....	199
5.2.3	PRC2-context dependent changes in gene expression inversely associates with differential gene expression	204
5.2.4	Reduction of SUZ12 expression follows complementation in MPNST cell line ...	206
5.3	Discussion	210
6	Chapter 6: Concluding remarks and future perspectives.....	214
6.1	Concluding remarks and future prospects	214
6.2	Optimise and validate enAsCas12a CRISPR screen workflow: ...	215
6.3	Cellular dependencies in different cell culture media	219
6.4	Potential genetic dependencies in MPNST models	223

6.5	PRC2 loss associated synthetic lethality in MPNST	227
6.6	Future Work	228
7	Supplemental information.....	230
	Bibliography	235

Abstract

Malignant peripheral nerve sheath tumours (MPNSTs) are aggressive neural-crest cell derived sarcoma with limited treatment options, and only margin-free surgery as a potentially curative treatment option. The dire prognosis of these patients warrants the need for novel therapeutics. The advancements in genome-engineering now allow for rapid functional interrogation of cancer genomes using genome-wide Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) knockout screens. This study aimed to optimise and utilise an enAsCas12a-based dual-targeting CRISPR library screen workflow to identify dependencies in MPNSTs in various context, including potential synthetic lethality relationships. During the optimisation of the CRISPR screen workflow, variability in initial guide RNA (gRNA) representation emerged as a significant confounding factor. This variability contributed significantly to discrepancies between replicate experiments and across different CRISPR libraries, leading to substantial alterations in observed gene effects and increasing the risk of false-positive results. Notably, this confounding factor was found to be prevalent and unaddressed in other external CRISPR screen datasets, including the entire DepMap dataset. Although further refinement in correcting for gRNA representation bias is possible, adjustments were made in this study to mitigate its impact while identifying robust genetic dependencies in MPNSTs across eight cell line models and 15 independent CRISPR screen experiments. These identified dependencies were either predicted to be druggable or involved in synthetic lethality relationships.

Key candidate genes included well-established cancer dependencies, such as MYC, aurora kinases, cyclin-dependent kinases, and MTOR, as well as several less well-characterized genes. Additionally, a more focused CRISPR screen was conducted to identify synthetic lethality partners associated with the loss of polycomb repressive complex 2 (PRC2) function, a common feature in MPNST. This was achieved by degrading PRC2 components using a targeted Proteolysis-targeting Chimera (PROTAC) approach in immortalized human Schwann cells. The candidate target genes identified in this study hold promise, but further functional and mechanistic validation is necessary to confirm their roles and therapeutic potential.

List of Figures

Figure 1.1. Models of genetic evolution in cancer.....	27
Figure 1.2. Driver variant accumulation during MPNST development.....	38
Figure 1.3. Deletion patterns on chromosome 17 affecting the NF1 gene and co-deletion of SUZ12.	40
Figure 1.4. Ras-driven transcription and epigenetic regulation in tumour progression..	44
Figure 3.1. Circos plots of copy number variation and single nucleotide variant data. 90-8, ST-8814, s462TY, STS26T, iHSC cells. (A-E)	85
Figure 3.2. Stability enAsCas12a expression is time and cell line dependent.	90
Figure 3.3. CRISPR library amplification shows minimal skewing of gRNA cassette distribution.....	93
Figure 3.4. CRISPR screen coverage is time dependent.....	98
Figure 3.5. Common essential gene recall is dependent on CRISPR screen effect size.	103
Figure 3.6. Chronos improves CRISPR screen data analysis based on common essential genes recall and replicate similarity.	106
Figure 3.7. Chronos improves CRISPR screen data analysis based on similarity between Humagne C and D enAsCas12a libraries.....	109
Figure 3.8. CRISPR screen outcomes correlate with higher gene copy numbers in a cell line dependent manner.	113
Figure 3.9. Quality of CRISPR screens in this study are of similar quality to DepMap screens.	115

Figure 3.10. CRISPR library choice can be a more important determinant of screen results than the cell line model.	120
Figure 3.11. CRISPR screen outcomes analysed by Chronos depend on the initial gRNA cassette representation in CRISPR libraries.	124
Figure 3.12. CRISPR screen outcomes depend on initial gRNA cassette representation in library in DepMap dataset.	126
Figure 3.13. Difference between replicate CRISPR screens and Humagne libraries depends on initial gRNA cassette representation.	127
Figure 3.14. Correction for the effect of initial gRNA cassette representation improves similarity between replicate CRISPR screens.	129
Figure 4.1. Comparative Analysis of Human Plasma-like Medium to Human Plasma Composition.	135
Figure 4.2. Culture medium context-dependent differential gene expression is cell line dependent.	139
Figure 4.3. Culture medium context-dependent differential gene expression shows consistent effects across cell lines.	143
Figure 4.4. DMEM medium context cell line transcriptomics show closer correlation with primary MPNST tumour samples compared to HPLM.	148
Figure 4.5. Cell line growth rate is medium context-dependent.	151
Figure 4.6. HPLM CRISPR screen durations capture maximum potential screen quality.	152
Figure 4.7. An initial reference time of four days retains HPLM and DMEM CRISPR screen data quality.	153

Figure 4.8. GSEA identifies differentially expressed and essential pathways and hits in medium context.	158
Figure 4.9. Culture medium-dependent differential hits do not correlate with medium context expressed genes.	163
Figure 4.10. Medium context-dependent gene effect intersection between cell line models.	167
Figure 4.11. Cell type and cell line dependent essential genes found in MPNST and non-MPNST cell line models.	176
Figure 4.12. Predicted synthetic lethality partners show higher gene essentiality in MPNST cell line models.	179
Figure 4.13. Essential genes across MPNST cell line models in this study and DepMap dataset include similar FDA-approved drug- and synthetic lethality prediction targets.	183
Figure 5.1. PROTAC UNC7700 designed to bind EED, degrade PRC2 components, and disrupt PRC2 functionality.	194
Figure 5.2. UNC7700 degrades PRC2 components and de-represses PRC2 target gene expression in iHSC.	198
Figure 5.3. UNC7700 CRISPR screen suggests PRC2 context-dependent hits in Humagne C and D library screens.	201
Figure 5.4. Potential PRC2 context-specific dependencies identified in an iHSC cell line model using Humagne C or D.	202
Figure 5.5. Potential PRC2 context-specific dependencies identified in an iHSC cell line model.	203

Figure 5.6. UNC7700-dependent differential hits correlate with UNC7700 context-	
independently expressed genes.	205
Figure 5.7. SUZ12 mutations targeting key domains highlighted.	207
Figure 5.8. SUZ12 reconstitution restores PRC2 function but reduces over time.	210

List of Tables

Table 1.1. Clinical trials for MPNST treatments with outcome.	42
Table 2.1. Materials and Reagents	52
Table 2.2. Plasmids used in this study.	54
Table 2.3. Commercial kits used in this study.....	54
Table 2.4. Software used in this study	55
Table 2.5. External datasets used for this study.	56
Table 2.6. Gene Blocks used in this study	58
Table 2.7. Primers used for CRISPR library validation.....	61
Table 2.8. PCR programme used for CRISPR library validation	61
Table 2.9. Cell lines used in this study.	63
Table 2.10. Cell line authentication	64
Table 2.11. Effective killing concentrations per cell line.....	65
Table 2.12. Antibodies.....	71
Table 2.13. Next-Generation-Sequencing primers.....	74
Table 2.14. Programme gRNA cassette region PCR amplification.	75
Table 3.1. Cell line models and genetics.....	86
Table 3.2. CRISPR screen time points and screen identifiers.	94
Table 4.1. GSEA results all cell lines combined.....	143
Table 4.2. Differential essential genes in all cell lines combined: HPLM vs. DMEM. .	158
Table 4.3. Medium-context dependent differential gene 'hits' common to STS26T, 90-8, ST-8814 and iHSC cell lines.	167

Table 4.4. Medium-context dependent differential essential genes common to 90-8 and ST-8814 cell lines.....	167
Table 4.5. Intersection of essential and non-essential genes from CRISPR screens in this study as displayed in Figure 4.11A-H.	176
Table 4.6. 90-8 and ST-8814 Humagne C or D hits, non-essential in STS26T and iHSC	176
Table 4.7. Gene hits and associated drugs.....	183
Table 5.1. Significantly enriched or depleted gene sets in iHSC cells treated with UNC7700 versus DMSO.	199
Table 5.2. Impact of Specific SUZ12 Mutations on PRC2 activity.	208
Supplementary table 7.1. GSEA in individual cell lines: HPLM vs. DMEM.	231

List of abbreviations

Abbreviation	Definition
ACSL	Acyl-CoA Synthetase Long Chain
ACTL	Actin-like Protein
ACTR	Actin-related Protein
ADA	Adenosine Deaminase
ADP	Adenosine Diphosphate
AEBP	Adipocyte Enhancer Binding Protein
AF	Allele Frequency
AKT	Protein Kinase B
ALDH	Aldehyde Dehydrogenase
ALL	Acute Lymphoblastic Leukemia
ALS	Amyotrophic Lateral Sclerosis
AM	Amplitude Modulation
ANE	Aneuploidy
AP	Alkaline Phosphatase
AQR	Aquaporin
ARHGAP	Rho GTPase-activating protein
ARID	AT-Rich Interaction Domain
ATN	Atypical Neurofibroma
AURKA	Aurora Kinase A
AURKB	Aurora Kinase B
BAGEL	Bayesian Analysis of Gene Essentiality Likelihood
BBAC	Biochemical Binding Affinity Constants

DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DNASE	Deoxyribonuclease
DNMT	DNA Methyltransferase
DPBS	Dulbecco's Phosphate Buffered Saline
DR	Direct Repeat
DSB	Double-Strand Break
ECM	Extracellular Matrix
EDTA	Ethylenediaminetetraacetic Acid
EED	Embryonic Ectoderm Development
EFS	Event-Free Survival
EGFP	Enhanced Green Fluorescent Protein
EGFR	Epidermal Growth Factor Receptor
EHMT	Euchromatic Histone Methyltransferase
EIF	Eukaryotic Initiation Factor
EM	Electron Microscopy
EMT	Epithelial-to-Mesenchymal Transition
EZH	Enhancer of Zeste Homolog
FACS	Fluorescence-Activated Cell Sorting
FASTQC	FASTQ Quality Control
FBS	Fetal Bovine Serum
FBXO	F-Box Protein
FDA	Food and Drug Administration
FLAG	Flag Tag (Peptide)

FTI	Farnesyltransferase Inhibitor
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GATK	Genome Analysis Toolkit
GSEA	Gene Set Enrichment Analysis
HER	Human Epidermal Growth Factor Receptor
HIV	Human Immunodeficiency Virus
HSC	Hematopoietic Stem Cell
IFITM	Interferon-Induced Transmembrane Protein
IHSC	Immortalised Human Schwann Cells
KRAS	Kirsten Rat Sarcoma Viral Oncogene Homolog
LOESS	Locally Estimated Scatterplot Smoothing
LOF	Loss of Function
LOH	Loss of Heterozygosity
LR	Likelihood Ratio
MAPK	Mitogen-Activated Protein Kinase
MEK	MAPK/ERK Kinase
MIT	Massachusetts Institute of Technology
MMR	Mismatch Repair
MOI	Multiplicity of Infection
MPNST	Malignant Peripheral Nerve Sheath Tumour
MRPL	Mitochondrial Ribosomal Protein Large Subunit
MRPS	Mitochondrial Ribosomal Protein Small Subunit
MSI	Microsatellite Instability
MTOR	Mechanistic Target of Rapamycin
MTREX	Mitochondrial RNA Exonuclease

MTRNR	Mitochondrial rRNA
MYC	Myelocytomatosis
NCI	National Cancer Institute
NDUFA	Oxidoreductase Subunit A
NDUFB	Oxidoreductase Subunit B
NDUFS	Oxidoreductase Subunit S
NEB	New England Biolabs
NEJM	New England Journal of Medicine
NF	Neurofibromin
NF1	Neurofibromatosis type 1
NGS	Next-Generation Sequencing
NHEJ	Non-Homologous End Joining
NRAS	Neuroblastoma RAS Viral Oncogene Homolog
NSCLC	Non-Small Cell Lung Cancer
PBS	Phosphate-buffered Saline
PBST	PBS with Tween-20
PC	Principal component
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PEI	Polyethylenimine
PIK	Phosphatidylinositol 3-Kinase
PIKFYVE	Phosphoinositide Kinase FYVE Finger-containing
PLK	Polo-like Kinase
POLD	DNA Polymerase Delta
POLE	DNA Polymerase Epsilon

POLR	RNA Polymerase
POMP	Proteasome Maturation Protein
PRC	Polycomb Repressive Complex
PROTAC	Proteolysis-targeting Chimera
PSMA	Proteasome 20S Subunit Alpha 1
PSMB	Proteasome Subunit Beta
PSMC	Proteasome Subunit ATPase
PSMD	Proteasome Subunit Non-ATPase
PTEN	Phosphatase and Tensin Homolog
QC	Quality Control
RPAP	RNA Polymerase II-associated Protein
RPMI	Roswell Park Memorial Institute Medium
SARC	Sarcoma
SARS	Seryl-TRNA Synthetase 1
SC	Stem Cell
SCLC	Small Cell Lung Cancer
SD	Standard Deviation
SDS	Sodium Dodecyl Sulfate
SMAD	Signalling Mothers Against Decapentaplegic
SNRNP	Small Nuclear Ribonucleoprotein
SNRPB	Small Nuclear Ribonucleoprotein Polypeptide B
SNRPD	Small Nuclear Ribonucleoprotein Polypeptide D
SNRPF	Small Nuclear Ribonucleoprotein Polypeptide F
SNV	Single Nucleotide Variant
SOC	Super Optimal Broth with Catabolite Repression

SPTBN	Spectrin Beta, Nonerythrocytic
STAR	Spliced Transcripts Alignment to a Reference
STR	Short Tandem Repeat
TALEN	Transcription Activator-like Effector Nuclease
TCA	Tricarboxylic Acid Cycle
TERT	Telomerase Reverse Transcriptase
TGF	Transforming Growth Factor
TOP	Topoisomerase
USA	United States of America
VCF	Variant Call Format
WGS	Whole Genome Sequencing
WHO	World Health Organization
ZNF	Zinc Finger Protein

Chapter 1: Introduction

1.1 Challenges in cancer treatment

In 2022, an estimated 10 million people died from cancer[2]. Cancer is fundamentally driven by evolutionary processes, where mutations and selective pressures enable cancer cells to proliferate and outcompete healthy cells. Efforts to control cancer through targeted therapies have uncovered a wide range of evolutionarily conserved and acquired mechanisms responsible for treatment resistance and tumour regrowth. Technological advancements have significantly deepened our understanding of the cellular, molecular, and environmental factors within the cancer ecosystem, resulting in successes such as leveraging the immune system against this disease. However, for most patients with advanced metastatic disease, current treatments primarily aim to suppress tumour burden rather than achieve curative outcomes. Although chronic therapies have modestly improved 3- and 5-year survival rates, these gains rarely lead to cures. This raises an important question: are our current models of cancer research, drug development, and clinical trials equipped to address the complex mutation-selection dynamics at the heart of cancer evolution? Many therapeutic strategies still adhere to the traditional focus on single-target, single-agent approaches, which often fail in the face of cancer's capacity for rapid evolutionary resistance. Despite high-profile initiatives such as "moonshot" programs and "grand challenges" that aim to encourage innovation[3], [4], the development of combination therapies, an approach that could more effectively address the multifaceted nature of cancer evolution, remains underexplored and often lacks the rigorous justification needed to overcome the challenges posed by the cancer evolutionary cycle. A shift toward more integrated,

evolutionary-based strategies is essential to develop the next generation of cancer therapies.

1.1.1 The cancer evolutionary cycle

In cancer, either single or several combined DNA variants act as drivers of uncontrolled cell division and growth. Several genetic and eventually chromosomal instability mechanisms (CIN), including error prone DNA repair and uncontrolled mitotic cell division, result in either gain of function (oncogene) mutations or loss of function (tumour suppressor gene) copies. Together, accumulation of these traits, and in particular larger chromosomal alterations, of which carriers are referred to as 'hopeful monsters'[5], inevitably accelerate genetic intra-tumour diversification between cells and leads to tumour formation, invasion and metastasis. This comes at a cost of purifying selection if the effect of CIN exceeds cell type specific limits that results in cell death. Although genetic selection and its relation to phenotypes is poorly understood, it can in part be explained by fusing together initially individually proposed models: branching evolution, punctuated evolution, neutral evolution between selection events, and, probably less frequently, linear evolution (Fig. 1.1)[5]. Ultimately, this genetic divergence, or heterogeneity, is compounded by co-existing epigenetic and proteomic divergence in cancer cells, also leading to intra-tumour heterogeneity at the expression and protein level. Genetic alterations at the cellular level are, however, not the only cause of intra-tumour heterogeneity. Tumour cell localisation and microenvironment also provide selective pressures, making intra-tumour heterogeneity dependent on multiple fixed and modifiable factors.

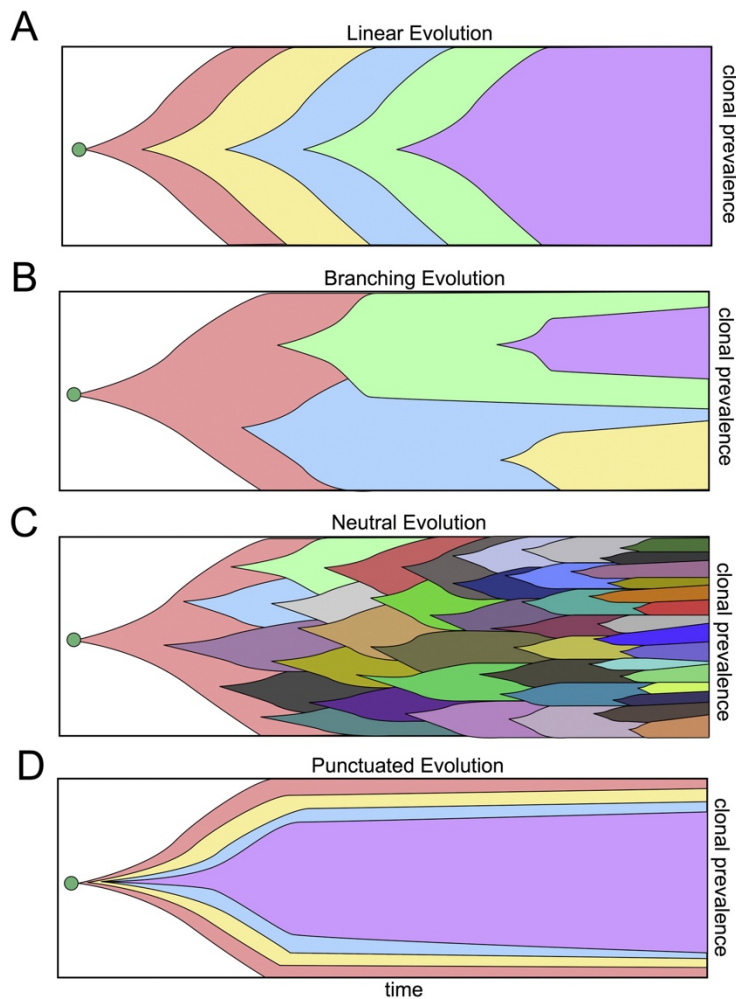


Figure 1.1. Models of genetic evolution in cancer. This figure illustrates four distinct models of genetic evolution observed in cancer development: A: Linear Evolution depicts a scenario where selective pressures allow only one dominant clone to survive and expand over time, leading to a relatively homogenous tumour. B: Branching Evolution reflects the nature of cancer's uncontrolled cell division, where multiple subclones simultaneously evolve, contributing to tumour heterogeneity. C) Neutral Evolution represents periods between strong selective events, where subclones accumulate mutations at a relatively constant rate, leading to a highly heterogeneous tumour without clear selective advantages for any particular

clone. D) Punctuated Evolution suggests that tumour diversification occurs abruptly following selective pressures, leading to a sudden expansion of new clones, in contrast to the more gradual processes depicted in the other models. This figure was adopted from another publication[6].

1.1.2 Cancer cells do not rely on one mechanism for growth

A key insight from cancer evolution is that the majority of accumulated genetic alterations arise from chromosomal instability (CIN)[7]. Consequently, cancer dependencies are more likely driven by combinations of alterations rather than by individual mutations. Historically, research has focused on identifying single driver mutations as promising drug targets. While therapies targeting single molecules, such as HER2 inhibitors for breast cancer or EGFR inhibitors for lung cancer, have shown efficacy, they are limited by the fact that only a small subset of patients respond to these

therapies, and many develop resistance due to further tumour diversification[8], [9]. To overcome these limitations, there is growing interest in targeting multiple dependencies simultaneously. Dual or multi-target therapies offer a promising alternative, as they are better suited to address the complex, multi-faceted survival mechanisms of cancer cells.

1.1.3 Plasticity and intra-tumour heterogeneity in cancer evolution

Not all cancer cells within a tumour are equal. A key principle derived from the evolution of cancer is that tumours possess multiple strategies, rather than a singular one, to survive selective pressures. Tumours are composed of various subpopulations or subclones, creating intra-tumour heterogeneity that mitigates the risk of dependence on any one cellular phenotype [10]. This heterogeneity is driven by intrinsic genetic factors, extrinsic environmental influences, spatial organization within the tumour, and patient-specific characteristics such as the germline genome. Clonal dominance within a tumour may not be evident initially but can emerge when selection pressures change, allowing certain subclones to become more prominent. This dynamic clonal evolution, in part, explains phenomena such as spontaneous cessation of tumour growth, changes in cellular differentiation, vascular access, and infiltration by stromal and immune cells [10].

Additionally, time acts as a "fourth dimension" in this heterogeneity: tumour cells are plastic to a certain degree, allowing them to adapt over time and develop capabilities to survive under new selective pressures [11]. This plasticity permeates all levels of intra-tumour heterogeneity but affects them differently. For instance, while genetic alterations tend to be more stable, epigenetic signatures and changes in the tumour microenvironment exhibit greater plasticity. A well-known example of cancer plasticity is epithelial-mesenchymal transition (EMT), which reflects the adaptability of tumour cells

at the epigenetic and microenvironmental levels. During EMT, tumour cells of epithelial origin acquire mesenchymal characteristics, initially thought to facilitate migration and metastasis [11]. However, EMT is also implicated in drug resistance by fostering dormancy and "cancer stemness" via the TGF- β pathway [11].

As a result of intra-tumour heterogeneity, certain cancer cells exhibit higher plasticity than others, making therapeutic strategies more challenging. In highly plastic areas, such as epigenetic targets, drug development can be less effective, though some phenotypic traits may remain stable enough to be targeted. Successful cancer treatments must either overcome or bypass this plasticity and heterogeneity by eradicating all cancer cell subpopulations or continually redirecting treatment to newly emerging subpopulations in the event of regrowth.

1.1.4 Clonal hierarchy

Clonal hierarchy within tumours is unevenly distributed across tumour tissues and is governed by a cellular hierarchy, with cancer stem cells (CSCs) often considered the root of tumour cell diversification and progression[12]. CSCs represent a small, undifferentiated population of tumour cells that exhibit stem-like properties, resembling pluripotent stem cells. Unlike differentiated cancer cells, CSCs possess the capacity for unlimited division, fuelling tumour growth through the production of progenitor cells that further expand the tumour mass. Local stem cell niche factors, regulating transcription, cell cycle activity, and specific signalling pathways, play a critical role in maintaining CSC function and self-renewal capacity.

CSCs are believed to be the origin of clonal subpopulations within tumours, with specific driver mutations arising from CSCs. This positions CSCs as major contributors

to both genetic and epigenetic intra-tumour heterogeneity. Various models have been proposed to describe the precise role and origins of CSCs [12]. Evidence suggests that CSCs may not only be the initial drivers of tumour growth but can also arise from non-stem cancer cells through a process of cellular plasticity [12]. This challenges the traditional view of a rigid clonal hierarchy, as stemness can be acquired in response to various intrinsic and extrinsic factors, including epigenetic modifications and microenvironmental interactions.

The concept of reversible plasticity in CSCs introduces both challenges and opportunities for therapeutic targeting. Although therapies aimed at eradicating CSCs could prevent or reverse tumour heterogeneity and resistance, the ability of non-stem cancer cells to acquire stem-like properties complicates these strategies [12], [13]. CSC-directed therapies currently in development target key signalling pathways involved in CSC maintenance, including Wnt, Notch, Hedgehog (Hh), TGF- β , JAK-STAT, PI3K, and NF- κ B [12], [13]. However, outcomes have been mixed, and there are obstacles to CSC-targeted therapies, particularly due to the limitations of *in vitro* and *in vivo* models. *In vitro* studies often fail to account for the critical role of the CSC niche, while immune-deficient murine models lack the interactions between CSCs and the host immune system, hindering the translatability of findings to clinical settings.

Despite these challenges, CSC-targeted strategies remain a promising approach to disrupting intra-tumour heterogeneity and preventing therapy resistance, and ongoing efforts to develop effective CSC-directed therapies, including those targeting the CSC niche and employing immunotherapies, hold potential for future breakthroughs.

1.2 Overcoming the cancer evolutionary cycle

To develop more effective cancer therapies, drug discovery must address two critical principles derived from cancer evolution. First, research models need to closely reflect the cancer subpopulations they aim to target, as intra-tumour heterogeneity means that different subclones may respond variably to treatment. Second, due to the genetic instability of cancer cells, therapies are unlikely to be successful if they rely on single targets. Instead, treatments should either focus on multiple mechanisms that capture all tumour subpopulations or target shared vulnerabilities across the tumour. Intra-tumour heterogeneity, spanning genetic, epigenetic, and microenvironmental factors, complicates treatment and necessitates strategies that can address the diverse subclonal populations within the tumour.

1.2.1 Immune checkpoint inhibition

Every tumour cell arises due to a mutation or genetic variant, enabling it to evade immune surveillance. Given this, reviving or amplifying an immune response is a promising strategy to bypass the challenges posed by intra-tumour heterogeneity. Immunotherapy offers the advantage of high specificity, as it leverages the immune system's natural capacity for antigen-directed cytotoxicity, allowing it to distinguish cancer cells from healthy tissue with greater precision.

One of the more established forms of immunotherapy is immune checkpoint inhibition. Immune checkpoint inhibitors (ICIs) work by targeting the mechanisms through which tumour cells corrupt immune checkpoint pathways[14]. The most well-known ICIs are PD-1/PD-L1 inhibitors, which block the interaction between the programmed cell death-1 protein (PD-1) and its ligand PD-L1, both of which are commonly overexpressed

in cancer cells to suppress immune responses[14]. Inhibitors targeting CTLA-4, a receptor on T cells that downregulates immune responses when activated, are also widely used. Other potential targets under investigation include LAG-3, TIM-3, ITIM, and VISTA[15]. ICIs are particularly effective in cancers with DNA mismatch repair (MMR) deficiencies, as these tumours tend to produce more mutations, leading to increased antigenicity and vulnerability to immune attack[16]. However, resistance to ICIs remains a challenge, primarily due to tumour heterogeneity in MMR dysfunction and the variable presence of infiltrating immune cells. Cancer cell plasticity also allows tumour cells to lose immune-targeted antigens, further complicating treatment[17]. To overcome these hurdles, combination therapies are being explored, such as combining immunotherapy with molecularly targeted treatments or increasing antigen expression through methods like low-dose radiotherapy[17]. Epigenetic modifications to induce the re-expression of tumour-specific antigens may also offer a promising strategy[17].

1.2.2 T-cell therapy in cancer

Chimeric antigen receptor T (CAR-T) cell therapy utilises engineered T-cells that recognise specific antigens on the surface of cancer cells, triggering a potent cytotoxic response. However, one of the primary limitations of CAR-T therapy is that it is largely restricted to targeting extracellular surface antigens, which are often absent or heterogeneously expressed in solid tumours, failing to address the conundrum of the cancer evolutionary cycle[18].

In contrast, T-cell receptor (TCR) engineered T-cells offer a broader range of target recognition. Unlike CAR-T cells, TCR-T cells can identify intracellular antigens that are processed and presented on the surface of cancer cells via major histocompatibility

complex (MHC) molecules[19]. This allows TCR-T cells to target a wider variety of cancer-specific proteins, including those derived from mutated or aberrantly expressed intracellular proteins. As a result, TCR-T therapies hold greater potential for treating solid tumours and cancers that lack abundant surface antigens. While CAR-T therapy has shown substantial success in haematological malignancies, particularly for CD19-positive cancers such as acute lymphoblastic leukaemia (ALL) and diffuse large B-cell lymphoma (DLBCL), TCR-T cell therapy expands the treatment landscape by overcoming the antigen limitations associated with solid tumours[20].

1.2.3 Cancer vaccines

Anti-cancer vaccines, utilising synthetic peptides, mRNA, vectors, or dendritic cells loaded with tumour-specific antigens, aim to prime the immune system to recognise and destroy cancer cells. Early efforts in this field were hampered by a focus on tumour-associated antigens (TAAs), which often failed to trigger a strong immune response because T cells that recognise these self-antigens are typically eliminated during immune development[21]. Tumour neoantigens (TNAs), on the other hand, are unique to cancer cells and elicit a more robust immune response with fewer off-target effects [21].

Although some cancer vaccines have shown efficacy as standalone treatments, they generally perform better when combined with other therapies, such as immune checkpoint inhibitors, which can amplify the immune response by "releasing the brakes" on immune activity[22]. This combination approach has revitalised interest in cancer vaccines, with numerous ongoing clinical trials investigating their potential[23]. Cancer vaccines targeting TNAs can provoke a stronger immune attack, addressing tumour heterogeneity by targeting multiple cancer subpopulations simultaneously[24]. While this

approach is still evolving, it holds promise for enhancing the immune system's ability to combat cancer on multiple fronts.

1.2.4 Synthetic lethality

Another opportunity to target the cancer evolutionary cycle involves exploiting the dependencies between driver or passenger mutations and associated 'partner' genes, a concept known as synthetic lethality. Synthetic lethality arises when the loss of one gene renders cells dependent on the function of another gene (or genes), such that inhibiting the second gene results in cell death. However, cells that retain the function of either gene remain unaffected. This concept provides a unique opportunity to target cancer-driving mutations through their synthetic lethal partners.

A paradigmatic example of synthetic lethality is the use of poly (ADP-ribose) polymerase (PARP) inhibitors in cancers harbouring *BRCA1* or *BRCA2* mutations [25], [26]. *BRCA1* and *BRCA2* are integral to the homologous recombination repair (HRR) pathway, responsible for repairing DNA double-strand breaks. In one of the multiple mechanism models that have been identified, the absence of functional *BRCA1* or *BRCA2*, cells become dependent on PARP-mediated base excision repair (BER) to resolve single-strand DNA breaks. Inhibition of PARP impairs the repair of these breaks, leading to the accumulation of double-strand breaks, which culminates in genomic instability and cell death. Clinical trials have demonstrated that PARP inhibitors significantly improve progression-free survival in patients with ovarian and breast cancers carrying somatic or germline *BRCA1/2* mutations [27], [28], [29]. Another notable example of synthetic lethality involves the redundancy between *ARID1A* and *ARID1B*, both of which are critical subunits of the SWI/SNF chromatin remodelling complex. This

complex modulates gene expression by using ATP to reposition nucleosomes, thereby regulating chromatin accessibility for transcription, DNA replication, and repair. The simultaneous loss of *ARID1A* and *ARID1B* impairs the essential chromatin remodelling function of the SWI/SNF complex, resulting in cell death. However, when only one of these subunits is lost, the other can compensate, maintaining complex functionality and preserving cell viability [30]. A final example of synthetic lethality can be observed in *KRAS*-mutant non-small cell lung cancer (NSCLC), where tumour cells exhibit dependency on the interplay between the proteasome pathway, Rho signalling, and NF- κ B signalling. The combined inhibition of the proteasome and NF- κ B pathways by Bortezomib, alongside inhibition of Rho signalling by Fasudil, induces tumour cell death [31]. In contrast, targeting one or two of these pathways individually is insufficient to induce this effect [40]. Together, these examples underscore the therapeutic potential of exploiting synthetic lethality in precision oncology.

Although these findings appear promising, it is important, however, to acknowledge that synthetic lethality relationships are often falsely presented as a binary concept, where targeting one of the pair is harmless, whereas targeting the pair is always lethal. The reality, however, is that synthetic lethality is often dependent on the wider genetic and epigenetic context of the cell[32]. The regularly observed resistance to PARP inhibition in breast cancer therapy highlights this complexity, where mechanisms including BRCA1/2 reversion mutations as well as non-reversion models, like the replication rescue model [33], highlight the broader need for caution when classifying synthetic lethality. A comprehensive understanding of the genetic dependencies underlying synthetic lethal pairs is crucial to overcoming the resistance observed in clinical settings across various targets.

1.3 Malignant Peripheral Nerve Sheath Tumours

Malignant peripheral nerve sheath tumours (MPNSTs) are aggressive neural crest and Schwann cell derived sarcoma with a poor prognosis. Approximately 60% of MPNST cases arise sporadically (37-58%) and 10% post-radiology [34], [35].

Around 32-52% of MPNST cases are associated with neurofibromatosis type 1 (NF1) [34], [35], where an autosomal dominant loss of function (LOF) mutation in *neurofibromin 1*, a Rat sarcoma virus (RAS) GTPase activating protein (RASGap) and regulator of the Rat sarcoma virus (RAS) signalling pathway, leads to neurofibroma formation primarily located across the peripheral nervous system (benign overgrowth of Schwann cells around peripheral nerves).

Clinically, patients with NF1 often present with multiple hallmark features that can aid in early diagnosis, including café-au-lait macules, axillary or inguinal freckling, and Lisch nodules of the iris[36]. Specifically, cutaneous neurofibromas are frequently observed, which, although benign, may cause significant morbidity depending on their size and location. Skeletal abnormalities such as scoliosis and tibial dysplasia are also common in NF1 patients, reflecting the disorder's systemic involvement[36]. These clinical manifestations contribute to the overall complexity of managing NF1, especially given the increased risk of malignant transformation into MPNSTs, which underscores the importance of vigilant monitoring and early intervention in this patient population.

Two specific subcategories of neurofibroma, called atypical neurofibroma (ATN) and plexiform neurofibroma (PN), can more frequently transform into MPNSTs[37], [38]. Both ATN and PN are usually deeper lesions and potentially invasive with the first having histopathological features including nuclear atypia, and the latter often located in neural

plexuses. MPNSTs in the context of NF1 are significantly more deadly than sporadic related cases [35], [39].

1.3.1 MPNST-driver gene variants

While loss of neurofibromin 1 per definition underpins all NF1-related MPNST cases, it is a hallmark of MPNST in general. A recent study showed bi-allelic loss in 88% (56/64) of MPNST tumour samples from NF1 patients, compared to 32% (10/31) in sporadic cases[40]. *CDKN2A/B* is also frequently mutated in MPNST with somatic biallelic inactivation found in 63% (40/64) of NF1-related cases and 55% (17/31) of sporadic MPNSTs[40]. Bi-allelic loss of *TP53* is less commonly found with double hit alterations found in about 21% of MPNST cases[41],[40].

1.3.2 Polycomb repressive complex 2 loss of function in MPNST

Another hallmark of MPNST is the loss of function of the polycomb repressive complex 2 (PRC2) by mutations of core components SUZ12 and EED (Fig. 1.2). Normally, PRC2 consists out of three core components SUZ12, EED, and EZH1 or EZH2, and functions as an epigenetic regulator by trimethylating histone 3 at position lysine 27 (H3K27ME3), but when mutated leads to open access chromatin and unregulated PRC2 target gene expression[42], [43]. Amongst these target genes are ones that are driven by RAS signalling, and so when PRC2 loss compounds *NF1* loss, which on itself accumulates active RAS, this leads to further increased RAS signalling[44], [43].

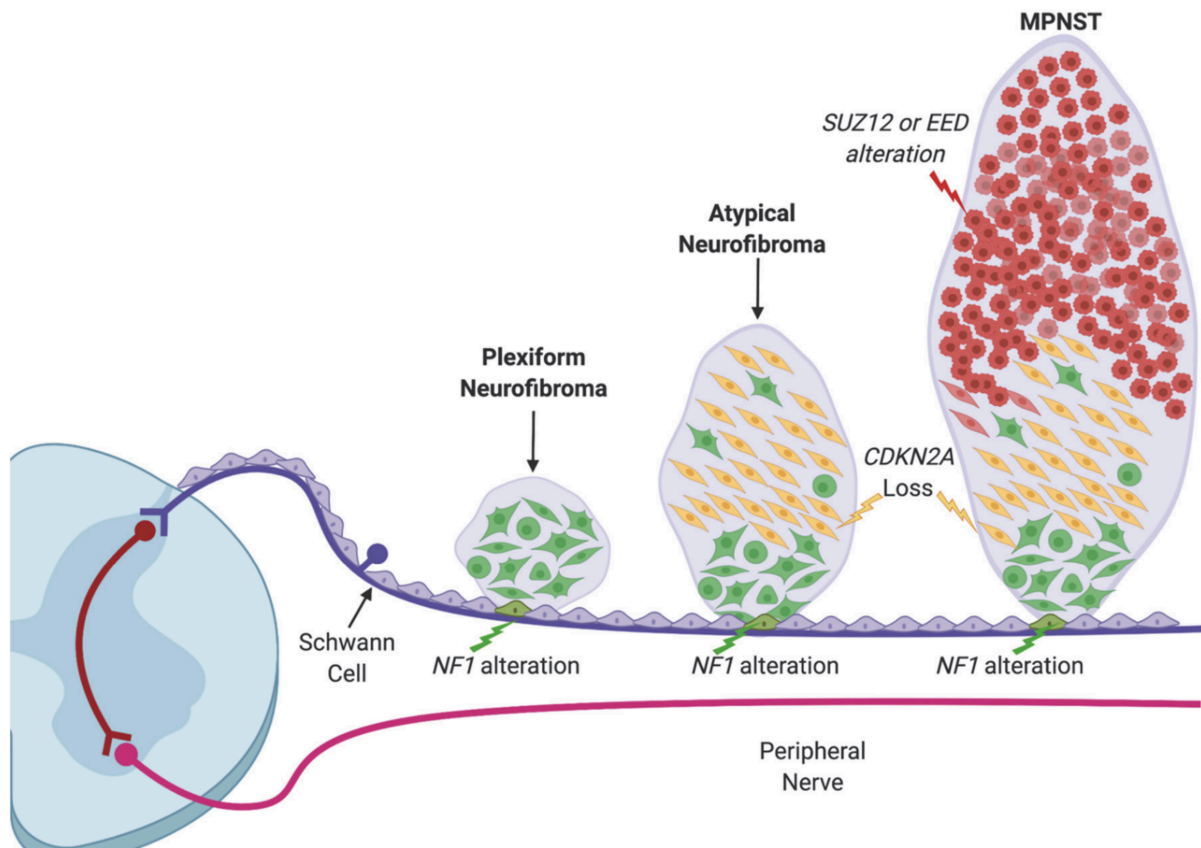


Figure 1.2. Driver variant accumulation during MPNST development. The figure depicts the sequential accumulation of genetic alterations in Schwann cells leading to the development of malignant peripheral nerve sheath tumours (MPNST). *NF1* inactivation typically represents the first genetic hit. Subsequent alterations in epigenetic regulators, such as *SUZ12* or *EED*, facilitate the malignant transformation to MPNST. This figure was adapted from [42].

EZH2 is the catalytic subunit that mediates the trimethylation of histone H3 at lysine 27 (H3K27me3), a key repressive mark associated with gene silencing [45]. EED binds to the methylated H3K27, which allosterically activates EZH2, enhancing its methyltransferase activity. *SUZ12* plays a critical structural role, interacting with all other core components to stabilize the complex and facilitate its assembly [45]. Beyond these core components, cofactors like JARID2 and AEBP2 are integral to PRC2's functionality [45]. JARID2, when methylated at K116, mimics a methylated histone tail, stimulating PRC2 by binding to EED. AEBP2 helps bridge the catalytic core and the structural components of PRC2 by interacting with both *SUZ12* and RBAP48. This connection

stabilizes the complex and ensures its proper conformation for chromatin binding. Together, these interactions ensure PRC2's role in maintaining repressive chromatin states and regulating gene expression during development and differentiation.

One study found loss-of-function somatic mutations in *SUZ12* or *EED* in $\pm 46\%$ of MPNST samples (24/52), with 29% (6/21), 78% (7/9) and 100% (5/5) loss specific to NF1-, sporadic- and irradiation-related cases, respectively[46]. In another study, they found biallelic loss of either *SUZ12* or *EED* in $\pm 51\%$ (26/51) of NF1 patients, although this may be an underestimation of PRC2 loss as they found loss of H3K27ME3 by immunohistochemistry (IHC) in samples for which they did not find biallelic *SUZ12* or *EED* loss, possibly because this was targeted exome-, rather than whole-genome sequencing[44]. Another review found loss of H3K27me3 in 55% (45/82) of MPNSTs (NF1 and sporadic) based on IHC data available, with *SUZ12* and *EED* being almost always mutually exclusively lost in 28% (27/95) and 17% (16/95), respectively[40].

Interestingly, germline loss of *neurofibromin 1* in NF1 patients via microdeletions spanning 1-1.4 mega bases (Mb) regularly co-deletes *SUZ12*, which is a neighbouring gene of neurofibromin 1 on chromosome 17q (Figure 1.3) [44]. Correlation between the larger size of the initiating deletion and the development of atypical neurofibroma and plexiform neurofibroma has been observed in these patients, suggesting co-regulatory pathways are established as the initiating driver event that eventually leads to MPNST[47]. NF1 patients harbouring such microdeletions often display severe clinical features absent in non-microdeletion *NF1* loss, including profound global developmental delay, cognitive disability, severe facial anomalies, and severe autism spectrum disorder symptoms[47].

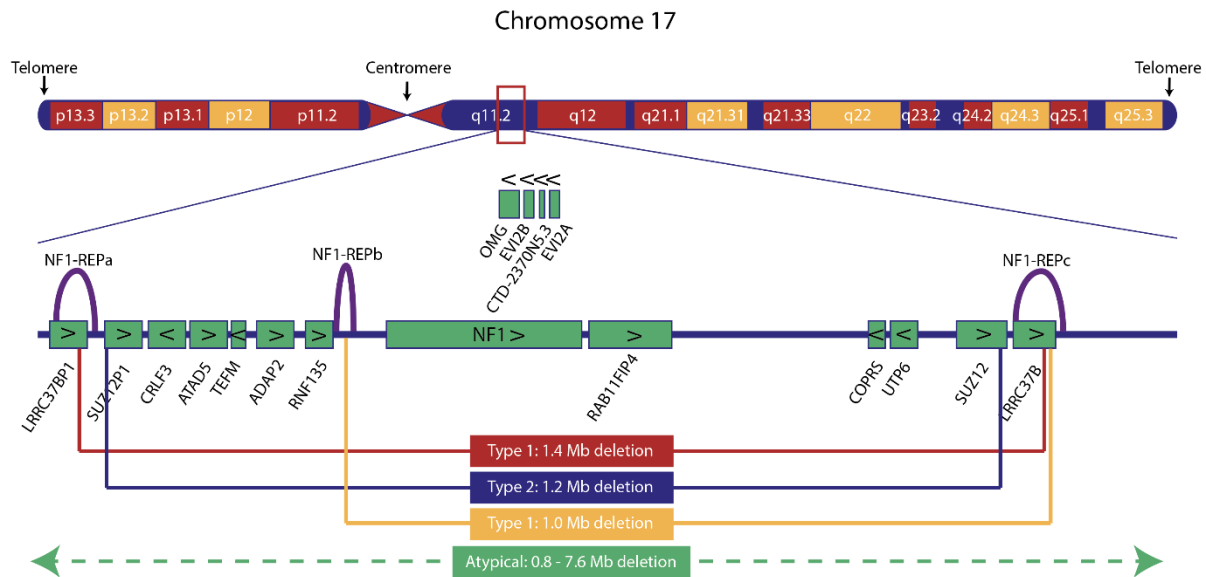


Figure 1.3. Deletion patterns on chromosome 17 affecting the NF1 gene and co-deletion of SUZ12.

The figure shows the structure of chromosome 17, specifically highlighting the NF1 gene region and associated deletions. NF1 is flanked by segmental duplications (NF1-REPa, NF1-REPb, NF1-REPC), which facilitate recurrent deletions of varying sizes. SUZ12, a critical epigenetic regulator, is located downstream of NF1. Due to its proximity, SUZ12 can be co-deleted with NF1 during large deletion events on chromosome 17. These co-deletions typically occur in Type 1, Type 2, and larger atypical deletions (spanning 0.8 to 7.6 Mb), leading to the loss of both NF1 and SUZ12.

Enhancer of zeste homolog 1 and 2 (*EZH1*, *EZH2*) are also core components of PRC2 but are rarely found mutated in MPNST[46]. Although initially *EZH1/2* was hypothesised to have a PRC2-independent function on which MPNST depends, one study provided convincing evidence that the real reason for this observation is that *EZH1* and *EZH2* are large functionally redundant, and so make it less likely for four gene copies to be lost compared to two[48].

PRC2 loss has been directly linked to MPNST progression and metastasis [48]. In MPNST cell line models, deletion of *Eed* or *SUZ12* led to increased invasion through collagen, clustering of cells, and upregulation of ECM remodelling enzymes, notably *Mmp9*, *LoxL2*, and *LoxL4*. Inhibiting these enzymes slowed migration, indicating a role for PRC2 loss in promoting metastasis. In mouse models, PRC2-deficient cells injected

via the tail vein formed larger lung metastases [48]. In orthotopic allograft models, PRC2 loss triggered fibrosis and heightened ECM enzyme activity, further associated with increased metastatic potential. Another study showed that *SUZ12* loss of function, in addition to *NF1* and *TP53* loss, led to earlier occurrence of MPNSTs and decreased survival in mouse models, compared to *NF1* and *TP53* loss alone [49].

More recently, MPNSTs have been classified based on comprehensive molecular profiling using whole-exome sequencing, single nuclei RNA sequencing, and genome-wide DNA methylation analysis[51]. This study found two distinct subtypes of MPNSTs, irrespective of context, where the MPNST-G1 subtype, as compared to MPNST-G2, exhibits more frequent PRC2 core component mutations, and a CpG island hypermethylation phenotype, accumulating in enrichment of specific pathway activation, in particular sonic hedgehog (SHH) signalling. In contrast, MPNST-G2 displays higher levels of WNT/ β -catenin/CCND1 signalling. MPNST-G1 subtype tumours are associated with a regression-free survival that is less than half that of the MPNST-G2 subtype.

1.3.3 Current therapeutic landscape and emerging targets in MPNST

Although our understanding of MPNST driver events and subtypes has been increasing over the last decade, this has not translated into novel therapeutic options, with the only proven effective therapy being complete surgical resection with free margins[52]. The five-year overall survival (OS) rate, event-free survival (EFS) rate and local recurrence (LR) rate remain poor and are estimated to be 49% (range, 16–63%), 37% (range 34–40.6%) and 38% (range 13–86%), respectively[53]. The incidence of MPNST is low with 1.46 per million person-years, with that being a major reason for why its understudied[54].

The benefit of systemic therapies for survival, including chemotherapy, has been investigated but convincing evidence for specific regimens remains absent. The SARC0006 Phase II clinical trial (Table 1.1) suggests a potential benefit from chemotherapy regimens in response rates as measured by WHO standards [54], where the trial patients were treated with two cycles ifosfamide and doxorubicin (IA) following by two cycles of ifosfamide and etoposide (IE), then underwent surgery and/or radiotherapy and had another two cycles of both IA and IE. In about 17.9 – 44.4% of cases this led to a partial response, but the trial was closed before reaching full accrual. Radiotherapy has shown to control MPNST locally and mentioned to be beneficial for tumour shrinkage before surgery but fails to improve OS rates [55], [56], [57]. More targeted therapies haven't shown any significant benefits to MPNST patients either, with barely any partial responses observed during phase II clinical trials, let alone improvement of survival rates (Table 1). Together, this shows the need for exploring new avenues of therapeutic targeting in MPNST.

Table 1.1. Clinical trials for MPNST treatments with outcome.

Study	Drug	Trial phase	Outcome	Reference
SARC023	Sirolimus and ganetespib	I/II	Stable/Progressive disease	[59]
S0330	Erlotinib	II	Stable/Progressive disease	[60]
SARC006	chemotherapy comprising doxorubicin hydrochloride and ifosfamide (IA) followed by etoposide and ifosfamide (IE)	II	after 4 cycles: complete response: 0, stable disease: 24, progressive disease: 4	[61]
ADVL1522	lorvotuzumab mertansine	II	Stable/Progressive disease	[62]
NCT01653028	alisertib	II	Stable/Progressive disease	[63]

NCT00245102	sorafenib	II	Stable/Progressive disease	[64]
-------------	-----------	----	----------------------------	------

Immunotherapies for MPNST have only recently been explored, and the anticipated therapeutic impact in the near future remains limited by several biological and clinical challenges. Specifically, MPNSTs are frequently classified as immune-cold tumours, a characteristic largely attributed to the loss of function in the epigenetic regulator polycomb repressive complex 2 (PRC2), suggesting low responsiveness to most immunotherapy strategies [64], [65]. However, a preclinical study has identified a potential method to sensitise MPNSTs to immunotherapy by activating the stimulator of interferon genes (STING) pathway, which enhanced the efficacy of immune checkpoint blockade in both mouse genetic models and human xenograft MPNST models [66]. Furthermore, the potential benefit of immune checkpoint blockade has been suggested in several case reports of PD-L1 expressing MPNSTs, providing additional support for its exploration in select cases [67], [68], [69]. Despite these promising early findings, immunotherapy for MPNST remains in its nascent stages and may only be effective in a limited subset of patients at this time.

1.3.4 Synthetic lethality in MPNST

Supposed synthetic lethality relationships have also been explored in MPNST. For example, PRC2 loss in MPNST cells increases their sensitivity to BRD4 inhibitors like JQ1, which disrupt the recruitment of transcription machinery by BRD4 by preventing it from binding to acetylated histones, which are more prevalent due to loss of the PRC2 H3K27 trimethylation function (Figure 1.4)[50]. While promising in preclinical models, resistance to BRD4 inhibitors remains a challenge, often due to persistent high levels of BRD4 protein that maintain tumour cell survival despite treatment [70]. Another study

compared dependencies in PRC2 wild type versus PRC2 loss MPNST cell lines during a short hairpin RNA (shRNA) screen targeting 565 epigenetic and chromatin regulator genes [71]. This group found that PRC2 loss sensitises cells to small-molecule inhibition of DNA methyltransferase 1 (DNMT1). Specifically, DNMT inhibition in PRC2 loss of function MPNST cell line models, increases retrotransposon expression, consequent viral mimicry response and cell death.

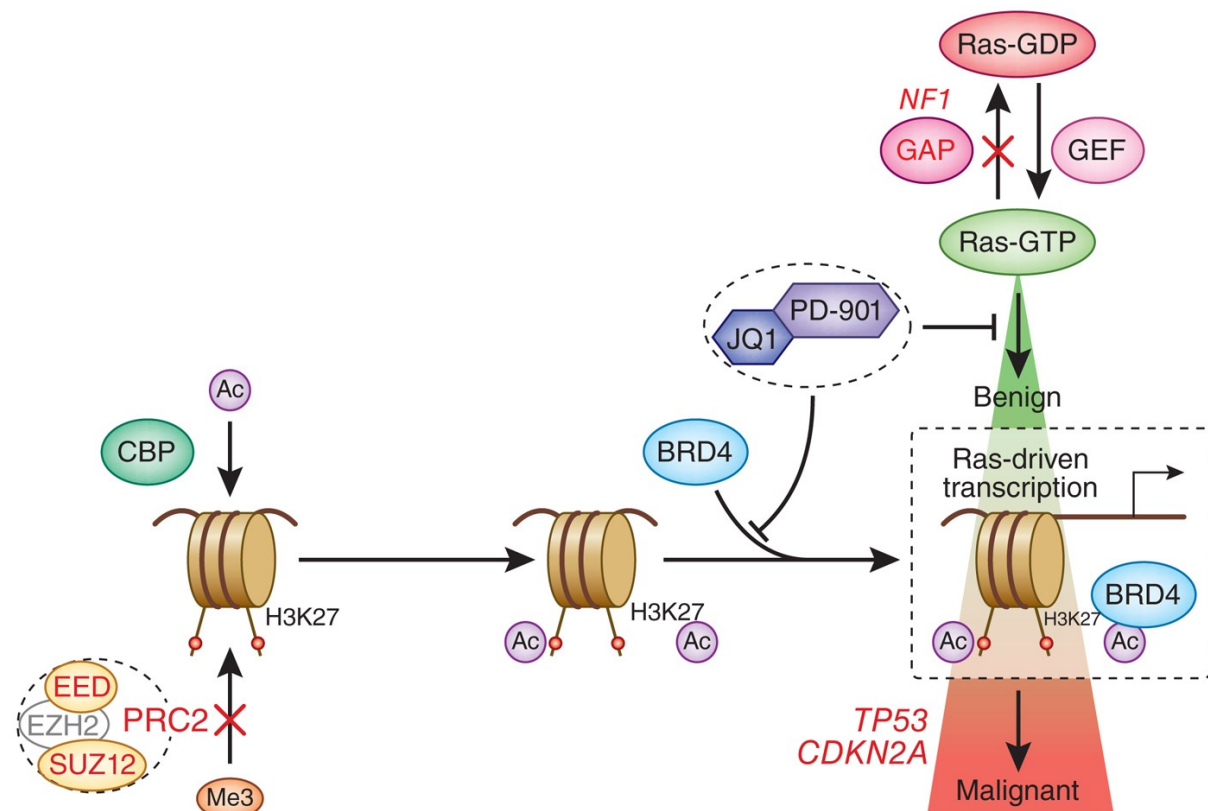


Figure 1.4. Ras-driven transcription and epigenetic regulation in tumour progression. This figure illustrates how loss of NF1 leads to hyperactivation of Ras signalling (Ras-GTP), promoting benign tumour growth. Epigenetic disruption of the PRC2 complex, involving SUZ12 and EED, results in loss of H3K27 trimethylation (H3K27Me3) and increased H3K27 acetylation, facilitating BRD4 recruitment and enhancing Ras-driven transcription. In MPNSTs, where NF1 loss has already increased Ras signalling, this effect is compounded by further Ras signalling due to PRC2 loss. Pharmacological inhibitors such as JQ1 (targeting BRD4) and PD-901 (targeting Ras) may modulate these pathways. This figure was adapted from elsewhere [43].

Importantly, screening for such targets in cancer research beyond MPNST has been aided by developments in Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) technology which allows for high-throughput genome-wide functional interrogations of genes for their vitality to the host cell. Genome-wide CRISPR knockout screens in most ways provide a superior alternative in terms of efficacy and efficiency compared to previous methodology, including shRNA, and thus are increasingly used to screen for cancer-specific dependencies, including synthetic lethality[73].

1.4 CRISPR

The CRISPR system comprises an endonuclease, often CRISPR associated nuclease 9 (Cas9), guided by a single guide RNA (sgRNA) that targets a specific sequence in the genome. The sgRNA is made up of a CRISPR RNA (crRNA) segment, which is a sequence complementary to a target DNA sequence of interest, fused to a tracrRNA segment that serves as a binding scaffold for the endonuclease. The crRNA segment is programmable via regular plasmid construction methods allowing researchers to direct the attached endonuclease to a target of interest. The number of locations that can be targeted is limited, however, because the endonuclease only binds to a target sequence in the presence of a so called protospacer adjacent motif (PAM), which can be unique for each different endonuclease. Depending on the endonuclease, the DNA is precisely cut at a specific distance from the PAM sequence. The resulting double strand breaks are repaired via the error prone, non-homologous end-joining (NHEJ) DNA repair mechanism, which causes disruption of the open reading frame in genes. The simplicity of programming sgRNAs and associated Cas proteins to precisely target specific genomic loci, makes CRISPR an ideal tool for many genome engineering applications.

1.4.1 Genome-wide CRISPR knockout screens

During CRISPR knockout screens, a library of unique sgRNAs targeting different genes is introduced into a number of target cells, where each cell receives one sgRNA causing the knockout of one gene which then potentially affects cell proliferation and viability. As the genes for making these gRNAs are usually integrated into the host genome using a lentivirus, a polymerase chain reaction (PCR) from the common regions flanking the variable crRNA region can be used to amplify, sequence and so detect the relative number of sgRNAs present in the total genomic DNA of the cell population. In this way CRISPR library screens can be pooled where the relative abundance of sgRNAs in a cell population can be determined without the need for isolated CRISPR knockout experiments. The pooling of CRISPR libraries and the relatively small-scale experiments this allows for have enabled the use of genome-wide CRISPR libraries, which contain an array of sgRNAs targeting a large portion or the entire protein-coding genome. The effectiveness as well as precision of CRISPR as a means for genome-wide functional screens is the reason behind its success, where previously established tools like small interference RNA screens induce more off-target effects and their interference is only temporary compared to CRISPR KO screens[74]. The use of CRISPR knockout screens has been well established as a means to identify druggable dependencies in various cancers *in vitro*, as well as *in vivo*[75], [76], [77], [78], [79].

Until relatively recently, genome-wide CRISPR library screens were only designed to work with *Streptococcus pyogenes* (SpCas9), mainly because it has a relatively high on-target editing efficiency and is not very restrictive in what PAM recognition sites it binds to (5'-NGG-3'). However, recent studies have shown that modifications to

Acidaminococcus sp. Cas12a (formerly known as Cpf1), referred to as enhanced AsCas12a (enAsCas12a), can advance its functionality to a similar level as SpCas9[80]. By single amino acid substitutions, the PAM target range of Cas12a was expanded beyond TTTV, enabling higher on target editing efficiency.

Cas12a has garnered significant attention as a CRISPR enzyme for genome-wide screens, primarily due to its ability to facilitate multiplexed targeting with relative ease[80]. The key advantage lies in the direct repeat (DR) sequences located between guide RNAs (gRNAs) and are only 20 base pairs long. Minor modifications to these DR regions have been shown to preserve the binding affinity of Cas12a while reducing sequence homology [79]. This reduction in homology is critical when multiple gRNA cassettes, each with distinct interspersed DR sequences, are assembled onto a single plasmid, thereby preventing unwanted bacterial recombination during lentiviral production. These DR modifications enable Cas12a to introduce multiple genome edits within a single cell using different gRNAs delivered on a single plasmid. This multiplexing approach, combined with the enhanced functionality of enAsCas12a, has already been successfully implemented in genome-wide CRISPR libraries Humagne A and B, followed by the slightly more refined versions Humagne C and D [80]. Each of these libraries comprises approximately 22,000 unique pairs of gRNAs delivered on a single construct, with each pair designed to target a distinct gene in the human genome. This heightened efficiency minimises the number of cells required for experimental coverage, streamlining the scale of CRISPR screens and making such experiments become more manageable and less labour-intensive [80].

As demonstrated by the Cancer Dependency Project (DepMap), a database harnessing publicly available data from over a thousand genome-wide CRISPR screens in different cancer cell lines, CRISPR screens enable the systematic interrogation of gene function at an unprecedented scale and have fundamentally transformed functional genomics[82]. Analysis from individual CRISPR screen as well as large data sets, like DepMap, are delivering discoveries that change our understanding of specific types of cancers and their dependencies. Specifically, a recent study from the Garnett group (Wellcome Sanger Institute, University of Cambridge, UK) integrated 930 genome-wide CRISPR screens from the DepMap data set, targeting 17,647 genes in cell lines representing 27 different cancer types[83]. Based on clinically informed dependency-marker associations with protein-protein interactions, they proposed 370 anti-cancer targets, many of which are novel to the cancer of interest[83]. More specifically, one study that utilised CRISPR screen data from DepMap revealed that loss of Werner syndrome ATP-dependent helicase (WRN) induces synthetic lethality in mismatch repair-deficient/microsatellite instability-high (dMMR/MSI-H) colorectal cancer cells[84]. This WRN dependency was confirmed across 60 diverse preclinical models, including those resistant to targeted therapies, chemotherapy, and immunotherapy, and highlighting WRN as a nearly universal therapeutic target in these tumours [83].

Although the previous exemplifies the potential power of genome-wide CRISPR screens, they come with limitations. First, CRISPR screens present with significant levels of noise that can only partially be reduced and, second, the prioritisation of hit selection is not systematically defined, leaving researchers aiming to correct for, and select based on, different variables that impact screen results [84], [82], [85].

1.4.2 Genome-wide CRISPR screens in MPNST cell line models

To date, no overarching study has been published on the use of genome-wide CRISPR knockout screens across multiple MPNST cell line models to identify common genetic dependencies. Moreover, synthetic lethality screens in isogenic MPNST cell lines, where specific driver mutations are systematically introduced or corrected, have never been conducted. Given the aggressive nature of MPNST and the paucity of effective treatments, there is an urgent need to explore these approaches to develop new interventions for these patients.

1.5 Aim of the study

The overall goal of this research is to optimise and apply advanced CRISPR screening methods, particularly using enAsCas12a-mediated multiplexed CRISPR libraries Humagne C and D, to investigate gene essentiality and synthetic lethality in MPNST cell line models. Through this, I aim to uncover potential novel therapeutic targets and improve the efficiency by which CRISPR screens identify such hits in both traditional and alternative media environments.

The specific objectives of this thesis are:

1. Optimise and validate enAsCas12a CRISPR screen workflow:

- Develop a genome-wide enAsCas12a CRISPR knockout screen workflow that integrates and, where possible, adjusts for key variables that influence screen outcomes.
- Systematically assess how these variables impact screen quality and the reliability of hit identification, improving our understanding of CRISPR screen results and the reproducibility across experiments.

2. Investigate differential dependencies in different cell culture media:

- Evaluate how the cellular growth environment affects gene essentiality by conducting parallel enAsCas12a CRISPR screens in DMEM and Human Plasma-Like Medium (HPLM).
- Identify candidate medium-specific essential genes and investigate how different metabolic contexts alter genetic dependencies, as reflected by cellular whole-transcriptomics.

3. Identify potential genetic dependencies and synthetic lethality targets in MPNST models:

- Identify and prioritise dependencies across MPNST cell line models, contexts and genome-wide enAsCas12a-based CRISPR libraries.
- Focus on discovering synthetic lethality relationships in immortalised Human Schwann cell and MPNST cells, particularly those with PRC2 loss, to identify potential therapeutic targets.

Chapter 2: Materials and Methods

2.1 Reagents and resources

Table 2.1. Materials and Reagents

Reagent	Description	Source/Catalogue Number
1X PBS	Phosphate-buffered saline	Sigma Life Sciences
2XYT Medium	Growth medium for bacterial culture	Made in house
AgeI-HF	High-fidelity restriction enzyme	NEB: R3552S
Ampicillin	Antibiotic used for bacterial selection	Thermo Fisher: CAS # 69-53-4
AO-DAPI Solution 13	Stain for cell counting and viability	Chemometec: 910-3013
Blasticidin	Antibiotic for selection of transfected cells	Invivogen: ant-bl-05
Bovine Serum Albumin (Blocking buffer)	Protein for blocking nonspecific binding	Merck: A9418
Bovine Serum Albumin (Standard)	Protein standard for assays	Thermo Fisher: 23209
Bradford Assay Reagent	Protein quantification reagent	Thermo Scientific: 23246
BrightComp eBeads™ Compensation Bead Kits	Reagent for FACS compensation	Invitrogen: A10514
Chemiluminescent detection reagents	Reagents for visualising Western blots	Merck: GERPN2235
Complete™, EDTA-free Protease Inhibitor Cocktail	Protease inhibitor for lysis buffer	Merck: 04693116001
CutSmart Buffer	Buffer for restriction enzyme digestion	NEB: B6004
DH5α competent cells	Competent cells for transformation	NEB: C2987H
DMEM	Dulbecco's Modified Eagle Medium for cell culture	Merck: D6429-24X500ML
DMSO	Cryoprotectant for cell freezing	Merck: D2650

EcoRI-HF	High-fidelity restriction enzyme	NEB: R3101S
EDTA	Chelating agent in FACS buffer	Merck: E8008
FastDigest Green Buffer	Buffer for restriction enzyme digestion	Thermo Fisher: B72
FBS	Fetal bovine serum	Merck: F2442
Gene Pulser/MicroPulser Electroporation Cuvettes, 0.1 cm gap	Cuvettes for electroporation	Bio-Rad: 1652089
Gibson Assembly Master Mix	Reagent for DNA assembly	NEB: E2611S
HCl	Reagent for pH adjustment	Merck: 2104-50ML
HPLM	Human plasma-like medium	Thermo Fisher: A4899101
LB Agar	Solid medium for bacterial culture	Made in house
LB Broth	Luria-Bertani broth for bacterial culture	Made in house
Mag-Bind® Blood & Tissue DNA HDQ 96 Kit	Kit for genomic DNA extraction	Omega Bio-Tek: M6399-00S
MegaX competent cells	Competent cells for high-efficiency transformation	Thermo Fisher: C640003
NaOH	Reagent for pH adjustment	Merck: 1055902500
NEB HiFi DNA assembly reaction	Reagent for high-fidelity DNA assembly	NEB: E2621S
NEBNext® Ultra™ II Q5® Master Mix	PCR reagent for high-fidelity amplification	NEB: M0544X
NotI	Restriction enzyme for DNA digestion	Thermo Fisher: FD0596
Parafilm	Sealing film for plates	Merck: HS234526B
PBS	1X Phosphate-buffered saline (PBS)	Thermo Fisher: 10010001
Penicillin	Antibiotic used in bacterial transformation and culture	Thermo Fisher: 11593027
Penicillin/Streptomycin	Antibiotic mix for cell culture	Thermo Fisher: 15140-122
PFA (Paraformaldehyde)	Fixative for immunostaining	Merck: 158127
Pierce Detergent Compatible Bradford Assay Kit	Protein quantification kit	Thermo Scientific: 23246
Polyethylenimine (PEI)	Transfection reagent	CAS # 9002-98-6
Polyvinylidene difluoride (PVDF) membranes	Membranes for protein transfer	Bio-Rad: 1704273
Precision Plus Protein Dual Color Standards	Protein ladder for SDS-PAGE	Bio-Rad: 1610394
RPMI	Roswell Park Memorial Institute, cell culture medium	Merck: R8758
Sall	Restriction enzyme for DNA digestion	Thermo Fisher: FD0644
SDS	Anionic surfactant used in SDS-PAGE	Bio-Rad: 1610732

SOC Medium	Medium for bacterial recovery after transformation	Made in house
Sodium Azide	Reagent for FACS buffer preparation	Merck: S2002
STBL3 competent cells	Competent cells for transformation	Thermo Fisher: C737303
T175 flasks	Cell culture flasks	Scientific Laboratory Supplies: 431080
T4 DNA Ligase	Enzyme used for DNA ligation	NEB: M0202S
TGX Precast 4-20% Gel	SDS-PAGE gel for protein separation	Bio-Rad: 5671094, 5671095
Triton X-100	Non-ionic surfactant for cell lysis	Merck: X100
TrypLE Express	Enzyme for cell detachment	Thermo Fisher: 12604-013
Turbo Transfer system	Equipment for Western blot protein transfer	Bio-Rad: 1704150
Tween-20	Non-ionic surfactant	Merck: 11332465001
UNC7700	Small molecule inhibitor for cell treatment	Collaborator

Table 2.2. Plasmids used in this study.

Plasmid Name	Recombinant DNA
<i>pRDA_052</i>	<i>EF-1α-PuroR</i>
<i>pRDA_174</i>	<i>enAsCas12a-3XHA-BSD</i>
<i>pRDA-SUZ12F86A-F90A</i>	<i>pRDA + SUZ12 F86A-F90A mutant</i>
<i>pRDA-SUZ12H466A</i>	<i>pRDA + SUZ12 H466A mutant</i>
<i>pRDA-SUZ12R196A</i>	<i>pRDA + SUZ12 R196A mutant</i>
<i>pRDA-SUZ12S583A</i>	<i>pRDA + SUZ12 S583A mutant</i>
<i>pRDA-SUZ12WT</i>	<i>pRDA + SUZ12 Wild Type (WT)</i>
<i>pRDA174-enAsCas12a-EGFP</i>	<i>enAsCas12a-T2A-EGFP</i>

Table 2.3. Commercial kits used in this study

Kit name	Description	Source/Catalogue Number
Mag-Bind® Blood & Tissue DNA HDQ 96 Kit	Kit for genomic DNA extraction	Omega Bio-Tek: M6399-00S
Qiagen Miniprep Kit	Kit for plasmid DNA isolation (small scale)	Qiagen: 27104
Qiagen Plasmid Maxi Kit	Kit for plasmid DNA isolation	Qiagen: 12162
Qiagen Plasmid Plus Midi Kit	Kit for plasmid DNA isolation	Qiagen: 12941
Qubit dsDNA Quantification Assay Kit	Reagent for DNA quantification	Invitrogen, Catalog number: Q32851
Zymo Gel DNA Recovery Kit	Kit for DNA purification from agarose gels	Zymo Research: D4001

Table 2.4. Software used in this study

Software/Library	Description	Source
AlphaFold2	Predicts protein structures, used for visualising specific protein domains.	[87]
AnnotationDbi	R package for annotation databases.	Bioconductor
apeglm	R package for statistical analysis of gene expression data.	Bioconductor
Burrows-Wheeler Aligner (BWA)	Aligns sequencing reads to the reference genome.	[88]
Chronos	Processes CRISPR screen data to calculate and correct gene effect scores.	[85]
circlize	R package for circular visualization of data.	Bioconductor
cluster	R package for cluster analysis.	Bioconductor
cowplot	R package for combining plots.	Bioconductor
DESeq2	Differential gene expression analysis.	[89]
devtools	R package for package development tools.	Bioconductor
dplyr	R package for data manipulation.	Bioconductor
edgeR	R package for differential expression analysis of RNA-seq data.	Bioconductor
EnhancedVolcano	R package for creating enhanced volcano plots for differential expression results.	Bioconductor
EnsDb.Hsapiens.v79	R package for genomic annotation using Ensembl data.	Bioconductor
FastQC	Quality control for raw sequencing data.	[90]
FeatureCounts	Gene quantification for RNA sequencing.	[91]
FlowJo (10.10.0)	Software for analysis of flow cytometry data.	[92]
FSA	R package for fisheries stock assessment, includes Dunn's test.	Bioconductor
Genome Analysis Toolkit (GATK)	Whole-genome sequencing analysis, including alignment, recalibration, and variant calling.	[93]
ggalluvial	R package for creating alluvial diagrams.	Bioconductor
ggcorrplot	R package for creating correlation matrix plots.	Bioconductor
ggplot2	R package for data visualization, creating a wide range of static and interactive graphics.	Bioconductor
ggrepel	R package for creating better labels in ggplot2.	Bioconductor
ggsignif	R package for adding significance levels to plots.	Bioconductor
ggvenn	R package for generating Venn diagrams using ggplot2.	Bioconductor
Glimma	R package for interactive visualization of gene expression data.	Bioconductor
GraphPad Prism	Software for statistical analysis, graphing, and data visualization.	https://www.graphpad.com/
GSEA (Gene Set Enrichment Analysis)	Functional enrichment analysis of differentially expressed genes.	[94]

hexbin	R package for hexagonal binning of data.	Bioconductor
IHW	R package for independent hypothesis weighting.	Bioconductor
MAGeCK	CRISPR screen data analysis, including gRNA count normalization and fold change calculation.	[95]
nortest	R package for normality tests.	Bioconductor
org.Hs.eg.db	R package with gene annotation data for Homo sapiens.	Bioconductor
pheatmap	R package for creating heatmaps.	Bioconductor
Picard tools	Marks duplicate reads in sequencing data.	[93]
Python	Programming language used for data analysis and bioinformatics.	https://www.python.org/
R	Statistical computing and graphics programming language.	https://www.r-project.org/
reshape2	R package for reshaping data.	Bioconductor
STAR	Aligns RNA sequencing reads to the reference genome.	[96]
tidyr	R package for data tidying.	Bioconductor
tidyverse	Collection of R packages for data science, including dplyr and ggplot2.	Bioconductor
Trimmomatic	Adapter removal and trimming of low-quality reads from RNA sequencing data.	[97]
UpSetR	R package for visualising intersecting sets, particularly for complex data sets.	Bioconductor
VennDiagram	R package for generating Venn diagrams.	Bioconductor
vsn	R package for variance stabilization normalization of microarray data.	Bioconductor

Table 2.5. External datasets used for this study.

Dataset	Source
Common Essential Genes	[98]
DepMap gRNA Count Matrix	DepMap; [98]
Druggable Genome Gene List	[100]
FDA Approved Druggable Gene List	FDA; [100]
MPNST Tumour Sample RNA-Seq Data	DbGap; [45]
Non-Essential Genes	DepMap; [101]
Predicted Synthetic Lethality Pairs	SynLeth Database 2.0I; [102]
Screen Gene Effect	DepMap; [98]

2.2 Molecular biology methods

2.2.1 Bacterial transformation

Competent cells (50 μ L) were combined with 1-2 μ L of plasmid DNA in a chilled tube and incubated on ice for 30 minutes. The mixture was subjected to a heat shock at 42°C for 1 minute, then returned to ice for 3 minutes. Following this, 100 μ L of LB broth with 100 μ g/mL penicillin was added, and the mixture was gently mixed. The cells were plated on LB agar containing 100 units/mL penicillin and incubated upside down overnight at 37°C, sealed with Parafilm.

2.2.2 Cloning pRDA174-enAsCas12a-EGFP

A *EGFP* insert, flanked by *NotI* and *Sall* restriction enzyme sites, was obtained from a DNA block (Gene block: *NotI*-*EGFP*-*Sall*, Table 2.6) and cloned into the *pRDA_174* vector, which was similarly linearised with *NotI* and *Sall*. Digestion reactions were performed in a total volume of 50 μ L, using 1 μ g of DNA, 1X Buffer, and 10 units each of *NotI* and *Sall* restriction enzymes, and incubated at 37°C for 2 hours, followed by heat inactivation at 65°C for 15 minutes. The digested products were separated by gel electrophoresis on a 1% agarose gel, and the appropriate bands were purified using a Zymo Gel DNA Recovery Kit. The purified *EGFP* insert was ligated into the linearised *pRDA_174* vector using T4 DNA Ligase (NEB) in various insert-to-vector ratios (1:2, 1:3, 1:5) in a total volume of 20 μ L and incubated at 16°C for 17 hours. The ligation products were transformed into STBL3 competent cells, and the bacteria were plated on LB agar containing 100 μ g/mL ampicillin. Colonies were picked, grown in LB medium with ampicillin overnight, and plasmid DNA was isolated using the Qiagen Plasmid Plus Midi Kit. Positive colonies were screened by Sanger sequencing.

Table 2.6. Gene Blocks used in this study

Name	Sequence
<i>Not1-EGFP-Sall</i>	GATGTCCCCGACTATGCCGCGGCCGCAGAGGGCAGAGGAAGTCTGCTAACAT GCGGTGACGTCGAGGAGAATCCTGGCCCAGTGAGCAAGGGCGAGGAGCTGTT CACCGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCAC AAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGA CCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCCGTGCCCTGGCCCACCCT CGTGACCACCCTGACCTACGGCGTGACGTGCTTCAGCCGCTACCCCGACCAC ATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGG AGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGT GAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGAC TTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACATAACA GCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAA CTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGACGTGCGCCGACCAC TACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCA CTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGAT CACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGG ACGAGCTGTACAAGGAATTCTAATAAGTCGACAATCAACCT
<i>AgeI-EGFP_targeting_gRNAs-EcoRI</i>	CTTTATATATCTTGTGGAAAGGACGAAACACCGGTAATTTCTACTCTTGTAG ATCGTCGCGCTCCAGCTCGACCAGGTAATTTCTACTATCGTAGATCTCAGGG CGGACTGGGTGCTCAGTTTTTTGAATTCGCTAGCAAGCTTGGCGTAACTAGA TCTTGA
<i>T2A-BSD</i>	ACAAGGACGACGACGACAAGGCGGCCGCAGAGGGCAGAGGAAGTCTGCTAAC ATGCGGTGACGTCGAGGAGAATCCTGGCCCAATGGCCAAGCCTTTGTCTCAA GAAGAATCCACCCTCATTGAAAGAGCAACGGCTACAATCAACAGCATCCCCA TCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCTCTCTAGCGACGGCCGCAT CTTCACTGGTGTCAATGTATATCATTTTACTGGGGGACCTTGTGCAGAACT CGTGGTGCTGGGCACTGCTGCTGCTGCGGCAGCTGGCAACCTGACTTGTATC GTCGCGATCGGAAATGAGAACAGGGGCATCTTGAGCCCCTGCGGACGGTGCC GACAGGTGCTTCTCGATCTGCATCCTGGGATCAAAGCCATAGTGAAGGACAG TGATGGACAGCCGACGGCAGTTGGGATTCGTGAATTGCTGCCCTCTGGTTAT GTGTGGGAGGGCTAAGTCGACAATCAACCTCTGGATTACA

2.2.3 Cloning pRDA052_EGFP_targeting_gRNAs

EGFP targeting gRNAs were designed using CRISPOR [103], resulting in the following two guide sequences: 5'-TCTCAGGGCGACTGGTGCTCAG-3' and 5'-CGTCGCGTCACGTCGACCAGG-3'. A gene block (*AgeI-EGFP_targeting_gRNAs-EcoRI*) containing these gRNAs was synthesised at IDT (www.IDT.com). The pRDA_052 plasmid and gene block were digested with AgeI-HF and EcoRI-HF (NEB) in a 20 µL reaction containing 1.5 µL plasmid DNA (~400 ng/µL), 0.5 µL each of AgeI-HF and EcoRI-HF, 2 µL CutSmart buffer, and 15.5 µL H₂O. The digested product was

run on an agarose gel, and the desired bands were extracted using a Zymo Gel DNA Recovery Kit, yielding 30 μ L at 115 ng/ μ L. Gibson assembly was performed using 100 ng of the digested pRDA_052 backbone and the gene block insert at backbone-to-insert ratios of 1:1 and 1:3. The assembly was carried out in a 15 μ L reaction containing 5 μ L NEB Gibson Assembly Master Mix, incubated at 50°C for 15 minutes. Transformation was performed by mixing 2.5 μ L of the Gibson assembly reaction with 50 μ L of competent cells, followed by plating on LB agar with 100 μ g/mL ampicillin. Colonies were picked 24 hours later, grown in LB overnight, and plasmid DNA was isolated using a Qiagen Miniprep Kit. Colonies were screened via Sanger sequencing.

2.2.4 Replacing mCherry with BSD in SUZ12 variant constructs

Five plasmids (*pRDA174_SUZ12WT / F86A-F90A / R196A / H466A / s583A*) were digested with NotI and Sall. Each plasmid was digested in a 50 μ L reaction mixture containing approximately 1000 ng of plasmid DNA, 1 μ L each of NotI and Sall FastDigest enzymes (Thermo Fisher), 25 μ L H₂O, and 3 μ L FastDigest Green Buffer. Incubations were carried out at 37°C for 1 hour, followed by gel electrophoresis (1% agarose) for 40 minutes at 130V. The DNA bands corresponding to the digested backbones were excised, and DNA was extracted using the Zymo Gel DNA Recovery Kit. For each plasmid, a 10 μ L NEB HiFi DNA assembly reaction was set up using the purified backbone and a *T2A-BSD* gene block with compatible ends for a DNA assembly reaction. Reactions were incubated at 50°C for 1 hour. Transformation was performed by mixing 2.5 μ L of the HiFi DNA assembly reaction with 50 μ L of competent DH5 α cells, followed by incubation on ice for 30 minutes, heat shock at 42°C for 1 minute, and recovery in 150 μ L LB medium. The mixture was plated on LB agar containing 100 μ g/mL ampicillin and

incubated overnight at 37°C. Colonies were picked, grown in LB medium with ampicillin, and plasmid DNA was isolated using the Qiagen Plasmid Plus Midi Kit. Positive colonies were screened by Sanger sequencing, confirming correct constructs without mutations in the area of interest.

2.2.5 CRISPR library amplification

CRISPR libraries Humagne C and D were diluted to achieve a concentration of 40 ng/ μ L in a final volume of 50 μ L. For each electroporation, 5 μ L (200 ng) of the diluted CRISPR pDNA library was added to 35 μ L of MegaX competent cells. Electroporation was performed in 0.1 cm cuvettes at 2.0 kV, 200 ohms, and 25 μ F. After electroporation, 465 μ L of SOC medium was immediately added, and the mixture was transferred to a round-bottom tube containing 2.5 mL of prewarmed SOC medium. This process was repeated three times for each library. The cells were then incubated in SOC medium at 37°C with shaking for 1 hour. Following recovery, the cultured cells were combined into a sterile tube, gently mixed, and a small sample was diluted and plated to estimate colony counts. The remaining recovered culture was added to 500 mL of 2XYT + AMP and incubated overnight at 37°C in a shaking incubator. The next day, colonies were counted, with 333 colonies for Library C and 402 colonies for Library D, both above the generally recommended threshold of 100 colonies per construct proceeding with library DNA purification[105]. Isolation and purification of the amplified CRISPR libraries was performed by using a Qiagen Plasmid Maxi Kit on the cultured bacteria.

2.2.6 CRISPR library validation

The sgRNA region was amplified using primers designed for sequencing with 41 bp cycles, requiring one of the sgRNA pairs to be within 41 bp of the Illumina adapter.

Primers with known annealing sites (P5_ARGON and P7_KERMIT) were used, incorporating a stagger region for library diversity and an index for sample barcoding. The sequences for the forward and reverse primers used in this process are provided in Table 2.7. Following PCR (Table 2.8), DNA products were analysed using gel electrophoresis on a 2% agarose gel to confirm the presence of the expected 350 bp bands. Sanger sequencing of the isolated bands verified the correct stagger and index regions corresponding to each sample. The validated samples were then quantified using Qubit for Illumina sequencing. The validated libraries were then used for lentiviral production and subsequent CRISPR screens.

Table 2.7. Primers used for CRISPR library validation

Primer Name	Sequence (5' to 3')
P5_ARGON_1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5_ARGON_2	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5_ARGON_3	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5_ARGON_4	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P7_KERMIT_1	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_KERMIT_2	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_KERMIT_3	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_KERMIT_4	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTG

Table 2.8. PCR programme used for CRISPR library validation

Stage	Temperature (°C)	Time	Number of Cycles
Initial Denaturation	98	03:00	1x
Denaturation	98	00:15	5x (Stage B)

Annealing	65	00:30	5x (Stage B)
Extension	72	00:30	5x (Stage B)
Denaturation	98	00:15	8x (Stage C)
Annealing	63	00:30	8x (Stage C)
Extension	72	00:30	8x (Stage C)
Denaturation	98	00:15	12x (Stage D)
Annealing	61	00:30	12x (Stage D)
Extension	72	00:30	12x (Stage D)
Final Extension	72	07:00	1x
Hold	4	Infinity	1x

2.3 Mammalian cell culture methods

2.3.1 Cell lines and cell culture methods

The mammalian cell lines used in this study are listed in Table 2.9. MPNST cell lines STS26T, ST88-14 and 90-8 were kindly provided by Professor Eduard Serra (Germans Trias i Pujol Research Institute, Barcelona, Spain), S462TY and the immortalised Human Schwann Cell line (iHSC; ipn02.3 1-λ) by Professor Nancy Ratner (Cincinnati Children's Hospital Medical Center, Cincinnati, USA), and LentiX Hek293T cell line by Professor Carvalho (University of Oxford, UK). Specifically, the iHSC line used in this study was made from the same population of Schwann cells as ipn02.3 2λ, developed by the group of Professor Wallace (University of Florida), but transduced with 1 instead of 2 μL of each virus (retro and lentivirus) carrying the *hTERT* and *mCdk4* immortalisation genes [105]. Cell lines tested negative for mycoplasma contamination by polymerase chain reaction (PCR) using primers 5'-YGCCTGVGTAGTAYRYWCGC-3' and 5'-GCGGTGTGTACAARMCCCGA-3'[107]. Cells were cultured in DMEM complemented with 10% FBS and Penicillin/Streptomycin (100 Units/mL) and passaged by first washing with Dulbecco's phosphate-buffered saline

(DPBS) (D8573-500 mL Sigma Life Sciences) and then treatment with TrypLE Express (CAT #: 12604-013, ThermoFisher Scientific) when reaching a confluency of 70-90%. Counting cells was performed with AO-DAPI Solution 13 following manufacturer's instructions. To quantify population doublings, a standard curve was established correlating percentage of cell growth to population doublings, where the initial time point is defined as 100% at 0 population doublings. Each subsequent doubling increases the population count by a factor of two, reflecting a linear progression of 1 population doubling per doubling in percentage. After each passage, the percentage increase in cell count relative to the previous time point was measured and converted into population doublings using this standard curve. For long-term storage, cells were frozen in approximately 5% DMSO in standard medium at -1 °C per second until -80 °C.

2.3.2 Cell line authentication

Target cell lines (Table 2.9) were authenticated by STR profiling at Eurofins Genomics using a panel of 16 loci, including the Amelogenin marker for sex determination (Table 2.10). Genomic DNA was extracted, and the STR profiles were generated using PCR amplification followed by capillary electrophoresis. The resulting profiles were compared against reference databases to confirm the identity of each cell line. All cell lines were verified as matching their expected profiles on Cellosaurus [108].

Table 2.9. Cell lines used in this study.

Cell line	Origin, Disease	Standard medium*	Reference
STS26T	Human melanoma, sporadic	DMEM + 10% FBS + Pen/Strep	[109]
ST-8814	Human MPNST, NF1	DMEM + 10% FBS + Pen/Strep	[110]
90-8	Human MPNST, NF1	DMEM + 10% FBS + Pen/Strep	[111]
S462TY	Human MPNST, NF1	DMEM + 10% FBS + Pen/Strep	[112]

iHSC	Human, no disease	DMEM + 10% FBS + Pen/Strep	[113]
LentiX Hek293T	Human Embryonic Kidney	DMEM + 10% FBS + Pen/Strep	Takara Bio, CAT # 632180

* DMEM: Dulbecco's Modified Eagle Medium, Pen/Strep: 100 Units/mL Penicillin Streptomycin (CAT #, 15140-122, ThermoFisher Scientific), FBS: fetal bovine serum (CAT #: F2442, Merck), RPMI: Roswell Park Memorial Institute (CAT #: R8758, Merck)

Table 2.10. Cell line authentication

Marker	STS26T	90-8	ST-8814	s462TY	iHSC	LentiX Hek293T
AMEL	X,X	X,X	X,Y	X,X	X,X	X,X
CSF1PO	10,13	12,12	9,12	12,12	10,11	11,12,13
D13S317	9,10	8,8	12,12	12,12	12,13	12,13,14,15
D16S539	12,13	9,9	13,13	11,13	12,12	9,14,13
D18S51	17,18	14,18	12,12	16,16	13,16	17,18
D19S433	14,14	13,16	13,14	14,14	14,14	17,18,19
D21S11	30,31	30,30	29,32.2	29,29	30.2,32.2	28,29,30.2,31.2
D2S1338	20,20	16,24	17,23	23,23	16,17	19,19
D3S1358	14,14	15,15	15,18	14,17	17,18	15,16,17
D5S818	11,12	11,11	12,13	12,12	8,11	8,9
D7S820	8,11	11,11	8,8	8,10	8,12	11,11
D8S1179	13,14	10,12	14,14	12,12	12,15	11,12,13,14,15
FGA	22,23	25,25	21,21	20,20	23,24	23,23
TH01	6, 9.3	6,6	9,9	7,8	6,7	7,9.3
TPOX	8,8	11,11	11,12	8,8	8,12	11,11
vWA	17,17	17,17	16,16	19,19	15,17	16,17,18,19,20

2.3.3 Puromycin and blasticidin kill curves

Cells were seeded on 96-well plates with a density of 1E4 cells per well and exposed to different concentrations of puromycin (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 µg/mL) and blasticidin (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20 µg/mL) with two or three replicates per condition. The cells were monitored for survival/death using microscopy over the course of five to seven days. The lowest concentration of puromycin or blasticidin that killed all the cells was used for further experiments (Table 2.11).

Table 2.11. Effective killing concentrations per cell line

Cell Line	Puromycin (µg/mL)	Blasticidin (µg/mL)
STS26T	1.7	6
90-08	1.7	2
ST-8814	3	4
s462TY	3	3
iHSC	1.7	3

2.3.4 UNC7700 treatment of iHSC cells

iHSC cells were treated with 10 µM UNC7700, achieved by adding a 1:1000 dilution of a 10 mM UNC7700 stock solution to the cell culture. A corresponding volume of 1:1000 DMSO was used for vehicle controls. The treatment medium, including either the drug or the vehicle, was replaced at least every four days. For RNA sequencing, cells were harvested, followed by centrifugation and washing with ice-cold PBS. The cell pellets were then snap-frozen using a dry ice-ethanol bath and stored at -80°C.

2.3.5 Lentivirus Production via PEI-Mediated Transfection

LentiX HEK293FT cells were cultured to 70-80% confluence in T175 flasks and transfected with a plasmid mixture containing 2.64 µg of pMD2.G, 5.28 µg of psPAX2, and 10.56 µg of the lentiviral transfer plasmid in 506.9 µL of serum-free DMEM. The plasmid solution was combined with 151.7 µL of PEI solution, vortexed, and incubated at room temperature for 10 minutes to allow for complex formation. The transfection mixture was then added to the cells in 25 mL of DMEM supplemented with 10% FBS. After 24 hours, the medium was replaced with fresh DMEM containing 10% FBS. Forty-eight hours post-transfection, the lentiviral supernatant was harvested by collecting and snap-freezing the supernatant for subsequent use.

2.3.6 Lentiviral Transduction

For general lentiviral transduction, target cells were seeded and transduced in suspension to achieve a transduction efficiency of approximately 30%. Depending on transgene different selection strategies were employed. For antibiotic resistance genes, the medium was replaced 24 hours post-transduction and puromycin or blasticidin added at a predetermined concentration depending on the cell line. Cells were exposed to the selection antibiotic for 72 hours after the selection pressure was removed and cells were expanded further for later experiments. Cells that received a fluorescence marker on the transgene were first expanded and then sorted for positives based on FACS. These populations were then expanded and monitored using FACS to evaluate whether they maintain sufficient expression over time.

2.3.7 CRISPR system validation

STS26T and STS26T-enAsCas12a-EGFP cells were seeded in 96-well plates at an estimated 50% confluency. Polybrene was added to each well at a final concentration of 10 µg/mL, followed by transduction with varying volumes (50, 25, 10, 5, 2.5, 1, 0 µL) of lentivirus containing the pRDA_052_EGFP_sgRNA plasmid, performed in replicates. Three replicate wells with STS26T-enAsCas12a-EGFP cells were not given any lentivirus. On day 8 after transduction, cells were washed with PBS, detached with TrypLE, and resuspended in FACS buffer (PBS + 10% FBS). FACS was conducted to assess the efficiency of *EGFP* knockout, comparing the treated conditions to controls. EGFP expression data were collected and analysed in FlowJo™ (version 10.10.0) [91].

2.3.8 Determination of transduction efficiency with CRISPR library

lentivirus

HEK293FT cells were seeded at 6.1111×10^6 cells per T175 flask across various conditions to determine the transduction efficiency. Eleven T175 flasks were seeded with 6.1111×10^6 cells each and transduced in suspension where one received no lentivirus (negative controls), and the others were transduced with 50, 100, two times 150, 250, 300, 500, and two times 1000 μL of lentivirus, followed by the addition of 25 μL polybrene (final concentration 8 $\mu\text{g}/\text{mL}$). Cells were incubated for 24 hours after which they were treated with a predetermined concentration of puromycin for 72 hours. One of two flasks that received 0, 150 and 1000 μL lentivirus were not treated with puromycin in order to evaluate lentiviral toxicity. Cells that reached confluence were split as necessary to maintain optimal growth conditions. On day 4 post-transduction, transduction efficiency was determined by counting the number of viable cells in each condition. After subtracting the negative control cell count, a standard curve was generated based on the cell counts for all conditions to extrapolate the viral volume required to achieve approximately 30% transduction efficiency, which was calculated to correspond to approximately 160 μL of virus per T175 flask for all target cell lines. For the controls where cells received lentivirus (150 and 1000 μL) but no puromycin, no toxicity was observed, and these conditions grew at the same rate as determined by cell counting.

2.3.9 Genome-wide CRISPR knockout screens

This workflow largely follows a previously described genome-wide CRISPR screen workflow described by Zhang's Group (Eli and Edythe L. Broad Institute of MIT and Harvard, USA)[105]. Cryopreserved cells were thawed and cultured until they reached a

density of $200\text{--}300 \times 10^6$ cells. Cell density and viability were assessed using acridine orange and DAPI (AO-DAPI) staining. Cells were then seeded at a minimum density of 6.12×10^6 cells per T175 flask, with the number of flasks adjusted to ensure a post-selection yield of at least 500 cells per construct within the CRISPR library, in 25 mL of DMEM supplemented with 10% FBS (antibiotic-free). To optimise lentiviral transduction, 25 μL of a 10 mg/mL polybrene solution was added to each flask, resulting in a final polybrene concentration of 10 $\mu\text{g/mL}$. A predetermined volume of 160 μL of Humagne C or D lentivirus was then added to each flask to achieve a transduction efficiency of approximately 30%. After 24 hours of incubation, the transduction medium was removed and replaced with fresh standard medium (Table 2.9), containing puromycin at a concentration optimised for the specific target cell line. The cells were incubated for an additional 72 hours at 37°C to allow for the selection of successfully transduced cells. After these 72 hours, and each subsequent passage, the cells were detached, counted, and the population doublings (PDs) were calculated. A multiple of 12×10^6 cells (usually three replicates) were pelleted, washed with cold PBS and stored at -80°C for later genomic DNA (gDNA) extraction, ensuring full coverage of over 500 cells (500x coverage for these libraries is approximately 10.5×10^6 cells) per CRISPR library construct per pellet. Of the remainder, a total of 12×10^6 cells ($> 500\text{x}$ coverage) were then reseeded across 9 $\times$ T175 flasks in fresh DMEM supplemented with 10% FBS. This process was repeated at least every four days until the cells reached approximately 14 population doublings, at which point a final harvest was conducted, and the screen was concluded. For CRISPR screens in HPLM, a similar workflow was utilised, except that at the first passage, right after puromycin selection, medium was switched from standard

medium (Table 2.9) to HPLM complemented with 10% FBS and 100 Units/mL Penicillin and Streptomycin. HPLM medium was replaced with every passage as with DMEM medium.

2.4 Biochemical methods

2.4.1 Preparation of Linear Polyethyleneimine (PEI) Solution

50 mg of linear polyethyleneimine (PEI) was dissolved in 45 mL of ultrapure water. The pH was initially adjusted to below 2.0 to ensure complete dissolution, then readjusted to 7.1 using 1 M NaOH. The solution was sterile-filtered through a 0.45 µm filter, aliquoted into 1 mL portions, and stored at -20°C. This PEI solution remained stable for up to 1 year and effective through five freeze-thaw cycles.

2.4.2 Genomic DNA extraction

Genomic DNA was isolated from up to 12E6 cultured cells using the Mag-Bind® Blood & Tissue DNA HDQ 96 Kit (Omega Bio-Tek). Cells were harvested and pelleted by centrifugation, followed by a wash in cold PBS (4°C) and resuspension in 180 µL cold PBS. A mastermix containing 230 µL AL Buffer and 20 µL Proteinase K Solution per sample was prepared and added to each cell sample. The samples were vortexed for 10 minutes to ensure proper mixing, followed by incubation at 55°C for 10 minutes in a shaking water bath. Then 320 µL HDQ Binding Buffer (diluted with 100% isopropanol) and 20 µL Mag-Bind® Particles HDQ were added. After vortexing for 10 minutes, the plate was placed on a magnetic separation device to magnetise the particles, and the cleared supernatant was aspirated. The particles were washed twice with 600 µL VHB Buffer (diluted with 100% ethanol), each time vortexing for 15 seconds and separating on the magnetic device. A final wash was performed with 600 µL SPM Buffer (diluted

with 100% ethanol), followed by a step to remove residual ethanol. The DNA was eluted by adding 200 μ L of nuclease-free water, vortexing for 5 minutes, and magnetising the particles to clear the eluted DNA.

2.4.3 Whole cell extract lysis

Cells were lysed using Triton X-100 lysis buffer, consisting of 1% Triton X-100, 200 mM NaCl, and 50 mM Tris-HCl (pH 7.4), supplemented with cOmplete™, EDTA-free Protease Inhibitor Cocktail (Merck, #4693132001). The lysis buffer was prepared by combining 0.5 mL Tris-HCl, 0.4 mL of 5 M NaCl, 100 μ L Triton X-100, and 0.4 mL of the 25X protease inhibitor cocktail, and adjusting the volume to 10 mL with water. Cells were harvested from a confluent 6-well plate wells, where they were first washed with ice-cold PBS, and lysed on ice in 120 μ L of lysis buffer. The lysate was incubated on ice for 1 hour and then harvested and placed into prechilled 1.5 mL collection tubes, followed by centrifugation at 12,000 rpm for 10 minutes at 4°C to clarify the lysate. The supernatant was collected and stored on ice.

2.4.4 Protein quantification

Protein concentration was quantified using the Pierce Detergent Compatible Bradford Assay Kit (Thermo Scientific, #23246). The assay was performed following the manufacturer's protocol. Bovine Serum Albumin (BSA) standards were prepared in a series of dilutions ranging from 2000 μ g/mL to 0 μ g/mL, as outlined in the provided standard curve preparation instructions. 10 μ L of each standard and unknown sample was pipetted into a 96-well plate, followed by the addition of 300 μ L of Pierce Detergent Compatible Bradford Assay Reagent. After a 10-minute incubation at room temperature, the absorbance was measured at 595 nm using a plate reader. The protein concentration

of the samples was determined by plotting a standard curve using the BSA standards and correcting for the blank measurements using GraphPad Prism version 10.0.0 for Mac, GraphPad Software, Boston, Massachusetts USA, www.graphpad.com".

2.4.5 SDS-PAGE and Western blotting

Protein samples were prepared in 20 μ L total volume, consisting of 3X loading buffer (Bio-Rad), approximately 10 μ g of protein, and lysis buffer to adjust the final volume. The samples were denatured at 65°C for 15 minutes and spun down at maximum speed for 1 minute, after which the supernatant was loaded onto a TGX precast 4-20% gel. A precision Plus Protein Dual Color Standards ladder (7 μ L) was used as a size reference. The gel was run at 300W, 250V, and 0.06 AMP. Following Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) in in TGS running buffer (25mM Tris pH 8.3, 192 mM glycine, 0.1% (w/v) SDS; 1610732; Bio-Rad), proteins were transferred to a polyvinylidene difluoride (PVDF) membranes (1704273; Bio-Rad) using the Turbo Transfer system under high molecular weight settings (2.5A, 23V, 9 minutes). Membranes were blocked for 1 hour in either 5% BSA or 5% skim milk in PBST (1X PBS, 0.1% Tween-20) at room temperature. Primary antibodies (Table 2.12) were incubated overnight at 4°C in the appropriate blocking buffer at recommended dilutions. Following primary antibody incubation, membranes were washed three times for 10 minutes each with PBST. Secondary antibodies were applied, followed by three additional 10-minute washes. Chemiluminescent detection reagents were applied, and the membranes were imaged using a chemiluminescence imager.

Table 2.12. Antibodies.

Target protein	Type	Source	Catalogue number	Species	Dilution
----------------	------	--------	------------------	---------	----------

EED	Primary	R&D systems	AF5827	Sheep	1:200
SUZ12	Primary	Cell signalling	D39F6	Rabbit	1:1000
EZH2	Primary	Cell signalling	D2C9	Rabbit	1:1000
H3K27Me3	Primary	Cell signalling	9733S	Rabbit	1:1000
GAPDH	Primary	Cell signalling	D4C6R	Mouse	1:1000
Anti sheep	Secondary	Abcam	ab6746	Rabbit	1:10000
Anti rabbit	Secondary	Cell signalling	7074	Goat	1:10000
Anti Mouse	Secondary	Cell signalling	7076	Horse	1:10000

2.4.6 Preparation of 4% paraformaldehyde (PFA) solution

A 4% paraformaldehyde (PFA) solution was prepared by adding 800 mL of 1X PBS to a glass beaker on a stir plate in a ventilated hood and heating to approximately 60°C. Forty grams of paraformaldehyde powder was added to the heated PBS, and the pH was slowly raised by adding 1 N NaOH dropwise until the solution cleared. Once fully dissolved, the solution was cooled, filtered, and the volume adjusted to 1 L with 1X PBS. The pH was rechecked and adjusted to approximately 6.9 with dilute HCl. The final solution was aliquoted and frozen at -20°C.

2.4.7 Immunolabelling for Flow cytometry

FACS buffer was made by mixing 50 mL FBS, 1.25 mL 2 M EDTA, and 1.25 mL 1 M sodium azide with 947.75 mL PBS. Approximately 7E6 cells were harvested per cell line, pelleted by centrifugation, washed twice with PBS, and resuspended in FACS buffer. Fixation was achieved by resuspending the cells in 100 µL of 4% paraformaldehyde (PFA) per 1 million cells, incubating for 15 minutes at room temperature, and subsequently washing twice with PBS. For cell permeabilization, a buffer containing 0.3% Triton X-100 in FACS buffer was prepared and added to the cells with 100 µL per sample. The cells requiring primary antibody staining were resuspended in 100 µL of FACS buffer with

0.1% Triton X-100, complemented with 1:100 dilution of primary antibody for 1 hour at room temperature in the dark. After incubation, cells were washed twice with FACS buffer. Secondary antibody staining was performed by resuspending cells in 100 μ L of FACS buffer containing a 1:200 dilution of secondary antibody and incubating for 30 minutes at room temperature in the dark. The cells were washed twice more with FACS buffer, resuspended in PBS, and analysed via flow cytometry using BrightComp eBeads™ Compensation Bead Kits (Invitrogen, Catalogue number A10514) to ensure optimal flow cytometry compensation for accurate analysis.

2.4.8 Preparation of Linear Polyethyleneimine (PEI) Solution

50 mg of linear polyethyleneimine (PEI) was dissolved in 45 mL of ultrapure water. The pH was initially adjusted to below 2.0 to ensure complete dissolution, then readjusted to 7.1 using 1 M NaOH. The solution was sterile-filtered through a 0.45 μ m filter, aliquoted into 1 mL portions, and stored at -20°C. This PEI solution remained stable for up to 1 year and effective through five freeze-thaw cycles.

2.4.9 PCR amplification of gRNA cassette region, isolation and NGS-sequencing.

Genomic DNA (gDNA) was extracted from CRISPR screen samples as previously described. PCR amplification of the gRNA regions was performed using the NEBNext® Ultra™ II Q5® Master Mix (NEB), optimised for high-fidelity amplification. The PCR master mix for each sample included 10 μ M concentrations of both forward and reverse primers, which were designed to be compatible with next-generation sequencing (NGS) workflows, allowing for efficient pooling (Table 2.13). The master mix also included 1X NEB Q5 High-Fidelity Master Mix and nuclease-free water. All available gDNA per sample

was used, with a maximum of 20 µg of gDNA per 100 µL reaction volume. The total reaction volume was a multiple of 100 µL, based on the amount of gDNA available for each sample. These reactions underwent a touchdown PCR programme (Table 2.14). Post-PCR, samples were run on a 1% agarose gel and appropriate bands isolated using a Zymoclean Gel DNA Recovery Kit (Zymo Research). Samples were quantified using a Qubit™ dsDNA Quantification Assay Kit (Invitrogen™, Catalogue number: Q32851) and sequenced at 500 reads per library construct (approximately 10.5E6 reads per sample) on an Illumina platform using 150 base pair paired-end reads.

Table 2.13. Next-Generation-Sequencing primers.

Name	Type	Sequence
P7_KER MIT_1	i7	CAAGCAGAAGACGGCATAACGAGATTCGCCTTAGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_KER MIT_2	i7	CAAGCAGAAGACGGCATAACGAGATCTAGTACGGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_KER MIT_3	i7	CAAGCAGAAGACGGCATAACGAGATGCTCAGGAGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_KER MIT_4	i7	CAAGCAGAAGACGGCATAACGAGATAGGAGTCCGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_KER MIT_5	i7	CAAGCAGAAGACGGCATAACGAGATTTCTGCCTGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_KER MIT_6	i7	CAAGCAGAAGACGGCATAACGAGATCATGCCTAGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_KER MIT_7	i7	CAAGCAGAAGACGGCATAACGAGATGTAGAGAGGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_KER MIT_8	i7	CAAGCAGAAGACGGCATAACGAGATCCTCTCTGGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P5_ARG ON_9	i5	AATGATACGGCGACCACCGAGATCTACACTATAGCCTACACTCTTTCCCT ACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5_ARG ON_10	i5	AATGATACGGCGACCACCGAGATCTACACATAGAGGCACACTCTTTCCCT ACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5_ARG ON_11	i5	AATGATACGGCGACCACCGAGATCTACACCCTATCCTACACTCTTTCCCT ACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5_ARG ON_12	i5	AATGATACGGCGACCACCGAGATCTACACGGCTCTGAACACTCTTTCCCT ACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5_ARG ON_13	i5	AATGATACGGCGACCACCGAGATCTACACAGGCGAAGACACTCTTTCCCT ACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG

P5_ARG ON_14	i5	AATGATACGGCGACCACCGAGATCTACACTAATCTTAACACTCTTTCCCT ACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5_ARG ON_15	i5	AATGATACGGCGACCACCGAGATCTACACCAGGACGTACACTCTTTCCCT ACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5_ARG ON_16	i5	AATGATACGGCGACCACCGAGATCTACACGTACTGACACACTCTTTCCCT ACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG

Table 2.14. Programme gRNA cassette region PCR amplification.

Stage	Temperature (°C)	Time	Number of Cycles
Initial Denaturation	98	03:00	1x
Denaturation	98	00:15	5x (Stage B)
Annealing	65	00:30	5x (Stage B)
Extension	72	00:30	5x (Stage B)
Denaturation	98	00:15	8x (Stage C)
Annealing	63	00:30	8x (Stage C)
Extension	72	00:30	8x (Stage C)
Denaturation	98	00:15	12x (Stage D)
Annealing	61	00:30	12x (Stage D)
Extension	72	00:30	12x (Stage D)
Final Extension	72	07:00	1x
Hold	4	Infinity	1x

2.4.10 Synthesis of UNC7700

UNC7700 was synthesised by linking EED226, an EED-binding inhibitor, to a von Hippel–Lindau (VHL) ligand, which targets proteins for degradation via the ubiquitin-proteasome system. EED226 inhibits the EED subunit of the PRC2 complex, while the VHL ligand enables selective degradation of the targeted proteins. The compound was developed as previously described by our collaborator, Lindsey I. James, at North Carolina Chapel Hill[1]. The compound was synthesised in bulk by WuXi AppTec (<https://www.wuxiapptec.com/>).

2.5 Bioinformatics and data analysis

2.5.1 Whole-genome sequencing analysis

Target cell lines STS26T, 90-8, ST-8814, s462TY and iHSC were analysed using the Genome Analysis Toolkit (GATK) following standard best practices[93]. Genomic DNA was extracted and sequenced by Azenta Life Sciences[114], yielding raw data in FASTQ format. The analysis using the GATK workflow was performed by Dr Mirvat Surakhy (University of Oxford) for cell lines 90-8, ST-8814 and s462TY, and by Dr Stuart Brown (University of Oxford) for STS26T and iHSC. The GATK pipeline began with the alignment of sequencing reads to the reference genome using the Burrows-Wheeler Aligner (BWA)[88]. Post-alignment, duplicate reads were marked with Picard tools[93], and local realignment around indels was performed to improve accuracy. Base quality score recalibration (BQSR) was then applied to correct systematic errors in base calling. Variant calling was conducted using the HaplotypeCaller in GATK[93], producing variant call format (VCF) files. Finally, variant filtering was applied to retain high-confidence variants based on standard quality metrics. All analyses utilised default GATK settings to ensure consistency and reliability across cell line assessments.

2.5.2 RNA sequencing analysis

RNA sequencing was performed on cell lines STS26T, 90-8, ST-8814, S462TY, and iHSC, sequenced at Azenta Life Sciences[114]. Quality control of the raw sequencing data was conducted using FastQC[90]. Adapter removal and trimming of low-quality reads were performed with Trimmomatic (v0.39)[97]. The cleaned reads were then aligned to the reference genome using STAR (version 2.7.8)[96]. For gene quantification, FeatureCounts (part of SUBREAD version 2.0.3)[91] was employed to generate raw

count matrices. Differential gene expression analysis was carried out using the DESeq2 workflow, following the guidelines outlined in the DESeq2 vignette available at Bioconductor[94]. Gene set enrichment analysis (GSEA) was performed utilising Signal2Noise metrics for HPLM versus DMEM conditions and 'Log₂ Ratio of Classes' for UNC7700 versus DMEM conditions, where the former can be used for comparisons between conditions with more than two replicates per condition and the latter for two replicates per condition. GSEA focussed on Hallmark, KEGG Legacy, and Reactome gene sets to assess functional enrichment of differentially expressed genes[94]. Only leading-edge genes responsible for significant enrichment or depletion were taken forward for hit prioritisation.

2.5.3 Standard CRISPR screen data analysis using Log₂(fold change)

FASTQ files, quality-checked using FASTQC, were processed with the MAGeCK count function as part of the MAGeCK RRA algorithm[95], aligning sequencing reads to a reference gRNA library to generate a count matrix reflecting gRNA abundance across different time points. The counts per gRNA were normalised using Reads Per Million (RPM). Subsequently, log₂ fold changes and downstream analysis were performed in R.

2.5.4 CRISPR screen data analysis using Chronos

The CRISPR screen count matrices were generated using the MAGeCK RRA count function as previously described. These count matrices were then processed using Chronos[85]. Read count data were imported and transposed to have guide RNAs (gRNAs) as columns and samples as rows. Sequence maps, guide-to-gene maps, and negative control gRNAs were also imported. The 'nan_outgrowths' function in Chronos was applied to correct for gRNAs exhibiting excessive outgrowth, which could bias the

analysis[85]. Chronos was run in a pretrained mode, with the model configured using parameters derived from training on the DepMap data from Humagne C and D CRISPR screens. Gene effect scores were calculated and subsequently corrected with Chronos for copy number variations (CNVs) effects using CNV data from the screened target cell lines as previously described. Both CNV-corrected and CNV-uncorrected gene effect matrices were utilised for downstream analysis, which was performed in R.

2.5.5 Common essential gene recall

Candidate essential genes were identified by ranking genes and selecting those within the bottom 15% of values, both for $\log_2$ (fold change) (LFC) and Chronos scores. The overlap between these selected genes and the list of common essential genes from the Trevor Hart group, institution(Table 2.5) was then calculated to determine recall[98]. The percentage of common essential genes identified within the bottom 15% of ranked genes served as the recall metric. In some instances, the arbitrary 15% threshold was replaced by a dynamic increasing threshold to assess recall across a broader range of gene rankings.

2.5.6 Use of DepMap in this study

In this study, DepMap was used as a reference for benchmarking the quality of CRISPR screen experiments, investigating findings at the technical level (specifically the impact of gRNA representation and its effect on Chronos gRNA ranking, as detailed in Chapter 3), and hit prioritization. DepMap also provided a list of commonly non-essential genes, which was used to further validate and prioritise hits (https://figshare.com/articles/dataset/Chronos_Gene_Fitness_Effect_Inference_fro

m_CRISPR/14067047). The specific datasets utilised from DepMap are listed in Table 2.5.

2.5.7 pDNA representation and gene effect correlation

To investigate the relationship between the initial pDNA representation and gene effects observed during CRISPR screens, the Chronos scores for each gene were first combined with the corresponding pDNA counts (initial gRNA representation) from the CRISPR libraries. These pDNA counts were normalised to reads per million (RPM) and were then used to categorise the genes into quantile-based bins. Within each bin, the mean absolute Chronos score across all replicates was calculated for each gene. The median of these mean absolute Chronos scores was then computed for each bin. The resulting medians were plotted against the corresponding pDNA count bins, and a locally estimated scatterplot smoothing (loess) curve was fitted to the median values to model the trend in the data[115]. A Spearman's rank correlation test was performed rather than a Pearson correlation to assess the monotonic relationship between the pDNA counts and the mean absolute Chronos scores, as the data did not meet the assumptions of normality.

2.5.8 Prioritisation of context-dependent hit genes

To identify context-dependent specific hits in the HPLM versus DMEM CRISPR screens, as well as in the UNC7700 versus DMSO CRISPR screens, differential Chronos scores were calculated by subtracting the Chronos score of one condition from that of the other. These differences were then ranked, and hit genes were prioritised based on their rank. For the HPLM versus DMEM screens, hit genes were selected primarily based on their effects in MPNST cell lines (90-8 and ST-8814), with additional consideration

for their presence across all screened cell lines (90-8, ST-8814, STS26T, and iHSC). Hits were identified separately in Humagne C and D CRISPR library screens. In the UNC7700 versus DMSO screens, which were conducted exclusively in iHSC cells, hit genes were required to be present in both Humagne C and D CRISPR library screens to minimise the likelihood of capturing noise rather than true biological signal, even though this approach might exclude potential hits detectable only with one library.

2.5.9 Protein structure and AlphaFold structural prediction

To investigate the functional significance of specific domains within the SUZ12 protein, residues F86, F90, H466, S583, and R196 were selected as targets for mutagenesis, aimed at disrupting key functional regions of the protein. These residues were highlighted within the structural context of SUZ12 using a high-resolution cryo-EM structure of the Polycomb Repressive Complex 2 (PRC2) bound to AEBP2 and RBBP4[45], with the exception of R196, which was not included in the cryo-EM model and was instead visualised using an AlphaFold-generated model[87].

2.5.10 Statistical analysis

Statistical analysis was performed on data generated from at least three independent experiments, with no samples excluded, unless otherwise specified. Statistical comparisons were conducted using the appropriate tests for the type of data, as mentioned in the figure legends (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Statistical analysis packages are listed in Table 2.4.

Chapter 3: CRISPR screen technical design variables impacting on the selection of hits

3.1 Introduction

Genome-wide CRISPR knockout screens offer a dual-edged potential in genomic research. On one hand, they enable relatively straightforward, high-throughput experiments with significant power to identify cellular dependencies across the genome. On the other hand, the resulting data often contains extensive noise, complicating the extraction of meaningful signals that warrant further validation.

The hope is to achieve clear-cut discrimination between true hits and irrelevant noise in CRISPR screens, yet the reality is often more complex[116], [85]. Hit prioritization is frequently non-standardised, and high noise levels complicate the process of discriminating true positive from false positive signals. The lack of perfect efficacy across gRNAs has often been the main suspect behind this noise, with one study having shown that prioritising gRNAs with high efficacy results in a better specificity in selecting true positives[117]. Another large-scale study also identified noise related to gRNA

efficacies, as well batch effects that include differences in CRISPR screen duration[118]. Despite advancements in CRISPR technology, such as improved CRISPR targeting efficiencies, reduced off-target effects, and improved tools for data analysis, CRISPR screens still exhibit a substantial level of noise that remains largely unaddressed, and ,to a certain extent, undiscussed. Therefore, effective hit prioritisation, involving the discrimination between true and false positive gene effects, requires a comprehensive understanding of CRISPR screen mechanics and the sources of associated noise and bias.

In this chapter, I will systematically address the various components that are critical to conducting and analysing CRISPR screens, with a focus beyond mere hit prioritization.

The specific aims of this chapter are:

1. **Benchmarking CRISPR Screen Analysis Tools:** I will conduct a comparative evaluation of two CRISPR screen analysis tools, one representing the current state-of-the-art (Chronos[81], [85]) and the other a standard method ($\log_2(\text{fold change})$ [81]).
2. **Comparative Analysis with DepMap Dataset:** To validate the quality and robustness of the CRISPR screens in this study, I will perform a comparative analysis using the DepMap dataset. This will involve applying standardised quality control metrics to assess the performance of the screens within a broader, established context.
3. **Evaluation of CRISPR Enzyme Activity:** I will examine the influence of CRISPR enzyme activity on the accurate identification of gene dependencies.
4. **Assessment of gRNA Efficacy:** I will analyse the relationship between gRNA efficacy and the commonly employed quality control metrics in CRISPR screens.

5. **Evaluate the impact of gRNA Representation in CRISPR Libraries:** I will explore the impact of gRNA representation within CRISPR libraries, focusing on how variations in gRNA abundance may bias screen results and affect hit identification.
6. **Integration of Data from Subsequent Screens:** I will integrate data from subsequent chapters' screens into the current analysis, offering a comprehensive view of screen performance and ensuring that the conclusions drawn are robust and generalisable across different experimental contexts.

By addressing these key areas, this chapter aims to establish a rigorous framework for the identification and prioritisation of novel cellular dependencies, setting the stage for the discovery of potential MPNST therapeutic targets.

3.2 Results

3.2.1 Developing and validating *enAsCas12a* expressing neural crest derived cell line models

In this study, a total of five distinct cell lines were subjected to CRISPR screen experiments, as detailed in Table 3.1. All these cell lines underwent whole-genome sequencing (WGS) and whole-transcriptome sequencing. The analysis of WGS data followed the standard Genome Analysis Toolkit (GATK) pipeline[119], which facilitated the identification of somatic mutations, and copy number alterations (Fig. 3.1). Across target cell lines 90-8, ST-8814, s462TY and STS26T, a region on the start of the long arm of chromosome 17, relatively close to the centromere, spanning a similar area containing the same genes was found lost. Specifically, genes *FLJ36000* and *MTRNR2L1* were found lost in 90-8, ST-8814, and s462TY in that area but not in STS26T, where no known (protein-coding) genes were located. *MTRNR2L1* is a pseudogene and

FLJ36000 is an uncharacterised RNA gene[120]. IHSC displayed no CNV alterations in this area. Further findings in individual cell lines were further interpreted and enriched by integrating transcriptomic data, Table 3.1.

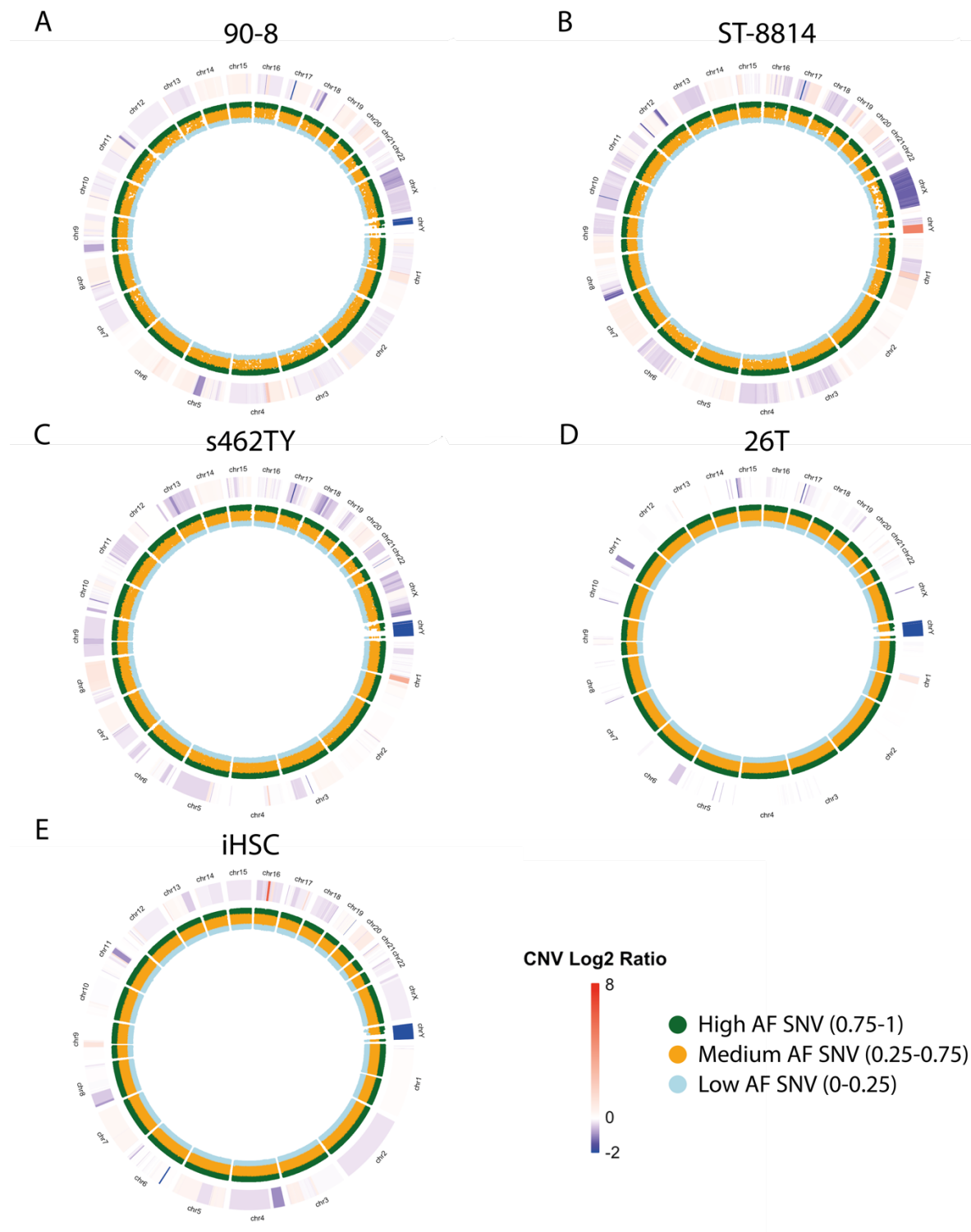


Figure 3.1. Circos plots of copy number variation and single nucleotide variant data. 90-8, ST-8814, s462TY, STS26T, iHSC cells. (A-E) Circos plots illustrating genomic alterations across five cell lines: 90-

8 (A), ST-8814 (B), s462TY (C), STS26T (D), and iHSC (E). The outer ring represents copy number variations (CNVs) with the red to blue colour range reflecting copy number changes. White indicates a copy number of two, red indicates higher copy number, and blue indicates a lower copy number. The inner three rings display single nucleotide variants (SNVs) categorised by allele frequency (AF): green for high AF (0.75-1), orange for medium AF (0.25-0.75), and light blue for low AF (0-0.25).

In the three MPNST cell lines 90-8, ST-8814 and s462TY, all were derived by others from patients with neurofibromatosis type 1, with expected biallelic loss of *NF1*. Notably, *SUZ12* was also disrupted in 90-8 and ST-8814. In 90-8, a single nucleotide variant (SNV) with an allele frequency (AF) of 96.1% was identified in *SUZ12*, potentially indicative of loss of heterozygosity and duplication (copy number neutral LOH). ST-8814 exhibited structural variation in *SUZ12* following a deletion, as evidenced by truncated RNA transcripts with no detectable alternate alleles.

In *tumour protein 53* (*TP53*), an SNV with an AF of 99% was identified in s462TY, while STS26T demonstrated copy number loss (CN = 1). *CDKN2A* exhibited copy number loss in the 90-8 and s462TY cell lines (CN = 1), whereas a gain was detected in the ST-8814 cell line (CN = 4). Additional details on *CDKN2A* genetic variation were referenced from another study [121]; however, these variations were not observed in the data generated during this study, Table 3.1. The complete genomic and genetic data on these target cell lines was used for various purposes during this study, including copy number variation correction during CRISPR screen data analysis, synthetic lethality prediction, and driver variant complementation strategies discussed in Chapter 3, 4 and 5, respectively.

Table 3.1. Cell line models and genetics.

Cell Line	Disease	NF1	CDKN2A	SUZ12	EED	TP53	Source
-----------	---------	-----	--------	-------	-----	------	--------

90-8	NF1-related MPNST	cn = 2, chr17:31235950 GGATCCTT>G, p.Asp1302TyrfsTer5, high impact, AF=97.6%	cn = 1, homozygous loss according to Magallón-Lorenz et al., 2023, iScience [122]	cn = 2, chr17:31993330 T>A, p.Tyr430Ter, high impact, AF=96.1%	N/A	N/A	[123]
ST-8814	NF1-related MPNST	cn = 1, chr17:31221856 C>CT, frameshift variant, high impact, AF=0.978	cn = 4, and SV according to Magallón-Lorenz et al., 2023, iScience [122]	cn = ?, chr17:31998662, SUZ12 BND; RNA-Seq: chr17:31,943,717, transcript terminates in intron between exon 3-4, LOH according to Magallón-Lorenz et al., 2023, iScience [122]	cn = 3	cn = 1	[124]
s462TY	NF1-related MPNST	cn = 1, chr17:31338739 C>A, stop gained, pathogenic, AF=0.985	cn = 1, High impact deletion; Strand -1; Exon 2/3; c.151-404_457+99 del; Transcript NM_000077.5; Protein P42771; and LOH according to Magallón-Lorenz et al., 2023, iScience [122]	cn = 1	cn = 1	cn = 1, TP53 (chr17:7676040 C>G), Arg110 Pro, missense variant, AF: 0.990	[125]
iHSC	Immortalised human Schwann cells	N/A	N/A	cn = 1	cn = 1	N/A	[106]
STS26T	Melanoma (misdiagnosed as	N/A	N/A	SNV and LOH according to	N/A	cn = 1, SNV and cn gain	[126]

	MPNST)			Magallón- Lorenz et al., 2023, iScience [122]		accordi ng to Magall ón- Lorenz et al., 2023, iScienc e [122]	
--	------------	--	--	---	--	---	--

cn = copy number, SV = structural variation, LOH: loss of heterozygosity, AF: allele frequency, NF1 = neurofibromatosis type 1, N/A = Not applicable.

The plasmid DNA (pDNA) CRISPR libraries for this study, Humagne C and D, were constructed without incorporating a gene for the expression of the Enhanced AsCas12a CRISPR (enAsCas12a) enzyme. Instead, the enAsCas12a enzyme was provided by John Doench (Eli and Edythe L. Broad Institute of MIT and Harvard) on a separate plasmid construct, pRDA174 (Addgene plasmid #136476)[80]. To facilitate the creation of stable cell lines expressing enAsCas12a, a T2A-EGFP tag was integrated into pRDA174, enabling Fluorescence-activated Cell Sorting (FACS). These modified constructs were introduced into target cell lines via lentiviral delivery.

3.2.2 Stability of enAsCas12a expression using GFP in cell lines

For successful CRISPR screening, it is essential that the target cells express the CRISPR-associated protein (Cas), particularly during the initial days following CRISPR library transduction (Chapter 3). This time frame is critical as it is when CRISPR-induced perturbations occur, typically reaching saturation around seven days after transfection[127]. However, given the frequent requirement for a substantial number of cells to initiate a screen (~35 million), cell cultures are typically expanded before the transduction of the CRISPR library. To assess whether target cell lines 90-8, ST-8814,

s462TY, iHSC, and STS26T maintained enAsCas12a-T2A-EGFP expression over time, cells were monitored over time using flow cytometry (Fig. 3.2A-B).

For ST-8814, s462TY, and iHSC cell lines, a time-dependent expanding population of cells was observed that expressed lower levels of EGFP expression compared to the initial levels. To avoid overestimating the population that effectively expresses enAsCas12a, gating was adjusted above this lower EGFP expressing population rather than using the intended negative control. Since there are no established thresholds for the required percentage of CRISPR enzyme expressing population within a screening target population, all cells were included in the subsequent CRISPR screening experiments.

To validate enAsCas12a activity, STS26T cells were transduced with a gRNA cassette targeting EGFP, which was inserted into a lentiviral construct analogous to the Humagne C and D CRISPR libraries, pRDA052 (Addgene plasmid #136474)[81]. Cells transduced with *EGFP*-targeting gRNAs exhibited a marked reduction in EGFP expression compared to enAsCas12a-T2A-EGFP expressing cells that did not receive the *EGFP*-targeting gRNAs (Fig. 3.2C). A sustained and lower level of EGFP expression was observed relatively long (15 days) after the initial *EGFP* knockout, likely due to the comparatively long half-life of EGFP (approximately 26 hours)[128].

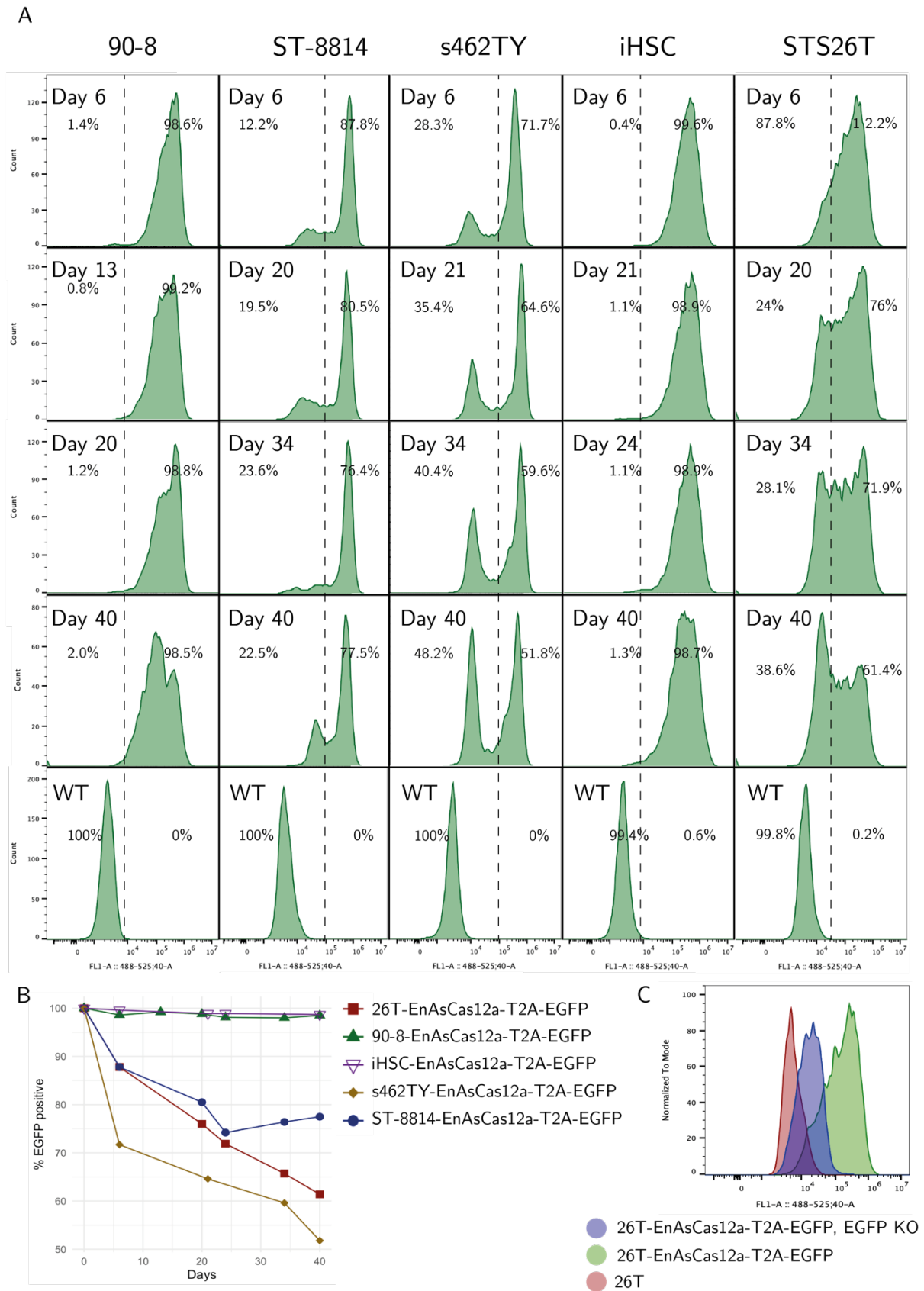


Figure 3.2. Stability enAsCas12a expression is time and cell line dependent. (A) Histograms of flow cytometry measuring EGFP in 90-8, ST-8814, s462TY, iHSC and STS26T cells complemented with

enAsCas12a-T2A-EGFP using lentiviral transgene delivery. Cells were monitored for up to 40 days post transduction. EGFP positivity was gated stringently well above negative control populations (WT) to be able to separate true positives from pseudo positives, especially for ST-8814, s462TY and STS26T cells. **(B)** Line plot with time series data as depicted showing the percentage of EGFP positive cells based on the gating strategy as depicted in (A) over time for 90-8, ST-8814, s462TY, iHSC and STS26T cells complemented with *enAsCas12a-T2A-EGFP*. Each data point represents one biological replicate. **(C)** Histogram showing shift in EGFP expression 15 days after transducing STS26T-*enAsCas12a-T2A-EGFP* cells with an *EGFP*-targeting gRNA cassette.

3.2.3 CRISPR library amplification and validation

Humagne C and D genome-wide dual-targeting CRISPR plasmid DNA (pDNA) libraries, kindly provided by the John Doench group at the Broad Institute of MIT and Harvard (USA), were transformed into electrocompetent cells for amplification, followed by lentiviral production (see Chapter 2). To assess changes in the distribution of individual gRNA cassettes containing gRNA pairs targeting a single gene, both the original and amplified CRISPR libraries underwent next-generation sequencing (NGS) after PCR amplification of the gRNA cassette regions. The MagEck RRA algorithm was utilised to determine the counts per gRNA cassette[95]. As the expectation was that there would be a normal distribution of gRNA representation in the library, it would be important to evaluate the statistical parameters of this distribution. One metric used to determine the level of skewing of the normal distribution within the post-amplification library that reflects the magnitude of the difference in gRNA representation across the library, is calculated as shown below[127].

$$\text{skew ratio} = \frac{\text{mean gRNA cassette count of top 10\% counted gRNA cassettes}}{\text{mean gRNA cassette count of bottom 10\% counted gRNA cassettes}}$$

As expected, both Humagne C and D libraries maintained a similar gRNA representation across libraries following amplification, with skew ratios being well below the recommended, but arbitrarily set, threshold of 10 as described in a study from Feng Zhang's group [127] (Eli and Edythe L. Broad Institute of MIT and Harvard)(Fig. 3.3A-B). Instances of substantial changes in gRNA representation were rare, and a Pearson correlation analysis between the original and amplified libraries demonstrated a nearly perfect linear relationship (Fig 3.3C-D), suggesting that the amplification process exerts minimal impact on the overall gRNA cassette distribution. Consequently, the amplification process was deemed successful without major skewing of gRNA representation and thus considered suitable for subsequent lentiviral production and CRISPR screening.

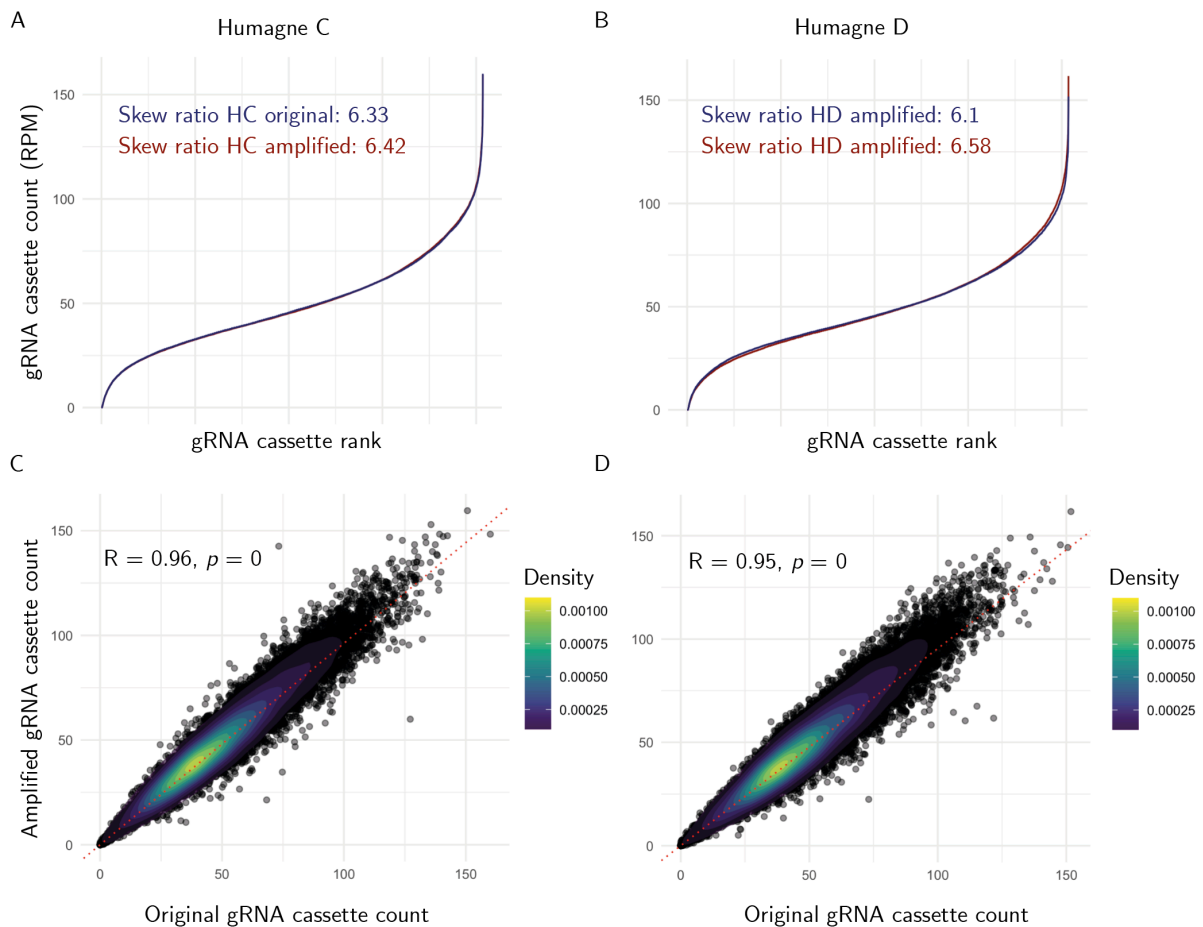


Figure 3.3. CRISPR library amplification shows minimal skewing of gRNA cassette distribution. (A) gRNA cassette distribution with normalised counts per gRNA cassette ranked for original and amplified plasmid DNA CRISPR libraries Humagne C. Skew ratio indicates ratio between mean top 10% gRNA cassette normalised counts divided by the mean bottom 10% gRNA cassette normalised counts. (B) Same as in (A) but for Humagne D. (C) Dot and density plot of relationship between the original and amplified gRNA cassette normalised count. A Spearman's correlation was performed and with Rho and p value displayed on the plot. The Rho value is further displayed by a red dashed line.

3.2.4 Summary of conducted CRISPR screens in enAsCas12a cell lines

Here, I conducted 28 independent genome-wide CRISPR knockout screens across the five cell lines described, utilising two types of culture media, DMEM and human plasma-like medium (HPLM, discussed in Chapter 4), and with or without DMSO or an EED-targeting proteolysis-targeting chimera (PROTAC; UNC7700, discussed in Chapter 5), Table 3.2.

Each condition model was screened with two separate genome-wide CRISPR libraries, Humagne C and D, which are proposed to function as genome-wide biological replicates [80]. For all screens, genomic DNA (gDNA) was extracted from cells harvested chronologically at multiple intermediate time points, in addition to the final time point, and processed for next-generation sequencing (NGS) to monitor the temporal representation of gRNAs.

To account for differences in growth rates between cell lines and ensure fair competition between cells with different CRISPR-based gene knockouts, final time points were standardised to 14 population doublings (PDs) rather than a fixed duration in days. However, exceptions were made for the second replicate screens of 90-8 in DMEM and STS26T in HPLM, for both Humagne C and D libraries. In these cases, the first screens were of poor quality, initially suspected to be due to the screen's duration. Consequently,

the second screens were extended beyond the standard time frame. Repeat experiments were also conducted for s462TY in DMEM, in an effort to enhance screen performance, particularly as measured by the recall of common essential genes, as discussed later in Chapter 3.

Although the screens STS26T.HPLM_R1, 90-8.DMEM_R1, and s462TY.DMEM_R2 are discussed in Chapter 3 for data analysis and quality control insights, they are excluded from the data analysis in Chapters 4 and 5, including hit prioritization.

Table 3.2. CRISPR screen time points and screen identifiers.

Humagne C screen ID	Humagne D screen ID	Humagne CD screen ID	cell line model identifier	cell line	days	medium	drug
pDNA_HC_amplified	pDNA_HD_amplified	pDNA_amplified	pDNA	NA	0	DMEM	NA
screen_2CD0_R1	screen_2DD0_R1	screen_2D0_R1	26T.DMEM_R1	STS26T	4	DMEM	NA
screen_2CD6.1_R1	screen_2DD6.1_R1	screen_2D6.1_R1	26T.DMEM_R1	STS26T	11	DMEM	NA
screen_2CD9.3_R1	screen_2DD9.3_R1	screen_2D9.3_R1	26T.DMEM_R1	STS26T	15	DMEM	NA
screen_2CD14_R1	screen_2DD14_R1	screen_2D14_R1	26T.DMEM_R1	STS26T	20	DMEM	NA
screen_2CH0_R1	screen_2DH0_R1	screen_2H0_R1	26T.HPLM_R1	STS26T	4	HPLM	NA
screen_2CH17d_R1	screen_2DH17d_R1	screen_2H17d_R1	26T.HPLM_R1	STS26T	17	HPLM	NA
screen_2CH0_R2	screen_2DH0_R2	screen_2H0_R2	26T.HPLM_R2	STS26T	4	HPLM	NA
screen_2CH16d_R2	screen_2DH16d_R2	screen_2H16d_R2	26T.HPLM_R2	STS26T	20	HPLM	NA
screen_2CH18_R2	screen_2DH18_R2	screen_2H18_R2	26T.HPLM_R2	STS26T	38	HPLM	NA
screen_8CD0_R1	screen_8DD0_R1	screen_8D0_R1	ST-8814.DMEM_R1	ST-8814	4	DMEM	NA
screen_8CD2.1_R1	screen_8DD2.1_R1	screen_8D2.1_R1	ST-8814.DMEM_R1	ST-8814	8	DMEM	NA
screen_8CD4.5_R1	screen_8DD4.5_R1	screen_8D4.5_R1	ST-8814.DMEM_R1	ST-8814	12	DMEM	NA
screen_8CD12_R1	screen_8DD12_R1	screen_8D12_R1	ST-8814.DMEM_R1	ST-8814	24	DMEM	NA
screen_8CD14_R1	screen_8DD14_R1	screen_8D14_R1	ST-8814.DMEM_R1	ST-8814	26	DMEM	NA
screen_8CH0_R1	screen_8DH0_R1	screen_8H0_R1	ST-8814.HPLM_R1	ST-8814	4	HPLM	NA
screen_8CH3_R1	screen_8DH3_R1	screen_8H3_R1	ST-8814.HPLM_R1	ST-8814	21	HPLM	NA

screen_8CH8.7_R1	screen_8DH8.7_R1	screen_8H8.7_R1	ST-8814.HPLM_R1	ST-8814	37	HPLM	NA
screen_8CH21d_R1	screen_8DH21d_R1	screen_8H21d_R1	ST-8814.HPLM_R1	ST-8814	25	HPLM	NA
screen_8CH14_R1	screen_8DH14_R1	screen_8H14_R1	ST-8814.HPLM_R1	ST-8814	49	HPLM	NA
screen_9CD0_R1	screen_9DD0_R1	screen_9D0_R1	90-8.DMEM_R1	90-8	4	DMEM	NA
screen_9CD14_R1	screen_9DD14_R1	screen_9D14_R1	90-8.DMEM_R1	90-9	29	DMEM	NA
screen_9CD0_R2	screen_9DD0_R2	screen_9D0_R2	90-8.DMEM_R2	90-10	4	DMEM	NA
screen_9CD2_R2	screen_9DD2_R2	screen_9D2_R2	90-8.DMEM_R2	90-11	8	DMEM	NA
screen_9CD4.4_R2	screen_9DD4.4_R2	screen_9D4.4_R2	90-8.DMEM_R2	90-12	12	DMEM	NA
screen_9CD6.6_R2	screen_9DD6.6_R2	screen_9D6.6_R2	90-8.DMEM_R2	90-13	16	DMEM	NA
screen_9CD9.5_R2	screen_9DD9.5_R2	screen_9D9.5_R2	90-8.DMEM_R2	90-14	21	DMEM	NA
screen_9CD14_R2	screen_9DD14_R2	screen_9D14_R2	90-8.DMEM_R2	90-15	29	DMEM	NA
screen_9CD28_R2	screen_9DD28_R2	screen_9D28_R2	90-8.DMEM_R2	90-16	52	DMEM	NA
screen_9CH0_R1	screen_9DH0_R1	screen_9H0_R1	90-8.HPLM_R1	90-17	4	HPLM	NA
screen_9CH26d_R1	screen_9DH26d_R1	screen_9H26d_R1	90-8.HPLM_R1	90-18	30	HPLM	NA
screen_9CH14_R1	screen_9DH14_R1	screen_9H14_R1	90-8.HPLM_R1	90-19	38	HPLM	NA
screen_iCD0_R1	screen_iDD0_R1	screen_iD0_R1	iHSC.DMEM_R1	iHSC	4	DMEM	NA
screen_iCD2.5_R1	screen_iDD2.5_R1	screen_iD2.5_R1	iHSC.DMEM_R1	iHSC	8	DMEM	NA
screen_iCD5_R1	screen_iDD5_R1	screen_iD5_R1	iHSC.DMEM_R1	iHSC	12	DMEM	NA
screen_iCD8.5_R1	screen_iDD8.5_R1	screen_iD8.5_R1	iHSC.DMEM_R1	iHSC	18	DMEM	NA
screen_iCD14_R1	screen_iDD14_R1	screen_iD14_R1	iHSC.DMEM_R1	iHSC	29	DMEM	NA
screen_iCH0_R1	screen_iDH0_R1	screen_iH0_R1	iHSC.HPLM_R1	iHSC	4	HPLM	NA
screen_iCH4.7_R1	screen_iDH4.7_R1	screen_iH4.7_R1	iHSC.HPLM_R1	iHSC	24	HPLM	NA
screen_iCH7.7_R1	screen_iDH7.7_R1	screen_iH7.7_R1	iHSC.HPLM_R1	iHSC	32	HPLM	NA
screen_iCH14_R1	screen_iDH14_R1	screen_iH14_R1	iHSC.HPLM_R1	iHSC	44	HPLM	NA
screen_iCD4d_D_R1	screen_iDD4d_D_R1	screen_iD4d_D_R1	iHSC.DMEM.DMSO_R1	iHSC	4	DMEM	DMSO
screen_iCD8d_D_R1	screen_iDD8d_D_R1	screen_iD8d_D_R1	iHSC.DMEM.DMSO_R1	iHSC	8	DMEM	DMSO
screen_iCD17d_D_R1	screen_iDD17d_D_R1	screen_iD17d_D_R1	iHSC.DMEM.DMSO_R1	iHSC	17	DMEM	DMSO
screen_iCD29d_D_R1	screen_iDD29d_D_R1	screen_iD29d_D_R1	iHSC.DMEM.DMSO_R1	iHSC	29	DMEM	DMSO
screen_iCD4d_U_R1	screen_iDD4d_U_R1	screen_iD4d_U_R1	iHSC.DMEM.UNC7700_R1	iHSC	4	DMEM	UNC7700
screen_iCD8d_U_R1	screen_iDD8d_U_R1	screen_iD8d_U_R1	iHSC.DMEM.UNC7700_R1	iHSC	8	DMEM	UNC7700

screen_iCD17d_U_R1	screen_iDD17d_U_R1	screen_iD17d_U_R1	iHSC.DMEM.UNC7700_R1	iHSC	17	DMEM	UNC7700
screen_iCD29d_U_R1	screen_iDD29d_U_R1	screen_iD29d_U_R1	iHSC.DMEM.UNC7700_R1	iHSC	29	DMEM	UNC7700
screen_sCD0_R1	screen_sDD0_R1	screen_sD0_R1	s462TY.DMEM_R1	s462TY	4	DMEM	NA
screen_sCD14_R1	screen_sDD14_R1	screen_sD14_R1	s462TY.DMEM_R1	s462TY	26	DMEM	NA
screen_sCD0_R2	screen_sDD0_R2	screen_sD0_R2	s462TY.DMEM_R2	s462TY	4	DMEM	NA
screen_sCD14_R2	screen_sDD14_R2	screen_sD14_R2	s462TY.DMEM_R2	s462TY	31	DMEM	NA

NA = not applicable

3.2.5 CRISPR screen duration affects gRNA dropouts

CRISPR screen experiments require sufficient coverage throughout the workflow in order to capture true signal and minimise noise[129], [104]. For genome-wide CRISPR knockout screens, a coverage of approximately 500 cells per gRNA construct is frequently adopted[127], along with 500 sequencing reads per construct during NGS. Errors arise with false positives if there is insufficient coverage at any stage, leading to the apparent premature dropout of gRNAs and an increase in undetected gRNAs during data analysis. While gRNA representation may decrease significantly due to gene effects, a complete 100% dropout is generally unexpected, mainly due to lack of CRISPR expression and targeting efficiency, even though there may be a lethal knockout phenotype[127].

At total of 106 CRISPR library screen time points were sequenced, including both the original and amplified CRISPR libraries. Of these, 14 time points had undetected gRNA cassettes exceeding the recommended threshold of 0.5% of the total library (for Humagne C and D this number would be 102), as recommended by a study from Feng Zhang's group (Eli and Edythe L. Broad Institute of MIT and Harvard)[127](Fig. 4.3A). The initial consideration was that this was due to insufficient NGS sequencing coverage. However, no significant correlation was found between the number of mapped reads per

sample and the number of undetected gRNAs (Fig. 4.3B). The proportion of mapped reads per sample relative to the total number of reads, which might indicate a PCR-related issue, also did not show a significant correlation with the number of undetected gRNAs (Fig. 4.3C).

Interestingly, these time points were relatively late, all occurring 21 days or more post-transduction. Although the threshold for undetected gRNAs is designed to account for early as well as late time points, CRISPR screens typically do not extend beyond 14 to 21 days[82]. However, in this study, CRISPR screens frequently included late time points that extended well beyond this standard duration (Table 3.2). A significant correlation was found between the number of days post-CRISPR library transduction and the number of undetected gRNAs for the time points in this study (Fig. 3.4D). While this trend is particularly pronounced at very late time points, it is also observable at time points under 21 days.

Although these late time points with relatively high numbers of undetected gRNAs may not independently provide sufficient coverage, Chronos, an algorithm used for CRISPR screen data analysis, integrates data from multiple time points within a single screen to generate one comprehensive effect score for each gene[130]. This algorithm will be further discussed in Chapter 3.

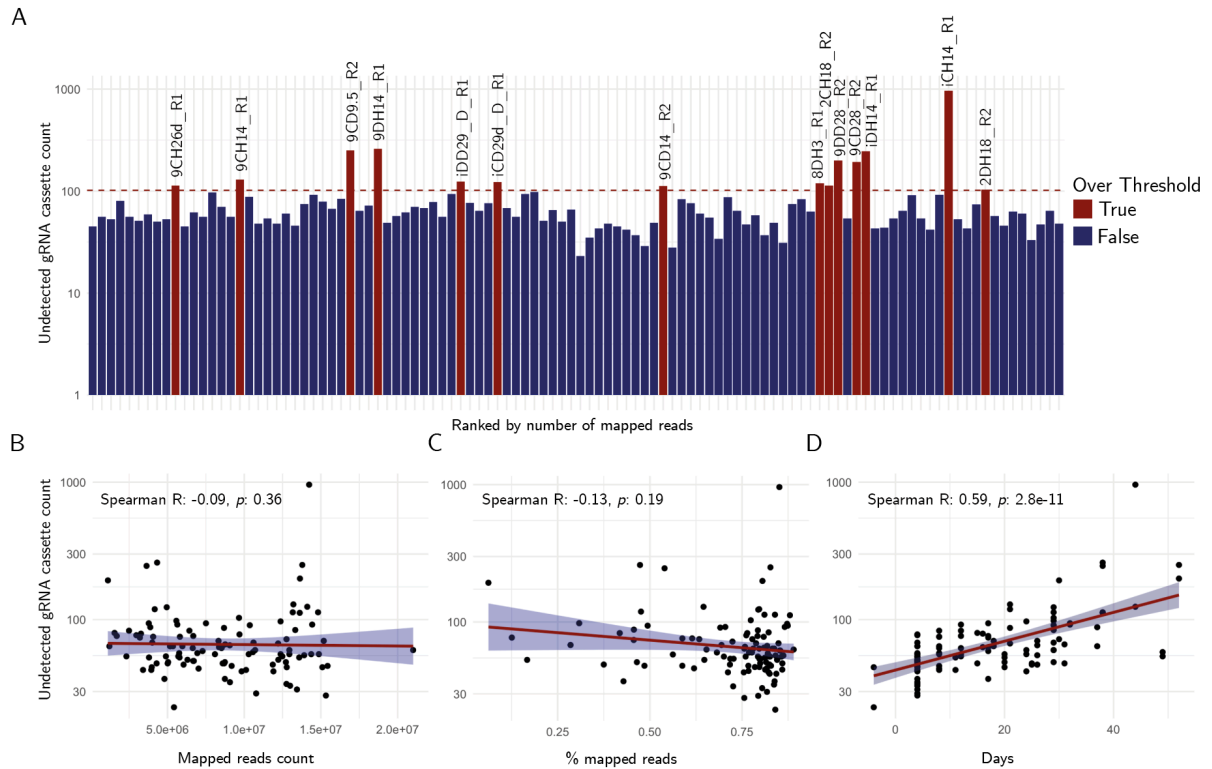


Figure 3.4. CRISPR screen coverage is time dependent. (A) Bar graph of undetected gRNA cassette counts per CRISPR screen sample. Horizontal red dashed line at 102 undetected gRNA cassettes represents a quality control threshold with samples surpassing it considered to have lost sufficient coverage by having a number of undetected gRNAs over the threshold (True) and are depicted as red bars. Samples that have a number of undetected gRNAs under the threshold (False) are depicted as blue bars. Screen identifiers are explained in Table 3.2. (B) Dot plot and linear regression model with 95% confidence interval of the relationship between the number of mapped reads and the number undetected gRNA cassettes per sample. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot. (C) Dot plot and linear regression model with 95% confidence interval of the relationship between the percentage of mapped reads of the total number of reads, and the number of undetected gRNA cassettes per sample. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot. (D) Dot plot and linear regression model with 95% confidence interval of the relationship between the number of days since CRISPR library transduction at which genomic DNA samples were collected, and the number of undetected gRNA cassettes per sample. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot.

3.2.6 Global effect size is a determinant for quality of CRISPR screen

data

Review of previously described CRISPR screen methodologies defined a number of key practical considerations when designing and analysing CRISPR screen data[129], [104]. The count of undetected gRNAs, can provide insights into the CRISPR screen coverage as well as the duration, as previously discussed. The count of undetected gRNAs, however, is not directly related to other critical factors that influence the outcomes of CRISPR screens. These include CRISPR nuclease system expression and activity, as well as the overall efficiency of the CRISPR library. When these parameters are also suboptimal, one would expect a reduction in CRISPR targeting efficiency, resulting in only a subset of cells achieving the desired knockout phenotype. As a result of the latter, the signal derived from a large population of cells, considered as 500 cells per library construct, is instead detected in a smaller subpopulation, thereby potentially reducing the overall statistical power and sensitivity of the screen. To visualise the impact of these multiple factors within a CRISPR screen, one can employ a measure of global effect size, which can be expressed as follows:

$$\text{global effect size} = \frac{\text{mean (positive } \log_2(\text{fold change}) \text{ values for all gRNAs)}}{\text{mean (absolute negative } \log_2(\text{fold change}) \text{ values for all gRNAs)}} / 2$$

This metric provides an aggregate view of the gene effect sizes in a whole-genome screen, balancing the contributions of both positively and negatively selected gRNAs. It serves as an indicator of the overall efficacy and of the CRISPR screen (time point), particularly in relation to CRISPR system performance and the efficiency of the library in

inducing the desired genetic perturbations across the population of cells. The effect size reflects the strength and range of changes in gRNA representation over time.

Although effect sizes can be represented by single numbers, they can also be visualised, as in Figure 3.5A. Notably, screens conducted in STS26T and s462TY exhibit the lowest effect sizes among all screens, observed in both Humagne C and D CRISPR library screens across different medium types. This is likely attributable to the previously discussed relatively low expression levels of EnAsCas12-T2A-EGFP. Additionally, the first replicate screen in 90-8 with Humagne C, and to a lesser extent with Humagne D (identifiers: screen_9CD14_R1 and screen_9DD14_R1), also demonstrated a relatively low effect size. This is likely due to the CRISPR libraries being transduced at a multiplicity of infection (MOI) that was initially believed to be 0.3, but later determined to be much higher at approximately 1.2. High MOI in this instance means that most cells would have more than one guide cassette, adding considerable noise to the system and minimising the effect size.

Another key parameter influencing the global screen effect size may be the duration of the screen. Extended screening periods allow for increased competition among cells, secondly impacted by different genetic perturbations, resulting in more pronounced changes in gRNA representation. This relationship was confirmed to be highly significant through a Spearman's correlation analysis[131], as shown in Figure 3.5B.

The global effect size is not only useful for diagnosing parameters that may have negatively impacted CRISPR screen outcomes, but it also correlates with more commonly used quality metrics in CRISPR screen data analysis, such as the recall of common essential genes.

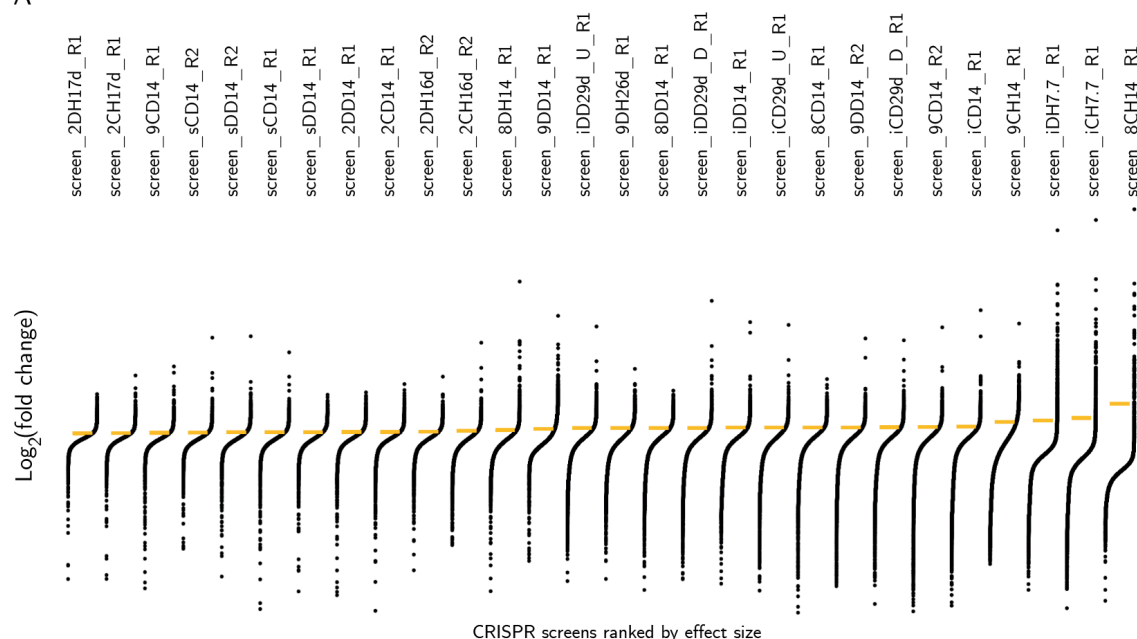
3.2.7 Recall of essential (lethality) genes can be applied as a quality control measure

Essential genes are those necessary for the survival of a cell. Common essential genes are a subset of essential genes that are essential across various cell lines, tissue types, and sometimes even across different species. The definition of these common essential genes has not been well defined and as a result has led to the emergence of different common essential gene reference lists[97], [132], [133]. These lists are derived from a variety of experimental approaches, including high-throughput genomic techniques such as RNAi and genome-wide CRISPR library knockout screens, as well as targeted knockout studies conducted both *in vitro* and *in vivo*.

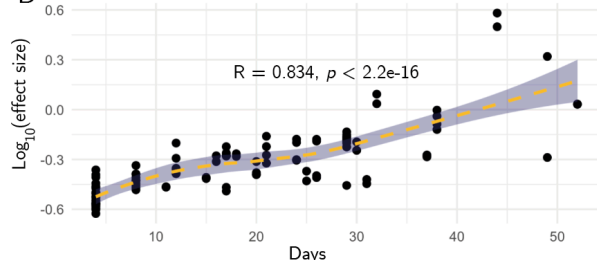
Common essential gene lists are frequently used as reference sets in the quality control analysis of genome-wide CRISPR screen datasets. Guides targeting genes from these common essential gene reference sets serve as positive controls. The recall rate of these common essential genes is one of the most critical quality metrics for genome-wide CRISPR screen data and has been employed to benchmark some of the most widely used CRISPR screen data analysis tools [94], [130], [134], [135]. One regularly used common essential gene list, established by Traver Hart's group (MD Anderson Cancer Center, Texas, USA), was derived from genome-wide CRISPR screens conducted in five different cell lines[98]. For inclusion in this list, common essential genes had to be hits in at least three of the five cell lines, resulting in a total of 1,580 genes. However, it may be apparent that the criteria for defining "common" essential genes based on their occurrence in only three out of five cell lines are arbitrary, as these genes are, per definition, not universally essential across all cell types. However, this list was utilised to benchmark the

performance of the Humagne C and D libraries by the group that developed these libraries, and so will also be used as a reference throughout this study, unless otherwise specified[81]. The global effect size of CRISPR screens in this study is significantly correlated with the recall of these 1,580 common essential genes (Fig. 3.5C), suggesting that both metrics reflect the quality of various parameters influencing CRISPR screens. Based on these findings, and the common essential gene recall rates of these screens, as discussed next, screens STS26T.HPLM_R1, 90-8.DMEM_R1, and s462TY.DMEM_R2 were kept for data analysis and quality control insights in this Chapter, but excluded from the data analysis in Chapters 4 and 5, including hit prioritization.

A



B



C

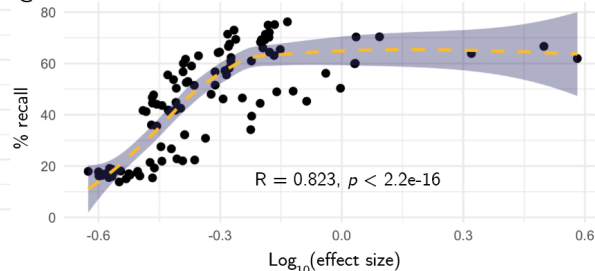


Figure 3.5. Common essential gene recall is dependent on CRISPR screen effect size. (A) Dot plot displaying the range of $\text{Log}_2(\text{fold change})$ values per gRNA, for each CRISPR screen in this study. Each screen is only represented by one late time point, which was selected based on the highest common essential genes recall from all time points in that screen[98]. Samples were ranked by effect size, arranged from low to high (from left to right, respectively). Screen identifiers are explained in Table 3.2. (B) Dot plot and locally estimated scatterplot smoothing (LOESS) with 95% confidence interval of the relationship between the effect size and percentage recall of common essential genes [136]. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot. (C) Dot plot and LOESS with 95% confidence interval of the relationship between the number of days after CRISPR library transduction and the effect size per screen. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot.

3.2.8 Chronos improves CRISPR screen data analysis

Chronos is a computational algorithm developed by the Broad Institute for the analysis of CRISPR screen data. It has been extensively validated across multiple large-scale CRISPR screening datasets and has been applied to a wide range of studies in both basic research and therapeutic development. It represents one of the most comprehensive algorithms for analysing CRISPR screen data. Chronos generates a single gene score by primarily evaluating changes in gRNA representation over time. This score is refined using various mechanistic parameters derived from the input data, including gRNA efficacy, screen quality, the growth rate of target cell lines, and the diversity of gene-independent DNA-cutting toxicity effects. Chronos can also integrate multiple time points, if available, to model the temporal change in gRNA representation. A key advantage of Chronos is its ability to account for the variable growth rates of different cell lines, which often necessitates harvesting at different time points. By incorporating these factors, Chronos minimises the risk of inflating effect sizes due to variations in CRISPR screen duration, making it particularly suitable for studies like these that involve distinct cell lines.

Furthermore, Chronos directly models gene effects at the read count level using a negative binomial noise model, which does not assume normal distribution of data, unlike methods that model log-fold change values with a Gaussian distribution[130]. This approach has been shown to enhance the accuracy of separating known common essential and non-essential genes, outperforming other CRISPR screen analysis methods.

Chronos also offers the flexibility of being run by individual users after pretraining on the DepMap dataset, provided they use one of the libraries included in the DepMap dataset: Avana, KY, or Humagne CD. This feature further enhances the data quality as measured by standard metrics, including the recall of common essential genes.

To independently benchmark Chronos, the algorithm was pretrained using DepMap-based Humagne CD parameters to analyse the data in this study. This approach was then compared to the traditional $\log_2(\text{fold change})$ values calculated per gene (Figure 3.6). The Chronos workflow incorporates all available time points, in contrast to the $\log_2(\text{fold change})$ method, which only uses single late time points. For this comparison, the single late time points for the $\log_2(\text{fold change})$ method were selected based on which had the best recall of common essential genes, as determined by selecting the bottom 15% ($N = 3053$) of genes[130].

Chronos demonstrated superior performance in recalling common essential genes for the vast majority of screens conducted in this study compared to the $\log_2(\text{fold change})$ method (Figure 3.6A). One screen, screen_8CD14_R1, showed an unexpected, substantially lower recall of common essential genes when analysed with Chronos, although it still maintained sufficient quality compared to other screens.

Similarity between replicate screens was also evaluated to determine whether Chronos improves the quality of CRISPR screen data. To assess this, the entire array of gene effects for each screen was compared to its biological replicate, conducted with the same library, using Spearman's correlation. The Rho values from comparisons based on the $\log_2(\text{fold change})$ method were then compared to those obtained using the Chronos method. The analysis revealed that the correlations between replicates analysed with Chronos significantly outperformed the correlations derived from the $\log_2(\text{fold change})$ method, suggesting that Chronos is superior in bringing replicates closer (Figure 3.6B, C).

The similarity between detected essential genes in replicate screens is particularly relevant when aiming to identify cancer-specific dependencies. To evaluate this, replicate screens were compared by overlapping the bottom 15% ($N = 3053$) of scoring genes using both $\log_2(\text{fold change})$ and Chronos (Fig 3.6D). The analysis revealed that Chronos significantly improved the overlap of essential genes identified across replicates. Similar to the improved recall of common essential genes, this suggests that Chronos may enhance the detection of true essential genes in screened cell lines.

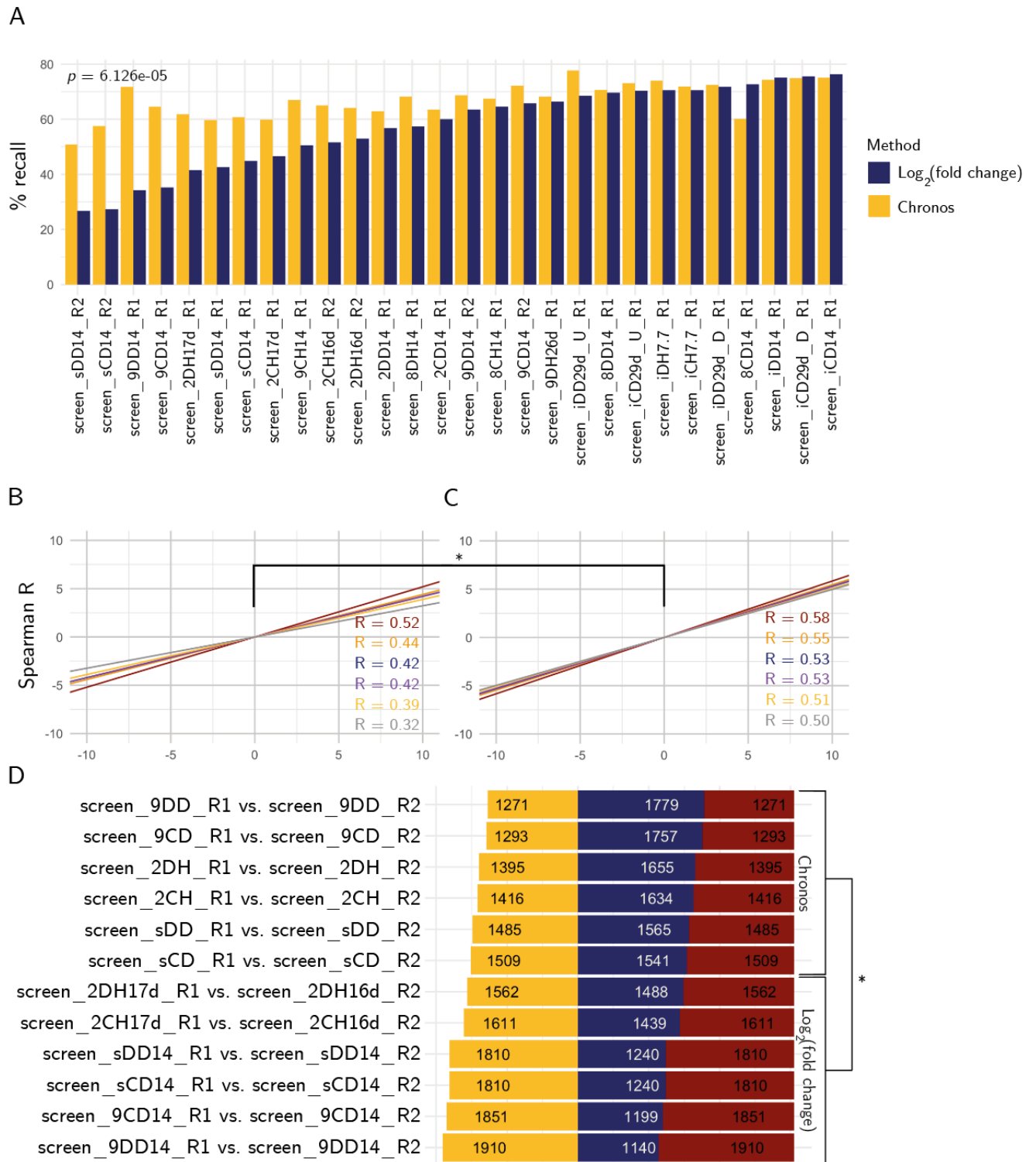


Figure 3.6. Chronos improves CRISPR screen data analysis based on common essential genes recall and replicate similarity. (A) Bar graph comparing common essential gene recall following log₂(fold change) (blue) or Chronos analysis (yellow). Log₂(fold change) time points were selected based on best recall per time point while Chronos includes all available time points. Screens were ranked based on the recall obtained using the log₂(fold change) analysis. Screen identifiers are explained in Table 3.2. **(B, C)** Line plot displaying Spearman's correlations between replicate CRISPR screens analysed using the log₂(fold change) (B) or

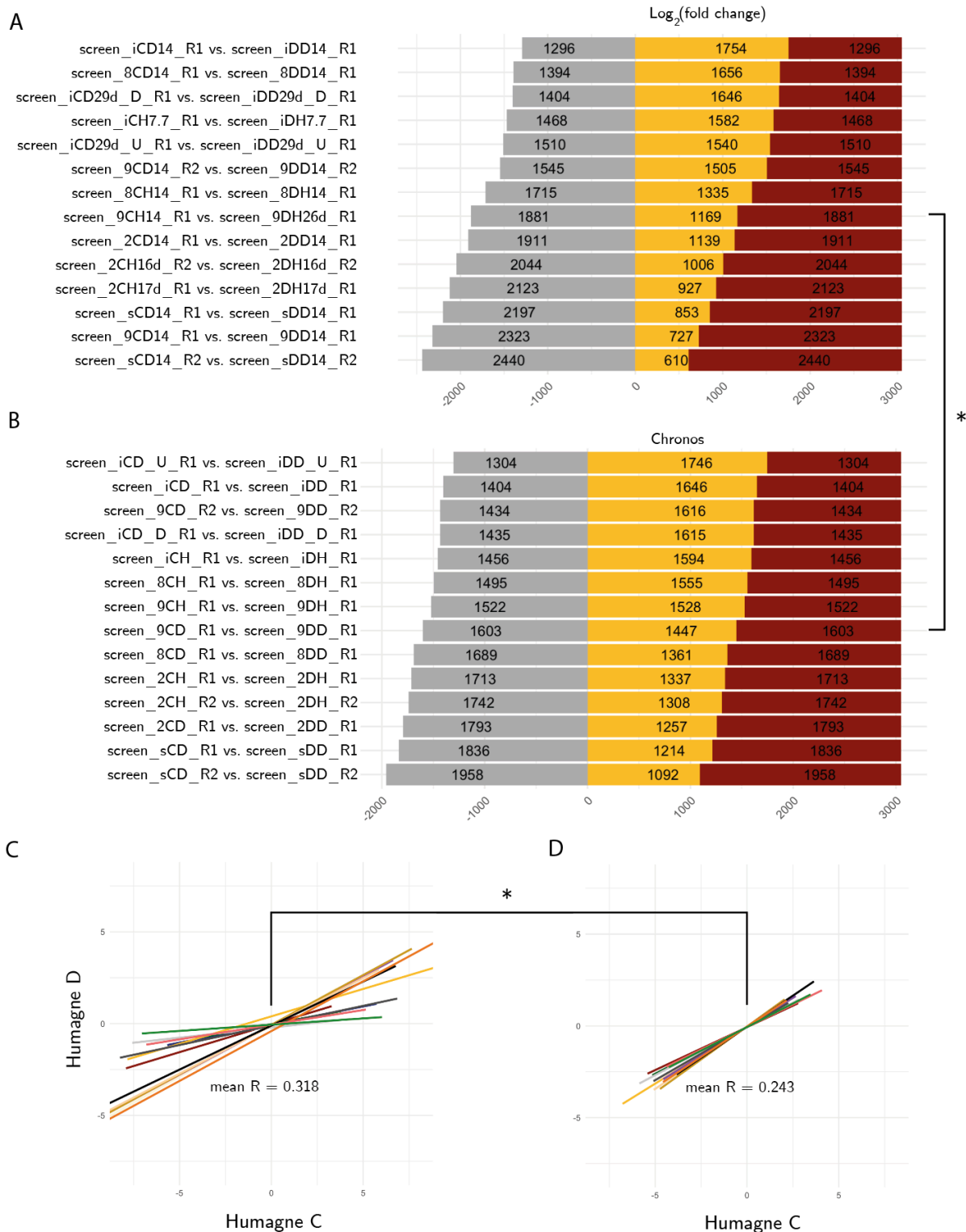
Chronos (C) method. Individual Rho (R) values per screen correlation are colour coded. A Wilcoxon signed-rank test was used to compare the distribution of Rho values between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS not significant. **(D)** Horizontal stacked bar chart displaying overlapping (blue) and non-overlapping (replicate 1: yellow, replicate 2: red) essential genes (bottom 15% genes) between replicate CRISPR screens analysed with $\log_2$ (fold change) or Chronos, ranked by number of overlaps. A Wilcoxon signed-rank test was used to compare the distribution of Rho values between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS not significant. Screen identifiers are explained in Table 3.2.

Humagne C and Humagne D are two independently produced dual-targeting CRISPR libraries, each containing a distinct set of gRNAs[81]. Unlike conventional CRISPR screen experiments that rely on repeating screens with the same CRISPR library to generate biological replicates, Humagne C and D offer a more advanced approach by utilising different gRNAs to target the same gene. This method allows these libraries to serve as more sophisticated biological replicates. Furthermore, Humagne C and D together have shown to outperform replicates of traditional CRISPR library screens, achieving similar or superior results while requiring fewer cells for coverage, thanks to the multiplexing of gRNA cassettes[81].

To assess the similarity in performance between Humagne C and D libraries, essential genes identified in each equivalent screen conducted with Humagne C and D, defined as the bottom 15% of genes ($N = 3053$)[130], were overlapped and analysed (Fig. 3.7A-B). This analysis was performed using both the $\log_2$ (fold change) and Chronos methods. With the $\log_2$ (fold change) method, nine out of fourteen screens exhibited more non-overlapping than overlapping genes, whereas Chronos reduced this number to eight. However, even with Chronos, 43–65% of essential genes identified in screens conducted within the same cell line model but using distinct libraries were still different.

The similarity between screens conducted within the same cell line model but using different libraries was also evaluated on a genome-wide scale by comparing them using Spearman's correlation (Fig. 3.7C, D). Chronos resulted in significantly higher correlations between screens compared to the $\log_2(\text{fold change})$ method, further demonstrating its enhanced reliability for cross-library comparisons.

In summary, Chronos was found to enhance the recall of common essential genes, improve the similarity between replicates, and increase the consistency between datasets derived from identical cell line models but obtained with different libraries. Based on these findings, Chronos was adopted as the standard method for CRISPR screen data analysis throughout this study, unless explicitly stated otherwise.



0.01, *** $p < 0.001$, NS not significant. Screen identifiers are explained in Table 3.2. **(C, D)** Line plot displaying Spearman's correlations between Humagne C and D CRISPR screens analysed using the $\log_2$ (fold change) (C) or Chronos (D) method. Mean Rho (R) values are indicated on the graphs.

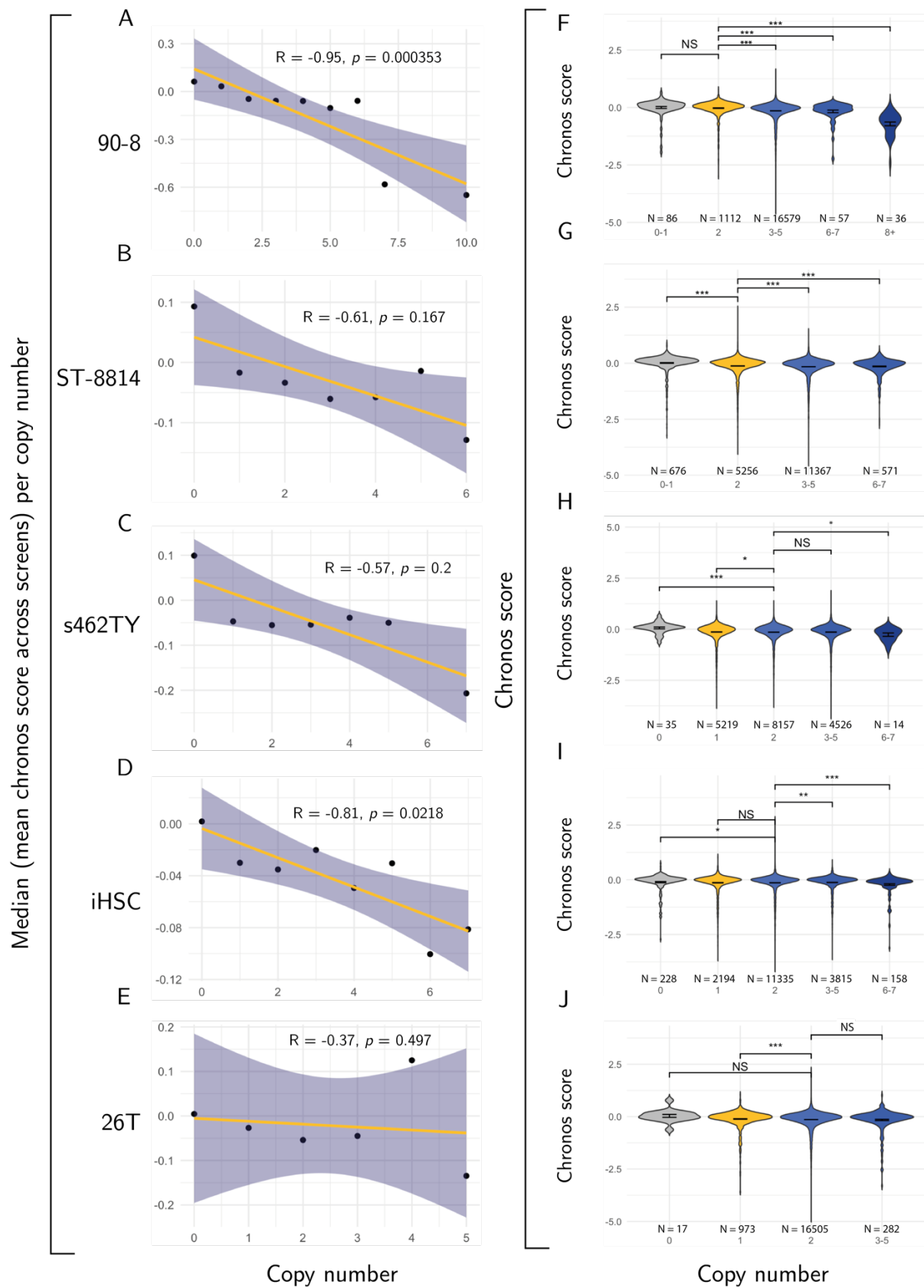
3.2.9 Copy number variation impact and correction

The CRISPR-Cas9 system continuously targets and cleaves sequences that match its guide RNAs (gRNAs). When these sequences are present in high copy numbers on the same chromosome, such as in cases of gene amplification with high copy number variation (CNV), the system can induce multiple double-strand breaks (DSBs). This accumulation of DSBs can lead to increased cellular toxicity and, ultimately, cell death. During CRISPR screens, a reduction in the representation of certain gRNAs can occur not due to the functional knockout of the targeted gene but rather as a consequence of the lethal effects of multiple DSBs. This phenomenon can generate false-positive signals, confounding the interpretation of screen results[137]. To mitigate this issue, analysis tools such as Chronos allow for the integration of custom CNV data, tailored to individual cell lines, to more accurately interpret CRISPR screen data and differentiate between true functional knockouts and artifacts caused by CNV-associated toxicity.

In order to correct for potential CNV effects in this study, all target cell lines were whole-genome sequenced (WGS) and CNV information was obtained using the GATK pipeline for genomic analysis followed by CNV calling using CNVkit [119], [138]. Chronos was first run without CNV correction to observe whether there is a correlation between high copy number and gRNA depletion for corresponding target genes as reflected in Chronos scores (Fig. 3.8A-E). A Spearman's correlation analysis revealed a significant negative relationship between increased copy numbers and Chronos scores in cell lines 90-8 and iHSC, but not for the other cell lines. To further explore the effect of different

CNV levels, Chronos scores were categorised based on CNV, and the different groups compared. Consistently, the mean Chronos score per CNV group decreases the higher the copy number gets. Compared to Chronos scores for genes with copy number 2, Chronos scores for genes with higher copy number values mostly have significantly lower scores (Fig. 3.8F-J). These observations suggest that gRNA depletion associated with CNVs may confound the functional interpretation of CRISPR screen results in these cell lines.

With this in mind, Chronos was provided with CNV data on each cell line to correct for this potential effect. Notably, while a marginal and sometimes statistically significant increase in Chronos scores reflects a correction for the impact of high copy number genes, there remains evidence of an inverse relationship between higher CNV levels and lower Chronos scores in certain instances (Fig. 3.8K-T). This observation may suggest a potential association between elevated CNV and gene essentiality, rather than the observed lower Chronos scores being solely attributable to cellular toxicity from DNA damage. Furthermore, although Chronos aims to account for the CNV effect, it may not achieve complete correction.



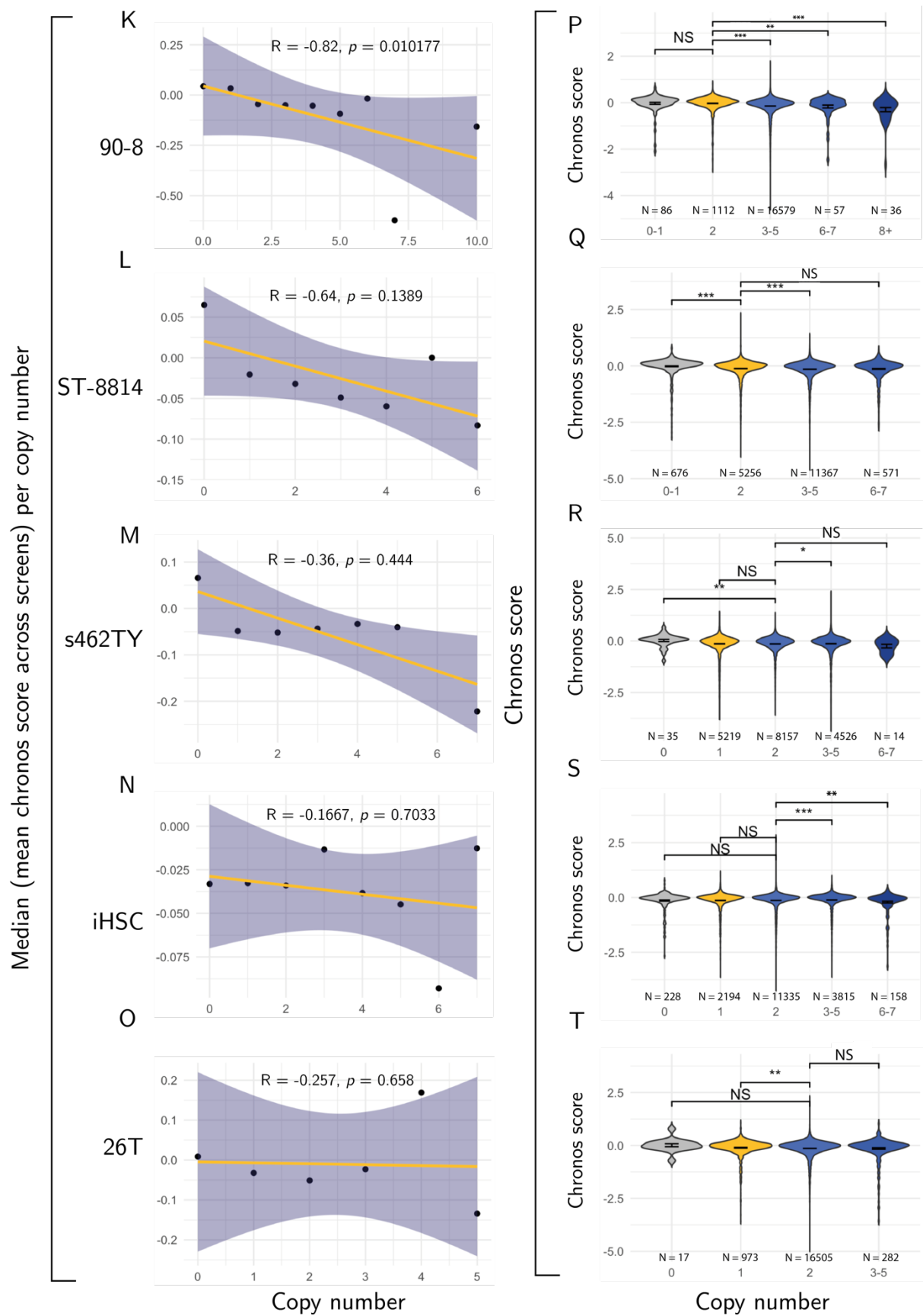


Figure 3.8. CRISPR screen outcomes correlate with higher gene copy numbers in a cell line dependent manner. (A) Dot plot and linear regression model with 95% confidence interval of the relationship between

copy number and the median of the mean Chronos score per gene across 90-8 screens, per copy number. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot. **(B)** Same as (A) but for ST-8814. **(C)** Same as (A) but for s462TY. **(D)** Same as (A) but for iHSC. **(E)** Same as (A) but for STS26T. **(F)** Violin plots of Chronos scores from genes with copy number = 2 and from genes with the highest possible copy, where at least three genes have this copy number. A Wilcoxon rank-sum test was used to compare the distribution of Chronos scores between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS not significant. **(G:J)** Same as (F) but for ST-8814, iHSC and STS26T, respectively. **(K-O)** Same as for (A-E) but then after CNV correction. **(P-T)** Same as for (F-J) but then after CNV correction.

3.2.10 Comparison of this study with DepMap screens

To assess the quality of the CRISPR screens performed in this study, I compared the common essential gene recall with that of the DepMap CRISPR screens [81] (Fig. 3.9). The most recent DepMap dataset includes analyses of screens with a single late time point. The DepMap dataset was generated by employing Chronos to combine Humagne C and D libraries into a single analysis, yielding one score per gene for both experiments. To ensure comparability, I adopted the same approach by analysing Humagne C and D as a single dataset for the screens in this study.

The mean common essential gene recall in our study (58.03%) was marginally higher than that of the DepMap dataset (56.93%), but this difference was not statistically significant. Notably, this trend remained consistent even after excluding screens that were considered too low in quality to proceed with for hit prioritization (screen_90-8.DMEM_R1, screen_s462TY.DMEM_R2, screen_26T.HPLM_R1), with the study mean slightly increasing to 58.52%.

A significant difference in recall was observed only after removing all screens performed in STS26T and s462TY cell line models, which express suboptimal levels of enAsCas12a (resulting in a study mean of 59.50%). While benchmarking the quality of

CRISPR screens against established datasets is valuable, it is important to note that there is no universally accepted threshold to discriminate high from low-quality screens. Given that all common essential gene recall rates in our study fell within the range observed in the DepMap dataset, no additional screens, but the previously mentioned STS26T.HPLM_R1, 90-8.DMEM_R1, and s462TY.DMEM_R2 screens, were excluded from the analysis.

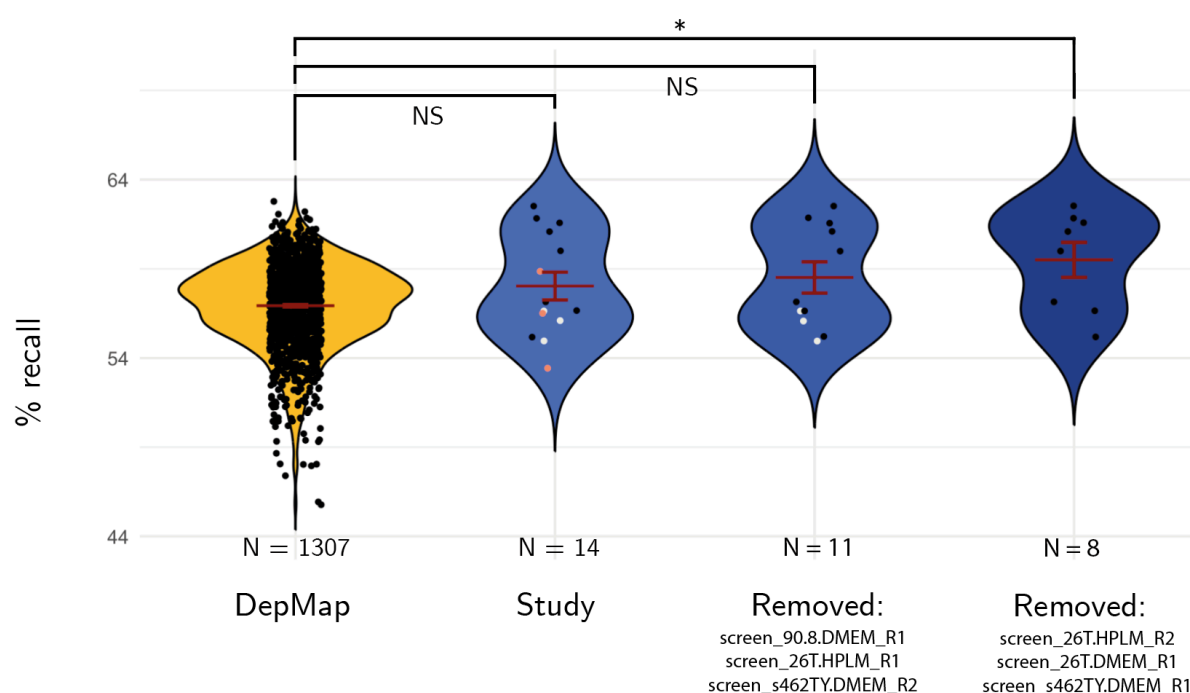


Figure 3.9. Quality of CRISPR screens in this study are of similar quality to DepMap screens. Violin plots displaying recall percentages per CRISPR screen experiment, indicated as black dots, as part of the DepMap dataset (grey violin) or this study (golden violin). During Chronos data analysis, Humagne C and D CRISPR screens were treated as a single library experiment to allow for a comparison with DepMap screens which were analysed in this way. Means and standard error of the mean error bars are indicated by horizontal red lines and vertical red brackets lines. A Wilcoxon rank-sum test was used to compare the distribution of percentage recall values between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS: not significant.

3.2.11 CRISPR library choice is major determinant for screen results.

Previously, it was shown that Humagne C and D, despite the highly significant correlation in performance, also lead to substantially different results. To investigate how different Humagne C and D libraries are, a principal component analysis (PCA) was performed on the Chronos datasets for each CRISPR library screen in this study (Fig. 3.10A). This shows that principal component 1 (PC1) can separate out screens based on the used library, with 19.7% of the variance explained. To evaluate whether this impact of library choice is related to Humagne C and D specifically, or whether other libraries show a similar clustering, a PCA was performed on the DepMap dataset, which contains data from CRISPR screens performed with Avana, KY and Humagne CD libraries (Fig. 3.10B). A similar observation was made for these screens where the first principal component very clearly separates all CRISPR screens based on the library that was used, with 12.85% of the variance in the data explained with this PC.

Importantly, both the screens from this study, as well as those in the DepMap dataset, encompass screens performed on identical cell line models. Specifically, the DepMap data includes 204 different cell lines that were screened using both the Avana and KY libraries. Despite this, the data clustered according to the library used rather than based on cell line models, medium context, or drug context, emphasising the significant influence of library selection on screen outcomes.

To reinforce this observation, screens conducted in the same cell line model but using different libraries (Humagne C versus Humagne D) were compared using Spearman's correlation. The resulting Rho values were then compared to correlations

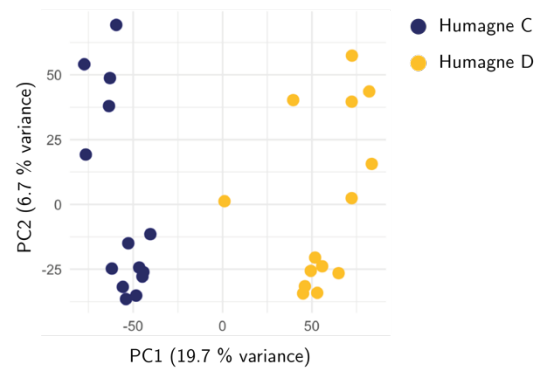
between screens in different cell line models performed with the same library (Humagne C versus Humagne C, and Humagne D versus Humagne D) (Fig. 3.10C). The analysis revealed that screens are more similar when the same library is used rather than when the same cell line model is used. These results strongly suggest that the choice of CRISPR library is a critical determinant of CRISPR screen outcomes, often outweighing the influence of the cell line model itself.

Chronos has the capability to analyse the Humagne C and Humagne D libraries either by treating them as part of a single experiment or as separate, independent experiments. When Chronos processes the screens as a unified experiment, it typically produces a Chronos score that falls between the individual scores derived from analysing Humagne C and Humagne D separately. In scenarios where there is a significant discrepancy between the Humagne C and D scores, Chronos tends to generate a score closer to the more neutral score of the two (Fig. 3.10D-E). This approach helps to mitigate the impact of off-target gene effects, especially when a single outlier gRNA shows a strong gene effect that is not supported by other gRNAs targeting the same gene.

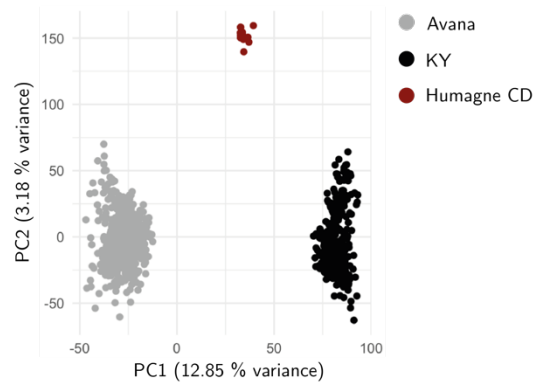
While off-target effects are a valid concern, previous research suggests that the frequency of gRNAs with off-target activity is relatively low in CRISPR screens, including those utilising enAsCas12a[81]. On the other hand, the proportion of gRNAs that fail to exhibit on-target activity is considerably higher, with a significant fraction in any given library unable to effectively cleave their intended targets. This is despite efforts exemplified in tools like CRISPRpick, CHOPCHOP, and CRISPOR which have been made to improve gRNA efficacy and reduce off-targeting by utilising advanced algorithms that

consider sequence specificity, predicted off-target sites, and gRNA secondary structure to optimise the design of guide RNAs [103], [139], [140]. In this study, analysis of the bottom 15% scoring genes (based on the mean score per gene across Humagne C or D screens) reveals a false positive rate of 4-5% (including off-target effects) compared to a false negative rate of 21-22% (including failure of on-target activity)(Fig. 3.10F-G). To avoid over-correcting for potential off-target effects, which are relatively infrequent, Chronos was run separately for Humagne C and D in this study to preserve true signals.

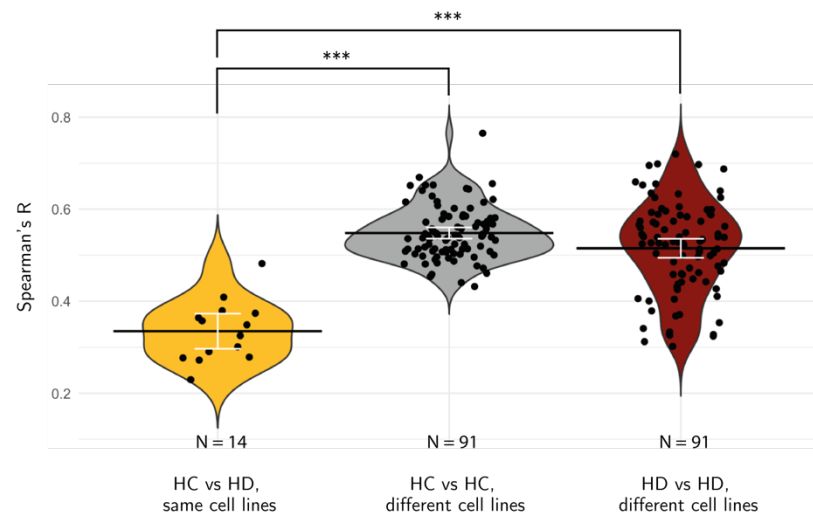
A



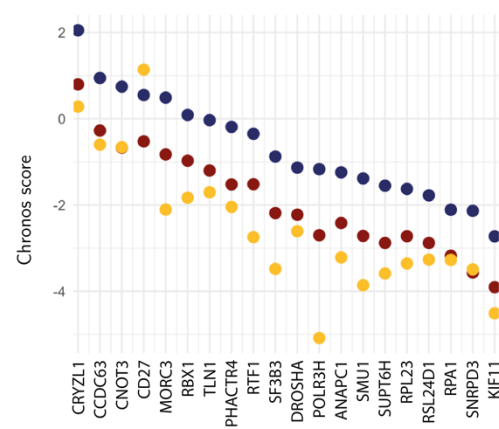
B



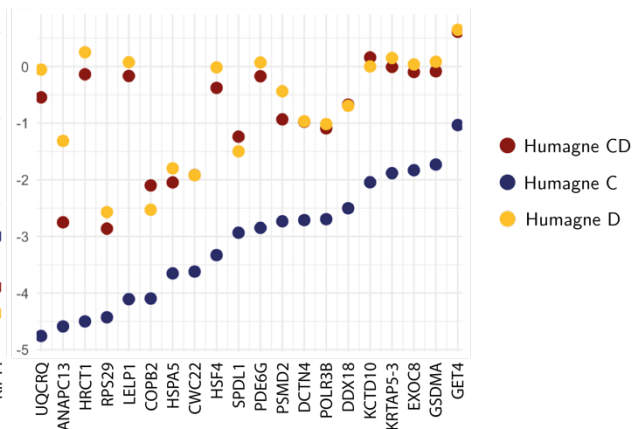
C



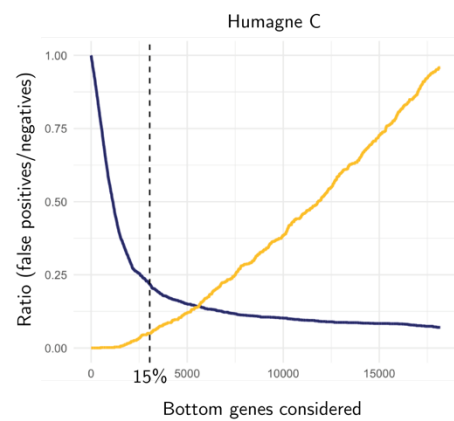
D



E



F



G

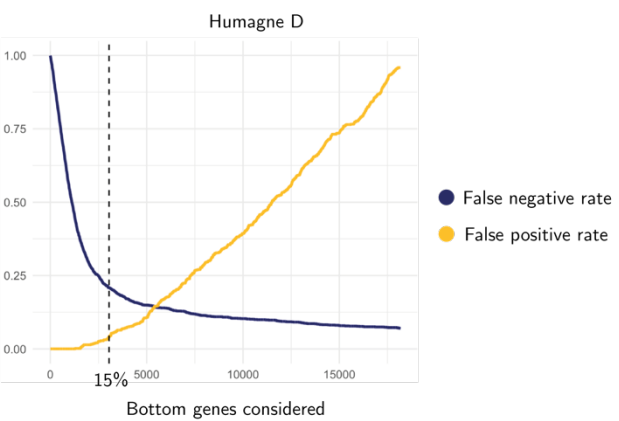


Figure 3.10. CRISPR library choice can be a more important determinant of screen results than the cell line model. (A) Dot plot displaying principal component analysis (PCA) on all CRISPR library screens performed ($N = 28$) in this study separately analysed for Humagne C (blue) and D (yellow) libraries. The variance explained by principal components is indicated on the side of each axis. PC1: principal component 1, PC2: principal component 2. (B) Dot plot displaying principal component analysis (PCA) on all CRISPR library screens from DepMap with dots colour coded according to library as indicated. The variance explained by principal components is indicated on the side of each axis. PC1: principal component 1, PC2: principal component 2. (C) Violin plots displaying Spearman's Rho (R) values of correlations between the same cell lines screened with Humagne C and Humagne D (golden violin) compared to all possible correlations between different cell lines and conditions screened with only Humagne C (grey violin) or Humagne D (red violin). Means per group are displayed by a horizontal black line and mean confidence intervals by white brackets. Wilcoxon rank-sum tests were performed to compare the distribution of Spearman's Rho (R) values between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS not significant. (D, E) Dot plot displaying Chronos scores per gene from the iHSC DMEM replicate 1 screen (top performing screen based on common essential gene recall) from Humagne C (blue), Humagne D (gold) and Humagne CD (red), where for the latter Chronos treated Humagne C and D as one experiment. Genes displayed are the top 20 (D) and bottom 20 genes (E) after subtracting Humagne D from Humagne C Chronos scores. (F, G) Line plots displaying the false negative rate of unrecalled common essentials (blue) and the false positive rate of negative control gRNAs (gold) calculated per number of bottom genes considered. Chronos scores were averaged per gene across screens.

3.2.12 Low initial gRNA representation leads to stronger gene effects during CRISPR screens.

To date, one major study has investigated why different CRISPR libraries yield such varying results. This comprehensive study compared hundreds of CRISPR screens conducted using the Avana library with screens performed on equivalent cell line models using the KY library [118], later integrated into the DepMap dataset. While this study did identify specific instances of discrepancies in predicted gRNA efficacy, these were not emphasised as the primary cause of the variations. The paper also examined differences in CRISPR screen duration between the two groups, revealing some distinctions in the effects of gene knockout over time, particularly between early and late gene effects of

knocked out essential genes. However, the study did not provide a definitive explanation for why different CRISPR libraries result in hugely different outcomes.

To explore whether a general relationship exists between differences in gRNA efficacy across libraries and their outcomes, Chronos scores from Humagne D were subtracted from those of Humagne C, then made absolute and correlated with the differences in gRNA efficacy scores between the two libraries (Humagne C subtracted by Humagne D) (Fig. 3.11A). Although the correlation observed is relatively weak, its statistical significance suggests that variations in gRNA efficacy contribute to the differences in Chronos scores between the libraries. However, despite this marginal trend in the data, it also appears that differences in gRNA efficacy alone do not fully account for the discrepancies observed between the CRISPR libraries.

Beyond the inherent differences in guide RNAs between libraries, these gRNAs are also produced independently, leading to variations in the representation of gRNA cassettes within each library, often referred to as the plasmid DNA count (pDNA count). This means that a gRNA targeting a specific gene in one library may have very low representation, whereas a gRNA from another library targeting the same gene might have an average or relatively high representation. Such differences in representation can potentially influence the outcomes of CRISPR screens and are attributable to the specific library (batch) used. Based on this, it was hypothesised that gRNAs with lower initial representation are more likely to exhibit stronger positive or negative changes compared

to gRNAs with average or high initial representation, and follows the law of large numbers, first described by Johann Bernoulli in 1713¹.

This potential effect is already evident in the data when analysed by using the $\log_2(\text{fold change})$ method, where higher initial guide representation appears to result in lower variance, as shown in Figure 3.11B-C (similar patterns were observed for all CRISPR screens in this study). In other words, gRNAs with lower associated pDNA counts tend to produce more pronounced positive or negative changes.

While this effect is observable in relatively unprocessed data, it becomes more pronounced when Chronos scores are categorised based on the initial gRNA representation (pDNA count) of the associated gRNAs. By segmenting the average Chronos scores per gene across screens into equally sized bins containing gRNAs with similar initial representations, and then taking their absolute values, one can examine the median score for each pDNA bin (Fig. 3.11D-E). This analysis reveals a highly significant and strong Spearman's correlation between lower pDNA bin numbers and more pronounced Chronos scores (both positive and negative), with correlation coefficients of $\text{Rho} = -0.98$ for both Humagne C and Humagne D.

As previously discussed, Chronos can analyse Humagne C and D screens as a single experiment. Since the two gRNA cassettes (a total of four gRNAs) from Humagne C and D are amplified separately, the effect of low pDNA representation in one library might be mitigated or even cancelled out by a normal representation in the other. To

¹ Bernoulli, J. (1713). *Ars Conjectandi*. Basel: Thurneysen Brothers. Johann Bernoulli first described the law of large numbers in this work, laying the groundwork for probability theory by demonstrating that as the sample size increases, the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed. This principle underlies the observed greater variability in gRNAs with lower initial representation in the study.

determine whether employing Chronos to analyse Humagne C and D as one experiment inadvertently reduces or corrects for this effect, these single Chronos scores per gene, coming from two different library screens, were binned according to the pDNA representations of either Humagne C or D (Fig 3.11F and G, respectively). Although the effect was reduced compared to analysing Humagne C and D separately, measured by Spearman's correlation, it remains strong (Humagne C Rho: -0.88, Humagne D Rho: -0.97) and highly significant. Thus, despite allowing Chronos to potentially and unintentionally dilute this effect by generating a single gene effect score that reflects data from two different libraries, with the score then mapped to only one of the two possible pDNA counts, the impact of pDNA representation remains highly significant.

To assess whether this pDNA representation effect was due to in-house CRISPR library amplification or was inherent to the original CRISPR library, average Chronos scores per gene across screens were binned according to the pDNA representation of the original, unamplified CRISPR library (Fig. 3.11H-I). The results closely mirrored those obtained using the amplified library pDNA counts.

Collectively, these results suggest that the representation of gRNA cassettes in the Humagne C and D CRISPR libraries significantly impacts CRISPR screen outcomes, with lower pDNA representation leading to stronger positive or negative effect scores, likely rooted in the original construction of the library.

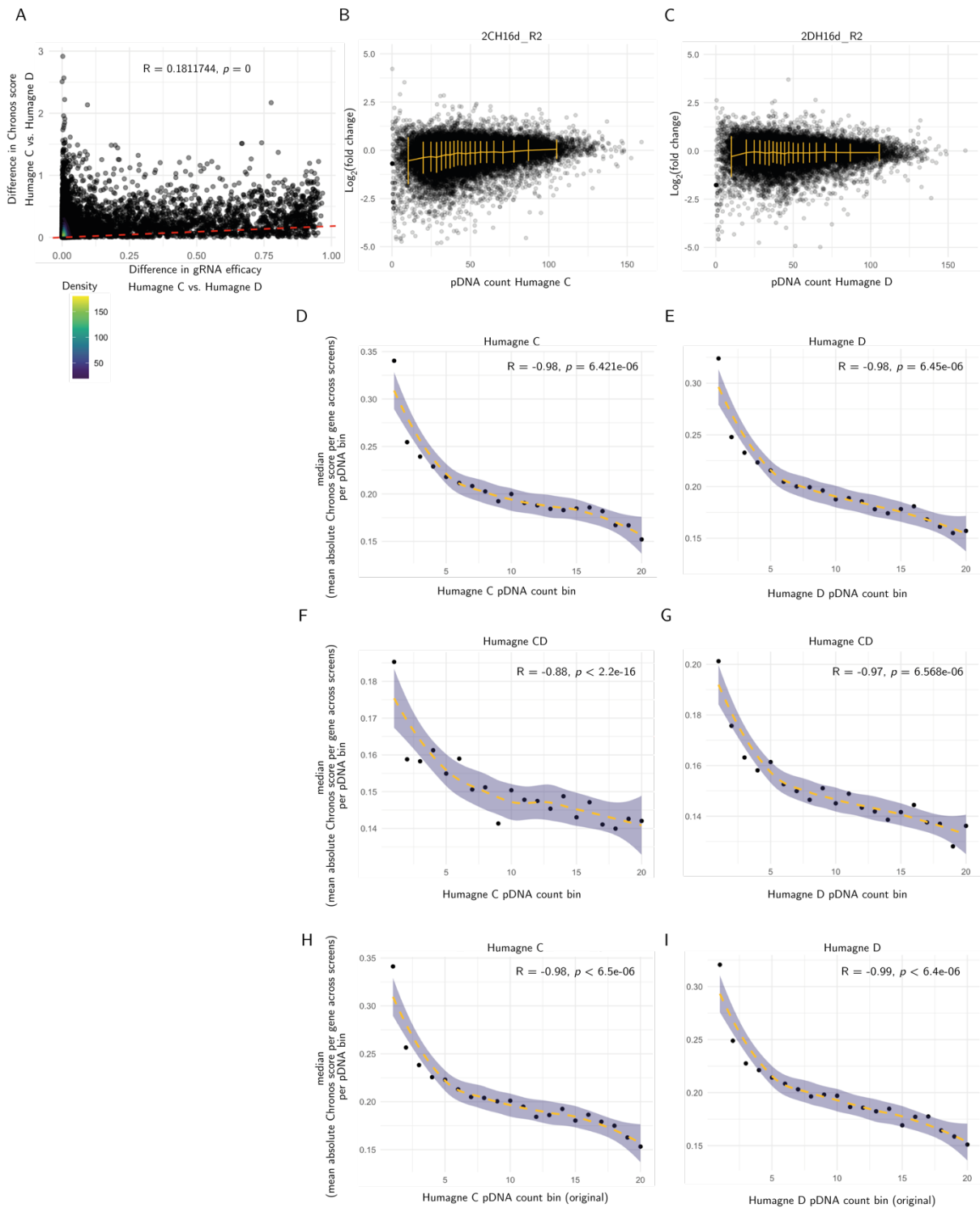


Figure 3.11. CRISPR screen outcomes analysed by Chronos depend on the initial gRNA cassette representation in CRISPR libraries. (A) Dot plot and density plot showing the relationship between the difference in gRNA efficacy and the difference in Chronos score between Humagne C and D. (B) Dot plots showing relationship between the initial gRNA cassette representation (pDNA count) and the $\text{Log}_2(\text{fold change})$ per gRNA cassette for this Humagne C screen in STS26T. Horizontal yellow line connects standard deviations of $\text{Log}_2(\text{fold change})$ values per equally sized bins of pDNA counts (vertical yellow lines). Similar data visualisation was observed for all other screens (data not shown). Screen identifiers are explained in

Table 3.2. **(C)** Same as (B) but for the Humagne D equivalent screen. **(D)** Dot plot and LOESS (yellow dashed line) with 95% confidence interval of the relationship between Humagne C pDNA count bins and the median of the mean absolute Chronos score per gene across Humagne C screens, per pDNA count bin. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot. **(E)** Same as for (D) but with Humagne D screens. **(F)** Dot plot and LOESS (yellow dashed line) with 95% confidence interval of the relationship between Humagne C pDNA count bins and the median of the mean absolute Chronos score per gene across Humagne CD (analysed by Chronos as if one library) screens, per pDNA count bin. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot. **(G)** Same as F but related to Humagne D pDNA count bins. **(H)** Dot plot and LOESS (yellow dashed line) with 95% confidence interval of the relationship between Humagne C original (before library amplification) pDNA count bins and the median of the mean absolute Chronos score per gene across Humagne C screens, per pDNA count bin. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot. **(I)** Same as H but related to Humagne D pDNA original (before library amplification) count bins.

To determine whether the observed effect is specific to the Humagne C and D libraries or is a general characteristic of most CRISPR libraries, the DepMap dataset was analysed using a similar methodology. The Avana and KY libraries typically contain four and five gRNA constructs per gene, respectively, whereas Chronos still generates a single score per gene. To determine if the pDNA effect persists despite the presence of multiple gRNA constructs per gene, an aspect that could potentially dilute the impact of outliers with low representation significantly, pDNA counts were averaged across constructs per gene. Subsequently, the mean absolute Chronos scores for each gene across screens were categorised into equally sized bins based on the averaged pDNA count. The median score for each bin was then computed.

My analysis across the Avana, KY, and Humagne CD screens in the DepMap dataset revealed the same effect, consistently, characterised by high Spearman's correlation coefficients (Rho) and significant p -values of nearing 0 (Fig. 3.12A-F). These

findings indicate that the inclusion of multiple gRNA constructs per gene in a library does not significantly attenuate the impact of the initial pDNA representation on the observed effects, and that the pDNA representation effect seems to impact most CRISPR libraries.

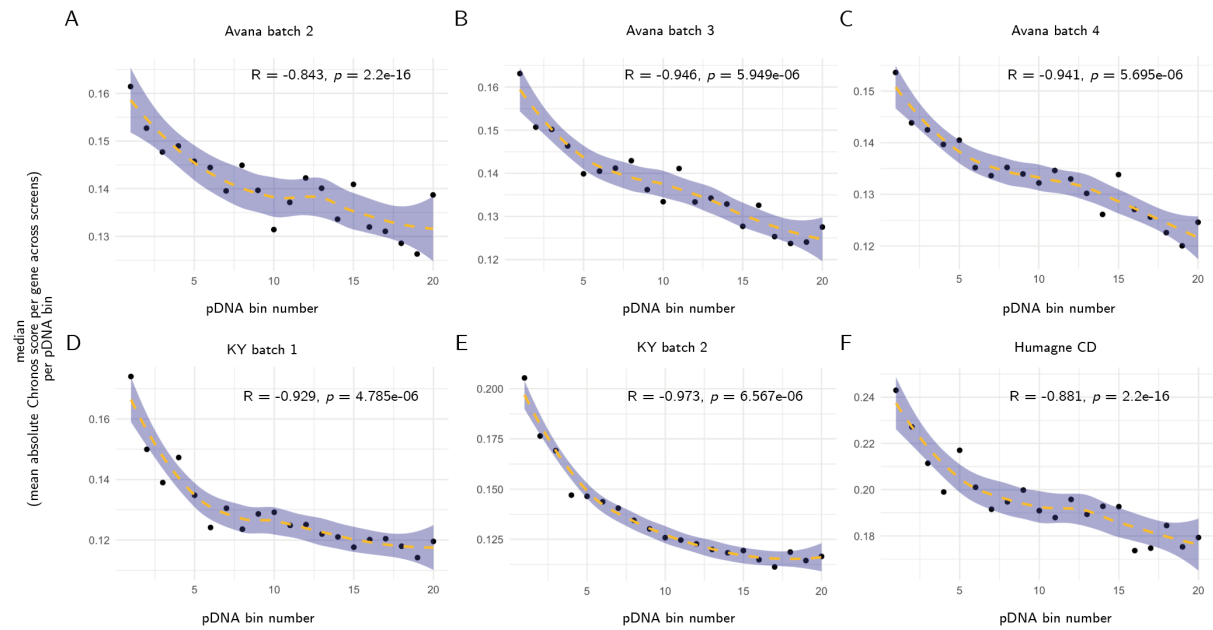


Figure 3.12. CRISPR screen outcomes depend on initial gRNA cassette representation in library in DepMap dataset. (A) Dot plot and LOESS (yellow dashed line) with 95% confidence interval of the relationship between DepMap's Avana batch 2 pDNA count bins and the median of the mean absolute Chronos scores per gene across DepMap's Avana batch 2 screens, per pDNA count bin. The pDNA counts were averaged per gene, because Avana contains four gRNA constructs per gene. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot. (B-F) Same as for (A) but for DepMap's Avana batch 3, Avana batch 4, KY batch 1, KY batch 2, and Humagne CD (only one batch).

To investigate the relationship between pDNA representation and the previously observed differences between CRISPR libraries, I calculated the absolute differences in Chronos scores between Humagne C and D libraries across various cell line models. These differences were averaged per gene across multiple screens, converted to absolute values, and subsequently binned according to associated pDNA counts for each gene in either the Humagne C or Humagne D library. The median of these values was used as the representative measure for each bin (Fig. 3.13A-B).

A Spearman's correlation analysis of these median values across pDNA bins revealed a very strong and highly significant correlation, confirming that pDNA representation is indeed a major factor influencing the variability in outcomes when different CRISPR libraries are used. Furthermore, differences observed between replicate screens (utilising the same library and cell line model) were similarly attributable to the pDNA effect to a comparable degree (Fig. 3.13C-D).

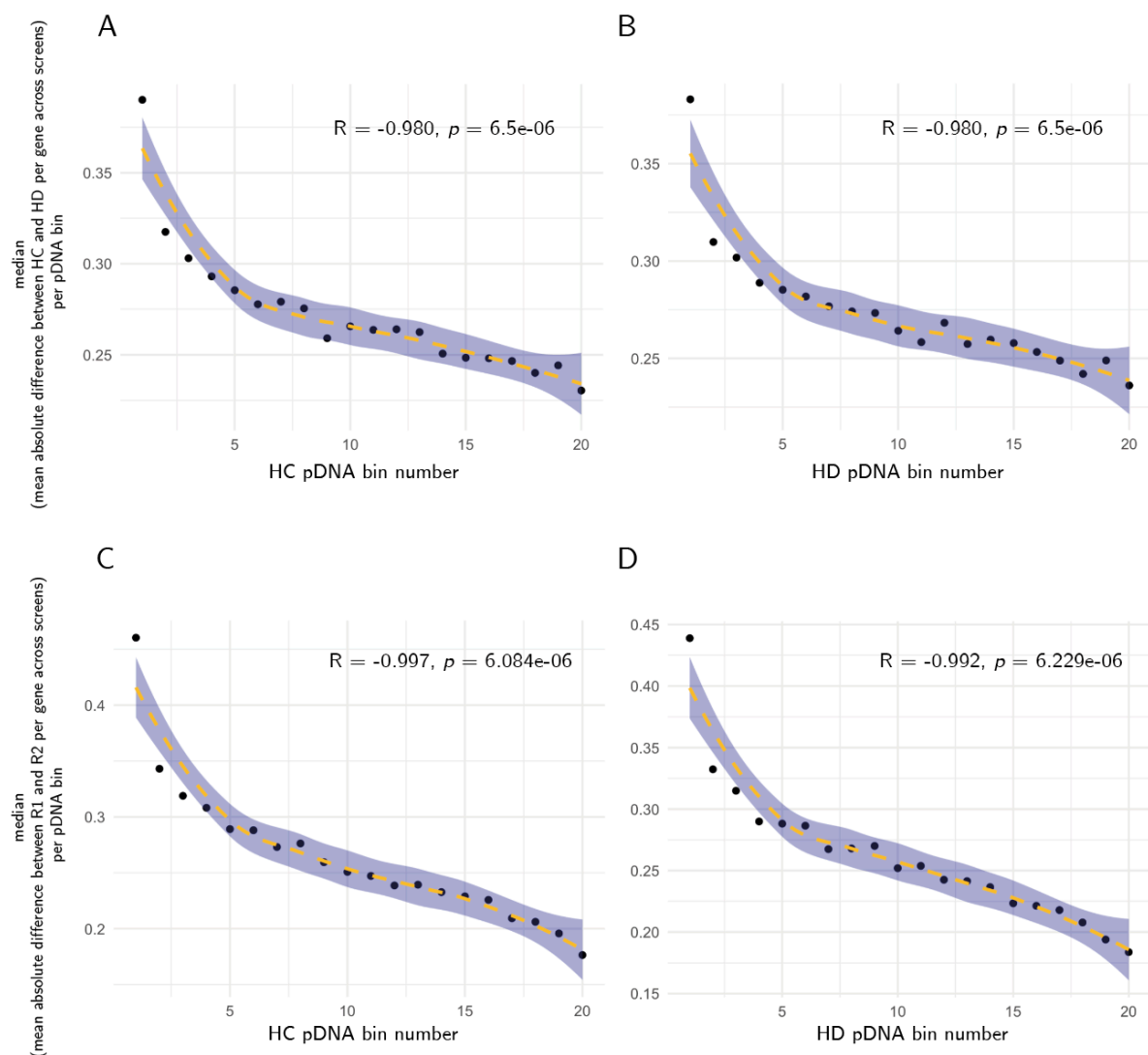


Figure 3.13. Difference between replicate CRISPR screens and Humagne libraries depends on initial gRNA cassette representation. (A) Dot plot and LOESS (yellow dashed line) with 95% confidence interval of the relationship between Humagne C pDNA count bins and the median of the mean absolute difference between Humagne C and D Chronos scores per gene across screens, per Humagne C pDNA count bin. A

Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot. **(B)** Same as (A) but related to Humagne D pDNA count bins. **(C)** Dot plot and LOESS (yellow dashed line) with 95% confidence interval of the relationship between Humagne C pDNA count bins and the median of the mean absolute difference between first and second replicate Chronos scores per gene across replicate Humagne C screens, per pDNA count bin. A Spearman's rank correlation test was performed to test for significance, with Rho (R) and p value displayed on the plot. **(D)** Same as in (C) but for Humagne D CRISPR screens.

3.2.13 Correction for the effect of initial gRNA representation

Previous analyses of data in this study and the DepMap data indicates a widespread problem with initial gRNA representation that results in potentially false positive signals. In an attempt to correct for this effect, a z-score was computed for each Chronos score relative to the distribution of scores within the same pDNA bin. This correction marginally, but significantly, improved the correlation between different replicates as well as the correlation between screens done in the same cell line model but with different libraries (Humagne C vs. D) (Fig. 3.14A-B). No significant improvement was observed for the recall of common essential genes (Fig. 3.14C). Together these findings suggest that there may be a moderate but significant benefit to calculating the z-score of Chronos scores per pDNA bin to correct for the effect related to initial pDNA representation. Consequently, the z-score normalised Chronos scores per pDNA bin will be used in all subsequent analyses, unless specified otherwise.

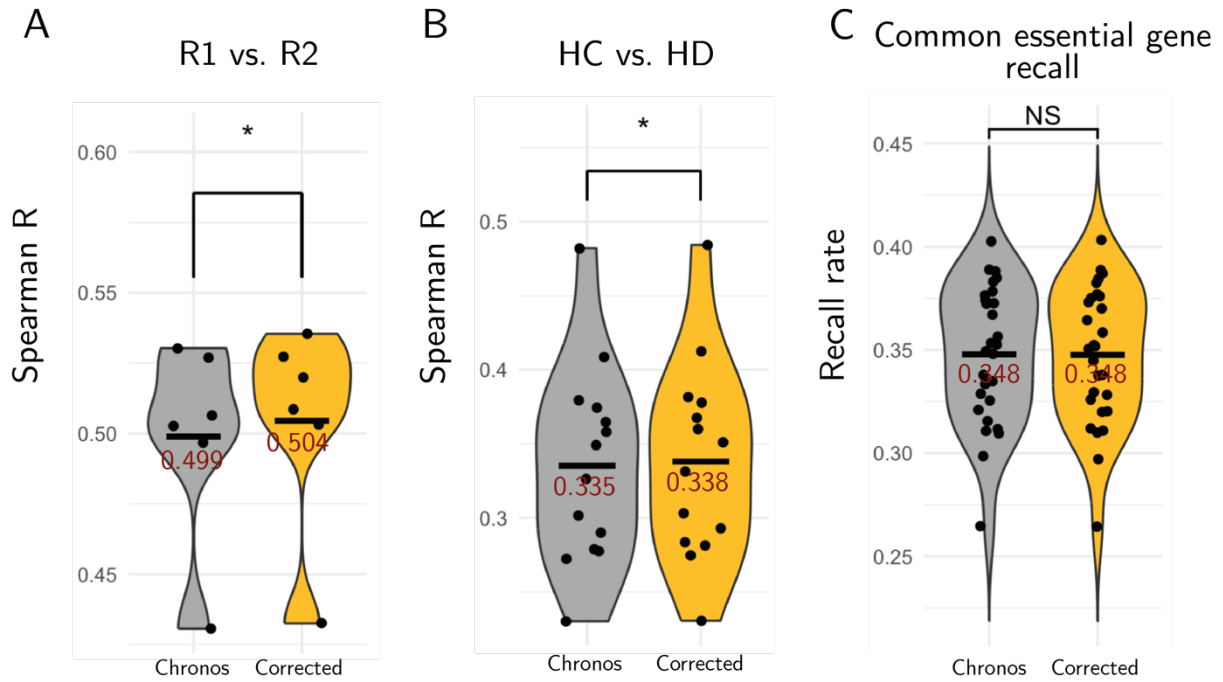


Figure 3.14. Correction for the effect of initial gRNA cassette representation improves similarity between replicate CRISPR screens. (A) Violin plots displaying Spearman's Rho (R) values of correlations between replicate screens analysed using Chronos (grey) and the Chronos score corrected for the initial gRNA cassette representation effect by calculating the z-score for Chronos scores per pDNA bin (golden violin). Means are represented by horizontal black lines and mean values indicated in red above. A Wilcoxon signed-rank test was performed to compare the distribution of Spearman's Rho (R) values between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS not significant. **(B)** Violin plots displaying Spearman's Rho (R) values of correlations between Humagne C and D screens in the same cell line model analysed using Chronos (Gray) and Chronos scores corrected for the initial gRNA cassette representation effect by calculating the z-score for Chronos values per pDNA bin (gold). Means are represented by horizontal black lines and mean values indicated in red above. A Wilcoxon signed-rank test was performed to compare the distribution of Spearman's Rho (R) values between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS not significant. **(C)** Violin plots displaying common essential recall rate values of all Humagne C and D screens performed in this study, analysed using Chronos (grey) and Chronos scores corrected for the initial gRNA cassette representation effect by calculating the z-score Chronos values per pDNA bin (gold). Means are represented by horizontal black lines and mean values indicated in red above. A Wilcoxon signed-rank test was performed to compare the distribution of Spearman's Rho (R) values between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS not significant.

3.3 Discussion

The results presented in this chapter provide an overview of the development and validation of the necessary components of enAsCas12a-based genome-wide CRISPR knockout screens, specifically tailored to provide a comprehensive methodology and insights to aid the identification of novel dependencies in MPNST cell line models. Additionally, I explored variables that may amplify noise and that may cause bias in CRISPR screen data.

The establishment of cell lines stably expressing enAsCas12a, the CRISPR enzyme, is critical for the success of CRISPR screens, where non-expressing cells limit gRNA targeting following CRISPR library transduction. In this study, cell lines were transduced and sorted in bulk rather than single cell-sorted, to allow for quicker culture expansion and limit silencing of enAsCas12a expression over time. While most cell lines utilised in this study sustained a relatively high level of enzyme expression throughout the necessary periods for cell expansion and subsequent CRISPR library transduction, this was not the case with STS26T and s462TY. These particular cell lines exhibited a significant proportion of cells with diminished or lost transgene expression over time. This decline in expression is hypothesised to be the primary cause of the reduced effect size and overall lower quality of screens in these cell lines, as evidenced by poorer recall of common essential genes.

The variability in transgene expression following lentiviral transduction is a well-documented epigenetic phenomenon [141]. A range of solutions has been proposed to mitigate this issue, with one prominent strategy being the incorporation of genomic insulators flanking the transgene. These insulators can minimise positional effects,

thereby stabilising transgene expression over extended periods [142]. Although an alternative approach involves delivering the CRISPR enzyme together with the CRISPR library, on the same lentiviral target plasmid, this method is not universally applicable, as not all libraries are designed for such co-delivery, as the increased plasmid size associated with this approach may negatively impact transduction efficiency [143].

Also, the amplification and validation of CRISPR libraries are critical components of the CRISPR screening pipeline, providing valuable insights into potential bottlenecks that can influence the outcomes of CRISPR screens. Notably, the parameters employed for library validation, especially the skew ratio, indicated that when the amplification of CRISPR libraries is executed well, the representation of gRNAs remains largely unchanged. This finding is particularly important, as it lays to rest concerns about whether any observed library bias could stem from errors in the in-house handling of these processes.

The implementation of the Chronos algorithm provides a significant advancement in CRISPR screen data analysis. Chronos not only enhances the recall of common essential genes but also improves the reproducibility of results across biological replicates. Additionally, it has demonstrated increased consistency between screens conducted in the same model using different libraries. However, despite these advantages, Chronos has its limitations. Persistent variability between different CRISPR libraries remains an issue that Chronos alone cannot fully address, underscoring the inherent challenges in interpreting CRISPR screen data.

While differences in gRNA efficacies contribute to the variability observed between CRISPR libraries, this study provides evidence that the initial representation of gRNAs in

the library (pDNA count) may play a more significant role. The initial gRNA representation effect was found to be strongly correlated with higher Chronos scores, indicating its partial responsibility for discrepancies between replicates and between different CRISPR libraries. Despite this, determining which gRNA effects are false positives remains challenging, as evidenced by the marginal improvements in data quality when z-scores for gene effects are calculated within the same pDNA bin.

Further refinement of the Chronos algorithm could involve modelling pDNA representation as a covariate and may provide more substantial improvements to the data as measured by common quality metrics. A recent study offers a more fundamental approach to addressing these issues, however, by demonstrating that key steps in CRISPR library production, such as the direction of oligo synthesis, the number of PCR cycles, and the temperature during cloning, can significantly improve library representation [144]. Although the primary aim of this study was to facilitate CRISPR screen scaling, given their labour-intensive nature, these techniques may also enhance downstream effects of pDNA representation. However, this approach does not provide a solution for the thousands of CRISPR screens that have already been conducted. For these existing datasets, the effects might still be mitigated through careful adjustments during the data analysis stage. Further research is required to determine to what extent pDNA representation is responsible for the impact on gene effect scores and whether this is truly quantifiable.

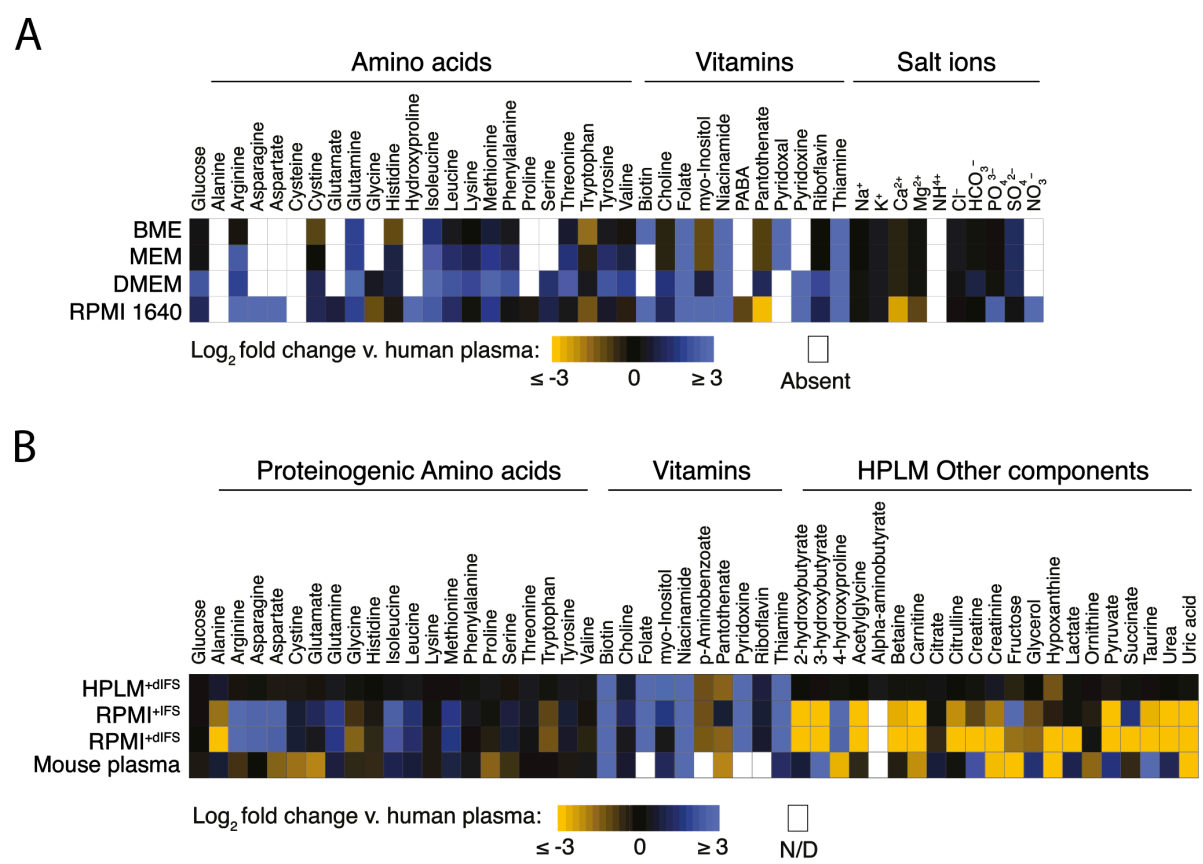
Chapter 4: Context dependent gene essentiality and hit prioritisation

4.1 Introduction

Genome-wide enAsCas12a-based CRISPR drop-out screen outcomes are influenced not only by intrinsic factors such as gRNA library efficiency and enAsCas12a expression levels but also by variables related to the specific cell line model being studied. A critical requirement for effective cancer drug discovery is the use of models that accurately represent the tumour being targeted for treatment. A major reason for the frequent failure of clinical trials is the low efficacy of drugs, where the therapeutic interventions do not yield the same outcomes in patients as observed in preclinical cancer models [145]. This underscores the significant discrepancies between current cancer models and actual patient tumours. With genome-wide CRISPR knockout screens, the cell model context is commonly distorted from physiological conditions and so biasing the identified cellular dependencies. These discrepancies arise from inherent limitations of the cell line models themselves, as well as from differences in the cellular microenvironment.

A key variable that can be modulated during high-throughput screens in cell line models is the medium context. Human Plasma-Like Medium (HPLM) has been developed to better mimic human blood plasma compared to conventional media such as either DMEM or RPMI [146]. HPLM is formulated with metabolites and ions, including glucose, amino acids, vitamins, and salts, at concentrations that closely reflect those found in human blood plasma (Fig. 4.1). This composition allows HPLM to provide a more physiologically relevant representation of the human metabolic environment. For instance,

HPLM has been shown to significantly impact cellular metabolism, such as the inhibition of de novo pyrimidine synthesis by uric acid—a metabolite present at physiological levels in HPLM but absent in DMEM and RPMI[147]. Despite its potential advantages, HPLM has been predominantly applied to cell lines derived from haematological cancers, with its use in cell lines from solid tumours remaining largely unexplored[146], [148]. Furthermore, only a single study has investigated the impact of HPLM on cellular dependencies in genome-wide CRISPR screens[148]. This study demonstrated that different medium contexts can alter cellular dependencies, which were linked to metabolism-related functions of the genes involved. Together, these findings underscore the importance of considering the medium context as a critical component of the cell line model to more accurately represent the cancer type targeted for treatment.



1640, compared to adult human plasma, using $\log_2$ -transformed fold changes. Components that are not found in a given medium are labelled as absent. **(B)** Heatmap illustrating the relative concentrations of specified components in the indicated media and mouse plasma relative to human plasma ($\log_2$ -transformed fold changes). N/D: fold change could not be determined. RPMI+IFS: RPMI 1640 with 5 mM glucose and 10% IFS. RPMI + dIFS: RPMI 1640 with 5 mM glucose and 10% dialyzed IFS. HPLM + dIFS: HPLM containing 10% dialyzed IFS (adapted from [146]).

Beyond the medium context, the genomic context also plays a crucial role in determining cellular gene dependencies. The loss of specific genes, either as driver mutations that promote cancer progression or as passenger mutations, creates vulnerabilities that lead to dependencies on other genes to maintain essential cellular functions. Understanding the impact of these genetic alterations is also vital for elucidating the origins of the specific cellular dependencies identified in CRISPR screens as this may guide more precise therapeutic strategies[149].

In this chapter, I investigate various contexts of the cancer cell line model that influence CRISPR screen outcomes. Target cell lines were characterised in HPLM compared to DMEM, utilising growth assays and transcriptomic analyses. This was complemented by genome-wide CRISPR screen analyses under both culture medium conditions to assess the impact of the medium context on screen results. Next, genetic information, specifically copy number variation, were incorporated into the hit prioritisation strategy by leveraging a recently established database of predicted human synthetic lethality partners. Finally, these findings were integrated with DepMap CRISPR screen datasets derived from MPNST cell line models and combined with information on known targeting drugs. Using this approach, candidate genes were identified for further validation as potential therapeutic targets for MPNST.

Importantly, the group that developed HPLM also created dialyzed heat-inactivated foetal bovine serum (IFS). However, for cost-effectiveness, FBS was used in place of dialysed IFS, as the differences between dialyzed IFS and FBS were relatively minor compared to the differences between the base culture media (DMEM or RPMI versus HPLM). Thus, the most substantial effects were anticipated to arise from the use of HPLM rather than from the serum choice.

Moreover, CRISPR screens in HPLM were not conducted for the s462TY cell line because, at the time of these experiments, the Chronos analysis tool had not yet been developed. Without Chronos, screens in DMEM for this cell line were considered too low in quality to justify continuing in HPLM. Nonetheless, screens performed in s462TY cell lines were incorporated into the hit prioritization analysis across all screens conducted in MPNST cell lines, including those utilising HPLM.

4.2 Results

4.2.1 Cell line-specific transcriptomic responses to HPLM culture

To characterise the STS26T, 90-8, ST-8814, and iHSC cell lines for the planned HPLM CRISPR screens, cells were cultured in HPLM and DMEM for 21 days, after which they were snap-frozen, and cell pellets underwent RNA extraction and paired-end mRNA sequencing. A principal component analysis (PCA) of the whole transcriptome successfully distinguished cells grown in HPLM from those in DMEM along the first principal component (PC1) for the STS26T and iHSC cell lines (Fig. 4.2A, D). PC1 accounted for 98% and 97% of the variance between the samples for STS26T and iHSC, respectively. However, for the 90-8 and ST-8814 cell lines, one biological replicate of cells grown in HPLM overlapped with the data points of cells grown in DMEM on the

PC1 axis, indicating less apparent distinction between the two media conditions for these replicates.

To visualise the differential gene expression between HPLM and DMEM conditions, volcano plots of Log₂ fold-change were generated (Figure 4.2E-H). In the STS26T and iHSC cell lines, a significant upregulation of 2,107 and 937 genes, respectively, was observed under HPLM conditions, characterised by a log₂(fold change) greater than 1 and an adjusted p-value below 0.05. Conversely, 1,065 and 521 genes were significantly downregulated in these cell lines under the same conditions. Notably, the 90-8 and ST-8814 cell lines exhibited markedly fewer differentially expressed genes, with only 20 and 162 genes showing significant upregulation, and 4 and 44 genes exhibiting significant downregulation, respectively. The top 10 over- and under expressed genes in HPLM for each cell line, visualised in heatmaps (Figure 4.2I-L), demonstrated no overlap across the different cell lines. Consistent with these findings, GSEA for each cell line (Supplemental Table 1) revealed a diverse and cell line-specific range of enriched and depleted gene sets. These observations align with previous reports from the group that originally developed HPLM, highlighting the distinct impact of HPLM on gene expression across various cell line models[146]. These results underscore the heterogeneous effects of HPLM on gene expression, emphasising the variability of cellular responses to this culture condition.

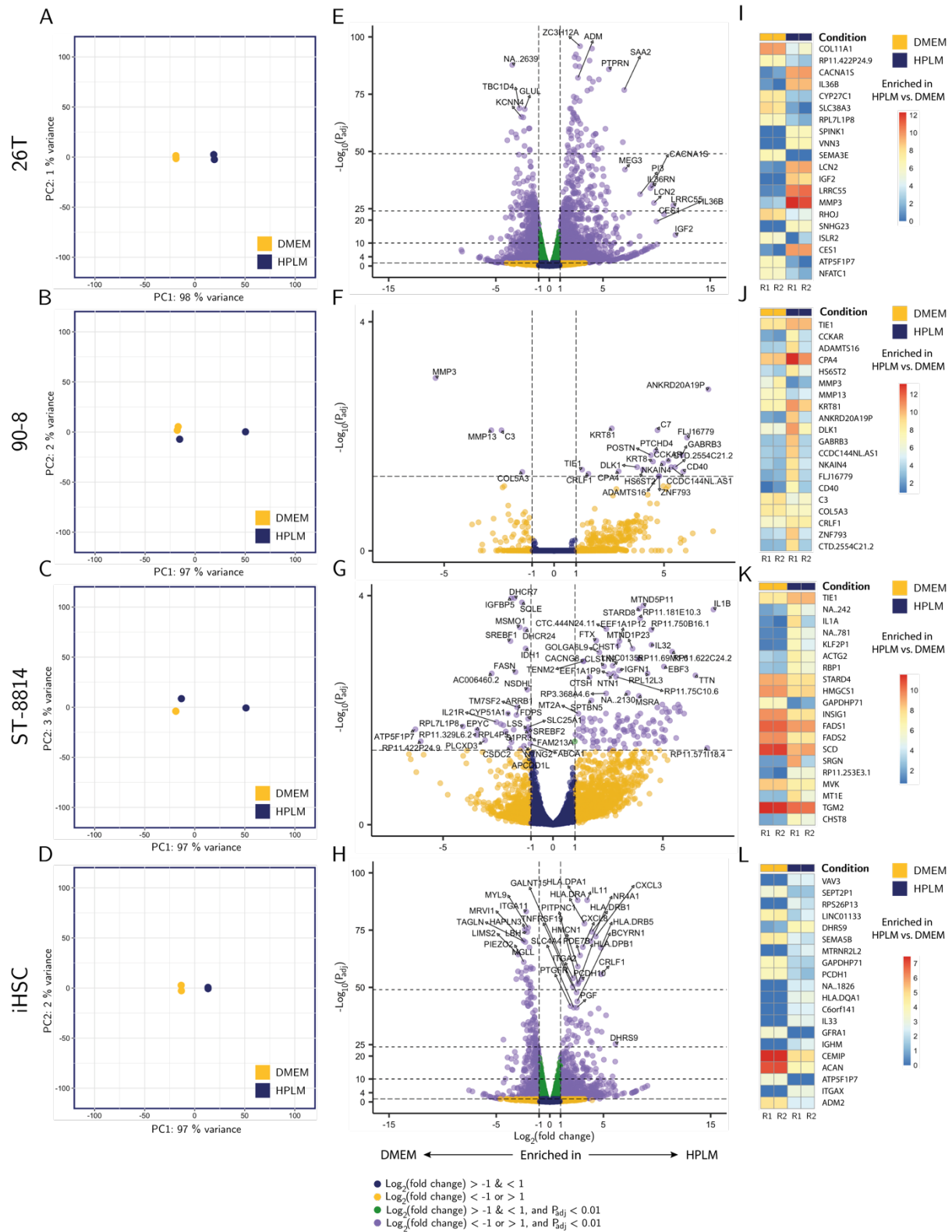


Figure 4.2. Culture medium context-dependent differential gene expression is cell line dependent. (A,B,C,D) Principal component analysis of target cell lines STS26T (A), 90-8 (B), ST-8814 (C), and iHSC (D) grown in DMEM (gold) and HPLM (blue). Each dot is one biological replicate. Variance per principal component is indicated on the side of each axis. PC1: principal component 1, PC2: principal component 2. (E-H) Volcano plots displaying genome wide mRNA transcriptomics of target cell lines grown

in HPLM versus DMEM with level and significance colour coded as indicated. Top differentially expressed genes are labelled by name with limited overlaps allowed. **(I-L)** heatmaps of the top 10 differentially over and under-expressed genes in the same target cell lines, in the same order, grown in HPLM versus DMEM. Each column is one biological replicate. R1: replicate 1, R2: replicate 2.

Principal Component Analysis (PCA) of transcriptomic profiles from target cell lines cultured in HPLM and DMEM suggested that while there are consistent shifts in gene expression patterns due to the culture medium, these shifts are relatively subtle (Fig. 4.3A). This also indicates that although the transcriptomic differences between cells grown in HPLM and DMEM are present they may not reflect a differential impact on cell function, supported to some extent by cell populations clustering primarily based on cell type rather than the culture medium.

Evaluating the expression levels of individual genes reveals specific transcriptomic changes between the two media conditions. As shown in Figure 4.3B, only 121 genes were differentially expressed, with 100 genes overexpressed and 21 under expressed in HPLM compared to DMEM, based on a $\log_2(\text{fold change})$ threshold of 1 and -1, respectively. Despite these significant changes, a limited number of differentially expressed genes were observed as shown by the heatmap of the top differentially expressed genes (Fig. 4.3C). These findings suggest that culturing cells in HPLM versus DMEM induces selected transcriptomic alterations that are cell line dependent rather than widespread alterations in the same gene signatures.

Some of the selected transcriptomic changes are reflected in the enrichment of specific gene sets in HPLM or DMEM, as identified by GSEA (Fig. 4.3D and Table 4.1). Notably, metabolic reprogramming is evident through various gene set enrichment in both culture media. In DMEM, gene sets related to lipid metabolism, particularly cholesterol

homeostasis, adipogenesis, and synthesis of unsaturated fatty acids gene sets, were enriched, suggesting that DMEM may facilitate a metabolic shift towards enhanced lipid storage.

In contrast, HPLM showed enrichment in gene sets associated with oxidative phosphorylation and respiratory electron transport, indicating enhanced mitochondrial activity and energy production. The elevated levels of oxidative phosphorylation in HPLM may be due to its lower glucose concentration (5 mM) compared to the 25 mM in high-glucose DMEM, the standard medium for these cell lines. The reduced glucose availability in HPLM may shift cellular metabolism toward a greater dependence on mitochondrial energy production [150].

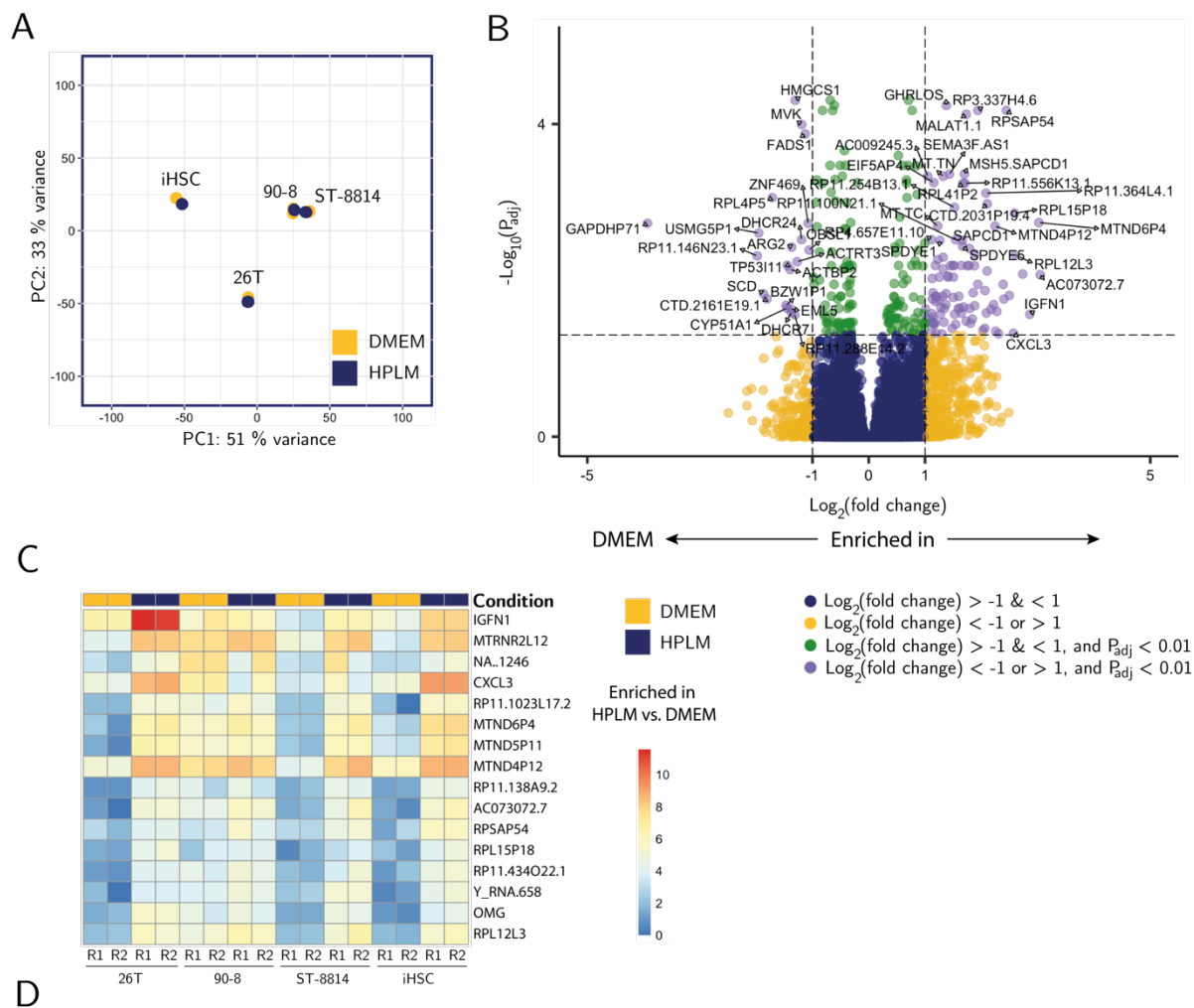


Figure 4.3. Culture medium context-dependent differential gene expression shows consistent effects across cell lines. Principal component analysis of target cell lines STS26T, 90-8, ST-8814 and iHSC grown in DMEM (gold) and HPLM (blue). Each dot is one biological replicate with two biological replicates for each condition for each cell line. Variance per principal component is indicated on the side of each axis. PC1: principal component 1, PC2: principal component 2. **(B)** heatmap of the top 16 differentially expressed genes in the same target cell lines grown in HPLM versus DMEM. Each column is one biological replicate. R1: replicate 1, R2: replicate 2. **(C)** Volcano plot displaying genome-wide mRNA transcriptomics of the same target cell lines, with two biological replicates each, grown in HPLM versus DMEM combined, with level and significance colour coded as indicated. Top differentially expressed genes are labelled by name with limited overlaps allowed. **(D)** Enrichment plots following GSEA of combined transcriptomics of STS26T, 90-8, ST-8814 and iHSC cells cultured in HPLM versus DMEM. Gene sets tested were Hallmarks, KeggLegacy and Reactome (Table 4.1).

Table 4.1. GSEA results all cell lines combined.

Gene set	Enriched in	FDR q value (<25%)
Hallmark_cholesterol_homeostasis	DMEM	0.024
Hallmark_peroxisome	DMEM	0.072
Hallmark_wnt_beta_catenin_signaling	DMEM	0.245
Hallmark_androgen_response	DMEM	0.214
Hallmark_adipogenesis	DMEM	0.194
Hallmark_mtorc1_signaling	DMEM	0.191
Kegg_ribosome	HPLM	0.14
Kegg_protein_export	HPLM	0.225
Kegg_hematopoietic_cell_lineage	HPLM	0.179
Kegg_type_i_diabetes_mellitus	HPLM	0.181
Kegg_parkinsons_disease	HPLM	0.155
Kegg_oxidative_phosphorylation	HPLM	0.155
Kegg_biosynthesis_of_unsaturated_fatty_acids	DMEM	0.055
Kegg_ppar_signaling_pathway	DMEM	0.078
Kegg_endometrial_cancer	DMEM	0.136

Kegg_inositol_phosphate_metabolism	DMEM	0.102
Kegg_mtor_signaling_pathway	DMEM	0.121
Kegg_citrate_cycle_tca_cycle	DMEM	0.103
Kegg_notch_signaling_pathway	DMEM	0.105
Kegg_aminoacyl_trna_biosynthesis	DMEM	0.111
Kegg_acute_myeloid_leukemia	DMEM	0.105
Kegg_steroid_biosynthesis	DMEM	0.115
Kegg_glycosaminoglycan_biosynthesis_heparan_sulfate	DMEM	0.15
Kegg_phosphatidylinositol_signaling_system	DMEM	0.141
Kegg_non_small_cell_lung_cancer	DMEM	0.131
Kegg_b_cell_receptor_signaling_pathway	DMEM	0.228
Kegg_fc_gamma_r_mediated_phagocytosis	DMEM	0.212
Kegg_lysinе_degradation	DMEM	0.209
Kegg_vegf_signaling_pathway	DMEM	0.208
Kegg_peroxisome	DMEM	0.205
Reactome_sars_cov_1_modulates_host_translation_machinery	HPLM	0.156
Reactome_srp_dependent_cotranslational_protein_targeting_to_membrane	HPLM	0.139
Reactome_eukaryotic_translation_elongation	HPLM	0.191
Reactome_activation_of_the_mrna_upon_binding_of_the_cap_binding_complex_and_eifs_and_subsequent_binding_to_43s	HPLM	0.183
Reactome_eukaryotic_translation_initiation	HPLM	0.249
Reactome_respiratory_electron_transport	HPLM	0.222
Reactome_response_of_eif2ak4_gcn2_to_amino_acid_deficiency	HPLM	0.204
Reactome_regulation_of_expression_of_slits_and_robo	HPLM	0.193
Reactome_nonsense_mediated_decay_nmd	HPLM	0.194
Reactome_complex_i_biogenesis	HPLM	0.206

Reactome_influenza_infection	HPLM	0.234
Reactome_selenoamino_acid_metabolism	HPLM	0.226
Reactome_sars_cov_1_host_interactions	HPLM	0.213
Reactome_respiratory_electron_transport_atp_synthesis_by_chemiosmotic_coupling_and_heat_production_by_uncoupling_proteins	HPLM	0.22

4.2.2 Medium-context dependent cell culture versus physiological transcriptomics

HPLM was originally developed to bridge the physiology gap between *in vitro* cellular models and the human diseases they are intended to represent. Unlike conventional cell culture media, HPLM more closely mimics the biochemical composition of human plasma, aiming to provide a more physiologically relevant environment for cell growth. However, despite offering a closer approximation to physiological conditions, it remains unclear whether HPLM effectively reprograms cellular models to more accurately reflect true human cellular physiology. This critical aspect was neither conclusively demonstrated in the original study introducing HPLM, nor has it been fully addressed in subsequent research. In addition, the selection events in cells grown primarily in DMEM, that are then switched to HPLM, may skew the impact of HPLM post-evolutionary selection of cells in culture.

To first assess whether HPLM genuinely enhances the physiological relevance of *in vitro* cell culture models when it comes to identifying novel MPNST dependencies, an optimal strategy would involve comparison of the overall cellular metabolism in these models with that of authentic MPNST cells. However, accurately measuring the myriad components of cellular metabolism remains challenging, despite some commendable efforts in the field [151], [152]. Genome-wide transcriptomics offers a more feasible

approach to explore this question, though it is important to recognise that transcriptomic data provide an indirect and potentially incomplete representation of the metabolic processes and protein functions that define cellular phenotypes. Despite this limitation, significant metabolic shifts have been demonstrated to correlate with substantial changes in gene expression[153], [151], which can be effectively captured using RNA sequencing technologies.

In this study, bulk genome-wide mRNA transcriptomic data from sixteen MPNST samples were retrieved from the database of Genotypes and Phenotypes (dbGaP) and compared with target cell lines cultured in both HPLM and DMEM media. Principal component analysis (PCA) revealed distinct clustering patterns among the cell lines and MPNST samples, independent of the growth medium (Fig. 4.4A). Notably, the MPNST cell lines 90-8 and ST-8814 clustered together along the first and second principal components but remained distinct from the MPNST tumour samples, which formed a separate cluster. Interestingly, the STS26T and iHSC cell lines also clustered together at a similar distance from the MPNST samples, comparable to the clustering of 90-8 and ST-8814, despite not being true MPNST models. It is worth noting that STS26T was initially diagnosed as an MPNST but is now believed to be of melanoma origin [122]. The iHSC line consists of human Schwann cells immortalised through the exogenous expression of telomerase reverse transcriptase and murine cyclin-dependent kinase[106]. Although these findings suggest that, despite not originating from MPNST, these cell lines may still be more similar to MPNST tumours than most other cell lines, it is important to recognise that PCA can sometimes oversimplify complex biological

variability, potentially masking important subtleties in the data that might influence clustering[154].

Regarding the influence of the growth medium, the HPLM context appeared to bring the ST-8814, STS26T, and iHSC cell lines slightly closer to the MPNST samples along the first principal component, which explains 37% of the variance in the data. However, this effect was only marginal and was not observed for the 90-8 cell line. In the second principal component, which accounts for 14% of the variance, the STS26T and iHSC cell lines demonstrated a closer overlap with the MPNST samples compared to the 90-8 and ST-8814 cell lines.

To extend the analysis beyond PCA, Spearman's correlations were calculated between the cell lines, grown in either DMEM or HPLM, and the MPNST samples, considering all possible combinations. The correlations between the cells cultured in DMEM and the MPNST samples were then compared to those between the HPLM-cultured cells and the MPNST samples (Fig. 4.4B-F). Interestingly, these results indicate that cell lines cultured in DMEM exhibited a marginal but statistically significant closer correlation to the MPNST samples compared to those cultured in HPLM. These findings suggest that, based on the transcriptomic data alone, DMEM might offer a slightly more physiological context. However, it is important to recognise that this is merely a reflection of the differences in cellular responses to the different media types rather than an absolute measure of physiological relevance.

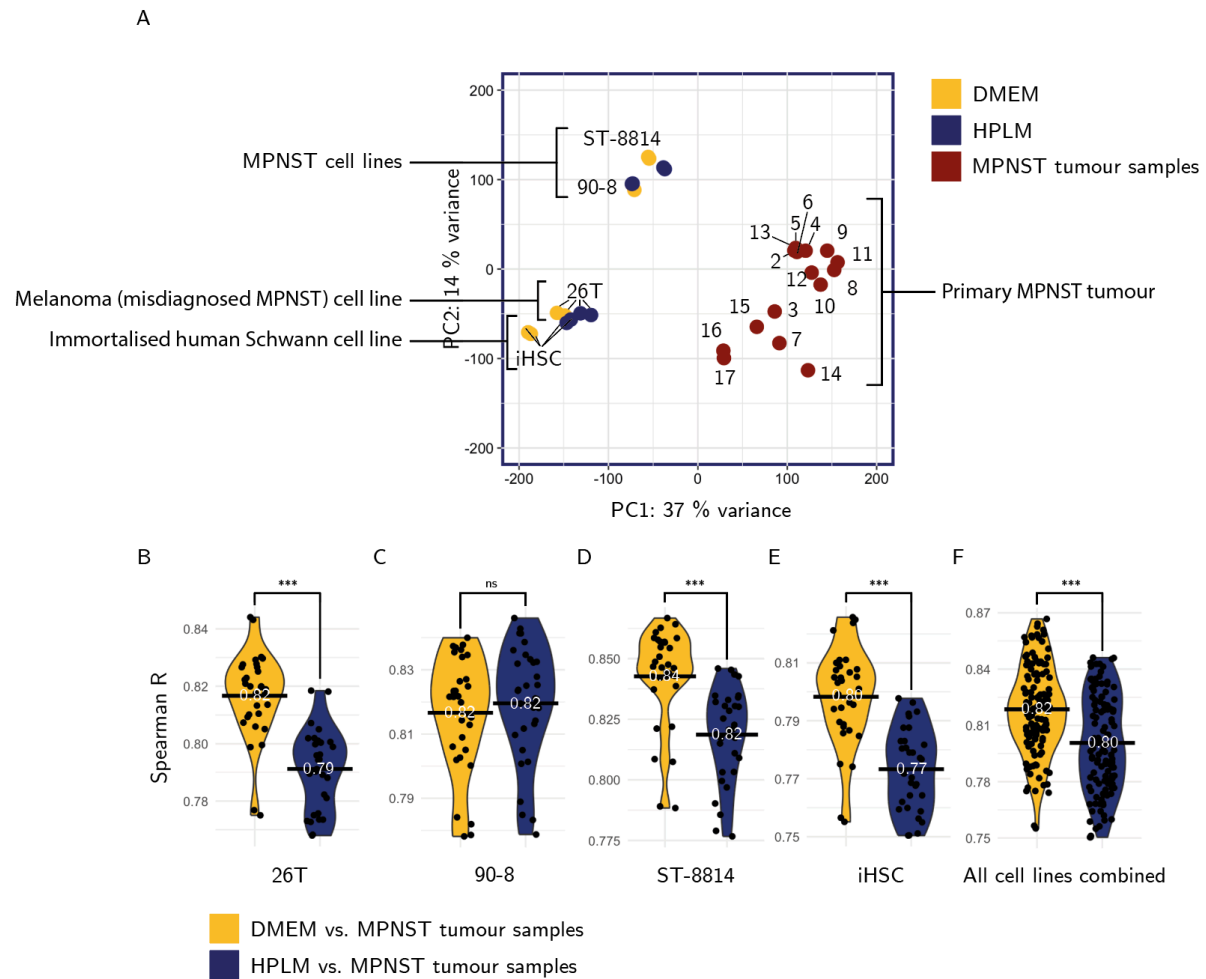


Figure 4.4. DMEM medium context cell line transcriptomics show closer correlation with primary MPNST tumour samples compared to HPLM. (A) Principal component analysis of target cell lines STS26T, 90-8, ST-8814 and iHSC grown in DMEM (gold) and HPLM (blue) and sixteen MPNST tumour samples (red) obtained from the database of Genotypes and Phenotypes (dbGaP). Each dot is one biological replicate with two biological replicates for each condition for each cell line and one for each tumour sample. Variance per principal component is indicated on the side of each axis. PC1: principal component 1, PC2: principal component 2. **(B)** Violin plots of Spearman's correlation Rho values for each possible comparison between STS26T cells grown in DMEM and HPLM, and MPNST tumour samples. A Wilcoxon rank-sum test was used to compare the distribution of Spearman's Rho between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Spearman R: Spearman's Rho values. **(C-E)** Same as (B) but with 90-8, ST-8814 and iHSC, respectively. **(F)** Violin plots of Spearman's correlation Rho values for each possible comparison between STS26T, 90-8, ST-8814 and iHSC cells grown in DMEM and HPLM, and MPNST tumour samples. A Wilcoxon rank-sum test was used to compare the distribution of Spearman's Rho between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Spearman R: Spearman's Rho values. Note: STS26T is since recently considered a melanoma cell line but was initially presented as an MPNST cell line.

4.2.3 HPLM culture population doubling rate alters the CRISPR screen workflow

To advance the genome-wide CRISPR screen experiments with STS26T, 90-8, ST-8814, and iHSC cultured in HPLM, their population doubling rates were compared to those in DMEM (Fig. 4.5A). Upon transitioning from DMEM to HPLM, most cell lines exhibited an adaptation phase characterised by population growth. Following this adjustment period, population doubling rates increased, with some lines, notably 90-8 and ST-8814, displaying exponential growth. Over a 40-day period, STS26T, ST-8814, and iHSC demonstrated slower population doubling rate in HPLM compared to DMEM. In contrast, 90-8 reached a comparable population doubling rate to that observed in DMEM after approximately 17 days, while iHSC achieved similar growth after about 35 days. These results indicate that during the initial weeks of culturing in HPLM, cells exhibit a significantly reduced population doubling rates compared to DMEM, consistent with previous reports[146]. Interestingly, this was not reflected by findings in differential gene expression analysis, further displaying the limitations of mRNA-wide transcriptomic analysis.

The information on population doubling rates was crucial for the planned CRISPR screen experiments. The reduced growth rate observed in HPLM posed significant challenges for scaling up cells in sufficient numbers for CRISPR library transduction while maintaining adequate levels of EnAsCas12 expression. Based on the established growth curves, it was determined that achieving the necessary cell numbers in HPLM would take approximately twice as long as in DMEM. Additionally, cells cultured in HPLM were found to be less resilient under selective pressure from puromycin, despite possessing a

resistance gene following CRISPR library transduction (data not shown). This limitation implied that performing puromycin selection in HPLM would require an impractically large number of cells, particularly given the intended number of eight screens in four target cell lines (Humagne C and D). To address these constraints, the medium switch to HPLM was implemented only after the puromycin selection stage, as outlined in Figure 4.5E. Although the effect of the medium switch may not be immediate, the integration of multiple late time points during Chronos data analysis will allow the capturing of specific effects attributable to differences in medium context.

To specifically capture changes in gRNA representation attributable to the HPLM-medium context, the cells harvested on the day of the medium switch were used as the baseline for gRNA representation during CRISPR screen data analysis. To ensure sufficient competition among perturbed cells under HPLM conditions, the cells were cultured to achieve at least the same number of population doublings as in the equivalent DMEM screens, which were generally limited to fourteen doublings. As previously mentioned, Chronos is particularly well-suited as an analysis tool for this experiment due to its ability to integrate variable time points and adjust for differences in cell line growth rates. Moreover, Chronos can detect and correct for the outgrowth of specific perturbed cell populations. For example, in this study, the knockout of PTEN resulted in exceptional growth rates, leading to this population constituting over 50% of the total cell population (data not shown).

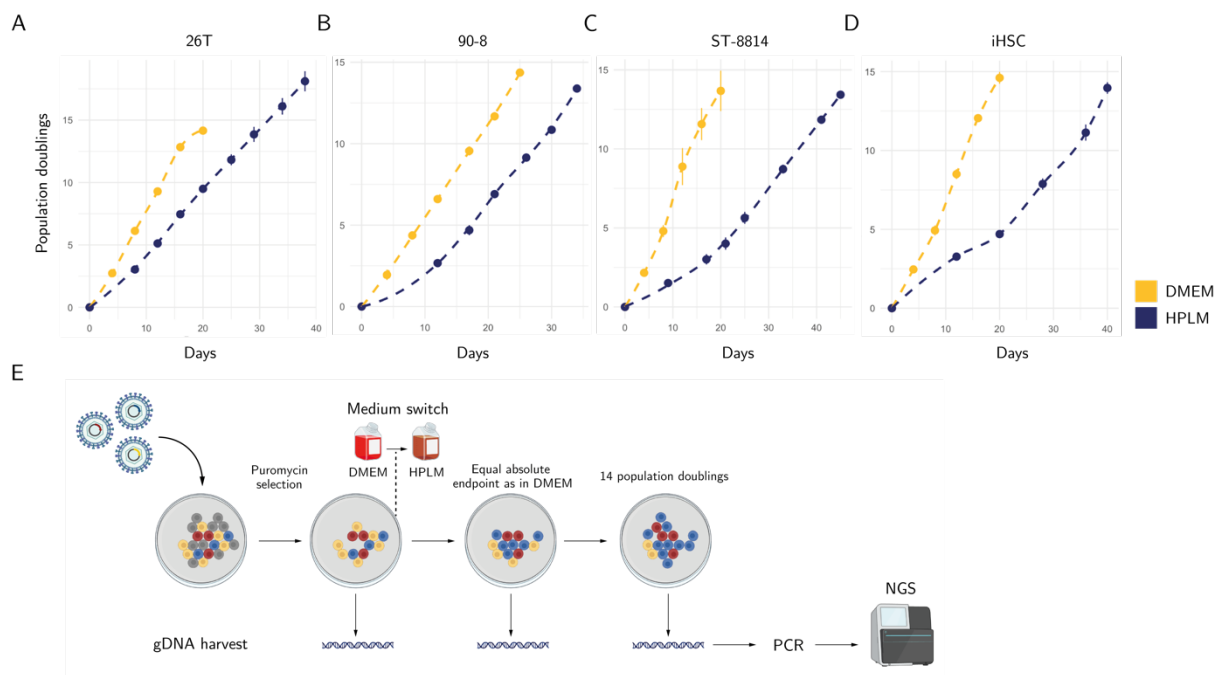


Figure 4.5. Cell line growth rate is medium context-dependent. (A-D) Growth curves of HPLM CRISPR screen target cell lines STS26T (A), 90-8 (B), ST-8814 (C) and iHSC (D) with population doublings (PDs) over time. Gold data points and dashed lines indicate PD measurements and Locally Estimated Scatterplot Smoothing (LOESS) of cells grown in DMEM, respectively. Blue data points and dashed lines indicate PD measurements and LOESS of cells grown in HPLM, respectively. Each dot represents two biological replicates with vertical lines representing standard deviation. **(E)** Schematic depiction of CRISPR screens performed in HPLM based on findings on differences in growth rate in HPLM compared to DMEM.

It was challenging to predict how the reduced growth rate in HPLM and potential differences in response to genetic perturbations would impact individual and CRISPR screen-wide effect sizes. The concern was that a lower effect size might result in the loss of signal from the hits I intended to identify. To evaluate whether the CRISPR screens conducted in HPLM captured the maximum potential true positive signal, as measured by the recall of common essential genes, one strategy I adopted was to utilise $\log_2(\text{fold change})$ values calculated for each gene across screen time points and ranked. The bottom ranked 15% of genes (relative comparison) were then assessed for their recall of common essential genes at each time point. The recall rates per screen were used as an assay to fit a quadratic polynomial regression model, which aimed to estimate the

maximum potential recall rate for each screen and determine whether the duration of the screen was sufficient to capture this optimal time point (Figure 4.6A-H). This analysis revealed that all HPLM screens extended beyond the predicted time point needed to achieve the maximum recall of essential genes. At least by these criteria, the adjustment appeared pragmatic and reasonable.

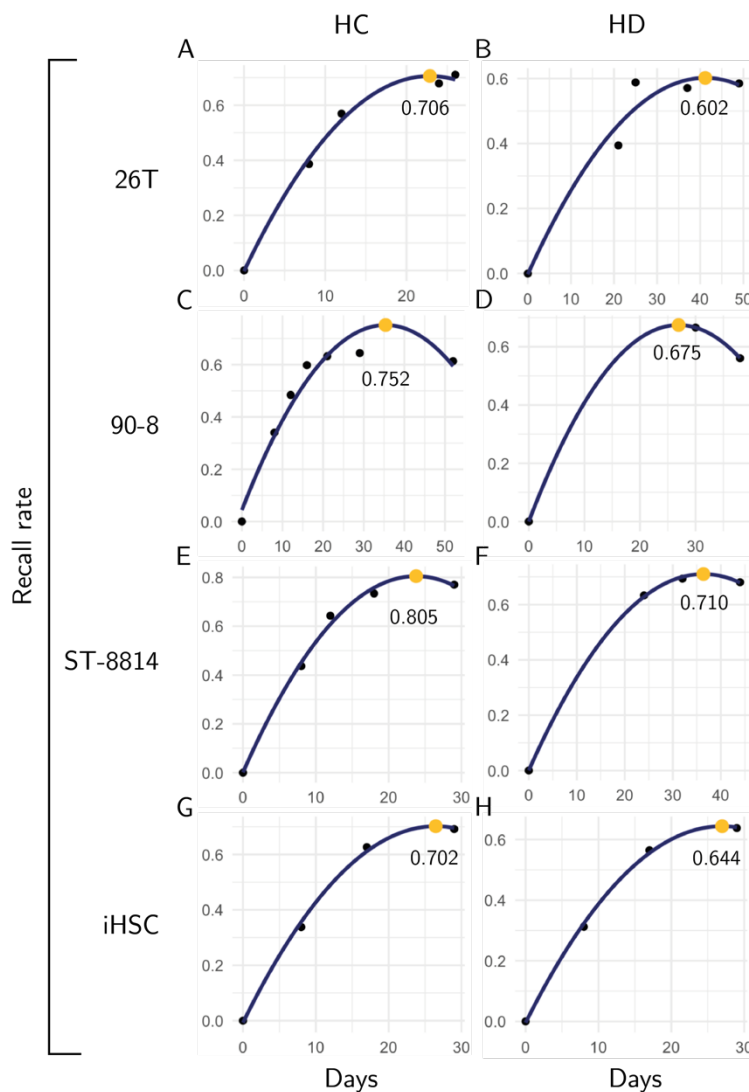


Figure 4.6. HPLM CRISPR screen durations capture maximum potential screen quality. (A) Dot plots of recall rates of common essential genes per STS26T Humagne C CRISPR screen time point based the 15% lowest $\log_2$ (fold change) values, representing change in gRNA representation, per time point. Recall rates are indicated by black dots. A quadratic polynomial regression model is represented by a blue line. The highest potential recall of common essential genes obtainable in this screen is indicated by a yellow dot and the corresponding recall rate provided below. (B) the same as in (A) but with Humagne D. (C, D) same as (A) and (B) but with 90-8. (E, F) same as (A) and (B) but with ST-8814. (G, H) same as (A) and (B) but with iHSC.

The CRISPR screen analysis for the DMEM screens used the pDNA representation in the CRISPR libraries as the early reference point. In contrast, for the HPLM screens, the earliest reference point available was four days post-transduction, immediately following puromycin selection. As previously demonstrated, the time elapsed since

CRISPR library transduction can influence the observed effect size in CRISPR screens, which, in turn, impacts their overall quality, as measured by the recall of common essential genes.

To ensure that the difference in reference time points does not compromise the quality of the CRISPR screen data, the recall rates of common essential genes from screens in this study, analysed using the four-day time point as an early reference, were compared to recall rates from the DepMap screens (Fig. 4.7). This comparison revealed no significant difference between the two groups, indicating that the applied workflow is capable of generating informative CRISPR screen data, regardless of the initial reference time point used.

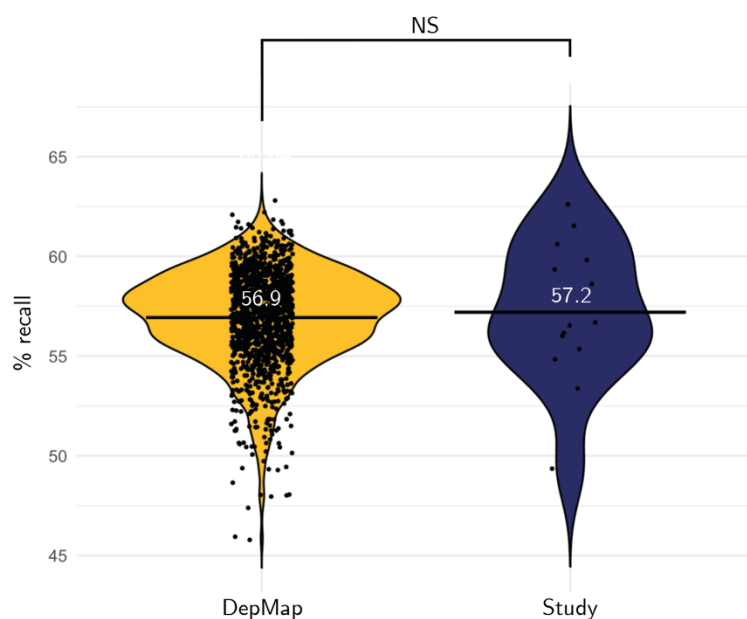


Figure 4.7. An initial reference time of four days retains HPLM and DMEM CRISPR screen data quality. Recall percentages per CRISPR screen experiment, indicated as black dots, as part of the DepMap dataset (golden violin) or this study (blue violin). For screens from this study, instead of the amplified CRISPR library pDNA representation, now the (late) four-day time point was used as the early reference time point during Chronos data analysis, as this was

when the medium was switch from DMEM to HPLM, for the HPLM screens. The DMEM screens were analysed in the same way with the four-day time point as the early reference time points. Chronos analysed Humagne C and D CRISPR screens as if coming from one CRISPR library screen experiment, to be able to include the Avana and KY datasets from DepMap. Means are indicated by horizontal black lines with values displayed above in white. A Wilcoxon rank-sum test was used to compare the distribution of % recall values between groups. Asterisks indicate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

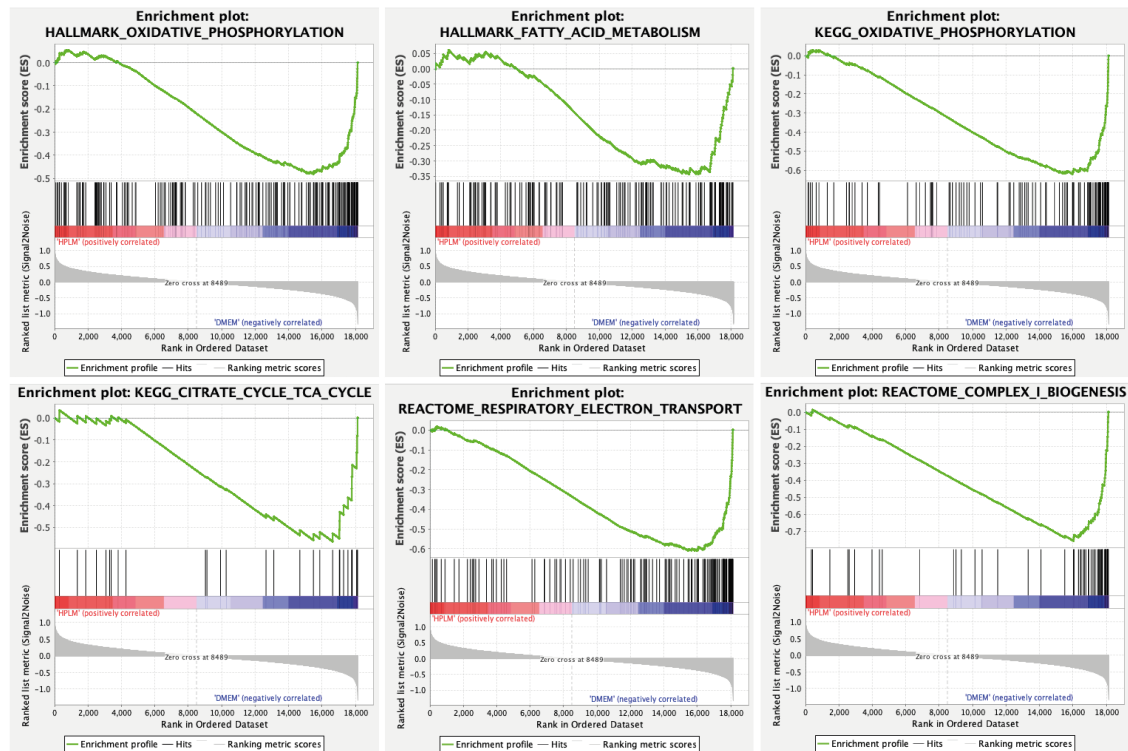
4.2.4 Medium-specific differential essentiality of gene sets linked to transcriptomic analysis

The previously conducted transcriptomic analysis identified significantly enriched gene sets in both DMEM and HPLM conditions. To determine whether CRISPR screens could uncover dependencies associated with these gene sets, GSEA was performed on the CRISPR screen data from cell lines cultured in DMEM and HPLM. For this analysis, data from Humagne C and D screens were combined to enhance statistical power. In this context, enrichment indicates relatively higher gene essentiality or a reduced growth effect, while depletion suggests the opposite.

The analysis revealed that specific pathways significantly enriched in DMEM, indicating higher essentiality in HPLM, were also overexpressed in HPLM, suggesting a correlation between higher expression levels and increased dependency on these gene sets (Fig. 4.8A and Table 4.2). For example, genes related to oxidative phosphorylation were more essential in HPLM, aligning with the idea that cells may depend more on mitochondrial energy production when glucose levels are lower[155]. Similarly, the increased essentiality of fatty acid metabolism in HPLM might be explained by its role in efficiently producing energy under these conditions[156]. These findings suggest a link between gene expression levels and cellular dependencies under different growth conditions, at least for these cellular processes. Notably, no gene sets were identified as more essential in DMEM that also exhibited elevated expression levels. Furthermore, there were no instances where higher gene set essentiality was associated with lower expression within the same medium. To pinpoint the specific genes within these gene sets driving the observed patterns, I overlaid the leading edge genes either significantly enriched or suppressed according to

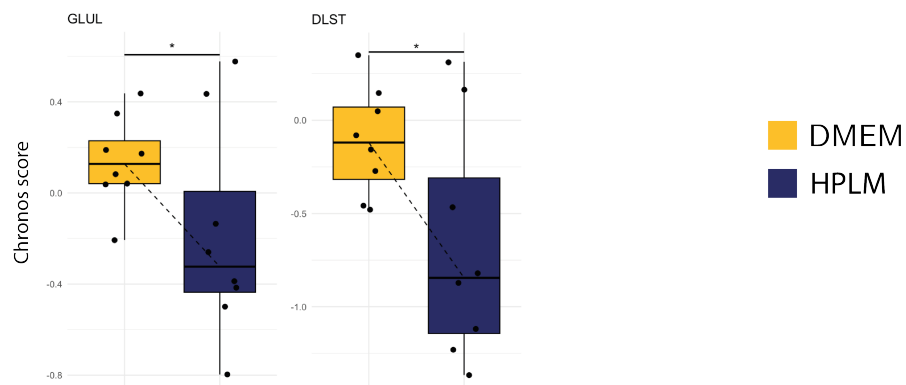
GSEA with those identified as differentially essential in the CRISPR screen experiments. This approach helps avoid scenarios where a pathway may be either differentially expressed or essential, but not all genes within its gene set list contribute to the observed effect. The overlaying of differential hits for the Hallmark fatty acid metabolism and KEGG oxidative phosphorylation gene sets are displayed in Figure 4.8B-D, as these gene sets contain the most distinct gene sets. It is important to note that while all genes display a significant difference in Chronos score, the genes may still be essential in both contexts. The difference may only indicate a higher dependency in one specific context compared to the other. Furthermore, some of these genes may be common essential genes and may suggest the difference in Chronos score is due to noise rather than true signal.

A



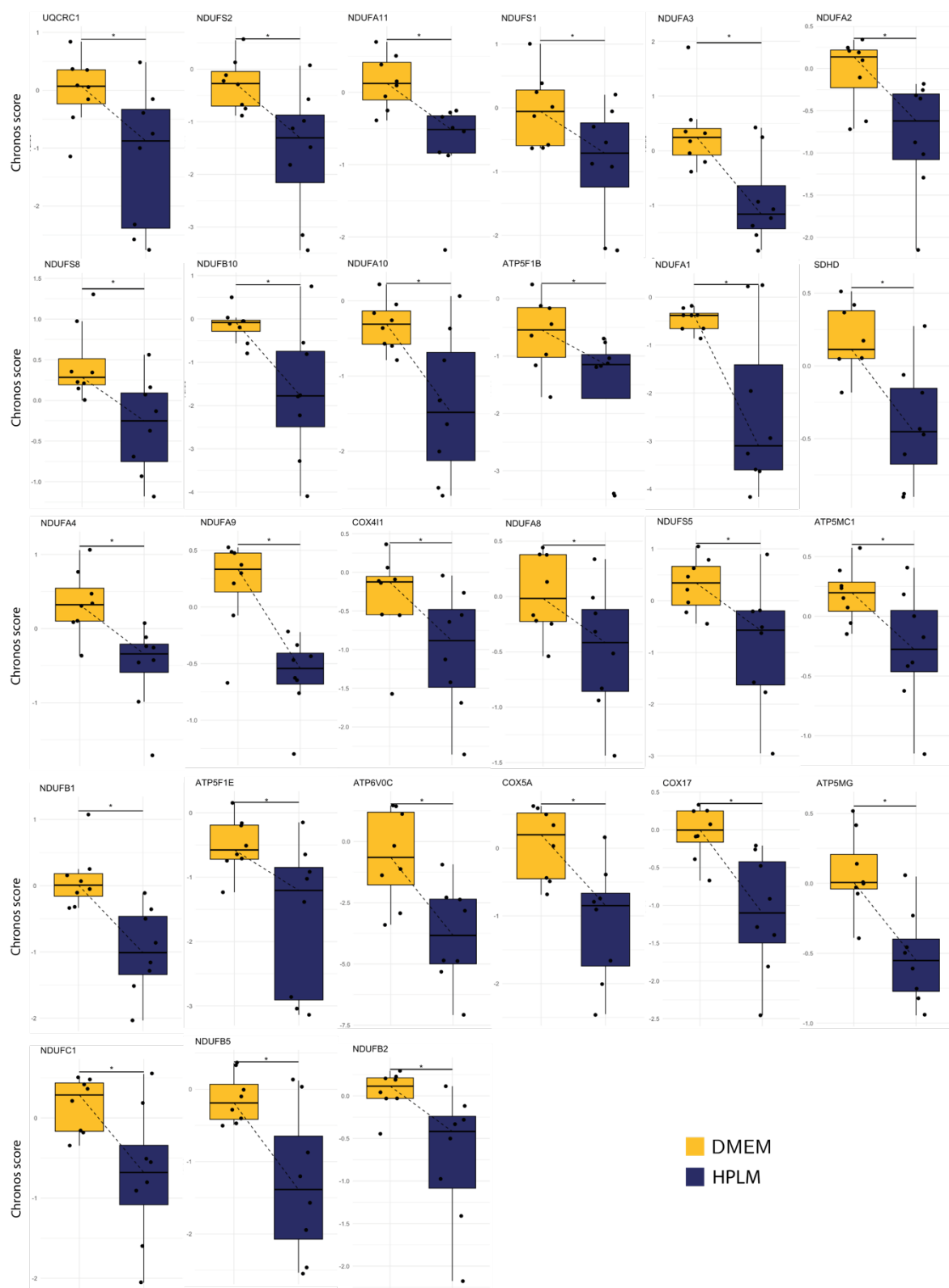
B

HALLMARK_FATTY_ACID_METABOLISM



C

Hallmarks_Oxidative_phosphorylation



D

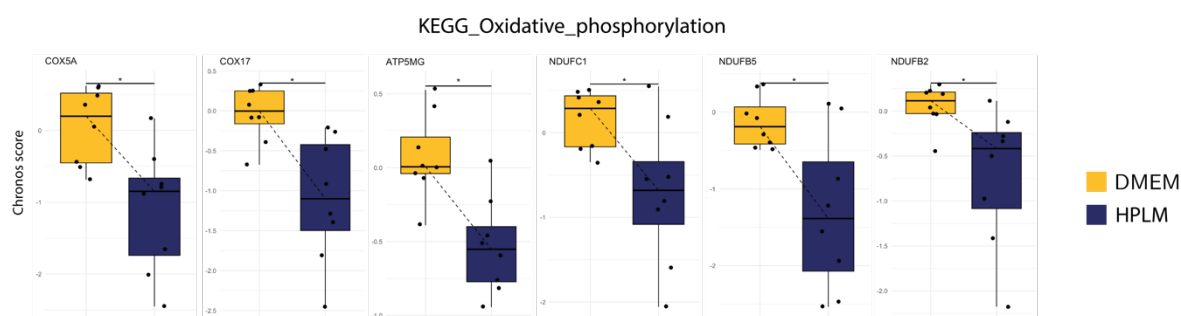


Figure 4.8. GSEA identifies differentially expressed and essential pathways and hits in medium context.(A) GSEA plots of significantly (FDR q value (<25%)) differentially expressed essential Hallmark and KeggLecacy gene sets, where a positive correlation reflects relatively higher Chronos scores for genes across HPLM versus DMEM CRISPR screens in this set. The complete list of significantly enriched sets in either medium is shown in Table 4.2. **(B-D)** The figure presents box plots of differential hits identified from the CRISPR screens conducted in HPLM (blue) versus DMEM (gold) media. These hits are part of gene sets that were found to be enriched in either HPLM or DMEM. Each box plot represents a differential hit, displaying eight Chronos scores per condition, derived from the STS26T, 90-8, s462TY, and iHSC cell lines used in the Humagne C and D screens. The black horizontal lines represent the means, while the vertical black lines indicate the data range. A Wilcoxon signed-rank test was performed to compare the distribution of Chronos scores between the groups. The signed-rank test was based on CRISPR screens performed in the same cell lines using the same library. Significance is denoted by asterisks: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4.2. Differential essential genes in all cell lines combined: HPLM vs. DMEM.

Gene set	Essential in	FDR q value (<25%)
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	DMEM	0.202
HALLMARK_OXIDATIVE_PHOSPHORYLATION	HPLM	0.002
HALLMARK_FATTY_ACID_METABOLISM	HPLM	0.095
KEGG_PROTEIN_EXPORT	DMEM	0.208
KEGG_PARKINSONS_DISEASE	HPLM	0.003
KEGG_OXIDATIVE_PHOSPHORYLATION	HPLM	0.001
KEGG_ALZHEIMERS_DISEASE	HPLM	0.004
KEGG_HUNTINGTONS_DISEASE	HPLM	0.004
KEGG_CITRATE_CYCLE_TCA_CYCLE	HPLM	0.059
REACTOME_TRNA_PROCESSING_IN_THE_NUCLEUS	DMEM	0.084
REACTOME_COMPLEX_I_BIOGENESIS	HPLM	0.013
REACTOME_AEROBIC_RESPIRATION_AND_RESPIRATORY_ELECTRON_TRANSPORT	HPLM	0.009

REACTOME_RESPIRATORY_ELECTRON_TRANSPORT	HPLM	0.007
REACTOME_COMPLEX_IV_ASSEMBLY	HPLM	0.195
REACTOME_ACTIVATION_OF_ANTERIOR_HOX_GENE_IN_HINDBRAIN_DEVELOPMENT_DURING_EARLY_EMBRYOGENESIS	HPLM	0.192
REACTOME_METABOLISM_OF_PORPHYRINS	HPLM	0.179
REACTOME_FORMATION_OF_ATP_BY_CHEMIOSMOTIC_COUPLING	HPLM	0.175
REACTOME_PHASE_2_PLATEAU_PHASE	HPLM	0.19
REACTOME_GLYOXYLATE_METABOLISM_AND_GLYCINE_DEGRADATION	HPLM	0.194
REACTOME_INTERLEUKIN_7_SIGNALING	HPLM	0.177
REACTOME_OXIDATIVE_STRESS_INDUCED_SENESCENCE	HPLM	0.226
REACTOME_ERBB2_REGULATES_CELL_MOTILITY	HPLM	0.233
REACTOME_RECOGNITION_AND_ASSOCIATION_OF_DNA_GLYCOSYLASE_WITH_SITE_CONTAINING_A_N_AFFECTED_PURINE	HPLM	0.222
REACTOME_MEIOTIC_RECOMBINATION	HPLM	0.213
REACTOME_BASE_EXCISION_REPAIR_AP_SITE_FORMATION	HPLM	0.202
REACTOME_INHIBITION_OF_DNA_RECOMBINATION_AT_TELOMERE	HPLM	0.193
REACTOME_CHROMATIN_MODIFICATIONS_DURING_THE_MATERNAL_TO_ZYGOTIC_TRANSITION_MZT	HPLM	0.188
REACTOME_MEIOSIS	HPLM	0.178
REACTOME_FORMATION_OF_THE_BETA_CATENIN_TCF_TRANSACTIVATING_COMPLEX	HPLM	0.18
REACTOME_ASSEMBLY_OF_THE_ORC_COMPLEX_AT_THE_ORIGIN_OF_REPLICATION	HPLM	0.173
REACTOME_PRC2_METHYLATES_HISTONES_AND_DNA	HPLM	0.173
REACTOME_RUNX1_REGULATES_GENES_INVOLVED_IN_MEGAKARYOCYTE_DIFFERENTIATION_AND_PLATELET_FUNCTION	HPLM	0.167
REACTOME_DNA_METHYLATION	HPLM	0.163
REACTOME_CONDENSATION_OF_PROPHASE_CHROMOSOMES	HPLM	0.162
REACTOME_REGULATION_OF_TLR_BY_ENDOGENOUS_LIGAND	HPLM	0.171
REACTOME_BUDDING_AND_MATURATION_OF_HIV_VIRION	HPLM	0.176
REACTOME_MEIOTIC_SYNAPSIS	HPLM	0.17
REACTOME_ERCC6_CSB_AND_EHMT2_G9A_POSITIVELY_REGULATE_RRNA_EXPRESSION	HPLM	0.176

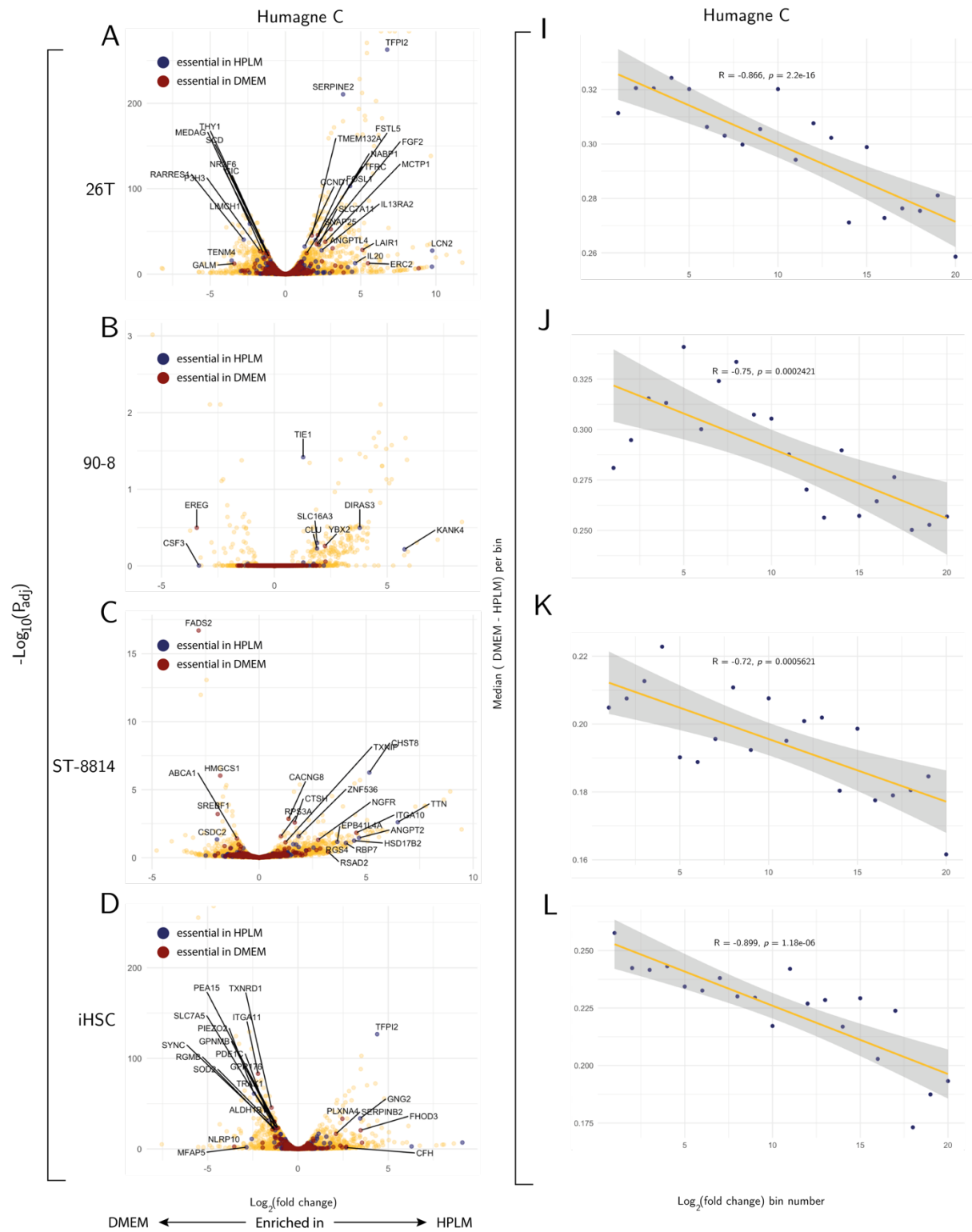
REACTOME_REGULATION_OF_ENDOGENOUS_RETR OELEMENTS_BY_PIWI_INTERACTING_RNAS_PIRNA S	HPLM	0.175
REACTOME_BASE_EXCISION_REPAIR	HPLM	0.19
REACTOME_SENESCENCE_ASSOCIATED_SECRETOR Y_PHENOTYPE_SASP	HPLM	0.191
REACTOME_RMTS_METHYLATE_HISTONE_ARGININ ES	HPLM	0.187
REACTOME_REGULATION_OF_ENDOGENOUS_RETR OELEMENTS_BY_THE_HUMAN_SILENCING_HUB_H USH_COMPLEX	HPLM	0.197
REACTOME_TRANSCRIPTIONAL_REGULATION_OF_ GRANULOPOIESIS	HPLM	0.21
REACTOME_FORMATION_OF_THE_NEPHRIC_DUCT	HPLM	0.227
REACTOME_DEFENSINS	HPLM	0.227
REACTOME_REGULATION_OF_INNATE_IMMUNE_R ESPONSES_TO_CYTOSOLIC_DNA	HPLM	0.229
REACTOME_SURFACTANT_METABOLISM	HPLM	0.224
REACTOME_REGULATION_OF_ENDOGENOUS_RETR OELEMENTS_BY_KRAB_ZFP_PROTEINS	HPLM	0.223
REACTOME_REGULATION_OF_PYRUVATE_DEHYDR OGENASE_PDH_COMPLEX	HPLM	0.23
REACTOME_MATERNAL_TO_ZYGOTIC_ TRANSITION_MZT	HPLM	0.235
REACTOME_DISEASES_OF_PROGRAMMED_ CELL_DEATH	HPLM	0.231
REACTOME_CITRIC_ACID_CYCLE_TCA_CYCLE	HPLM	0.245
REACTOME_STIMULI_SENSING_CHANNELS	HPLM	0.248

4.2.5 Inverse associations between medium-specific gene essentiality and differential gene expression

Some gene sets were identified as both differentially expressed and essential, meaning these genes are critical for cellular viability and exhibit changes in expression between different conditions (HPLM vs. DMEM). A portion of these genes showed a direct association between their essentiality and their expression levels, meaning changes in gene essentiality corresponded with significant changes in their expression. However, for the majority of genes, this association was not observed, indicating a more complex relationship between gene essentiality and gene expression profiles.

To further explore this relationship, the top 5% of differentially essential genes (those more or less essential in either HPLM or DMEM) were visualised alongside their corresponding differential gene expression data (Fig. 4.9A-H). While a small subset of these essential genes displayed significant differences in expression between the two conditions, the majority of genes clustered near the baseline expression level, suggesting minimal correlation between their essentiality and expression. This lack of correlation implies that gene essentiality does not always directly correspond to changes in transcriptional output, pointing to additional regulatory mechanisms or post-transcriptional factors influencing gene essentiality under different growth conditions.

Further analysis categorised genes based on absolute differential expression, and this uncovered a significant and strong inverse correlation between medium-dependent differences in essential genes and those differentially expressed genes with Spearman's Rho values of -0.87, -0.75, -0.72, and -0.90 for the STS26T, 90-8, ST-8814, and iHSC cell lines, respectively (Fig 4.9I-P). These findings challenge the expectation by some that higher differential expression would generally be associated with increased gene essentiality in specific contexts[157]. The results align with the idea that essential genes are under tight regulatory control and exhibit stable expression, for example as a likely protective mechanism to minimise the risks posed by their dysregulation[158]. Further support for this relationship is evidenced by another study which utilised HPLM as the culture medium during CRISPR screens and similarly found no significant correlation between differential gene expression and gene dependency[148]. Together, these findings suggest that gene essentiality is mostly determined by factors beyond simply the level of gene expression.



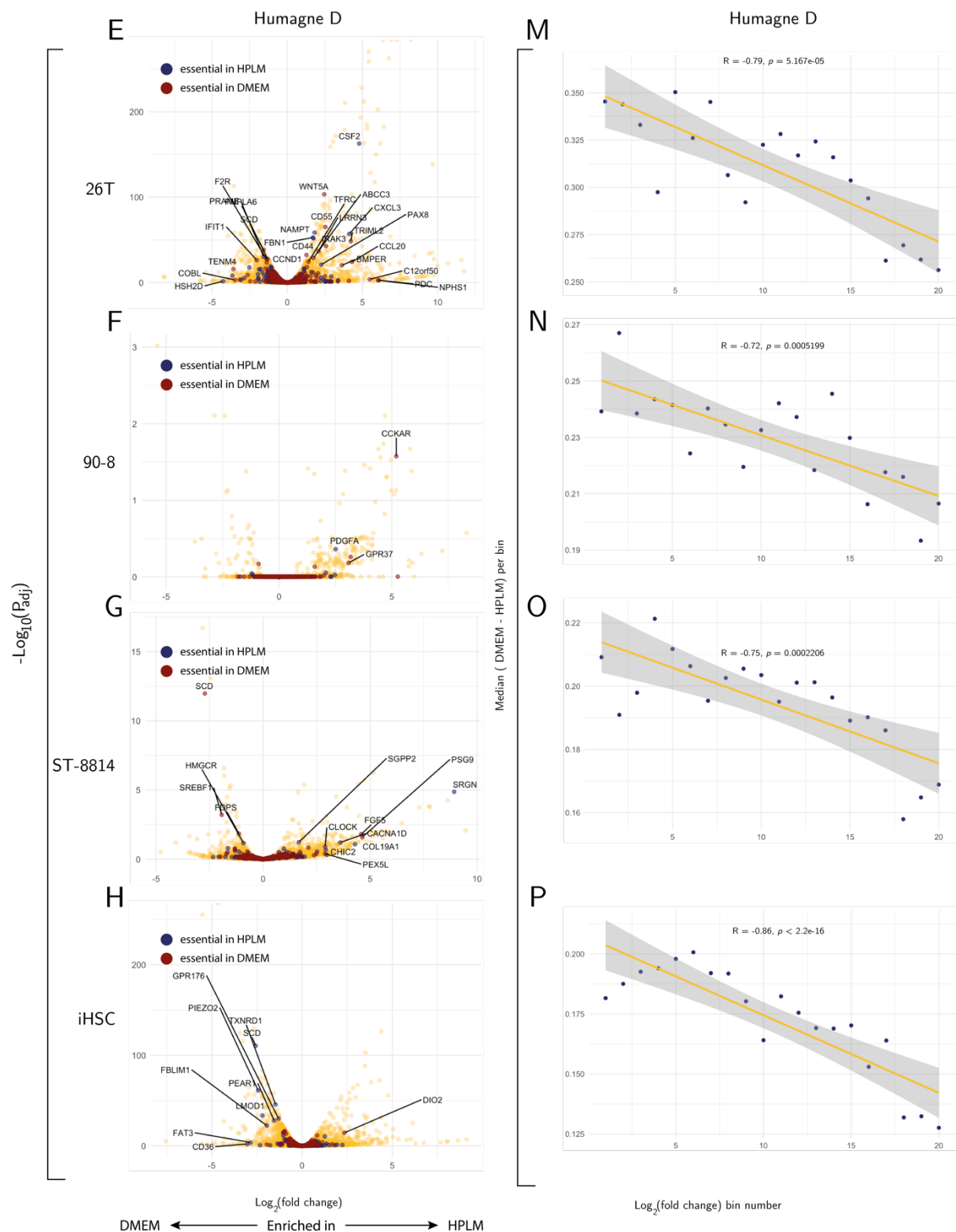


Figure 4.9. Culture medium-dependent differential hits do not correlate with medium context expressed genes. (A-H) Volcano plots displaying differential gene expression of STS26T, 90-8, ST-8814, and iHSC cells cultured in HPLM versus DMSO, with top and bottom 5% differential Humagne C (A-D) and D (E-H) library screen hit genes (DMEM – HPLM) highlighted in blue (essential in HPLM) and red (essential in DMEM). The maximum overlap limits the labelled differential hits displayed. (I-P) Dot plot and linear

regression model depicting the relationship between $\log_2(\text{fold change})$ gene expression bin number under medium-specific conditions and the differential hits identified from Humagne C (I-L) and D (M-P) library screens across the STS26T, 90-8, ST-8814, and iHSC cell lines. Chronos scores were calculated by subtracting HPLM values from DMEM for each gene, followed by grouping these scores into 20 equally sized $\log_2(\text{fold change})$ bins (blue dots). The median Chronos score was then determined for each bin. Bins were created by taking the absolute $\log_2(\text{fold change})$ values. A Spearman's rank correlation test was conducted to evaluate the significance of the observed relationship, with a linear regression line (gold) and a 95% confidence interval indicating the trend across median values per bin (blue). Spearman's Rho and p-values are reported on each plot.

4.2.6 Hit prioritisation of medium-dependent essential genes

To systematically identify genes that exhibit medium-dependent effects upon knockout across multiple cell lines, including those shared between MPNST cell lines 90-8 and ST-8814 but independent of medium-dependent transcriptomic variation, I ranked the differences in Chronos scores between HPLM and DMEM (DMEM – HPLM) for individual cell lines. These ranked differences were then overlapped based on the top (indicating lower Chronos scores in HPLM screens) or bottom (indicating lower Chronos scores in DMEM screens) genes, and counted (Fig. 4.10A, B, E, F). Using an arbitrary threshold of the top 500 differential gene effects, I identified an overlap of 2 and 15 genes with consistently lower scores in HPLM compared to DMEM across all cell lines for Humagne C or D libraries, respectively (Fig. 4.10C, G). For DMEM, 7 and 13 genes were found to have lower scores for Humagne C and D libraries, respectively (Fig. 4.10D, H). When focusing on the overlapping genes between the MPNST cell lines 90-8 and ST-8814, additional differential gene effects were observed. No common hits were found between the Humagne C and D library screens. Genes identified as essential across all cell lines per library are listed in Table 4.3, while those shared between 90-8 and ST-8814 are detailed in Table 4.4.

PRELID3B was one of the two genes consistently found to have lower gene effect scores in HPLM in Humagne C screens across all four cell lines. Notably, *PRELID3B* encodes *SLMO2*, a protein involved in the transport of phosphatidylserine, playing a crucial role in lipid metabolism[159]. This observation aligns with the results from the GSEA, suggesting that HPLM induces a cellular shift towards reliance on alternative energy sources. This shift may be driven by the more physiologically relevant glucose levels in HPLM compared to the high glucose content in DMEM, which could influence cellular metabolic dependencies.

COP9 Signalosome Subunit 6 (*COPS6*) was another gene identified with consistently lower Chronos scores across all four tested cell lines in the Humagne C screens. Notably, *COPS6* has been implicated in oxidative phosphorylation in previous studies, further supporting the notion that cells in HPLM may exhibit increased dependency on this metabolic pathway [160]. Interestingly, *COPS6* is not included in any Hallmarks, KEGG Legacy, or Reactome gene sets and highlights the important consideration that while specific studies have established connections between certain genes and cellular processes, the gene sets used in these analyses are not exhaustive. The absence of a gene from a particular set does not preclude its involvement in the associated biological pathway of that gene set, and in-/exclusivity of these gene sets may be highly context-dependent.

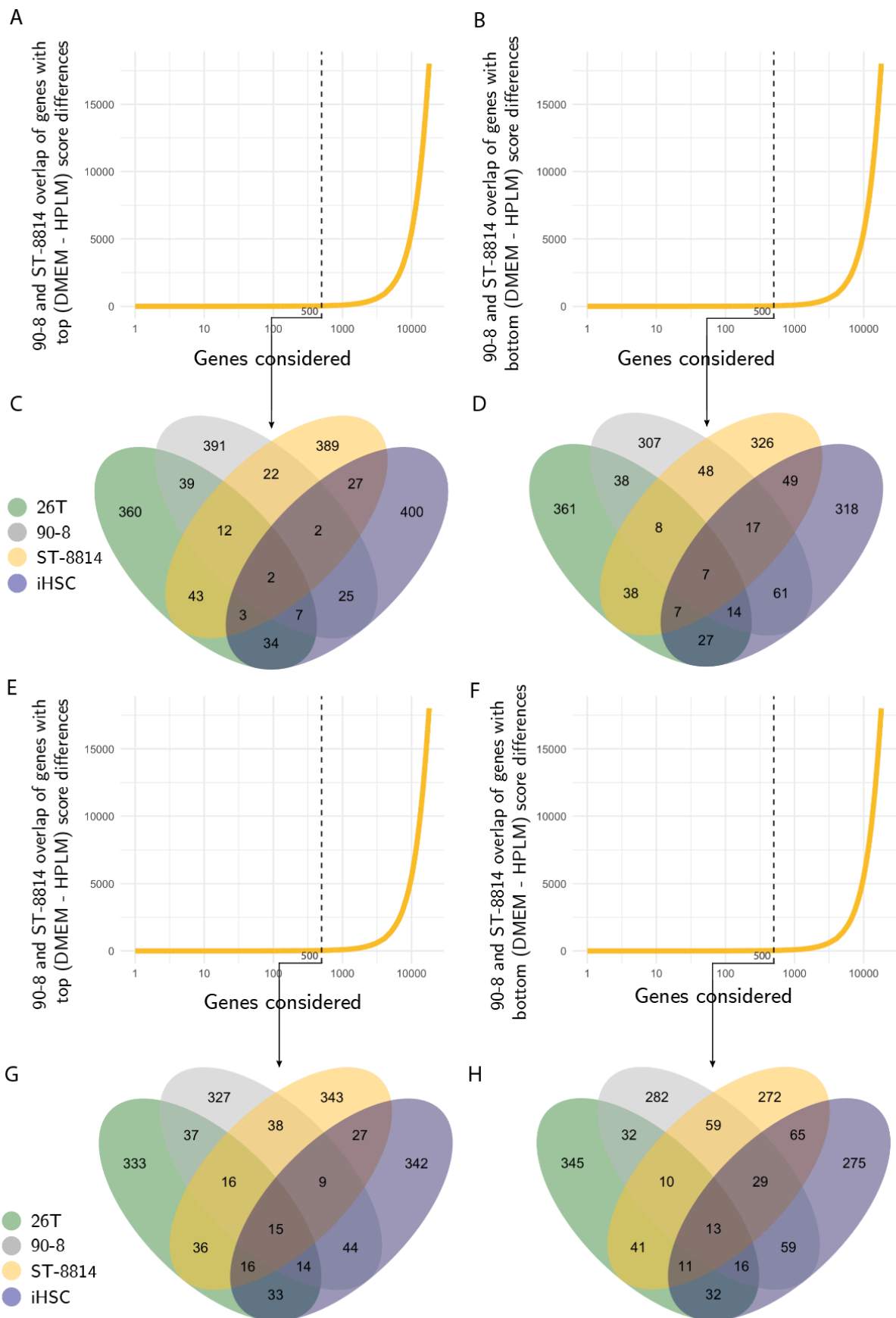


Figure 4.10. Medium context-dependent gene effect intersection between cell line models. (A-B) Line plots showing the overlap counts at an increasing threshold of differential gene effects considered between cell lines. Differential gene effects were determined by subtracting HPLM Chronos scores from DMEM (DMEM – HPLM) per screen and ranking these values, where the top value genes represent genes with higher Chronos scores in HPLM and bottom value genes representing genes with higher Chronos scores in DMEM. **(C-D)** Venn diagrams displaying the intersections between the top 500 differential essential genes in HPLM (C) or DMEM (D) per cell line, identified during Humagne C CRISPR screens. **(E-F)** Same as (A-F) but then for Humagne D.

Table 4.3. Medium-context dependent differential gene 'hits' common to STS26T, 90-8, ST-8814 and iHSC cell lines.

Lower HPLM scores in HC	Lower DMEM scores in HC	Lower HPLM scores in HD	Lower DMEM scores in HD
PRELID3B	SF3B1	SRSF2	SRPRB
ATP6V0C	TAF1D	CDK9	UTP18
	DDX47	SAP30BP	POLE
	SNRPD2	RPS5	MTREX
	CENPN	FCF1	PTDSS1
	PPP1R12A	COPS6	U2SURP
	TXNRD1	GINS1	ARPC2
		H2BC11	CEBPZ
		TMED4	NCAPD2
		CFAP298	MED7
		PAM16	TRAPPC11
		SPC25	PAFAH1B1
		ZNHIT2	PSMB3
		POLR2E	
		IFITM3	

HC: Humagne C library, HD: Humagne D library

Table 4.4. Medium-context dependent differential essential genes common to 90-8 and ST-8814 cell lines.

HPLM essentials HC	DMEM essentials HC	HPLM essentials HD	DMEM essentials HD
CNOT3	EFTUD2	SRSF2	SNRPF
PRELID3B	RPL36	CDK9	CDCA8
DUT	SRSF1	SPC25	BCAS2
RAD21	ISCU	H2BC11	PAFAH1B1
HELB	PSMD11	SARS1	DDX10
KARS1	COPB2	CDC123	ACTL6A
ATP6V0C	SNRPD2	RPS5	PSMB3

STIL	PSMD6	RPAP1	SRSF3
SAFB2	WDR3	DYNLRB1	SCAP
SART1	HNRNPH1	SAP30BP	RBM39
RBM14	POLR1B	COPS6	COX6C
NDUFS2	RPL5	RPS9	SREBF1
RPL30	SRSF3	GIN51	RPL13
BCS1L	COPA	POLR2E	PCYT1A
CAPZB	SUDS3	NDUFA1	PSMA5
ACTR10	CKAP5	NELFB	MAX
CDC20	SF3B1	IFITM3	MAK16
RPL3	INTS4	LYRM4	POLE
LRPPRC	PNN	SMC4	TIPIN
MAF	FBXO5	NDUFB5	COPS5
URB1	CDK7	HAUS5	CEBPZ
OR2M4	TAF1D	SEC13	OR6C6
NDUFB9	RAN	NDUFB4	AQR
SV2C	CUL1	PCBP2	TERF2
PDRG1	SRP19	ALDH3B1	FBXO5
MCM3AP	CENPN	TSEN2	RAN
POLD3	PSMA1	UTP4	SMC5
RBM25	PPP1R12A	PSMC3	NUP133
NSF	ASCL4	ZNF841	ZNF681
CDK12	NUP205	LIG1	POLR3D
GNGT2	EIF2S1	CFAP298	RPL31
NDUFA4	ITGA10	SNRPB	MCM5
SASS6	KNL1	DPPA5	EIF4A3
TRMT1	SCAP	NUBPL	PSMD7
RPL32	NIPBL	PAM16	TARS1
GTPBP4	DDX27	NDUFA5	RTF1
GIN51	HCFC1R1	GIN52	NOC3L
LMOD3	RPS19	SRP54	SRPRB
	ACSL4	NDUFA3	UXT
	PSMD1	CDC45	NACA
	CEBPZ	USP24	NCAPD2
	SREBF1	NOB1	COPZ1
	PPP1CB	HCCS	NFYB
	BUB3	COX5A	EIF2B3
	CENPH	ATP5ME	RUVBL1
	PRPF38A	DBF4B	POLD3
	SIN3A	DDX42	FNBP1L
	IL2RB	UQCR10	TSEN34
	BTBD2	PDE12	RPL23

	SMC2	FCF1	CCT6A
	GOLT1B	CEP43	ACTR3
	VCP	RPS15A	MYRFL
	NSMCE3	ERCC2	THAP7
	BORA	SLC35B2	PIKFYVE
	PGGT1B	RPL6	TADA2B
	RPL7A	EFCAB2	TRNT1
	XRN2	NMD3	RPS8
	RRP12	RBM17	ATP6V1B2
	EXOSC8	UTP11	RPL7A
	MRPL49	MRPS36	MBTPS1
	POMP	TG	SCD
	CHEK1	TFB1M	TRAPPC11
	TXNRD1	EIF1AX	TOP3A
	WDR12	CREBRF	CAPN11
	NOP58	NRDE2	CENPH
	CENPI	TMED4	CNOT1
	TUT1	RPL32	TAF1A
	SMAD4	HCFC1	YME1L1
	C2orf68	SLC25A26	CKAP5
	ABCE1	ATP6V1H	SUPT20H
	CNNM3	DUT	SS18L2
	RPP14	H3C6	CCNB1
	RANGAP1	CES5A	MTREX
	PGGHG	NSA2	CRCP
	HYPK	ZNHIT2	EIF2S2
	ERCC6L	INTS4	TOR1A
	PRMT1	RNASEH2A	SHKBP1
	TPRKB	MED1	TAF1B
	DDX47		EIF3D
	DRD5		ZEB1
			TARS2
			PPP2CA
			FAM98B
			DNAJC14
			NHLRC2
			SPDL1
			SNRNP200
			EVX2
			PTDSS1
			GCN1
			RAD17

			NOP2
			TIMM44
			MYO3B
			EIF3I
			MED6
			MED7
			GTF2E1
			UTP18
			RSPH1
			RPL21
			NAA50
			U2SURP
			CTR9
			STRC
			PRPF8
			TSPEAR
			DNASE1L3
			ARPC2
			HNRNPH1
			PXYLP1

HC: Humagne C, HD: Humagne D

4.2.7 Hit prioritisation of MPNST-specific medium-independent

essential genes

While medium-context dependent hits are relevant because they may present opportunities for targeted interventions within specific environments and subcellular contexts, another goal of this study was to identify targets that are independent of context. Such context-independent hits would be potentially more universally applicable and more reliable as therapeutic targets, ensuring efficacy across different physiological conditions.

Previously, to accurately determine differences between HPLM and DMEM hits, the initial reference time point was set at four days to capture gene effects specifically related to either HPLM or DMEM. However, it's important to note that the first four

days also contribute to the overall screen effect size, as demonstrated previously in Figure 3.4D (Chapter 3). In order to select dependencies in MPNST regardless of medium context, it was desirable to maximise the detection of significant signals. Thus, for this purpose, the early time point for all screens was reverted to the original pDNA representation in the respective CRISPR library, consistent with the approach used in the analyses described in Chapter 3.

A potential approach is to hypothesise that MPNSTs may harbour unique vulnerabilities distinct from those of other cancers and, perhaps, even from (immortalised) human Schwann cells. To investigate this hypothesis, I aimed to identify MPNST-dependent genes in cell lines while ensuring these genes are not essential in the STS26T and iHSC control cell lines. Instead of applying a binary threshold, such as classifying the bottom 15% of genes as essential and the top 85% as non-essential, it may be more effective to employ a tiered system that categorises genes into three distinct groups: essential genes, non-essential genes, and an intermediate group that lies between the bottom 15% and (for example) the top 66.7% of genes (based on an arbitrary threshold). After implementing this tiered approach, the essential genes identified in MPNSTs can be overlaid with the non-essential genes from the control group, ensuring a substantial difference in gene effect scores between the target and control cell lines. This method allows for the identification of specific vulnerabilities in MPNSTs that could be exploited for targeted treatments, thereby minimising the risk of off-target effects.

In this study, there are two control cell lines: iHSC and STS26T. The iHSC cell line, while an immortalised Schwann cell, may have acquired characteristics more similar to MPNSTs due to its immortalisation process, though it lacks the common driver

mutations typically associated with MPNSTs. The melanoma cell line STS26T, which was previously misdiagnosed as an MPNST, serves as another control that might be relatively closely related to MPNSTs. This is because melanocytes, like Schwann cells, are derived from neural crest cells, suggesting some level of functional overlap between these cell types[161]. Melanocytes share certain neural crest-derived traits with Schwann cells, which are involved in the support and insulation of neurons[161]. Furthermore, studies have demonstrated that Schwann cell progenitors can differentiate into melanocytes, further narrowing the functional and developmental link between MPNSTs and melanomas [161], [162], [163], [164], [165]. This relationship underscores the relevance of using both iHSC and STS26T as control cell lines, given their similarities and differences with MPNSTs. These controls may help to more precisely identify MPNST-specific vulnerabilities by accounting for potential overlaps in cellular function and lineage.

To identify intersections of essential genes in different MPNST cell lines and non-essential genes in control cell lines across medium contexts where possible (s462TY screens were not conducted in HPLM), the bottom 15% of genes were considered essential and the top 66.7% of genes were considered non-essential. The basis of these seemingly arbitrary cutoffs is the assumption that approximately 15% of all genes are essential in any given cell line with a margin of around 20% between the bottom 15% and the top 66.7% deemed sufficient to discriminate the two populations[85]. These MPNST essential and STS26T and/or iHSC non-essential genes were then overlaid across different cell lines screened with either the Humagne C or D libraries (Fig. 4.11A-H).

No overlapping genes were found when overlaying essential genes from the MPNST cell lines 90-8, ST-8814, and s462TY with non-essential genes identified in iHSC and STS26T, using data from both Humagne C and Humagne D library screens, (data not shown). A possible explanation for this could be the significantly lower quality of the s462TY screens as measured by the general effect size and recall of common essential genes, which may not have fully captured the complete set of essential genes, thereby limiting its potential to reveal overlapping hits.

However, when the s462TY cell line data was left aside, three genes were identified as essential in both 90-8 and ST-8814 cell lines and non-essential in iHSC and STS26T cell lines in the Humagne C screens (Fig 4.11A, Table 4.5, Table 4.6). In the Humagne D screens, four additional genes were found (Fig 4.11B, Table 4.5, Table 4.6). These are proposed targets hits that are specifically essential for MPNSTs but not for the control cell lines.

One of the three genes identified using the Humagne C library is Spectrin β II, non-erythrocytic 1 (*SPTBN1*). SPTBN1 is a crosslinking and scaffold protein that plays a crucial role in shaping the cytoskeleton, facilitating organelle formation, and organising transmembrane proteins. Notably, neurofibromin 2 (NF2) interacts with SPTBN1 and depends on SPTBN1 for its binding to the cytoskeleton. In this way NF2 and SPTBN1 bridge signals transduced from the extracellular matrix via transmembrane proteins and the cytoskeleton to mediate contact inhibition of proliferation[166]. Loss of NF2 function can also lead to unchecked Schwann cell proliferation, potentially culminating in cancer formation[167], [168]. Additionally, disruption of spectrin proteins, including those encoded by *SPTBN1*, has been linked to demyelination *in vivo*[166]. However,

interestingly, substantial knockdown of *SPTBN1* in Schwann cells does not appear to result in widespread cellular dysfunction, suggesting that *SPTBN1* may not be essential in non-cancerous Schwann cells[169]. This is further supported by the fact that *SPTBN1* has not been identified in common essential gene lists[98], [81], and disruption of *SPTBN1* in other cancer cell lines rarely exhibits a significant negative effect, as evidenced by DepMap gene dependency data[170]. The relative rarity of *SPTBN1* being essential, combined with its potential role in Schwann cell morphology and function, makes it an interesting candidate gene for further exploration and validation. In contrast, the other two genes identified, *FBXO42* and *CDK8*, are more established oncogenes, with well-documented roles in cancer development and progression [171], [172], [173], [174].

When considering non-essential genes identified in iHSC but not necessarily in STS26T, 12 and 14 genes were found using Humagne C and Humagne D, respectively. Conversely, when focusing solely on non-essential genes in STS26T, 8 and 14 genes were identified as specifically essential in 90-8 and ST-8814 with Humagne C and D, respectively.

Although the s462TY screens were of lower quality, they should not be disregarded, as they could still reveal critical dependencies in MPNSTs. Additionally, it is important to recognise that dependencies are often shared between various cancers. Thus, while identifying a highly specific target for MPNST could uncover a key mechanism specific to this cancer type, which might be valuable for therapeutic exploration, a drug that can target multiple cancer types is always desirable. With this in mind, essential genes identified across all MPNST cell lines using either Humagne C or D were overlaid, disregarding the STS26T and iHSC control cell lines, and revealed 992

and 976 common potentially essential genes. Although many of these genes hold potential interest, the list remains too extensive for practical selection of candidates for further validation. Further hit prioritization will be discussed next to narrow down this list to the most promising targets for subsequent experimental validation.

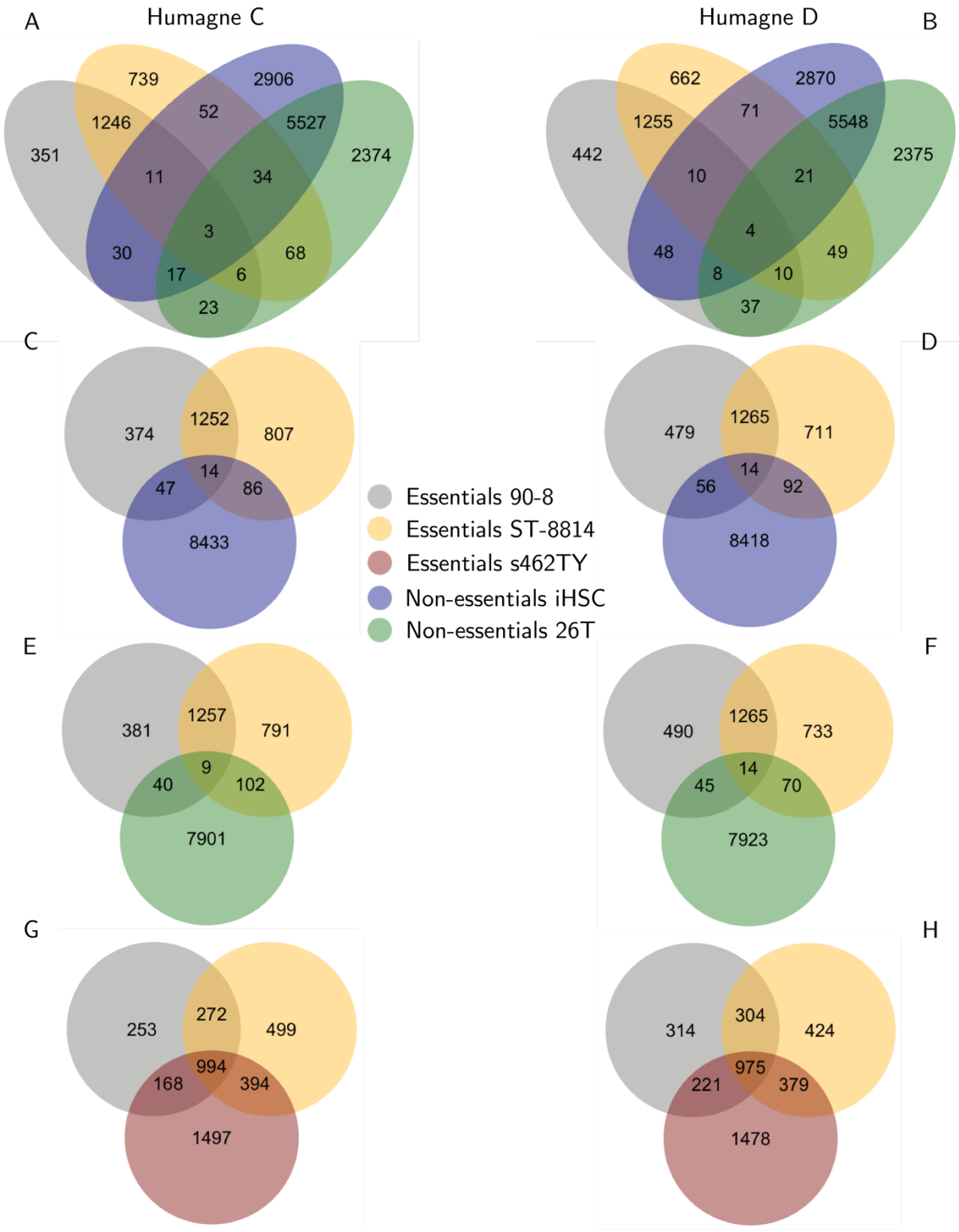


Figure 4.11. Cell type and cell line dependent essential genes found in MPNST and non-MPNST cell line models. (A-H) Venn diagrams displaying different intersections of essential genes and non-essential genes found in unbiased genome-wide CRISPR library screens in both DMEM and HPLM medium with Humagne C or D. The bottom 15% (N = 3053) genes based on the Chronos z-score per pDNA bin were considered (cell line-specific) essential genes and the bottom 66.7% genes (arbitrary) were considered non-essential. Essential genes for 90-8, ST-8814 and non-essentials for STS26T and iHSC had to be observed in both DMEM and HPLM screens, to be considered medium independent observations. Note: essential gene lists for s462TY are only based on CRISPR screens in DMEM.

Table 4.5. Intersection of essential and non-essential genes from CRISPR screens in this study as displayed in Figure 4.11A-H.

Humagne	C	D	C	D	C	D
Essential in 90-8	●	●	●	●	●	●
Essential in ST-8814	●	●	●	●	●	●
Essential in s462TY	○	○	○	○	○	○
Non-essential in iHSC	●	●	●	●	○	○
Non-essential in 26T	●	●	○	○	●	●

SPTBN1	UBE2D3	SPECC1L	SPECC1L	TRAPPC12	SAFB
FBXO42	NUSAP1	ANKRD52	HERC2	ALG10B	EXOC4
CDK8	HDLBP	ZNF142	UBE2D3	SPTBN1	C22orf15
	CCDC6	SNX1	NUSAP1	IL36G	UBE2D3
		CLPX	CORO1C	PTPN11	PTPN11
		SPTBN1	DGUOK	FBXO42	NUSAP1
		FBXO42	ZDHHC8	NIPAL2	DYNC1LI1
		LEO1	PLAA	SERAC1	ARHGAP5
		ERMARD	HDLBP	CDK8	HTATSF1
		C2orf76	GLT8D1		HDLBP
		MATR3	CCDC6		CCDC6
		ANXA4	ITGB4		NAPRT
		SMARCAL1	POU5F1		INCENP
		CDK8	TTC9B		LPAR1

Note: intersections between 90-8, ST-8814 and s462TY were not included as these numbers were too large.

Table 4.6. 90-8 and ST-8814 Humagne C or D hits, non-essential in STS26T and iHSC

Gene	Function	Drug
SPTBN1	Cytoskeletal structure and cell signalling	Not directly targeted by drugs
FBXO42	Involved in protein ubiquitination and degradation	Not directly targeted by drugs
CDK8	Promotes G1/S transition through gene expression regulation	Senexin A, B, and others
UBE2D3	Ubiquitin-conjugating enzyme, protein degradation	Not directly targeted by drugs

NUSAP1	Microtubule binding, regulates mitosis	Not directly targeted by drugs
HDLBP	RNA-binding, lipid transport	Not directly targeted by drugs
CCDC6	Tumour suppression, DNA damage response	Not directly targeted by drugs

4.2.8 Prediction of synthetic lethality associates with lower gene effect scores

Validation experiments following the identification of CRISPR screen hits are comprehensive and time-consuming endeavours. Before selecting specific hits for validation, it's crucial to thoroughly assess all available data to make informed decisions. This approach increases the likelihood of validating true hits rather than inadvertently pursuing noise.

Beyond its intrinsic value in understanding cancer biology, identifying synthetic lethality pairs in MPNSTs serves as a strategic method for prioritising hits. The rationale behind this approach is strengthened by the growing repository of data on potential synthetic lethality pairs and the development of advanced prediction tools[102], [175], [176]. By leveraging synthetic lethality as a criterion, one can narrow down the list of potential targets, thereby focusing on those that are more likely to yield meaningful results in subsequent validation experiments. This strategy not only aligns with the overarching goal of identifying synthetic lethality pairs in MPNSTs but also provides a scientifically sound basis for prioritising hits, ultimately enhancing the efficiency and success of the validation process.

One of the datasets providing predictions for synthetic lethality is SynLethDB 2.0, an enhanced version of the original SynLethDB[103]. This database comprises 35,944 documented synthetic lethality interactions, which have been predicted based on various levels of evidence, including drug inhibition screens, RNAi, and CRISPR experiments.

SynLethDB 2.0 encompasses well-understood and validated mechanisms, such as the relationship between BRCA1/BRCA2 defects and sensitivity to PARP inhibition[103]. In addition to these well-characterised interactions, SynLethDB 2.0 also includes less well-understood, or even predicted relationships between genes, which are based on unbiased screening experiments [177]. This extensive dataset may allow prioritisation potential hits beyond the scope of already thoroughly investigated interactions, opening up new potential avenues for exploration in synthetic lethality.

To evaluate whether the predicted synthetic lethality targets for the cell lines studied also corresponded with lower Chronos scores, a comparative analysis was conducted. First, copy number variation (CNV) data was filtered to identify genes with a copy number of 0, which are indicative of potential gene loss in the target cell lines. Predicted synthetic lethality pairs for these genes were then extracted from SynLethDB 2.0. Subsequently, these predicted synthetic lethality pair genes were overlaid with their associated Chronos scores, derived from the CRISPR screen data. For comparison, the gene effects of genes without predicted synthetic lethality, were used as the control group.

In all the 90-8 and ST-8814 screens (DMEM and HPLM media, Humagne C and Humagne D libraries), Chronos scores for genes in the synthetic lethality predicted group were significantly lower compared to the control group (Fig. 4.12A-D, F-I). Although the s462TY screens showed a consistent, albeit marginal, reduction in Chronos scores for the synthetic lethality group, the differences were not statistically significant (Fig 4.12E, J). This may be related to the smaller overall effect size observed in s462TY screens. Together, these findings suggest a relationship between synthetic lethality predictions

from SynLethDB 2.0 and lower gene effect scores in CRISPR screens, at least in 90-8 and ST-8814, supporting the utility of these predictions for identifying key dependencies in cancer cell lines.

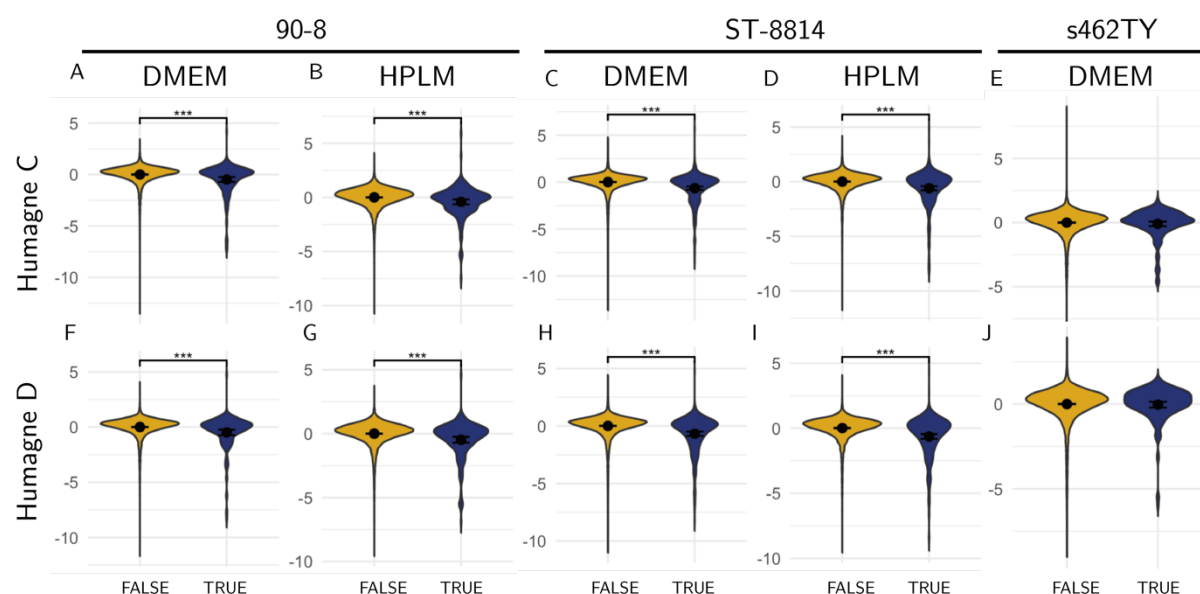


Figure 4.12. Predicted synthetic lethality partners show higher gene essentiality in MPNST cell line models. (A-J) Violin plots of Chronos scores for genes for which synthetic lethality was predicted using the SynLethDB 2.0 (TRUE) versus not (FALSE) in 90-8 (A,B,F,G), ST-8814 (C,D,H,I) and s462TY (E,J) MPNST cell line models in Humagne C (A-E) and D (F-J) CRISPR library screens. A Wilcoxon rank-sum test was used to compare the distribution of Chronos scores between groups where synthetic lethality was predicted versus not. Asterisks indicate a significant difference: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4.2.9 Integration with external MPNST CRISPR screens suggests synthetic lethality and druggable gene targets

The DepMap database provides extensive CRISPR screen and CNV data for numerous cell lines, including five MPNST cell lines. By integrating this data with the CRISPR screens from our study, the initial 994 and 975 hit genes identified using Humagne C and D can be refined further. Synthetic lethality predictions can then be applied to these refined gene lists.

To ensure the most stringent selection of hits, genes were required to be consistently identified as hits (bottom 15%) across all different screens, including DMEM, HPLM (if performed), Humagne C and D, as well as all MPNST DepMap screens. This selection process encompassed three different CRISPR libraries, eight different MPNST cell lines, and two types of medium. The next selection criterion was that the genes had to be predicted as synthetic lethality hits in all cell lines except MPNST724. MPNST724 was excluded from the synthetic lethality prediction requirement, because no synthetic lethality relationships were predicted, which was believed to be attributed to a potential lack of data on lost genes rather than an absence of genetic dependencies in this cell line. Using this approach, *MYC* and *NAE1* were identified as "MPNST-wide" essential genes, as well as predicted to be lethal when knocked out, based on synthetic lethality predictions derived from CNV data (with the exception of MPNST724) (Fig. 4.13, Table 4.7).

MYC, a transcription factor responsible for targeting and activating genes stimulating proliferation, is a well-known common dependency in many cancers [178]. Despite its molecular properties that make it nearly impossible to target directly [179], a recent Phase I clinical trial tested an 'indirect' *MYC* inhibitor called OMO-103, a 91 amino acid protein that can interfere with *MYC* dimerization to *MAX* and so prevent their DNA transcription activation properties. This study has shown signs of clinical benefit in some patients [180]. Moreover, *MYC* overexpression has been found associated with MPNST tumorigenesis as well [181].

NAE1 has only relatively recently garnered attention in cancer research, where it has been found to be overexpressed in various cancers, including breast cancer [182].

NAE1 is involved in regulating key cellular processes, such as the suppression of apoptosis and senescence, which are important during tumour progression.

Both *MYC* and *NAE1* have emerged as candidate hit genes worth exploring and validating as potential targets for MPNST. Although both genes are predicted to have a relatively large number of potential synthetic lethality partners [102], the observed gene effects could be explained by a minimal subset of lost genes. For *NAE1*, these could minimally include *ARHGAP36*, *CDKN2A*, and *DAZ3*. For *MYC*, the lost genes could minimally include *KDM5D*, *ARHGAP6*, *GRK1*, *HDLBP*, and *CXXC1*. Mechanistic validation is required to determine which of these lost genes may truly be responsible for the predicted synthetic lethality effects observed in each cell line. This validation will help to confirm the potential relevance of *MYC* and *NAE1* as therapeutic targets in MPNST.

Beyond synthetic lethality predictions, another approach is to examine the intersection between MPNST essential genes identified through CRISPR screens and known druggable protein targets. For example, the United States Food and Drug Administration (FDA) has a curated list of approved drugs, which can be used to determine which gene targets from the CRISPR screen in this study are targetable with existing FDA-approved drugs. Through this approach, superoxide dismutase 1 (*SOD1*) and, once again, *MYC* were identified as such targets (Fig. 4.13, Table 4.7).

Notably, the drug registered for targeting *SOD1* is an antisense oligonucleotide that binds to *SOD1* mRNA, thereby reducing its synthesis. This drug is currently being studied in the context of treating amyotrophic lateral sclerosis (ALS) patients with *SOD1* mutations (Table 4.7)[183].

Another relevant effort involved a study that constructed a comprehensive and up-to-date set of druggable human genes, referred to as the “druggable genome.” This list includes bio-therapeutics, expanding upon previously proposed gene target lists, and incorporates drug-binding predictions based on sequence and structural similarities between known and predicted drug targets[100].

Intersecting this comprehensive druggable genome with the proposed MPNST CRISPR screen hits from this study, along with DepMap MPNST cell line screens, revealed an additional 23 druggable hits (Fig. 4.13, Table 4.7). Many of these are well-known cancer targets, including Aurora kinases A and B (AURKA, AURKB), cyclin-dependent kinases (CDK1, CDK4, CDK6), CHEK1, and MTOR. Additionally, PLK1, TOP2A, KIF11, and EIF4E are currently being explored as targets for cancer therapy [184], [185], [186], [187]. This intersection highlights potential targets for therapeutic intervention in MPNST, using drugs already being investigated or developed for other cancer types.

Together, these druggable genes (or those predicted to be druggable) that were also found to be essential across MPNST cell line models from CRISPR screens propose a list of hits for validation. This approach offers the potential for a more rapid translation into novel therapeutic options for MPNST patients, leveraging existing or developing therapies that target these MPNST dependencies.

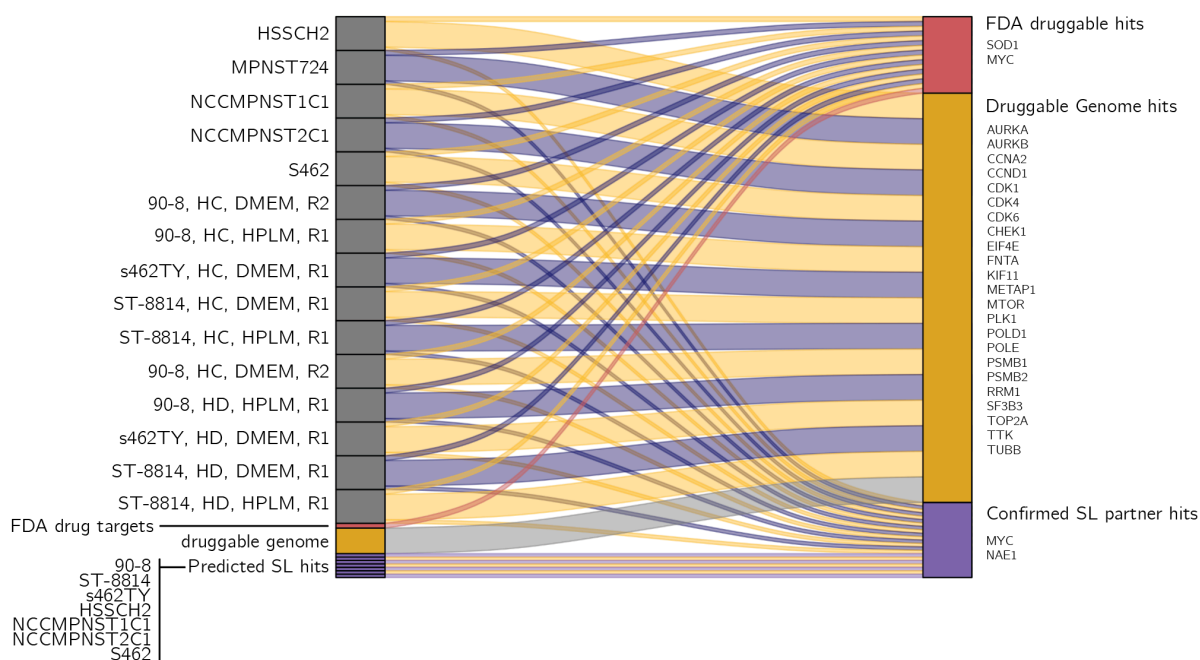


Figure 4.13. Essential genes across MPNST cell line models in this study and DepMap dataset include similar FDA-approved drug- and synthetic lethality prediction targets. Sanky flow diagram depicting the intersection of CRISPR screen hits found in eight different MPNST cell lines and in fifteen independent experiments, with the FDA druggable proteome, the “druggable genome” and predicted synthetic lethality partners for these cell lines. Experimental and copy number data for the HSSCH2, MPNST724, NCCMPNST1C1, NCCMPNST1C2, and S462 cell lines was obtained from DepMap. The FDA druggable proteome list was obtained from the FDA webpage[188]. The druggable genome list was obtained from the supplemental information in the original study[100]. Synthetic lethality partners were predicted using [SynLethDB 2.0](#). Note: MPNST724 had no predicted synthetic lethality partners and so no MPNST724-specific synthetic lethality prediction list was used as a criterion for this analysis. SL: synthetic lethality, FDA: Food and Drug Administration (United States of America).

Table 4.7. Gene hits and associated drugs

Gene	Known drug/predicted	Reference
AURKA	Alisertib, AT9283, Danusertib, ENMD-2076, Hesperadin, Reversine, SNS-314, Tozasertib, ZM447439	[189]
AURKB	AT9283, Baracertib, Hesperadin, Reversine, SNS-314, Tozasertib, ZM447439	[189], [190], [191]
CCNA2	predicted	[192]
CCND1	predicted	[192]
CDK1	BAY 1000394, Dinaciclib, Flavopiridol, RO-3306	[193]
CDK4	Abemaciclib, BAY 1000394, Flavopiridol, palbociclib, PF-06873600, ribociclib, Trilaciclib	[193]
CDK6	abemaciclib, palbociclib, PF-06873600, ribociclib, Trilaciclib	[193]
CHEK1	AZD7762, GDC-0575 , LY2603618, PF-477736, Prexasertib, SRA737	[194]

EIF4E	ribavirin	[195]
FNTA	BMS-214662, FTI-277, L-778,123, Lonafernib, Manumycin A, Sarasar, Tipifarnib	[196], [197]
KIF11	4SC-205, Dimethylenastron, Filanesib, Ispinesib, Litronesib, Monastrol, SB-743921	[198]
METAP1	predicted	[192]
MTOR	AZD8055, BEZ235, Everolimus, GDC-0980, INK128, OSI-027, PP242, Ridaforolimus, Sirolimus, Temsirolimus, Torin 1, Torin 2	[199]
MYC	OMO-103	[180]
PLK1	BI 2536, GSK461364, HMN-214, NMS-P937, Onvansertib, Rigosertib, Volasertib	[200]
POLD1	Predicted	[192]
POLE	Predicted	[192]
PSMB1	Predicted	[192]
PSMB2	Predicted	[192]
RRM1	Predicted	[192]
SF3B3	FR901464, meayamycin, pladienolide B, Spliceostatin A	[201], [202], [203]
SOD1	Tofersen ¹	[204]
TOP2A	doxorubicin, etoposide, Mitoxantrone	[205], [206]
TTK	BAY 1161909, CCT271850, Mps-BAY2a, NMS-P715	[207]
TUBB	Predicted	[192]

4.3 Discussion

In this chapter, I explored the impact of cellular context for gene dependency by varying medium conditions, specifically HPLM versus DMEM. Using two sgRNA libraries and enAsCas12a, I identified potential MPNST-wide essential genes that could serve as potential candidate therapeutic targets. Ultimately, I propose candidate genes for further validation, also prioritising those predicted to be synthetic lethality partners or those that are druggable, to enhance their potential as viable therapeutic targets.

Cells cultured in HPLM, in comparison to the regular medium, DMEM, displayed different growth rates. Although cells growing in HPLM managed to catch up partially, it still took at least more than approximately two days before a steady growth rate was achieved. This suggests that the phenotypic effect of switching medium is not immediate,

but rather a shift that happens over time, and so it was key to integrate multiple late time points during CRISPR screen data analysis.

Target cell lines grown in HPLM or DMEM were also characterised using transcriptomic analysis and growth assays. These analyses suggest considerable variability in transcriptomic responses between cell lines with the two different media. Specifically, distinct effects were observed across the two tested media types when analysing gene sets. Cells cultured in HPLM exhibited enrichment of specific gene sets using GSEA, including those involved in oxidative phosphorylation and fatty acid oxidation. These gene sets also demonstrated higher essentiality in HPLM compared to DMEM during CRISPR screens. A selection of genes associated with these enriched pathways were found to also differed significantly between HPLM and DMEM on an individual level. To confirm whether these individual genes are truly essential in this context, further mechanistic validation is required.

Interestingly, although HPLM was designed to more accurately reflect physiological conditions, this was not consistently reflected by the transcriptomic data. In some cases, correlations between cells grown in DMEM and actual MPNST samples were stronger than those between cells grown in HPLM and MPNST samples. This suggests that, at the transcriptomic level, HPLM does not necessarily bring the cellular phenotype closer to true MPNST pathophysiology. However, it is important to consider that the MPNST samples used for comparison were bulk-sequenced, potentially reflecting the transcriptomics of the broader tissue environment, including fibroblasts, immune cells, and stromal cells, rather than being specific to MPNST cancer cells. It is also important to note that cultured cancer cell lines do not necessarily strongly correlate with the cancer

it aims to represent. Specifically, one study integrated several large-scale cell line and bulk-tumour RNA-Sequencing datasets, accumulating data across 1,249 cell lines, 12,236 independent tumour samples and 37 different cancer types, and found that cell lines cluster together more closely with each other regardless of origin, rather than with their associated tumour type[208]. It would require high-coverage single cell RNA and WGS sequencing to determine whether the difference is attributable to pollution of additional non-cancerous cell types. Of course, different types of cancer models, including organoids, may better represent the cancer of interest, but are less suitable for high-throughput experiments[209]. Additionally, transcriptomic data provides only a partial view of the cellular phenotype, potentially overlooking critical aspects at the protein and wider metabolic levels. Although recent efforts in single-cell sequencing-based metabolomics efforts are commendable[151], [152], high-throughput methods, including mass spectrometry, that directly measure metabolic parameters may be more reliable[210].

Therefore, while DMEM appears to exhibit slightly better correlation with MPNST samples than HPLM, further research is needed to determine which medium context is truly more reflective of physiological conditions. It is also possible that while certain medium contexts may better represent physiology on a whole-cell level, specific pathways could be more accurately reflected in media that do not fully recapitulate the physiological environment at the cellular level. Following this rationale, it may still be worth exploring different media types to study particular subcellular processes and phenotypes.

The observation that medium-dependent differentially expressed genes do not correspond to a significantly higher number of different gene dependencies: rather, the opposite, that further suggests that transcriptomics may not be the most effective

approach for understanding how medium context influences cellular phenotype. With this observation in mind, medium-dependent essential genes were investigated independently of transcriptomic analyses. This approach led to the identification of potential medium-dependent hits in MPNST cell lines, as well as across various other cell lines. These findings reinforced the essential pathways already recognised within specific medium contexts, although none of the identified genes were part of any curated gene sets that showed differential essentiality in particular medium conditions.

In order to find genes that were essential regardless of medium context, MPNST CRISPR screen hits from this study were intersected with CRISPR screen data on MPNST cell lines from DepMap. Hit prioritisation included selection based on whether hit genes were predicted to be synthetic lethality partners of lost genes in these cell lines. This was performed by utilising SynlethDB 2.0[103], a list of gene pairs predicted or experimentally proven to have a synthetic lethality relationship. As previously discussed, confirming a true synthetic lethality relationship is challenging, as they often turn out to be context-dependent[211], and so this tool should be seen more as a means to prioritise hits rather than assuming such a relationship. Another means to prioritise hits was to see whether CRISPR screen hit genes were (predicted to be) druggable with small molecules or regular drugs, or even FDA approved drugs, as this would enable rapid translation to clinical benefit to MPNST patients.

Using these two approaches, two genes were found to be synthetic lethality predicted genes and 25 were druggable targets across the studied MPNST cell lines (Fig 4.13 and Table 4.6). Of the druggable targets, AURKA, AURKB, CDK1, CDK4, CDK6, MTOR, and PLK1, were also found targets in a recent drug screening in organoids derived

from MPNST patients[212]. This study collected 194 samples from 126 patients across 24 sarcoma subtypes and leveraged a high-throughput drug-screening workflow to test various agents to repurpose these to treat sarcomas. None of the drugs they tested were targeting any of the other hit genes as displayed in Figure 4.13. Together, these findings add to the increasing evidence that Aurora kinases, CDKs, MTOR and PLK1 are candidate targets for MPNST treatment [212], [213], [214], [215], [216].

Importantly, during the SARC023 trial, the mTOR inhibitor sirolimus was combined with the Hsp90 inhibitor ganetespib to treat refractory MPNST. While the combination was tolerable, it did not demonstrate significant clinical benefits in patients. While further investigation into the mechanisms behind resistance to mTOR inhibition in MPNSTs remains essential, future experimental validation efforts may be more productively directed toward targeting PLK1, CDKs, and Aurora kinases.

This curated list of candidate genes represents a promising starting point for identifying therapeutic targets in MPNST, but further experimental validation is essential to confirm their potential. For genes predicted to have synthetic lethality relationships, an important next step will be to perform complementation assays using inducible systems. By selectively inducing or repressing expression of the candidate genes alongside their predicted synthetic lethality partners, it will be possible to assess whether co-inactivation results in cell death, thereby confirming or refuting these relationships in the context of the tested cell line. To ensure the robustness of these findings, it will be important to test these relationships across multiple MPNST cell line models, as well as perhaps across non-MPNST models, to assess whether these interactions are robust across different cell types or if they are highly context dependent.

For druggable targets, validation will involve a combination of both genetic and pharmacological approaches. Inducible gene knockouts using CRISPR/Cas9 or other gene-editing technologies will be essential for directly assessing the essentiality of these genes. Simultaneously, pharmacological assays using small-molecule inhibitors or FDA-approved drugs targeting these proteins will help determine whether pharmacological inhibition mirrors the effects of genetic knockouts, ensuring a comprehensive validation of these druggable targets.

For *in vivo* validation, various mouse models, including NF1-associated MPNST models or xenografts, can be employed to evaluate the efficacy of targeting these genes. These models will provide critical insight into whether genetic or pharmacological inhibition can suppress tumour growth, advancing the therapeutic potential of these targets toward clinical application.

Chapter 5: Polycomb Repressive Complex 2-context (PRC2) dependent synthetic lethality

5.1 Introduction

Polycomb Repressive Complex 2 (PRC2) loss is a well-documented driver of malignant peripheral nerve sheath tumour (MPNST) initiation and progression, and is associated with poor progression-free survival rates [217]. Despite its significant role in MPNST tumorigenesis, PRC2 loss has been relatively understudied, largely because it is not a common feature in other cancers. Consequently, there are limited therapeutic strategies aimed at targeting PRC2 loss. Specifically, PRC2 loss of function may induce synthetic lethality relations that can be targeted for therapeutic purposes.

Previous research has shown that the loss of Polycomb Repressive Complex 2 (PRC2) sensitises malignant peripheral nerve sheath tumour (MPNST) cells to inhibition of Bromodomain-containing protein 4 (BRD4), a prominent member of the bromodomain and extra-terminal domain (BET) protein family [49]. BRD4 functions by binding to acetylated lysine residues on histone tails, where it facilitates the recruitment and stabilization of transcriptional machinery, including RNA polymerase II and positive transcription elongation factor b (P-TEFb), to initiate and maintain transcriptional activity [218]. In the context of PRC2 deficiency in MPNST, reduced levels of H3K27Me3 shift the histone modification balance towards H3K27 acetylation, resulting in increased expression of PRC2 target genes. BRD4 inhibitors, such as JQ1, disrupt this transcriptional activation, showing synergistic inhibitory effects on RAS signalling pathways in MPNST cell line models, with promising efficacy in both *in vitro* and xenograft models [71]. However, despite these encouraging preclinical outcomes, resistance to BRD4 inhibitors remains a major hurdle in clinical applications, largely due to compensatory mechanisms and feedback loops that maintain BRD4 activity and

support cellular survival, underscoring the need for further research to address this emerging challenge [70].

To identify alternative therapeutic strategies for targeting cells with PRC2 loss, a more recent study utilised RNA interference (RNAi) screens in MPNST cell line models, revealing that PRC2 loss induces sensitivity to the disruption of DNA methyltransferase 1 (DNMT1) function, which they validated through CRISPR/Cas9-mediated knockout and pharmacological inhibition experiments. This study also initially employed non-isogenic cell line models with PRC2 loss or wild-type PRC2 to identify PRC2 context-specific dependencies[219]. The outcome was only a partial (approximately 2-fold) sensitivity in cell xenograft models, even though PRC2 dependency appeared consistent. An alternative and more rigorous approach would involve comparing essential gene dependencies in isogenic cell lines with or without functional PRC2. This approach may eliminate confounding variables associated with genetic differences between cell lines, allowing for a direct comparison of PRC2-dependent vulnerabilities. Notably, high-throughput screens in isogenic cell lines that differ solely in their PRC2 status have not yet been attempted and was the impetus for the experiments in this chapter.

One approach to identifying context-specific dependencies of PRC2 involves disrupting its function in a cell line that possesses functional PRC2. This can be achieved either by generating a stable cell line in which one or more PRC2 components are knocked out, for example by using CRISPR, or through protein-targeting methods such as drug inhibition or targeted degradation using PROTACs. Creating cell lines with homozygous loss-of-function mutations can be a time-intensive process, particularly when requiring the expression of *enAsCas12a-T2A-EGFP*. This is because *enAsCas12a-T2A-EGFP*

complementation must be performed after PRC2 knockout generation, given that prolonged expression of enAsCas12a-T2A-EGFP often leads to its silencing over time. In contrast, disrupting PRC2 function through drug inhibition or targeted degradation offers an acute and reversible approach that can be highly selective. This approach also facilitates direct comparisons between isogenic cell lines, even at the level of enAsCas12a-T2A-EGFP expression, as the phenotype can be rapidly induced by simply adding the drug, bypassing the need for complex genetic manipulations.

Recently, a proteolysis-targeting chimera (PROTAC) known as UNC7700 was developed to specifically target and degrade the PRC2 component EED (Fig. 5.1)[1]. UNC7700 binds to the WD40 aromatic cage of EED on one end and to Cullin-RING Ligase (CRL) 2 Von Hippel-Lindau protein (VHL) (CRL2^{VHL}) on the other, linked by a cis-cyclobutane linker. This PROTAC induces the ubiquitination and subsequent proteasomal degradation of EED, EZH2, and, to a lesser extent, SUZ12, and so disrupts PRC2 function. UNC7700 has demonstrated anti-proliferative effects in diffuse large B-cell lymphoma (DLBCL) cell lines.

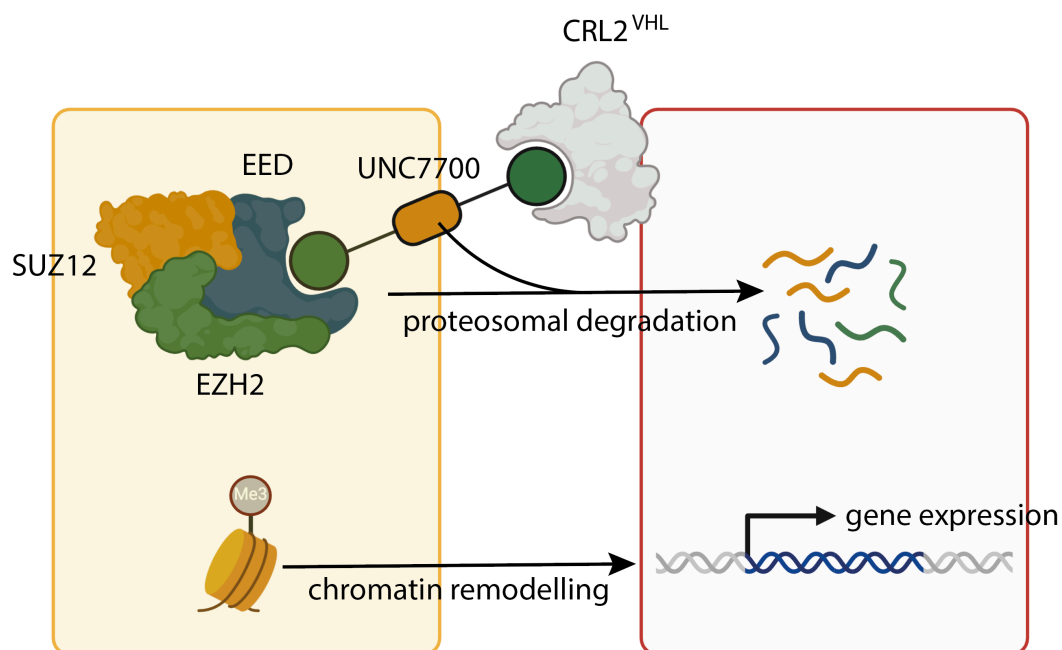


Figure 5.1. PROTAC UNC7700 designed to bind EED, degrade PRC2 components, and disrupt PRC2 functionality. Depiction of UNC7700 PROTAC binding to EED in PRC2 leading to Cullin-RING Ligase (CRL) 2 Von Hippel-Lindau protein (VHL) mediated ubiquitination and subsequent proteasomal degradation. Upon PRC2 loss, target genes are derepressed via chromatin remodelling leading to gene expression. Adapted from a figure in [1].

In this chapter, I will investigate the use of UNC7700 to disrupt PRC2 function and assess its impact on cellular dependencies in immortalised human Schwann cells (iHSC). This investigation will include both biochemical and transcriptomic analyses following UNC7700 treatment, as well as genome-wide CRISPR knockout screens in the context of functional versus disrupted PRC2 functionality. These experiments demonstrate a streamlined method for uncovering dependencies that arise specifically from PRC2 loss, providing a foundation for potential therapeutic targets that can be further validated experimentally.

5.2 Results

5.2.1 EED-targeting PROTAC UNC7700 degrades PRC2 components and derepresses target genes.

To optimise the treatment of iHSC cells with UNC7700, a dose-response experiment was conducted using a range of concentrations (0, 1 nM, 10 nM, 0.1 μ M, 1 μ M, 10 μ M) to assess the degradation of PRC2 components. The cells were treated for eight days, with drug and medium being refreshed on day four. The 10 μ M concentration produced the clearest reduction in PRC2 components, as demonstrated by reduction of EED, SUZ12 and EZH2 in western blots (Figure 5.2A). While the 1 μ M concentration also resulted in a marked decrease in PRC2 components, particularly EED and EZH2, the extent of reduction appeared considerably less compared to the 10 μ M concentration. Notably, the decrease in H3K27me3 levels was more pronounced at the 10 μ M concentration, although this technique used does not allow for absolute quantification. Subsequently, the optimal duration of UNC7700 treatment was investigated to determine the time required for a significant reduction in PRC2 components and H3K27 trimethylation. After one day of treatment, reductions in EED and EZH2 levels were already apparent, with further decreases observed with prolonged exposure (Figure 5.2B). By day four, the levels of EZH2 and EED appeared to stabilise at lower levels. H3K27me3 showed a substantial reduction only after two days of treatment, with the maximal effect observed after eight days.

To investigate the downstream effects of PRC2 disruption, iHSC cells were treated with 10 μ M UNC7700 for eight days. Following treatment, the cells were snap-frozen, and mRNA was extracted for next-generation sequencing (paired end RNA-Seq).

Principal component analysis (PCA) of the transcriptomic data revealed hardly any separation between UNC7700-treated cells and DMSO controls along the first principal component (PC1), even when this accounted that accounted for 59% of the total variance (Figure 5.2C).

A heatmap of the top differentially expressed genes only showed genes that are overexpressed in the UNC7700-treated condition, with none showing overexpression observed in the DMSO controls (Figure 5.2D). This pattern is consistent with the known role of PRC2 as a transcriptional repressor. Here, disruption of PRC2 function by UNC7700 leads to de-repression of its target genes. Notably, among the genes overexpressed following UNC7700 treatment, are LRCH2 and BRINP3 that have previously been identified as established PRC2 targets[220]. The extent of gene expression changes induced by UNC7700 treatment is further illustrated by the volcano plot in Figure 5.2E, where the vast majority of significantly differentially expressed genes were upregulated, again reflecting the expected gene de-repression due to PRC2 inhibition.

To explore the biological processes associated with these differentially expressed genes, GSEA was performed. Despite the pronounced differences in gene expression between UNC7700-treated cells and DMSO controls, as highlighted in the volcano plot (Figure 5.2E), only six gene sets from the Hallmark, Kegg Legacy, and Reactome collections were significantly enriched in the UNC7700 treatment group (Table 5.1).

The PI3K-AKT-mTOR signalling pathway was amongst the enriched pathways and may be relevant in the context of PRC2 loss induced by UNC7700. This pathway is crucial for regulating cellular processes such as proliferation, differentiation, and cancer

progression[221]. In Schwann cells, PI3K-AKT-mTOR signalling is also involved in cellular migration, myelination, and demyelination, underscoring its broader significance beyond traditional oncogenesis-related functions[222]. Moreover, a previous study has shown that dual mTORC1/2 inhibition exerts anti-proliferative effects in neurofibroma and MPNST cells, suggesting a potential therapeutic avenue[223].

These results point to a possible connection between PRC2 loss and enhanced mTOR signalling, suggesting that MPNST initiation or progression could, in part, rely on activation of the mTOR signalling pathway following PRC2 disruption. This finding opens-up new avenues for research into the role of mTOR signalling in PRC2-deficient contexts and its implications for tumorigenesis.

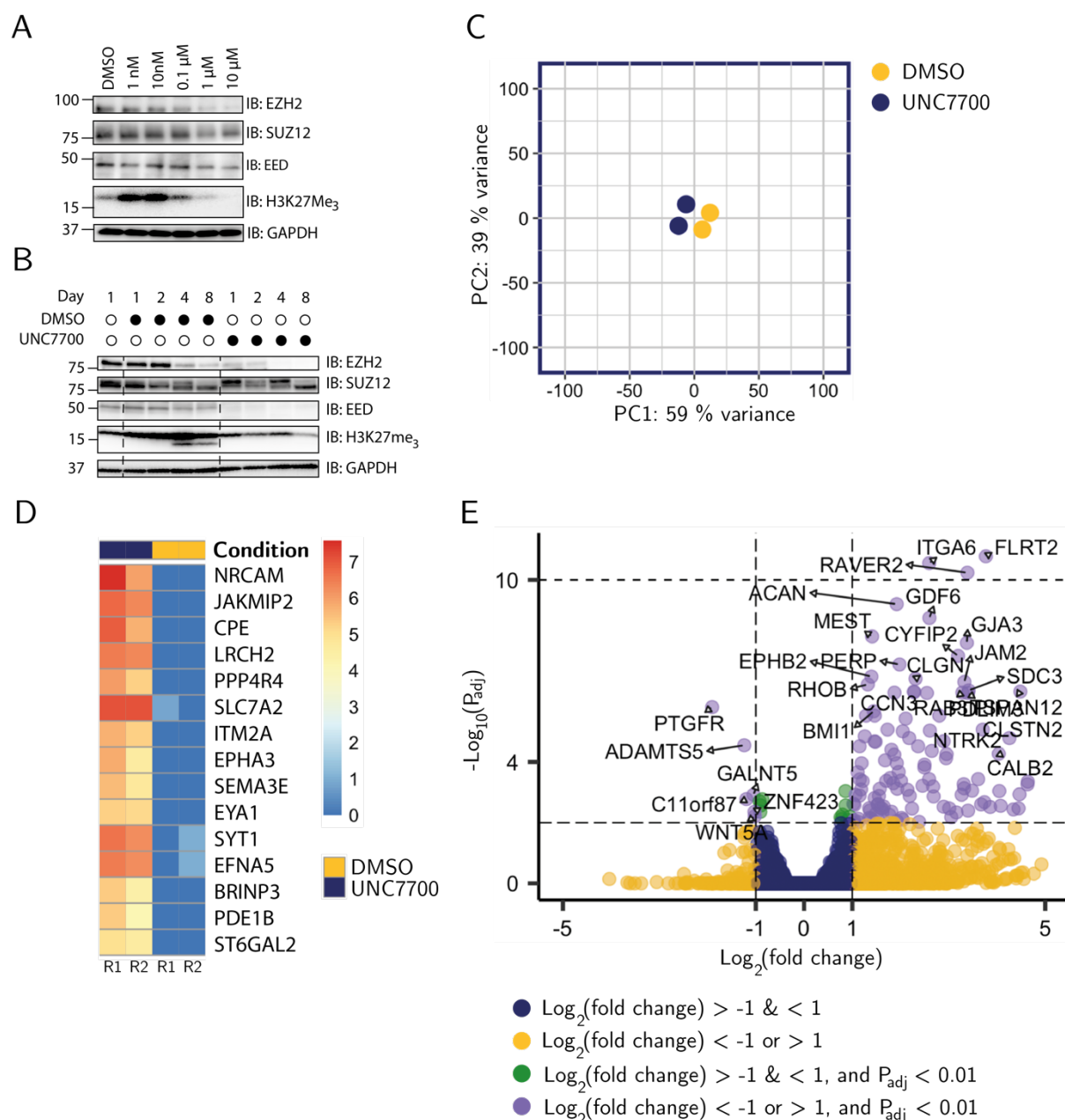


Figure 5.2. UNC7700 degrades PRC2 components and de-represses PRC2 target gene expression in iHSC. Levels of PRC2 core components (EZH1, EZH2, SUZ12, and EED), and PRC2 activity (H3K27ME3) in iHSC cells treated with UNC7700 or DMSO, visualised by Western blotting. Whole-cell extracts were prepared from exponentially growing cells treated with either UNC7700 (10 μ M) or DMSO for up to 8 days. EZH1, EZH2, SUZ12, EED, and H3K27ME3 were detected by anti-EZH2, -SUZ12, -EED, and -H3K27ME3, respectively. GAPDH, which was used as a loading control, was detected by anti-GAPDH. IB: immunoblot. **(B)** Same as (A) but including whole-cell extracts of cells that were treated with DMSO for up to 8 days to evaluate a DMSO-specific effect on PRC2 components and functionality. **(C)** Heatmap of top 15 differentially expressed genes between iHSC cells treated with UN7700 and DMSO for eight days. Colour indicates Log₂(fold change). R1: replicate 1, R2: replicate 2. **(D)** Principal component analysis of transcriptomics in iHSC cells treated with UNC7700 (10 μ M) or DMSO. The variance explained

by principal components is indicated on the side of each axis. PC1: principal component 1, PC2: principal component 2. **(E)** Volcano plot representing whole-transcriptome differential gene expression analysis. Levels of fold change and significance are colour coded as indicated. Significantly differentially expressed genes are labelled while allowing limited overlaps.

Table 5.1. Significantly enriched or depleted gene sets in iHSC cells treated with UNC7700 versus DMSO.

Gene set collection	Gene set	Enriched with	FDR (<25%)
Hallmark	HALLMARK_PI3K_AKT_MTOR_SIGNALING	UNC7700	0.193
	HALLMARK_GLYCOLYSIS	UNC7700	0.193
	HALLMARK_PROTEIN_SECRETION	DMSO	0.229
KEGG Legacy	KEGG_CHRONIC_MYELOID_LEUKEMIA	UNC7700	0.247
	KEGG_TIGHT_JUNCTION	UNC7700	0.245
	KEGG_PROSTATE_CANCER	UNC7700	0.236
Reactome	None		

5.2.2 CRISPR screens reveal candidate PRC2 context-specific essential genes

To explore whether PRC2 context influences dependencies in iHSCs, separate CRISPR screens were conducted using the Humagne C and D libraries in cells treated with either UNC7700 (10 μ M) or DMSO (as a control at an equivalent volume). The treatment regimen with UNC7700 or DMSO began eight days prior to CRISPR library transduction. During the CRISPR screen, culture media were refreshed at least every four days with fresh UNC7700 or DMSO being added accordingly. During CRISPR screens, intermediate and final time point samples were processed for sequencing and analysed (Chronos) following the workflow described in Chapter 3.

To identify differential genetic dependencies between the UNC7700 and DMSO conditions, Chronos scores from both conditions were visualised in a correlation dot plot

for both the Humagne C and D screens separately (Fig. 5.3A, B). The 10 genes exhibiting the most substantial differences in Chronos scores, showing higher dependency under the UNC7700 treatment, are highlighted in Figure 5.4. Notably, among these genes is Phosphoinositide 3-kinase regulatory subunit 4 (PIK3R4), a component of the PI3K-AKT-mTOR signalling pathway. However, these selected observations, including the one concerning PIK3R4, were only detected in either the Humagne C or D screens individually, and so are based on single biological replicates. Therefore, these findings warrant further validation, ideally first in a high-throughput manner before proceeding with more labour-intensive focused experimental validation.

Furthermore, certain hits exhibited consistent differential effects between UNC7700 and DMSO treatments across both the Humagne C and D screens, indicating these may be more reliable at this stage of the investigation (Fig. 5.3C). Depending on the thresholds set, a discrete set of genes with differential effects between UNC7700 and DMSO conditions can be identified, showing consistent gene effects in both Humagne C and D CRISPR screens. Using an arbitrary threshold of a difference in Chronos score of < -1.75 or > 1.75 in both Humagne C and D (Fig. 5.3C), 19 genes were identified as having differential effects. Of these, six genes, DCTN2, LRR1, NOC3L, RPP21, SRSF11, and TOP1, demonstrated substantially lower Chronos scores in the UNC7700 condition compared to DMSO in both screens (Fig. 5.5).

Among these, topoisomerase 1 (TOP1) emerges as a potentially interesting candidate for further investigation. TOP1 is known to alleviate DNA supercoiling by transiently breaking and re-ligating a single DNA strand, thereby facilitating replication and transcription. In cancer, progression is sometimes dependent on TOP1-mediated

transcription-associated mutagenesis [224], and TOP1 inhibitors, like irinotecan, Topotecan, and Camptothecin, are widely employed in cancer treatment, targeting the elevated TOP1 expression on which certain cancers rely for their high DNA metabolism[225], [226], [227], [228]. Moreover, TOP1 inhibition has been shown to induce context-dependent lethality in MYC-driven cancers by exacerbating R-loop accumulation, already elevated due to MYC dysregulation, resulting in genomic instability and selective death of MYC-transformed cells[229]. None of the other six genes have drugs targeting the associated protein. It is important to note that while all genes display a difference in Chronos score, the genes may still be essential in both contexts. The difference may only indicate a higher dependency in one specific context compared to the other. Furthermore, some of these genes may be common essential genes and may suggest the difference in Chronos score is due to noise rather than true signal.

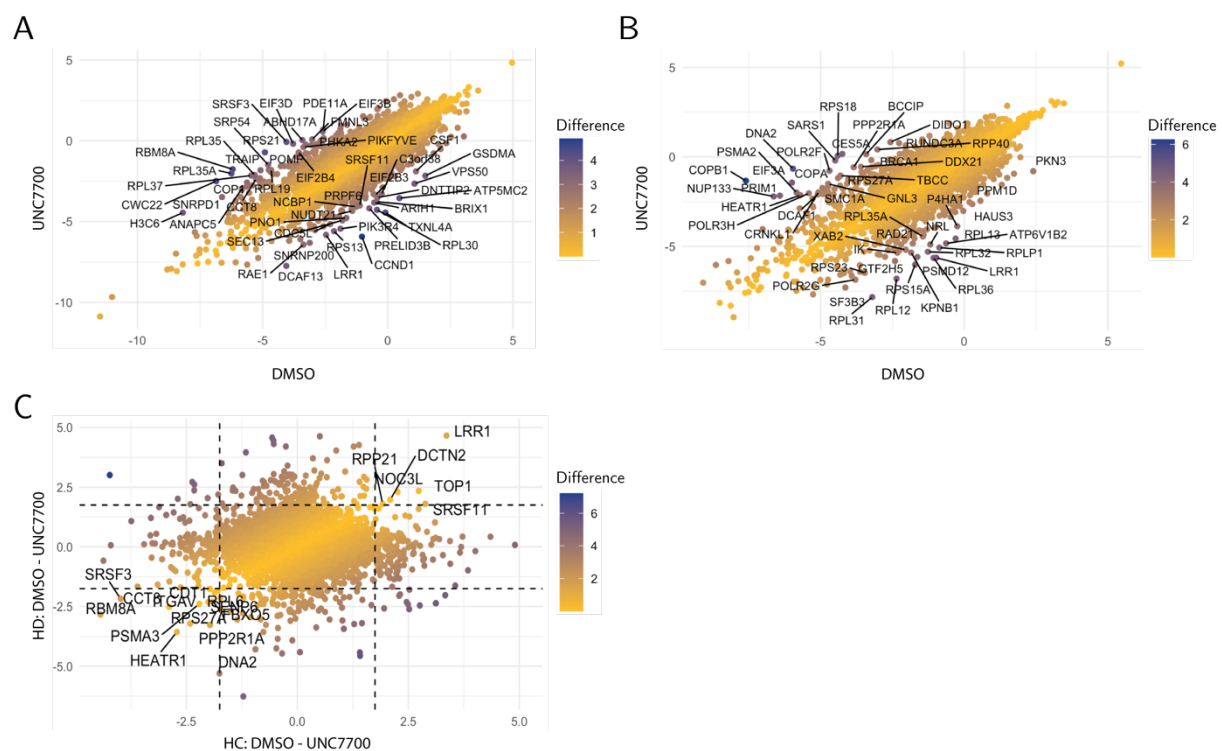


Figure 5.3. UNC7700 CRISPR screen suggests PRC2 context-dependent hits in Humagne C and D library screens. (A-B) Dot plot displaying the relationship between Chronos z-scores per pDNA bin from

UNC7700 and DMSO Humagne C (A) and Humagne (D) CRISPR library screens. Differences in Chronos z-scores per pDNA bin between UNC7700 and DMSO are indicated by colour, with bigger differences turning from yellow to more blue. Differential gene hits are labelled while allowing limited overlapping. (C) Dot plot displaying the relationship between Humagne C and Humagne D differential gene hits. Thresholds for common hits between Humagne C and D were set at below -1.75 and above 1.75. Differences in values from Humagne C and D screens are indicated by colour, with bigger differences turning from yellow to more blue.

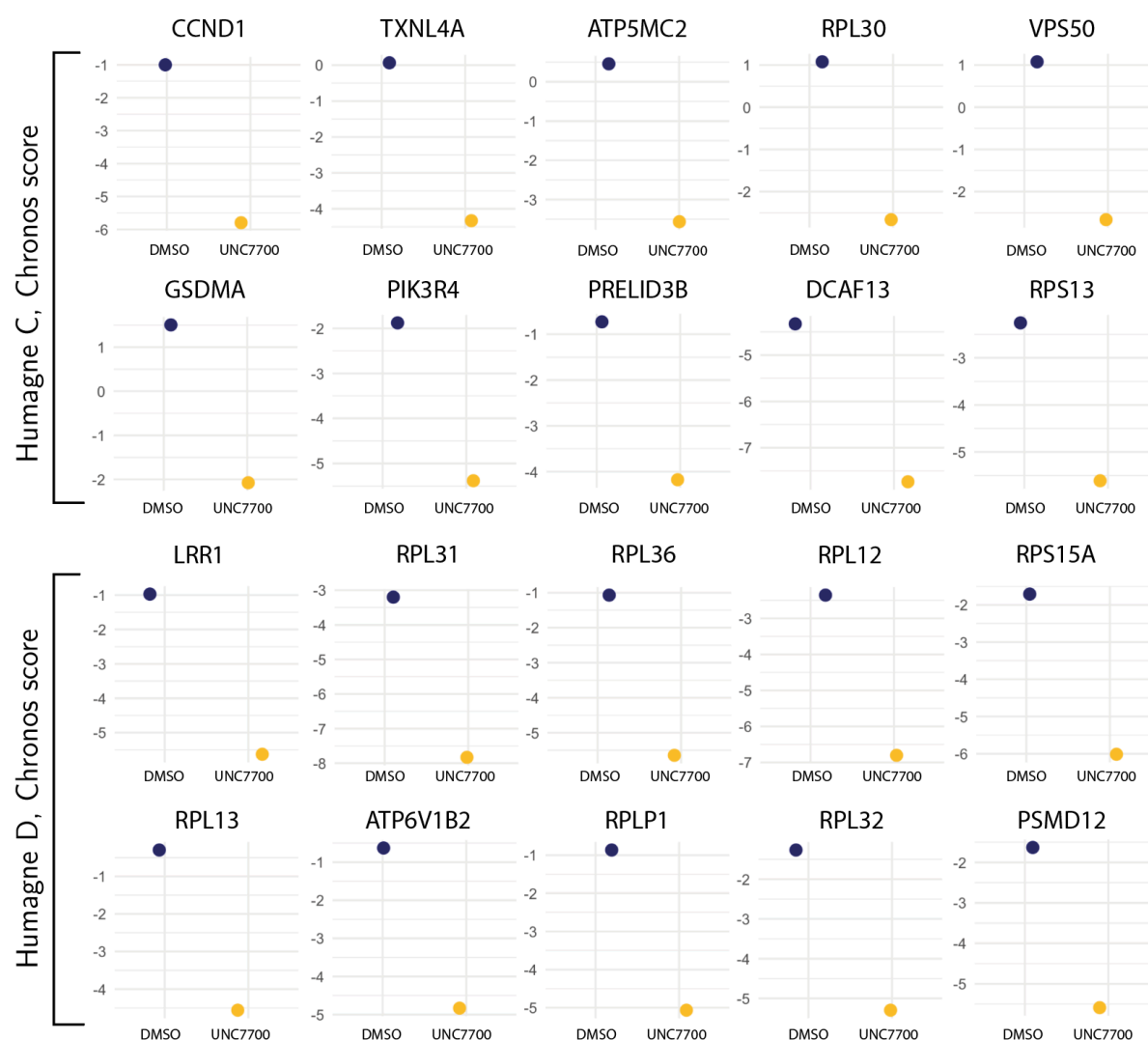


Figure 5.4. Potential PRC2 context-specific dependencies identified in an iHSC cell line model using Humagne C or D. Dot plots of differential hits identified from the UNC7700 vs. DMSO Humagne C or D CRISPR screens. Each dot represents one biological replicate.

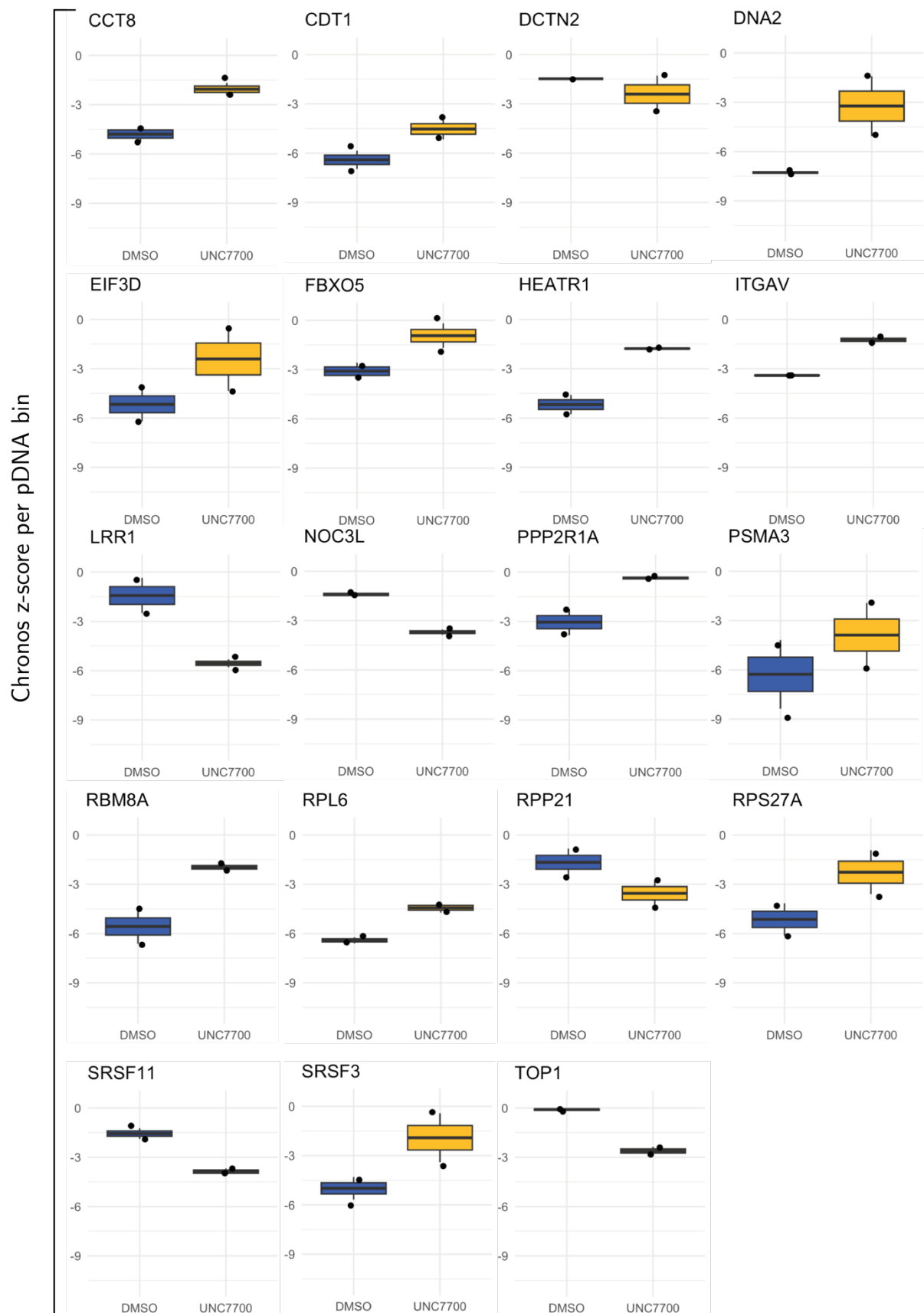


Figure 5.5. Potential PRC2 context-specific dependencies identified in an iHSC cell line model. Box plots of differential hits identified from the UNC7700 vs. DMSO Humagne C and D CRISPR screens. Each

hit is represented by a box plot with two Chronos z-scores per pDNA bin per condition derived from Humagne C and D screens. Means are represented by black horizontal lines and vertical black lines indicate the data range.

5.2.3 PRC2-context dependent changes in gene expression inversely associates with differential gene expression

To assess whether there was a correlation between differential gene expression and differential gene essentiality in the iHSC cells treated with either UNC7700 or DMSO, differential gene essentiality was highlighted on volcano plots of differential gene expression (Fig. 5.6A, B). Despite the clear effect of PRC2 target gene derepression, which resulted in an absolute increase in the number of essential genes in the UNC7700-treated context that were also differentially expressed, the proportion of such genes remained, perhaps unexpectedly, low.

To statistically evaluate the relationship between differential gene expression and differential gene essentiality, absolute values of the differential gene expression $\log_2$ (fold change) were calculated, and genes were categorised into equally sized bins based on these absolute values. The median of these absolute Chronos scores per bin were then used to perform a Spearman's correlation analysis. This analysis revealed an inverse relationship (Fig. 5.6 C, D): lower differential gene expression was associated with higher differential gene essentiality, mirroring observations made in cells cultured in HPLM versus DMEM (Fig. 4.9I-P, Chapter 4).

While in some instances differential gene expression does appear to correlate with differential gene essentiality[230], [231], these results, along with those previously described in Chapter 4 (Fig. 4.9I-P), suggest that differential gene expression is limited

in predicting differential cellular dependencies. Rather, it appears that differential gene essentiality is more strongly associated with genes that are under tighter regulatory control of gene expression. This underscores the complexity of the relationship between gene expression and essentiality, suggesting that factors beyond mere changes in transcript levels play significant roles in determining gene essentiality in specific contexts.

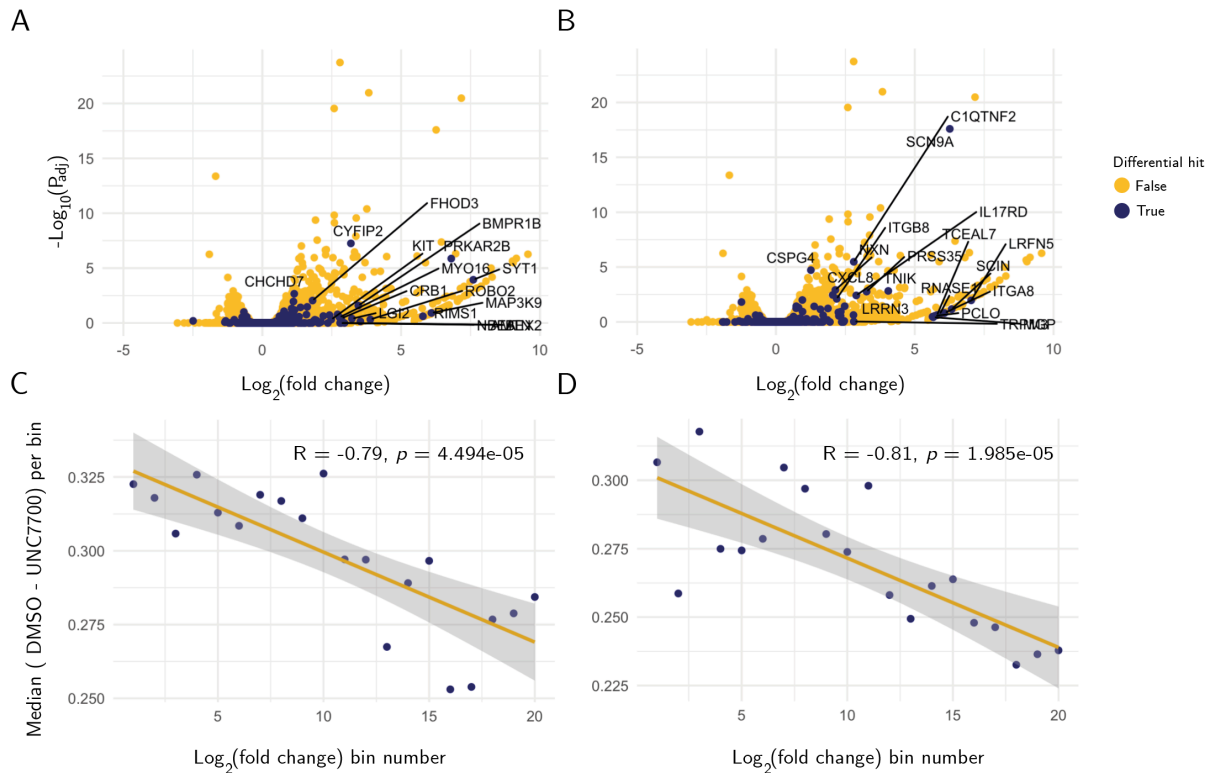


Figure 5.6. UNC7700-dependent differential hits correlate with UNC7700 context-independently expressed genes. (A) Differential gene expression analysis of cells treated with UNC7700 versus DMSO, with top and bottom 5% differential Humagne C library screen hit genes (DMSO – UNC7700) highlighted in blue. The maximum overlap limits the labelled differential hits displayed. (B) Same as in (A) but for differential Humagne D library screen hit genes. (C) Dot plot and linear regression model of the relationship between the $\text{log}_2(\text{fold change})$ gene expression bin number and differential Humagne C library screen hit genes calculated by taking the median of UNC7700-subtracted DMSO chronos scores per gene per bin. A Spearman's rank correlation test was performed to test for significance, with a linear regression line (gold) with a 95% confidence interval indicating the trend along median values per bin (blue). (D) Same as in E but for differential Humagne D library screen hit genes.

5.2.4 Reduction of SUZ12 expression follows complementation in

MPNST cell line

Another isogenic approach to identify PRC2 context-specific dependencies using CRISPR screens is by complementing back *SUZ12* in MPNST cell lines where *SUZ12* was lost. The comparison between dependencies in the MPNST cell lines where *SUZ12* was lost and in which it was complemented back may suggest such PRC2-context specific essential genes.

Additionally, *SUZ12* function as part of the PRC2 complex has not been completely uncovered as yet, and specific *SUZ12* mutations may retain partial and specific PRC2 functionality. With these points in mind, a set of four different versions of *SUZ12* were designed based on previous studies that show a specific functional effect of these mutations in PRC2 complex formation as well as the interaction with co-factors (Fig. 5.7, Table 5.2.).

The first variant involves the substitution of phenylalanine residues F86 and F90 with alanine. These residues are located within an alpha helix of *SUZ12* that is critical for interaction with JARID2 [232]. Previous studies demonstrated that these mutations abolish *SUZ12*'s ability to bind JARID2, a cofactor that, in conjunction with PRC2, is essential for establishing repressive chromatin states and mediating gene silencing. Another variant, R196A, was engineered to disrupt PRC2 dimerization, an interaction fundamental for the structural integrity and function of the PRC2 complex [233]. The third variant, H466A, involves the disruption of a C2H2 zinc finger domain within *SUZ12*, which is critical for maintaining the structural stability of the complex, and its disruption is likely to affect PRC2's ability to interact with its target sites on chromatin [234], [235].

The final variant, S583A, impairs the recruitment of PRC2 to chromatin for the trimethylation of histone H3 at lysine 27 (H3K27me3). The phosphorylation of this serine residue is important for the recruitment of PRC2 to its genomic targets, and its alteration significantly reduces the trimethylation activity [236].

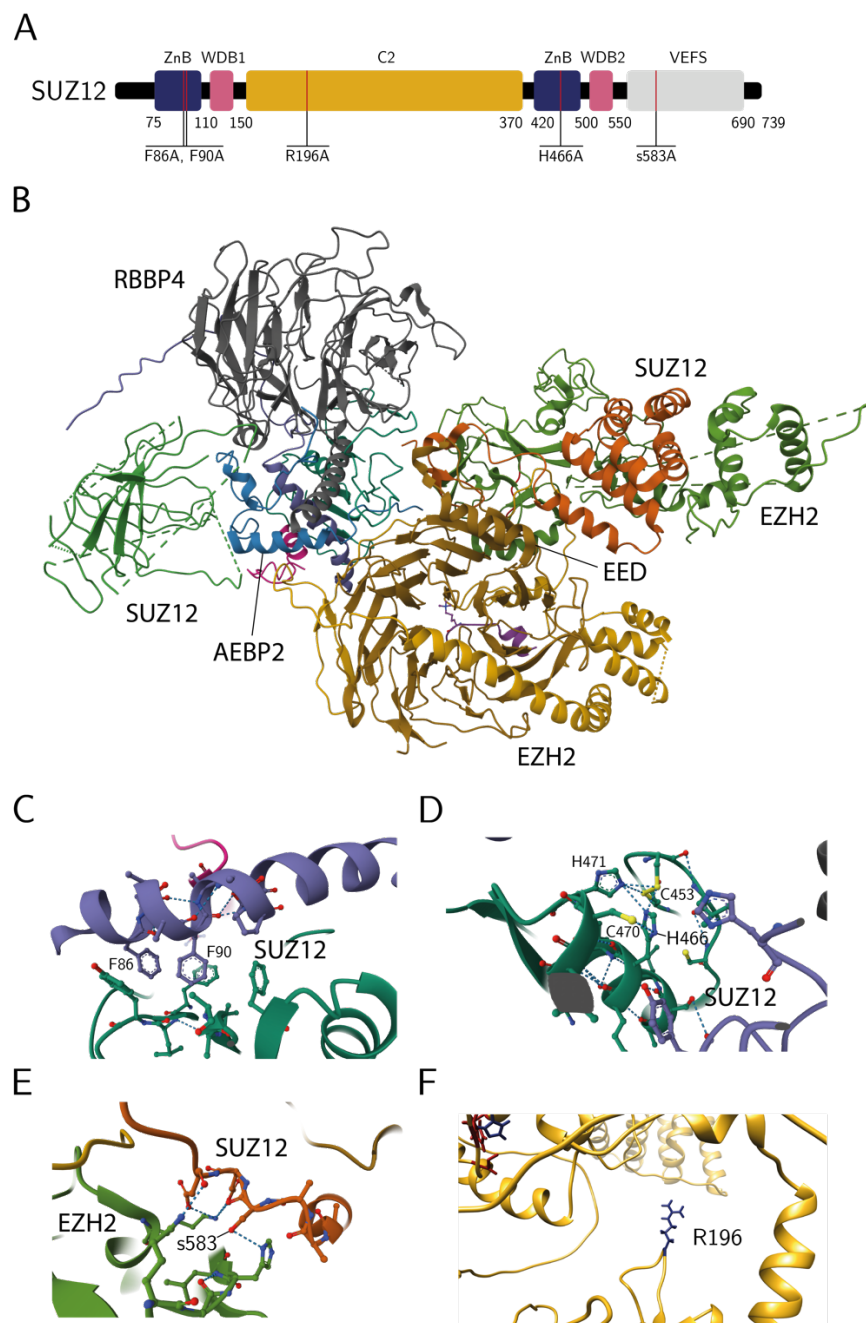


Figure 5.7. SUZ12 mutations targeting key domains highlighted. Schematic depiction of SUZ12 protein domains with an overview of designed mutations targeting SUZ12 functionality. **(B)** Cryo-EM structure of PRC2 binding AEBP2 and RBBP4 [115]. **(C-E)** Intended mutation locations F86 and F90, H466, and

s583, respectively, highlighted in the Cryo-EM PRC2 structure from (B). **(F)** Intended mutation location R196 highlighted in an AlphaFold-based model of SUZ12 (the previously mentioned CRYO-EM model does not include R196 in the predicted structure)[87].

Table 5.2. Impact of Specific SUZ12 Mutations on PRC2 activity.

Mutation	Effect	Reference
F86A-F90A	abolishes JARID2 binding	Chen et al., 2018, Molecular cell [232]
R196A	disrupts PRC2 dimerisation	Chen et al., 2018, Molecular cell [233]
H466A	C2H2 zinc finger disruption	Mueller et al., 2020, Nucleic Acid Research [234]
s583A	impairs PRC2 recruitment for H3K27 trimethylation	Gong et al., 2022, Nature Communications [236]

The design of the SUZ12-complementation plasmids was based on the previously successful *pRDA-174* plasmids, which had effectively delivered the *enAsCas12a-T2A-EGFP* construct. For the purpose of complementing *SUZ12*, instead of using a fluorescent tag for selection, a blasticidin resistance gene was incorporated. This choice was made to enable large-scale transduction and selection, as FACS-based cell sorting, though effective, is too costly and inefficient for selecting large numbers of positive cells. Following lentiviral production, these constructs, including both mutated versions of *SUZ12* and a wild-type *SUZ12* construct, were transduced into the 90-8-*enAsCas12a-T2A-EGFP* target cell line at a multiplicity of infection (MOI) of approximately 0.3 (Fig. 5.8A). After blasticidin selection and subsequent sample preparation, co-immunoprecipitation confirmed the re-expression of all SUZ12 variants except H466A and their binding to PRC2 core components, EED and EZH2 (Figure 5.8B). Additionally, complementation of all SUZ12 variants, except H466A, restored PRC2 function, as indicated by the recovery of H3K27me3 levels. Notably, the H466A variant, which carries a zinc finger mutation, not only resulted in reduced SUZ12 levels as detected by Western

blotting but also led to decreased levels of EED and EZH2, the mechanisms of which remain incompletely understood.

To assess the stability of SUZ12 expression over time, immunostaining was performed on these cells at various time points after transduction and selection, with the parental cell line serving as a negative control and LentiX Hek293T cells as a positive control. Unfortunately, within nine days, the percentage of SUZ12-positive cells had dropped below 1%, making it impractical to study a functional PRC2 phenotype using CRISPR screens (data not shown). To address this reduction in SUZ12 expression, the target cells were transduced at a higher MOI of approximately 3, in hopes that a significant number of cells would integrate at least three copies of the SUZ12 gene. However, despite this high MOI, a rapid decline in SUZ12 expression was observed again; within six days, the proportion of SUZ12-positive cells fell from 65% to 1.4% (Fig 5.8C). This reduction persisted even with longer observation periods. A parallel decline in H3K27me3 levels was also noted. Using an alternative selection marker, mCherry, followed by FACS-based cell sorting, resulted in a similar reduction in SUZ12 expression over time (data not shown). This trend was consistent across all SUZ12 variants (except H466A which already displayed absent levels of SUZ12) with an increasing number of cells losing SUZ12 expression over time. Collectively, these findings show that the complementation of SUZ12 was more complex than anticipated. Potential explanations and solutions to address this challenge will be addressed in the discussion below.

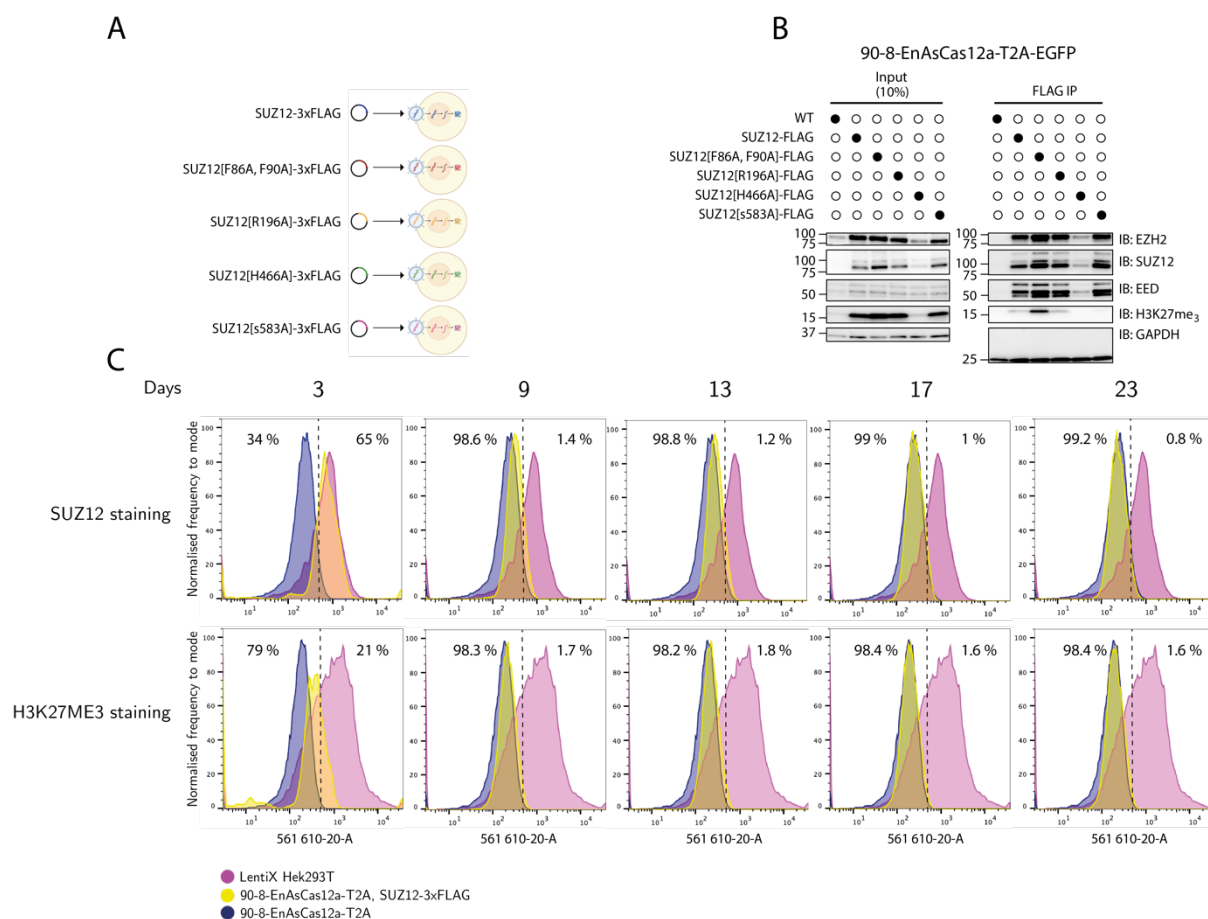


Figure 5.8. SUZ12 reconstitution restores PRC2 function but reduces over time. Schematic representation of the constructed SUZ12 variants and their lentiviral transduction into target cell lines. **(B)** Whole cell extracts of 90-8-enAsCas12a-T2A-EGFP with or without (WT) complemented SUZ12 variants represented in (A), were prepared from exponentially growing cells, solubilised with detergent, and subjected to immunoprecipitation with an anti-FLAG antibody. Captured proteins were visualised by Western blotting. EZH2, SUZ12, EED and H3K27ME3 proteins were detected with anti-EZH2, -SUZ12, -EED and -H3K27ME3, respectively. GAPDH, used as a loading control was detected with anti-GAPDH. IB: immunoblot. **(C)** Frozen pelleted LentiX Hek293T, 90-8-enAsCas12a-T2A-EGFP, 90-8-enAsCas12a-T2A-EGFP cells complemented with SUZ12-3xFLAG, from different time points, were fixed, permeabilised and immunostained with anti-SUZ12 or anti-H3K27ME3 as primary and Alexa 594nm goat anti-rabbit as secondary antibodies. Immunostaining was measured using flow cytometry.

5.3 Discussion

In this chapter, I investigated the effects of PRC2 degradation in iHSCs using a PROTAC, UNC7700, which targets the EED subunit, leading to the degradation of PRC2 components. Through transcriptomic analysis and CRISPR screens, I explored the

subsequent changes in gene expression and cellular dependencies associated with PRC2 loss.

Differential gene expression analysis confirmed that PRC2 degradation resulted in the de-repression of presumed PRC2 target genes, consistent with PRC2's established role as a transcriptional repressor. This finding substantiates the effectiveness of UNC7700 in functionally inhibiting PRC2 activity. GSEA further revealed a significant enrichment of the PI3K-AKT-mTOR signalling pathway among the differentially expressed genes in cells treated with UNC7700. This suggests a potential mechanistic link between PRC2 loss and enhanced mTOR signalling in Schwann cells and, perhaps, its relation to MPNST development. However, whether PRC2 loss independently augments both RAS and PI3K-AKT-mTOR pathways, or whether these pathways are interconnected, as has been proposed previously[237], warrants further investigation.

CRISPR screens comparing PRC2 wild-type and PRC2-deficient contexts revealed numerous potential differential genetic dependencies. Notably, in the Humagne C screens, PIK3R4 knockout was significantly more lethal in PRC2-disrupted iHSCs compared to DMSO-treated controls, indicating an increased reliance on PI3K-AKT-mTOR signalling in the absence of PRC2. Given the potential for noise in these library-specific findings and as these observations were only observed in one CRISPR screen, further high-throughput experimental validation is necessary before proceeding with downstream mechanistic validation. However, differential hits identified in both the Humagne C and D libraries, such as DNA topoisomerase I (*TOP1*), which was more lethal in iHSCs treated with UNC7700, are probably more robust. TOP1, involved in resolving DNA supercoils to facilitate transcription and replication, has been implicated in the treatment

of MYC-driven cancers. This may be relevant to MPNSTs, particularly since MYC inhibition has already been identified as essential across MPNST cell lines, that all lost PRC2 function, in CRISPR screens in this study (Chapter 4).

In these CRISPR screen experiments, I also examined the relationship between context-dependent differential gene expression and context-dependent essential genes. Consistent with the findings in Chapter 4, an inverse relationship was observed between differentially expressed genes and differential gene essentiality. This suggests that differentially expressed genes may not directly indicate which genes are essential in a given context, highlighting the complexity of this relationship beyond what can be inferred from transcriptomic analysis alone.

Finally, I aimed to explore the context-specific lethality of PRC2 by reintroducing *SUZ12* into MPNST cell line models. Based on prior studies of *SUZ12* function and its interactions within PRC2 and with other factors, five distinct *SUZ12* transgenes were introduced into MPNST cell lines via lentiviral delivery. Three of these variants involved mutations in key residues critical for *SUZ12* interactions with other PRC2 components, while a fourth variant targeted the C2H2 zinc finger domain, anticipated to disrupt *SUZ12*'s tertiary structure.

Following *SUZ12* complementation, a marked reduction in *SUZ12* expression was observed over time, resulting in a decline in the proportion of *SUZ12*-positive cells within the population. One potential explanation is that *SUZ12* re-expression reestablishes PRC2 function, leading to gene repression via H3K27 trimethylation, including possible silencing of the transgene itself. Alternatively, the reintroduction of exogenous *SUZ12* may lead to expression levels that are toxic to MPNST cell lines, which have adapted to its absence,

resulting in cell death upon re-expression. If this is the case, utilising an inducible expression system to titrate SUZ12 levels could mitigate this toxicity. Another possibility is gene silencing independent of PRC2 function. However, the rapid decline in SUZ12 expression, in contrast to the moderate decline in enAsCas12a-T2A-EGFP expression in the same cell lines (delivered via the same construct under the same promoter), suggests that this silencing would at least be partially related to PRC2 function or intrinsic to the *SUZ12* transgene itself.

Chapter 6: Concluding remarks and future perspectives

6.1 Concluding remarks and future prospects

Despite efforts from the scientific community, no systemic treatments have been proven to significantly improve the lives of MPNST patients to this day[238]. Conventional, focussed drug discovery approaches, that only target one pathway with a limited scope of targets, are inefficient or, worse, more likely to meet therapy resistance due to targeting protein/gene that can be compensated for easily by the cancer. This has sparked interest in targeting synthetic lethality gene pairs where one is lost during cancer development, usually a driver variant as this means more subclones are affected, and the other gene drugged, which then leads to cell death[239]. Depending on the type of alteration, a lost gene is unlikely to be regained by a cancer as part of selection pressure when its synthetic lethality partner is targeted with a drug, and so resistance to targeting true synthetic lethality is less likely[33].

The main aims of this thesis were to optimise an enAsCas12a-based, multiplexed, genome-wide CRISPR knockout library screen workflow to identify MPNST context-dependent and independent essential genes, including potential synthetic lethality relationships, and aid the discovery of novel therapeutic targets for MPNST patients. By assessing the variables that impact on genome-wide CRISPR screen results, I was able to provide a list of potential candidate gene hits that need further validation to evaluate whether they are targets that provide clinical benefit to MPNST patients. The main conclusions regarding the aims of this thesis are summarised below.

6.2 Optimise and validate enAsCas12a CRISPR screen workflow:

Based on the monitoring of expression levels of the enAsCas12a-T2A-EGFP transgene in target cell lines for this study, two cell lines, STS26T and s462TY, displayed an increasingly large negative enAsCas12a expression population over time (Chapter 3).

Generally, screens performed in these cell lines were also of substantial lower quality, as measured by the global CRISPR screen effect size as well as the recall of common essential genes. Although the lack of enAsCas12a expressing cells seems like a sound theoretical explanation for the observed lower quality, this may also be explained by unexplained intrinsic factors related to the target cell lines. Whether this is due to the reduced expression or not, may be challenging to extrapolate, where the data of this study only shows two such cell lines with both reduced enAsCas12a expression and lower screen quality, and the ultimate comparison would be with the same target cells but with a 100% of the population expressing enAsCas12a to see if the quality improves.

The duration of CRISPR screens also proved to be a significant factor that influences the global effect size and common essential gene recall as quality measures (Chapter 3). While extended screening periods can provide more time for genetic perturbations to manifest into differences in gRNA representation, they also introduce additional sources of noise, particularly in the form of gRNA dropout. In this study, screen samples that had a marked increase in undetected gRNAs were all isolated at time points that extended beyond the conventional 14 population doublings. And although an extended time-period strongly correlates with an increased effect size, this does not always correspond to a better common essential gene recall (Fig. 3.5, Chapter 3), suggesting that longer durations can also result in lower quality screens. Although computational tools such as Chronos can integrate data from multiple late time points and mitigate some of the issues associated with very late time points, for instance by correcting for outgrowths of specific gRNA effects, there remains a need to balance the trade-off between extended screening durations and the quality of the readout. Based on

the findings in this study I would propose that durations in CRISPR screens do not need to extend beyond 14 population doublings, especially if one does not correct for the associated consequences during data analysis.

The choice of CRISPR library was also found to heavily impacts the results of CRISPR screens (Chapter 3). In fact, across experimental variables in this study, from medium context to drug treatments, the choice of CRISPR library is the biggest determinant of outcome. This was also found in external datasets from DepMap, where three different CRISPR libraries were used across many overlapping cell line models. Yet outcomes across these experiments clustered together based on CRISPR library use, rather than anything else. Previous studies suggest that this is in part due to differences in gRNA efficacies[85], [117], and this was confirmed by a trend observed in data from this study where bigger differences in gene effects between libraries correlated significantly with bigger differences in gRNA efficacies. However, this correlation was only small ($R = 0.18$, Fig. 3.11A, Chapter 3), and was deemed unlikely to explain the strong differences in outcomes observed between different CRISPR libraries. Importantly, it would require validation, possibly with the use of NGS, to determine the true efficacy of gRNAs and their on- or off-target effects.

As CRISPR libraries are produced independently as much as they are designed differently, the representation of gRNAs in each CRISPR library is also dramatically different. Following the law of large numbers², I hypothesised that gRNAs with a lower

² Bernoulli, J. (1713). *Ars Conjectandi*. Basel: Thurneysen Brothers. Johann Bernoulli first described the law of large numbers in this work, laying the groundwork for probability theory by demonstrating that as the sample size increases, the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed. This principle underlies the observed greater variability in gRNAs with lower initial representation in the study.

initial representation would display larger gene effect sizes, and should be deemed noise or, at least, an exaggeration of a potentially realistic effect. In order to evaluate whether this was true, CRISPR screen data from this study was analysed by categorising gene effects across equally sized bins based on the initial representation of associated gRNAs to see whether gene effects from bins with gRNAs that had a lower initial representation displayed stronger gene effect scores compared to bins with gRNAs that had a higher initial representation (Chapter 3). The measured correlation of this trend was found to be very strong and highly significant. To see whether this was inherent to Humagne C and D, or whether other CRISPR library datasets display a similar relationship, data from DepMap was split based on CRISPR library and gene effect scores split across equally sized bins based on the average initial representation for gRNAs targeting the same gene. Again, lower initial representation of associated gRNAs strongly and highly significantly correlated with more dramatic gene effect scores. I then also confirmed that the initial gRNA representation strongly and highly significantly correlated with differences between CRISPR libraries and suggests to be of bigger importance than gRNA efficacy based on the statistics from the analyses of these data sets. Additionally, differences between replicate screens with the same library are in part the result of the effect of initial gRNA representation, where differences in gene effects correlate strongly and highly significantly with gRNAs with lower initial representation. Together, these analyses suggest that initial gRNA representation is responsible for differences observed between CRISPR libraries and even between replicate screens performed with the same library. Moreover, lower initial gRNA representation leads to more dramatic gene effect scores and so risks the

selection of presumed hit genes for downstream validation experiments based on false-positive observations.

In an attempt to correct for this effect, z-scores were calculated for gene effect scores per initial gRNA representation bin (Chapter 3). In this way, gene effect scores in lower pDNA bins are adjusted more strongly due to a bigger variance of scores in that group. Although this score adjustment did result in a significant improvement of similarity between Humagne C and D screens and between replicate screens, the effect was only marginal. Furthermore, the quality of the screens as measured by the recall of common essential genes was not improved. It was surprising to find that this correction did not improve the data more than anticipated, considering the strength of the relationship between lower initial gRNA representation and stronger gene effect scores, and so, to a certain extent, this issue remains unaddressed. More sophisticated modelling of this effect may be required to more substantially improve the data and perhaps such correction algorithms need to be incorporated into Chronos or other CRISPR screen analysis tools for optimal performance. Eventually, the only way to properly address this issue may be by preventing it. By producing a library that creates a more equal gRNA representation from the start, for instance by optimisation of oligo synthesis, PCR, and the temperature during cloning[144], this effect may not occur in the first place.

6.3 Cellular dependencies in different cell culture media

Many cancer models fail to accurately capture the complexity of the cancers they aim to represent, as demonstrated by the significant discrepancies observed between therapeutic effects in *in vitro* studies and clinical trials[145]. To bridge this gap, Human Plasma-Like Medium (HPLM) was developed to better mimic physiological conditions

compared to conventional cell culture media like DMEM. In this study, HPLM was used in comparison to standard DMEM to identify dependencies in MPNST cell line models that are medium-dependent or independent (Chapter 4). Although HPLM mimics plasma, it is still relevant to solid tumours, like MPNSTs, where leaky vascularisation results in a tumour microenvironment of which the basis is, in large, plasma [240].

Transcriptomic analysis of cells cultured in HPLM versus DMEM revealed that the majority of gene expression changes were specific to each cell line, indicating considerable variability in how different cell types respond to physiologically relevant culture conditions. However, GSEA uncovered limited but specific metabolic shifts that were consistent across all cell lines, particularly an increase in oxidative phosphorylation and the synthesis of unsaturated fatty acids in cells grown in HPLM. These findings suggest a metabolic reprogramming, where cells in HPLM become more reliant on mitochondrial energy production, possibly due to the physiologically relevant glucose concentrations in HPLM compared to the super-physiological levels present in DMEM.

While HPLM contains more physiologically relevant concentrations of nutritional compounds, this may not result in more accurate cancer models. To address whether this is the case, bulk RNA sequencing data from cell lines cultured in HPLM and DMEM were compared with RNA sequencing data from MPNST samples (Chapter 4). Spearman correlation analysis revealed that cells grown in DMEM had a marginally, but statistically significant, stronger correlation with MPNST tumour samples than those grown in HPLM. Although transcriptomics provides insight into cellular states, it is only one aspect of a cell's overall physiology, and likely does not fully capture the metabolic shifts that may occur with changes in culture medium. Therefore, transcriptomic comparisons may

not completely reflect the metabolic differences that arise from the use of HPLM versus standard media. To gain a more accurate understanding of whether HPLM improves the physiological relevance of cancer models, further research should incorporate direct metabolomic measurements, like mass spectrometry, rather than relying solely on RNA sequencing-based methods [210], [240].

The notion that transcriptomics provides a limited reflection of metabolomics, or the overall cellular phenotype, is further reinforced by the finding that medium-specific genetic dependencies inversely correlated with differentially expressed genes. Despite the presence of strong changes in gene expression, in some instances, these shifts did not enrich for essential genes. In fact, genes whose expression is tightly regulated were found to be enriched for medium-specific dependencies. A similar phenomenon, unrelated to switching culture medium, was observed during treatment with UNC7700, an EED-targeting PROTAC that induces degradation of PRC2, an epigenetic repressor complex with hundreds of target genes (Chapter 5). In this case, the differentially expressed genes, which were mostly upregulated, also inversely correlated with genetic dependencies specific to normal or disrupted PRC2 function. This experiment demonstrates that although the loss of PRC2 function leads to upregulation of specific genes, as expected, it does not necessarily imply that the cell becomes dependent on the increased expression of these genes. Rather, these gene expression changes are likely symptomatic of the altered cellular context, rather than integral to core cellular functions. These findings suggest that manipulating cell models to reflect a specific phenotype, whether through changes in cell culture media, epigenetic modulation, or even genetic editing, may be more challenging than often assumed, as these models did not naturally evolve to mirror

such conditions. However, theoretically, robust relationships, such as context-independent synthetic lethality, should still emerge when manipulating cell models to study these, regardless of their inherent metabolic, genetic, or epigenetic backgrounds.

Despite the general observation that changes in gene expression do not typically align with differential genetic dependencies, there were notable exceptions. For instance, elevated levels of oxidative phosphorylation in cells grown in HPLM, as identified by GSEA, were also associated with increased dependencies on these gene sets, as revealed by multiple genome-wide CRISPR screen experiments (Chapter 4). Although these findings require further mechanistic validation, they could be related to the lower glucose levels in HPLM that increases the cellular reliance on genes involved in maintaining energy homeostasis, such as those governing oxidative phosphorylation. This could indicate that the metabolic shift induced by HPLM fosters a greater dependency on mitochondrial energy production pathways under conditions of reduced glycolytic flux.

When examining the top 500 differential genetic dependencies independent of transcriptomic changes for each cell line, two genes, *PRELID3B* and *COPS6*, were found to be common across all four cell lines, with lower gene effect scores in HPLM (Chapter 4). As previously discussed in Chapter 4, both genes play roles in energy metabolism, suggesting that the alterations in energy metabolism observed in cells cultured in HPLM may induce novel genetic dependencies compared to those in DMEM. Notably, neither *PRELID3B* nor *COPS6* were part of the pathways enriched in HPLM according to GSEA, underscoring that gene sets identified through pathway analysis may not be all-encompassing.

6.4 Potential genetic dependencies in MPNST models

MPNSTs are aggressive tumours with limited treatment options, and there is an urgent need to uncover vulnerabilities that could be exploited for therapy. By targeting genes that are crucial for tumour cell survival but not for normal cell function, precision therapies can be developed with minimal off-target effects. One approach is the identification of MPNST-specific dependencies through genome-wide CRISPR screens. In this context, the identification of genetic dependencies across different contexts holds particular significance, as these genes would represent robust therapeutic targets across varying physiological environments.

A comprehensive approach was taken in this study to discern such targets, utilising both DMEM and HPLM media to simulate different medium-context conditions (Chapter 4). This medium-independent perspective aimed to reveal essential genes that maintain their critical role in MPNST survival regardless of environmental context. Importantly, the decision to include multiple MPNST cell lines, 90-8, ST-8814, and s462TY, alongside two control lines, iHSC (immortalised human Schwann cells) and STS26T (melanoma cells), ensured that the analysis could differentiate between cancer-specific dependencies and shared lineage traits. Notably, melanocytes and Schwann cells, share a neural crest origin, and these controls helped to refine the search for genes uniquely essential to MPNSTs.

Genes *SPTBN1*, *FBXO42*, and *CDK8* were found to be essential in both 90-8 and ST-8814 cell lines but non-essential in the control lines (Chapter 4). The identification of *SPTBN1* (spectrin β II) is particularly interesting due to its role in cytoskeletal architecture and its interaction with neurofibromin 2 (NF2), a key tumour suppressor in

Schwann cells. NF2 functions by linking the cytoskeleton to transmembrane proteins, mediating contact inhibition of cell proliferation. Loss of NF2 function is a well-established driver of unchecked Schwann cell proliferation, leading to tumorigenesis in the context of neurofibromatosis type 2[166]. Interestingly, despite its role in Schwann cell biology, *SPTBN1* was not found to be essential in non-cancerous Schwann cells, suggesting it represents a cancer-specific vulnerability[166].

FBXO42 and *CDK8*, may also represent interesting targets (Chapter 4). *FBXO42* is involved in protein ubiquitination, a process critical for maintaining cellular protein homeostasis, while *CDK8* is a cyclin-dependent kinase with well-documented roles in cancer progression through regulation of transcription and cell cycle transitions[171], [241], [172], [173]. Importantly, *CDK8* has already been investigated as a druggable target, with small-molecule inhibitors, such as Senexin A and B, showing promise in preclinical studies of gastrointestinal cancers, breast cancer, prostate cancer and melanoma[242]. Thus, while *SPTBN1* represents a relatively novel target, *FBXO42* and *CDK8* offer more established therapeutic avenues for potential MPNST treatment.

An important focus of this study was the identification of targets that can be rapidly translated into therapeutic strategies for MPNSTs. The FDA-approved drug list and the Druggable Genome dataset provide information on genes that are known to be targeted by existing drugs or predicted to be druggable based on structural and functional similarities to known targets, respectively. By integrating these datasets with results from CRISPR screens internally performed for this study, as well as with data from the DepMap database, a resource that offers extensive datasets, including gene dependency profiles across multiple cancer cell lines, including MPNST cell line models, it was possible to

further refine the list of potential MPNST-specific vulnerabilities (Chapter 4). This integration allowed for a more rigorous selection of targets based on MPNST cell line models beyond the experimental scope of this study alone.

This approach uncovered several druggable targets, which include AURKA, CDK1, CDK4, MTOR, CHEK1 and PLK1, all of which are currently being explored for cancer therapy. For example, inhibitors targeting Aurora kinases and CDKs are already under investigation in clinical trials for various cancers, including sarcomas (Chapter 4). The identification of these targets in MPNST presents an opportunity to repurpose existing therapies for this cancer. Among the remaining relevant Druggable Genome hits, KIF11, a key player in mitotic spindle formation, has been targeted by inhibitors like ispinesib, showing efficacy in several cancer types, including glioblastoma, meningioma and even MPNST, *in vitro* and *in vivo* [242], [243], [244]. One study showed that KIF11 inhibition using ispinesib or ARRY-520 significantly reduced MPNST cell viability, particularly in cell lines like S462 and ST88-14. These models also demonstrated enhanced sensitivity to combined targeting of KIF11 and BRD4 (using JQ1), resulting in a strong synergistic antitumour effect, making this combination a potential therapeutic strategy for MPNST. FNTA, another hit gene, is involved in protein prenylation, a process essential for the function of oncogenic proteins like RAS and could also be a valuable target for MPNST treatment, particularly with farnesyltransferase inhibitors such as tipifarnib[246]. PSMB1 and PSMB2, both part of the proteasome complex, have been indirectly targeted in other cancers through proteasome inhibitors like bortezomib, making them interesting candidates for validation in MPNST models as well[247], [248].

Another important aspect of this study involved leveraging synthetic lethality to refine and prioritise potential therapeutic targets. Synthetic lethality occurs when the simultaneous inactivation of two genes leads to cell death, whereas the inactivation of either gene alone is tolerated. By integrating synthetic lethality predictions from the SynLethDB 2.0 database with gene dependency data from CRISPR screens [102], it was possible to narrow down the list of candidate genes in an attempt to select those most likely to yield effective results in therapeutic settings (Chapter 4). This prioritisation was in part validated by observing that predicted synthetic lethality partners in MPNST cell lines were associated with significantly lower gene effect scores, indicating higher dependency in the tumour context. Importantly, there may be a bias in this where the predicted synthetic lethality partners from SynLethDB 2.0[103], which was used for these predictions, may be skewed to contain genes more generally essential in cancer, rather than specifically related to synthetic lethality.

Of particular note, *MYC* and *NAE1* emerged as high-priority candidates through this synthetic lethality approach in genome-wide CRISPR knockout screens in MPNST cell line models from this study as well as external datasets present in DepMap (Chapter 4). *MYC*, a transcription factor that regulates genes involved in cell proliferation, is a well-known oncogene implicated in numerous cancers, including MPNSTs[178]. Although direct targeting of *MYC* is challenging due to its "undruggable" molecular nature, recent advances, such as the development of the indirect *MYC* inhibitor OMO-103, offer promising new avenues for therapeutic intervention[179], [180]. Similarly, *NAE1*, involved in protein neddylation, has gained attention for its role in cancer biology, with its overexpression linked to tumour progression[182], [249], [250].

6.5 PRC2 loss associated synthetic lethality in MPNST

Polycomb Repressive Complex 2 (PRC2) loss is a defining feature of many MPNSTs, leading to the de-repression of PRC2 target genes and promoting tumorigenesis. Given the challenge of directly targeting PRC2 loss therapeutically, synthetic lethality approaches provide an alternative strategy by identifying gene dependencies that arise specifically in PRC2-deficient cells. The use of the EED-targeting PROTAC, UNC7700, enabled efficient degradation of PRC2 components, offering a platform to investigate context-specific vulnerabilities in iHSC cells.

Through transcriptomic analysis, PRC2 disruption was found to enhance the activation of the PI3K-AKT-mTOR signalling pathway, which has implications for cell proliferation and survival in Schwann cells (Chapter 5). This observation aligns with earlier findings of mTOR involvement in MPNST biology [214], [223], suggesting a potential avenue for therapeutic intervention. Further studies are needed to explore the potential mechanistic relationship between PRC2 loss and increased mTOR signalling.

Genome-wide CRISPR screens in this context identified several potential PRC2-dependent synthetic lethality targets (Chapter 5). Among these, TOP1 (topoisomerase 1) emerged as a promising candidate, demonstrating increased lethality in PRC2-deficient cells in both Humagne C and D screens. Given the established role of TOP1 in DNA replication and transcription, and the availability of clinical inhibitors like irinotecan, further validation of TOP1 inhibition in MPNST models seems warranted. Additionally, potential hits from Humagne C and D screens were also presented separately, although these will require additional biological replicates to confirm their robustness.

6.6 Future Work

The next steps in validating the identified target genes will involve a combination of both genetic and pharmacological methods. To directly assess the essentiality of these genes, inducible knockout systems, such as CRISPR/Cas9 or other gene-editing technologies, can be employed to selectively disrupt these target genes in MPNST models. In parallel to genetic validation, pharmacological inhibition using small-molecule inhibitors, or existing FDA-approved drugs where applicable, will be crucial to evaluate if the therapeutic effects seen with genetic knockout can be mirrored by drug treatment. Specifically, the validation of predicted synthetic lethality-related hit genes can be further investigated by selective knockout or complementation of synthetic lethality partner genes in addition to disruption of the selected hit gene.

For *in vivo* validation, both genetically engineered mouse models of MPNST, like the previously described *NF1*^{+/-} *TP53*^{+/-} *SUZ12*^{+/-} mouse model [49], or xenograft models may be instrumental in assessing the impact of target inhibition on tumour growth and progression. These animal models provide a more relevant physiological context, enabling the evaluation of both genetic and pharmacological inhibition in a more physiological tumour microenvironment.

By integrating these genetic, pharmacological, and *in vivo* approaches, future studies will move towards translating these findings into potential clinical applications, hopefully offering a wider panel of treatment options for MPNST patients.

To conclude, in this study I optimised an enAsCas12a-based genome-wide CRISPR screen workflow, with a focus on reducing false positive hit selection, in order to identify novel dependencies in MPNST. First, an unbiased approach was taken where cellular

dependencies were identified in these screens without a manipulated genetic context, allowing the uncovering of targets in a panel of genetically diverse MPNST models. Data from these screens was integrated with CRISPR screen data from DepMap and hit genes prioritised based on druggability or predicted synthetic lethality, and eventually provided a list of known as well as novel candidate genes that warrant further investigation. Second, this optimised workflow was used to identify potential synthetic lethality relationships involving PRC2 loss by manipulating PRC2 context with an EED-targeting PROTAC. These candidate genes may suggest PRC2 context-specific dependencies that may involve effective therapeutic targets for patients with MPNST.

Supplemental information

Supplementary table 7.1. GSEA in individual cell lines: HPLM vs. DMEM.

Cell line	Gene set collection	Gene set	enriched in	FDR q value (<25%)
STS2 6T	Hallmark	None		
	KEGG Legacy	Kegg_biosynthesis_of_unsaturated_fatty_acids	DMEM	0.238
	Reactome	Reactome_regulation_of_cholesterol_biosynthesis_by_srebp_srebf	DMEM	0.160
90-8	Hallmark	Hallmark_tnfa_signaling_via_nfkb	DMEM	0.181
	KEGG Legacy	Kegg_biosynthesis_of_unsaturated_fatty_acids	DMEM	0.238
	Reactome	Reactome_regulation_of_cholesterol_biosynthesis_by_srebp_srebf	DMEM	0.23
ST-8814	Hallmark	Hallmark_myc_targets_v2	DMEM	0.177
		Hallmark_mtorc1_signaling	DMEM	0.177
		Hallmark_oxidative_phosphorylation	DMEM	0.177
		Hallmark_cholesterol_homeostasis	DMEM	0.177
	KEGG Legacy	Kegg_ribosome	HPLM	0.155
		Kegg_ether_lipid_metabolism	DMEM	0.207
		Kegg_steroid_biosynthesis	DMEM	0.181
		Kegg_propanoate_metabolism	DMEM	0.172
		Kegg_biosynthesis_of_unsaturated_fatty_acids	DMEM	0.168
		Kegg_systemic_lupus_erythematosus	DMEM	0.165
		Kegg_citrate_cycle_tca_cycle	DMEM	0.164
		Kegg_butanoate_metabolism	DMEM	0.162
		Kegg_pyruvate_metabolism	DMEM	0.162
		Kegg_glutathione_metabolism	DMEM	0.161
		Kegg_amyotrophic_lateral_sclerosis_als	DMEM	0.171
		Kegg_glycolysis_gluconeogenesis	DMEM	0.169
		Kegg_galactose_metabolism	DMEM	0.186
		Kegg_notch_signaling_pathway	DMEM	0.183
		Kegg_porphyrin_and_chlorophyll_metabolism	DMEM	0.181
		Kegg_one_carbon_pool_by_folate	DMEM	0.183
		Kegg_pentose_phosphate_pathway	DMEM	0.181
		Kegg_glycosaminoglycan_degradation	DMEM	0.219
	Reactome	Reactome_cholesterol_biosynthesis	DMEM	0.178
		Reactome_b_wich_complex_positively_regulates_rna_expression	DMEM	0.178
		Reactome_formation_of_the_beta_catenin_tcf_transactivating_complex	DMEM	0.178
		Reactome_transcriptional_regulation_by_small_rnas	DMEM	0.178
		Reactome_maternal_to_zygotic_transition_mzt	DMEM	0.178

		Reactome_runx1_regulates_transcription_of_genes_involved_in_differentiation_of_hscs	DMEM	0.178
		Reactome_hcmv_late_events	DMEM	0.178
		Reactome_rna_polymerase_i_promoter_escape	DMEM	0.178
		Reactome_ercc6_csb_and_ehmt2_g9a_positively_regulate_rna_expression	DMEM	0.178
		Reactome_hdacs_deacetylate_histones	DMEM	0.178
		Reactome_inhibition_of_dna_recombination_at_telomere	DMEM	0.178
		Reactome_activation_of_gene_expression_by_srebf_srebp	DMEM	0.178
		Reactome_base_excision_repair	DMEM	0.178
		Reactome_positive_epigenetic_regulation_of_rna_expression	DMEM	0.178
		Reactome_activation_of_anterior_hox_genes_in_hindbrain_development_during_early_embryogenesis	DMEM	0.183
		Reactome_dna_damage_telomere_stress_induced_senescence	DMEM	0.183
		Reactome_condensation_of_prophase_chromosomes	DMEM	0.182
		Reactome_nr1h2_and_nr1h3_mediated_signaling	DMEM	0.182
		Reactome_prc2_methylates_histones_and_dna	DMEM	0.182
		Reactome_activation_of_smo	DMEM	0.185
		Reactome_rmts_methylate_histone_arginines	DMEM	0.185
		Reactome_oxidative_stress_induced_senescence	DMEM	0.185
		Reactome_gene_silencing_by_rna	DMEM	0.184
		Reactome_pre_notch_expression_and_processing	DMEM	0.187
		Reactome_diseases_of_programmed_cell_death	DMEM	0.19
		Reactome_transcriptional_regulation_of_granulopoiesis	DMEM	0.189
		Reactome_sirt1_negatively_regulates_rna_expression	DMEM	0.189
		Reactome_assembly_of_the_orc_complex_at_the_origin_of_replication	DMEM	0.188
		Reactome_hdms_demethylate_histones	DMEM	0.188
		Reactome_formation_of_senescence_associated_heterochromatin_foci_sahf	DMEM	0.188
		Reactome_hcmv_early_events	DMEM	0.187
		Reactome_met_promotes_cell_motility	DMEM	0.187
		Reactome_metal_ion_slc_transporters	DMEM	0.187
		Reactome_recognition_and_association_of_dna_glycosylase_with_site_containing_an_affected_purine	DMEM	0.187
		Reactome_regulation_of_cholesterol_biosynthesis_by_srebp_srebf	DMEM	0.186
		Reactome_negative_epigenetic_regulation_of_rna_expression	DMEM	0.186
		Reactome_hats_acetylate_histones	DMEM	0.186

		Reactome_dna_methylation	DMEM	0.186
		Reactome_base_excision_repair_ap_site_formati on	DMEM	0.185
		Reactome_replacement_of_protamines_by_nucle osomes_in_the_male_pronucleus	DMEM	0.185
		Reactome_activated_pkn1_stimulates_transcripti on_of_ar_androgen_receptor_regulated_genes_ klk2_and_klk3	DMEM	0.185
		Reactome_rna_polymerase_i_transcription	DMEM	0.188
		Reactome_chromatin_modifications_during_the_ maternal_to_zygotic_transition_mzt	DMEM	0.188
		Reactome_meiotic_recombination	DMEM	0.188
		Reactome_met_activates_ptk2_signaling	DMEM	0.189
		Reactome_nr1h3_nr1h2_regulate_gene_expressi on_linked_to_cholesterol_transport_and_efflux	DMEM	0.189
		Reactome_hcmv_infection	DMEM	0.189
		Reactome_nonhomologous_end_joining_nhej	DMEM	0.189
		Reactome_dna_double_strand_break_response	DMEM	0.188
		Reactome_estrogen_dependent_gene_expression	DMEM	0.188
		Reactome_ire1alpha_activates_chaperones	DMEM	0.189
		Reactome_class_i_peroxisomal_membrane_prote in_import	DMEM	0.19
		Reactome_regulation_of_pten_gene_transcriptio n	DMEM	0.191
		Reactome_protein_localization	DMEM	0.191
		Reactome_citric_acid_cycle_tca_cycle	DMEM	0.191
		Reactome_runx1_regulates_genes_involved_in_ megakaryocyte_differentiation_and_platelet_func tion	DMEM	0.193
		Reactome_constitutive_signaling_by_akt1_e17k_ in_cancer	DMEM	0.194
		Reactome_metabolism_of_steroids	DMEM	0.195
		Reactome_amyloid_fiber_formation	DMEM	0.195
		Reactome_integrin_signaling	DMEM	0.195
		Reactome_amino_acid_transport_across_the_pl asma_membrane	DMEM	0.197
		Reactome_mtorc1_mediated_signalling	DMEM	0.203
		Reactome_interleukin_7_signaling	DMEM	0.204
		Reactome_branched_chain_amino_acid_cataboli sm	DMEM	0.203
		Reactome_processing_of_dna_double_strand_br eak_ends	DMEM	0.203
		Reactome_dna_replication_pre_initiation	DMEM	0.206
		Reactome_epigenetic_regulation_of_gene_expres sion	DMEM	0.205
		Reactome_binding_and_uptake_of_ ligands_by_scavenger_receptors	DMEM	0.206
		Reactome_diseases_associated_with_n_glycosyla tion_of_proteins	DMEM	0.205

		Reactome_sumoylation_of_chromatin_organization_proteins	DMEM	0.206
		Reactome_mecp2_regulates_neuronal_receptors_and_channels	DMEM	0.207
		Reactome_mismatch_repair	DMEM	0.207
		Reactome_nicotinamide_salvaging	DMEM	0.212
		Reactome_mitotic_prophase	DMEM	0.212
		Reactome_the_phototransduction_cascade	DMEM	0.212
		Reactome_transcriptional_regulation_by_runx1	DMEM	0.212
		Reactome_signaling_by_met	DMEM	0.213
		Reactome_deposition_of_new_cenpa_containing_nucleosomes_at_the_centromere	DMEM	0.214
		Reactome_export_of_viral_ribonucleoproteins_from_nucleus	DMEM	0.214
		Reactome_transcriptional_regulation_of_white_adipocyte_differentiation	DMEM	0.214
		Reactome_rho_gtpases_activate_pkns	DMEM	0.215
		Reactome_plasma_lipoprotein_clearance	DMEM	0.216
		Reactome_rna_polymerase_iii_transcription_termination	DMEM	0.217
		Reactome_fatty_acid_metabolism	DMEM	0.224
		Reactome_runx2_regulates_osteoblast_differentiation	DMEM	0.224
		Reactome_mitochondrial_trna_aminoacylation	DMEM	0.234
		Reactome_esr_mediated_signaling	DMEM	0.235
		Reactome_syndecan_interactions	DMEM	0.236
		Reactome_signaling_by_notch	DMEM	0.236
		Reactome_g2_m_dna_damage_checkpoint	DMEM	0.236
		Reactome_diseases_of_carbohydrate_metabolism	DMEM	0.236
		Reactome_notch3_intracellular_domain_regulates_transcription	DMEM	0.237
		Reactome_non_integrin_membrane_ecm_interactions	DMEM	0.238
		Reactome_peroxisomal_lipid_metabolism	DMEM	0.245
		Reactome_meiotic_synapsis	DMEM	0.245
		Reactome_signaling_by_nuclear_receptors	DMEM	0.244
		Reactome_transcriptional_regulation_by_mecp2	DMEM	0.244
iHSC	Hallmark	Hallmark_unfolded_protein_response	DMEM	0.141
		Hallmark_adipogenesis	DMEM	0.141
	KEGG Legacy	Kegg_lysozyme	DMEM	0.173
		Kegg_acute_myeloid_leukemia	DMEM	0.173
		Kegg_porphyrin_and_chlorophyll_metabolism	DMEM	0.173
		Kegg_prostate_cancer	DMEM	0.21
		Kegg_pyruvate_metabolism	DMEM	0.203
	Reactome	None		

Bibliography

- [1] F. M. Bashore *et al.*, "PROTAC Linkerology Leads to an Optimized Bivalent Chemical Degradator of Polycomb Repressive Complex 2 (PRC2) Components," *ACS Chem Biol*, vol. 18, no. 3, pp. 494–507, Mar. 2023, doi: 10.1021/ACSCHEMBIO.2C00804/SUPPL_FILE/CB2C00804_SI_003.XLSX.
- [2] F. Bray *et al.*, "Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA Cancer J Clin*, vol. 74, no. 3, pp. 229–263, May 2024, doi: 10.3322/CAAC.21834.
- [3] "Home | Cancer Grand Challenges." Accessed: Oct. 01, 2024. [Online]. Available: <https://www.cancergrandchallenges.org/>
- [4] "Cancer MoonshotSM - NCI." Accessed: Oct. 01, 2024. [Online]. Available: <https://www.cancer.gov/research/key-initiatives/moonshot-cancer-initiative>
- [5] S. Turajlic, A. Sottoriva, T. Graham, and C. Swanton, "Resolving genetic heterogeneity in cancer," Jul. 01, 2019, *Nature Publishing Group*. doi: 10.1038/s41576-019-0114-6.
- [6] A. Davis, R. Gao, and N. Navin, "Tumor evolution: Linear, branching, neutral or punctuated?," *Biochim Biophys Acta Rev Cancer*, vol. 1867, no. 2, pp. 151–161, Apr. 2017, doi: 10.1016/j.bbcan.2017.01.003.
- [7] S. F. Bakhoun and L. C. Cantley, "The Multifaceted Role of Chromosomal Instability in Cancer and Its Microenvironment," Sep. 06, 2018, *Cell Press*. doi: 10.1016/j.cell.2018.08.027.
- [8] S. M. Swain, M. Shastry, and E. Hamilton, "Targeting HER2-positive breast cancer: advances and future directions," *Nature Reviews Drug Discovery* 2022 22:2, vol. 22, no. 2, pp. 101–126, Nov. 2022, doi: 10.1038/s41573-022-00579-0.
- [9] R. S. Iyer *et al.*, "Drug-resistant EGFR mutations promote lung cancer by stabilizing interfaces in ligand-free kinase-active EGFR oligomers," *Nature Communications* 2024 15:1, vol. 15, no. 1, pp. 1–21, Mar. 2024, doi: 10.1038/s41467-024-46284-x.
- [10] A. Marusyk, M. Janiszewska, and K. Polyak, "Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance," Apr. 13, 2020, *Cell Press*. doi: 10.1016/j.ccell.2020.03.007.
- [11] A. Pérez-González, K. Bévant, and C. Blanpain, "Cancer cell plasticity during tumor progression, metastasis and response to therapy," *Nature Cancer* 2023 4:8, vol. 4, no. 8, pp. 1063–1082, Aug. 2023, doi: 10.1038/s43018-023-00595-y.
- [12] L. Yang *et al.*, "Targeting cancer stem cell pathways for cancer therapy," Dec. 01, 2020, *Springer Nature*. doi: 10.1038/s41392-020-0110-5.
- [13] X. Chu *et al.*, "Cancer stem cells: advances in knowledge and implications for cancer therapy," *Signal Transduct Target Ther*, vol. 9, no. 1, Dec. 2024, doi: 10.1038/S41392-024-01851-Y.
- [14] C. Robert, "A decade of immune-checkpoint inhibitors in cancer therapy," *Nature Communications* 2020 11:1, vol. 11, no. 1, pp. 1–3, Jul. 2020, doi: 10.1038/s41467-020-17670-y.

- [15] S. Qin, L. Xu, M. Yi, S. Yu, K. Wu, and S. Luo, "Novel immune checkpoint targets: Moving beyond PD-1 and CTLA-4," Nov. 06, 2019, *BioMed Central Ltd.* doi: 10.1186/s12943-019-1091-2.
- [16] I. Vitale, E. Shema, S. Loi, and L. Galluzzi, "Intratumoral heterogeneity in cancer progression and response to immunotherapy," *Nat Med*, vol. 27, no. 2, pp. 212–224, Feb. 2021, doi: 10.1038/s41591-021-01233-9.
- [17] M. J. Topper, M. Vaz, K. A. Marrone, J. R. Brahmer, and S. B. Baylin, "The emerging role of epigenetic therapeutics in immuno-oncology," Feb. 01, 2020, *Nature Research*. doi: 10.1038/s41571-019-0266-5.
- [18] R. C. Sterner and R. M. Sterner, "CAR-T cell therapy: current limitations and potential strategies," *Blood Cancer Journal* 2021 11:4, vol. 11, no. 4, pp. 1–11, Apr. 2021, doi: 10.1038/s41408-021-00459-7.
- [19] S. P. Foy *et al.*, "Non-viral precision T cell receptor replacement for personalized cell therapy," *Nature* 2022 615:7953, vol. 615, no. 7953, pp. 687–696, Nov. 2022, doi: 10.1038/s41586-022-05531-1.
- [20] M. Narkhede, A. Mehta, S. M. Ansell, and G. Goyal, "CAR T-cell therapy in mature lymphoid malignancies: clinical opportunities and challenges," *Ann Transl Med*, vol. 9, no. 12, pp. 1036–1036, Jun. 2021, doi: 10.21037/ATM-20-5546.
- [21] A. D. Waldman, J. M. Fritz, and M. J. Lenardo, "A guide to cancer immunotherapy: from T cell basic science to clinical practice," Nov. 01, 2020, *Nature Research*. doi: 10.1038/s41577-020-0306-5.
- [22] J. Kleponis, R. Skelton, and L. Zheng, "Fueling the engine and releasing the break: combinational therapy of cancer vaccines and immune checkpoint inhibitors," Sep. 01, 2015, *Cancer Biology and Medicine*. doi: 10.7497/j.issn.2095-3941.2015.0046.
- [23] U. Sahin *et al.*, "An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma," *Nature*, vol. 585, no. 7823, pp. 107–112, Sep. 2020, doi: 10.1038/s41586-020-2537-9.
- [24] A. M. D'Alise *et al.*, "Adenoviral vaccine targeting multiple neoantigens as strategy to eradicate large tumors combined with checkpoint blockade," *Nature Communications* 2019 10:1, vol. 10, no. 1, pp. 1–12, Jun. 2019, doi: 10.1038/s41467-019-10594-2.
- [25] H. E. Bryant *et al.*, "Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase," *Nature* 2005 434:7035, vol. 434, no. 7035, pp. 913–917, Apr. 2005, doi: 10.1038/nature03443.
- [26] H. Farmer *et al.*, "Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy," *Nature* 2005 434:7035, vol. 434, no. 7035, pp. 917–921, Apr. 2005, doi: 10.1038/nature03445.
- [27] P. C. Fong *et al.*, "Inhibition of Poly(ADP-Ribose) Polymerase in Tumors from BRCA Mutation Carriers," *New England Journal of Medicine*, vol. 361, no. 2, pp. 123–134, Jul. 2009, doi: 10.1056/NEJMOA0900212/SUPPL_FILE/NEJM_FONG_123SA1.PDF.
- [28] B. Kaufman *et al.*, "Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation," *Journal of Clinical Oncology*, vol. 33, no. 3, pp. 244–250, Jan. 2015, doi: 10.1200/JCO.2014.56.2728/ASSET/IMAGES/ZLJ00315-4867-T0A1.JPEG.

- [29] K. A. Gelmon *et al.*, "Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: A phase 2, multicentre, open-label, non-randomised study," *Lancet Oncol*, vol. 12, no. 9, pp. 852–861, Sep. 2011, doi: 10.1016/S1470-2045(11)70214-5.
- [30] K. C. Helming *et al.*, "ARID1B is a specific vulnerability in ARID1A-mutant cancers," *Nature Medicine* 2014 20:3, vol. 20, no. 3, pp. 251–254, Feb. 2014, doi: 10.1038/nm.3480.
- [31] M. S. Kumar *et al.*, "The GATA2 Transcriptional Network Is Requisite for RAS Oncogene-Driven Non-Small Cell Lung Cancer," *Cell*, vol. 149, no. 3, pp. 642–655, Apr. 2012, doi: 10.1016/J.CELL.2012.02.059.
- [32] S. Li *et al.*, "Development of synthetic lethality in cancer: molecular and cellular classification," *Signal Transduction and Targeted Therapy* 2020 5:1, vol. 5, no. 1, pp. 1–14, Oct. 2020, doi: 10.1038/s41392-020-00358-6.
- [33] X. Li and L. Zou, "BRCAness, DNA gaps, and gain and loss of PARP inhibitor–induced synthetic lethality," *J Clin Invest*, vol. 134, no. 14, Jul. 2024, doi: 10.1172/JCI181062.
- [34] B. S. Ducatman, B. W. Scheithauer, D. G. Piepgras, H. M. Reiman, and D. M. Ilstrup, "Malignant Peripheral Nerve Sheath Tumors A Clinicopathologic Study of 120 Cases", doi: 10.1002/1097-0142.
- [35] C. C. H. Stucky *et al.*, "Malignant Peripheral Nerve Sheath Tumors (MPNST): The Mayo Clinic experience," *Ann Surg Oncol*, vol. 19, no. 3, pp. 878–885, Mar. 2012, doi: 10.1245/S10434-011-1978-7/FIGURES/3.
- [36] K. Jett and J. M. Friedman, "Clinical and genetic aspects of neurofibromatosis 1," *Genetics in Medicine* 2010 12:1, vol. 12, no. 1, pp. 1–11, Dec. 2009, doi: 10.1097/gim.0b013e3181bf15e3.
- [37] M. M. Miettinen *et al.*, "Histopathologic evaluation of atypical neurofibromatous tumors and their transformation into malignant peripheral nerve sheath tumor in patients with neurofibromatosis 1—a consensus overview," *Hum Pathol*, vol. 67, pp. 1–10, Sep. 2017, doi: 10.1016/j.humpath.2017.05.010.
- [38] K. Staser, F. C. Yang, and D. W. Clapp, "Pathogenesis of plexiform neurofibroma: Tumor-stromal/hematopoietic interactions in tumor progression," 2012, *NIH Public Access*. doi: 10.1146/annurev-pathol-011811-132441.
- [39] D. E. Porter, V. Prasad, L. Foster, G. F. Dall, R. Birch, and R. J. Grimer, "Survival in Malignant Peripheral Nerve Sheath Tumours: A Comparison between Sporadic and Neurofibromatosis Type 1-Associated Tumours," *Sarcoma*, vol. 2009, 2009, doi: 10.1155/2009/756395.
- [40] I. Cortes-Ciriano *et al.*, "Genomic Patterns of Malignant Peripheral Nerve Sheath Tumor (MPNST) Evolution Correlate with Clinical Outcome and Are Detectable in Cell-Free DNA," *Cancer Discov*, vol. 13, no. 3, pp. 654–671, Mar. 2023, doi: 10.1158/2159-8290.CD-22-0786/712001/AM/GENOMIC-PATTERNS-OF-MALIGNANT-PERIPHERAL-NERVE.
- [41] R. M. Verdijk, M. A. Den Bakker, H. J. Dubbink, W. C. J. Hop, W. N. M. Dinjens, and J. M. Kros, "TP53 Mutation Analysis of Malignant Peripheral Nerve Sheath Tumors," *J Neuropathol Exp Neurol*, vol. 69, no. 1, pp. 16–26, Jan. 2010, doi: 10.1097/NEN.0B013E3181C55D55.

- [42] X. Zhang, B. Murray, G. Mo, and J. F. Shern, "The role of polycomb repressive complex in malignant peripheral nerve sheath tumor," Mar. 01, 2020, *MDPI AG*. doi: 10.3390/genes11030287.
- [43] A. Baude, A. M. Lindroth, and C. Plass, "PRC2 loss amplifies Ras signaling in cancer," *Nature Genetics* 2014 46:11, vol. 46, no. 11, pp. 1154–1155, Oct. 2014, doi: 10.1038/ng.3124.
- [44] T. De Raedt *et al.*, "PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies," *Nature* 2014 514:7521, vol. 514, no. 7521, pp. 247–251, Aug. 2014, doi: 10.1038/nature13561.
- [45] V. Kasinath *et al.*, "Structures of human PRC2 with its cofactors AEBP2 and JARID2," *Science (1979)*, vol. 359, no. 6378, pp. 940–944, Feb. 2018, doi: 10.1126/SCIENCE.AAR5700/SUPPL_FILE/AAR5700S1.MOV.
- [46] W. Lee *et al.*, "PRC2 is recurrently inactivated through EED or SUZ12 loss in malignant peripheral nerve sheath tumors," *Nat Genet*, vol. 46, no. 11, pp. 1227–1232, Nov. 2014, doi: 10.1038/ng.3095.
- [47] H. Kehrer-Sawatzki, U. Bätzner, J. Krämer, J. Lewerenz, and C. Pfeiffer, "The NF1 microdeletion syndrome: early genetic diagnosis facilitates the management of a clinically defined disease," *JDDG: Journal der Deutschen Dermatologischen Gesellschaft*, vol. 20, no. 3, pp. 273–277, Mar. 2022, doi: 10.1111/DDG.14707.
- [48] M. Wassef *et al.*, "EZH1/2 function mostly within canonical PRC2 and exhibit proliferation-dependent redundancy that shapes mutational signatures in cancer," *Proc Natl Acad Sci U S A*, vol. 116, no. 13, pp. 6075–6080, Mar. 2019, doi: 10.1073/PNAS.1814634116/SUPPL_FILE/PNAS.1814634116.SD01.XLSX.
- [49] Q. R. Brockman *et al.*, "PRC2 loss drives MPNST metastasis and matrix remodeling," *JCI Insight*, vol. 7, no. 20, Oct. 2022, doi: 10.1172/JCI.INSIGHT.157502.
- [50] T. De Raedt *et al.*, "PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies," *Nature* 2014 514:7521, vol. 514, no. 7521, pp. 247–251, Aug. 2014, doi: 10.1038/nature13561.
- [51] S. Suppiah *et al.*, "Multiplatform molecular profiling uncovers two subgroups of malignant peripheral nerve sheath tumors with distinct therapeutic vulnerabilities," *Nature Communications* 2023 14:1, vol. 14, no. 1, pp. 1–15, May 2023, doi: 10.1038/s41467-023-38432-6.
- [52] C. Yao *et al.*, "Malignant Peripheral Nerve Sheath Tumors: Latest Concepts in Disease Pathogenesis and Clinical Management," *Cancers (Basel)*, vol. 15, no. 4, Feb. 2023, doi: 10.3390/CANCERS15041077.
- [53] Z. Cai, X. Tang, H. Liang, R. Yang, T. Yan, and W. Guo, "Prognosis and risk factors for malignant peripheral nerve sheath tumor: a systematic review and meta-analysis," *World J Surg Oncol*, vol. 18, no. 1, Dec. 2020, doi: 10.1186/s12957-020-02036-x.
- [54] J. E. Bates, C. R. Peterson, S. Dhakal, E. J. Giampoli, and L. S. Constance, "Malignant peripheral nerve sheath tumors (MPNST): A SEER analysis of incidence across the age spectrum and therapeutic interventions in the pediatric population," *Pediatr Blood Cancer*, vol. 61, no. 11, pp. 1955–1960, Nov. 2014, doi: 10.1002/PBC.25149.

- [55] C. S. Higham *et al.*, "SARC006: Phase II Trial of Chemotherapy in Sporadic and Neurofibromatosis Type 1 Associated Chemotherapy-Naive Malignant Peripheral Nerve Sheath Tumors," *Sarcoma*, vol. 2017, 2017, doi: 10.1155/2017/8685638.
- [56] D. Wang *et al.*, "Significant Reduction of Late Toxicities in Patients With Extremity Sarcoma Treated With Image-Guided Radiation Therapy to a Reduced Target Volume: Results of Radiation Therapy Oncology Group RTOG-0630 Trial," *Journal of Clinical Oncology*, vol. 33, no. 20, p. 2231, Jul. 2015, doi: 10.1200/JCO.2014.58.5828.
- [57] L. Sloan, S. A. Terezakis, J. O. Blakeley, B. Slobogean, and L. R. Kleinberg, "Long-Term Outcomes of Radiation Therapy (RT) in the Management of Malignant Peripheral Nerve Sheath Tumors (MPNST) in Patients with Neurofibromatosis Type 1 (NF1)," *International Journal of Radiation Oncology*Biophysics*, vol. 102, no. 3, pp. e474–e475, Nov. 2018, doi: 10.1016/j.ijrobp.2018.07.1357.
- [58] J. Kahn *et al.*, "Radiation Therapy in Management of Sporadic and Neurofibromatosis Type 1-Associated Malignant Peripheral Nerve Sheath Tumors," *Front Oncol*, vol. 4, no. NOV, 2014, doi: 10.3389/FONC.2014.00324.
- [59] A. Kim *et al.*, "Targeting Refractory Sarcomas and Malignant Peripheral Nerve Sheath Tumors in a Phase I/II Study of Sirolimus in Combination with Ganetespib (SARC023)," *Sarcoma*, vol. 2020, 2020, doi: 10.1155/2020/5784876.
- [60] K. H. Albritton *et al.*, "Phase II study of erlotinib in metastatic or unresectable malignant peripheral nerve sheath tumors (MPNST).," *Journal of Clinical Oncology*, vol. 24, no. 18_suppl, pp. 9518–9518, Jun. 2006, doi: 10.1200/JCO.2006.24.18_SUPPL.9518.
- [61] "Study Results | Combination Chemotherapy in Treating Patients With Stage III or Stage IV Malignant Peripheral Nerve Sheath Tumors | ClinicalTrials.gov." Accessed: Sep. 29, 2024. [Online]. Available: <https://clinicaltrials.gov/study/NCT00304083?cond=Malignant%20Peripheral%20Nerve%20Sheath%20Tumor&term=Malignant%20Peripheral%20Nerve%20Sheath%20Tumors&intr=Drug&aggFilters=phase:2,results:with,status:com&rank=4&tab=results>
- [62] "Study Results | Lorvotuzumab Mertansine in Treating Younger Patients With Relapsed or Refractory Wilms Tumor, Rhabdomyosarcoma, Neuroblastoma, Pleuropulmonary Blastoma, Malignant Peripheral Nerve Sheath Tumor, or Synovial Sarcoma | ClinicalTrials.gov." Accessed: Sep. 29, 2024. [Online]. Available: <https://clinicaltrials.gov/study/NCT02452554?cond=Malignant%20Peripheral%20Nerve%20Sheath%20Tumor&term=Malignant%20Peripheral%20Nerve%20Sheath%20Tumors&intr=Drug&aggFilters=phase:2,results:with,status:com&rank=5&tab=results#outcome-measures>
- [63] M. A. Dickson *et al.*, "Phase II study of MLN8237 (Alisertib) in advanced/metastatic sarcoma," *Annals of Oncology*, vol. 27, no. 10, pp. 1855–1860, Oct. 2016, doi: 10.1093/ANNONC/MDW281.
- [64] "Study Results | Sorafenib in Treating Patients With Metastatic, Locally Advanced, or Recurrent Sarcoma | ClinicalTrials.gov." Accessed: Sep. 29, 2024.

- [Online]. Available:
<https://clinicaltrials.gov/study/NCT00245102?cond=Malignant%20Peripheral%20Nerve%20Sheath%20Tumor&term=Malignant%20Peripheral%20Nerve%20Sheath%20Tumors&intr=Drug&aggFilters=phase:2,results:with,status:com&page=2&rank=15&tab=results#outcome-measures>
- [65] J. Yan *et al.*, "Tumor-intrinsic PRC2 inactivation drives a context-dependent immune-desert microenvironment and is sensitized by immunogenic viruses," *J Clin Invest*, vol. 132, no. 17, Sep. 2022, doi: 10.1172/JCI153437.
 - [66] M. Høland *et al.*, "Transcriptomic subtyping of malignant peripheral nerve sheath tumours highlights immune signatures, genomic profiles, patient survival and therapeutic targets," *EBioMedicine*, vol. 97, Nov. 2023, doi: 10.1016/j.ebiom.2023.104829.
 - [67] B. N. Somatilaka *et al.*, "STING activation reprograms the microenvironment to sensitize NF1-related malignant peripheral nerve sheath tumors for immunotherapy," *J Clin Invest*, Mar. 2024, doi: 10.1172/JCI176748.
 - [68] M. Payandeh, M. Sadeghi, and E. Sadeghi, "Complete Response to Pembrolizumab in a Patient with Malignant Peripheral Nerve Sheath Tumor: The First Case Reported," *J Appl Pharm Sci*, vol. 7, no. 10, pp. 182–184, Oct. 2017, doi: 10.7324/JAPS.2017.71026.
 - [69] B. C. Özdemir *et al.*, "Deep Response to Anti-PD-1 Therapy of Metastatic Neurofibromatosis Type 1-Associated Malignant Peripheral Nerve Sheath Tumor With CD274/PD-L1 Amplification," *JCO Precis Oncol*, no. 3, pp. 1–6, Dec. 2019, doi: 10.1200/PO.18.00375/SUPPL_FILE/DS_PO.18.00375.DOCX.
 - [70] L. E. Davis, L. A. Nicholls, H. M. Babiker, J. Liao, and D. Mahadevan, "PD-1 Inhibition Achieves a Complete Metabolic Response in a Patient with Malignant Peripheral Nerve Sheath Tumor," *Cancer Immunol Res*, vol. 7, no. 9, pp. 1396–1400, 2019, doi: 10.1158/2326-6066.CIR-19-0072.
 - [71] J. M. Cooper *et al.*, "Overcoming BET inhibitor resistance in malignant peripheral nerve sheath tumors," *Clinical Cancer Research*, vol. 25, no. 11, pp. 3404–3416, Jun. 2019, doi: 10.1158/1078-0432.CCR-18-2437/15312/AM/OVERCOMING-BET-INHIBITOR-RESISTANCE-IN-MALIGNANT.
 - [72] A. J. Patel *et al.*, "PRC2-Inactivating Mutations in Cancer Enhance Cytotoxic Response to DNMT1-Targeted Therapy via Enhanced Viral Mimicry," *Cancer Discov*, vol. 12, no. 9, p. 2120, Sep. 2022, doi: 10.1158/2159-8290.CD-21-1671.
 - [73] B. Evers, K. Jastrzebski, J. P. M. Heijmans, W. Grenrum, R. L. Beijersbergen, and R. Bernards, "CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes," *Nature Biotechnology* 2016 34:6, vol. 34, no. 6, pp. 631–633, Apr. 2016, doi: 10.1038/nbt.3536.
 - [74] B. E. Housden and N. Perrimon, "Comparing CRISPR and RNAi-based screening technologies," *Nature Biotechnology* 2016 34:6, vol. 34, no. 6, pp. 621–623, Jun. 2016, doi: 10.1038/nbt.3599.
 - [75] M. B. Dong *et al.*, "Systematic Immunotherapy Target Discovery Using Genome-Scale In Vivo CRISPR Screens in CD8 T Cells," *Cell*, vol. 178, no. 5, pp. 1189–1204.e23, Aug. 2019, doi: 10.1016/J.CELL.2019.07.044.

- [76] R. T. Manguso *et al.*, "In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target," *Nature* 2017 547:7664, vol. 547, no. 7664, pp. 413–418, Jul. 2017, doi: 10.1038/nature23270.
- [77] K. Tzelepis *et al.*, "A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia," *Cell Rep*, vol. 17, no. 4, pp. 1193–1205, Oct. 2016, doi: 10.1016/J.CELREP.2016.09.079.
- [78] N. A. Thompson *et al.*, "Combinatorial CRISPR screen identifies fitness effects of gene paralogues," *Nature Communications* 2021 12:1, vol. 12, no. 1, pp. 1–11, Feb. 2021, doi: 10.1038/s41467-021-21478-9.
- [79] H. M. Lee *et al.*, "A CRISPR-drug perturbational map for identifying compounds to combine with commonly used chemotherapeutics," *Nature Communications* 2023 14:1, vol. 14, no. 1, pp. 1–18, Nov. 2023, doi: 10.1038/s41467-023-43134-0.
- [80] B. P. Kleinstiver *et al.*, "Engineered CRISPR–Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing," *Nature Biotechnology* 2019 37:3, vol. 37, no. 3, pp. 276–282, Feb. 2019, doi: 10.1038/s41587-018-0011-0.
- [81] P. C. DeWeirdt *et al.*, "Optimization of AsCas12a for combinatorial genetic screens in human cells," *Nat Biotechnol*, 2020, doi: 10.1038/s41587-020-0600-6.
- [82] "DepMap: The Cancer Dependency Map Project at Broad Institute." Accessed: Aug. 15, 2024. [Online]. Available: https://depmap.org/portal/download/data_page_citation/
- [83] C. Pacini *et al.*, "A comprehensive clinically informed map of dependencies in cancer cells and framework for target prioritization," *Cancer Cell*, vol. 42, no. 2, pp. 301–316.e9, Feb. 2024, doi: 10.1016/J.CCELL.2023.12.016.
- [84] G. Picco *et al.*, "Werner helicase is a synthetic-lethal vulnerability in mismatch repair– deficient colorectal cancer refractory to targeted therapies, chemotherapy, and immunotherapy," *Cancer Discov*, vol. 11, no. 8, pp. 1923–1937, Aug. 2021, doi: 10.1158/2159-8290.CD-20-1508/674141/AM/WERNER-HELICASE-IS-A-SYNTHETIC-LETHAL.
- [85] J. M. Dempster *et al.*, "Chronos: a cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects," *Genome Biol*, vol. 22, no. 1, pp. 1–23, Dec. 2021, doi: 10.1186/S13059-021-02540-7/FIGURES/6.
- [86] J. M. Dempster *et al.*, "Agreement between two large pan-cancer CRISPR-Cas9 gene dependency data sets," *Nature Communications* 2019 10:1, vol. 10, no. 1, pp. 1–14, Dec. 2019, doi: 10.1038/s41467-019-13805-y.
- [87] J. Jumper *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature* 2021 596:7873, vol. 596, no. 7873, pp. 583–589, Jul. 2021, doi: 10.1038/s41586-021-03819-2.
- [88] H. Li and R. Durbin, "Fast and accurate short read alignment with Burrows-Wheeler transform," *Bioinformatics*, vol. 25, no. 14, pp. 1754–1760, Jul. 2009, doi: 10.1093/BIOINFORMATICS/BTP324.
- [89] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2," *Genome Biol*, vol. 15, no. 12, pp. 1–21, Dec. 2014, doi: 10.1186/S13059-014-0550-8/FIGURES/9.

- [90] "Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data." Accessed: Sep. 25, 2024. [Online]. Available: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- [91] Y. Liao, G. K. Smyth, and W. Shi, "featureCounts: an efficient general purpose program for assigning sequence reads to genomic features," *Bioinformatics*, vol. 30, no. 7, pp. 923–930, Apr. 2014, doi: 10.1093/BIOINFORMATICS/BTT656.
- [92] "FlowJo™ Version 10.10.0. Ashland, OR: Becton, Dickinson and Company; 2023", [Online]. Available: <https://www.flowjo.com/>
- [93] G. Van der Auwera, B. O'Connor, and Safari, "Genomics in the Cloud: Using Docker, GATK, and WDL in Terra," *O'Reilly Media*, p. 300, 2020, Accessed: Sep. 25, 2024. [Online]. Available: <https://www.oreilly.com/library/view/genomics-in-the/9781491975183/>
- [94] A. Subramanian *et al.*, "Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles," *Proc Natl Acad Sci U S A*, vol. 102, no. 43, pp. 15545–15550, Oct. 2005, doi: 10.1073/PNAS.0506580102/SUPPL_FILE/06580FIG7.JPG.
- [95] W. Li *et al.*, "MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens," *Genome Biol*, vol. 15, no. 12, p. 554, Dec. 2014, doi: 10.1186/s13059-014-0554-4.
- [96] A. Dobin *et al.*, "STAR: ultrafast universal RNA-seq aligner," *Bioinformatics*, vol. 29, no. 1, p. 15, Jan. 2013, doi: 10.1093/BIOINFORMATICS/BTS635.
- [97] "Releases · usadellab/Trimmomatic." Accessed: Sep. 25, 2024. [Online]. Available: <https://github.com/usadellab/Trimmomatic/releases>
- [98] T. Hart *et al.*, "High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities," *Cell*, vol. 163, no. 6, pp. 1515–1526, Dec. 2015, doi: 10.1016/j.cell.2015.11.015.
- [99] "Data | DepMap Portal." Accessed: Sep. 26, 2024. [Online]. Available: https://depmap.org/portal/data_page/?tab=allData
- [100] C. Finan *et al.*, "The druggable genome and support for target identification and validation in drug development," *Sci Transl Med*, vol. 9, no. 383, Mar. 2017, doi: 10.1126/SCITRANSLMED.AAG1166.
- [101] "Table of Pharmacogenomic Biomarkers in Drug Labeling | FDA." Accessed: Sep. 26, 2024. [Online]. Available: <https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling>
- [102] "Chronos: Gene Fitness Effect Inference from CRISPR." Accessed: Sep. 26, 2024. [Online]. Available: https://figshare.com/articles/dataset/Chronos_Gene_Fitness_Effect_Inference_from_CRISPR/14067047
- [103] J. Wang *et al.*, "SynLethDB 2.0: a web-based knowledge graph database on synthetic lethality for novel anticancer drug discovery," *Database (Oxford)*, vol. 2022, 2022, doi: 10.1093/DATABASE/BAAC030.
- [104] J. P. Concordet and M. Haeussler, "CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens," *Nucleic Acids Res*, vol. 46, no. Web Server issue, p. W242, Jul. 2018, doi: 10.1093/NAR/GKY354.

- [105] J. Joung *et al.*, "Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening," *Nat Protoc*, vol. 12, no. 4, pp. 828–863, Apr. 2017, doi: 10.1038/NPROT.2017.016.
- [106] H. Li, L. J. Chang, D. R. Neubauer, D. F. Muir, and M. R. Wallace, "Immortalization of human normal and NF1 neurofibroma Schwann cells," *Laboratory Investigation* 2016 96:10, vol. 96, no. 10, pp. 1105–1115, Sep. 2016, doi: 10.1038/labinvest.2016.88.
- [107] U. CC and D. HG, "Detecting mycoplasma contamination in cell cultures by polymerase chain reaction," *Methods Mol Biol*, vol. 731, pp. 93–103, 2011, doi: 10.1007/978-1-61779-080-5_8.
- [108] A. Bairoch, "The cellosaurus, a cell-line knowledge resource," *Journal of Biomolecular Techniques*, vol. 29, no. 2, pp. 25–38, Jul. 2018, doi: 10.7171/JBT.18-2902-002.
- [109] D. WK, L. JB, F. JA, S. HD, and O. P, "Radiosensitivity in vitro of human soft tissue sarcoma cell lines and skin fibroblasts derived from the same patients," *Int J Radiat Biol*, vol. 63, no. 2, pp. 191–198, 1993, doi: 10.1080/09553009314550251.
- [110] T. W. Glover, C. K. Stein, E. Legius, L. B. Andersen, A. Brereton, and S. Johnson, "Molecular and cytogenetic analysis of tumors in von recklinghausen neurofibromatosis," *Genes Chromosomes Cancer*, vol. 3, no. 1, pp. 62–70, Jan. 1991, doi: 10.1002/GCC.2870030111.
- [111] E. Legius *et al.*, "TP53 mutations are frequent in malignant NFI tumors," *Genes Chromosomes Cancer*, vol. 10, no. 4, pp. 250–255, Aug. 1994, doi: 10.1002/GCC.2870100405.
- [112] Y. Y. Mahller *et al.*, "Tissue Inhibitor of Metalloproteinase-3 via Oncolytic Herpesvirus Inhibits Tumor Growth and Vascular Progenitors," *Cancer Res*, vol. 68, no. 4, pp. 1170–1179, Feb. 2008, doi: 10.1158/0008-5472.CAN-07-2734.
- [113] H. Li, L.-J. Chang, D. R. Neubauer, D. F. Muir, and M. R. Wallace, "Immortalization of human normal and NF1 neurofibroma Schwann cells," *Laboratory Investigation* 2016 96:10, vol. 96, no. 10, pp. 1105–1115, Sep. 2016, doi: 10.1038/labinvest.2016.88.
- [114] "Azenta Life Science | Azenta Life Sciences." Accessed: Sep. 25, 2024. [Online]. Available: <https://www.azenta.com/>
- [115] W. G. Jacoby, "Loess:: a nonparametric, graphical tool for depicting relationships between variables," *Elect Stud*, vol. 19, no. 4, pp. 577–613, Dec. 2000, doi: 10.1016/S0261-3794(99)00028-1.
- [116] R. M. Meyers *et al.*, "Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells," *Nat Genet*, vol. 49, no. 12, pp. 1779–1784, Dec. 2017, doi: 10.1038/NG.3984.
- [117] B. S. Park, H. Jeon, S. G. Chi, and T. Kim, "Efficient prioritization of CRISPR screen hits by accounting for targeting efficiency of guide RNA," *BMC Biol*, vol. 21, no. 1, Dec. 2023, doi: 10.1186/S12915-023-01536-Y.
- [118] J. M. Dempster *et al.*, "Agreement between two large pan-cancer CRISPR-Cas9 gene dependency data sets," *Nat Commun*, vol. 10, no. 1, Dec. 2019, doi: 10.1038/s41467-019-13805-y.

- [119] G. Van der Auwera, B. O'Connor, and Safari, "Genomics in the Cloud: Using Docker, GATK, and WDL in Terra," *O'Reilly Media*, p. 300, 2020, Accessed: Aug. 15, 2024. [Online]. Available: <https://www.oreilly.com/library/view/genomics-in-the/9781491975183/>
- [120] "FLJ36000 Gene - GeneCards | FLJ36000 RNA Gene." Accessed: Oct. 03, 2024. [Online]. Available: <https://www.genecards.org/cgi-bin/carddisp.pl?gene=FLJ36000>
- [121] M. Magallón-Lorenz *et al.*, "Deep genomic analysis of malignant peripheral nerve sheath tumor cell lines challenges current malignant peripheral nerve sheath tumor diagnosis," *iScience*, vol. 26, no. 2, p. 106096, Feb. 2023, doi: 10.1016/j.isci.2023.106096.
- [122] M. Magallón-Lorenz *et al.*, "Deep genomic analysis of malignant peripheral nerve sheath tumor cell lines challenges current malignant peripheral nerve sheath tumor diagnosis," *iScience*, vol. 26, no. 2, p. 106096, Feb. 2023, doi: 10.1016/J.ISCI.2023.106096.
- [123] E. Legius *et al.*, "TP53 mutations are frequent in malignant NFI tumors," *Genes Chromosomes Cancer*, vol. 10, no. 4, pp. 250–255, Aug. 1994, doi: 10.1002/GCC.2870100405.
- [124] J. E. Reynolds, J. A. Fletcher, C. H. Lytle, L. Nie, C. C. Morton, and S. R. Diehl, "Molecular characterization of a 17q11.2 translocation in a malignant schwannoma cell line," *Hum Genet*, vol. 90, no. 4, pp. 450–456, Dec. 1992, doi: 10.1007/BF00220476/METRICS.
- [125] Y. Y. Mahller *et al.*, "Tissue Inhibitor of Metalloproteinase-3 via Oncolytic Herpesvirus Inhibits Tumor Growth and Vascular Progenitors," *Cancer Res*, vol. 68, no. 4, pp. 1170–1179, Feb. 2008, doi: 10.1158/0008-5472.CAN-07-2734.
- [126] D. WK, L. JB, F. JA, S. HD, and O. P, "Radiosensitivity in vitro of human soft tissue sarcoma cell lines and skin fibroblasts derived from the same patients," *Int J Radiat Biol*, vol. 63, no. 2, pp. 191–198, 1993, doi: 10.1080/09553009314550251.
- [127] J. Joung *et al.*, "Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening," *Nat Protoc*, vol. 12, no. 4, pp. 828–863, Apr. 2017, doi: 10.1038/nprot.2017.016.
- [128] P. Corish and C. Tyler-Smith, "Attenuation of green fluorescent protein half-life in mammalian cells," *Protein Engineering, Design and Selection*, vol. 12, no. 12, pp. 1035–1040, Dec. 1999, doi: 10.1093/PROTEIN/12.12.1035.
- [129] J. G. Doench, "Am I ready for CRISPR? A user's guide to genetic screens," *Nature Reviews Genetics* 2017 19:2, vol. 19, no. 2, pp. 67–80, Dec. 2017, doi: 10.1038/nrg.2017.97.
- [130] J. M. Dempster *et al.*, "Chronos: a cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects," *Genome Biol*, vol. 22, no. 1, pp. 1–23, Dec. 2021, doi: 10.1186/S13059-021-02540-7/FIGURES/6.
- [131] P. Schober and L. A. Schwarte, "Correlation coefficients: Appropriate use and interpretation," *Anesth Analg*, vol. 126, no. 5, pp. 1763–1768, May 2018, doi: 10.1213/ANE.0000000000002864.

- [132] I. Bartha, J. Di Iulio, J. C. Venter, and A. Telenti, "Human gene essentiality," *Nature Reviews Genetics* 2017 19:1, vol. 19, no. 1, pp. 51–62, Oct. 2017, doi: 10.1038/nrg.2017.75.
- [133] V. A. Blomen *et al.*, "Gene essentiality and synthetic lethality in haploid human cells," *Science*, vol. 350, no. 6264, pp. 1092–1096, Nov. 2015, doi: 10.1126/SCIENCE.AAC7557.
- [134] T. Hart and J. Moffat, "BAGEL: A computational framework for identifying essential genes from pooled library screens," *BMC Bioinformatics*, vol. 17, no. 1, pp. 1–7, Apr. 2016, doi: 10.1186/s12859-016-1015-8.
- [135] J. Winter *et al.*, "CRISPRAnalyzeR: Interactive analysis, annotation and documentation of pooled CRISPR screens," *bioRxiv*, p. 109967, Feb. 2017, doi: 10.1101/109967.
- [136] T. Hart, M. Chandrashekhar, M. Aregger, D. Durocher, S. Angers, and J. M. Correspondence, "High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities," 2015, doi: 10.1016/j.cell.2015.11.015.
- [137] R. M. Meyers *et al.*, "Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells," *Nat Genet*, vol. 49, no. 12, pp. 1779–1784, Dec. 2017, doi: 10.1038/ng.3984.
- [138] E. Talevich, A. H. Shain, T. Botton, and B. C. Bastian, "CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing," *PLoS Comput Biol*, vol. 12, no. 4, p. e1004873, Apr. 2016, doi: 10.1371/JOURNAL.PCBI.1004873.
- [139] "CRISPick."
- [140] T. G. Montague, J. M. Cruz, J. A. Gagnon, G. M. Church, and E. Valen, "CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing," *Nucleic Acids Res*, vol. 42, no. Web Server issue, Jul. 2014, doi: 10.1093/NAR/GKU410.
- [141] A. Cabrera *et al.*, "The sound of silence: Transgene silencing in mammalian cell engineering," *Cell Syst*, vol. 13, no. 12, pp. 950–973, Dec. 2022, doi: 10.1016/J.CELS.2022.11.005.
- [142] A. Pérez-González and E. Caro, "Benefits of using genomic insulators flanking transgenes to increase expression and avoid positional effects," *Scientific Reports* 2019 9:1, vol. 9, no. 1, pp. 1–11, Jun. 2019, doi: 10.1038/s41598-019-44836-6.
- [143] N. P. Sweeney and C. A. Vink, "The impact of lentiviral vector genome size and producer cell genomic to gag-pol mRNA ratios on packaging efficiency and titre," *Mol Ther Methods Clin Dev*, vol. 21, p. 574, Jun. 2021, doi: 10.1016/J.OMTM.2021.04.007.
- [144] S. J. Heo *et al.*, "Compact CRISPR genetic screens enabled by improved guide RNA library cloning," *Genome Biol*, vol. 25, no. 1, pp. 1–20, Dec. 2024, doi: 10.1186/S13059-023-03132-3/FIGURES/6.
- [145] T. J. Hwang, D. Carpenter, J. C. Lauffenburger, B. Wang, J. M. Franklin, and A. S. Kesselheim, "Failure of investigational drugs in late-stage clinical development and publication of trial results," *JAMA Intern Med*, vol. 176, no. 12, pp. 1826–1833, Dec. 2016, doi: 10.1001/jamainternmed.2016.6008.

- [146] J. R. Cantor *et al.*, "Physiologic Medium Rewires Cellular Metabolism and Reveals Uric Acid as an Endogenous Inhibitor of UMP Synthase," *Cell*, vol. 169, no. 2, pp. 258-272.e17, Apr. 2017, doi: 10.1016/j.cell.2017.03.023.
- [147] J. R. Cantor *et al.*, "Physiologic Medium Rewires Cellular Metabolism and Reveals Uric Acid as an Endogenous Inhibitor of UMP Synthase," *Cell*, vol. 169, no. 2, pp. 258-272.e17, Apr. 2017, doi: 10.1016/J.CELL.2017.03.023/ATTACHMENT/2DBD023D-FB4C-485B-AC88-4EA1EA829FD2/MMC3.XLSX.
- [148] N. J. Rossiter *et al.*, "CRISPR screens in physiologic medium reveal conditionally essential genes in human cells," *Cell Metab*, vol. 33, no. 6, pp. 1248-1263.e9, Jun. 2021, doi: 10.1016/J.CMET.2021.02.005/ATTACHMENT/A07D043D-77FB-40C5-8225-645E1CE9FF5F/MMC7.XLSX.
- [149] A. Tsherniak *et al.*, "Defining a Cancer Dependency Map," *Cell*, vol. 170, no. 3, pp. 564-576.e16, Jul. 2017, doi: 10.1016/J.CELL.2017.06.010.
- [150] M. P. Bayona-Bafaluy, J. Montoya, and E. Ruiz-Pesini, "Oxidative phosphorylation system and cell culture media," *Trends Cell Biol*, vol. 31, no. 8, pp. 618–620, Aug. 2021, doi: 10.1016/J.TCB.2021.05.003.
- [151] T. Hu *et al.*, "Single-cell spatial metabolomics with cell-type specific protein profiling for tissue systems biology," *Nature Communications 2023 14:1*, vol. 14, no. 1, pp. 1–20, Dec. 2023, doi: 10.1038/s41467-023-43917-5.
- [152] Y. Hou, H. Yao, and J. M. Lin, "Recent advancements in single-cell metabolic analysis for pharmacological research," *J Pharm Anal*, vol. 13, no. 10, p. 1102, Oct. 2023, doi: 10.1016/J.JPHA.2023.08.014.
- [153] Y. Huang *et al.*, "Characterizing cancer metabolism from bulk and single-cell RNA-seq data using METAFIux," *Nature Communications 2023 14:1*, vol. 14, no. 1, pp. 1–16, Aug. 2023, doi: 10.1038/s41467-023-40457-w.
- [154] J. Lever, M. Krzywinski, and N. Altman, "Principal component analysis," *Nature Methods 2017 14:7*, Jun. 2017, Accessed: Sep. 22, 2024. [Online]. Available: <https://www.nature.com/articles/nmeth.4346>
- [155] M. P. Bayona-Bafaluy, J. Montoya, and E. Ruiz-Pesini, "Oxidative phosphorylation system and cell culture media," *Trends Cell Biol*, vol. 31, no. 8, pp. 618–620, Aug. 2021, doi: 10.1016/J.TCB.2021.05.003.
- [156] P. C. Calder, "Fatty Acids: Metabolism," *Encyclopedia of Food and Health*, pp. 632–644, Sep. 2015, doi: 10.1016/B978-0-12-384947-2.00276-2.
- [157] Y. Zhao *et al.*, "CEDA: integrating gene expression data with CRISPR-pooled screen data identifies essential genes with higher expression," *Bioinformatics*, vol. 38, no. 23, p. 5245, Dec. 2022, doi: 10.1093/BIOINFORMATICS/BTAC668.
- [158] H. Chen *et al.*, "New insights on human essential genes based on integrated analysis and the construction of the HEGIAP web-based platform," *Brief Bioinform*, vol. 21, no. 4, p. 1397, Jul. 2020, doi: 10.1093/BIB/BBZ072.
- [159] X. Liu *et al.*, "SLMO2 is a potential prognostic and immunological biomarker in human pan-cancer," *Scientific Reports 2024 14:1*, vol. 14, no. 1, pp. 1–15, Jan. 2024, doi: 10.1038/s41598-024-51720-5.
- [160] H. ZHOU, W. SUN, and J. ZOU, "Analysis of expression profiles and prognostic value of COP9 signalosome subunits for patients with head and neck squamous

- cell carcinoma," *Oncol Lett*, vol. 22, no. 5, Nov. 2021, doi: 10.3892/OL.2021.13064.
- [161] C. D. Van Raamsdonk and M. Deo, "Links between Schwann cells and melanocytes in development and disease," *Pigment Cell Melanoma Res*, vol. 26, no. 5, pp. 634–645, Sep. 2013, doi: 10.1111/PCMR.12134.
 - [162] A. Graham, "Melanocyte Production: Dark Side of the Schwann Cell," *Current Biology*, vol. 19, pp. R1116–R1117, doi: 10.1016/j.cub.2009.10.063.
 - [163] L. Hou and W. J. Pavan, "Transcriptional and signaling regulation in neural crest stem cell-derived melanocyte development: do all roads lead to Mitf?," *Cell Research 2008 18:12*, vol. 18, no. 12, pp. 1163–1176, Nov. 2008, doi: 10.1038/cr.2008.303.
 - [164] E. Nitzan, E. R. Pfaltzgraff, P. A. Labosky, and C. Kalcheim, "Neural crest and Schwann cell progenitor-derived melanocytes are two spatially segregated populations similarly regulated by Foxd3," *Proc Natl Acad Sci U S A*, vol. 110, no. 31, pp. 12709–12714, Jul. 2013, doi: 10.1073/PNAS.1306287110/SUPPL_FILE/PNAS.201306287SI.PDF.
 - [165] I. Adameyko *et al.*, "Schwann Cell Precursors from Nerve Innervation Are a Cellular Origin of Melanocytes in Skin," *Cell*, vol. 139, no. 2, pp. 366–379, Oct. 2009, doi: 10.1016/J.CELL.2009.07.049/ATTACHMENT/9BAEBD72-1304-4C0D-8956-C67A79672C0C/MMC1.PDF.
 - [166] D. R. Scoles *et al.*, "Neurofibromatosis 2 tumour suppressor schwannomin interacts with β II-spectrin," *Nature Genetics 1998 18:4*, vol. 18, no. 4, pp. 354–359, 1998, doi: 10.1038/ng0498-354.
 - [167] H. Morrison *et al.*, "The NF2 tumor suppressor gene product, merlin, mediates contact inhibition of growth through interactions with CD44," *Genes Dev*, vol. 15, no. 8, p. 968, Apr. 2001, doi: 10.1101/GAD.189601.
 - [168] H. N. Vasudevan *et al.*, "Functional interactions between neurofibromatosis tumor suppressors underlie Schwann cell tumor de-differentiation and treatment resistance," *Nature Communications 2024 15:1*, vol. 15, no. 1, pp. 1–13, Jan. 2024, doi: 10.1038/s41467-024-44755-9.
 - [169] K. Susuki *et al.*, "Schwann cell spectrins modulate peripheral nerve myelination," *Proc Natl Acad Sci U S A*, vol. 108, no. 19, pp. 8009–8014, May 2011, doi: 10.1073/PNAS.1019600108/SUPPL_FILE/PNAS.201019600SI.PDF.
 - [170] "SPTBN1 DepMap Gene Summary." Accessed: Aug. 31, 2024. [Online]. Available: <https://depmap.org/portal/gene/SPTBN1?tab=dependency>
 - [171] P. Hoellerbauer *et al.*, "FBXO42 activity is required to prevent mitotic arrest, spindle assembly checkpoint activation, and lethality in glioblastoma and other cancers," *bioRxiv*, p. 2022.11.29.518420, Dec. 2022, doi: 10.1101/2022.11.29.518420.
 - [172] R. Firestein *et al.*, "CDK8 is a colorectal cancer oncogene that regulates β -catenin activity," *Nature*, vol. 455, no. 7212, p. 547, Sep. 2008, doi: 10.1038/NATURE07179.
 - [173] K. Fukasawa *et al.*, "CDK8 maintains stemness and tumorigenicity of glioma stem cells by regulating the c-MYC pathway," *Oncogene 2021 40:15*, vol. 40, no. 15, pp. 2803–2815, Mar. 2021, doi: 10.1038/s41388-021-01745-1.

- [174] X. Yin *et al.*, "Unveiling the impact of CDK8 on tumor progression: mechanisms and therapeutic strategies," *Front Pharmacol*, vol. 15, p. 1386929, Mar. 2024, doi: 10.3389/FPHAR.2024.1386929/BIBTEX.
- [175] Y. Zhu, Y. Zhou, Y. Liu, X. Wang, and J. Li, "SLGNN: synthetic lethality prediction in human cancers based on factor-aware knowledge graph neural network," *Bioinformatics*, vol. 39, no. 2, Feb. 2023, doi: 10.1093/BIOINFORMATICS/BTAD015.
- [176] J. Wang *et al.*, "Computational methods, databases and tools for synthetic lethality prediction," *Brief Bioinform*, vol. 23, no. 3, pp. 1–22, May 2022, doi: 10.1093/BIB/BBAC106.
- [177] C. J. Lord and A. Ashworth, "PARP Inhibitors: The First Synthetic Lethal Targeted Therapy," *Science*, vol. 355, no. 6330, p. 1152, Mar. 2017, doi: 10.1126/SCIENCE.AAM7344.
- [178] R. Dhanasekaran, A. Deutzmann, W. D. Mahauad-Fernandez, A. S. Hansen, A. M. Gouw, and D. W. Felsher, "The MYC oncogene — the grand orchestrator of cancer growth and immune evasion," *Nature Reviews Clinical Oncology* 2021 19:1, vol. 19, no. 1, pp. 23–36, Sep. 2021, doi: 10.1038/s41571-021-00549-2.
- [179] C. Wang *et al.*, "Alternative approaches to target Myc for cancer treatment," *Signal Transduction and Targeted Therapy* 2021 6:1, vol. 6, no. 1, pp. 1–14, Mar. 2021, doi: 10.1038/s41392-021-00500-y.
- [180] E. Garraalda *et al.*, "MYC targeting by OMO-103 in solid tumors: a phase 1 trial," *Nature Medicine* 2024 30:3, vol. 30, no. 3, pp. 762–771, Feb. 2024, doi: 10.1038/s41591-024-02805-1.
- [181] B. N. Somatilaka, A. Sadek, R. M. McKay, and L. Q. Le, "Malignant Peripheral Nerve Sheath Tumor: Models, Biology, and Translation," *Oncogene*, vol. 41, no. 17, p. 2405, Apr. 2022, doi: 10.1038/S41388-022-02290-1.
- [182] L. Zhou, Y. Jiang, Q. Luo, L. Li, and L. Jia, "Neddylation: A novel modulator of the tumor microenvironment," *Mol Cancer*, vol. 18, no. 1, pp. 1–11, Apr. 2019, doi: 10.1186/S12943-019-0979-1/FIGURES/5.
- [183] T. M. Miller *et al.*, " Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS ," *New England Journal of Medicine*, vol. 387, no. 12, pp. 1099–1110, Sep. 2022, doi: 10.1056/NEJMOA2204705/SUPPL_FILE/NEJMOA2204705_DATA-SHARING.PDF.
- [184] R. E. A. Gutteridge, M. A. Ndiaye, X. Liu, and N. Ahmad, "Plk1 inhibitors in cancer therapy: From laboratory to clinics," *Mol Cancer Ther*, vol. 15, no. 7, pp. 1427–1435, Jul. 2016, doi: 10.1158/1535-7163.MCT-15-0897/10637/P/PLK1-INHIBITORS-IN-CANCER-THERAPY-FROM-LABORATORY.
- [185] J. Meng, Y. Wei, Q. Deng, L. Li, and X. Li, "Study on the expression of TOP2A in hepatocellular carcinoma and its relationship with patient prognosis," *Cancer Cell Int*, vol. 22, no. 1, pp. 1–18, Dec. 2022, doi: 10.1186/S12935-021-02439-0/FIGURES/12.
- [186] Y. Zhou *et al.*, "KIF11 is upregulated in colorectal cancer and silencing of it impairs tumor growth and sensitizes colorectal cancer cells to oxaliplatin via p53/GSK3 β signaling," *J Cancer*, vol. 12, no. 12, p. 3741, 2021, doi: 10.7150/JCA.52103.

- [187] A. C. Hsieh and D. Ruggero, "Targeting eukaryotic translation initiation factor 4E (eIF4E) in cancer," *Clinical Cancer Research*, vol. 16, no. 20, pp. 4914–4920, Oct. 2010, doi: 10.1158/1078-0432.CCR-10-0433/AM/TARGETING-EIF4E-IN-CANCERTARGETING-EIF4E-IN-CANCER.
- [188] "Table of Pharmacogenomic Biomarkers in Drug Labeling | FDA." Accessed: Aug. 25, 2024. [Online]. Available: <https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling>
- [189] B. A. Helfrich *et al.*, "Barasertib (AZD1152), a Small Molecule Aurora B Inhibitor, Inhibits the Growth of SCLC Cell Lines In Vitro and In Vivo," *Mol Cancer Ther*, vol. 15, no. 10, pp. 2314–2322, Oct. 2016, doi: 10.1158/1535-7163.MCT-16-0298.
- [190] A. H. Kovacs, D. Zhao, and J. Hou, "Aurora B Inhibitors as Cancer Therapeutics," *Molecules*, vol. 28, no. 8, p. 3385, Apr. 2023, doi: 10.3390/MOLECULES28083385/S1.
- [191] N. A. Borah and M. M. Reddy, "Aurora Kinase B Inhibition: A Potential Therapeutic Strategy for Cancer," *Molecules*, vol. 26, no. 7, 2021, doi: 10.3390/MOLECULES26071981.
- [192] "DGIdb." Accessed: Aug. 25, 2024. [Online]. Available: <https://www.dgidb.org/>
- [193] P. Łukasik, I. Baranowska-bosiacka, K. Kulczycka, and I. Gutowska, "Inhibitors of Cyclin-Dependent Kinases: Types and Their Mechanism of Action," *Int J Mol Sci*, vol. 22, no. 6, pp. 1–23, Mar. 2021, doi: 10.3390/IJMS22062806.
- [194] K. Jiang, M. Deng, W. Du, T. Liu, J. Li, and Y. Zhou, "Functions and inhibitors of CHK1 in cancer therapy," *Med Drug Discov*, vol. 22, p. 100185, Jun. 2024, doi: 10.1016/J.MEDIDD.2024.100185.
- [195] K. A. Urtishak *et al.*, "Targeting EIF4E signaling with ribavirin in infant acute lymphoblastic leukemia," *Oncogene*, vol. 38, no. 13, p. 2241, Mar. 2019, doi: 10.1038/S41388-018-0567-7.
- [196] A. Mullard, "The FDA approves a first farnesyltransferase inhibitor," *Nat Rev Drug Discov*, vol. 20, no. 1, p. 8, Jan. 2021, doi: 10.1038/D41573-020-00213-X.
- [197] S. Figarol *et al.*, "Farnesyltransferase inhibition overcomes oncogene-addicted non-small cell lung cancer adaptive resistance to targeted therapies," *Nature Communications* 2024 15:1, vol. 15, no. 1, pp. 1–15, Jun. 2024, doi: 10.1038/s41467-024-49360-4.
- [198] R. Chamariya and V. Suvarna, "Role of KSP Inhibitors as Anti-cancer Therapeutics: An Update," *Anticancer Agents Med Chem*, vol. 22, no. 14, pp. 2517–2538, Jan. 2022, doi: 10.2174/1871520622666220119093105.
- [199] E. S. Ali *et al.*, "Recent advances and limitations of mTOR inhibitors in the treatment of cancer," *Cancer Cell International* 2022 22:1, vol. 22, no. 1, pp. 1–16, Sep. 2022, doi: 10.1186/S12935-022-02706-8.
- [200] S. Su, G. Chhabra, C. K. Singh, M. A. Ndiaye, and N. Ahmad, "PLK1 inhibition-based combination therapies for cancer management," *Transl Oncol*, vol. 16, p. 101332, Feb. 2022, doi: 10.1016/J.TRANON.2021.101332.
- [201] D. Kaida *et al.*, "Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA," *Nat Chem Biol*, vol. 3, no. 9, pp. 576–583, 2007, doi: 10.1038/NCHEMBIO.2007.18.

- [202] T. Carvalho, S. Martins, J. Rino, S. Marinho, and M. Carmo-Fonseca, "Pharmacological inhibition of the spliceosome subunit SF3b triggers exon junction complex-independent nonsense-mediated decay," *J Cell Sci*, vol. 130, no. 9, pp. 1519–1531, May 2017, doi: 10.1242/JCS.202200.
- [203] T. Yamano *et al.*, "Splicing modulator FR901464 is a potential agent for colorectal cancer in combination therapy," *Oncotarget*, vol. 10, no. 3, p. 352, Jan. 2019, doi: 10.18632/ONCOTARGET.26564.
- [204] R. J. Mead, N. Shan, H. J. Reiser, F. Marshall, and P. J. Shaw, "Amyotrophic lateral sclerosis: a neurodegenerative disorder poised for successful therapeutic translation," *Nature Reviews Drug Discovery* 2022 22:3, vol. 22, no. 3, pp. 185–212, Dec. 2022, doi: 10.1038/s41573-022-00612-2.
- [205] "DNA Topoisomerase II α (TOP2A) inhibitors doxorubicin and etoposide (VP-16) up-regulate Fatty Acid Synthase (FAS) gene expression in SK-Br3 breast cancer cells: In vitro evidence for a 'functional amplicon' involving FAS, Her-2/neu and TOP2A genes | Cancer Research | American Association for Cancer Research." Accessed: Aug. 25, 2024. [Online]. Available: https://aacrjournals.org/cancerres/article/65/9_Supplement/834/522159/DNA-Topoisomerase-II-TOP2A-inhibitors-doxorubicin
- [206] M. Abu Saleh, M. Solayman, M. M. Hoque, M. A. K. Khan, M. G. Sarwar, and M. A. Halim, "Inhibition of DNA Topoisomerase Type II α (TOP2A) by Mitoxantrone and Its Halogenated Derivatives: A Combined Density Functional and Molecular Docking Study," *Biomed Res Int*, vol. 2016, 2016, doi: 10.1155/2016/6817502.
- [207] J. Elsner *et al.*, "Structure-Guided Optimization Provides a Series of TTK Protein Inhibitors with Potent Antitumor Activity," *J Med Chem*, vol. 64, no. 17, pp. 12670–12679, Sep. 2021, doi: 10.1021/ACS.JMEDCHEM.1C00635/SUPPL_FILE/JM1C00635_SI_002.CSV.
- [208] A. Warren *et al.*, "Global computational alignment of tumor and cell line transcriptional profiles," *Nature Communications* 2021 12:1, vol. 12, no. 1, pp. 1–12, Jan. 2021, doi: 10.1038/s41467-020-20294-x.
- [209] P. W. Nagle, J. T. M. Plukker, C. T. Muijs, P. van Luijk, and R. P. Coppes, "Patient-derived tumor organoids for prediction of cancer treatment response," *Semin Cancer Biol*, vol. 53, pp. 258–264, Dec. 2018, doi: 10.1016/J.SEMCANCER.2018.06.005.
- [210] S. Alseekh *et al.*, "Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices," *Nature Methods* 2021 18:7, vol. 18, no. 7, pp. 747–756, Jul. 2021, doi: 10.1038/s41592-021-01197-1.
- [211] C. J. Ryan, I. Bajrami, and C. J. Lord, "Synthetic Lethality and Cancer – Penetrance as the Major Barrier," *Trends Cancer*, vol. 4, no. 10, pp. 671–683, Oct. 2018, doi: 10.1016/J.TRECAN.2018.08.003/ASSET/9F642C67-880E-4910-ADA5-FBB8F9382332/MAIN.ASSETS/GR3.SML.
- [212] A. Al Shihabi *et al.*, "The landscape of drug sensitivity and resistance in sarcoma," *Cell Stem Cell*, vol. 31, no. 10, pp. 1524–1542.e4, Oct. 2024, doi: 10.1016/J.STEM.2024.08.010.

- [213] J. L. Kohlmeyer *et al.*, "CDK4/6-MEK inhibition in MPNSTs causes plasma cell infiltration, sensitization to PD-L1 blockade, and tumor regression," *Clin Cancer Res*, vol. 29, no. 17, p. 3484, Sep. 2023, doi: 10.1158/1078-0432.CCR-23-0749.
- [214] A. Schulte *et al.*, "Combined Targeting of AKT and mTOR Inhibits Proliferation of Human NF1-Associated Malignant Peripheral Nerve Sheath Tumour Cells In Vitro but not in a Xenograft Mouse Model In Vivo," *Int J Mol Sci*, vol. 21, no. 4, Feb. 2020, doi: 10.3390/IJMS21041548.
- [215] J. Guo *et al.*, "Polo-like kinase 1 as a therapeutic target for malignant peripheral nerve sheath tumors (MPNST) and schwannomas," *Am J Cancer Res*, vol. 10, no. 3, p. 856, 2020, Accessed: Oct. 03, 2024. [Online]. Available: [/pmc/articles/PMC7136923/](#)
- [216] A. V. Patel *et al.*, "Ras-Driven Transcriptome Analysis Identifies Aurora Kinase A as a Potential Malignant Peripheral Nerve Sheath Tumor Therapeutic Target," *Clin Cancer Res*, vol. 18, no. 18, p. 5020, Sep. 2012, doi: 10.1158/1078-0432.CCR-12-1072.
- [217] S. Suppiah *et al.*, "Multiplatform molecular profiling uncovers two subgroups of malignant peripheral nerve sheath tumors with distinct therapeutic vulnerabilities," *Nature Communications* 2023 14:1, vol. 14, no. 1, pp. 1–15, May 2023, doi: 10.1038/s41467-023-38432-6.
- [218] M. Yang, K. Liu, P. Chen, H. Zhu, J. Wang, and J. Huang, "Bromodomain-containing protein 4 (BRD4) as an epigenetic regulator of fatty acid metabolism genes and ferroptosis," *Cell Death & Disease* 2022 13:10, vol. 13, no. 10, pp. 1–17, Oct. 2022, doi: 10.1038/s41419-022-05344-0.
- [219] S. Guil and M. Esteller, "PRC2 Loss and DNMT Inhibition Boost Viral Mimicry in Cancer," *Cancer Discov*, vol. 12, no. 9, pp. 2020–2022, Sep. 2022, doi: 10.1158/2159-8290.CD-22-0733.
- [220] A. P. Bracken, N. Dietrich, D. Pasini, K. H. Hansen, and K. Helin, "Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions," *Genes Dev*, vol. 20, no. 9, p. 1123, May 2006, doi: 10.1101/GAD.381706.
- [221] A. Glaviano *et al.*, "PI3K/AKT/mTOR signaling transduction pathway and targeted therapies in cancer," *Mol Cancer*, vol. 22, no. 1, pp. 1–37, Dec. 2023, doi: 10.1186/S12943-023-01827-6/TABLES/1.
- [222] A. Ishii, M. Furusho, and R. Bansal, "Mek/ERK1/2-MAPK and PI3K/Akt/mTOR signaling plays both independent and cooperative roles in Schwann cell differentiation, myelination and dysmyelination," *Glia*, vol. 69, no. 10, pp. 2429–2446, Oct. 2021, doi: 10.1002/GLIA.24049.
- [223] J. Varin *et al.*, "Dual mTORC1/2 inhibition induces anti-proliferative effect in NF1-associated plexiform neurofibroma and malignant peripheral nerve sheath tumor cells," *Oncotarget*, vol. 7, no. 24, pp. 35753–35767, Jan. 2016, doi: 10.18632/ONCOTARGET.7099.
- [224] M. A. M. Reijns *et al.*, "Signatures of TOP1 transcription-associated mutagenesis in cancer and germline," *Nature* 2022 602:7898, vol. 602, no. 7898, pp. 623–631, Feb. 2022, doi: 10.1038/s41586-022-04403-y.
- [225] Y. Pommier and A. Thomas, "New Life of Topoisomerase I Inhibitors as Antibody-Drug Conjugate Warheads," *Clinical Cancer Research*, vol. 29, no. 6, pp. 991–993, Mar. 2023, doi: 10.1158/1078-0432.CCR-22-

- 3640/715749/AM/NEW-LIFE-OF-TOPOISOMERASE-I-INHIBITORS-AS-ANTIBODY.
- [226] E. Martino *et al.*, "The long story of camptothecin: From traditional medicine to drugs," *Bioorg Med Chem Lett*, vol. 27, no. 4, pp. 701–707, Feb. 2017, doi: 10.1016/J.BMCL.2016.12.085.
 - [227] J. Garst, "Topotecan: An evolving option in the treatment of relapsed small cell lung cancer," *Ther Clin Risk Manag*, vol. 3, no. 6, p. 1087, 2007, Accessed: Sep. 16, 2024. [Online]. Available: /pmc/articles/PMC2387299/
 - [228] K. I. Fujita, Y. Kubota, H. Ishida, and Y. Sasaki, "Irinotecan, a key chemotherapeutic drug for metastatic colorectal cancer," *World J Gastroenterol*, vol. 21, no. 43, p. 12234, Nov. 2015, doi: 10.3748/WJG.V21.I43.12234.
 - [229] P. Lin *et al.*, "Topoisomerase 1 Inhibition in MYC-Driven Cancer Promotes Aberrant R-Loop Accumulation to Induce Synthetic Lethality," *Cancer Res*, vol. 183, no. 24, pp. 4015–4029, Dec. 2023, doi: 10.1158/0008-5472.CAN-22-2948/730183/P/TOPOISOMERASE-1-INHIBITION-IN-MYC-DRIVEN-CANCER.
 - [230] Y. T. Liang, H. Luo, Y. Lin, and F. Gao, "Recent advances in the characterization of essential genes and development of a database of essential genes," *iMeta*, vol. 3, no. 1, p. e157, Feb. 2024, doi: 10.1002/IMT2.157.
 - [231] C. Tudose, J. Bond, and C. J. Ryan, "Gene essentiality in cancer is better predicted by mRNA abundance than by gene regulatory network-inferred activity," *NAR Cancer*, vol. 5, no. 4, pp. 1–13, Oct. 2023, doi: 10.1093/NARCAN/ZCAD056.
 - [232] S. Chen, L. Jiao, M. Shubbar, X. Yang, and X. Liu, "Unique Structural Platforms of Suz12 Dictate Distinct Classes of PRC2 for Chromatin Binding," *Mol Cell*, vol. 69, no. 5, p. 840, Mar. 2018, doi: 10.1016/J.MOLCEL.2018.01.039.
 - [233] S. Chen, L. Jiao, X. Liu, X. Yang, and X. Liu, "A Dimeric Structural Scaffold for PRC2-PCL Targeting to CpG Island Chromatin," *Mol Cell*, vol. 77, no. 6, p. 1265, Mar. 2020, doi: 10.1016/J.MOLCEL.2019.12.019.
 - [234] A. L. Mueller *et al.*, "The geometric influence on the Cys2His2 zinc finger domain and functional plasticity," *Nucleic Acids Res*, vol. 48, no. 11, pp. 6382–6402, Jun. 2020, doi: 10.1093/NAR/GKAA291.
 - [235] S. Aizenshtein-Gazit and Y. Orenstein, "DeepZF: improved DNA-binding prediction of C2H2-zinc-finger proteins by deep transfer learning," *Bioinformatics*, vol. 38, no. Supplement_2, pp. ii62–ii67, Sep. 2022, doi: 10.1093/BIOINFORMATICS/BTAC469.
 - [236] L. Gong, X. Liu, L. Jiao, X. Yang, A. Lemoff, and X. Liu, "CK2-mediated phosphorylation of SUZ12 promotes PRC2 function by stabilizing enzyme active site," *Nat Commun*, vol. 13, no. 1, Dec. 2022, doi: 10.1038/S41467-022-34431-1.
 - [237] M. C. Mendoza, E. E. Er, and J. Blenis, "The Ras-ERK and PI3K-mTOR Pathways: Cross-talk and Compensation," *Trends Biochem Sci*, vol. 36, no. 6, p. 320, Jun. 2011, doi: 10.1016/J.TIBS.2011.03.006.
 - [238] N. Somaiah, B. Paudyal, R. E. Winkler, B. A. Van Tine, and A. C. Hirbe, "Malignant Peripheral Nerve Sheath Tumor, a Heterogeneous, Aggressive Cancer with Diverse Biomarkers and No Targeted Standard of Care: Review of the

- Literature and Ongoing Investigational Agents," *Target Oncol*, vol. 19, no. 5, pp. 665–678, Sep. 2024, doi: 10.1007/S11523-024-01078-5/TABLES/6.
- [239] A. Huang, L. A. Garraway, A. Ashworth, and B. Weber, "Synthetic lethality as an engine for cancer drug target discovery," Jan. 01, 2020, *Nature Research*. doi: 10.1038/s41573-019-0046-z.
- [240] "Significance of Blood Vessel Leakiness in Cancer1 | Cancer Research | American Association for Cancer Research." Accessed: Jan. 26, 2025. [Online]. Available: <https://aacrjournals.org/cancerres/article/62/18/5381/509091/Significance-of-Blood-Vessel-Leakiness-in-Cancer1>
- [241] W. Timp and G. Timp, "Beyond mass spectrometry, the next step in proteomics," *Sci Adv*, vol. 6, no. 2, p. eaax8978, Jan. 2020, doi: 10.1126/sciadv.aax8978.
- [242] X. Yin *et al.*, "Unveiling the impact of CDK8 on tumor progression: mechanisms and therapeutic strategies," *Front Pharmacol*, vol. 15, p. 1386929, Mar. 2024, doi: 10.3389/FPHAR.2024.1386929/BIBTEX.
- [243] M. Venere *et al.*, "The Mitotic Kinesin KIF11 is a Central Driver of Invasion, Proliferation, and Self Renewal in Glioblastoma," *Sci Transl Med*, vol. 7, no. 304, p. 304ra143, Sep. 2015, doi: 10.1126/SCITRANSLMED.AAC6762.
- [244] G. Gampa *et al.*, "Enhancing Brain Retention of a KIF11 Inhibitor Significantly Improves its Efficacy in a Mouse Model of Glioblastoma," *Scientific Reports 2020 10:1*, vol. 10, no. 1, pp. 1–13, Apr. 2020, doi: 10.1038/s41598-020-63494-7.
- [245] E. Terribas *et al.*, "KIF11 and KIF15 mitotic kinesins are potential therapeutic vulnerabilities for malignant peripheral nerve sheath tumors," *Neurooncol Adv*, vol. 2, no. Supplement_1, pp. i62–i74, Jun. 2020, doi: 10.1093/NOAJNL/VDZ061.
- [246] R. Tang *et al.*, "Multiplexed screens identify RAS paralogs HRAS and NRAS as suppressors of KRAS-driven lung cancer growth," *Nat Cell Biol*, vol. 25, no. 1, p. 159, Jan. 2023, doi: 10.1038/S41556-022-01049-W.
- [247] R. F. H. Walter *et al.*, "Bortezomib sensitivity is tissue dependent and high expression of the 20S proteasome precludes good response in malignant pleural mesothelioma," *Cancer Manag Res*, vol. 11, p. 8711, 2019, doi: 10.2147/CMAR.S194337.
- [248] W. He *et al.*, "PSMB2 plays an oncogenic role in glioma and correlates to the immune microenvironment," *Sci Rep*, vol. 14, no. 1, p. 5861, Dec. 2024, doi: 10.1038/S41598-024-56493-5.
- [249] X. Xia *et al.*, "Neddylation of HER2 Inhibits its Protein Degradation and promotes Breast Cancer Progression," *Int J Biol Sci*, vol. 19, no. 2, p. 377, 2023, doi: 10.7150/IJBS.75852.
- [250] D. Liu, X. Che, and G. Wu, "Deciphering the role of neddylation in tumor microenvironment modulation: common outcome of multiple signaling pathways," *Biomarker Research 2024 12:1*, vol. 12, no. 1, pp. 1–24, Jan. 2024, doi: 10.1186/S40364-023-00545-X.