

Genetic screening to investigate treatment resistance in MLL-rearranged leukemia



Maximilian G Jamilly
Lincoln College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2019

« Est-ce que nous voyons la cent-millième partie de ce qui existe ? Tenez, voici le vent... l'avez-vous vu et pouvez-vous le voir ! Il existe, pourtant ! »
— *Le Horla*, Guy de Maupassant

Acknowledgements

I would like to thank past and present members of the Fulga lab – Markus, David, Áron, Alice, Andrew, Hector, Muhammad, Oana, Clara, Yale, Quentin and Tony – for their insight and support over the last three years. Yale was a selfless mentor, dive buddy and friend. His scientific curiosity is endless and his comments on this thesis were invaluable. I am especially grateful to the Ribonucleotide Teens: Alice, athlete and embroiderer extraordinaire, who has helped me throughout and provided generous and helpful feedback on the entire thesis; Andrew "Le Bron" Ramos, meme connoisseur; and Hector S. Barbosa, the PCR doctor, without whom life at the WIMM would not have been the same.

I am also grateful to members of the Milne Lab – Phill, Nick, Laura, I-Jun, Dmitra, Joe, Tom, Aini and Alastair – for welcoming me into their tissue culture and always being willing to donate their reagents and their time. Phill deserves a special mention for teaching me the subtleties and frustrations of THP-1 cell culture; Nick, for helping me with NGS; and Joe, for his bioinformatic expertise.

I would like to thank the members of my thesis committee – Prof Andrew Bassett, Prof Adam Mead, Prof Jan Rehwinkel and Dr Philip Hublitz – for their time, as well as Prof Marella de Bruijn for serving as my Director of Graduate Studies, Prof Ahmed Ahmed for examining my transfer of status and Prof Antonis Papachristodoulou and Prof Robert Carlisle for running the Synthetic Biology Centre for Doctoral Training.

I owe much to my early scientific mentors, especially Dr Vicky Stubbs and Dr Andrew Szydlo from Highgate School who started me on this journey. I fondly remember Charles Ross, my long-time friend, mentor, polymath and fellow-Cholmeleian.

For guiding me into the world of CRISPR screening, I am very grateful to Dr Alice Mayer and Dr Ryan Beveridge. Ryan was critical to the success of the screening project and always generous with his time and expertise, not to mention the excellent virus preps.

I am fortunate to have had the constant support of my family: my father, Jonathan; my sister, Joanna; my aunt, Kim; and my uncles, Michel and DJ. They mean the world to me.

I am grateful to my friends in Oxford and beyond, especially Blake and Beth who are penpals *par excellence*. Luis helped me through many long days in the

library with friendship and good humour: I am indebted to him and Christian for their comments on my thesis.

Finally, I owe a debt of gratitude to my two supervisors, Prof Tudor Fulga and Prof Thomas Milne, for their skilful guidance over the past three years. They have taught me a great deal.

Declaration

This thesis and the work presented herein are my own unless otherwise specified. All work was carried out at the Weatherall Institute of Molecular Medicine, University of Oxford. This thesis is submitted for the degree of DPhil in Medical Sciences and has not been submitted for any other degree.

The following experiments were conducted in collaboration with others:

1. All lentivirus preps were produced by Ryan Beveridge (Viral Production Facility, Weatherall Institute of Molecular Medicine, Oxford)
2. The TKOv3 plasmid library was expanded and sequenced by Ryan Beveridge. I analysed the sequencing data.
3. Nicholas Crump (Molecular Hematology Unit, Weatherall Institute of Molecular Medicine, Oxford) assisted with quantifying amplicon libraries for NextSeq.

The analysis in Figure 5.6B was performed by Joe Harman (Molecular Hematology Unit, Weatherall Institute of Molecular Medicine, Oxford).

This work was funded by the University of Oxford and the EPSRC & BBSRC Centre for Doctoral Training in Synthetic Biology (grant EP/L016494/1).

Publications

Works published during this thesis:

1. Knapp, D.J.H.F., Michaels, Y.S., **Jamilly, M.**, Ferry, Q.R. V., Barbosa, H., Milne, T.A., and Fulga, T.A. (2019). Decoupling tRNA promoter and processing activities enables specific Pol-II Cas9 guide RNA expression. *Nat. Commun.* 10, 1490.

Abstract. Spatial/temporal control of Cas9 guide RNA expression could considerably expand the utility of CRISPR-based technologies. Current approaches based on tRNA processing offer a promising strategy but suffer from high background. Here, to address this limitation, we present a screening platform which allows simultaneous measurements of the promoter strength, 5', and 3' processing efficiencies across a library of tRNA variants. This analysis reveals that the sequence determinants underlying these activities, while overlapping, are dissociable. Rational design based on the ensuing principles allowed us to engineer an improved tRNA scaffold that enables highly specific guide RNA production from a Pol-II promoter. When benchmarked against other reported systems this tRNA scaffold is superior to most alternatives, and is equivalent in function to an optimized version of the Csy4-based guide RNA release system. The results and methods described in this manuscript enable avenues of research both in genome engineering and basic tRNA biology.

International conferences attended:

1. Pooled CRISPR screening to investigate resistance to the leukemia drug venetoclax. **Max Jamilly**, Ryan Beveridge, Thomas A Milne, Tudor A Fulga. *Poster*. Keystone Symposia meeting on Genome Engineering: From Mechanisms to Therapies. Feb 2019.

Abstract. The ability of malignant cells to evade apoptosis is a key feature of leukemia. The intrinsic apoptosis pathway has, therefore, become a major focus for targeted cancer therapies. In some blood cancers, oncogenic mutations lead to overexpression of antiapoptotic factors such as BCL-2. Venetoclax, a BH3-mimetic BCL-2 inhibitor, is clinically approved for

treatment of chronic lymphocytic leukemia (CLL). CLL depends heavily on BCL-2, whereas acute myeloid and lymphoid leukemia (AML and ALL) also depend on additional antiapoptotic factors such as BCL-XL and MCL-1. Correspondingly, early clinical data indicate that AML/ALL patients are refractory to venetoclax monotherapy but may respond to combination therapies. We reasoned that a better understanding of the mechanism for venetoclax resistance in AML/ALL would enable the development of a broader range of combination therapies. We set out to conduct an unbiased survey of all protein-coding genes involved in venetoclax resistance in an AML cell line. We began by generating CRISPR/Cas9 knockout cell lines for each of six apoptotic genes and characterised their response to venetoclax. Next, we conducted a pooled CRISPR knockout experiment by transducing cells with a lentiviral sgRNA library targeting the same six genes. We cultured transduced cells under venetoclax selection, then sequenced to determine the final sgRNA distribution. This showed that loss of antiapoptotic and proapoptotic proteins increases sensitivity and resistance to venetoclax respectively. Now we are conducting a pooled whole-genome CRISPR knockout screen based on cell proliferation under venetoclax selection. This powerful approach will allow us to reverse-engineer the underlying causes of resistance to a blockbuster leukemia drug and search for venetoclax combination therapies in AML/ALL. Pooled CRISPR drug screening is an exciting technology for discovering targeted therapies for previously untreatable diseases.

Abstract

Despite the promise of targeted cancer therapies, some subsets of acute myeloid leukemia (AML) and acute lymphoid leukemia (ALL) continue to present a dismal prognosis. Translocations in the *mixed lineage leukemia* (*MLL*) gene generate MLL fusion proteins and cause *MLL*-rearranged (MLLr) leukemia, an aggressive subtype of AML and ALL which is common in infants and children. MLLr leukemia is resistant to conventional chemotherapy. I used CRISPR/Cas9-mediated mutagenesis to investigate miRNA function and BH3 mimetic drug sensitivity in MLLr leukemia. I began by testing a published model for the role of a non-coding RNA in MLLr leukemia. The MLL-AF9 fusion protein activates the transcription factors *HOXA9* and *MEIS1*, which are both predicted targets of the oncogenic miRNA miR-196b. I used dual-sgRNA CRISPR/Cas9 knockout to generate and characterise a miR-196b null mutant MLL-AF9 cell line, demonstrating that miR-196b ablation reduces colony-forming capacity but does not affect expression of *HOXA9* or *MEIS1*. Next, I investigated the genetic basis for venetoclax resistance in MLLr leukemia. Venetoclax is a potent inhibitor of the anti-apoptotic protein BCL-2 and is an effective monotherapy in CLL but not in AML or ALL. I optimised a system of pooled CRISPR/Cas9 editing to compare the contributions of BCL-2 family proteins to venetoclax resistance in MLL-AF9 cells. Finally, I used this system to conduct a pooled genome-wide CRISPR/Cas9 screen in MLL-AF9 cells. My results suggest that the 26S proteasome mediates venetoclax resistance in MLLr leukemia. I provide evidence that the proteasome inhibitor bortezomib synergises with venetoclax to kill MLL-AF9 cells *in vitro*. The work presented in this thesis demonstrates that CRISPR-Cas9 screening is a powerful tool for investigating the causes of resistance to targeted cancer therapy.

Contents

List of Figures	xvii
List of Tables	xix
List of Abbreviations	xxi
1 Introduction	1
1.1 Motivation	2
1.2 Epidemiology and etiology of MLLr leukemia	2
1.2.1 <i>MLL</i> -rearranged leukemia is associated with poor treatment outcomes	2
1.2.2 <i>MLL</i> fusion proteins cause leukemogenesis by disrupting transcriptional regulation	4
1.3 Homeobox genes and miRNAs: key regulators in MLLr leukemia . .	10
1.3.1 Homeobox genes are involved in development and hematopoiesis	10
1.3.2 <i>HOXA9</i> and <i>MEIS1</i> are drivers of MLLr leukemogenesis . .	11
1.3.3 <i>MLL</i> fusion proteins specifically upregulate expression of <i>HOXA9</i> and <i>MEIS1</i>	13
1.3.4 Oncogenic and tumour-suppressor miRNAs are associated with progression of <i>MLL</i> -rearranged leukemia	15
1.3.5 miR-196b is an MLLr leukemia-associated oncomir which may interact with <i>HOXA</i> genes and <i>MEIS1</i>	16
1.3.6 There is no accepted model for the interaction between homeobox genes and miR-196b in <i>MLL</i> -rearranged leukemia . . .	19
1.3.7 THP-1 cells are a model system for <i>MLL</i> -AF9 AML	20
1.3.8 MLLr leukemia treatments face many challenges	21
1.4 Programmed cell death in MLLr leukemia	22
1.4.1 Evading apoptosis is a hallmark of many cancer types	22
1.4.2 <i>BCL-2</i> is a key regulator of programmed cell death	23
1.4.3 <i>MLL</i> fusion genes disrupt intrinsic apoptosis	26
1.4.4 Venetoclax is a highly effective <i>BCL-2</i> inhibitor drug	26
1.4.5 Many AML/ALL patients are resistant to venetoclax	30

1.4.6	Treatment of MLLr leukemia with venetoclax needs rational combination therapies	31
1.5	Genome-wide CRISPR screening	33
1.5.1	RNAi and CRISPR are powerful technologies for forward and reverse genetics in mammalian cells	33
1.5.2	CRISPR/Cas9 is used for pooled genome-wide screening	38
1.5.3	Pooled CRISPR/Cas9 screens enable target discovery in leukemia	40
1.6	Summary	42
1.6.1	Thesis overview	42
1.6.2	Non-technical summary	43
2	Methods and materials	47
2.1	Cloning	47
2.2	miR-196b knockout THP-1 cell lines	48
2.3	Human cell culture and electroporation	49
2.4	Cell viability assays	49
2.5	CRISPR knockout validation	51
2.6	Lentiviral transduction	54
2.7	Screening optimisation	55
2.8	Pooled miniscreen	56
2.9	Genome-wide CRISPR screen	61
2.10	Data analysis	67
3	Investigating the role of miR-196b in THP-1 cell proliferation	69
3.1	Introduction	69
3.2	Results	70
3.2.1	CRISPR knockout of miR-196b in THP-1 cells	70
3.2.2	Generating and validating miR-196b knockout THP-1 cell lines	71
3.2.3	Phenotypic analysis of miR-196b knockout THP-1 cells	73
3.2.4	A fluorescent reporter system for measuring miR-196b activity	75
3.2.5	The link between miR-196b and intrinsic apoptosis	77
3.3	Discussion	78
3.3.1	miRNA sequences can be deleted using dual-sgRNA CRISPR	78
3.3.2	miR-196b promotes clonogenicity of THP-1 cells	80
3.3.3	<i>HOXA9</i> , <i>MEIS1</i> and <i>FAS</i> are not regulated by miR-196b in THP-1 cells	80
3.3.4	miR-196b knockout does not affect venetoclax resistance in THP-1 cells	83
3.3.5	The role of ncRNAs in MLLr leukemia remains uncertain	83

4	Intrinsic apoptosis and venetoclax resistance	85
4.1	Introduction	85
4.2	Results	87
4.2.1	Venetoclax resistance in leukemia cell lines	87
4.2.2	Optimising transduction and puromycin selection in THP-1 cells	87
4.2.3	CRISPR knockout of BCL-2 family genes in THP-1 cells	91
4.2.4	Venetoclax miniscreen in THP-1 cells	93
4.3	Discussion	101
4.3.1	THP-1 cells are an effective model for clinical resistance to VEN	102
4.3.2	Intrinsic apoptosis genes contribute to VEN resistance in MLLr leukemia	103
4.3.3	Pooled CRISPR screens are feasible in THP-1 cells	105
5	Genome-wide screening to discover venetoclax combination therapies	109
5.1	Introduction	109
5.1.1	Experimental design of CRISPR screens	110
5.1.2	Screening coverage	113
5.1.3	Analysing pooled CRISPR screens	114
5.2	Results	114
5.2.1	Genome-wide CRISPR screen	114
5.2.2	Finding druggable targets and combination therapies	121
5.3	Discussion	128
5.3.1	Maintaining sufficient coverage is essential in pooled CRISPR screens	128
5.3.2	Screen design impacts discovery rates	129
5.3.3	Analysis of pooled screens has not been standardised	131
5.3.4	Mitochondrial apoptosis may control VEN resistance in THP-1 cells	133
5.3.5	VEN–BTZ combination therapy may overcome VEN resistance in MLLr leukemia	134
5.3.6	Dropout hits suggest other potential combination therapies for AML/ALL	137
5.3.7	CRISPR screens offer new perspectives on VEN sensitivity	138
5.3.8	A high-throughput pipeline for combination therapy discovery	140

6	Discussion and future directions	141
6.1	A model for non-coding RNA and apoptosis in MLLr leukemia . . .	142
6.1.1	ncRNA in MLLr leukemia: driver or passenger?	142
6.1.2	A proposed pipeline for functional identification of oncogenic miRNAs	144
6.1.3	Current models for apoptosis in cancer are inadequate . . .	145
6.1.4	New approaches to treating MLLr leukemia	148
6.2	Pipelines for high-content drug resistance screening in primary cells	151
6.2.1	Pooled CRISPR screening methodologies are constantly im- proving	152
6.2.2	Automation will enable parallel high-throughput CRISPR screening	155
6.2.3	p53: a double-edged sword for CRISPR screening and cancer treatment	157
6.2.4	The future of drug resistance screening	159
6.3	Concluding remarks	160
 Appendices		
A	Oligonucleotide sequences	165
B	Plasmid maps	169
C	Western blots	173
References		177

List of Figures

1.1	11q23 translocations generate MLL fusion proteins	9
1.2	The <i>MIR196B</i> locus lies within the <i>HOXA</i> cluster	17
1.3	A model for miR-196b and HOXA9/MEIS1/FAS in MLL-AF9 leukemia	20
1.4	Intrinsic and extrinsic apoptosis	23
1.5	Intrinsic apoptosis and the BCL-2 family	27
1.6	Methods for gene perturbation using RNAi or CRISPR	35
1.7	Lentiviral vector-mediated RNAi or CRISPR	38
1.8	Pooled CRISPR screening	40
2.1	qPCR primer validation	53
2.2	NGS PCR1 optimisation	65
3.1	CRISPR-KO of miR-196b	72
3.2	Generating miR-196b KO cell lines	74
3.3	Proliferation and gene expression in miR-196b KO THP-1 cells	76
3.4	Venetoclax sensitivity of miR-196b KO THP-1 cells	77
3.5	Chromatin landscape surrounding <i>MIR196B</i>	79
4.1	The intrinsic apoptosis pathway	86
4.2	Comparing venetoclax resistance in leukemia cell lines	88
4.3	Puromycin selection and lentiviral titre	89
4.4	Observed and predicted lentiviral titre	92
4.5	Generating dual-sgRNA knockouts of BCL-2 family genes	94
4.6	Proliferation of BCL-2 family knockout THP-1 cells in VEN	95
4.7	IC ₅₀ (VEN) of BCL-2 family knockout THP-1 cells	96
4.8	Pooled CRISPR-KO miniscreen targeting BCL-2 family	97
4.9	Initial distribution of miniscreen sgRNAs	98
4.10	Fitness effects of miniscreen sgRNAs without selection	99
4.11	Effect of venetoclax selection on miniscreen sgRNA abundance	101
5.1	Screening with the TKOv3 library	116
5.2	Genome-wide screening data analysis	118
5.3	Comparing sgRNA read counts from replicate screens	119

5.4	Genes affecting VEN resistance in THP-1 cells	122
5.5	sgRNA-level selection for key depleted and enriched genes	123
5.6	Functional validation of screening hits	124
5.7	Functional analysis of screening hits	125
5.8	Venetoclax-bortezomib combination assays	127
6.1	A proposed pipeline for oncomir discovery	146
6.2	MLL fusion proteins and apoptosis	149
6.3	High-throughput screening for leukemia drug discovery	160

List of Tables

1.1	Common MLL translocation partner genes	7
2.1	Antibodies used for western blotting	54
2.2	PCRs for pooled screen NGS library preparation	59
2.3	Primers for pooled miniscreen NGS library preparation	60
2.4	Primers for TKOv3 plasmid NGS library preparation	62
2.5	PCRs for TKOv3 plasmid NGS library preparation	63
2.6	NGS primers for genome-wide screening	65
4.1	IC ₅₀ (VEN) values reported in the literature	103
5.1	Popular libraries for genome-wide CRISPR-KO screens	111
5.2	Screening steps critical for library coverage.	113
5.3	Genome-wide screening coverage	115

List of Abbreviations

11q23		chromosome 11 band 23
AGO		Argonaute family
ALL		acute lymphoblastic leukemia
AML		acute myeloid leukemia
BAGEL		Bayesian analysis of gene essentiality
BH		BCL-2 homology
bp		base pairs
BTZ		bortezomib
CLL		chronic lymphocytic leukemia
CNV		copy-number variation
CRISPR	. . .	clustered regularly interspaced short palindromic repeats
CRISPR-KO	.	CRISPR knockout
CRISPRa	. . .	CRISPR activation
CRISPRi	. . .	CRISPR inhibition
CTG		CellTiter-Glo
dCas9		dead Cas9
DISC		death-inducing signalling complex
DMSO		dimethyl sulfoxide
DSB		double strand break
dsDNA		double stranded DNA
EAP		elongation-assisting proteins
EFS		event-free survival
EGFP		enhanced green fluorescent protein
FACS		fluorescence-activated cell sorting
FDR		false discovery rate

FISH	fluorescence in-situ hybridisation
FPG	fusion partner gene
gDNA	genomic DNA
GFP	green fluorescent protein
GO	gene ontology
GRN	gene regulatory network
H3K4me3	histone 3 lysine 4 trimethyl
HDR	homology-dependent repair
HOX	homeobox
HSC	hematopoietic stem cell
ITD	internal tandem duplication
Kb	kilobases
kDa	kilodaltons
LEDGF	lens epithelium-derived growth factor
LMI	LEDGF-Menin interaction
LSC	leukemic stem cell
mAb	monoclonal antibody
MAGeCK	model-based analysis of genome-wide CRISPR/Cas9 knockout
MFI	median fluorescent intensity
miRNA, miR	micro RNA
MLL	<i>mixed lineage leukemia</i>
MLLr	MLL-rearranged
MOI	multiplicity of infection
MOM	mitochondrial outer membrane
MOMP	mitochondrial outer membrane permeabilisation
MPAL	mixed-phenotype acute leukemia
MRE	miRNA response element
NGFR	nerve growth factor receptor
NGS	next-generation sequencing
NHEJ	non-homologous end joining
nt	nucleotide

PAM		protospacer-adjacent motif
PBS		phosphate-buffered saline
PCR		polymerase chain reaction
PHD		plant homedomain
pre-miRNA	. .	precursor miRNA
pri-miRNA	. .	primary miRNA
R/R		relapsed/refractory
RACE		rapid amplification of cDNA ends
RISC		RNA-induced silencing complex
RNA pol II	. .	RNA polymerase II
RNAi		RNA interference
RRA		robust rank aggregation
RT-qPCR		quantitative reverse-transcription PCR
SD		standard deviation
SEM		standard error of the mean
sgRNA		single guide RNA
shRNA		short hairpin RNA
siRNA		small interfering RNA
SpCas9		<i>S. pyogenes</i> Cas9
ssDNA		single stranded DNA
t-AML		therapy-related AML
T7E1		T7 endonuclease
TALEN		TALE effector nuclease
TKOv3		Toronto knockout version 3
TU		transducing units
UPR		unfolded protein response
UTR		untranslated region
VEN		venetoclax
WT		wild type
ZFN		zinc-finger nuclease
ZIP		zero interaction potency

1

Introduction

Contents

1.1	Motivation	2
1.2	Epidemiology and etiology of MLLr leukemia	2
1.3	Homeobox genes and miRNAs: key regulators in MLLr leukemia	10
1.4	Programmed cell death in MLLr leukemia	22
1.5	Genome-wide CRISPR screening	33
1.6	Summary	42

In this thesis, I use CRISPR/Cas9-based genetic screening technology to investigate resistance to targeted anticancer therapies in *MLL*-rearranged (MLLr) leukemia. I focus on two hallmarks of MLL-rearranged leukemia cells: loss of control of *HOX* gene expression and dysfunctional mitochondrial apoptosis.

Firstly, I use CRISPR to generate knockout cell lines and challenge a published model describing the interaction between *HOX* genes and an oncogenic miRNA in MLLr leukemia. Secondly, I expand this CRISPR platform to conduct a reverse genetic screen testing the effect of known mitochondrial apoptosis genes on resistance to venetoclax, a blockbuster leukemia drug. Thirdly, I conduct a genome-wide forward CRISPR screen to discover genes involved in venetoclax resistance. Finally, I test a candidate combination therapy predicted to overcome venetoclax resistance.

This introduction begins with a survey of acute leukemia, focusing on the causes and effects of *MLL* rearrangement. I introduce homeobox genes and oncogenic miRNAs and describe theories and outstanding questions about their role in MLLr leukemia. Next I turn to intrinsic apoptosis, a hallmark of cancer whose misregulation is characteristic of acute leukemia. After an overview of the main regulatory proteins involved in intrinsic apoptosis and their relevance to MLLr leukemia, I describe recent efforts to treat leukemia using BH3 inhibitor drugs. Then I review the successes and failures of BH3 inhibitor drugs in MLL leukemia. Finally, I introduce genome-wide CRISPR/Cas9 screening as a powerful new technology for reverse engineering the causes of cancer drug resistance.

1.1 Motivation

The motivation for this thesis was threefold:

1. Develop platforms for forward and reverse genetic screening to investigate gene and protein regulation in MLLr leukemia
2. Understand the contribution of miR-196b to tumorigenesis and proliferation in MLLr leukemia
3. Identify and inhibit proteins involved in clinical resistance to an apoptosis-inducing drug in MLLr leukemia

1.2 Epidemiology and etiology of MLLr leukemia

1.2.1 *MLL*-rearranged leukemia is associated with poor treatment outcomes

Leukemia arises due to aberrant development of hematopoietic stem cells (HSCs)¹. Acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL) are two highly aggressive subtypes of leukemia associated with incorrect cell-fate specification in the myeloid or lymphoid lineages. In the USA, ALL is the most frequent cause of death among all cancer types in patients younger than 20 years old².

Approximately 10% of all cases of ALL and AML are associated with translocations in the *MLL* (*mixed lineage leukemia*) gene³. Despite the advent of modern targeted anticancer therapies, AML/ALL patients with *MLL* translocations still face a particularly dismal prognosis⁴.

MLL translocations are found in >70% of *de novo* infant AML and ALL cases⁵ but they are much rarer in adult *de novo* acute leukemia, accounting for only 3-4% of adult cases of AML and 3-7% of adult cases of ALL⁶. *MLL* translocations are more common in therapy-related leukemia: 10% of adult therapy-related AML (t-AML) patients have *MLL* translocations, normally caused by treatment with topoisomerase II inhibitors⁷⁻¹⁰. Sometimes t-AML also arises in infants and children. Some patients with *MLL*-rearranged (MLLr) leukemia have the symptoms of single-lineage AML or ALL; some display symptoms and biomarkers of both diseases (mixed phenotype acute leukemia, MPAL); and some exhibit lineage switching between AML and ALL over time^{4,11}.

The most significant adverse prognostic factor in MLLr leukemia is age¹². Infants below the age of one with MLLr AML or ALL have a particularly poor prognosis, perhaps due to a combination of inferior dosing schedules and fundamental differences in gene expression in infant hematopoietic cells¹².

Treatment outcomes for infant ALL have not significantly improved in the past two decades and *MLL* translocations significantly worsen the prognosis¹³. The Interfant Study Group is a collaborative global effort to coordinate clinical trials in infant ALL¹⁴. Interfant-99, a phase III trial which began recruitment in 1999, established the current standard of care for ALL in infants younger than one year¹⁵. Infants are treated with a hybrid chemotherapy schedule that combines treatment strategies for AML and ALL. Some infants in a high-risk group also respond favourably to hematopoietic stem cell transfer. Interfant-06, a follow-on from Interfant-99, was designed to test the benefit of early treatment intensification (aimed at preventing relapse) and additional AML-like chemotherapy. Unfortunately, across the cohort of 651 infants, outcomes with the Interfant-06 protocol were no

better than Interfant-99. Six-year overall survival was 74% for wild-type *MLL* patients compared to only 36% for *MLL*-rearranged patients.

Results from a US-based study group, the Children's Oncology Group (COG), reflect a similar situation. Among 147 infants with ALL subjected to a typical treatment regime, infants older than 90 days experienced a five-year event-free survival rate (EFS) of ~44% for MLLr ALL, compared to ~69% for non-MLLr ALL. EFS for infants younger than 90 days with MLLr ALL was only ~16%¹⁶. MLLr patients are also more likely to suffer relapse than non-MLLr patients^{4,17}.

In AML, unlike ALL, the prognostic impact of *MLL* translocations is sometimes favourable. Factors such as patient karyotype, patient age and treatment regime are associated with a wide range of survival rates. Many studies have disagreed about the prognostic value of different MLLr karyotypes, probably due to statistical flaws arising from the small number of patients within each study cohort. Thus, overall five-year EFS for *de novo* pediatric non-MLLr AML is ~55%¹⁸, compared to five-year EFS of up to ~92% for some *MLL* translocations or only ~11% for others¹⁹. As with ALL, the success rate of AML treatment has steadily improved over time but some MLLr karyotypes remain highly refractory²⁰. Will a better understanding of the downstream effects of *MLL* rearrangement lead to improved therapies for adult and pediatric AML and ALL?

1.2.2 MLL fusion proteins cause leukemogenesis by disrupting transcriptional regulation

Lysine methyltransferase 2A (*KMT2A*), also known as mixed lineage leukemia (*MLL* or *MLL1*), is the human homologue of the *Drosophila* gene *trithorax*²¹. Wild-type *MLL* encodes a large (~500 kDa) multidomain nuclear protein which is posttranslationally cleaved by Taspase1 to yield two subunits, the *N*- and *C*-terminal subunits^{22,23}. These subunits constitutively combine to form a holocomplex which engages in a vast array of interactions with DNA, chromatin and other proteins to mediate both epigenetic modification and direct transcriptional regulation (Figure 1.1B)^{4,21,22,24,25}.

The 320 kDa *N*-terminal subunit contains an LEDGF-Menin interaction (LMI) domain, three DNA-binding AT hook domains and the CxxC domain^{21,23}. The LMI domain recruits Menin and lens epithelium-derived growth factor (LEDGF) to bind active chromatin^{23,26}. The CxxC domain binds unmethylated CpG and indirectly recruits activating ubiquitinases²⁷. After the CxxC domain comes the breakpoint region, where MLL translocations occur, then four plant homeodomain (PHD) fingers which bind trimethylated histone 3 lysine 4 (H3K4me3) and lastly a bromodomain^{23,28}. The CxxC domain and PHD fingers mediate recruitment of wild-type MLL to the *HOXA9* locus (discussed below)²⁷. In MLL fusion proteins, the PHD fingers and bromodomain are always lost²³.

The 180 kDa *C*-terminal subunit of wild-type MLL is involved in chromatin modification²³. Firstly, a transactivation domain recruits histone acetyltransferases^{22,25}. Secondly, a highly conserved SET domain has histone 3 lysine 4 (H3K4) methyltransferase activity at *HOX* gene loci²⁹. In turn, H3K4me binds MLL via the *N*-terminal PHD fingers, leading to positive feedback after initial MLL recruitment^{23,28}. Inactivating the H3K4 methyltransferase activity of the SET domain is not lethal in mice but loss of the entire *C*-terminal subunit is lethal, suggesting that the MLL holocomplex has an essential role in embryogenesis that goes beyond its methyltransferase activity³⁰. It is probable that wild-type MLL has a global role in positive transcriptional regulation as well as a more specific role in positive regulation of *HOX* genes during embryogenesis and adult hematopoiesis^{4,21,24}.

The *MLL* locus lies on the long arm of chromosome 11 at band 23 (11q23). Chromosomal rearrangements at 11q23 almost always lead to rearrangements of *MLL*³¹, which forms in-frame fusions with at least 84 different genes^{32,33} (Figure 1.1A). Studies of monozygotic twins have been a useful tool for understanding the natural history of infant MLLr leukemia³⁴. *MLL* rearrangements normally arise *in utero* early during fetal development and lead to leukemia with very low latency^{35,36}.

MLL appears to be so susceptible to translocation because the 11q23 locus is also involved in genome topology. In most patient samples, *MLL* breakpoints are concentrated between two scaffold-associated regions. Moreover, the *MLL* locus

contains an unusually high density of DNA cleavage sites and topoisomerase II binding sites³⁷. This may explain the frequency of *MLL* rearrangements in t-AML following topoisomerase II inhibitor treatment^{8,38}. Correspondingly, topoisomerase II inhibitors induce *MLL* breakpoints *in vitro* in malignant cell lines and healthy monocytes³⁹. An internal RNA polymerase II promoter in the 11q23 locus has also been associated with increased translocation³². It is unclear which of these features has the greatest effect on the frequency of *MLL* breakpoints³². More generally, it is not known whether high clinical incidence of MLLr fusion proteins is due to the fact that 11q23 translocations are very common or that the resulting fusion proteins are highly oncogenic and therefore selected.

The most common *MLL* translocation partner genes are shown in Table 1.1. The frequency of each *MLL* fusion protein varies considerably with age and MLLr leukemia subtype³³. AF4, AF9, ENL and AF10 are all nuclear proteins, although other less common fusion partners are cytoplasmic proteins⁴. The leukemogenic activity of MLLr fusion proteins is dictated by the protein domains retained from wild-type *MLL* – normally the LMI domain, AT hook and CxxC – and by the activity of the fusion partner (Figure 1.1B)^{21,23}. A landmark structure-function analysis study of *MLL* demonstrated that the LMI and CxxC domains are necessary for *MLL*-ENL-mediated transformation of murine myeloid cells⁴⁰. A later study showed that these two domains are sufficient for the leukemogenic activity of *MLL*-AF4⁴¹.

It has proven difficult to ascertain the mechanistic differences between different *MLL* fusion proteins, and harder still to connect the genotype of MLLr leukemia patients to their clinical symptoms. Studies in mouse and human models have long contradicted each other. For example, it remains unclear which HSC compartment is responsible for initiation of *MLL*-AF4 ALL^{42–44}. Furthermore, 11q23 translocations do not always cause *MLL* fusions. Instead, some translocations lead to in-frame partial tandem duplications (PTDs) of *MLL* exons 5–12. This occurs in ~7.5% of AML cases, primarily in adults⁴⁵.

MLL fusion proteins are remarkably leukemogenic. Retroviral transduction of *MLL*-ENL immortalises myeloid progenitor cells⁴⁶. Ectopic expression of *MLL*-AF9⁴⁷

translocation partner gene	translocation	% frequency	
		MLLr ALL	MLLr AML
<i>AF4</i> (<i>AFF1</i>)	t(4;11)(q21;q23)	60	1
<i>AF9</i> (<i>MLLT3</i>)	t(9;11)(p22;q23)	12	29
<i>ENL</i> (<i>MLLT1</i>)	t(11;19)(q23;p13.3)	18	4
<i>AF10</i> (<i>MLLT10</i>)	t(10;11)(p12;q23)	3	15

Table 1.1: The most common translocation partner genes in MLLr leukemia. Fusions between MLL and over 70 different translocation partner genes have been described in MLLr leukemia³². The frequency of the four most common MLL fusions, shown here, differs between MLLr ALL and AML³². Percentages represent adult, child and infant cases combined.

is sufficient for leukemic transformation in mouse models. MLL-AF4 is sufficient for initiation⁴⁸ and necessary for maintenance⁴⁹ of leukemia. Clonal analysis of infant MLLr AML patients reveals a lower rate of somatic coding mutations than almost all other human cancers, although it remains controversial whether MLL fusion proteins alone drive the initiation and maintenance of malignant hematopoiesis. The most common additional somatic mutations observed in MLLr AML are in *FLT3*, *KRAS*⁵⁰, *NRAS* and components of the PI3K pathway⁵¹. These proto-oncogenes are all involved in growth control signalling.

Yeast two-hybrid studies⁵² and biochemical purification⁵³ showed that AF4, AF9 and ENL belong to a large multiprotein complex of ENL-associated proteins or elongation-assisting proteins (EAPs)^{53,54}. In association with MLL fusion proteins, the EAPs appear to induce transcriptional activation by two key mechanisms (Figure 1.1C).

Firstly, the EAPs promote RNA polymerase II transcription. Key members of the EAPs include the transcription elongation factors pTEFb (which phosphorylates the C-terminal domain of RNA pol II) and ELL (itself another MLL translocation partner)^{21,54–57}. Secondly, the EAPs induce activating chromatin modifications. AF4, AF9, ENL and AF10 all interact with the H3K79 methyltransferase DOT1L, which enhances transcription by catalysing H3K79 dimethylation^{53,54,58–63}. DOT1L or Menin inhibition prevents MLL fusion protein-induced leukemogenesis in mouse models^{21,64,65}. Maethner et al. demonstrated that, alongside activation of DOT1L

and pTEFb, MLL-ENL also indirectly inactivates the Polycomb proteins, preventing the formation of repressive heterochromatin⁶⁶. DOT1L, AF9 and ENL also appear to interact with components of the Wnt signalling pathway, which has been implicated in leukemogenesis⁵⁷. It is unclear whether the two functions of the EAP correspond to two different subcomplexes, sometimes separately referred to as the super-elongation complex (SEC) and DOT1L-complex (DotCom)^{21,23,66}. In general, there remains considerable controversy surrounding the composition and function of the many protein complexes associated with MLL fusions, as discussed comprehensively in several excellent recent reviews^{4,21,23,62,67}. Many basic questions remain unanswered about mechanistic differences between common MLL fusion proteins.

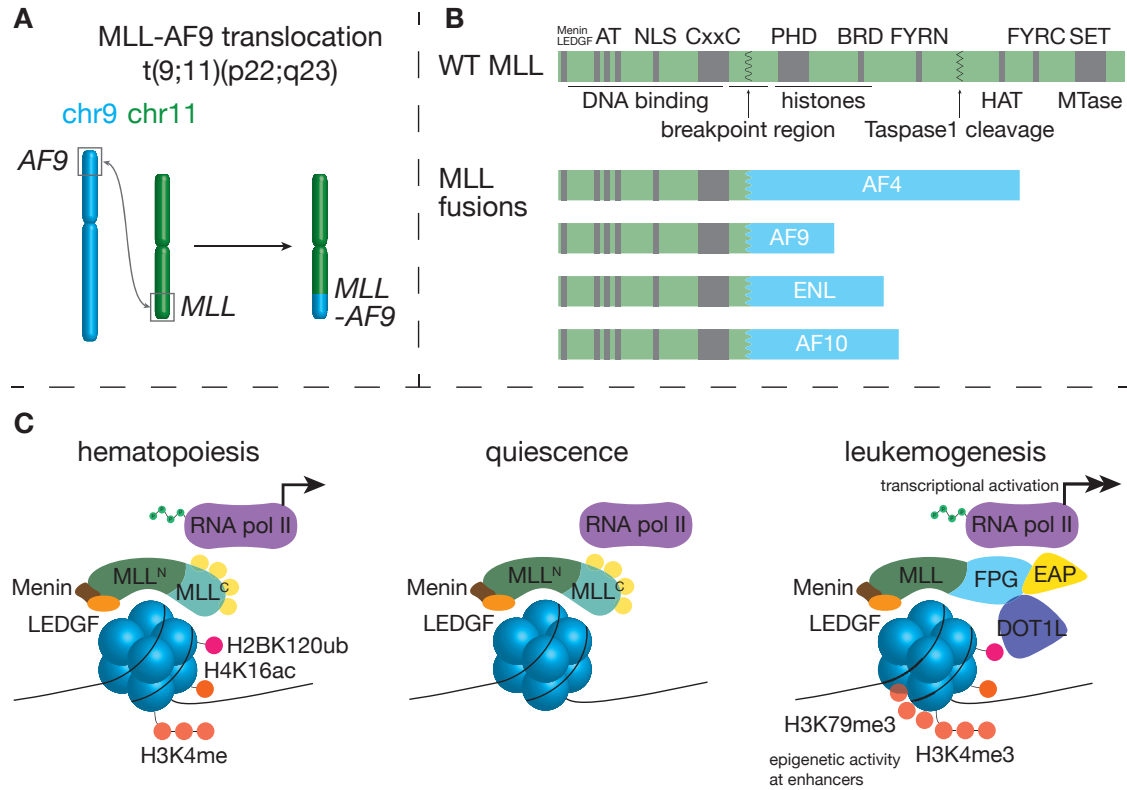


Figure 1.1: 11q23 translocations generate MLL fusion proteins. (A) MLL-AF9 is a common MLL fusion protein created when the t(9;11)(p22;q23) translocation leads to an in-frame fusion between the 5' end of *MLL* and the 3' end of *AF9* (*MLLT3*). The resulting MLL-AF9 fusion protein is found in MLLr AML and ALL patients. Other in-frame translocations are generated via similar mechanisms (see Table 1.1). (B) Simplified representation of functional domains in wild-type (WT) MLL protein and common fusions. Functions include DNA binding, histone binding, interaction with the transcription machinery and catalytic activity. At the *N*-terminus are Menin- and LEDGF-binding motifs, DNA-binding AT hook domains and the CxxC domain which binds unmethylated CpG and recruits activating ubiquitinases. The NLS is a group of nuclear localisation sequences. The breakpoint region is the site of common translocations. In MLL fusion proteins, most domains closer to the *C*-terminus than the breakpoint region are lost. The plant homeodomain (PHD) fingers interact with methylated chromatin; the bromodomains (BRD), with acetylated chromatin. FYRN and FYRC mediate heterodimerisation between the *N*- and *C*-terminal subunits after Taspase1 cleavage. A region between the Taspase1 cleavage site and FYRC interacts with three histone acetyltransferases (HATs). The *C*-terminal SET domain has methyltransferase (MTase) activity. *Adapted from Krivtsov et al.*²¹ (C) Simplified cartoon showing wild-type and malignant activity of MLL and MLL fusion proteins. During normal hematopoiesis, the *N*-terminal subunit (MLL^N) interacts with actively transcribed DNA and, via Menin and LDGF, with active chromatin. The *C*-terminal domain (MLL^C) interacts with HATs. Hyperphosphorylated RNA polymerase II is recruited, leading to transcription. During quiescence, MLL is repressed. During leukemogenesis, MLL fusion proteins (including a fusion partner gene (FPG)) interact with the elongation-assisting proteins (EAP) and H3K79 methyltransferase DOT1L to promote active chromatin, especially at *HOXA9* and *MEIS1* loci.

1.3 Homeobox genes and miRNAs: key regulators in MLLr leukemia

1.3.1 Homeobox genes are involved in development and hematopoiesis

In 1921, Morgan and Bridges postulated the existence of the homeotic genes, a set of genetic regulators that would control the entire *Drosophila* body plan during development⁶⁸. Thirty years later, researchers characterised the *bithorax* complex, a family of eight genes which control the patterning of *Drosophila* embryos⁶⁹. Structural analysis showed that all the *bithorax* genes, and many others besides, share a characteristic ~180bp sequence which encodes a DNA-binding domain⁷⁰. In reference to Morgan and Bridges' homeotic genes, the 180bp sequence was named the homeobox sequence; the domain was named the homeodomain. Comparative genomics has since demonstrated that genes containing the homeobox sequence, known as homeobox genes, are widely conserved among vertebrates and invertebrates, where they are involved in processes essential for embryonic development such as differentiation, apoptosis and angiogenesis.

In mammals, the *HOX* genes are a special subset of 39 highly conserved homeobox genes arranged in four distinct genomic clusters, *HOXA*, *HOXB*, *HOXC* and *HOXD*⁷¹. Another important subset of homeobox genes is the three amino acid loop extension (TALE) family. *MEIS1* and *PBX3* are TALE-family cofactors of *HOXA9*^{72,73}.

Homeobox genes such as *HOXA9* and *MEIS1* play a central role in normal hematopoiesis. Both are highly expressed in immature human and murine HSCs and their expression declines during lineage specification^{74,75}. Of all the *HOX* genes, deletion of *HOXA9* has the greatest impact on hematopoiesis⁷⁶. *HoxA9*^{-/-} mice are viable^{77,78} whereas homozygous deletion of *Meis1* is embryonic lethal⁷⁹. *HoxA9* knockout prevents transplanted HSCs from repopulating irradiated mice⁸⁰, although crossing *HoxA9*^{-/-} mice with *MLL-AF9* mice still leads to leukemogenesis⁸¹. Loss-of-function mutations affecting either *HoxA9* or *Meis1* both lead to similar

defects in mice⁷⁶. Wild-type MLL is essential for normal steady-state expression of *HOX* genes²¹. In *Mll* knockout mice, *Hox* gene expression initiates normally but is not maintained^{5,29,57,70,82–84}.

1.3.2 *HOXA9* and *MEIS1* are drivers of MLLr leukemogenesis

MLLr leukemia is a disease of HSC fate specification. In normal HSCs, wild-type MLL regulates *HOXA9* and *MEIS1*, which in turn control the ‘stemness’ of HSCs⁵⁷. It is unsurprising, therefore, that two of the key targets of MLL fusion proteins in malignant leukemia should be *HOXA9* and *MEIS1*⁵⁴.

HOXA9 and *MEIS1* expression is strongly associated with poor prognosis in MLLr leukemia. Moreover, functional studies have shown that *HOXA9* and *MEIS1* are both key members of the gene regulatory network (GRN) downstream of MLL fusion proteins⁷⁶. In a classic study, Armstrong and coworkers expressed MLL-AF9-GFP to transform common myeloid progenitor cells and showed that *HOXA9* and *MEIS1* belong to the ‘self-renewal-associated signature’, a specific transcriptional program of genes which are overactivated in leukemia stem cells (LSCs) to maintain self renewal⁴⁷.

Many lines of evidence have since supported the hypothesis that upregulation of *HOXA* genes is the main mechanism by which MLL fusion proteins initiate or maintain malignant hematopoiesis. Ectopic expression of *HOXA9* and *MEIS1* in human and mouse models is sufficient to transform HSCs^{71,85–88}. Conversely, inhibition or knockout of *HOXA9* or *MEIS1* prevents leukemogenesis in MLL-AF9 mouse models and reverses MLL fusion protein-induced HSC immortalisation⁷¹. In patient samples, coactivation of *HOXA9* and *MEIS1* is a hallmark of human MLLr leukemia⁸⁷.

Homeobox transcription factors are highly promiscuous and their target networks are complex. Early studies showed that HOXA9-MEIS1 complexes indirectly increase expression of the transcription factor c-MYB^{89,90}. In recent years, a model has emerged where HOXA9-MEIS1 dimers promote hematopoiesis by recruiting

the transcription factor CREB1 and its cofactors CBP and PBX3 to cis-regulatory enhancer elements, where they induce cell-type specific transcriptional repression or activation⁷¹. Candidate HSC-specific factors include PU.1 and C/EBP α ⁷¹. ChIP-seq experiments have identified thousands of promoter-distal, intergenic HOXA9-binding sites. Validated targets include well-known AML/ALL-associated genes such as *MYB*, *RUNX1*, *LMO2* and *BCL-2*⁹¹.

The wild-type function of HOXA9 depends on both the core homeodomain and a divergent transcriptional activation domain at the *N*-terminus, whereas in leukemia, most HOX-specific phenotypes depend on the homeodomain alone⁹². It remains to be fully determined whether HOXA9 and MEIS1 have the same activity at the same target genes before and after transformation and how they interact⁹¹. Interestingly, recent experiments in myeloid and lymphoblastic cells transformed with MLL-AF9 or directly with HOXA9/MEIS1 revealed two distinct classes of HOXA9-binding enhancers⁹³. The ‘physiological enhancers’ had similar density of the activating epigenetic mark H3K4me1 in transformed and untransformed cells, whereas ‘*de novo* enhancers’ gained H3K4me1 in transformed cells. Correspondingly, HOXA9, CEBP and RUNX1 were more abundant at *de novo* enhancers. HOXA9 binding recruited the methyltransferases KMT2C (MLL3) and KMT2D (MLL4) to *de novo* but not physiological enhancers, which proved necessary for HOXA9/MEIS1-mediated leukemogenesis. CEBP appeared to bind DNA before HOXA9 at physiological enhancers but after HOXA9 at *de novo* enhancers. The set of genes under the control of *de novo* enhancers included *BCL2* (see subsection 1.4.2) and *IKZF2* (IKAROS); the set was enriched for genes involved in embryonic development. Taken together, these results suggest that after leukemic transformation, HOXA9 targets an additional set of genes and promotes their transcription by a mechanism distinct from its normal activity⁹³. It is unclear why transformation directs HOXA9 towards *de novo* enhancers. Perhaps additional epigenetic marks guide HOXA9 in malignant HSCs to reactivate genes inactivated during early development.

A recent study by Slany and coworkers revealed the genome-wide binding pattern of HOXA9⁹⁴. HOXA9 ChIP experiments had previously been unsuccessful

because HOXA9 quickly degrades in cell extracts before immunoprecipitation can be performed. Using a combination of chemical protease inhibition and CRISPR/Cas9 knockout, they determined that three lysosomal granule proteases were responsible for HOXA9 degradation. To overcome this longstanding technical barrier, they performed HoxA9 ChIP-seq experiments in HSCs taken from mice harbouring homozygous deletions of all three granule proteases. This yielded a ten-fold increase in the number of detectable HoxA9 targets compared to wild-type cells. A similar improvement was seen using MV4-11 cells (MLL-AF4), which have naturally low granule protease activity. HoxA9 binding was detected at known HoxA9 targets such as CEBP, RUNX1 and PU.1. Next they used nascent RNAseq with a tamoxifen-inducible HoxA9 construct to focus on genes whose transcription depends on HoxA9. Among these genes, they found a core set of 65 genes previously identified as targets of MLL-ENL⁹⁵. ChIP-seq showed that transcription of HoxA9 targets correlated with H3K27ac abundance at corresponding enhancers, although the majority of HoxA9-binding enhancers were not modified in response to HoxA9 expression. Interestingly, HoxA9 drove a positive feedback loop by binding enhancers for *HoxA1* and *HoxA6* and binding the *HoxA9* promoter, promoting transcription of all three genes. The authors also reported that Meis1 transcription was under control of HoxA9, which is surprising in light of the fact that both HoxA9 and Meis1 are necessary for leukemogenesis. They show that expression of Pbx3, a critical cofactor of Meis1^{72,73}, independently controlled Meis1 expression. Overall, they proposed that the cell-cycle genes *CDK6*, *CCND1* and *TERC*, all species-conserved targets of HOXA9, are key to the self-renewing proliferative signalling induced by HOXA9 activity.

1.3.3 MLL fusion proteins specifically upregulate expression of *HOXA9* and *MEIS1*

Specific activation of *HOX* genes by MLL fusion proteins requires the presence of wild-type MLL⁹⁶. Although the C-terminal SET domain in wild-type MLL promotes *HOX* gene activation by catalysing H3K4 methylation²⁹, loss of the SET domain in MLL fusion proteins does not prevent activation of *HOX* genes⁹⁷.

Milne et al. proposed that wild-type MLL is specifically recruited to *HOXA9* via interactions between the CxxC domain and PAF elongation complex and between the PHD fingers and H3K4me2/3²⁷. Only after wild-type MLL has activated the *HOXA9* locus can MLL-AF9 bind via interactions between CxxC, PAF and CpG (but not via PHD, which MLL-AF9 lacks).

Once MLL fusion proteins have been recruited to *HOX* loci, the DotCom complex assembles. DOT1L-mediated H3K79 methylation appears to inhibit heterochromatin formation and lead to transcriptional activation^{21,58,59,70}. *DOT1L* knockout reduces *HOXA9* and *MLL* expression and is essential for self-renewal in MLL-AF9 leukemia stem cells but not normal HSCs^{98,99}. Finally, pharmacological inhibition of DOT1L decreases *HOXA9* and *MEIS1* expression in an MLL-AF4 rat xenograft model¹⁰⁰. Work is ongoing to elucidate the downstream effects of DOT1L and other cofactors and to explain the specific upregulation of *HOXA9* and *MEIS1*⁵⁹. Menin is also essential for MLL fusion-mediated leukemogenesis^{26,101} and *HOXA9* expression¹⁰².

Homeobox genes are also upregulated in some non-MLLr leukemia subtypes. *HOXA5*, *HOXA9*, *HOXA10* and *MEIS1* are overexpressed in T-cell ALL (T-ALL) and in AML associated with the CALM (Clathrin-Assembly protein-like Lymphoid-Myeloid leukaemia)-AF10 fusion^{103–105}.

In summary, wild-type MLL is involved in upregulating *HOXA9* and *MEIS1*. These transcription factors dimerise to direct the assembly of a *cis*-binding transcriptional regulatory complex which has cell-type specific effects, including the maintenance of self-renewal capacity in HSCs during the early stages of normal hematopoiesis. In MLLr leukemia, MLL fusion proteins cause loss of regulation of *HOXA9* and *MEIS1*, leading to malignant hematopoiesis. The target network of *HOXA9* and *MEIS1* in normal hematopoiesis is poorly understood, as is the way this network changes during malignant hematopoiesis.

However, there is more to the regulation of MLLr leukemia than coding genes. In chapter 3, I focus on an intriguing piece of the MLLr puzzle: miR-196b.

1.3.4 Oncogenic and tumour-suppressor miRNAs are associated with progression of MLL-rearranged leukemia

microRNAs (miRNAs), a conserved family of 18-25nt non-coding RNAs found in all metazoan cells, play a significant role in the complex regulatory network that drives normal hematopoiesis¹⁰⁶.

Human cells express over 2500 mature miRNAs which regulate the expression of most proteins^{107,108}. The biogenesis of miRNAs begins with primary miRNAs (pri-miRNAs), which are hairpin-containing RNA polymerase II transcripts¹⁰⁹. These transcripts arise either from intergenic miRNA loci, where they are independently transcribed, or more frequently from intronic loci where they may be independently transcribed or released by intron splicing from mRNAs^{109,110}. The Microprocessor complex recognises and cleaves a pri-miRNA to yield a precursor miRNA (pre-miRNA) which is exported from the nucleus¹⁰⁹. In the cytosol, Dicer nuclease cleaves the pre-miRNA to produce a short miRNA duplex which is loaded, in humans, into an Argonaute (AGO) protein before one of the two strands is discarded¹⁰⁹. The AGO-miRNA ribonucleoprotein complex recruits other factors to form the RNA-induced silencing complex (RISC)¹⁰⁹. miRNA-RISC binds mRNA transcripts whose sequence is complementary to the seed sequence of the miRNA (nucleotides 2-8)¹⁰⁹. The primary mechanisms for miRNA-mediated gene repression in mammals appear to be transcript destabilisation and translational inhibition^{109,111,112}. Transcript destabilisation may occur co-translationally¹¹³. There is also evidence that some AGO proteins can directly cleave transcripts¹⁰⁹.

The development and progression of acute leukemia, characterised by a loss of control of hematopoiesis, is closely linked to misregulated miRNA activity¹¹⁴. Dicer is essential for HSC maintenance and lineage specification¹¹⁵. AML and ALL both display distinct miRNA expression patterns which can be used for prognosis, even though the target networks of these miRNAs are largely unknown¹¹⁶⁻¹¹⁸.

1.3.5 miR-196b is an MLLr leukemia-associated oncomir which may interact with *HOXA* genes and *MEIS1*

miR-196a1, miR-196a2 and miR-196b are three conserved miRNAs that share the same seed sequence and are all overexpressed in AML despite a general background of miRNA downregulation in cancer^{118–120}. miR-196b is specifically overexpressed in MLLr leukemia and miR-196b expression is negatively correlated with survival in AML patients¹²¹.

Experimental evidence, mostly from mouse models, points to a functional link between MLL fusion proteins and miR-196b. *Mll*-null mouse embryonic fibroblasts exhibit reduced miR-196b expression, which can be rescued by ectopic expression of *Mll*^{84,122}. MLL-AF9-transformed mouse primary bone marrow cells overexpress miR-196b; addition of antagomir-196b (which represses miR-196b) reduces proliferation^{84,123}. Expression of miR-196b in hematopoietic progenitor cells isolated from mouse bone marrow appears to block differentiation and increase cell ‘stemness’⁸⁴. Recent work showed that therapeutic inhibition of miR-196b inhibits survival in a murine AML model¹²⁴.

Human models agree with results from mice. THP-1 and MONOMAC-6 cells (both MLL-AF9) have greater expression of miR-196b, *HOXA9* and *MEIS1* than non-MLLr KASUMI-1 cells¹²¹. The same is observed in human CD34⁺ umbilical cord blood cells after transformation with MLL-AF9¹²¹.

Predicting and validating miRNA targets is a complex task¹²⁵. miR-196b is no exception. Nonetheless, several lines of evidence indicate that the miR-196 family interacts with *HOXA* cluster genes. Firstly, the *MIR196B* locus lies on chromosome 7 in an intergenic region between *HOXA9* and *HOXA10* (Figure 1.2). This linkage group is highly conserved and is thought to contribute to a remarkable hierarchical pattern of *HOX* gene expression observed throughout bilaterians called posterior prevalence¹²⁶. Secondly, *HOXC8* and *HOXA7* each contain four highly conserved near-consensus miR-196b target sites and *HOXA9* contains two. Thirdly, a classic study concluded that miR-196b directs cleavage of *HOXB8*, *HOXC8*, *HOXD8* and *HOXA7*, providing some of the earliest evidence of miRNA-targeted mRNA

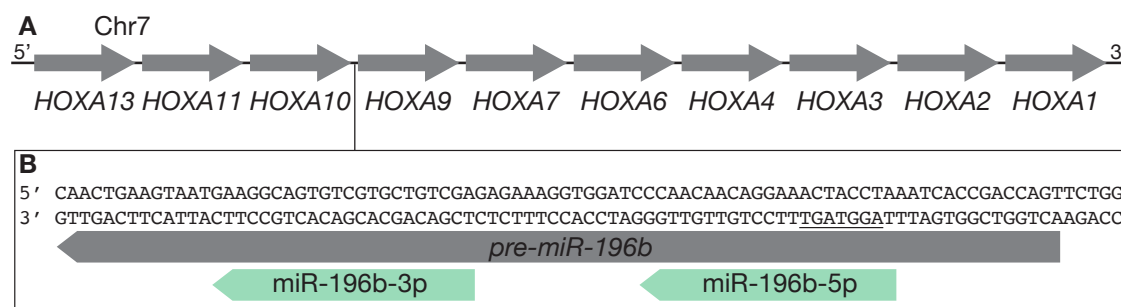


Figure 1.2: The *MIR196B* locus lies within the *HOXA* cluster. (A) *MIR196B* is expressed from a distinct transcript with an autonomous promoter located in the intergenic region between *HOXA10* and *HOXA9*. The order of the *HOXA* genes on chromosome 7 corresponds to their spatiotemporal expression patterns. The 5' *HOX* genes (starting with *HOXA13*) are responsible for posterior patterning; the 3' *HOX* genes, for anterior patterning. **(B)** Pri-miR-196b, which is transcribed from the opposite strand to the *HOXA* genes, is processed to yield pre-miR-196b (grey). In turn, this is processed to produce mature miR-196b-5p (referred to in this thesis as miR-196b) and the antisense miRNA miR-196b-3p (miR-196b*). The seed sequence of miR-196b-5p is underlined.

cleavage¹²⁷. The authors of this study used 5'-RACE (rapid amplification of cDNA ends) to detect the products of miR-196-directed cleavage of *HOXB8* mRNA in mouse embryos. Additionally, using firefly luciferase reporter constructs in HeLa cells, they demonstrated that miR-196b targets the 3' UTR of *HOXA-D* genes. Nonetheless, they neither tested whether miR-196b targets *HOX* genes during physiological human development, nor proposed a mechanism for repression.

Determining whether MLL fusion proteins directly promote transcription of miR-196 family loci has proven difficult because the *MIR196B*, *HOXA9* and *HOXA10* loci are in such close proximity. Fantini et al. used 5' RACE to conclude that miR-196b derives from a 2671bp primary transcript which maps to the intergenic region between *HoxA10* and *HoxA9* in mice¹²⁸. They found a putative TATA box and Initiator sequence at the 5' end of this transcript. Thus, transcription of pri-miR-196b is likely to be driven by an autonomous promoter. According to the posterior prevalence model, only 'posterior' HOX proteins should activate miR-196b transcription – that is, HOX proteins responsible for posterior patterning such as HOXD10-13 and HOXA9. However, another interpretation of this model suggests that HOX proteins whose loci are 5' of the *MIR196B* locus (with respect to the direction of the *HOX* genes) should promote miR-196b, thus excluding HOXA9¹²⁶.

Based on electrophoretic mobility shift assays using HEK293T nuclear extracts, conserved promoter-proximal homeodomain binding sequences in the pri-miR-196b transcript appeared to bind HOXD proteins and the related transcription factor CDX2. Deletion of some of these sites prevented binding. ChIP experiments suggested that the putative *MIR196B* promoter is active during embryogenesis (based on H3K4me3, H3K27ac and H3K27me2 profiles) and binds CDX2 and HOXD13 but not HOXA9. The sites binding CDX2 and HOXD13 were necessary for miR-196b transcription in the neural tube during chick embryogenesis. This lends support to the notion that regulation of *HOX* genes is closely linked to regulation of miR-196b in normal development.

Evidence for a physical interaction between MLL fusion proteins and miR-196b-RISC remains elusive. Mueller et al. used ChIP-qPCR to track the *MIR196B* locus and the 5'- and 3'-terminal exons of the *HOXA9* locus after switching off MLL-ENL expression in an inducible cell line⁵⁴. They found that *HOXA9* transcription and density of the activating H3K79me2 mark at the *HOXA9* and *MIR196B* loci fell by ~80% within 3 days. For 10 days, RNA pol II remained bound but inactive at the *HOXA9* locus, corroborating previous evidence that MLL-ENL is necessary for transcriptional elongation. Li et al. also used ChIP-qPCR in THP-1 and MONOMAC-6 cells to show that MLL fusion binding in a CpG island upstream of *MIR196B* recruits RNA polymerase II, leading to elevated expression of miR-196b¹²¹. An alternative possibility is that miR-196 production increases indirectly due to MLL fusion protein-mediated transcription of neighbouring genes^{97,129}. Finally, the *HOXA9* positive feedback loop described by Slany and coworkers shows that *HOXA9* does activate several regions close to the *MIR196B* locus, although these regions do not appear to overlap with the miR-196b promoter described by Fantini et al.^{94,128}

Many other miRNAs have also been functionally linked to MLL fusion proteins, including miR-10a, miR-9 and miR-17~92¹³⁰. Similar to miR-196b, miR-9 inhibition inhibits Mll-activated immortalisation of mouse bone-marrow progenitor cells¹³⁰.

1.3.6 There is no accepted model for the interaction between homeobox genes and miR-196b in MLL-rearranged leukemia

The evidence above suggests that miR-196b is responsible for tuning *HOX* gene expression during normal HSC development and that miR-196b is associated with proliferation in MLLr leukemia. But the relationship between miR-196b and its target genes is poorly understood in both normal and malignant hematopoiesis. Is miR-196b overexpression a cause or effect of transformation and self-renewal in MLLr cells? To what extent do results in murine and human models reflect real disease states? If miR-196b has a functional role in MLLr leukemia, does this change over time and is it distinct from or entwined with the misregulation of *HOXA9* and *MEIS1*?

Chen and coworkers proposed the first hypothesis to explain the interplay between *HOXA9* and miR-196b in MLLr leukemia (referred to here as the ‘Chen model’, Figure 1.3)¹²¹. Interpreting evidence from mouse models, they focused on an apparent contradiction: *HOXA9* and *MEIS1* are critical targets for activation by MLL fusion proteins and they are overexpressed after transformation, but miR-196b, which is also upregulated after transformation, is predicted to repress *HOXA9* and *MEIS1*¹²¹.

In their model, Chen and coworkers included a third predicted target gene of miR-196b: the tumour-suppressor protein FAS. FAS is a cell-surface death receptor which contains an intracellular death domain to trigger extrinsic apoptosis (see subsection 1.4.1). *FAS* downregulation in AML patients is a predictor of poor survival. *Fas* expression is downregulated in the peripheral blood of MLL-AF9 mice, but not in the absence of miR-196b. Conversely, *Fas* overexpression in mouse primary MLL-AF9 bone marrow cells reduces proliferation.

According to the Chen model, miR-196b normally prevents overexpression of *HOXA9* and *MEIS1* during the early stages of HSC lineage specification, while also weakly repressing factors such as *FAS*. In MLL-rearranged cells, *HOXA9* and *MEIS1* are so abundantly overexpressed that they escape miR-196b-mediated

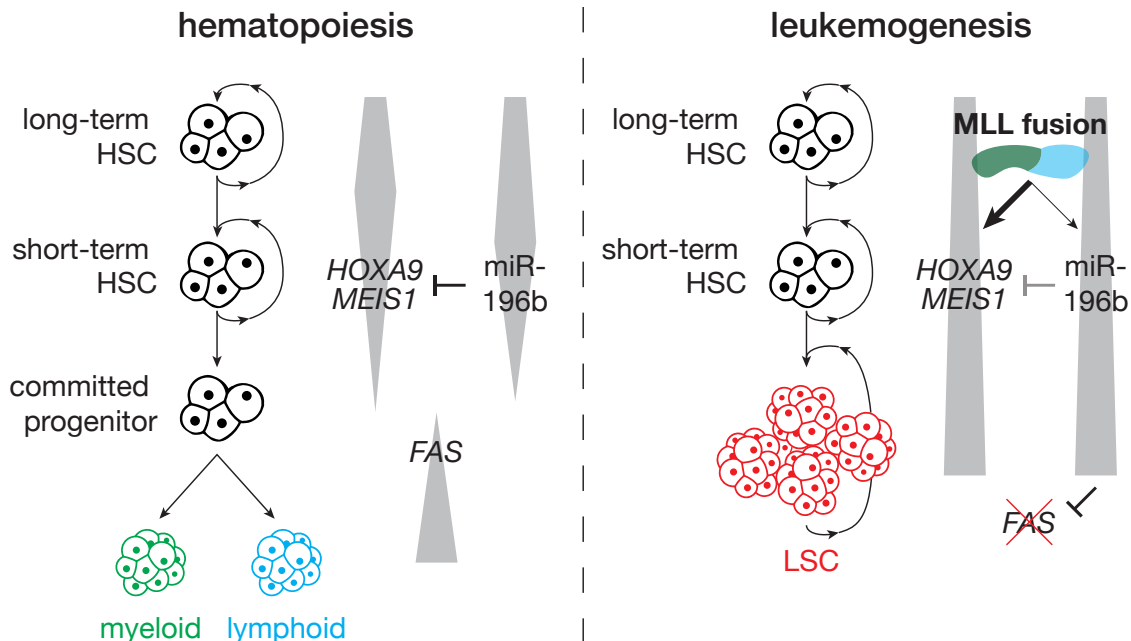


Figure 1.3: A model for miR-196b and HOXA9/MEIS1/FAS in MLL-AF9 leukemia. Chen and coworkers proposed a model to explain the role of the oncomir miR-196b, the homeobox genes *HOXA9* and *MEIS1* and the death receptor *FAS* in MLLr leukemia. During normal hematopoiesis, wild type MLL and other transcription factors drive expression of *HOXA9* and *MEIS1*, which allows long-term hematopoietic stem cells (HSCs) to maintain their self-renewal capacity. miR-196b represses *HOXA9* and *MEIS1* to allow short-term HSCs to mature into committed myeloid/lymphoid progenitors. Expression of miR-196b and *HOXA9/MEIS1* decreases in committed progenitors, allowing *FAS* expression to increase. *FAS* contributes to differentiation along the myeloid or lymphoid lineage. In malignant hematopoiesis, MLL fusion proteins upregulate *HOXA9* and *MEIS1* more than they upregulate miR-196b. *HOXA9* and *MEIS1* overexpression leads to self-renewal. Increased miR-196b activity prevents expression of *FAS*, also contributing to self-renewal. Instead of committing to the myeloid or lymphoid lineage, progenitors develop into leukemia stem cells (LSCs).

repression. Increased levels of miR-196b do, however, lead to repression of *FAS*, thereby preventing differentiation. This model leaves many questions unanswered, such as whether *FAS* is the only miR-196b target gene involved in fate specification. My efforts to validate the Chen model are described in chapter 3.

1.3.7 THP-1 cells are a model system for MLL-AF9 AML

Stable cell lines established for each of the most common 11q23 translocations have become workhorses of *in vitro* studies of MLLr AML/ALL¹³¹. I use one of these to test the Chen model. Most of these cell lines are monoclonal; their

characteristic chromosomal abnormalities are stable and they are differentiation-arrested along the HSC lineage¹³².

The THP-1 cell line was established from peripheral blood monocytes isolated from a one-year-old boy with MLL-AF9 AML¹³³. Validation using cytogenetic techniques such as spectral karyotyping, fluorescence in situ hybridisation (FISH) and RT-qPCR confirmed the expression of MLL-AF9, as well as other secondary rearrangements which remain poorly understood¹³⁴. THP-1 cells have Fc and C3B receptors but they lack surface and cytoplasmic immunoglobulins¹³³. A *CSNK2A1-DDX39B* fusion may affect cell cycle control and apoptosis¹³⁵. THP-1 cells are also heterozygous for a frameshift mutation in the tumour suppressor gene *TP53* (p.Arg174fs3/c.520_545del26)¹³⁶ and an activating point mutation in the oncogene *NRAS* (p.Gly12Asp/c.35G>A)¹³⁷.

MLL-AF9 is the most common fusion protein in adult and pediatric AML; in infant AML, MLL-AF9 and MLL-AF10 have similar frequency³². The Chen model was developed in THP-1 cells¹²¹. The work described in this thesis uses THP-1 cells almost exclusively. Whilst recent work in the Milne lab has established that the transcriptional programs in MLL-AF9 and MLL-AF4 cell lines have much in common, it is important when attempting to generalise results to remember that MLL-AF9 and MLL-AF4 are each associated with several clinically distinct subtypes of MLLr leukemia. The transcriptional networks downstream of MLL-AF4 and MLL-AF9, although similar, are not identical¹³⁸. This corresponds to the differences in leukemogenesis, self-renewal capacity and lineage plasticity¹³⁹ between MLL-AF9 and MLL-AF4 cells. THP-1 cells only reflect MLL-AF9 AML; in any case, they are a secondary cell line and may differ significantly from primary monocytes¹⁴⁰.

1.3.8 MLLr leukemia treatments face many challenges

Treatment of MLLr leukemia leads to very poor overall survival rates and frequent relapse. Our limited understanding of the downstream targets of MLL fusion proteins makes them difficult to drug. A better understanding of the *HOXA9/MEIS1/FAS*/miR-196b axis could lead to more effective targeted therapies

for MLLr leukemia. The role of *FAS* is especially intriguing because it indicates that MLLr leukemia is resistant to apoptosis. In section 1.4, I explore this theme in more detail to ask whether apoptosis could offer a promising therapeutic avenue to overcome the treatment resistance of MLLr leukemia.

1.4 Programmed cell death in MLLr leukemia

1.4.1 Evading apoptosis is a hallmark of many cancer types

Apoptosis is a highly regulated form of programmed cell death which plays an important role in healthy embryonic development, tissue homeostasis and the response to cell damage¹⁴¹. Unlike necroptosis and pyroptosis, in which plasma membranes rupture and the cytosol spills out, apoptosis is a carefully controlled process. When a cell commits to apoptosis, the nucleus shrinks and chromatin becomes fragmented, the cell shrinks as the cytoskeleton is dismantled, and cellular components are compartmentalised in vesicles which are removed by macrophages in a process called efferocytosis¹⁴².

There are two distinct apoptotic signalling pathways: extrinsic apoptosis (also known as the death factor pathway) and intrinsic apoptosis (also known as mitochondrial apoptosis) (Figure 1.4). Both pathways converge in the cytosol by activating cysteine aspartyl proteases (caspases), the final effectors of apoptosis. Caspase-3 cleaves over 1300 cellular substrates to initiate apoptosis¹⁴¹. Here I focus on intrinsic apoptosis.

Since cancer cells accumulate DNA and protein damage as a result of rapid proliferation, they must activate prosurvival pathways and inactivate apoptotic pathways in order to avoid cell death^{143,144}. Key signalling pathways inactivated in cancer include the pro-apoptotic TP53 pathway and the pro-survival IL-3 pathway¹⁴⁴. A better understanding of how cancer cells evade apoptosis – and, especially in MLLr leukemia, how apoptosis can be reactivated – holds great promise for targeted cancer therapy.

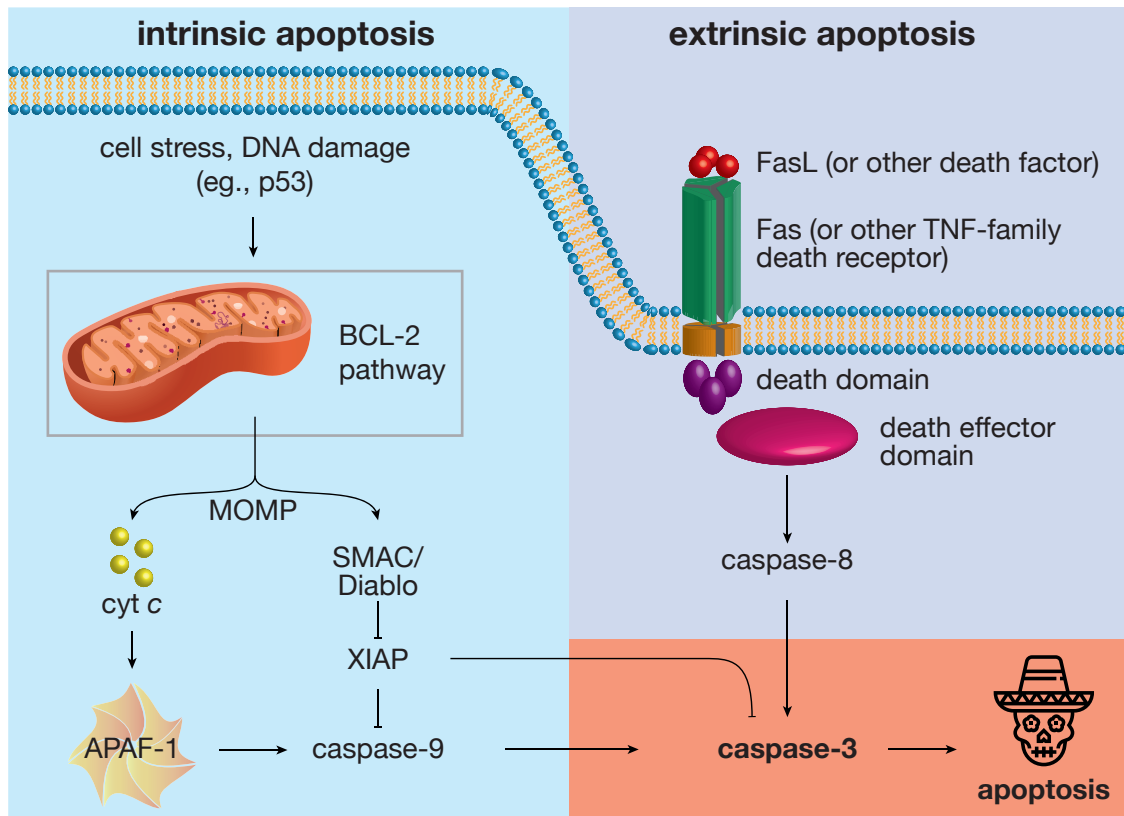


Figure 1.4: Intrinsic and extrinsic apoptosis. There are two distinct pathways for programmed cell death in eukaryotes. In intrinsic apoptosis (*left*), internal cell stress signalling activates the mitochondrial BCL-2 pathway. This leads to mitochondrial outer membrane permeabilisation (MOMP) and the release of downstream cytoplasmic effectors which activate caspases. In extrinsic apoptosis (*right*), extracellular ligands bind tumour necrosis factor (TNF)-family receptors in the plasma membrane (including FAS), activating intracellular signalling pathways which activate caspases. Caspase-3 initiates cell death. Mitochondrion cartoon reused with permission from Wikimedia Commons.

1.4.2 BCL-2 is a key regulator of programmed cell death

Mitochondrial apoptosis is regulated by the BCL-2 family, a well-studied group of more than 20 signalling proteins whose subcellular localisation, abundance, physical interactions, structure and stability change over time in response to intrinsic stimuli^{145,146}. Mitochondrial apoptosis culminates in mitochondrial outer membrane permeabilisation (MOMP) – the ‘point of no return’¹⁴⁵. Pores in the MOM allow proteins such as cytochrome *c* and SMAC/DIABLO to diffuse out of the intermembrane space and activate downstream effectors such as caspases. Live-cell imaging recently showed that mitochondrial DNA is also extruded through

pores during MOMP¹⁴⁷.

BCL-2 family proteins are characterised by a group of four different helical dimerisation domains called BH1-4 (BCL-2 homology)¹⁴⁸. The most important domain, shared by all BCL-2 family proteins, is BH3. BCL-2 family proteins can be grouped into three types according to their function and structure^{145,146}:

1. **BH3-only proteins.** These pro-apoptotic proteins (including BID, BAD, BIM, NOXA, PUMA) only contain the BH3 domain. The relative concentration of each BH3-only protein changes in response to different types of cell stress¹⁴⁶. Some BH3-only proteins have independent non-apoptotic roles^{149,150}.
2. **Anti-apoptotic proteins.** These multidomain proteins (BCL-2, MCL-1, BCL-X_L, BCL-W, BFL-1) all contain BH1, BH2, BH3 and BH4. They act to prevent apoptosis.
3. **Pore-forming proteins.** BAX and BAK (and, in the endoplasmic reticulum, BOK) contain BH1, BH2 and BH3. MOMP occurs when they form symmetric homodimers which oligomerise and create pores¹⁵¹.

The BCL-2 family proteins also have a role in the membrane of the endoplasmic reticulum, where they may regulate the unfolded protein response and autophagy. Autophagy is a distinct cell death pathway which continuously renews organelles¹⁴⁵.

There are three distinct models describing how the BCL-2 family proteins interact to cause MOMP¹⁴⁶. The earliest two models were based on biochemical characterisation of the relative binding affinities of each BCL-2 family protein in solution. According to the *direct activation* model, there are two classes of BH3-only proteins: activators, which bind all five anti-apoptotic proteins; and sensitisers, which each bind only a subset of anti-apoptotic proteins¹⁵². Sensitisers have a greater affinity than activators for the anti-apoptotic proteins. Apoptotic signalling activates BH3-only proteins. The sensitisers bind BH3-binding domains on anti-apoptotic proteins, thus displacing activators. Released activators promote BAX/BAK oligomerisation to initiate MOMP. The anti-apoptotic proteins do not interact with BAX/BAK directly.

The *displacement model* differs from the direct activation model by positing that BAX and BAK are inhibited not by the BH3-only proteins but by anti-apoptotic factors. According to the displacement model, both sensitiser and activator BH3-only proteins bind only the anti-apoptotic factors. BAX/BAK constitutively oligomerise. Once BH3-only proteins bind anti-apoptotic proteins and liberate BAX/BAK, BAX/BAK oligomerise and initiate MOMP.

The most recent model, known as the *embedded together* model, combines results from the earlier two models (Figure 1.5). It is based on empirically determined relative concentrations and binding affinities of the BCL-2 family proteins in the lipid environment where they actually act: the mitochondrial outer membrane (MOM)¹⁴⁶. In this model, the BH3-only proteins recruit anti-apoptotic proteins to the MOM, where anti-apoptotic proteins sequester activator BH3-only proteins *and* BAX/BAK. This interaction is mediated by hydrophobic BH3-binding grooves on anti-apoptotic proteins and BH3 domains on BH3-only and pore-forming proteins. Sensitiser BH3-only proteins displace activators and BAX/BAK from the BH3-binding domains of anti-apoptotic proteins, leading to MOMP¹⁵³. *In vivo* knockout studies and *in vitro* assays with purified proteins and lipid bilayers confirm the dual functions of BH3-only proteins and anti-apoptotic proteins^{145,146}.

The embedded together model presents a picture of the BCL-2 family proteins in continuous flux between the cytosol and MOM. In the cytosol, post-translational modifications regulate the BH3-only proteins; in the MOM, lipid interactions cause conformational changes which activate both BH3-only proteins and anti-apoptotic proteins^{145,146,154}. Some BH3-only proteins such as tBID adopt different conformations when binding to anti-apoptotic proteins or pore-forming proteins¹⁵⁵. The abundant nonfamilial interactions of BCL-2 family proteins with cancer signalling pathways are still being characterised¹⁵⁶. Interestingly, there is evidence that extrinsic apoptosis can activate intrinsic apoptosis. Activated death receptors in the plasma membrane oligomerise and recruit signalling molecules to form the death-induced silencing complex (DISC), which activates caspase-8¹⁵⁷. Cytosolic BID is cleaved by caspase-8 to produce active tBID¹⁵⁷.

Many questions remain unanswered about the dynamic equilibrium of BCL-2 family proteins in the MOM. Studying the key effectors of apoptosis is difficult in apoptosis-resistant immortalised cell lines. For example, it is unclear whether BAX and BAK are functionally equivalent and whether BAX/BAK homodimers form before insertion into the lipid bilayer^{145,158}. Live-cell imaging of photoactivatable BAX-GFP suggested that cytosolic BAX is constantly shuttled between the cytosol and mitochondria and that, in the MOM, BAX associates exclusively with BCL-X_L¹⁵⁹. BH3-only proteins in apoptotic cells may activate BAX to reduce its rate of dissociation from BCL-X_L. The difference in localisation between BAX and BAK coincides with the observation that BAX, but not BAK, has a role in pore formation in the membrane of the endoplasmic reticulum¹⁴⁵.

1.4.3 MLL fusion genes disrupt intrinsic apoptosis

The link between HOXA9 and BCL-2 suggests that loss of control of mitochondrial apoptosis could be necessary for MLLr leukemia. Brumatti et al. found that reducing *HOXA9* expression in a myeloid progenitor cell line led to reduced expression of BCL-2 but not MCL-2 or BCL-X_L¹⁶⁰. Conversely, *BCL-2* knockout prevented HOXA9-induced immortalisation¹⁶⁰.

Anti-apoptotic proteins have several noncanonical interactions which may contribute to their role in MLLr leukemia and drug resistance. MCL-1 has a regulatory role in mitochondrial fatty acid metabolism¹⁶¹. Extramitochondrial BCL-2, MCL-1 and MCL-X_L all act as prosurvival transcription factors in cancer and are perhaps involved in promoting their own expression¹⁶².

1.4.4 Venetoclax is a highly effective BCL-2 inhibitor drug

As the molecular mechanism for intrinsic apoptosis was pieced together, attention turned to developing therapies to reactivate apoptosis in tumour cells. The most promising strategy was to develop selective small-molecule BH3 mimetic drugs that competitively inhibit anti-apoptotic proteins.

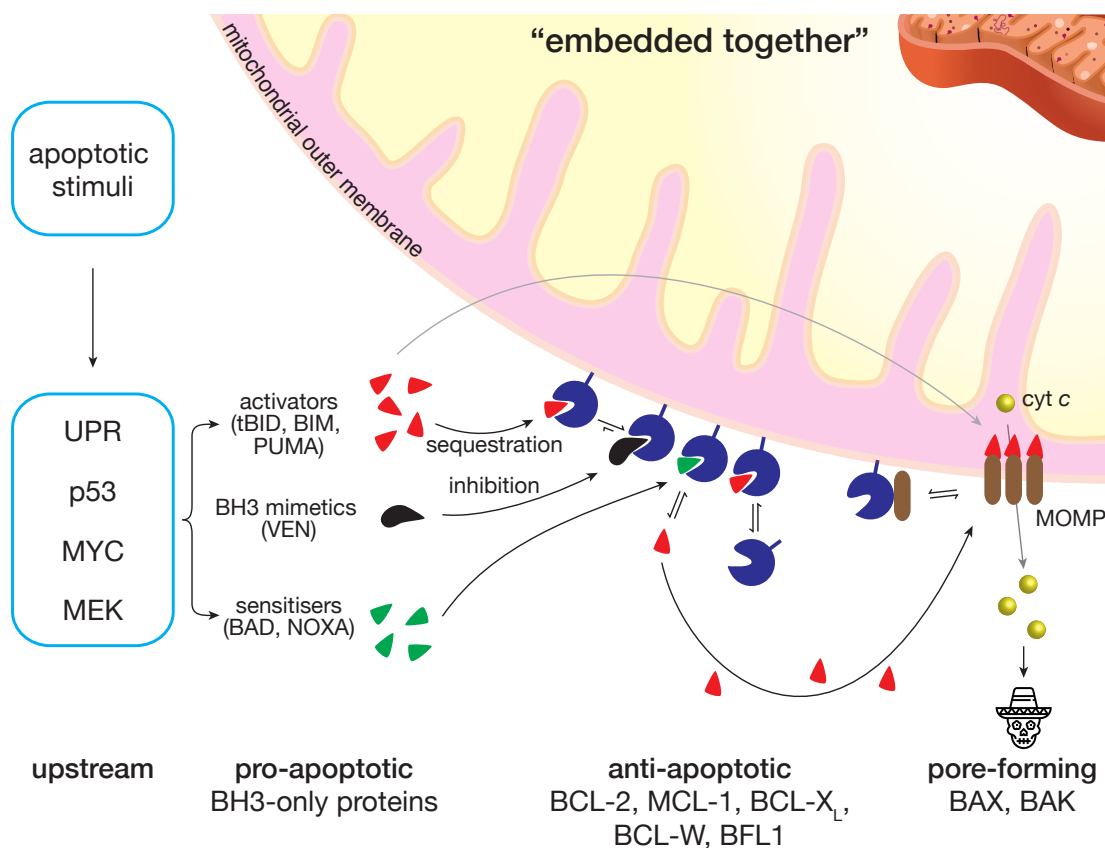


Figure 1.5: Intrinsic apoptosis and the BCL-2 family. Simplified representation of the embedded together model of intrinsic apoptosis. Apoptotic stimuli are transduced by upstream signalling pathways such as the unfolded protein response (UPR). These lead to activation of BH3-only proteins. All BH3-only proteins bind and inhibit the anti-apoptotic proteins, releasing the pore-forming proteins into the mitochondrial outer membrane. The sensitizer BH3-only proteins bind anti-apoptotic proteins more strongly than the activator BH3-only proteins, displacing the activator BH3-only proteins. Activator BH3-only proteins also directly stabilise the pore-forming proteins in the outer membrane. BAX/BAK hetero-oligomerisation creates pores which lead to mitochondrial outer membrane permabilisation (MOMP).

The first major BH3 mimetic drug was ABT-737, an inhibitor of BCL-2, BCL-X_L and BCL-W developed by Abbott Laboratories (now AbbVie). ABT-737 was designed using nuclear magnetic resonance-guided structure-based design to conform to the BH3-binding domain of anti-apoptotic proteins. It was effective at inducing apoptosis in cell lines and animal models, although oral bioavailability and solubility were poor and treatment led to decreased platelet survival because platelet lifespan depends on BCL-X_L abundance^{163–166}. Konopleva et al. showed that ABT-737 reactivated intrinsic apoptosis in AML cell lines, primary cells and a mouse AML

model¹⁶⁷. They reported that cell lines with high MCL-1 expression were resistant to ABT-737 and siRNA knockdown of *MCL-1* relieved resistance¹⁶⁷. Shortly thereafter, Letai, Armstrong and coworkers demonstrated similar efficacy and resistance in ALL cell lines and primary cells¹⁶⁸. Using BH3 profiling, a technique in which purified mitochondria are challenged with BH3-only peptides *in vitro* and their response is measured by detecting cytochrome *c* release, they argued that some cells were ‘primed’ for antagonism by different BH3 mimetics depending on which anti-apoptotic proteins were overexpressed. This raised the possibility of using BH3 profiling to predict patient responses to BH3 mimetic drugs.

Next came navitoclax (ABT-263), a BH3 mimetic with the same targets as ABT-737 but superior oral bioavailability and solubility¹⁶⁹. Combination therapy with navitoclax and rapamycin led to impressive overall response rates in mouse models of lymphoma¹⁷⁰. Unfortunately, like ABT-737, navitoclax treatment reduced platelet count. Consequently, phase I trials with navitoclax revealed severe dose-limiting toxicity due to thrombocytopenia¹⁶⁹. Although a subsequent study showed that navitoclax was effective in MLLr ALL xenograft models¹⁷¹, clinical interest in navitoclax waned because it was clear that more specific BH3 mimetics were needed in order to spare platelets and reduce toxicity. As with ABT-737, navitoclax resistance correlated with MCL-1 expression¹⁷¹. Since the clinical effects of both ABT-737 and navitoclax seemed to depend primarily on their inhibition of BCL-2, the goal soon became to develop a specific BCL-2 inhibitor drug¹⁷².

Other BH3-mimetic inhibitors of BCL-2 family anti-apoptotic proteins were equally unsuccessful: obataclax, which inhibited BCL-2, MCL-1, BCL-X_L and BCL-W, showed promise in MLLr AML but failed to achieve objective response in phase II trials^{173,174}. An antisense oligonucleotide BCL-2 inhibitor, oblimersen sodium, reached phase III trials but also had poor efficacy¹⁷⁵.

A breakthrough came thanks to ingenious structure-guided re-engineering of navitoclax to avoid inhibition of BCL-X_L and thus to spare platelets. This resulted

in the first highly potent selective inhibitor of BCL-2: venetoclax (ABT-199, GDC-0199, Venclexta; referred to here as VEN)¹⁷⁶. VEN is orally bioavailable, binds BCL-2 with high affinity and does not affect human platelets^{172,176}.

Early studies showed that VEN specifically induces intrinsic apoptosis in MLL-AF4 RS4;11 cells and other AML and ALL cell lines, but not in *Bax/Bak* double knockout mouse embryonic fibroblasts¹⁷⁶. As a single agent, VEN had high *in vivo* efficacy in an RS4;11 xenograft model¹⁷⁶. In primary cells from CLL patients, VEN sensitivity was correlated with BCL-2 expression¹⁷⁶. A first-in-human trial in three refractory CLL patients suggested that VEN had less platelet toxicity and better antileukemic effect than navitoclax¹⁷⁶.

In 2016, 32 years after the discovery of *BCL-2*, VEN became the first clinically approved anti-apoptotic cancer therapy. Based on overall response rates of ~80%, VEN received FDA approval as a breakthrough therapy for treatment of relapsed/refractory chronic lymphocytic leukemia (R/R CLL) with 17p13 deletion^{177,178}. Prior to the approval of VEN, CLL patients with chromosome 17p13 deletion were highly chemoresistant. The 17p13 band contains the *TP53* locus and p53 is necessary for chemotherapy-induced activation of pro-apoptotic proteins. VEN bypasses p53 deletion by inhibiting BCL-2.

The phase I trial which led to approval of VEN for 17p-deleted R/R CLL also demonstrated a response rate of ~80% in ordinary R/R CLL patients. The subsequent phase III MURANO trial led to overall approval of VEN for R/R CLL in combination with the anti-CD20 monoclonal antibody (mAb) rituximab¹⁷⁹. In the same indication, VEN has also been combined with the newer anti-CD20 mAb obinutuzumab, the Bruton tyrosine kinase inhibitor ibrutinib, or all three^{180–182}. Results are also promising in non-Hodgkin lymphoma and multiple myeloma¹⁸³.

VEN/rituximab combination therapy may be so effective because anti-CD20 therapy decreases MCL-1 expression¹⁸⁴. Taken together with the observation that MCL-1 expression was correlated with resistance to navitoclax and ABT-737, this suggests that successful treatment of other hematological malignancies using VEN will depend on a better understanding of intrinsic apoptosis.

1.4.5 Many AML/ALL patients are resistant to venetoclax

Despite stunning results in CLL, especially for the 17p-deleted subgroup, the efficacy of VEN in standard or MLLr AML/ALL was more modest^{185,186}. A phase II trial with VEN monotherapy in R/R AML patients achieved a response rate of only 15.5%, although the safety profile was considerably better than that of navitoclax¹⁸⁷. Primary cells from AML patients with cytogenetic abnormalities were resistant to VEN *in vitro*¹⁸⁸. Compared to its predecessors, the high selectivity of VEN improved its efficacy, but also increased the risk of resistance in patients who did not have elevated BCL-2 expression. CLL is unusual in its singular addiction to BCL-2¹⁸⁹.

The most common known mechanism for VEN resistance in AML/ALL is high expression of MCL-1 or BCL-X_L¹⁹⁰. MCL-1 sequesters BIM that is released from BCL-2 after VEN treatment¹⁹¹. This was first observed in cultured cells treated with navitoclax, then later in CLL patients treated with VEN¹⁹⁰. *TP53* and *MLL3* mutations also frequently accumulate in the same patients^{192,193}. A pathway activation screen in AML cell lines found that MCL-1 and BCL-X_L were the principal mediators of VEN resistance¹⁹⁴. A key group of patients with increased MCL-1 and BCL-X_L activity are those with *FLT3* internal tandem duplications (FLT3-ITD). Unsurprisingly, these patients are highly VEN resistant¹⁹⁵.

Rational combination therapies are a powerful strategy for overcoming VEN resistance in AML/ALL. For example, *NPM1* is a gene commonly mutated in non-MLLr AML^{196,197}. NPM1 protein has several roles, including stabilising p53 and antagonising MDM2 (which normally tags p53 for degradation) to mediate the cell stress response and promote intrinsic apoptosis. Mutations in *NPM1* and the MAPK signalling pathway both stabilise MCL-1. Reasoning that increased MCL-1 would sequester BH3-only proteins and prevent apoptosis, Konopleva and coworkers evaluated the MAPK inhibitor cobimetinib in combination with VEN in *NPM1/NRAS* mutant OCI-AML3 cells and primary cells¹⁹⁸. They found that MAPK inhibition reduced MCL-1 abundance and synergised with VEN to promote apoptosis¹⁹⁸. Along similar lines – and reflecting the success of VEN in combination

with anti-CD20 mAbs in CLL – recent work showed that combination of an MDM2 inhibitor with VEN was effective in cell lines and patient samples even with high MCL-1 expression¹⁹⁹. VEN resistance in FLT3-ITD patients can be overcome with FLT3 inhibitor combination therapy¹⁹⁵.

1.4.6 Treatment of MLLr leukemia with venetoclax needs rational combination therapies

Clinical trials have not yet decisively established whether MLL rearrangements tend to make patients more or less likely to respond to VEN treatment, although it is known that MLL-AF4 leukemias are sensitive to VEN²⁰⁰. VEN monotherapy seems unusually effective in patient xenograft models of infant MLLr ALL²⁰¹. It is possible that MLL fusion proteins directly or indirectly activate anti-apoptotic factors, leading to a specific impact on VEN sensitivity depending on the MLL fusion partner. Rational combination therapies could, therefore, be tailored to the unique molecular landscape of MLLr leukemia.

The interaction of MLLr proteins with intrinsic apoptosis signalling presents an attractive axis for selectively killing LSCs while sparing healthy HSCs. Several elements of MLLr fusion complexes are druggable, such as Menin and the SET domain^{202,203}. Moreover, VEN could be combined with drugs that inhibit the epigenetic activity of MLL fusion protein binding partners²⁰⁴. This concept has been demonstrated in non-MLLr leukemia: a small phase Ib trial of VEN with the hypomethylating agents 5-azacytidine or decitabine or the polymerase inhibitor cytarabine in elderly patients with previously untreated AML achieved a promising complete remission rate of ~67%, leading to accelerated approval for this combination^{205,206}. Similarly, promising preliminary phase I results for VEN in combination with cytarabine or the topoisomerase II inhibitor idarubicin in children with R/R AML were released at ASCO 2019 (NCT03194932). Perhaps this approach will work in MLLr leukemia as well: DNA methyltransferase inhibitors such as 5-azacytidine also reduce MCL-1 expression in MLL-AF4 primary cells, although the mechanism remains unclear²⁰⁷.

One line of evidence of a link between MLL fusion proteins and apoptosis comes from recent work by Mueschen and coworkers on the proto-oncogenic transcription factor BCL6, which is associated with poor clinical outcomes in MLLr B-ALL²⁰⁸. Transformation of murine pro-B cells with MLL-AF4 or MLL-ENL led to *BCL6* overexpression. ChIP-seq demonstrated that MLL-AF4 was bound to the *BCL6* promoter, which was enriched for H3K79me2, the activating histone mark catalysed by DOT1L. BCL6 appeared to be essential for maintenance of B-ALL cells in a murine model of MLL-ENL B-ALL. *BCL6* deletion or inactivation led to an increase in expression of BIM (*BCL2L11*) and, crucially, ChIP-seq indicated that MLL fusion proteins and BCL6 both bind to the *BCL2L11* locus. Taken together, these results indicate that MLL fusion proteins can specifically downregulate pro-survival signalling by BH3-only proteins.

Intriguingly, VEN synergises with the DOT1L inhibitor EPZ-5676 in t(4;11) SEM cells, in which BCL-2 is overexpressed²⁰⁰. DOT1L inhibition reduces expression of BCL-2 but not MCL-1, BIM, BAX, or BCL-X_L²⁰⁰. This suggests that MLL-AF4-DOT1L activity specifically increases BCL-2 expression and corroborates evidence that MLL-AF4-DOT1L promotes transcription of *HOXA9* which, in turn, is known to upregulate the *BCL2* locus^{91,160,200}. Furthermore, *HOX* gene expression predicts VEN sensitivity in MLLr AML patients^{209,210}. Finally, MLL-AF4 physically interacts with the tumour suppressor gene *FAF1* (Fas-associated factor 1) which has been implicated in intrinsic apoptosis²¹¹.

Two intriguing studies suggest that the preferential dependence of different cell lines on anti-apoptotic factors is not the only factor affecting drug resistance. HCT116 Octa-KO cells are a cell line in which all eight key BH3-only proteins (BIM, BID, PUMA, NOXA, HRK, BMF, BIK and BAD) have been deleted²¹². BH3-mimetic inhibitors of MCL-1 and BCL-X_L lead to the same degree of BAX activation and apoptosis in HCT116 Octa-KO cells as in wild-type HCT116 cells, suggesting that BH3 mimetics do not act solely by inhibiting BH3-only proteins from interacting with anti-apoptotic proteins²¹³.

In light of the Chen model (Figure 1.3), which posits a connection between HOXA9/MEIS1 and miR-196b in MLLr leukemia, it is interesting to speculate whether miR-196b may contribute to VEN resistance (or perhaps sensitivity) if it directly or indirectly affects the expression of BCL-2 family proteins. A similar phenomenon exists in CLL, where BCL-2 overexpression can arise due to loss of the BCL-2-targeting miRNAs miR-15 and miR-16²¹⁴. There is no direct evidence linking miR-196b to VEN resistance.

Tumour heterogeneity may also be a significant factor contributing to VEN resistance in MLLr leukemia. Even within a clonal population, cellular responses to environmental stimuli may vary. This fractional response depends on the density of BAX/BAK at the MOM and the overall cellular density of mitochondria^{215,216}.

Apart from combinations of VEN with traditional chemotherapeutic agents, current efforts are focused on coinhibition of BCL-2 and other anti-apoptotic proteins – that is, achieving similar activity to the earlier BH3 mimetics without the dose-limiting toxicity. Several small-molecule MCL-1 and BCL-X_L inhibitors have been recently described^{183,190,217,218}. Combined with dynamic BH3 profiling, these hold great promise but there is still a significant unmet need in MLLr leukemia.

In summary, designing venetoclax combination therapies for MLLr leukemia requires a better understanding of how MLL fusion proteins affect the signalling pathways that regulate intrinsic apoptosis. In this thesis, I explore these pathways using forward and reverse CRISPR screening, which I introduce in section 1.5.

1.5 Genome-wide CRISPR screening

1.5.1 RNAi and CRISPR are powerful technologies for forward and reverse genetics in mammalian cells

The abundance of genomic sequencing data belies the fact that functional annotation of the human genome is still incomplete²¹⁹. There are, broadly, two strategies for discovering the function of coding and non-coding genes: forward and reverse genetics.

Forward genetics is an unbiased approach in which a panel of genetic variants is screened for association with a phenotype of interest²²⁰. The variants are either

pre-existing in a population, such as in disease linkage studies or GWAS, or the variants are experimentally induced by perturbing a random or systematic set of genes. In both cases, genes associated with the target phenotype are identified by the screen and investigated further.

Reverse genetics, on the other hand, begins with a known gene of interest²²⁰. That gene is perturbed and the resulting phenotype is studied in order to learn about the function of the gene.

Classical tools for genetic perturbation, such as chemical or transposon-mediated mutagenesis, were relatively slow and experimentally challenging²²¹. Soon after its discovery two decades ago, RNA interference (RNAi) became a popular tool for forward and reverse genetics (Figure 1.6C). As an experimental technique, RNAi uses short interfering RNAs (siRNAs) for targeted gene knockdown, harnessing existing cellular RNA-induced silencing mechanisms (see subsection 1.3.4)^{222–224}. Compared to previous techniques, RNAi is highly programmable and siRNAs are simple to deliver to many model systems. However, RNAi can only be used for knockdown (not gain of function) and the extent of RNAi-induced knockdown varies between genes²²⁵.

CRISPR (clustered regularly interspaced palindromic repeats)/Cas9 has supplanted RNAi as the method of choice for targeted genetic screening. Wild-type *S. pyogenes* Cas9 (SpCas9) is a programmable nuclease which induces double strand breaks (DSBs) at genetic loci specified by the sequence of a single guide RNA (sgRNA)²²⁵. The target locus must be complementary to the sgRNA sequence and followed by a protospacer-adjacent motif (PAM). Double strand breaks are repaired by one of two mechanisms: homology-directed repair (HDR), which uses a template sequence for precise repair; or non-homologous end joining (NHEJ), which is error-prone and introduces insertion/deletion (indel) mutations at the site of repair²²⁵. Indels in an early exon of a protein-coding gene tend to cause loss of function. I refer to editing using wild-type SpCas9 as CRISPR-KO (knockout) (Figure 1.6A) in order to distinguish it from CRISPRa (activation) and CRISPRi (inhibition) (Figure 1.6B), in which fusions of catalytically inactive dead Cas9 (dCas9) with effector domains can be used for gene activation and inhibition respectively^{225,226}.

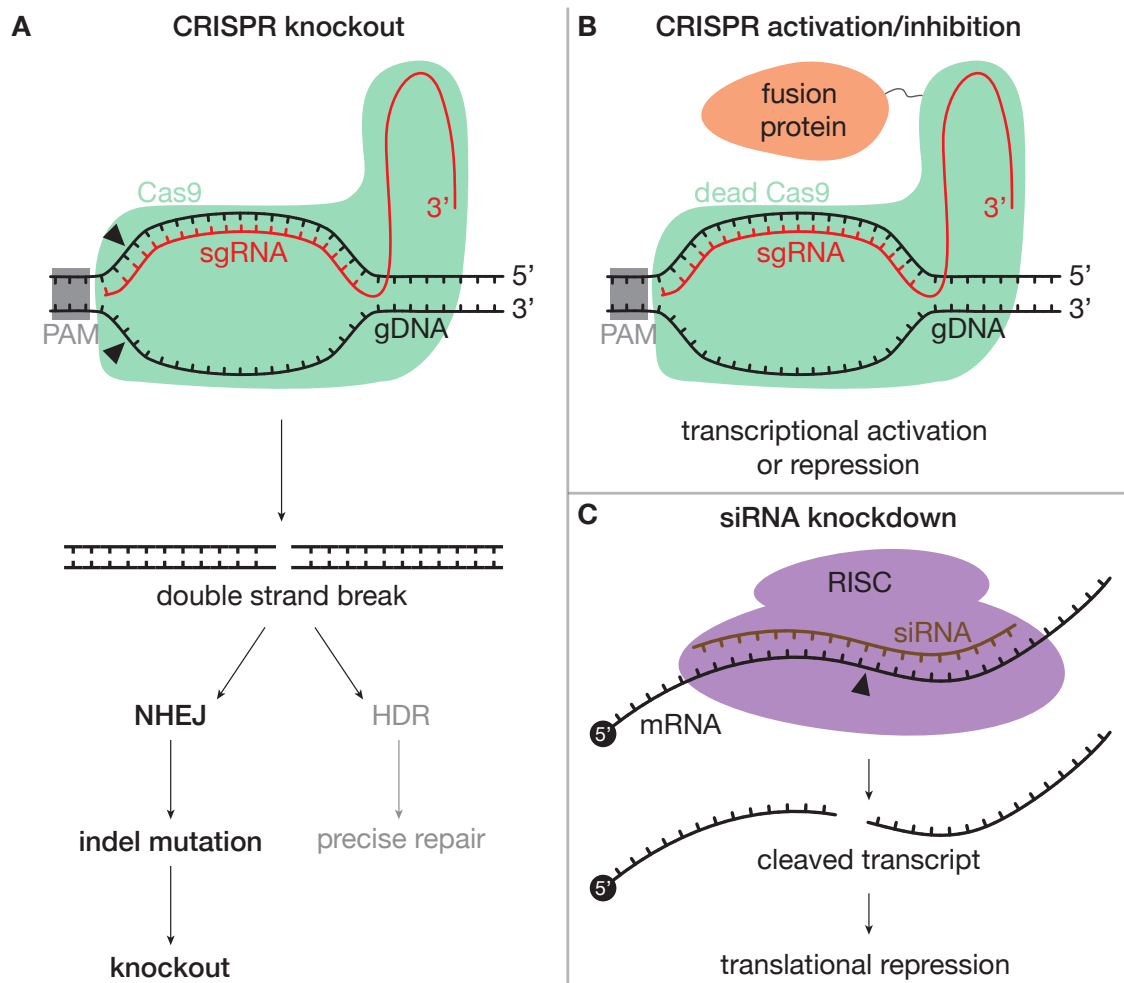


Figure 1.6: Methods for gene perturbation using RNAi or CRISPR. (A) In CRISPR knockout (CRISPR-KO), the Cas9-sgRNA complex searches the genomic DNA (gDNA) for a sequence complementary to the sgRNA upstream of a protospacer-adjacent motif (PAM). The sgRNA hybridises to the target strand, melting the gDNA duplex. The target strand and non-target strand are each cleaved by a separate nuclease domain. The cleavage sites (black triangles) are 3nt from the PAM. The resulting double-strand break is repaired by one of two mechanisms: homology-dependent repair (HDR) or non-homologous end joining (NHEJ). NHEJ frequently results in indel mutations which cause either frameshift mutations or protein disruption, knocking out the gene. Typically the sgRNA is designed to target the first exon of a coding gene. (B) In CRISPR activation/inhibition (CRISPRa/i), wild-type Cas9 is replaced with catalytically inactive dead Cas9 (dCas9). dCas9 is fused to one of many activation or repression complexes which promote or repress transcription by chromatin modification or direct interaction with RNA polymerase. (C) siRNA knockdown is primarily post-transcriptional. siRNAs in the cytoplasm are loaded into the RNA-induced silencing complex (RISC) which binds to mRNA transcripts that are complementary to the siRNA. The transcript is cleaved and destroyed, preventing translation.

In chapter 3 and chapter 4, I use CRISPR-KO for reverse genetic screens. The main advantage of using RNAi or CRISPR for reverse screening compared to other genome engineering technologies such as zinc finger nucleases (ZFNs) or TALE effector nucleases (TALENs) is that RNAi and CRISPR are readily programmable²²⁷. siRNAs and sgRNAs are straightforward to design and synthesise because each sequence corresponds (at least in theory) to a specific perturbation.

RNAi and CRISPR also have tremendous benefits for forward screening. There are two types of forward genetic screen: arrayed or pooled. In general terms, an arrayed screen involves separately inducing one mutation per cell population, whereas a pooled screen involves inducing a mixture of mutations in the same population. In an arrayed screen, the cellular phenotype of each population is measured separately and the populations are compared. In a pooled screen, the whole population is selected according to its phenotype and the distribution of perturbations is recovered by next-generation sequencing (NGS) after selection. Since each siRNA or sgRNA acts like a molecular barcode for the perturbation that it causes, multiple siRNAs or sgRNAs can be combined in a single population to enable a pooled forward screen^{225,228}. The siRNAs or sgRNAs can then be recovered after the screen by NGS.

Pooled screens are simpler and less resource-intensive than arrayed screens because shRNAs (which encode siRNAs) or sgRNAs can be synthesised, cloned and delivered as a pool, and cells can be cultured in bulk. These advantages have made it practicable to perform pooled *genome-wide* RNAi or CRISPR screens, which target all known coding or non-coding genes^{225,228–230}. Nonetheless, one disadvantage of pooled screens is that they are restricted to a smaller range of phenotypes than arrayed screens.

Genome-wide pooled CRISPR screens have become more popular than RNAi screens because CRISPR/Cas9 offers more robust on-target activity, fewer off-target effects and greater versatility^{225,228}. Empirical comparisons show that CRISPR-KO screens identify essential genes more effectively than RNAi screens^{231,232}. For the

remainder of section 1.5, I focus on CRISPR screens, although many aspects of the design and analysis of RNAi screens are very similar.

There are several non-viral strategies for delivering RNAi or CRISPR components, including microinjection, lipid transfection and electroporation. These have been extensively reviewed elsewhere^{233,234}. Pooled screens use lentiviral transduction because lentivirus can be accurately titred to ensure that each cell receives no more than one lentivirus (Figure 1.7). This precise control over copy number avoids perturbing multiple genes within the same cell and improves consistency of transgene expression. Successive generations of attenuated recombinant lentivirus based on HIV-1 have become established tools for gene delivery^{235–238}.

To generate lentivirus, the viral genome is transfected into a producer cell line such as HEK293T, in which the virus replicates. For safety (since HIV-1 is a human pathogen), recombinant lentiviral vector systems divide the viral genome into several plasmids. Most of the elements essential for viral packaging are delivered to the producer cells *in trans* relative to the transgene – that is, on separate plasmids. The separate plasmids are known as helper plasmids. The lentivirus thus produced contains the transgene but lacks many other essential genes, making it replication-incompetent. Third-generation recombinant lentiviral systems, as used in this thesis, typically comprise a transfer plasmid (containing the transgene to be integrated), envelope plasmid and one or two packaging plasmids. 48 hours after transfection, lentiviral particles are harvested and purified from the cell supernatant²³⁹. The resulting lentiviral preparation is used to transduce cells. Expression of sgRNAs and shRNAs in lentiviral constructs is typically driven by an RNA pol III promoter such as the U6 promoter²³⁹. After transcription, sgRNAs are loaded into SpCas9 (or equivalent). shRNAs are processed in the nucleus and exported into the cytosol where they are cleaved to generate an siRNA duplex. One of these strands is loaded into RISC.

A separate advantage of using lentiviral transduction for forward and reverse screening is that some cell lines, including monocytic lines such as THP-1 in which the cGAS/STING pathway is active, are highly resistant to plasmid transfection²⁴⁰.

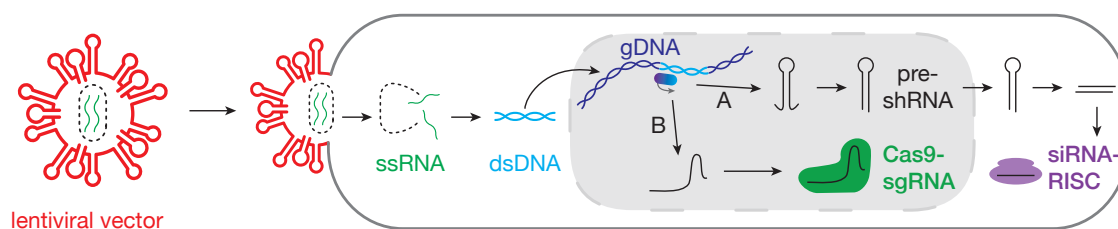


Figure 1.7: Lentiviral vector-mediated RNAi or CRISPR. Recombinant lentiviral vectors are used to deliver sgRNAs or shRNAs to mammalian cells for pooled screens. The lentiviral vector (red) contains a transgene (green) encoded in single stranded RNA (ssRNA). During transduction, the vector binds and enters the target cell. The capsid disassembles and the released ssRNA is reverse transcribed to generate proviral double stranded DNA (dsDNA). The proviral DNA is randomly integrated into the host genomic DNA (gDNA). The integrated construct is transcribed. It is common to use the RNA pol III U6 promoter to drive sgRNA or shRNA transcription. **(A)** For RNAi, the shRNA gene is transcribed to produce a primary shRNA. The primary shRNA is processed by Drosha to yield a precursor shRNA (pre-shRNA) that is exported from the nucleus. The resulting shRNA is cleaved by Dicer to generate an siRNA duplex, one strand of which is loaded into RISC. **(B)** For CRISPR/Cas9-mediated gene perturbation, the sgRNA is transcribed and processed, then loaded into Cas9 (or variants). Cas9 can be expressed from the same construct as the sgRNA (although normally driven by a separate EF1 α promoter) or from an alternative site (see subsection 1.5.2). Simplified cartoon; not to scale.

1.5.2 CRISPR/Cas9 is used for pooled genome-wide screening

The design of a pooled CRISPR screen revolves around three components: a model system, a perturbation and a selection assay²⁴¹. The choice of specific components depends on the purpose and context of the screen. In this section, I introduce how CRISPR screens work. In chapter 5, I provide a detailed description of the practical aspects of pooled CRISPR-KO screening, including delivery, sgRNA libraries, screen design and analysis.

A CRISPR screen begins with a model system: a population of primary or secondary cells which can be selected according to the phenotype of interest²⁴¹. A simple model system such as secondary cells cultured *in vitro* (as used in this thesis) allows for easier selection and readout, although it may be less representative of the real biological system. Improved practical techniques mean that screens are increasingly being conducted in primary cells or *in vivo*^{241,242}.

For a given model system, a pooled CRISPR screen involves three key steps (Figure 1.8):

1. **Transduction.** CRISPR screens use Cas endonuclease and an sgRNA library to introduce pooled genetic perturbations in the model system²²⁵. Different Cas variants can be used for CRISPR-KO/a/i screens and many different sgRNA libraries are available for different target gene sets. sgRNA libraries can also be custom designed. As discussed above, sgRNA libraries are delivered using lentivirus.

In ‘single-vector’ screens, the sgRNA and Cas9 are codelivered in the same lentiviral construct. In ‘dual-vector’ screens, a lentiviral construct containing only the sgRNA is transduced into a cell line that constitutively or inducibly expresses Cas9²⁴³. In both cases, the lentiviral construct contains an antibiotic resistance marker so that untransduced cells can be removed from the population by antibiotic selection. The result is a mixed population of cells, each containing a single gene perturbation.

2. **Selection.** Next, the pooled edited population is cultured *in vitro* or *in vivo*. The cells are divided into treatment and control groups (for example, drug/no drug) and selected based on proliferation (viability screening) or another phenotype (phenotypic screening). In a phenotypic screen, the cells must be physically separated (for example, by FACS). A special subset of viability screening is essentiality screening (discussed below), in which proliferation of untreated cells is both the phenotype of interest and the selection method^{225,242–245}.

3. **NGS.** Each gene perturbation corresponds to a specific genomically integrated sgRNA. Cells are harvested at the start and end of selection (and sometimes at intermediate timepoints). Genomic DNA is harvested from each cell sample and subjected to NGS. The distribution of sgRNAs in the population corresponds to the distribution of perturbations. By comparing the distribution of sgRNAs at each timepoint, it is possible to identify sgRNAs which were

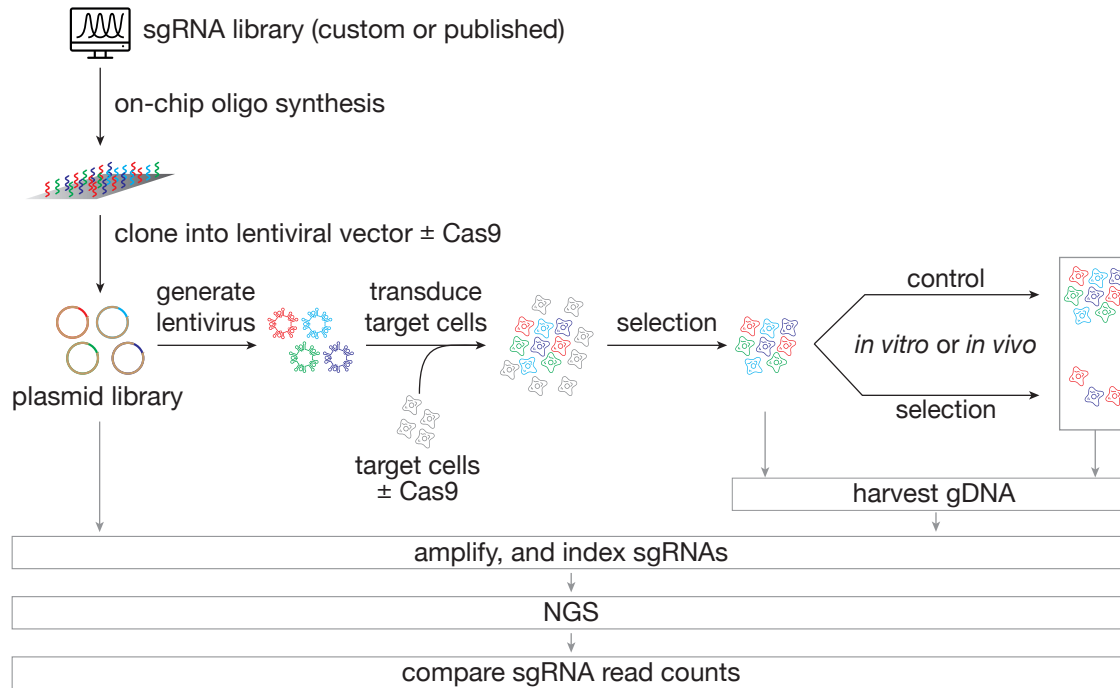


Figure 1.8: Pooled CRISPR screening. Generalised summary of pooled CRISPR screening. A custom or published sgRNA library is synthesised and cloned into a lentiviral vector to make a plasmid pool. In a single-vector screen, the lentiviral vector expresses both the sgRNA and SpCas9 (or a variant for CRISPRa/i). In a dual-vector screen, the lentiviral vector expresses only the sgRNA; a separate lentiviral vector is used to establish a constitutive or inducible Cas9-expressing target cell line. In both cases, the sgRNA-containing vector library is validated by next-generation sequencing (NGS), then transduced into helper cells (normally HEK293T or HEK293FT) to generate lentivirus. The lentivirus is titred and used to transduce target cells at low multiplicity of infection. In a single-vector screen, target cells may be primary or secondary. Untransduced cells are removed based on a selection marker expressed by the sgRNA-containing vector. Some transduced cells are harvested. The remainder are cultured *in vitro* or *in vivo*. The cells are divided into treatment and control groups (for example, drug/no drug) and subjected to viability screening or phenotypic screening. Cells are harvested again at the end of the screen. Genomic DNA (gDNA) is extracted from all harvested cells and the sgRNAs are amplified and sequenced. Comparing sgRNA read counts between the start and end of the screen indicates the effect of specific perturbations on the phenotype of interest.

enriched or depleted under selection. These correspond to perturbations which affect the phenotype of interest.

1.5.3 Pooled CRISPR/Cas9 screens enable target discovery in leukemia

CRISPR screens hold great potential for discovering novel therapies for MLLr leukemia²⁴⁶. Early genome-wide CRISPR screens were essentiality screens, which

sought to identify cell type-specific sets of genes essential for cell survival²⁴⁷. Genes essential in cancer cells but not in healthy cells could be valuable therapeutic targets for selective killing of cancer cells.

Some essentiality screens focused on leukemia cell lines, demonstrating the power of CRISPR screening for investigating hematologic malignancy. For example, Sabatini and colleagues performed a seminal screen in two CML lines and two Burkitt's lymphoma lines²⁴⁸. Of 33 essential genes specific to the CML lines, the top two were *BCR* and *ABL*, the two translocation partners in the common BCR-ABL fusion protein. Similarly, essential genes specific to the Burkitt's lymphoma lines included commonly translocated genes and *MEF2B*, an activator of *BCL6* (see subsection 1.4.6).

Soon after, Yusa and colleagues published a genome-wide essentiality screen in five AML cell lines, including MOLM-13 (MLL-AF9) and MV4-11 (MLL-AF4)²⁴⁹. Two non-AML lines served as a control. *DOT1L* and *BCL-2* were depleted only in the AML lines, consistent with the model that these genes are specific to the oncogenic transcriptional program of acute leukemia. Some genes were depleted only in the MLLr AML lines, such as *HDAC3*, a histone deacetylase which is a known binding partner of MLL fusion genes. Interestingly, the histone acetyltransferase gene *KAT2A* was depleted in MOLM-13, OCI-AML2 and OCI-AML3 but not MV4-11, suggesting that *KAT2A* has a role in MLL-AF9 AML but not MLL-AF4 AML. This was confirmed by lentiviral CRISPR-KO of *KAT2A* in MLL-AF4- and MLL-AF9-transformed FLT3-ITD mice. MOLM-13 cells treated with a *KAT2A* inhibitor drug exhibited reduced H3K9 and H3K27 acetylation at the *KAT2A* locus, reduced *HOXA9* and *MEIS1* expression and increased apoptosis. Although this study did not confirm a mechanistic link between *KAT2A*, *HOXA9* and intrinsic apoptosis, the results, taken together, strongly suggest that CRISPR-KO screening could yield significant insight into the regulation of MLLr leukemia.

Erb et al. performed a related genome-wide CRISPR screen in MV4-11 cells²⁵⁰. Top dropout hits included *HOXA9*, *MEIS1* and *ENL*. Although the transcriptional activator *ENL* is a common MLL fusion partner (Table 1.1), subsequent experiments

in this study showed that wild-type ENL is a specific driver of proliferation in MLL-AF4 leukemia. They found that the YEATS chromatin binding domain in ENL (also found in AF9) bound preferentially to leukemogenic genes such as *MEIS1* and was critical for recruitment of the EAP and DOT1L.

Finally, Wallace et al. set out to search for MLLr-associated oncomirs by performing an essentiality screen in MV4-11 cells using a genome-wide library that targets both protein-coding genes and miRNAs²⁵¹. They restricted their analysis to 197 conserved miRNAs whose expression was confirmed by RNAseq, demonstrating that miR-150-mediated repression of p53 contributes to proliferation in this cell line. It is unclear whether this regulatory interaction exists in other MLL-AF4 cells. Protein-coding dropouts included BCL-2 and HOXA9.

A logical step beyond an essentiality screen is a drug screen (or chemogenetic interaction screen), in which cells are selected with a drug so that resistance-causing genes are depleted and sensitising genes are enriched²⁵². Like essentiality screens, drug screens offer insight into novel biological pathways, whilst focusing on the specific issue of drug resistance. Drug screens using RNAi^{228,253,254}, CRISPR-KO²⁵⁵ and CRISPRa²⁵⁶ have been used to find genes which contribute to clinical drug resistance. I performed a CRISPR-KO screen with VEN selection, described in chapter 5.

1.6 Summary

1.6.1 Thesis overview

1. In chapter 3 of this thesis, I describe a lentiviral CRISPR-KO platform for miRNA deletion. I establish and characterise miR-196b^{-/-} THP-1 cell lines. My results contradict the Chen model regarding the role of miR-196b and HOXA9/MEIS1/FAS in MLLr leukemia.
2. In chapter 4, I describe experiments using VEN to investigate treatment resistance in MLL-AF9 leukemia. I perform a reverse genetic screen with lentiviral CRISPR-KO to validate the role of BCL-2 family genes in VEN

resistance. Meanwhile, I optimise methods for genome-wide CRISPR screening in THP-1 cells.

3. In chapter 5, I describe a TKOv3 genome-wide CRISPR screen with VEN selection in THP-1 cells. This sheds light on the pathways underlying VEN resistance in MLL-AF9 leukemia and indicates promising VEN combination therapies.

1.6.2 Non-technical summary

Mixed-lineage leukemia is a particularly aggressive form of blood cancer which is very common in young children and rare in adults. Only about 30% of infants diagnosed with mixed-lineage leukemia survive beyond their fifth birthday. Although recent innovations in cancer therapy have improved the prospects of many leukemia patients, mixed-lineage leukemia is still very difficult to treat.

Our genes contain the instructions for cell growth and development. Genes are like the individual entries in an encyclopedia. Within each cell, the genes are arranged in 23 chromosomes, just as an encyclopedia is arranged in several volumes. Mixed-lineage leukemia arises when one of these chromosomes breaks and becomes rearranged. This rearrangement disrupts a gene which normally controls the development of blood cells. Blood cells containing the disrupted gene lose control of their development and become cancer cells. My doctoral research focused on understanding what goes wrong in cells when this gene is disrupted, and investigating why this particular gene seems to make mixed lineage leukemia so resistant to most cancer drugs.

For most of my research, I used an exciting new technology called CRISPR. CRISPR allows researchers to make very precise changes to our genes – like editing a specific word in a specific entry of the encyclopedia. If we don't understand what a gene does, we can use CRISPR to turn the gene off and see what happens. It would be difficult to do this in real leukemia patients, so instead I used leukemia cells growing in the lab.

Most genes, called coding genes, are blueprints for making proteins. These proteins interact with each other (and with the genes) to direct the cell's activities. Some genes, however, do not make proteins. Instead, these *non*-coding genes directly control other genes. The result is a huge and very complex network of interactions inside the cell. In the case of mixed-lineage leukemia, there is controversy about whether or not the non-coding genes are important. I used CRISPR to delete a particular non-coding gene in mixed-lineage leukemia cells. Deleting the gene did not have the effects I expected, suggesting that previous theories about its role in the disease may have been wrong.

Next I focused on mixed-lineage leukemia patients' resistance to a new cancer drug called venetoclax. Venetoclax has been remarkably effective at treating some kinds of leukemia but, disappointingly, not mixed-lineage leukemia. Venetoclax has a very interesting mechanism of action. The body contains trillions of cells and, over time, these cells can become damaged or exhausted. Many cells recognise when they are damaged and commit suicide in a process known as programmed cell death. When healthy cells turn into cancer cells, they normally switch off programmed cell death so that they can keep growing and growing, even as they become increasingly damaged. Venetoclax works by switching programmed cell death back on in cancer cells so that they die.

We have some ideas about why venetoclax is not very effective in patients with mixed-lineage leukemia. It is possible, for example, that cells with mixed-lineage leukemia switch off programmed cell death in many different ways at once. Maybe venetoclax only resets one of the switches while the others remain off. If we could find out what those other switches are, then perhaps we could use different cancer drugs in combination with venetoclax to reset all the switches at once and restore normal programmed cell death.

One way to find the right switches would be to test every single switch, one by one. Using drugs, this would be as difficult and laborious as it sounds. But the switches that control programmed cell death are proteins and the proteins are encoded by genes. Instead of drugs, which flip the switches at the protein level, we

can use CRISPR in the lab to flip the switches at the gene level. A remarkable new kind of CRISPR technology called CRISPR screening lets you turn off every single coding gene one by one in a very efficient way. I did this in mixed-lineage leukemia in the hope of finding a gene which made cells more sensitive to venetoclax.

Eventually I found a group of genes which seem to be switched off especially by mixed-lineage leukemia cells so that they can avoid programmed cell death. These genes are involved in the proteasome, a cellular machine that cleans up damaged or unwanted proteins. I tried simultaneously treating cells with venetoclax and a drug that targets the proteasome. Fortunately, this seemed to overcome venetoclax resistance. The next step is to test this combination in patients with mixed-lineage leukemia. Going forward, I plan to use CRISPR screening to understand and overcome resistance to many other promising cancer drugs in cancer types which are difficult to treat.

2

Methods and materials

Contents

2.1	Cloning	47
2.2	miR-196b knockout THP-1 cell lines	48
2.3	Human cell culture and electroporation	49
2.4	Cell viability assays	49
2.5	CRISPR knockout validation	51
2.6	Lentiviral transduction	54
2.7	Screening optimisation	55
2.8	Pooled miniscreen	56
2.9	Genome-wide CRISPR screen	61
2.10	Data analysis	67

Oligonucleotides (including primers) are listed in Appendix A. Plasmid maps are shown in Appendix B.

2.1 Cloning

Unless otherwise indicated, all oligos were ordered from Integrated DNA Technologies (lyophilised, standard desalting) and resuspended to 100 nM in ddH₂O. All enzymes used for cloning were supplied by New England Biolabs and used according to the manufacturer's instructions. All plasmids were verified by Sanger sequencing (Eurofins Genomics).

pL-CRISPR.EFS.GFP cloning

The lentiviral construct pL-CRISPR.EFS.GFP (Addgene plasmid #57818, a gift from Benjamin Ebert; Appendix B), was used to generate stable GFP⁺ miR-196b knockout THP-1 cell lines. As described previously, pL-CRISPR.EFS.GFP was linearised by digestion with BsmBI and dephosphorylated with Antarctic phosphatase, then agarose gel purified using the QIAquick Gel Extraction Kit (QIAGEN)²⁵⁷. sgRNA sequences KO1 and KO2 were optimised using CRISPOR. The scramble sequence was taken from Wang et al.²⁴⁸. Oligo pairs were annealed and phosphorylated using T4 PNK to generate double-stranded sgRNA inserts with overhangs suitable for BsmBI. Insert and vector were ligated for two hours at room temperature using T4 DNA ligase, then transformed by heat shock into Subcloning Efficiency DH5 α Competent Cells (Thermo Fisher). Plasmid DNA was purified using the QiaPrep Spin Miniprep Kit or Plasmid Midi Kit (QIAGEN).

lentiCRISPRv2 constructs

Apart from the GFP⁺ miR-196b knockout cells, all CRISPR knockouts in THP-1 cells were generated using the lentiviral construct lentiCRISPRv2 (Addgene plasmid #52961, a gift from Feng Zhang; Appendix B). sgRNA sequences were cloned into lentiCRISPRv2 as described previously, using the same strategy as for pL-CRISPR.EFS.GFP²⁵⁸.

2.2 miR-196b knockout THP-1 cell lines

72 hours after transduction with pL-CRISPR.EFS.GFP lentivirus, GFP⁺ THP-1 cells were sorted using a Sony SH800Z Cell Sorter and 100 μ m sorting chip. Untransduced cells were used to set FSC-A/SSC-A and FSC-A/FSC-H sorting gates. 24 hours after sorting, cells were resuspended in MethoCult H4434 Classic (Stemcell Technologies) and plated in technical triplicates of 210 cells/mL in 1.2 mL per plate. After incubating for 14 days, 40 single colonies per plate were picked using a P20 pipette tip into complete media and expanded.

14 days later, genomic DNA was extracted and the target *MIR196B* locus was PCR-amplified using 196B_T7-1_F/R and subsequently screened by T7E1 assay, Sanger sequencing and TaqMan assay. Sequencing traces were analysed using TIDE and, where necessary for transheterozygotes, deconvoluted using CodonCode Aligner v7.1.2^{259,260}.

2.3 Human cell culture and electroporation

Cell culture

Low-passage THP-1, HL-60 and U-937 cells (ATCC) were incubated at 37 °C in 95% air/5% CO₂ in RPMI 1640 (Thermo Fisher) supplemented with 1x Glutamax and 10% (THP-1, U-937) or 20% (HL-60) v/v sterile Fetal Bovine Serum (FBS) (“complete RPMI”). Cell density was maintained between 5×10^5 and 1.5×10^6 cells/mL by passaging every 48-72 hours. Cells were mycoplasma tested every two months.

Cells were cryopreserved in complete media v/v 5% (HL-60, U-937) or 7.5% (THP-1) DMSO (Thermo Fisher).

Electroporation

THP-1 cells were electroporated using a square-wave electroporation device (Fischer) and 4mm electroporation cuvettes (GeneFlow) using 12 µg midiprep plasmid DNA for 1×10^6 cells in 100 µL serum-free RPMI 1640. Cells were pulsed for 10 ms at 330 mV, allowed to recover for 5 minutes and then diluted to 1×10^6 cells/mL in prewarmed complete media. Typical cell survival after 72h was 20%.

2.4 Cell viability assays

Drug-based viability assays

Venetoclax (Stratech), etoposide (Cayman), A-1210477 (MedChemExpress) and bortezomib (MedChemExpress) were resuspended in DMSO (Thermo Fisher) according to the manufacturer’s instructions and stored at −80 °C. Thawed aliquots were never refrozen and activity of stock venetoclax was tested every six months in

THP-1 cells. Cells were treated with a range of drug concentrations in technical triplicates at a density of 2.5×10^5 cells/mL (2.5×10^4 cells per well in 96-well plate). For combination assays, drugs were pre-mixed using serial dilutions in 96-well plates. Etoposide (Cayman) 100 μ M was included as a positive control for apoptosis. DMSO v/v 0.66% was used as a negative (vehicle-only) control. After 48 hours, cells were harvested and analysed by Annexin-V/PI or CellTiter-Glo assay as indicated in the text.

Annexin-V/PI assay

The annexin-V/PI cell viability assay was conducted as described previously^{176,188}. For harvesting, cells were washed once in PBS (Thermo Fisher) and resuspended in 50 μ L annexin-V binding buffer (Biolegend) containing FITC-conjugated annexin-V (Biolegend) and propidium iodide (PI, ChemoMetec). After incubation on ice in the dark for 15 minutes, cells were diluted to a final volume of 150 μ L in annexin-V binding buffer and analysed using an Attune NxT flow cytometer (Thermo Fisher) with autosampler, obviating the need for cell counting beads. Healthy (DMSO-only) and single-stain apoptotic (etoposide-treated) cells were used to set an annexin-V⁻/PI⁻ gate.

CellTiter-Glo assay

The CellTiter-Glo Luminescent Cell Viability Assay 1.0 (Promega) was conducted as described previously using approximately 2.5×10^5 cells/mL²⁴³. The reagent was diluted 1:4 in PBS before use as recommended by Joung et al.²⁴³. Plates were imaged on a Promega GloMax luminometer.

IC₅₀ and synergy calculation

IC₅₀ was determined in GraphPad Prism by plotting an unconstrained least-squares regression of $\log_{10}[\text{drug}]$ vs. % viability. Where statistical tests were performed, IC₅₀ was compared to control using the Extra sum of squares F-test with $p = 0.01$.

The SynergyFinder Bioconductor package was used to analyse data from drug combination assays²⁶¹. Synergy scores were plotted with *seaborn* in Python.

Liquid proliferation assay

To determine their long-term proliferation in suspension culture, THP-1 cells were resuspended in complete media with 8 μ M venetoclax (MedChemExpress) at a density of 1×10^5 cells/mL with 2×10^4 cells/well in technical triplicate. Cells were seeded in venetoclax-free complete media as control. After five days, cell viability was analysed by CellTiter-Glo assay.

Colony assay

To determine their colony-forming capacity, THP-1 cells were resuspended in MethoCult H4434 Classic (Stemcell Technologies) and plated in technical triplicates of 210 cells/mL in 1.2 mL per plate. After incubating for 14 days, plates were scanned and colonies digitally counted using Fiji (ImageJ). Biological replicates were prepared on separate days.

2.5 CRISPR knockout validation

T7E1 assay

For T7 endonuclease I assays²⁶², genomic DNA was extracted from 1×10^6 cells by washing in PBS and incubating for 60min at 37 °C in 500 μ L proteinase-supplemented lysis buffer (25 mL Tris HCl pH8 1 M, 50 μ L EDTA 11 M, 0.25 μ L Tween 20, 17.2 mL MQ H₂O), a kind gift from Héctor S. Barbosa. CRISPR target loci were amplified from genomic DNA by PCR with Phusion High-Fidelity PCR Master Mix with GC Buffer (Thermo Fisher) and primers designed using CHOP-CHOP^{263,264}. Products were purified using the MinElute PCR Purification Kit (QIAGEN) and verified by Sanger sequencing (Eurofins). Purified PCR products were reannealed, incubated with T7 endonuclease I (NEB) according to the manufacturer's instructions and visualised by agarose gel electrophoresis with GelRed Nucleic Acid Gel Stain (Biotium).

TaqMan assay

TaqMan assays were conducted according to the manufacturer's TaqMan Small RNA Assays protocol. Small RNA was extracted from 1×10^6 PBS-washed THP-1 cells using the miRNeasy Mini Kit (QIAGEN). Reverse transcription probes for miR-196b-5p (TaqMan assay ID 002215) and RNU6B (001093) were combined into a single RT reaction as described previously with MultiScribe Reverse Transcriptase (Thermo Fisher). qPCR was performed on a CFX384 real-time thermocycler (Bio-Rad) with TaqMan universal PCR Master Mix II (no UNG) (Thermo Fisher). miRNA abundance fold-change was calculated using the $\Delta\Delta C_t$ method with the small nucleolar RNA RNU6B and the ubiquitous miRNA miR-17 as references and five scramble-KO cell lines as control^{265–267}.

miRNA activity fluorescent reporter assay

The fluorescent miRNA reporter plasmids (pBrown_Bi_Dir_2xMRE, Appendix B) were a generous gift from Yale Michaels²⁶⁸. They were cloned from the parent vector AB.pCCL.sin.cPPT.GFP.miR-17-3p.sensor.PGK.dNGFR.WPRE, a gift from Brian Brown (Addgene plasmid #85866). An NGFR-only plasmid was used to determine background GFP fluorescence. The normalised median fluorescent intensity of EGFP was calculated using the following formula:

$$\text{normalised MFI}_{EGFP} = \frac{MFI_{EGFP}(\text{test}) - MFI_{EGFP}(\text{NGFR-only})}{MFI_{EGFP}(\text{scramble}) - MFI_{EGFP}(\text{NGFR-only})}$$

Target gene RT-qPCR

RNA for RT-qPCR was extracted from 1×10^6 PBS-washed THP-1 cells per sample using the RNeasy Mini Kit (QIAGEN). All primers were validated by melt curve analysis and standard-curve assay (Figure 2.1) using template cDNA generated from WT THP-1 cells. Gene expression fold-change was calculated using the $\Delta\Delta C_t$ method with *GAPDH* as reference gene and five scramble-KO cell lines as control.

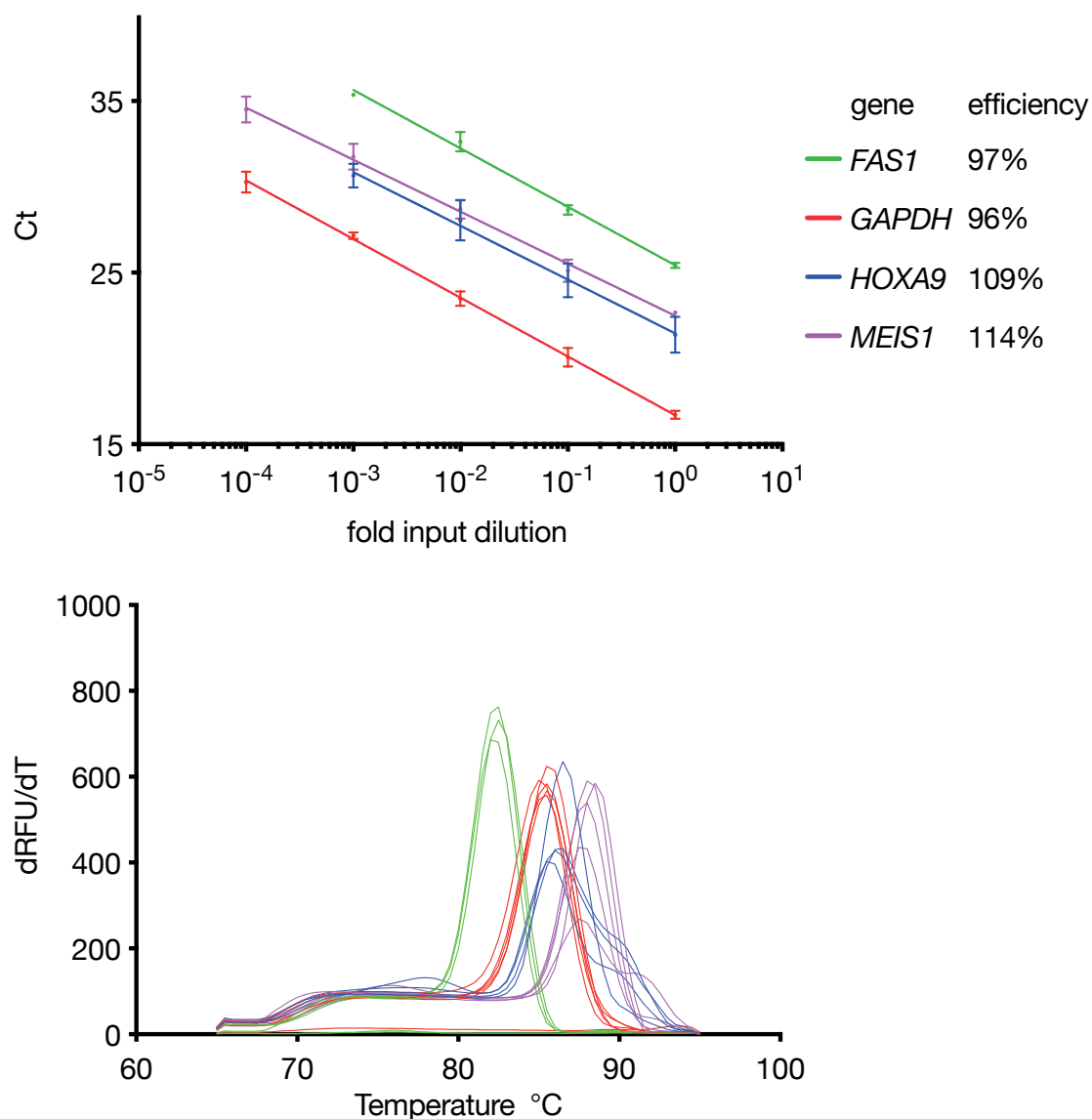


Figure 2.1: qPCR primer validation. Lines were fitted using semilog least squares regression ($R^2 > 0.95$). Efficiency is calculated as $(10^{\frac{-1}{slope}} - 1) * 100$.

Western blotting

See Table 2.1 for details of antibodies used for western blotting. THP-1 whole-cell lysates were prepared using 1×10^6 cells in NuPage LDS Sample Buffer (Invitrogen NP0007) supplemented with protease inhibitor according to the manufacturer's protocol and then sonicated twice (five cycles of 30s on, 30s off) with a Bioruptor Plus sonicator. Samples were run in NuPAGE Novex 4-12% Bis-Tris 1.5mm protein gels (Invitrogen NP0336) alongside Precision Plus Protein Kaleidoscope Standards

Reagent	Supplier	Supplier code	Purpose	Dilution
anti-vinculin	Abcam	ab129002	primary Ab	1:10,000
anti-BIM	Abcam	ab32158	primary Ab	1:1000
Goat Anti-Rabbit	Abcam	ab6721	secondary Ab	1:10,000
anti-BAX	Cell Signaling	2772	primary Ab	1:1000
anti-BAK D2D3	Cell Signaling	6947	primary Ab	1:1000
anti-MCL-1	Cell Signaling	5453	primary Ab	1:1000
anti-BCL-XL	Cell Signaling	2764	primary Ab	1:1000

Table 2.1: Antibodies used for western blotting. All primary mAbs were from rabbit and all secondary mAbs were from goat.

(Bio-Rad 1610375) at 150V using pre-chilled NuPage MOPS SDS running buffer (Invitrogen NP0001). Protein was transferred to 0.2 μ m nitrocellulose membrane using an iBlot 2 semi-dry transfer device and iBlot 2 Transfer Stacks (Invitrogen IB23001) according to a program optimised for small proteins: 20V 1min, 23V 4min, 25V 4min.

Membranes were blocked for 90min in 5% milk powder (Marvel) w/v TBS-T (Tris-buffered saline v/v 0.1% Tween-20). After blocking, membranes were washed for three cycles of five minutes each in TBS-T and incubated overnight at 4°C with primary antibody diluted in 5% BSA Heat Shock Fraction V (Sigma Aldrich) w/v TBS-T. Where appropriate, primary antibodies were combined in the same incubation. Membranes were washed as before and incubated for 60min at room temperature with horseradish peroxidase-conjugated goat anti-rabbit secondary antibody. Finally, membranes were washed as before and imaged using Pierce ECL (Thermo Fisher) and a BioRad ChemiDoc. Uncropped blots are shown in Appendix C.

2.6 Lentiviral transduction

Lentivirus production

Lentivirus was produced by Ryan Beveridge in the WIMM Viral Production Facility. HEK293T cells were plated onto 8x15cm tissue culture treated dishes at 1.2×10^7 cells per dish. The next morning, the cells were transfected with the TKOv3 plasmid

library and lentiviral helper plasmids psPAX2 and pMD2.G following the ratios outlined in²³⁹ to a total of 30 µg per dish (16.67:8.33:5.0) using PEIPro (Polyplus Transfection) at 1 µL per 1 µg DNA following the manufacturer's guidelines. Six hours after transfection, the media was removed and replaced with 15 mL fresh complete DMEM. 48 hours after transfection the viral harvest was collected and stored at 4 °C and 15 mL fresh DMEM was added. 72 hours after transfection the second viral harvest was collected, combined with harvest one, centrifuged at 2000g for 5 minutes at 4 °C and filtered using a 0.45 µm cellulose acetate vacuum filter (Sigma CLS430768-12EA). After filtration the virus was aliquoted into 1 mL or 10 mL aliquots and stored at −80 °C.

Lentiviral transduction

Unless otherwise indicated, THP-1 cells were transduced with lentivirus at MOI=0.3 in six-well plates with 3×10^6 cells/well in a total volume of 3 mL/well by centrifugation at 800g for 60 minutes. Cells were incubated for 2 hours after transduction and then diluted to 5×10^5 cells/mL in complete media.

Puromycin selection

72 hours after transduction, cells transduced with lentiCRISPRv2 were resuspended in complete media with 2 µg/mL puromycin dihydrochloride (Thermo Fisher) at a density of 2.5×10^5 cells/mL with 2.5×10^4 cells/well in technical triplicate. Cells were seeded in complete media without puromycin as control. 72 hours later, cells were analysed by CellTiter-Glo assay. Survival in puromycin relative to no-puromycin control was plotted in GraphPad Prism to calculate multiplicity of infection.

2.7 Screening optimisation

Optimising puromycin selection

pLenti CMV GFP Puro (658-5) (Addgene plasmid #17448, a gift from Eric Campeau and Paul Kaufman²⁶⁹) is a third generation lentiviral vector expressing PuroR and

GFP. This construct was used to optimise the puromycin concentration such that % survival under puromycin selection would accurately reflect % transduction. For five biological replicates, cells were transduced with five independent serial dilutions of pLenti CMV GFP Puro. After 72 hours, cells were selected with puromycin as indicated above at 0 µg/mL, 0.5 µg/mL, 1 µg/mL or 2 µg/mL. After 72 hours, all cells were analysed by CellTiter-Glo assay as indicated above; control-treated cells were analysed using an Attune NxT flow cytometer to determine % GFP⁺.

Determining viral titre

All lentiCRISPRv2 viral preps were titred in THP-1 cells prior to screening experiments. Briefly, three independent replicate serial dilutions of virus were prepared, each with six dilutions (1, 0.5,..., 0). THP-1 cells were transduced as described above. After three days, cells were selected with or without puromycin as described above and analysed by CellTiter-Glo assay. Cell survival vs. viral dilution was plotted and a least squares linear regression was fitted to the linear portion of the plot. This best fit was analysed to determine viral titre as described previously²⁷⁰. See chapter 4 for further details.

2.8 Pooled miniscreen

The protocol for the pooled miniscreen was adapted from protocols previously described by the Zhang Lab²⁴³.

Pooled miniscreen: library cloning and lentivirus production

The pooled miniscreen library included a total of 16 sgRNAs: two sgRNAs per gene for each of seven protein-coding genes, one sgRNA targeting a silent site (*AAVS1*) and one non-targeting sgRNA (sgCTRL-1). For each of the protein-coding genes and *AAVS1*, the two sgRNAs (or, for *AAVS1*, the one sgRNA) with the highest sgRNA score were selected from the TKOv3 library²⁷¹. sgCTRL-1 was derived from a published pooled CRISPR essentiality screen²⁴⁸. Each sgRNA

sequence was individually cloned into lentiCRISPRv2 as described above. The resulting plasmids were pooled at equimolar ratio and used to generate a single pool of lentivirus as described above. *BCL2*-targeting sgRNAs were not included in the miniscreen library.

Pooled miniscreen: lentiviral transduction and selection

For the pooled miniscreen, 7×10^7 THP-1 cells were transduced with pooled lentivirus six days before T0 at MOI=0.3 by spinfection in a 50mL Falcon tube. Following transduction, cells were diluted to 5×10^5 cells/mL in complete media and cultured for three days.

Three days before T0, at least 1.92×10^8 cells were selected with puromycin to eliminate untransduced cells. A small fraction of cells was reserved to confirm the MOI as described above.

Pooled miniscreen: selection

On T0, after confirming MOI=0.3, three technical replicates of 3×10^6 cells (equivalent to a coverage of 2×10^5 cells/sgRNA) were harvested and washed in ice-cold PBS. Cell pellets were dried and snap-frozen on dry ice for storage at -80°C . The remaining cells were passaged into one of two conditions, DMSO-only or VEN (three replicates per condition). For the DMSO-only condition, 2.8×10^6 cells were passaged at a density of 5×10^5 cells/mL in complete media v/v 0.074% DMSO. For the VEN condition, 5.6×10^6 cells were passaged at a density of 1×10^6 cells/mL in complete media containing $7.4\mu\text{M}$ VEN.

Cells were cultured for a total of five passages (15 days). At each passage, 3×10^6 cells per technical replicate were harvested and stored as described above. The remaining cells were centrifuged at $100g$ for 10 minutes and resuspended in fresh DMSO-only media at 5×10^5 cells/mL or fresh VEN media at 1×10^6 cells/mL.

Pooled miniscreen: genomic DNA harvesting and NGS library preparation

The pooled miniscreen generated 33 cell pellets (three technical replicates from T0 and three technical replicates for each of two conditions from T1-T5). gDNA was harvested using the DNeasy Blood and Tissue Kit (QIAGEN) according to the manufacturer's protocol.

Two consecutive PCR steps, summarised in Table 2.2, were used (1) to amplify sgRNA sequences and (2) to add sequencing adaptors and indices. All PCR was performed using Kapa HiFi HotStart ReadyMix (Roche) and a Bio-Rad C1000 thermal cycler. Forward primers, shown in Table 2.3, included stagger sequences to improve NGS cluster generation. For PCR1, 0.5 µg Nanodrop-quantified input was used per sample. Assuming that the mass of genomic DNA in a diploid THP-1 cell is 6 pg, this equates to an input of $(0.5 \mu\text{g} \div 6 \text{ pg per cell} =) 8.3 \times 10^4$ cells. For the miniscreen sgRNA pool this ensures a coverage of $(8.3 \times 10^4 \text{ cells} \div 16 \text{ sgRNA} =) \sim 5200$ cells/sgRNA. After PCR1, the size of the product (600bp) was verified by gel electrophoresis. The samples were pooled at approximately equal concentrations based on relative band intensities (analysed using Bio-Rad Image Lab v5.2.1). 5 µL unpurified pooled PCR1 product was used as input for PCR2. The PCR2 product was gel purified twice using the QIAquick Gel Extraction Kit with MinElute columns (QIAGEN). After checking with Sanger sequencing, the product was further purified using Agencourt AMPure XP beads (Beckman Coulter) and quantified using the Qubit dsDNA HS assay.

PCR1		
Step	Temp (°C)	Time (s)
initial denaturation	98	30
	98	10
x25	66	30
	72	15
final extension	72	120
hold	10	∞

PCR2		
Step	Temp (°C)	Time (s)
initial denaturation	98	30
	98	10
x10	55	30
	65	15
final extension	65	300
hold	10	∞

Table 2.2: PCRs for pooled screen NGS library preparation. Two-step PCR was used to prepare amplicon libraries from gDNA for next-generation sequencing. The same strategy was used for both the miniscreen and the genome-wide screen. In PCR1, a targeted amplicon containing the sgRNA sequence was amplified from sufficient gDNA input for a coverage of at least 120 cells/sgRNA. In PCR2, sequencing adaptors and sample-specific indices were added to each amplicon using a low cycle number to reduce PCR bias.

PCR1		
	F	R
all samples	M_0143_v2.1-F1 Moffat	M_0144_v2.1-R1 Moffat
PCR2		
	F	R
T0-A	D506 ATCGA	D701
T0-B	D506 ATCGA	D702
T0-C	D506 ATCGA	D704
T1A-A	D501	D701
T1A-B	D501	D702
T1A-C	D501	D704
T1D-A	D501	D705
T1D-B	D501	D706
T1D-C	D501	D707
T2A-A	D502 A	D701
T2A-B	D502 A	D702
T2A-C	D502 A	D704
T2D-A	D502 A	D705
T2D-B	D502 A	D706
T2D-C	D502 A	D707
T3A-A	D503 GA	D701
T3A-B	D503 GA	D702
T3A-C	D503 GA	D704
T3D-A	D503 GA	D705
T3D-B	D503 GA	D706
T3D-C	D503 GA	D707
T4A-A	D504 CGA	D701
T4A-B	D504 CGA	D702
T4A-C	D504 CGA	D704
T4D-A	D504 CGA	D705
T4D-B	D504 CGA	D706
T4D-C	D504 CGA	D707
T5A-A	D505 TCGA	D701
T5A-B	D505 TCGA	D702
T5A-C	D505 TCGA	D704
T5D-A	D505 TCGA	D705
T5D-B	D505 TCGA	D706
T5D-C	D505 TCGA	D707

Table 2.3: Primers for pooled miniscreen NGS library preparation.

Pooled miniscreen: next-generation sequencing

The pooled library was quantified using the Kapa Library Quantification Kit (Roche), combined with 10% PhiX Control V3 DNA (Illumina) and sequenced on an Illumina MiSeq with a paired-end 150-cycle MiSeq Nano kit using dual-end indexing: 151 forward reads + 8 index reads, 151 reverse reads + 8 index reads.

Pooled miniscreen: data analysis

Raw fastq files were transferred from Illumina Basespace Sequencing Hub to the University of Oxford Computational Biology Research Group servers for downstream analysis. Quality control was performed using MultiQC²⁷². Next, forward and reverse reads were merged using BBMerge with default quality filtering. >98% of reads were successfully merged. Raw sgRNA counts were generated using CRISPRessoCount in Python v2.7 with a quality score threshold of 30²⁷³. sgRNA counts for each sample were compiled and plotted using *matplotlib*, *pandas* and *seaborn* in Python v3.0.

2.9 Genome-wide CRISPR screen

The protocol for genome-wide screening was adapted from protocols previously described by the Moffat Lab²⁷⁴ and Zhang Lab²⁴³. See chapter 4 for an account of how the screening protocol was optimised for THP-1 cells.

TKOv3 library amplification, validation and lentivirus production

The Toronto human knockout pooled library (TKOv3) in a lentiCRISPRv2 backbone was a gift from Jason Moffat (Addgene #90294). The pooled plasmid library was amplified by electroporation in Lucigen Endura electrocompetent cells at a minimum coverage of 250x as described previously²⁷⁴.

The amplified library was prepared for Illumina amplicon sequencing by two-step PCR (Table 2.4, Table 2.5). PCR1 was performed using Phusion High-Fidelity PCR Master Mix with GC Buffer (Thermo Fisher). After PCR1, 45 μ L of the

PCR1			
	F		R
plasmid library	M_0143_v2.1-F1 Moffat		M_0144_v2.1-F1 Moffat
PCR2			
	F		R
plasmid library	D701		D501

Table 2.4: Primers for TKOv3 plasmid NGS library preparation

PCR product was purified using 81 μ L of Mag-Bind TotalPure NGS beads (VWR M1378-00) following the manufacturer’s instructions. PCR2 was performed using Kapa HiFi HotStart ReadyMix (Roche). The product was size-verified using the Bioanalyzer (Agilent) and sequenced using NextSeq 500/550 High Output Kit v2.5 (75 Cycles) (Illumina 20024906).

Reads were trimmed using *cutadapt* v1.8.3 in Python v2.7 with an error threshold of 30.0% in single-end mode²⁷⁵. Guide counts were determined using *count_spacers.py* in Python v2.7²⁴³.

After the plasmid library had been validated by sequencing, it was used to prepare and titre lentivirus as described above.

Genome-wide screen: lentiviral transduction and selection

The genome-wide screen was conducted in two independent consecutive biological replicates (screen 1 and screen 2). For each replicate, THP-1 cells were transduced and puromycin-selected as a single pool, then separated at T0 (the first day of screening) into the control group (‘DMSO’) and treatment group (‘VEN’). Six days before T0, at least 5.76×10^8 THP-1 cells were transduced with TKOv3 lentivirus at multiplicity of infection MOI=0.3 by transduction in two batches of 16 six-well plates. Following transduction, cells were diluted to 5×10^5 cells/mL in complete media, pooled into four T175 flasks per batch and cultured for three days.

Three days before T0, at least 9.43×10^8 cells were selected with puromycin to eliminate untransduced cells. Briefly, cells from each transduction batch were separately pooled and counted. A small fraction of cells was reserved to measure

PCR1		
Step	Temp (°C)	Time (s)
initial denaturation	98	30
	98	10
x25	66	30
	72	15
final extension	72	120
hold	4	∞

PCR2		
Step	Temp (°C)	Time (s)
initial denaturation	98	180
	98	30
x8	55	30
	72	30
final extension	72	300
hold	4	∞

Table 2.5: PCRs for TKOv3 plasmid NGS library preparation.

the MOI of each transduction batch as described above. The remaining cells were resuspended in complete medium containing 2 $\mu\text{g}/\text{mL}$ puromycin at a density of 2.5×10^5 cells/mL, separated into 16 T175 flasks per transduction batch and cultured for three days.

Genome-wide screen: selection

On T0, after confirming MOI=0.3 for each transduction batch, all cells were pooled and counted. 1.8×10^7 cells (for a minimum coverage of 250 cells/guide) were harvested and washed in ice-cold PBS. Cell pellets were dried and snap-frozen on dry ice for storage at -80°C . The remaining cells were passaged into two conditions by centrifugation at 100g for ten minutes and resuspension in fresh media. For the DMSO condition, at least 2.7×10^7 cells were passaged at a density of 5×10^5 cells/mL in complete media v/v 0.074% DMSO in one T175 flask. For the VEN condition, at least 4.72×10^8 cells were passaged at a density of 1×10^6 cells/mL in complete media containing 7.4 μM VEN distributed across four T175 flasks.

Cells were cultured for a total of six passages (18 days). At each passage, cells from the DMSO condition were counted and at least 2.7×10^7 cells were resuspended at a density of 5×10^5 cells/mL in fresh complete media v/v 0.074% DMSO. Cells from the VEN condition were pooled, counted and at least 2.25×10^8 cells were passaged at a density of 1×10^6 cells/mL in fresh complete media containing 7.4 μ M VEN. At the final passage (T6), 1.8×10^7 cells were harvested from both the DMSO-only and VEN condition.

Genome-wide screen: genomic DNA extraction and NGS library preparation

Amplicon libraries were prepared and sequenced separately for screen 1 and screen 2. For each screen, genomic DNA was extracted from one harvested cell pellet per condition (T0, DMSO and VEN) using the QIAamp DNA Blood Maxi Kit (QIAGEN) as previously described²⁷⁴. gDNA was quantified using the Qubit dsDNA HS assay and Qubit Fluorometer 2.0 (ThermoFisher) according to the manufacturer's protocol.

The PCR steps for library preparation for the genome-wide screen were essentially the same as those for the pooled miniscreen (Table 2.2), except different indexing primers (Table 2.6) were used. See chapter 4 for a description of how this protocol was optimised. In PCR1, 51 μ g of input gDNA per condition was distributed across 34 individual 50 μ L reactions, each containing 1.5 μ g input DNA. Figure 2.2 shows how gDNA input amount was optimised. Assuming that the mass of genomic DNA in a diploid THP-1 cell is 6 pg, this equates to an input of (51 μ g / 6 pg per cell =) 8.5×10^6 cells²⁷⁶. Given the size of the TKOv3 library, this ensures a coverage of (8.5×10^6 cells / 71090 sgRNAs =) 120 cells/sgRNA. Reaction products were pooled for each condition and the size of the product (600bp) was verified by agarose gel electrophoresis. The product was column-purified using the MinElute PCR Purification Kit (QIAGEN) and 5 μ L was used as input for PCR2, in which each condition was labelled with a unique sequencing index. PCR2 product was gel purified twice using the QIAquick Gel Extraction Kit with MinElute

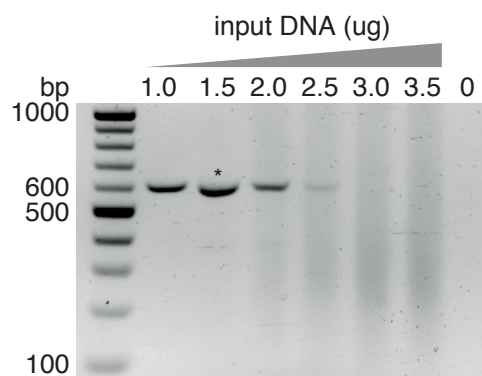


Figure 2.2: NGS PCR1 optimisation. Amplifying from ~50 µg genomic DNA ensures coverage greater than 100 cells/sgRNA. In order to minimise the number of individual 50 µl PCR reactions, it is necessary to determine the maximum possible gDNA input that does not inhibit polymerase activity. The optimum input is marked *. This depends on the purity of input gDNA.

PCR1		
	F	R
all samples	M_0143_v2.1-F1 Moffat	M_0144_v2.1-F1 Moffat
PCR2		
	F	R
T0	D705	D501
T6 DMSO-only	D707	D501
T6 VEN	D706	D501

Table 2.6: NGS primers for genome-wide screen.

columns (QIAGEN). After checking with Sanger sequencing, the product was further purified using Agencourt AMPure XP beads (Beckman Coulter) and quantified using the Qubit dsDNA HS assay.

Genome-wide screen: next-generation sequencing

Each uniquely indexed library was quantified using the Kapa Library Quantification Kit (Roche). T0, T6 DMSO-only and T6 VEN were pooled at equimolar ratio and then pooled with a high-diversity library (a gift from Nicholas Crump) at 3:7 ratio. The final pool was sequenced on an Illumina NextSeq with a paired-end 150-cycle (high) NextSeq kit using single-end indexing: 78 forward reads +

8 index reads, 78 reverse reads.

Genome-wide screen: data analysis

Raw fastq files were transferred from Illumina Basespace Sequencing Hub to the University of Oxford Computational Biology Research Group servers for downstream analysis. Quality control was performed using MultiQC²⁷². Next, forward and reverse reads were merged using BBMerge with default quality filtering. >95% of reads were successfully merged. Raw sgRNA counts were generated using CRISPEssCount in Python v2.7 with a quality score threshold of 30²⁷³.

```
> BAGELp -i [input_fold_changes] -o [essential_genes]
          -e [nonessential_genes] -c 1
```

Raw guide counts were analysed with MAGECK v0.5.7 using the following command²⁷⁷.

```
> mageck test -k [input_read_counts] -t [test_read_counts]
          -c [control_read_counts] -n [output_file_name]
          --cnv-norm [THP-1_CNV_data]
          --cell-line THP1_HAEMATOPOIETIC_AND_LYMPHOID_TISSUE
          --additional-rra-parameters "--permutation 10000 "
```

THP-1 gene copy numbers derived from Affymetrix SNP6.0 arrays were downloaded from the Cancer Cell Line Encyclopedia²⁷⁸.

Volcano plots, rank plots and gene ontology plots were generated with MAGECK-Flute v1.3.1 and Bioconductor in R v3.6.0²⁷⁹. Scatter plots, Venn diagrams and correlation and regression analyses were generated with the packages *matplotlib*, *pandas* and *seaborn* in Python v3.0. Data were also visualised interactively in MAGECK-VISPR in Python v2.7²⁸⁰. GO analysis was performed using ShinyGo with Bonferroni multiple comparison correction^{281,282}.

Validating TKOv3 hits

Hits from the TKOv3 whole-genome screen were validated using sgRNA sequences drawn from the TKOv3 and Brunello screening libraries as indicated in the text. sgRNAs were individually cloned into lentiCRISPRv2. Cells were transduced and selected as described above. Cells were recovered in complete media for three days

before selection with 7.4 μ M or 20 μ M VEN or DMSO-only control in technical triplicates. After 48 hours, cell viability was measured by CellTiter-Glo assay. For biological triplicates, the entire experiment was repeated on different days, using three independent lentivirus preps from the same plasmid stock.

2.10 Data analysis

Unless otherwise stated, all graphs were generated using GraphPad Prism v7.0a or v8.0a or with the packages *matplotlib*, *pandas*, Phyla textAdjust²⁸³ and *seaborn* in Python. Data were managed in Jupyter notebooks using Anaconda virtual environments running Python v2.7 or v3.0. All flow cytometry data were analysed using FlowJo v10.4.2.

3

Investigating the role of miR-196b in THP-1 cell proliferation

Contents

3.1	Introduction	69
3.2	Results	70
3.3	Discussion	78

3.1 Introduction

In this chapter, I explore the role of miR-196b in creating and maintaining HSC-like gene expression in MLLr leukemia. The *MIR196B* locus lies on chromosome 7 within an intergenic region between *HOXA9* and *HOXA10* (Figure 3.1A). *MIR196B* expresses a bicistronic pri-miRNA which is processed to yield two mature miRNAs, miR-196b-3p and miR-196b-5p (henceforth referred to as miR-196b)²⁸⁴. Here, I develop a lentiviral CRISPR-based strategy for inactivating miRNA precursor genes and I apply this to miR-196b in an MLL-AF9 cell line. Having generated and functionally validated miR-196b knockout cell lines, I study their clonogenicity, drug resistance and gene expression to test the Chen model, a proposed model for the role of miR-196b in MLLr leukemia. I consider whether the target network of

miR-196b could be investigated using a pooled CRISPR screen and I shed light on the status of miR-196b as an oncogenic miRNA.

3.2 Results

3.2.1 CRISPR knockout of miR-196b in THP-1 cells

Previous work has established that CRISPR/Cas9 provides an effective means for repressing miRNA expression²⁸⁵. Indels within the Microprocessor/Dicer processing sequences or seed sequence prevent the biogenesis or activity of the mature miRNA. I reasoned, therefore, that knockout of the mature sequence of miR-196b within *MIR196B* would result in loss of miR-196b activity. I selected two target sites in the *MIR196B* locus proximal to the miR-196b seed sequence (Figure 3.1B). I analysed the corresponding sgRNAs (KO1 and KO2) with CRISPOR and found that they both had high specificity and efficiency scores²⁸⁶.

pL-CRISPR.EFS.GFP (Figure 3.1C) is a second-generation lentiviral vector expressing an sgRNA under the PolIII-driven U6 promoter and SpCas9 under the PolII-driven EF1 α promoter. EGFP is coupled to SpCas9 with a P2A cleavage site and coexpressed as a selection marker^{257,287}. I cloned sgRNA sequences KO1 and KO2 into pL-CRISPR.EFS.GFP to generate two separate variants targeting *MIR196B*: pL.CRISPR.EFS.GFP_MIR-196b-KO-1 and pL.CRISPR.EFS.GFP_MIR-196b-KO2. A non-targeting sgRNA sequence was used to generate a control plasmid, pL.CRISPR.EFS.GFP_scramble. Plasmids were used to make lentivirus which I transduced individually into THP-1 cells before isolating EGFP⁺ cells via FACS (Figure 3.1E).

To assess Cas9-mediated editing, I amplified the *MIR196B* locus by PCR, melted and reannealed the product and digested with T7 endonuclease, an enzyme which recognises and cleaves DNA heteroduplexes (Figure 3.1D)²⁶². This demonstrated that both KO1 and KO2 led to indel generation, although KO1 was more efficient than KO2.

I further characterised cleavage with KO1 and KO2 using Sanger sequencing. Non-homologous end joining (NHEJ)-mediated repair of double strand breaks leads to a distribution of indel sizes, typically from -10bp to +2bp²⁸⁸. These mutations remain difficult to predict²⁸⁹. Sequencing an amplified target locus from a non-clonal edited population should generate a single trace if the locus has not been edited (or two traces for a heterozygous locus), but a convoluted trace if the locus has been edited. I used the TIDE assay, which analyses such convoluted traces in comparison to control traces, to confirm that both KO1 and KO2 generated a DSB 2bp upstream of the protospacer-adjacent motif (PAM), as expected. A representative TIDE trace for KO1 is shown in Figure 3.1F²⁶⁰.

3.2.2 Generating and validating miR-196b knockout THP-1 cell lines

An issue with single-sgRNA knockout of miRNA loci is that small indels, even in the seed sequence, may not fully ablate activity of the mature miRNA. In this case, a minimum deletion of 3bp would be necessary for KO1 to disrupt the seed sequence²⁸⁸. Residual miRNA activity may mask the mutant phenotype or make the knockout phenotype harder to detect. Dual-sgRNA editing is an alternative strategy in which two sgRNAs are coexpressed to generate two simultaneous DSBs, leading to deletions of up to several kilobases between the cleavage sites^{290,291}. This offered a straightforward method to delete the entire seed sequence of miR-196b. Besides improving the degree of miRNA knockdown, dual-sgRNA editing has the advantage of generating clones which can be rapidly screened by PCR because they contain larger deletions.

Having validated that KO1 and KO2 were independently effective, I cotransduced pL.CRISPR.EFS.GFP_MIR-196b-KO-1 and pL.CRISPR.EFS.GFP_MIR-196b-KO2 lentivirus into THP-1 cells (Figure 3.2B). After sorting EGFP⁺ cells, I plated them at single-cell density in MethoCult methylcellulose medium and incubated until clonal colonies had formed. I picked individual colonies and screened them by PCR based on the size of the *MIR196B* target region relative to unedited

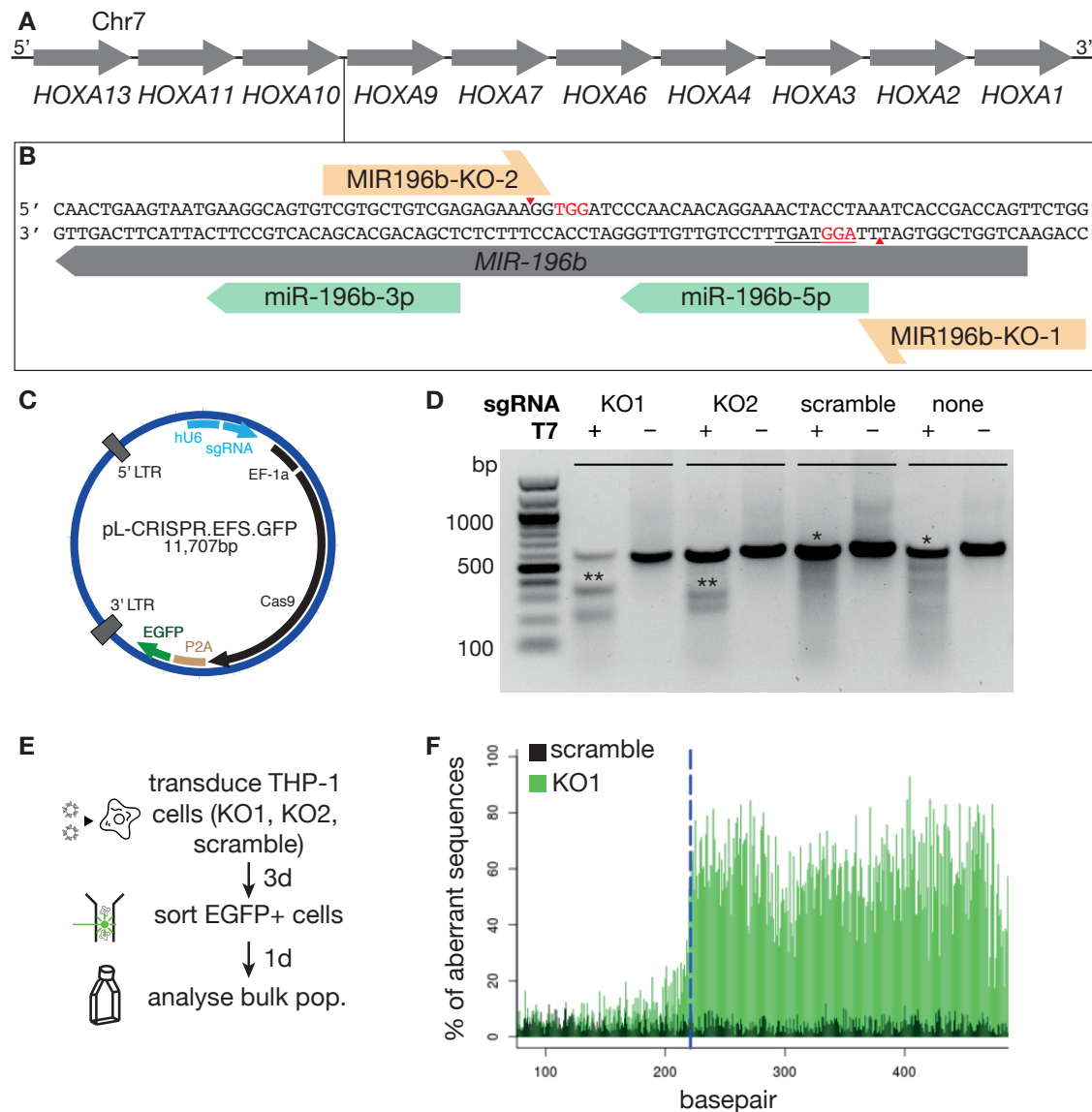


Figure 3.1: CRISPR-KO of miR-196b. (A) The *MIR196B* locus is located in the *HOX* gene cluster on chromosome 7 between *HOXA9* and *HOXA10*. (B) Pri-miR-196b is processed into two mature miRNAs (green). This chapter focuses on miR-196b-5p. MIR196b-KO-1 and MIR196b-KO-2 (orange) are two sgRNAs complementary to spacers near the seed sequence of miR-196b-5p (underlined). Predicted Cas9 cleavage sites (red triangles) are 2bp upstream of the protospacer-adjacent motifs (red). (C) Map of pL-CRISPR.EFS.GFP parent vector. (D) Agarose gel to show T7E1 assay on *MIR196B* locus from THP-1 cells edited using pL-CRISPR.EFS.GFP lentivirus expressing sgRNA KO1, KO2 or non-targeting scramble control or unedited cells. Edited loci are cleaved by T7E1 (**); unedited loci are not cleaved (*). (E) Schematic: single-sgRNA CRISPR editing. (F) TIDE assay output based on Sanger sequencing trace of *MIR196B* locus amplified from THP-1 cells edited with KO1. Dashed line indicates predicted CRISPR cleavage site.

control. Sanger sequencing confirmed that four clones contained a deletion of the entire mature miR-196b sequence (Figure 3.2A). In parallel, I generated five clonal lines from scramble-edited THP-1 cells to control for clonal variation. A TaqMan assay demonstrated that miR-196b expression in the four KO clones was reduced relative to scramble control clones by a mean fold-change of 708 ($\Delta\Delta C_t=9.47$, $SD=0.65$) (Figure 3.2C). miR-196b was present at very high amplification cycles but this was probably due to background effects. Since expression of miR-196b was comparable in all five scramble clones, I used scramble 1 as a control line for all subsequent experiments.

The TaqMan assay suggested that there is minimal clonal heterogeneity in the expression of miR-196b and miR-17 in THP-1 cells. Based on the scramble clones, I found that the normal expression of miR-196b was 99.4-fold lower than that of miR-17 ($\Delta\Delta C_t=6.64$, $SD=0.26$). I used published miRNA expression data to confirm that miR-196b is expressed in THP-1 cells, but at a lower rate than miR-17. Ranked by miR-196b expression, THP-1 cells are in the upper quartile of all 169 human cell lines listed in miRmine, a curated database of miRNA expression²⁹². The median reported expression of miR-196b in THP-1 cells is $\log_2 \text{RPM} = 8.91$, compared to $\log_2 \text{RPM} = 12.11$ for miR-17.

3.2.3 Phenotypic analysis of miR-196b knockout THP-1 cells

The miR-196b TaqMan assay confirmed the successful knockout of miR-196b in the KO clones. Next I set out to explore the resulting phenotype in terms of clonogenicity and gene expression. Colony forming assays measure the clonogenic capacity of hematopoietic cells²⁹³. Based on studies using antagomir-196b in mice⁸⁴, I predicted that removing miR-196b would decrease colony formation. As expected, the colony forming assay revealed that miR-196b KO cells formed colonies around 50% less than scramble control cells (Figure 3.3A).

Deletion of a miRNA should derepress its target genes. According to the Chen model, miR-196b contributes to the enhanced proliferation of MLLr leukemia

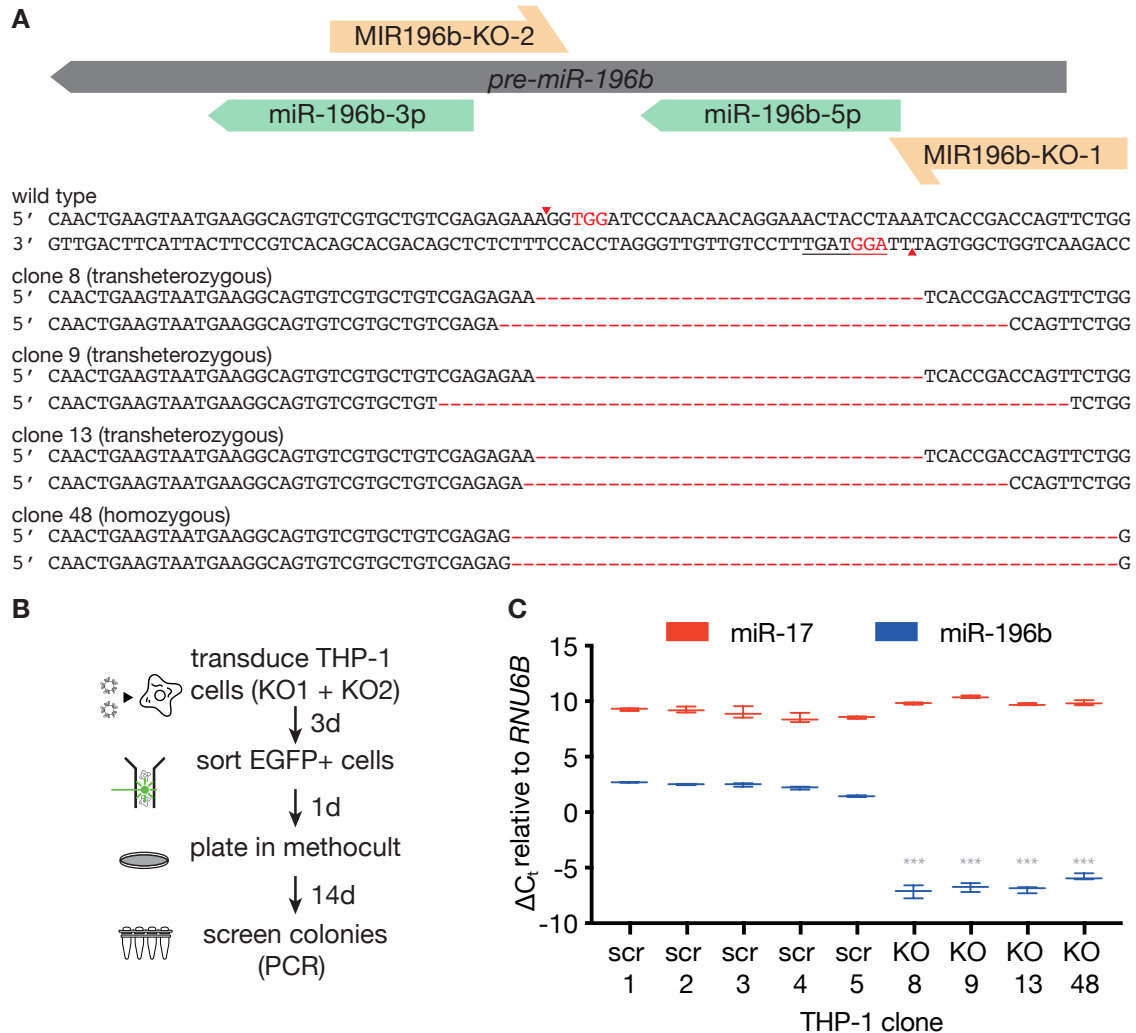


Figure 3.2: Generating miR-196b KO cell lines. (A) 5' sequences of both alleles of *MIR196B* for each of four THP-1 cell lines edited with KO1+KO2. Deletions are marked with dashes. Wild type sequence shown for reference; predicted Cas9 cleavage sites (red triangles) are 2bp upstream of the protospacer-adjacent motifs (red). NB that the terminal 4bp of the deleted region in clone 48 extends beyond the 5' end of *MIR196B*. The 5' end of the *HOXA10* ORF is over 1kb downstream. (B) Schematic: generating dual-sgRNA-edited clonal cell lines. (C) TaqMan qPCR assay on THP-1 miR-196b KO clones and scramble control clones. Bars show mean $\pm$ SD. n=3 biological replicates per cell line. Mean miR-196b expression in each KO line was compared to mean miR-196b expression across all scramble control lines using ordinary one-way ANOVA with Dunnett's multiple comparisons test. *** p<0.001.

cells by silencing the expression of *FAS*¹²¹. The same study suggests that miR-196b may also interact with *HOXA9* and *MEIS1*. All three genes have been bioinformatically predicted as targets of miR-196b and they have a well-studied role in MLLr leukemia¹²¹. I tested this hypothesis by examining whether expression of these three genes was altered in the miR-196b KO clones (Figure 3.3B). Surprisingly, differences in expression between these genes were negligible and inconsistent between the four clones. *HOXA9* and *FAS* were more variable than *MEIS1*. *HOXA9* and *FAS* were upregulated ~10-fold in clone 9 but downregulated ~5-fold in clone 48. Thus, although ablation of miR-196b reduced the colony-forming capacity of THP-1 cells, this did not appear to be mediated by changes in expression of the key putative target genes of miR-196b.

3.2.4 A fluorescent reporter system for measuring miR-196b activity

There are several possible explanations for the unexpected observation that miR-196b knockout did not affect expression of its top predicted targets. Either miR-196b is not active in wild type THP-1 cells or, if miR-196b is active, it does not normally target *HOXA9*, *MEIS1* and *FAS*. There are many ways in which miR-196b could affect proliferation independently of *HOXA9*. I set out to test the activity of endogenous miR-196b in wild type THP-1 using a fluorescent reporter system.

The miRNA reporter construct shown in Figure 3.3C contains a bidirectional CMV promoter driving simultaneous expression of EGFP and the nerve growth factor receptor (NGFR)²⁹⁴. There are two miRNA response elements (MREs) in the 3' untranslated region of EGFP, such that an active miRNA targeting those MREs will repress EGFP expression. NGFR, a cell-surface receptor that is not endogenously expressed in THP-1 cells, acts as an internal control. I transiently transfected reporter constructs containing MREs perfectly complementary to the seed sequences of miR-196b, miR-17 (positive control) or *C. elegans* miR-67 (negative control) and analysed them by flow cytometry (Figure 3.3D-E). Relative to the negative control, silencing mediated by miR-196b was detectable (~15% reduction in EGFP median

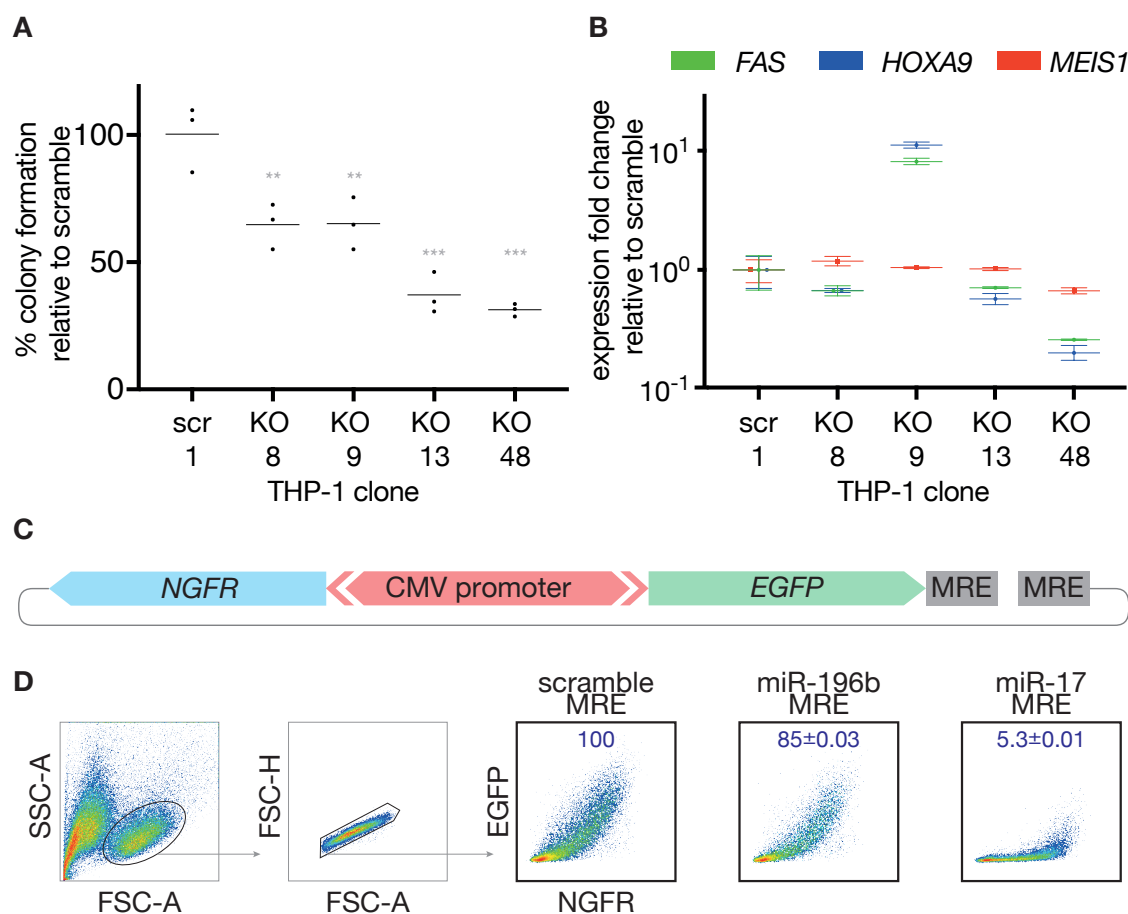


Figure 3.3: Proliferation and gene expression in miR-196b KO THP-1 cells. (A) Colony forming assay to compare colony-forming capacity of miR-196b KO clones to scramble control. n=3 biological replicates. Bar shows mean. KO line means were compared to scramble control mean using ordinary one-way ANOVA with Dunnett's multiple comparisons test. ** p<0.01, *** p<0.001. (B) RT-qPCR for *HOXA9*, *MEIS1* and *FAS*, the top predicted target genes of miR-196b. Bars show mean ± SEM. n=3 biological replicates. (C) Key components of miRNA sensor plasmid. NGFR: nerve growth factor receptor. MRE: miRNA response element. By cloning two MREs with sequence perfectly complementary to the seed sequence of the endogenous miRNA of interest, the expression of EGFP relative to NGFR control indicates the level of activity of the miRNA of interest. (D) I transiently transfected THP-1 cells with reporter constructs containing MREs for *C. elegans* miR-67 (scramble control), miR-196b or miR-17. Left plots: gating strategy. Right plots: the normalised median fluorescent intensity of EGFP relative to scramble control (inset statistic, ± SD) indicates the level of activity of the target miRNA. n=3 biological replicates; representative plots shown.

fluorescent intensity) but considerably weaker than the effect observed with miR-17 (~95% reduction). This suggested that miR-196b is active to some degree in THP-1 cells even though it did not appear to target *HOXA9*, *MEIS1* or *FAS*.

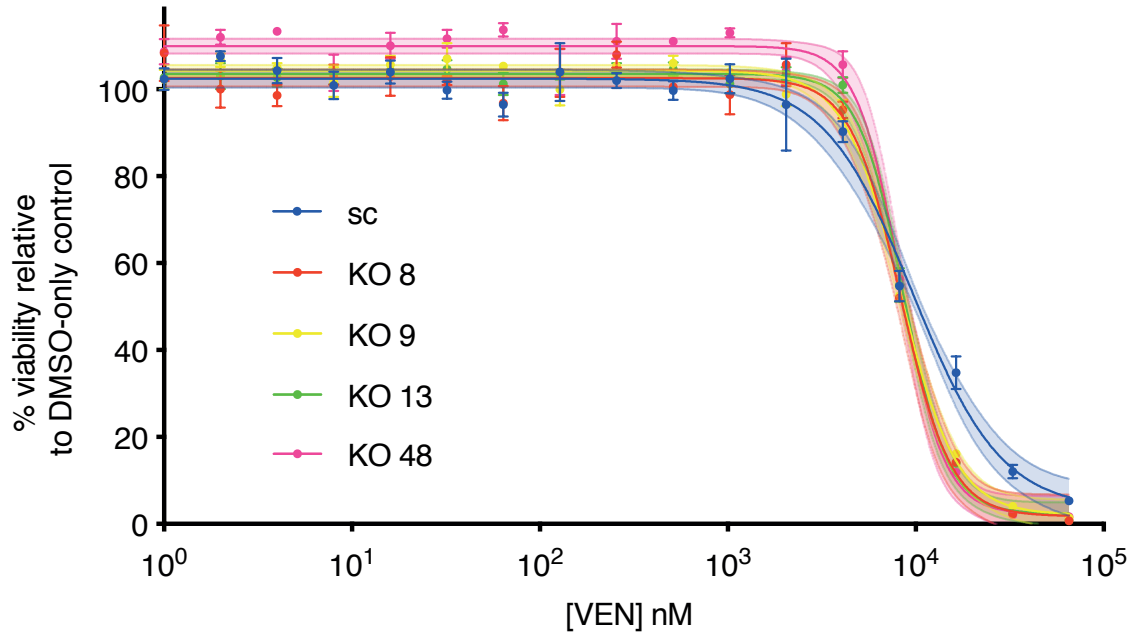


Figure 3.4: VEN sensitivity of miR-196b KO THP-1 cells. miR-196b KO cell lines and scramble control were treated with different concentrations of VEN. After 48h, I measured cell viability using CellTiter-Glo and normalised against DMSO-only (vehicle-only) control. $n=3$ biological replicates. Dots show mean; error bars show SD. Lines show unconstrained four-parameter sigmoidal least-squares regression best fit; shaded region shows 99% confidence interval. For all curves, $R^2 > 0.99$.

3.2.5 The link between miR-196b and intrinsic apoptosis

Recent evidence suggests that miR-196b may contribute to maintenance of the HSC-like state by inhibiting apoptosis¹²⁴. FAS is a TNF-family cell-surface receptor which triggers the extrinsic apoptosis pathway and it has a key role in regulating the growth and activity of immune cells²⁹⁵. *FAS* expression is a valuable diagnostic and prognostic biomarker in AML and ALL^{296–298}. I hypothesised, therefore, that miR-196b knockout would increase the sensitivity of THP-1 cells to the BCL-2 inhibitor drug venetoclax (VEN). I exposed the KO clones to varying concentrations of VEN for 48 hours and measured their viability relative to vehicle-only (DMSO-only) control using the CellTiter-Glo luminescent viability assay, which detects ATP as an indicator of metabolic activity (Figure 3.4). The miR-196b KO clones did not appear to be more sensitive to VEN than control.

3.3 Discussion

In summary, in this chapter I show that:

1. The *MIR196B* locus can be cleaved using lentiviral CRISPR/Cas9 in THP-1 cells.
2. miR-196b ^{-/-} THP-1 cells have very low expression of mature miR-196b.
3. miR-196b ^{-/-} THP-1 cells proliferate less than wild type cells.
4. miR-196b knockout does not strongly affect *HOXA9*, *MEIS1* or *FAS* expression.
5. miR-196b is active, albeit weakly, in wild type THP-1 cells.
6. miR-196b knockout does not affect sensitivity to venetoclax.

3.3.1 miRNA sequences can be deleted using dual-sgRNA CRISPR

To my knowledge, this chapter describes the first stable THP-1 cell lines containing a deletion of the entire sequence of miR-196b-5p. I validated this deletion in four clonal cell lines using Sanger sequencing, which confirmed that the mature sequence of miR-196b-5p (but not of neighbouring miR-196b-3p) had been removed. TaqMan assays further confirmed knockout of mature miR-196b. In summary, it is highly likely that miR-196b was successfully deleted in the KO clones. Other researchers have since described similar dual-sgRNA CRISPR deletion of other miRNAs^{299,300}.

There is a risk that dual-sgRNA deletion could affect the expression of neighbouring genes. In particular, disrupting the secondary structure of the pri-miR-196b transcript probably prevents proper biogenesis of miR-196b-3p. A miR-196b-3p probe would have been an informative control for the TaqMan assay. There is a low risk that miR-196b deletion affected the expression of *HOXA9* or *HOXA10* because *MIR196B* is in an intergenic region, far from either gene; besides, *MIR196B* is on the opposite strand to the neighbouring *HOX* genes. However, the chromatin landscape

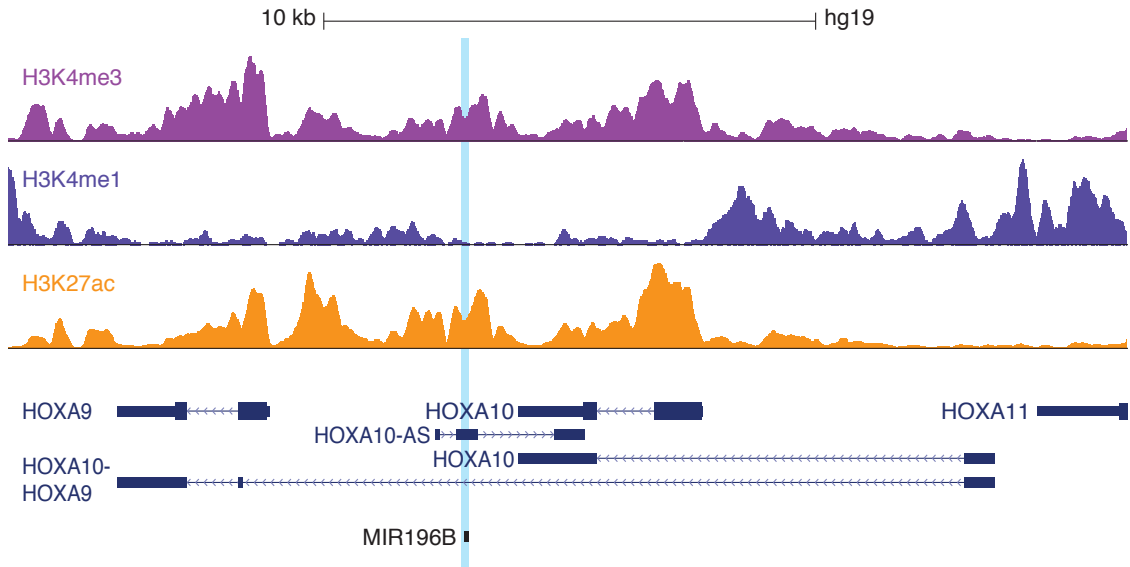


Figure 3.5: Chromatin landscape surrounding *MIR196B*. ChIP-seq data from SEM cells (MLL-AF4) provided by I-Jun Lau and Joe Harman. The intergenic region between *HOXA9* and *HOXA10*, which includes the *MIR196B* locus, contains local peaks for H3K4me3 and H3K27ac. These may have been disrupted by dual-sgRNA CRISPR editing, although a small deletion is unlikely to have a significant effect.

in this region does reveal moderate density of the activating marks H3K4me3³⁰¹ and H3K27ac (data provided by Milne group, Figure 3.5). The predicted promoter for *MIR196B* is between $\sim 780\text{bp}$ ¹²⁸ and $\sim 400\text{bp}$ (predicted using miRStart³⁰²) upstream of the transcription start site. The deleted region in clone 48 does extend 4bp beyond the *MIR196B* locus but this is very unlikely to have affected *HOXA10* expression.

Given that the *MIR196B* promoter has been identified, an alternative strategy would have been to delete this promoter, although the pri-miR-196b transcript described by Fantini et al. may not be the only transcript that yields miR-196b. In any case, such a strategy would be difficult to extend to other miRNAs because many are intronic or their promoters are difficult to identify¹²⁸.

Using dual-sgRNA CRISPR to knock out a miRNA has other limitations. In particular, off-target cleavage could lead to unexpected gene inactivation, confounding my analysis of the miR-196b KO phenotype. Alternative approaches exist such as miRNA sponges or antagomirs although these have their own disadvantages, such as lack of specificity and incomplete or transient knockdown^{303,304}. Either of

these techniques could be used to validate the miR-196b knockout phenotype in THP-1 cells.

Another issue is ploidy: this technique may delete some but not all copies of a miRNA locus. THP-1 cells are predicted to be diploid for chromosome 17 but genomic instability is common in MLLr cell lines. There was no evidence of unedited copies of *MIR196B* but NGS would be necessary to confirm this. Three of the clones were trans-heterozygous deletions while one was a homozygous deletion; the TaqMan assay showed that miR-196b expression decreased in all clones.

Unfortunately it was not possible to establish a fluorescent reporter system to measure miR-196b activity in the GFP⁺ KO clones. This would be a valuable tool for future experiments to confirm that the miR-196b KO lines do not have any residual active miR-196b.

3.3.2 miR-196b promotes clonogenicity of THP-1 cells

Popovic and coworkers concluded that miR-196b was ‘necessary’ for MLLr-driven immortalisation of HSCs⁸⁴. They found that overexpression of miR-196b in MLL-AF9 transformed mouse bone marrow progenitor cells led to increased clonogenicity, while inhibition led to reduced survival. I validated this result using a colony-forming assay in miR-196b KO cells, observing that loss of miR-196b reduced clonogenicity. It should be noted that the cells did form colonies (albeit more slowly), so in THP-1 cells miR-196b may be important but not strictly necessary for clonogenesis. Future experiments could investigate this further by measuring the serial replating efficiency of THP-1 cells, their engraftment in a mouse model or their readiness to differentiate in vitro.

3.3.3 *HOXA9*, *MEIS1* and *FAS* are not regulated by miR-196b in THP-1 cells

The results of Popovic and coworkers, as independently validated by my colony-forming assay, are difficult to understand given that *HOXA9* contains a consensus miR-196b response element. If *HOXA9* promotes proliferation and miR-196b

represses *HOXA9*, then a simple model would predict that ablation of miR-196b should derepress *HOXA9* and increase proliferation. It is, however, possible that miR-196b only tunes the expression of *HOXA9* and *MEIS1* during normal hematopoiesis. It is well known that both *HOXA9* and miR-196b are upregulated by MLL fusion proteins and their expression is strongly correlated in early AML/ALL patient samples. Chen and coworkers proposed a model where *HOXA9* and *MEIS1* are too strongly activated by MLL fusions to be silenced by miR-196b; instead, elevated miR-196b silences other lineage-specific targets such as *FAS*, a tumour-suppressor gene¹²¹.

Using RT-qPCR, I measured expression of *HOXA9*, *MEIS1* and *FAS* in the miR-196b KO cells. *MEIS1* expression was unchanged in all four cell lines. Surprisingly, *HOXA9* and *FAS* expression increased in clone 9 and decreased in clone 48. It would be interesting to compare gene expression more broadly in these two cell lines using RNAseq. Screening a larger set of miR-196b KO clones would also indicate the normal clonal variation in the expression of these three genes. Taken together, these results are at odds with the Chen model, in particular because miR-196b knockout does not appear to affect *FAS*¹²¹. The fluorescent reporter assay suggested that miR-196b is indeed active in THP-1 cells, although at a considerably lower level than miR-17. It is not possible to determine whether the lower relative activity of miR-196b compared to miR-17 can be fully explained by the difference in expression according to the TaqMan assay. It should be noted that miRNA-mediated silencing of a reporter may not perfectly reflect silencing of endogenous targets.

Before we can properly understand the role of miR-196b in MLLr leukemia, we must elucidate its target network. Recent results corroborate my observation that miR-196b does not strongly interact with *HOXA9*, *MEIS1* or *FAS* in THP-1 cells, but does regulate other targets. Meyer et al conducted an unbiased miR-196b target screen in THP-1 cells by transiently transfecting biotinylated miRNA mimics¹²⁴. They pulled down miRNA-RISC-target complexes and identified them using microarray analysis and RT-qPCR. The most promising target is *CDKN1B*, which encodes the cell cycle protein p27^{kip1}. It is intriguing that deletion of p27^{kip1} in

normal hematopoietic progenitor cells reduces proliferation³⁰⁵, whereas inhibition of p27^{kip1} in leukemia cells increases proliferation^{124,306}. This suggests that miR-196b contributes to the maintenance of leukemia rather than leukemogenesis.

Pooled CRISPR screening could offer a complementary approach to the miRNA-RISC pulldown screen. Indeed, originally I set out to conduct a pooled CRISPR screen for miR-196b target genes. Previous work in the Fulga group had shown that miRNA recognition elements (MREs) in the 3' UTR of target genes could be disrupted using single-sgRNA editing with SpCas9^{307,308}. Building on this work, I proposed a screening pipeline called MRE-CRISPR. I planned to predict the top 20 000 miR-196b MREs and construct a custom pooled sgRNA library targeting these sequences. A CRISPR-KO screen with selection based on proliferation or clonogenicity would reveal which genes were essential targets of miR-196b in MLLr leukemia. For each MRE, several control sgRNAs targeting other sites within the same 3' UTR would detect the baseline contribution of that UTR to the phenotype of interest. Each MRE-targeting sgRNA would have to be normalised against its control sgRNAs. Following the MRE-CRISPR screen, I could further interrogate the 3' UTRs of genes of interest using the GenERA pipeline³⁰⁸. In a GenERA screen, a pooled library of sgRNAs tiles the entire 3' UTR and the effect of each sgRNA on transcript abundance is measured by NGS, creating a detailed profile of all MREs in a given UTR.

It is unlikely that deletion of a miR-196b response element would have a greater effect on proliferation than deletion of miR-196b itself. The change in proliferation induced by deletion of miR-196b represents, therefore, the greatest effect size that I could reasonably expect in a pooled CRISPR-KO screen targeting miR-196b response elements. miR-196b deletion did reduce proliferation of THP-1 cells, but perhaps not to a sufficient extent for robust growth-based selection in a CRISPR screen. Perhaps the target networks of other miRNAs would have a greater impact on proliferation. Given this limitation, I chose not to conduct a CRISPR screen for miRNA target genes. Instead, I designed a genome-wide CRISPR screen targeting protein-coding sequence – as described in chapter 4 and chapter 5.

3.3.4 miR-196b knockout does not affect venetoclax resistance in THP-1 cells

Lastly, I speculated that miR-196b may modulate targets involved in apoptotic signalling, even if *FAS* itself was unaffected by miR-196b knockout. THP-1 cells are known to be resistant to the leukemia drug venetoclax, which induces apoptosis by inhibiting the antiapoptotic factor BCL-2. I used a standard cell viability assay to measure venetoclax resistance in the miR-196b knockout cells but found no effect.

3.3.5 The role of ncRNAs in MLLr leukemia remains uncertain

Targeting miRNAs has become an important and valuable strategy in cancer therapy but, as the results described in this chapter demonstrate, much remains unclear³⁰⁹. A possible explanation for the discrepancy between published results and my data is that murine models are an inaccurate reflection of human AML/ALL, or that ectopic overexpression of miR-196b produces non-physiological results. Stable knockout of miR-196b in other human cell lines would shed light on this issue.

Several other miRNAs also warrant further attention in MLLr leukemia. For example, Bhayadia et al showed that expression of miR-193b is significantly decreased in pediatric MLLr AML patients^{310,311}. They transformed miR-193b^{+/+} and miR-193b^{-/-} mouse bone marrow cells with *Hoxa9* and *Meis1*. On transplantation, miR-193b^{-/-} cells exhibited increased engraftment and expansion and were associated with shorter survival. Furthermore, they showed that ectopic expression of miR-193b reduces proliferation of human AML cell lines and primary cells. They validated miR-193b-directed repression of the MAPK pathway but found that relieving this repression was insufficient to negate the effects of miR-193b. miR-193b must, therefore, have other, unpredicted targets. The true significance of miR-193b in initiation or maintenance of MLLr leukemia is still unknown.

Since proposing their model for miR-196b, Chen and coworkers have also implicated miR-22³¹², miR-150³¹³, miR-126³¹⁴ and miR-9³¹⁵ in MLLr leukemia. In addition, it appears that the homeobox gene PBX3 is an essential cofactor of

HOXA9 and MEIS1 in MLLr leukemogenesis^{316,317}. Each of these miRNAs (or several) could be the basis for a pooled MRE-targeting CRISPR screen. Establishing better human models of MLLr leukemia and building precise coding and non-coding gene regulatory networks will be crucial to dissecting the role of homeobox genes.

Bioinformatic prediction of miRNA targets is unreliable. Even with novel prediction approaches³¹⁸ and highly integrative databases³¹⁹, most miRNA-target interactions have an inconsequential effect on gene expression³²⁰. In addition, many miRNA-target interactions appear to be noncanonical (that is, they do not involve the seed sequence) yet the seed sequence is the focus of most target prediction algorithms³²¹. Future studies must emphasise functional *in situ* validation of predicted miRNA-target interactions, ideally *in vivo*.

Many questions remain unanswered about the role of miR-196b in MLL-AF9 cells such as THP-1. I continued my efforts to understand treatment resistance in MLLr leukemia, again using THP-1 cells as a model but focusing instead on protein-coding genes. This is the subject of the following two chapters.

4

Intrinsic apoptosis and venetoclax resistance

Contents

4.1	Introduction	85
4.2	Results	87
4.3	Discussion	101

4.1 Introduction

In this chapter, I continue to investigate the causes of treatment resistance in MLLr leukemia, focusing on the ability of malignant cells to evade apoptosis. I use a recently developed approach – lentiviral CRISPR knockout (CRISPR-KO) – to offer an explanation for clinical observations of treatment resistance in MLLr leukemia. In parallel, I optimise methods for a whole-genome CRISPR-KO screen in THP-1 cells.

The overexpression of antiapoptotic factors such as BCL-2 is a hallmark of many leukemias, making the intrinsic apoptosis pathway a major focus for targeted cancer therapies (Figure 4.1). Venetoclax (VEN), a BH3-mimetic BCL-2 inhibitor, is clinically approved for treatment of chronic lymphocytic leukemia (CLL)^{176,190}. CLL depends heavily on BCL-2 overexpression, unlike AML and ALL which depend on a

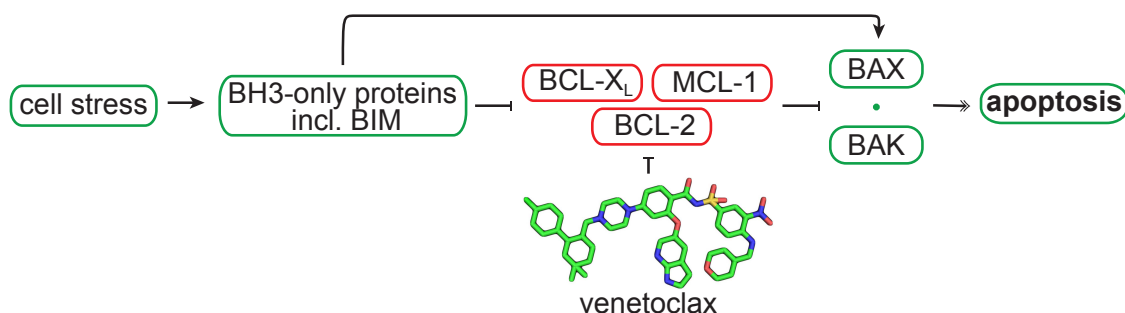


Figure 4.1: The intrinsic apoptosis pathway. Simplified representation of key regulators involved in intrinsic apoptosis. Green: proapoptotic factors. Red: antiapoptotic factors. Cell stress activates BH3-only pro-apoptotic factors such as BIM (*BCL2L11*), BID and BAD. BH3-only proteins inhibit the anti-apoptotic factors BCL- X_L (*BCL2L1*), BCL-2 and MCL-1, which in turn inhibit BAX and BAK (*BAK1*), the immediate effectors of mitochondrial apoptosis. BH3-only proteins also directly activate BAX and BAK. Active BAX-BAK heterodimers lead to depolarisation of the mitochondrial membrane and release of cytochrome *c* into the cytosol, which activates caspases and triggers apoptosis. Venetoclax (inset: chemical structure) is a BH3-mimetic which competitively inhibits BCL-2.

combination of several antiapoptotic factors including BCL-2, BCL- X_L and MCL-1. Correspondingly, early clinical data indicate that AML/ALL patients are refractory to VEN monotherapy but may respond to combination therapies¹⁹⁰. MLL-AF4 ALL appears to be relatively sensitive to VEN, although the molecular basis for this sensitivity remains unclear^{200,201}. I reasoned that a better understanding of the mechanism for VEN resistance in an MLLr leukemia model would enable the development of a broader range of combination therapies.

In this chapter, I describe results which identify THP-1 cells as a model for clinical VEN resistance. Next I optimise viral transduction and puromycin selection for screening experiments. Then I use dual-sgRNA lentiviral CRISPR to knock out key components of the intrinsic apoptosis pathway, also known as the *BCL-2* gene family. I demonstrate that loss of anti-apoptotic and pro-apoptotic proteins increases sensitivity and resistance to VEN and that this can be measured using cell proliferation as a readout. I further confirm this result using a small-scale pooled lentiviral ‘miniscreen’. The miniscreen serves as a pilot experiment for the whole-genome screen described in chapter 5, in which I investigate whether other genes beyond the *BCL-2* family can affect VEN resistance.

4.2 Results

4.2.1 Venetoclax resistance in leukemia cell lines

I selected THP-1 cells as an effective model for VEN resistance. THP-1 cells are a well characterised, long-standing model for clinical MLL-AF9 leukemia¹³³ (subsection 1.3.7). I confirmed the VEN resistance of THP-1 cells by screening them alongside two myeloid leukemia cell lines, HL-60 (non-MLLr AML) and U-937 (*PICALM-AF10* histiocytic lymphoma). After treating with VEN for 48 hours, I stained cells with Annexin-V/PI and used flow cytometry to measure cell viability (Figure 4.2). In line with published data, HL-60 cells were considerably more sensitive to VEN than either THP-1 or U-937 cells¹⁸⁸. THP-1 cells were the most resistant to VEN ($IC_{50}=7.4\mu M$).

4.2.2 Optimising transduction and puromycin selection in THP-1 cells

In order to reliably dissect the genes involved in VEN resistance using lentiviral CRISPR-KO, I set out to optimise transduction and puromycin selection in THP-1 cells. The integrating portion of lentiCRISPRv2 (Figure 4.3B) contains a puromycin resistance marker so that untransduced cells can be removed by puromycin selection. The selection step serves a dual purpose: (1) to remove untransduced cells and (2) to determine multiplicity of infection and ensure that cells are transduced with only one sgRNA per cell (see below).

Firstly, I determined the optimum puromycin concentration, defined as the lowest concentration sufficient to kill untransduced cells but not transduced cells. In pooled screens, untransduced cells do not directly affect the sgRNA distribution but they should still be eliminated because untransduced cells reduce screening coverage and may alter the growth dynamics of transduced cells. For arrayed KO experiments where edited cells are analysed in bulk, it is essential to remove untransduced cells because they could confound measurement of the KO phenotype.

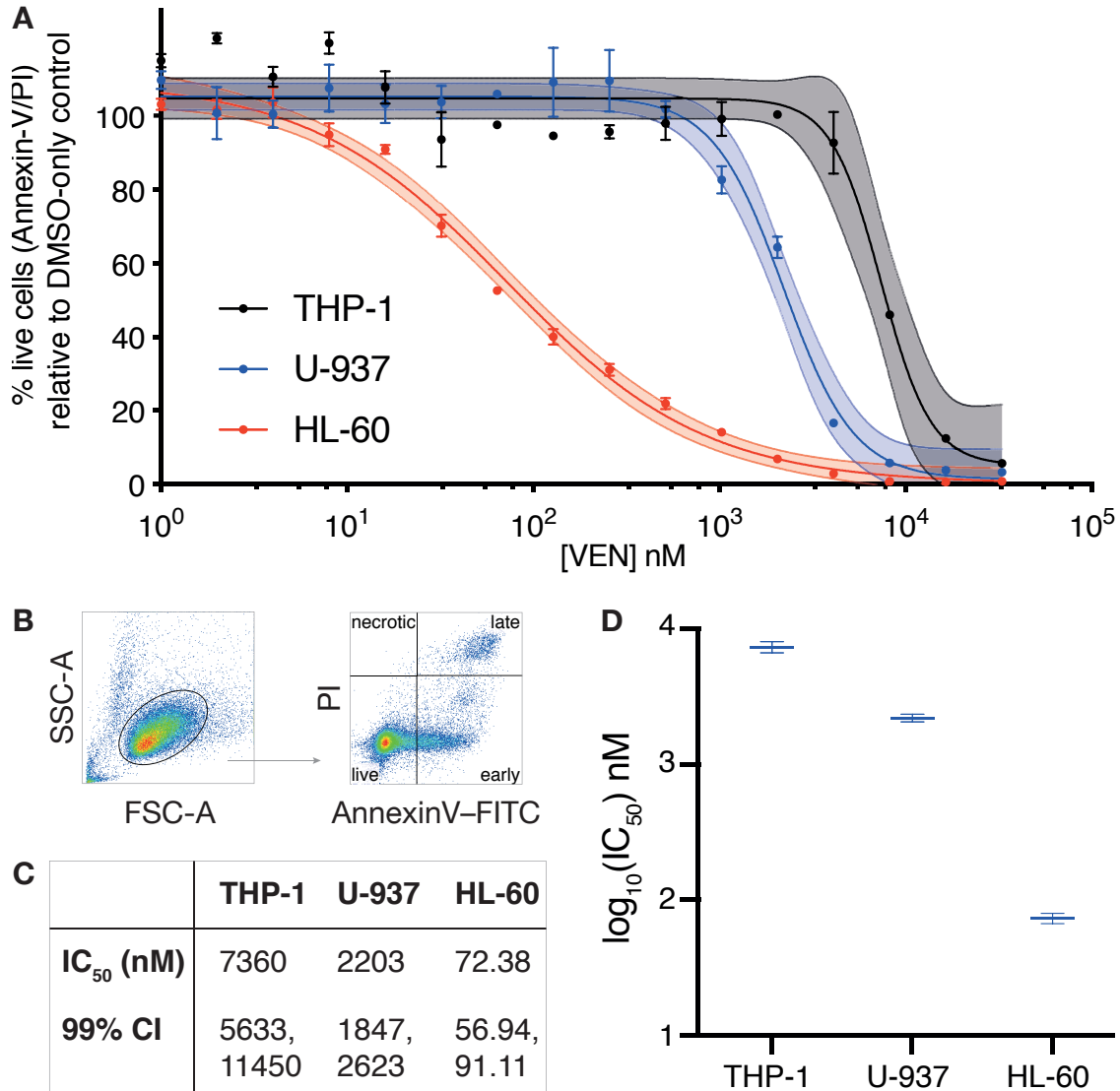


Figure 4.2: Comparing venetoclax resistance in leukemia cell lines. (A) THP-1 (*MLL;AF9*), U-937 (*CALM;AF10*) and HL-60 cells were cultured for 48 hours at different concentrations of VEN. Their viability was determined using Annexin-V/PI staining and flow cytometry. Dots show mean; error bars show SD. Lines show unconstrained four-parameter sigmoidal least-squares regression best fit; shaded region shows 99% confidence interval. For all curves, $R^2 > 0.99$. $n=5$ biological replicates. (B) Flow cytometry gating for (A). PI: propidium iodide; early/late: early/late apoptotic. (C-D) Relative IC₅₀ values, defined as the half-maximal inhibitory concentration of venetoclax, determined from best-fit curves in (A). Bars show mean $\pm$ 99% fit confidence interval (CI).

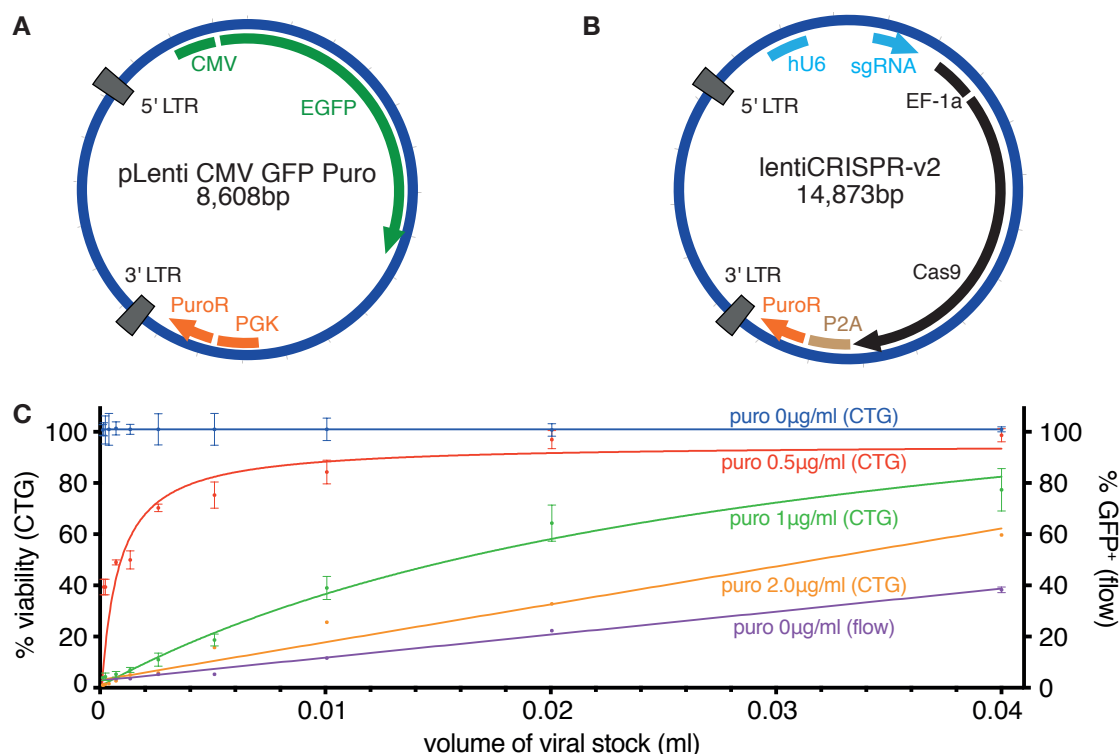


Figure 4.3: Puromycin selection and lentiviral titre. (A) pLenti CMV GFP Puro is a third-generation lentiviral vector expressing both *PuroR* and GFP. (B) lentiCRISPRv2 is a third-generation lentiviral vector which I used for KO experiments. (C) Cells were transduced with a serial dilution of pLenti CMV GFP Puro and selected with puromycin (puro). After 72 hours, viability was determined using the CellTiter-Glo assay (CTG, left axis) or %GFP⁺ was determined by flow cytometry (flow, right axis). n=5 biological replicates. Points show mean $\pm$ SD. Curves show unconstrained robust regression hyperbolae ($R^2 > 0.95$).

At the optimum puromycin concentration, the proportion of cells surviving selection should be the same as the proportion of cells transduced. To measure transduction independently of survival, I used pLenti CMV GFP Puro, a plasmid expressing both GFP and *PuroR* (Figure 4.3). After transduction, I selected cells at different puromycin concentrations and measured their viability, or cultured them without puromycin and used flow cytometry to measure the proportion of GFP⁺ cells. 2 μ g/mL puromycin was most effective at removing untransduced cells. I used this concentration for all subsequent experiments. I also established that selection was most effective at a cell density of 2.5×10^5 cells/mL for 72 hours.

The multiplicity of infection (MOI) is a mathematical quantity equal to the ratio of viral transducing units (TU) to the number of cells. The titre of a preparation

of lentivirus (in TU/mL) depends on the concentration of viral particles and their activity. The greater the MOI, the higher the likelihood that a cell will be transduced by multiple lentiviruses²⁷⁰. Since it is very unlikely that two lentiviruses in the pool will contain the same sgRNA sequence, a doubly transduced cell would harbour two separate gene knockouts, confounding the screen.

It is important, therefore, to ensure that cells are transduced at low MOI in order to minimise the number of multiply transfected cells. The probability P of a cell being infected n times at a given MOI can be modelled using a Poisson distribution such that:

$$P_{n>0} = 1 - P_{n=0} = 1 - e^{-\text{MOI}}.$$

For a given cell population, $P_{n>0}$ is equivalent to the fraction of transduced cells or, after ideal puromycin selection, the fraction of cells that survive. More precisely:

$$P_n = \text{MOI}^n \cdot \frac{e^{-\text{MOI}}}{\Gamma(n+1)}$$

where $\Gamma(z)$ is the gamma function (Figure 4.4A). I transduced cells at MOI=0.3, at which MOI the model predicts that ~74% of cells will be untransduced, ~22% will be transduced with one virus and only ~3% will be transduced with two or more viruses.

In order to ensure transduction at MOI=0.3, I measured the lentiviral titre by transducing THP-1 cells with a serial dilution of lentiCRISPRv2 lentivirus. Where c is the total number of cells at transduction and V is the volume of viral stock in mL, the titre can be determined as follows:

$$\text{MOI} = -\ln(1 - P_{n>0})$$

$$\text{titre} = c \cdot \frac{\text{MOI}}{V}$$

$$\Rightarrow \text{titre} = \frac{-\ln(1 - P_{n>0}) \cdot c}{V}$$

I used a linear approximation for $P_{n<0}$ at low lentivirus concentrations to estimate the titre (Figure 4.4B). To validate this approach, I replotted the survival data based on the estimated MOI – calculated from estimated titre – and found that the empirical data closely matched theoretical predictions, confirming that the estimated titre was reliable (Figure 4.4C). I used this method to determine the titre of all lentivirus preps for screening.

4.2.3 CRISPR knockout of BCL-2 family genes in THP-1 cells

The BCL-2 family proteins regulate intrinsic apoptosis. I individually knocked out six of these genes (*BCL2*, *MCL1*, *BCL-X_L*, *BAX*, *BAK* and *BIM*) in THP-1 cells to understand their relative contributions to VEN resistance. For each gene, I selected from the TKOv3 pooled screening library the two sgRNAs with the highest sgRNA sequence scores. (See chapter 5 for a detailed description of the TKOv3 library.) I cloned them into lentiCRISPRv2, a third-generation lentiviral vector which expresses an sgRNA, SpCas9 and a puromycin resistance marker (Figure 4.3B). The two plasmids per gene were pooled and used to generate lentivirus. Combining two sgRNAs per gene increased the likelihood of successful knockout. An sgRNA targeting the safe-harbour locus *AAVS1* served as a silent targeting control; the scrambled sgRNA described in chapter 3 served as a non-targeting control. I transduced cells at MOI=0.3 and selected them with puromycin, then validated sgRNA activity using the T7E1 assay (Figure 4.5A-B). The *MCL-1* target locus

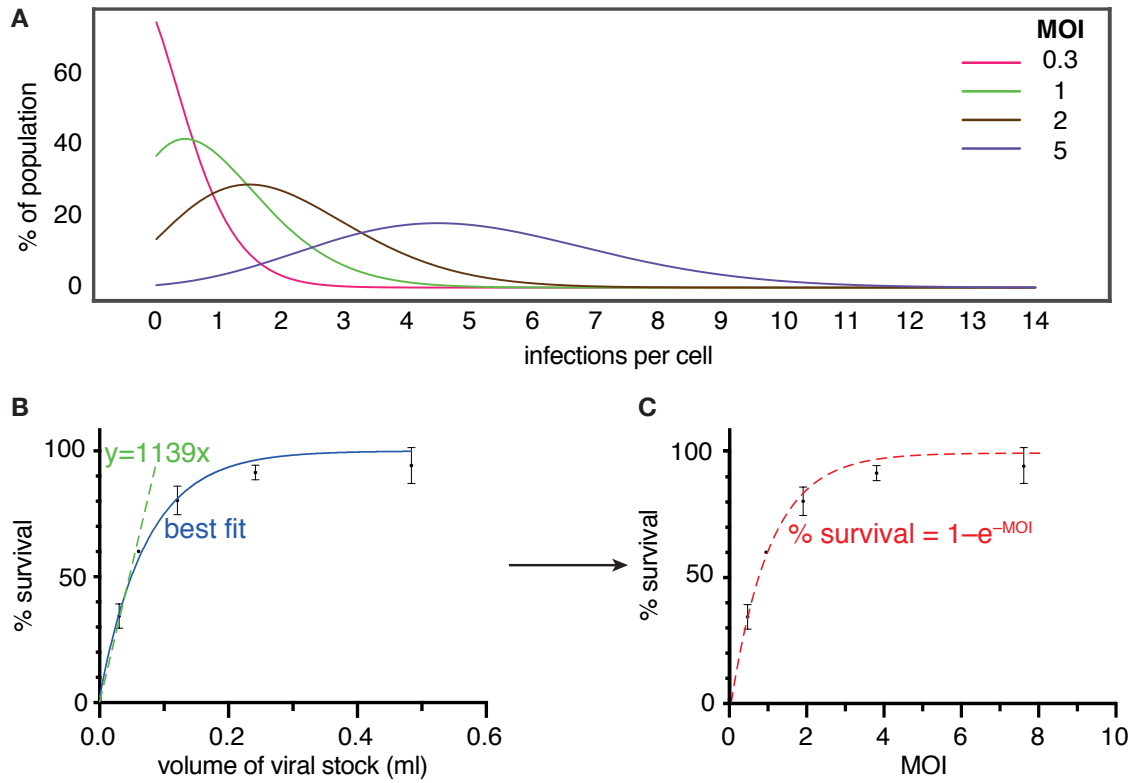


Figure 4.4: Observed and predicted lentiviral titre. (A) Probability density distributions to show probability of n infections per cell at different multiplicity of infection (MOI), based on a Poisson distribution: $P_n = \text{MOI}^n \cdot \frac{e^{-\text{MOI}}}{\Gamma(n+1)}$ where $\Gamma(z)$ is the gamma function. At MOI=0.3, almost all cells are transduced by ≤ 1 lentivirus. (B) THP-1 cells were transduced with a serial dilution of lentiCRISPRv2 lentivirus and selected with 2 $\mu\text{g/mL}$ puromycin. $n=2$ biological replicates. Points show mean $\pm$ SD. A linear approximation closely fits the data at low concentrations of virus, allowing simple estimation of the titre. Blue curve: least squares best fit of the form $y = 100(1 - e^{-Kx})$ ($R^2 = 0.99$). Green dashed line: least-squares best fit of the form $y = mx$. (C) Using the linear approximation to estimate titre, the volume of viral stock is converted to MOI and the data are replotted. They closely fit the theoretical predicted relationship between MOI and % survival. Red dashed line is described by $1 - e^{-\text{MOI}}$.

could not be amplified. For some genes, the two target loci were too close to each other to be amplified separately. In these cases, the T7E1 assay reported the combined activity of both sgRNAs. Judging from target loci which could be individually amplified, it was apparent that sgRNAs had varying on-target efficiency. Nonetheless, western blotting confirmed partial or complete protein knockout in all cases (Figure 4.5C), including MCL-1. Double bands were detected for BAK, BIM and MCL-1. These were likely due to different protein isoforms present in the cell, reflected in similar blots in the literature¹⁹⁴.

In addition to the BCL-2 family genes shown in Figure 4.5, I also targeted two highly essential genes drawn from the PICKLES database: *POLR2A*, encoding the largest subunit of RNA Pol II; and *TUBB*, encoding β -Tubulin³²². Again, I combined two sgRNAs per gene from the TKOv3 library. Knockouts of *POLR2A* and *TUBB* were lethal: none of the THP-1 cells transduced with these sgRNAs survived puromycin selection.

Next I investigated the VEN resistance of the BCL-2 family knockout cells by plating them at a VEN concentration equal to the IC_{50} established for wild-type THP-1 cells ($[VEN]=7.4\text{ }\mu\text{M}$) for five days (~ 2 doublings). I compared their proliferation to control-treated cells using the CellTiter-Glo assay (Figure 4.6). As expected, drug treatment at IC_{50} reduced the growth of cells edited with *AAVS1*-targeting or scramble sgRNA by $\sim 50\%$. Growth decreased by only $\sim 25\%$ in BAX KO cells, corresponding to the role of BAX as an apoptotic effector, although no change was seen for the other proapoptotic proteins, BAK and BIM. By contrast, knockout of the anti-apoptotic factors MCL-1 and BCL-X_L slightly increased VEN sensitivity. Finally, I measured the cytotoxicity of VEN in each of the edited lines using the Annexin-V/PI assay (Figure 4.7). Only BAX and BAK knockout appeared to change the 48-hour VEN sensitivity of THP-1 cells. Taken together, these experiments confirmed that VEN sensitivity depends not only on BCL-2 but also on other BCL-2 family proteins.

4.2.4 Venetoclax miniscreen in THP-1 cells

Based on the arrayed lentiCRISPRv2 knockouts described in Figure 4.5, it seemed likely that cell proliferation would provide an effective readout for VEN sensitivity in a pooled screen with VEN selection. Before conducting a genome-scale screen (chapter 5), I designed a pilot experiment to explore whether such a high-throughput study would be feasible. I expected the pilot pooled screen to mirror the results of the arrayed screen. As described above, I had already optimised transduction and puromycin selection with lentiCRISPRv2 in THP-1 cells; but the pilot screen

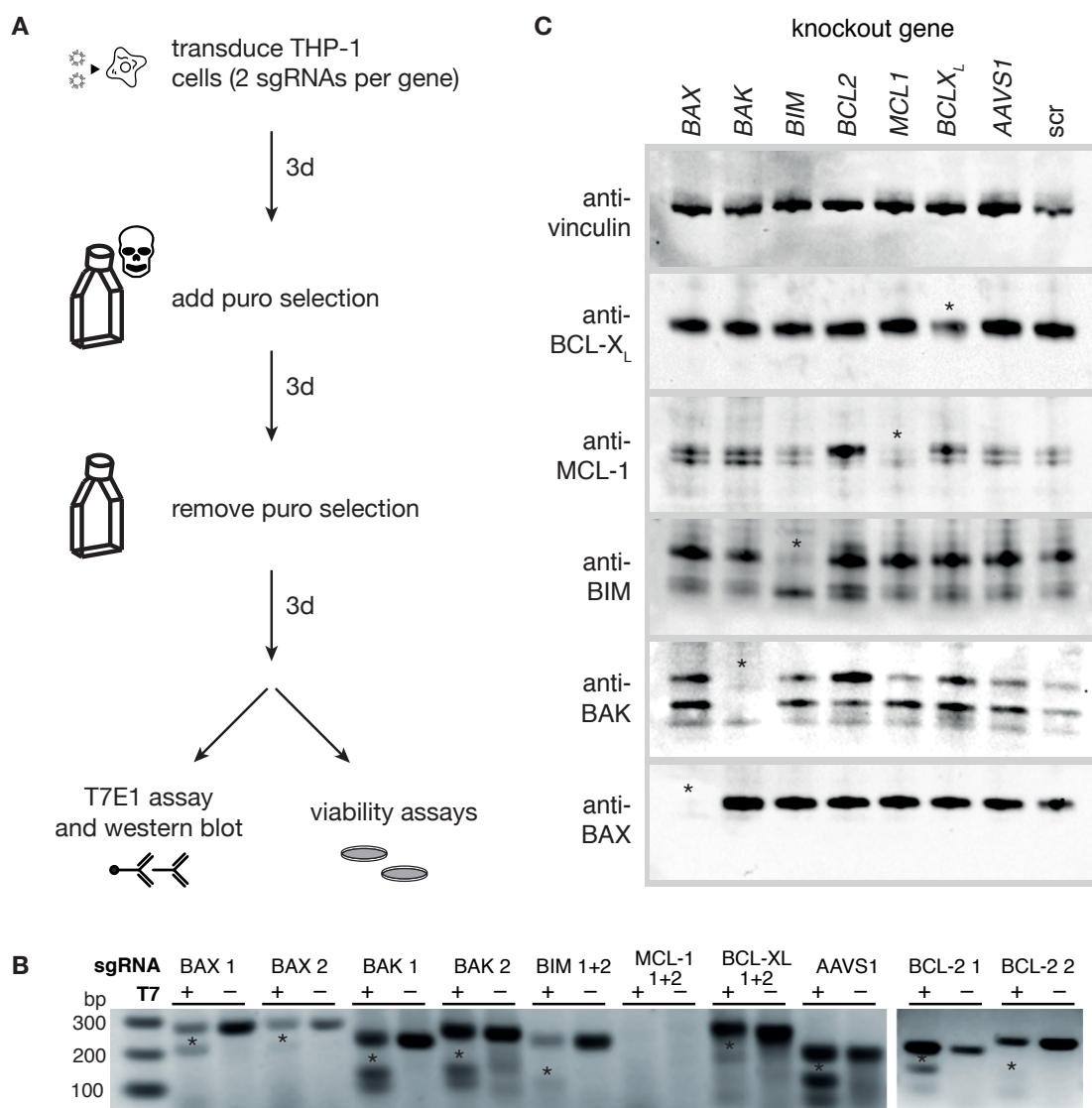


Figure 4.5: Generating dual-sgRNA knockouts of BCL-2 family genes in THP-1 cells. (A) Dual-sgRNA knockout with lentiCRISPRv2. (B) Agarose gel to show T7E1 assay for two sgRNAs per target gene (separately or combined). Cleavage products indicating edited loci are marked (*). (C) Cropped western blots to validate BCL-2 family protein knockouts. Knockout bands are marked (*). Vinculin was used as a control (representative blot shown). Uncropped blots are shown in Appendix C.

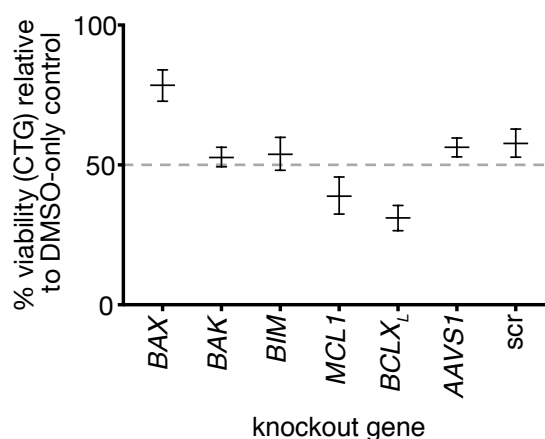


Figure 4.6: Proliferation of BCL-2 family knockout THP-1 cells in VEN. Dual-sgRNA knockout THP-1 cells were cultured in 8 μ M VEN for five days. CTG: CellTiter-Glo. Bars show mean $\pm$ SD. n=3 biological replicates.

would be an ideal opportunity to develop optimised protocols for the entire whole-genome screening workflow.

For this pilot screen (referred to here as the ‘miniscreen’), I used the same subset of TKOv3 sgRNAs described in Figure 4.5. Thus, the final pool of 16 plasmids used to generate lentivirus contained one non-targeting scramble sgRNA, one *AAVS1* sgRNA and two sgRNAs per gene for each of *MCL1*, *BCL-X_L*, *BAX*, *BAK*, *BIM*, *AAVS1*, *POLR2A* and *TUBB*. I did not include *BCL2*-targeting sgRNAs because in the absence of BCL-2, any activity of VEN would be off-target.

The miniscreen was designed as a standard pooled lentiviral CRISPR viability screen (Figure 4.8). I harvested gDNA at every passage in order to measure the dynamics of sgRNA distribution over time, and thus to determine the optimal duration for a whole-genome screen²⁴³. I counted cells at transduction and at each subsequent passage and harvest to ensure that library representation was >500 cells per sgRNA at all times. I harvested gDNA in triplicate from transduced cells on day 0 and from the DMSO-only and VEN treatment arms every three days, also in triplicate, for a total of five passages. It was important to ensure that dead cells were removed at each passage to exclude them from the harvested gDNA. Using cell viability staining to count dead cells, I found that density gradient centrifugation was no more effective than standard centrifugation at removing dead cells. I extracted

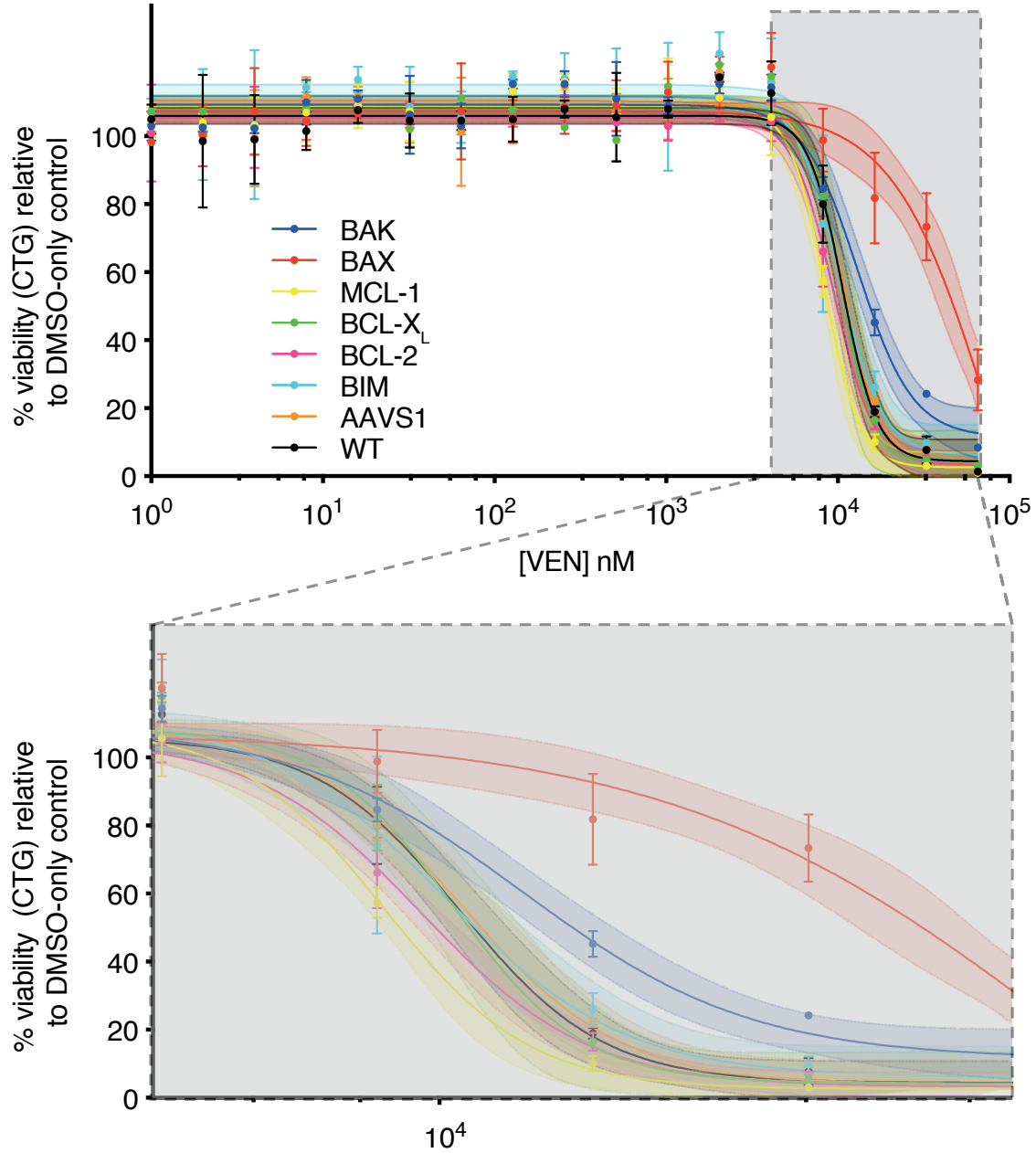


Figure 4.7: IC_{50} (VEN) of BCL-2 family knockout THP-1 cells. (A) Non-clonal knockout THP-1 cells were cultured for 48 hours at different concentrations of VEN and analysed using the CellTiter-Glo assay to compare the effect of BCL-2 family genes on VEN sensitivity. IC_{50} (VEN) in BAX KO cells was $\sim 78 \mu\text{M}$. (B) Enlarged detail from (A). Dots show mean; error bars show SD. Lines show unconstrained four-parameter sigmoidal least-squares regression best fit; shaded region shows 99% confidence interval. For all curves, $R^2 > 0.99$. $n=3$ biological replicates.

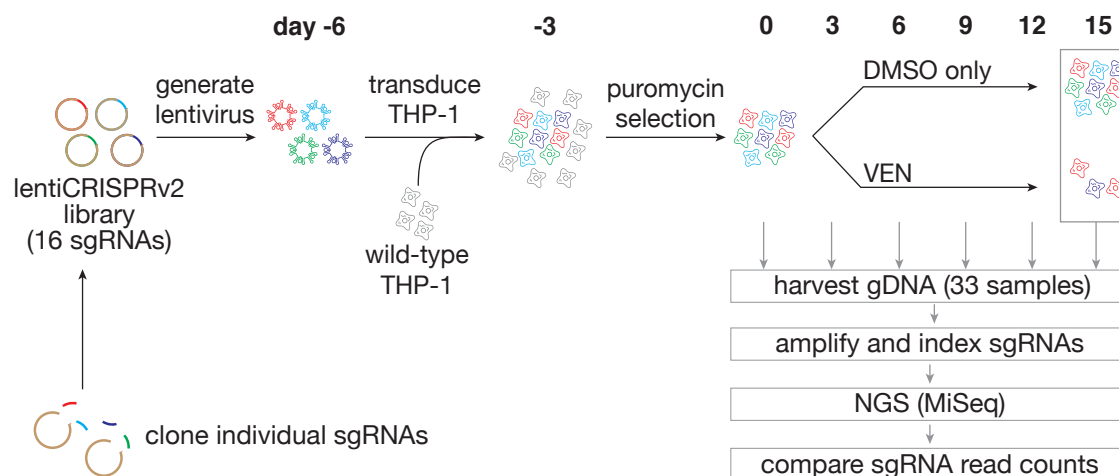


Figure 4.8: Pooled CRISPR-KO miniscreen targeting BCL-2 family. A pooled VEN miniscreen in THP-1 cells served as a pilot study for the whole-genome screen. The pool contained 16 sgRNAs: one sgRNA targeting *AAVS1*, one scramble control sgRNA, and two sgRNAs targeting each of *BAX*, *BAK*, *MCL-1*, *BIM*, *BCL-X_L*, *POLR2A* and *TUBB*. I transduced THP-1 cells with pooled lentivirus and, three days later, applied puromycin selection to kill untransduced cells. On day 0, puromycin selection was removed and gDNA was harvested in triplicate; the remaining cells were cultured for another fifteen days under constant venetoclax selection (at IC₅₀) or DMSO control. gDNA was harvested every three days. sgRNAs were amplified from gDNA, indexed, pooled and sequenced.

gDNA from the harvested cells, PCR-amplified the integrated sgRNA region and indexed and pooled each sample for NGS. The total number of aligned reads was 347 767, equivalent to a mean of 659 reads/sgRNA across all samples.

For quality control, I began by analysing the relative abundance of all 16 sgRNAs on day 0: that is, after transduction and puromycin selection but before the cells were divided into control and treatment arms (Figure 4.9). The lentiviral pool had been generated from a hand-mixed plasmid pool so I expected some variation in sgRNA abundance. Concordance between individual replicates was strong but some sgRNAs were five-fold more abundant than others. In order to overcome this skew, for further analysis I normalised each sgRNA against its abundance on day 0.

Figure 4.10 shows how the relative abundance of each sgRNA changed over time in the DMSO-only control arm. An sgRNA whose abundance decreases in the absence of selection must be targeting an essential gene (notwithstanding any off-target effects). As expected, both *TUBB* sgRNAs were lost over time, as was

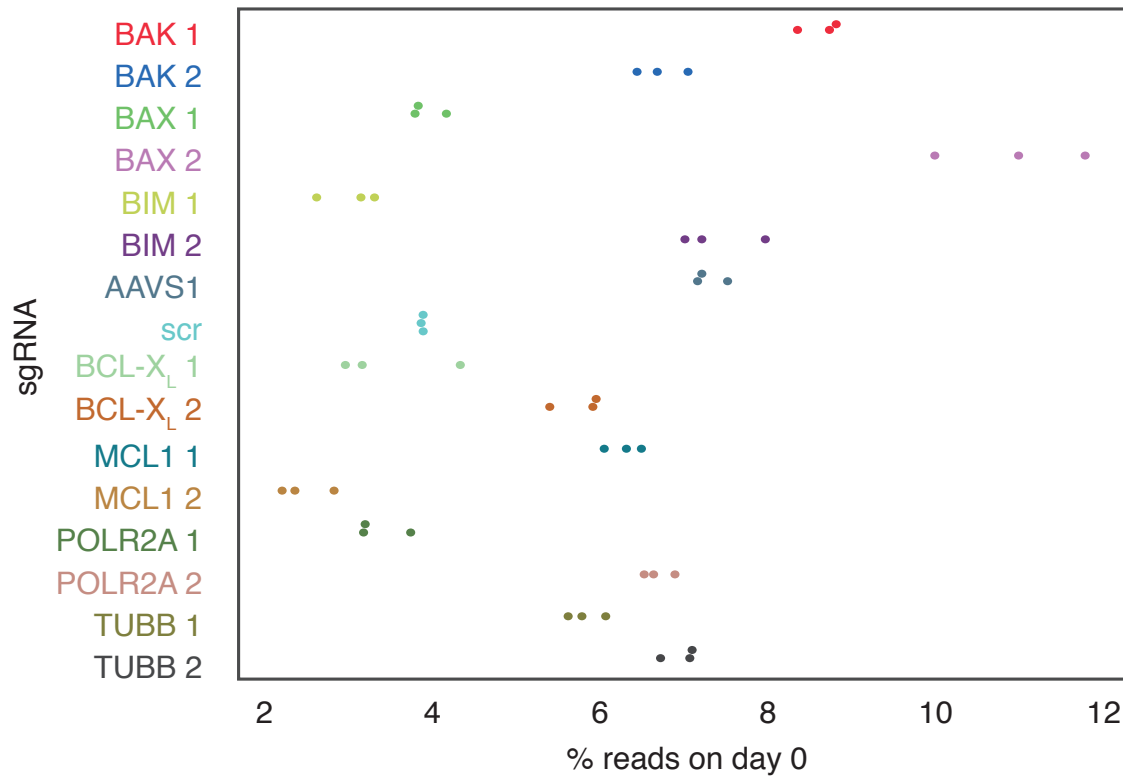


Figure 4.9: Initial distribution of miniscreen sgRNAs. Swarm plot to show relative read count (% of total sgRNA reads) on day 0 for each of 16 sgRNAs in the miniscreen. $n=3$ replicate gDNA harvests from one biological sample.

one of the POLR2A sgRNAs. It is likely that the other POLR2A sgRNA was not active. None of the other genes in the miniscreen is known to be essential, except for *MCL-1* which may be weakly essential in THP-1 cells^{278,322}.

Unlike sgRNAs targeting essential genes, those targeting non-essential genes should not be lost over time. In fact, I observed that most other sgRNAs were enriched, including the scramble control but not the AAVS1 control, with the notable exception of one of the BCL- X_L sgRNAs. This was probably due to a confounding effect of a small sgRNA pool: when some sgRNAs are negatively selected, the remaining sgRNAs appear more abundant due to drift. It is also possible that knockout of pro-apoptotic genes leads to enrichment.

Under VEN selection, the changes in relative sgRNA abundance partly corroborated earlier results from the arrayed screen (Figure 4.11). As expected, BAX and BAK-targeting sgRNAs were enriched relative to DMSO-only control, in line with the earlier observation that BAX and BAK knockout mutants were resistant to

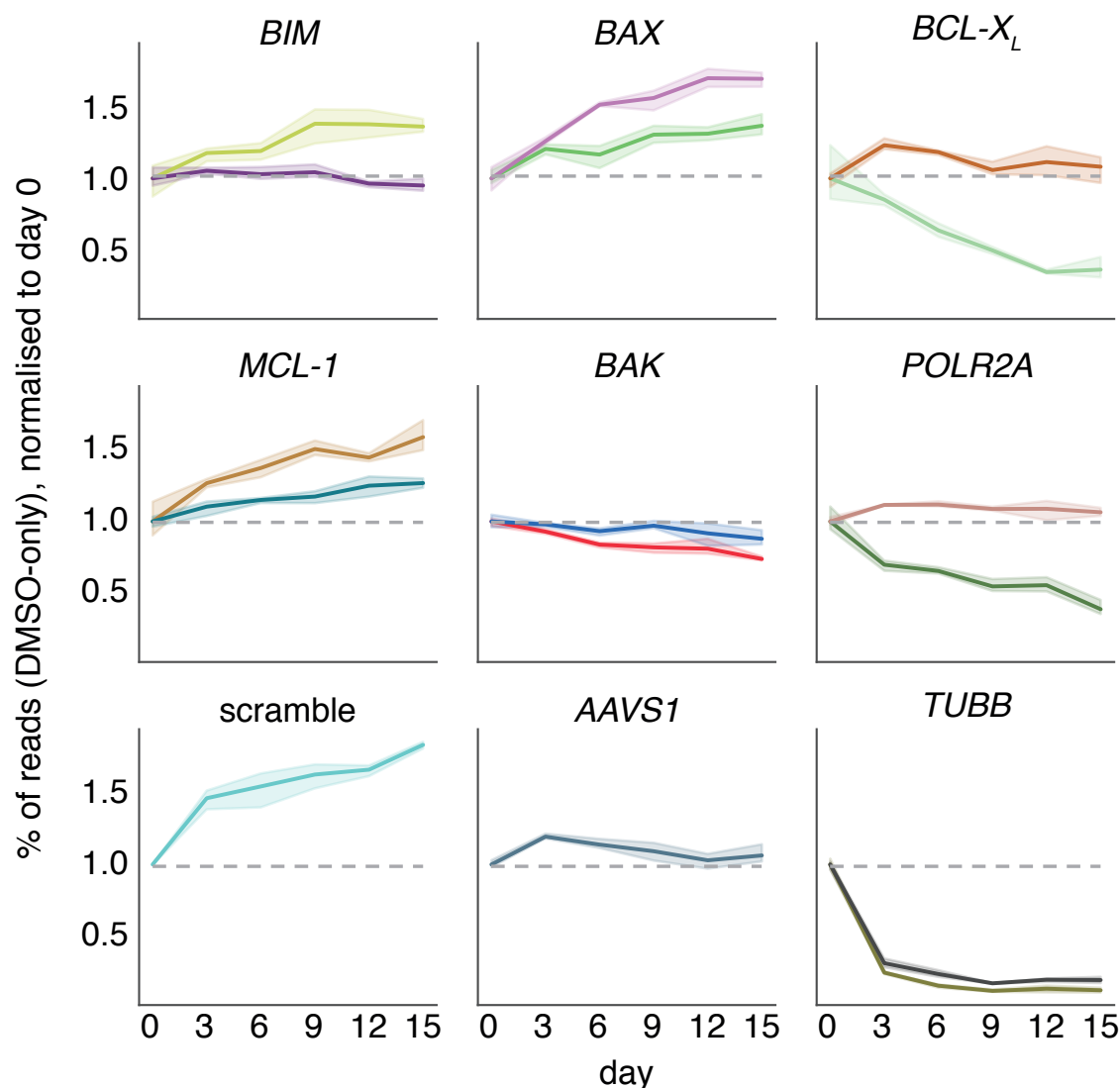


Figure 4.10: Fitness effects of miniscreen sgRNAs without selection. On day 0 of the miniscreen, cells were divided into two arms: control (DMSO-only) and treatment (VEN selection). Subplots show the abundance of each sgRNA at each timepoint relative to the abundance of that sgRNA on day 0 for the control arm only. Subplots are separated by gene; line colours correspond to Figure 4.9. Points show mean; bands show 95% confidence interval. n=3 replicate gDNA harvests per timepoint from one biological sample.

VEN. Similarly, MCL-1 and BCL-X_L sgRNAs were depleted; and POLR2A and AAVS1 were unchanged. It is unclear why one BIM sgRNA was also depleted, as was the scramble control. The TUBB sgRNAs were strongly enriched at early passages before also being depleted. Again, this may be an artefact of the small number of sgRNAs in the pool.

Assuming that cells were to be treated at the wild-type IC₅₀(VEN), the miniscreen demonstrated that five passages (15 days) is a suitable duration for a pooled screen with VEN in THP-1 cells. Although most trends observed in Figure 4.11 were already established after six days, I reasoned that more subtle resistance or sensitivity phenotypes may take longer to be detectable.

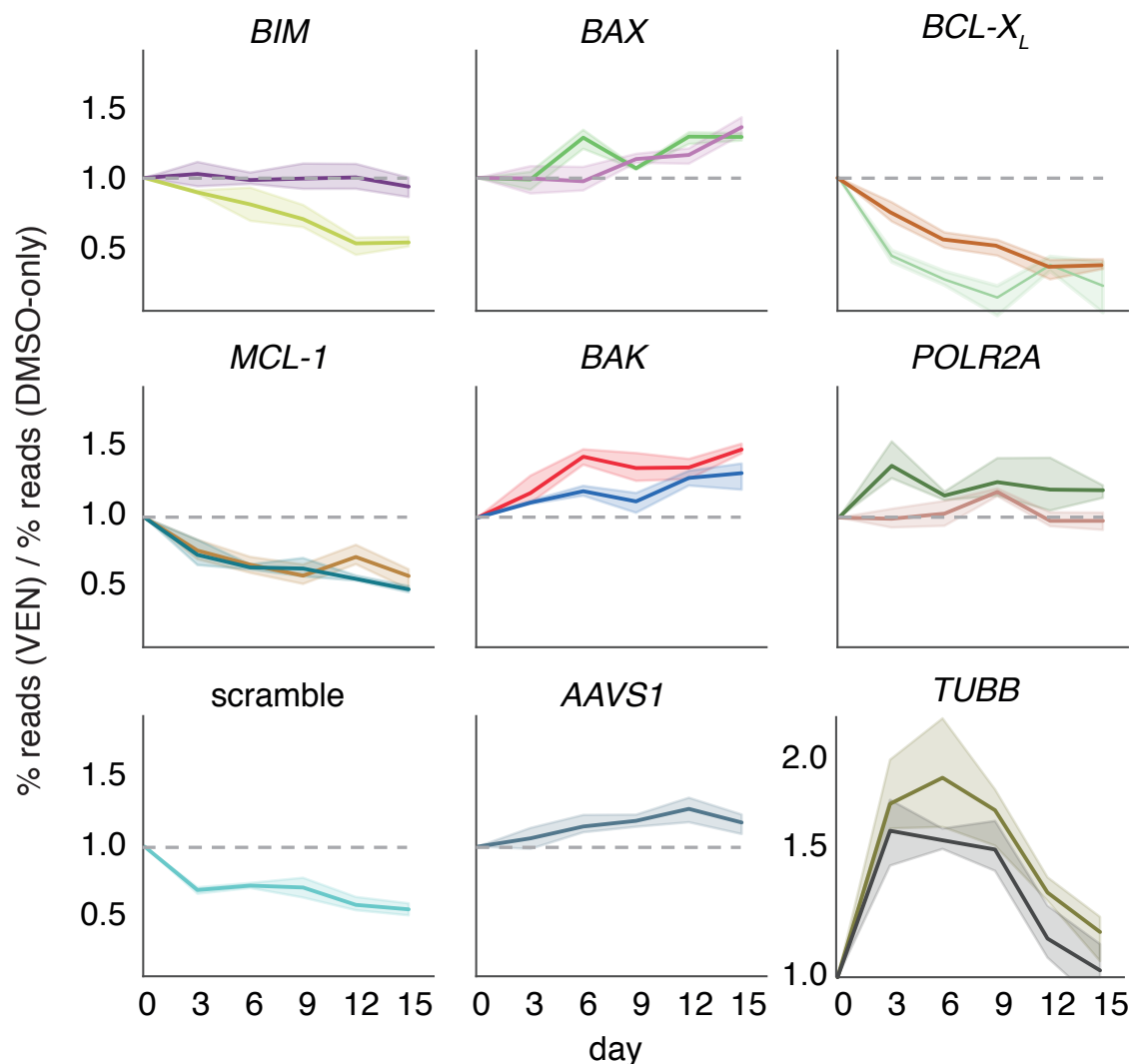


Figure 4.11: Effect of venetoclax selection on miniscreen sgRNA abundance. Subplots show the abundance of each sgRNA under VEN selection relative to the mean abundance of that sgRNA under control (DMSO-only) selection at the same timepoint. NB different scale for *TUBB*. Subplots are separated by gene; line colours correspond to Figure 4.9. Points show mean; bands show 95% confidence interval. n=3 replicate gDNA harvests from one biological sample per timepoint.

4.3 Discussion

In summary, in this chapter I show that:

1. THP-1 cells are more resistant to VEN than U-937 and HL-60 cells.
2. BCL-2 family genes can be individually knocked out using lentiCRISPRv2 and puromycin selection.

3. BAX KO cells have greater viability and proliferate more quickly than control under VEN selection.
4. MCL-1 and BCL-X_L KO cells proliferate more slowly than control under VEN selection.
5. Pooled lentiviral CRISPR viability screens with VEN selection are feasible in THP-1 cells.

4.3.1 THP-1 cells are an effective model for clinical resistance to VEN

Most studies report that the IC₅₀ of VEN varies significantly between AML/ALL cell lines. In order to identify a tractable cell line for screening experiments, I used the standard Annexin-V/PI apoptosis assay to compare the IC₅₀ of VEN in three AML/ALL cell lines: THP-1, U-937 and HL-60 (Figure 4.2).

The IC₅₀ values reported here match those reported in similar studies (Table 4.1). All lie within the typical range of VEN sensitivity in leukemia lines¹⁷⁶. It is worth noting that, in some studies, U-937 cells appear to be more resistant to VEN than THP-1. The variation in reported IC₅₀ values probably arises because cell lines are highly sensitive to culture conditions. In pharmacology, IC₅₀ refers to the half-maximal binding concentration for a specific target-ligand interaction. In oncology, however, the term is ill-defined and is typically used to refer to the relative or absolute half-maximal cytotoxic concentration, as here, or occasionally the half-maximal growth inhibitory concentration (also known as GI₅₀). However IC₅₀ be defined, methods for measuring and calculating it are poorly standardised. I chose the approach most relevant for conducting a whole-genome screen.

Other features of THP-1 cells make them a pragmatic choice for CRISPR knockout screens. Apart from the t(9;11)(p21;q23) translocation, THP-1 cells have relatively few karyotypic abnormalities and are near-diploid¹³⁴. Stable copy number is important for reliable whole-genome screening (see chapter 5 for more details). THP-1 cells, like U-937 and HL-60, exhibit typical expression levels of BCL-2,

study	[323]	[188]	[327]	[191]	[328]	[329]	[198]	[330]	[213]
HL-60	0.1	0.002	–	–	–	–	–	–	–
THP-1	2.4	0.08	>2	>2	>4	8	40	10	>1
U-937	13.5	1	>4	>8	>4	–	–	–	>1
method	AV/PI (48h)				CTG (24h)		CTG (72h)	PS	

Table 4.1: IC₅₀(VEN) values reported in the literature. Values are in nM. AV/PI: Annexin-V/PI. CTG: CellTiter-Glo. PS: phosphatidylserine externalisation.

BCL-X_L and MCL-1³²³. BCL-2 expression and VEN sensitivity are positively correlated, albeit weakly, in AML cell lines¹⁸⁸.

I reasoned that using a cell type that is highly resistant to VEN would make it easier to detect sensitising genes. In retrospect, it may have been more sensible to choose a more sensitive cell line in order to detect resistance-causing genes as well. This would be an interesting basis for future experiments. Another alternative would have been to select VEN-sensitive THP-1 subclones and use these instead of wild-type THP-1 cells. This would have had the added advantage of overcoming genetic heterogeneity in the population^{324,325}.

It is interesting to ask whether the *in vitro* doses of VEN used in my experiments are physiologically relevant. Based on PK/PD data used to support the MURANO trial³²⁶, C_{max}(VEN)≈2 mg. Given M_w(VEN)=868.44 g/mol, C_{max}(VEN)≈2 μM. Thus THP-1 and U-937 cells (but not HL-60) exhibit an IC₅₀(VEN) in the same order of magnitude as the effective clinical concentration of VEN. A similar IC₅₀(VEN) is observed in primary cells from AML patients³²³.

4.3.2 Intrinsic apoptosis genes contribute to VEN resistance in MLLr leukemia

The western blots in Figure 4.5C confirm that I successfully targeted five key BCL-2 family proteins using lentiCRISPRv2. For some targets, especially BCL-X_L, residual protein was detected. Even for BAX, it is likely that a low level of expression remained. Differences in knockout levels may have been due to observed antibody affinity, sgRNA on-target efficiency or the essentiality of the target gene.

The method for empirically determining viral titre was optimised (Figure 4.3, Figure 4.4) to avoid transducing more than one sgRNA per cell. Figure 4.3 shows that selection using 2 $\mu\text{g}/\text{mL}$ puromycin slightly overestimates viral titre. This further reduces the risk of multiple transduction. For a more precise comparison between pLenti CMV GFP Puro and lentiCRISPRv2, ideally the same promoter would drive expression of EGFP in both plasmids and PuroR in pLenti CMV GFP Puro. Nonetheless, observed titre matched theoretical predictions (Figure 4.4C), suggesting that pLenti CMV GFP Puro provides good insight into puromycin selection with lentiCRISPRv2 and that the Poisson distribution is an effective model for quantifying lentiviral infection. These optimised conditions were used for the arrayed knockout experiment, the miniscreen and the whole-genome screen.

The effect of BCL-2 family gene knockouts on VEN sensitivity (Figure 4.6) corresponded in part with published literature and predictions based on the known functions of these genes in intrinsic apoptosis. VEN competitively inhibits BIM from binding to BCL-2. The liberated BIM is instead sequestered by MCL-1 and possibly BCL-X_L, in line with the clinical observation that MCL-1 expression is inversely correlated with VEN sensitivity^{188,191,327}. I observed that MCL-1 or BCL-X_L knockout weakly sensitised THP-1 cells to VEN. This agrees with several similar lines of evidence: THP-1 cells cultured in 2.5 μM VEN acquire resistance by increasing expression of MCL-1 and BCL-X_L³³⁰; siRNA knockdown of MCL-1 in OCI-AML3, a relatively VEN-resistant cell line, sensitises the cells to VEN¹⁸⁸; and small-molecule inhibitors of MCL-1 and BCL-X_L sensitise MOLM-13 cells to VEN^{194,218,331,332}. Similarly, the strong sensitising role of BAX but not BAK (Figure 4.6) corroborates AML patient expression data³³³ and agrees with experiments in VEN-resistant murine cell lines³³⁴.

It was surprising that BIM knockout did not affect VEN sensitivity. This is perhaps explained by the observation that VEN induces apoptosis even in the absence of MCL-1, BCL-X_L and all eight BH3-only proteins, suggesting that the mechanism of action of VEN is independent of BIM²¹². Dual MCL-1/BCL-X_L

inhibition also indicates that BIM is dispensable^{213,335}. Further mechanistic research will fully reveal the role of BIM in VEN-induced apoptosis.

The five-day VEN treatment (Figure 4.6) measured a combination of cytotoxicity and growth inhibition. A pooled screen would be more likely to detect small differences in VEN sensitivity if knockouts affected cytotoxicity rather than proliferation: dead cells are depleted more quickly than slow-growing cells. Although the Annexin-V/PI assay found that only BAK and BAX mutants differed substantially from control, this experiment demonstrated that some BCL-2 family mutants have an elevated $IC_{50}(\text{VEN})$ (Figure 4.7). Moreover, it offered a partial explanation for the difference between BAX and BAK in Figure 4.6, underlining that they are not functionally equivalent. It should be noted that the cells were non-clonal so these IC_{50} values may not reflect the true IC_{50} of VEN in the corresponding mutant lines. It may also have been simpler to edit cells with a single sgRNA, rather than a pool of two sgRNAs per gene.

In summary, BAX, MCL-1 and BCL-X_L seemed to make the greatest contribution to VEN sensitivity in THP-1 cells. Lentiviral CRISPR is a simple and effective method for BCL-2 family knockout in THP-1 cells. Moreover, this approach allowed me to confirm the activity of several TKOv3 sgRNAs before using them in the miniscreen. This is a sensible first step in any whole-genome CRISPR screening workflow.

Using lentiviral CRISPR, future experiments should focus on the effect of combined knockouts (such as MCL-1/BCL-X_L) or other poorly understood anti-apoptotic proteins such as BFL-1, an antiapoptotic protein that exerts its function by sequestering BH3-only proteins, mainly BIM, BID, PUMA, and NOXA¹⁴⁵. BFL-1 overexpression has been linked to resistance to ABT-737 and VEN in lymphoma^{336,337}.

4.3.3 Pooled CRISPR screens are feasible in THP-1 cells

Having shown that TKOv3 sgRNAs targeting the BCL-2 family were effective in THP-1 cells, I set out to reproduce these results in a pooled miniscreen as a pilot

experiment for the genome-wide screen described in chapter 5.

Analysis of sgRNA enrichment and depletion in the DMSO-only condition demonstrated that sgRNAs targeting some essential genes are rapidly depleted from the pool as the cells expressing them die (Figure 4.10). One sgRNA targeting *POLR2A* was not selected, presumably because it does not efficiently knock out the target gene. But if one of two sgRNAs targeting *POLR2A* did not work, then some of the cells edited at MOI=0.3 with these sgRNAs in the arrayed knockout experiments should have survived. Either MOI was too high and most cells were transduced with both sgRNAs or, more likely, any healthy cells edited with the faulty sgRNA died due to low cell density. This could be solved by comparing activity of the two sgRNAs separately *in vitro*.

Several genes, including the negative controls, appear to have been confounded by the effects of a very small pool size, making the miniscreen results difficult to analyse in isolation. Nonetheless, analysis of the VEN selection arm did partly reproduce the trends observed in the arrayed knockout experiments, in particular the depletion of sgRNAs targeting MCL-1 and BCL-X_L, as well as enrichment of sgRNAs targeting BAX (Figure 4.11). Including BCL-2 in the miniscreen might have allowed me to measure non-specific effects of VEN, although none are predicted to occur at IC₅₀.

It is likely that using a larger sgRNA pool – targeting more genes with more sgRNAs per gene – would have significantly improved the reliability of the miniscreen. As I find in chapter 5, many of these issues are absent with a genome-wide sgRNA library. Yet given limited resources, the miniscreen was an effective proof of concept and showed that five passages are sufficient for measurable changes in sgRNA abundance. For future drugs or phenotypes, I would conduct the miniscreen with up to 100 sgRNAs. At this scale, lentivirus could easily be prepared as a pool or individually and the risk of drift would be much reduced. It may also be informative to run miniscreens at different drug concentrations such as IC₂₀ and IC₈₀. I explore these questions further in my analysis of the whole-genome screen in chapter 5.

In conclusion, small-scale CRISPR knockout experiments in arrayed or pooled format confirm that MCL-1 and BCL-X_L are promising targets for VEN combination

therapies. But I hypothesised that other genes which interact more distantly with the BCL-2 family may also be functionally relevant to VEN resistance in MLLr leukemia. I turned to whole-genome screening as an unbiased method for discovering such genes.

5

Genome-wide screening to discover venetoclax combination therapies

Contents

5.1	Introduction	109
5.2	Results	114
5.3	Discussion	128

5.1 Introduction

Analysis of AML/ALL patients confirms that venetoclax resistance cannot be explained by BCL-2 family genes alone³³³. MLLr patients may respond marginally better than non-MLLr patients to VEN monotherapy, but the overall failure of rational VEN combination therapies in AML/ALL indicates that genes beyond the canonical mitochondrial apoptosis pathway must be involved^{200,201}. In this chapter, I describe a genome-wide CRISPR screen in THP-1 cells designed to discover genes that could mediate resistance or sensitivity to VEN treatment and to predict potential VEN combination therapies.

Many lines of evidence show that the response to BH3 mimetic drugs such as VEN does not only depend on proteins in the known mitochondrial apoptosis

pathway. For example, it has been shown in t(4;11) SEM cells that MLL-AF4 increases BCL-2 expression²⁰⁰. A likely explanation is that MLL-AF4 recruits the histone methyltransferase DOT1L, which increases expression of HOXA9. In turn, HOXA9 activates BCL-2¹⁶⁰. Accordingly, the DOT1L inhibitor EPZ5676 synergises with VEN in SEM cells²⁰⁰.

In another example, BH3-mimetic inhibitors of MCL-1 and BCL-X_L lead to the same degree of BAX activation and apoptosis in HCT116 Octa-KO cells (in which all eight BH3-only proteins have been deleted) as in wild-type HCT116 cells. This suggests that the BH3 proteins do not directly contribute to VEN resistance^{212,213}.

In chapter 4, I used CRISPR-KO to inactivate apoptotic genes in THP-1 cells. Although reverse screening is a powerful approach for investigating known genes of interest, discovering novel VEN resistance genes requires a forward screening approach. Building on the methods developed in chapter 4, in this chapter I turn to genome-wide screening as an unbiased technique to discover VEN resistance genes. This workflow is applicable not only to VEN, but also to other cases of clinical drug resistance.

5.1.1 Experimental design of CRISPR screens

Genome-wide CRISPR screens are not all created equal. Many aspects of experimental design depend on the target phenotype and candidate genes²⁴¹. Having already optimised lentiviral transduction and selection in THP-1 cells and having demonstrated that cell growth was a reliable readout for VEN resistance, I based the design of the genome-wide screen on that of the miniscreen described in chapter 4.

The screen was designed to identify two types of hits: genes whose knockout causes drug sensitivity (‘dropout genes’) and genes whose knockout causes drug resistance (‘enriched genes’). In a pooled CRISPR-KO screen, each gene is targeted by several sgRNAs delivered via integrating lentivirus. Most loss-of-function mutations are silent with respect to a given drug resistance phenotype. For some genes, however, loss of function increases drug resistance or sensitivity. By exposing growing cells to a drug, cells containing sgRNAs which increase drug sensitivity

library	GeCKOv2 ²⁵⁸	Yusa ²⁴⁹	Brunello ³³⁹	TKOv3 ²⁷¹
target scope	coding and noncoding	coding	coding	coding
total target genes	20 914	18 009	19 114	18 056
total sgRNAs (excl. controls)	122 411	90 709	77 441	71 090
sgRNAs/gene	6	~5	~4	~4

Table 5.1: Popular libraries for genome-wide CRISPR-KO screens in human cells.

will proliferate more slowly, whereas cells containing sgRNAs which increase drug resistance will proliferate more quickly. Thus, sensitising sgRNAs are depleted from the pool while resistance-causing sgRNAs are enriched. A gene whose cognate sgRNAs are mostly depleted is scored as a dropout gene; a gene whose sgRNAs are enriched is an enriched gene.

I reasoned that dropout genes could encode promising targets for VEN combination therapies, whilst enriched genes would be useful for understanding the broader mechanism of mitochondrial apoptosis in MLL-rearranged leukemia. Dropout screening is more challenging than enrichment screening because the negatively selected sgRNAs can easily drop out of the population before they reach statistical significance. Compared to an enriched gene, achieving the same false discovery rate for a dropout gene requires more sgRNAs per gene, greater fold representation of each sgRNA in the pool (also known as coverage) or a longer screen duration. Most tools for analysing CRISPR screens are designed to detect both enriched and dropout hits, although this can lead to a high degree of false positives³³⁸.

There are several genome-wide sgRNA libraries widely used for human CRISPR-KO screens: GeCKOv2²⁵⁸, Yusa²⁴⁹, Wang/Sabatini²⁴⁷, Avana and Brunello³³⁹, and TKOv3^{271,340} (Table 5.1). Each library is designed to optimise classification sensitivity and specificity while minimising the size of the library. High sensitivity and specificity translate to better identification of true positive and true negative hits.

The performance of a library depends on the sgRNA sequences it contains. The ideal sgRNA for a pooled CRISPR-KO screen has:

1. high on-target specificity and activity;

2. DSB repair outcomes biased towards gene knockout;
3. high knockout rate (indels lead to protein loss-of-function);
4. low off-target activity.

sgRNA performance probably depends on a combination of primary sgRNA sequence, target gene sequence and chromatin context but the underlying processes are not fully understood²⁸⁶. Nonetheless, sequence-based sgRNA prediction tools have been successful: the *in silico* efficiency of an sgRNA is a good predictor of its *in vivo* knockout activity, regardless of cell type^{241,341–343}. The first CRISPR-KO libraries were built from the bottom up using a set of heuristic sgRNA design principles. For example, the original GeCKOv1 library was constructed using a simple off-target prediction algorithm without accounting for on-target activity²⁵⁵. GeCKOv2 was designed in a similar way and, like its predecessor, it was hampered by poor on-target efficiency and several non-specific sgRNAs^{243,258,339}. More recent libraries such as Brunello are optimised for on-target efficiency and knockout efficiency³³⁹. They combine heuristic sgRNA design rules with functional validation, achieved by benchmarking sgRNA performance against other libraries or reference sets of known essential genes.

For the screen described in this chapter, I used the TKOv3 library (Figure 5.1B), a widely used genome-wide sgRNA library targeting only protein-coding genes. The original TKOv1 library, which contains six sgRNAs per gene, was based on similar design rules to Avena and Brunello³⁴⁰ and validated against a set of reference genes identified from a panel of pooled shRNA screens. In subsequent work, Moffat and coworkers applied the Bayesian Analysis of Gene Essentiality (BAGEL) algorithm to a meta-analysis of twelve CRISPR essentiality screens in various cell lines. They determined an expanded set of essential genes and improved their sgRNA design rules to generate TKOv3, a refined sequenced-optimised library with only four sgRNAs per gene²⁷¹. TKOv3 is a single-vector library: both SpCas9 and the sgRNA are delivered in the same vector, lentiCRISPRv2 (Figure 4.3B).

screen stage	coverage metric
amplification of plasmid library	plasmids per sgRNA
lentiviral transduction	cells per sgRNA
cell culture and drug selection	cells per sgRNA
gDNA harvesting	genomes per sgRNA
PCR amplification of sgRNA sequences	templates per sgRNA
NGS	reads per sgRNA

Table 5.2: Screening steps critical for library coverage.

CRISPR screens are experimentally challenging. It is preferable to use a smaller library since this reduces the practical burden and allows higher throughput. The more sgRNAs per gene in a library or the better each sgRNA is designed, the more robust the library is to subsampling caused by low coverage or too few replicates. The TKOv3 and Brunello libraries both use four sgRNAs per gene. This is a good compromise between discovery rate and library size: for the TKOv3 library, Moffat and coworkers found that expanding the library to include more than four sgRNAs per gene did not significantly improve the discovery rate. Similarly, they found that addition of a second replicate improved discovery by only $\sim 10\%$; addition of a third replicate, by $< 5\%$ more. In an independent analysis based on the GeCKOv1 library, Yusa and coworkers found that using > 4 sgRNAs per gene led to quickly diminishing returns³³⁸. Similarly, with four sgRNAs per gene, one biological replicate was sufficient to recover most hits³³⁸.

5.1.2 Screening coverage

For a pooled CRISPR screen using any library, the most important aspect of experimental design is coverage. The impact of coverage on reproducibility was first described for shRNA screens³⁴⁴. At low coverage, there is a risk that sgRNAs may become enriched or depleted in the pool due to random drift rather than positive or negative selection. If this occurs for several sgRNAs targeting the same gene, there is a risk of false positives or false negatives. A minimum coverage of 100x to 500x is normally recommended at every stage of the screen^{241,243}. These stages are summarised in Table 5.2.

5.1.3 Analysing pooled CRISPR screens

In this chapter, I describe an analysis workflow which includes Model-based Analysis of Genome-wide CRISPR/Cas9 Knockout (MAGeCK) to compare sgRNA read counts from different time points and conditions (Figure 5.2)^{277,279,280}. Several other algorithms have been developed or repurposed for analysing pooled CRISPR screens, including EdgeR³⁴⁵, DESeq³⁴⁶ (both originally for RNA-Seq expression data) and RIGER³⁴⁷ and RSA³⁴⁸ (both for microarray-based RNA interference screens). Compared to these algorithms, MAGeCK has several advantages: lower false positive and false negative rate, better consistency between replicates, and robustness against the effects of low sequencing depth or low numbers of sgRNAs per gene.

Recently, several end-to-end analysis platforms have been released, such as caRools³⁴⁹/CRISPRAnalyzeR³⁵⁰, PinAPL-Py³⁵¹, ENCoRE³⁵², CRISPRMatch³⁵³, JACKS³⁵⁴, MAGeCK-VISPR and MAGeCK-Flute. These platforms implement the standard algorithms within a more user-friendly interface, but at the cost of less precise control over the analysis.

In this chapter, I run a pooled CRISPR screen to find protein-coding genes which increase the sensitivity or resistance of THP-1 cells to VEN. I validate hits by single-sgRNA CRISPR-KO with VEN selection and find four druggable components of the 26S proteasome which contribute to VEN resistance. Finally, I use THP-1 cells as a model to test a VEN-proteasome inhibitor combination therapy.

5.2 Results

5.2.1 Genome-wide CRISPR screen

To begin, we amplified the TKOv3 library from bacterial stock (Figure 5.1A-B), then subjected the library to NGS at a coverage of ~200 reads per sgRNA (Figure 5.1C). The skew ratio was 3.5, where skew ratio is defined as the ratio of the sum of read counts for the top 10% of sgRNAs to the sum of read counts for the bottom

sample	T0	DMSO	VEN
transduction coverage (cells/sgRNA)	250	250	250
mean minimum cell culture coverage (cells/sgRNA)	–	361	1330
cell harvesting coverage	250	250	250
PCR coverage (cells/sgRNA)	120	120	120
mean aligned NGS reads (M)	14	12.6	13.2
mean aligned sgRNA coverage (reads/sgRNA)	201	176	188
genes undetected (%)	0.58	17	0.59

Table 5.3: Genome-wide screening coverage. Summary statistics for the genome-wide screen. Mean of two biological replicates.

10% of sgRNAs ranked by frequency. A skew value <10 is taken to indicate that a library has been faithfully amplified²⁴³.

To conduct the screen, I titred the TKOv3 lentivirus and used it to infect wild-type THP-1 cells at $\text{MOI}=0.3$ (Figure 5.1A). 72 hours later, I treated cells with puromycin to kill any untransduced cells. After a further 72 hours, I removed puromycin selection, ensuring that sufficient cells survived for a transduction coverage of 250 cells/sgRNA (Table 5.3). I harvested genomic DNA from a minimum of 1.8×10^7 cells (250 cells/sgRNA) to capture the initial distribution of sgRNAs (‘T0’). I divided the cells into DMSO-only (control) and VEN conditions and passaged them for 18 days (six passages). At each passage, I resuspended the cells in fresh medium containing DMSO or $7.4 \mu\text{M}$ VEN (IC_{50} , Figure 4.2). On day 3, ~50% of cells in the VEN condition survived VEN treatment; by day 18, ~75% of cells survived. Cell viability in the DMSO-only condition always exceeded 90%. On day 18, I harvested cells from the DMSO and VEN conditions, again from a minimum of 1.8×10^7 cells. After extracting gDNA, I used a minimum of 120 cells/sgRNA/sample as template to generate PCR-amplified libraries of sgRNAs for NGS. I conducted two fully independent biological replicates of the screen. For all samples, $\geq 70\%$ of NGS reads passing filter aligned with sgRNA sequences. The aligned reads represented the library at a coverage of ≥ 176 reads/sgRNA.

I combined tools to build a custom data analysis pipeline, summarised in Figure 5.2A. MAGeCK-RRA interprets sgRNA counts in four stages, described

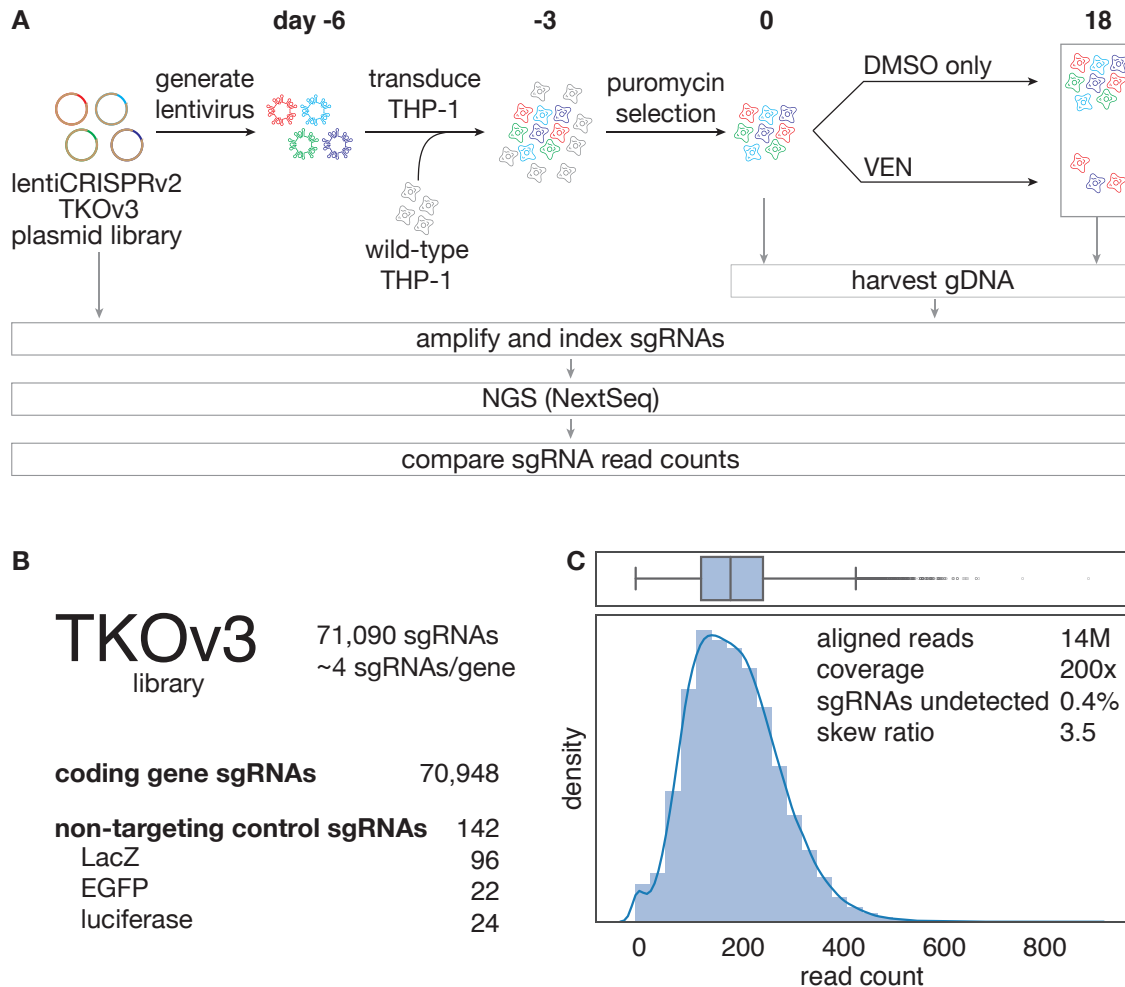


Figure 5.1: Screening with the TKOv3 library. (A) Simplified schematic for the genome-wide CRISPR screen. The TKOv3 plasmid library in lentiCRISPRv2, amplified from Addgene stocks and validated by NGS, was transfected into HEK293T cells to create pooled lentivirus. Lentivirus was titred and transduced at MOI=0.3 into THP-1 cells. Puromycin selection was applied to kill untransduced cells. Genomic DNA was harvested before splitting cells into VEN selection and DMSO-only (control) conditions which were maintained for six passages (18 days), after which gDNA was harvested from both conditions. Positive and negative selection of sgRNAs was measured by next-generation amplicon sequencing. (B) Summary of the TKOv3 pooled CRISPR screening library. (C) After amplification, the plasmid pool was sequenced to determine TKOv3 library representation. *Bottom:* histogram to show distribution of sgRNA read counts for all sgRNAs in the TKOv3 library. Curve: Gaussian kernel density estimate. *Top:* Tukey boxplot showing distribution of sgRNA read counts. Whiskers show 1.5xIQR; dots show outliers.

in detail below: (1) read-count normalisation, (2) variance modelling, (3) sgRNA ranking and (4) gene ranking (Figure 5.2B)²⁷⁷.

1. **Read-count normalisation.** In the first stage, the read count of each sgRNA in each sample is normalised against the median read count of that sgRNA across all samples. Normalising in this way accounts for differences in the overall number of read counts between different samples.
2. **Variance modelling.** The next step is to estimate the read-count variance for each sgRNA in each condition. MAGeCK builds a ‘mean-variance’ model based on the sample variance of all the control sgRNAs, then adjusts the variance of each sgRNA based on its mean. The variance cannot be calculated directly because there are typically few or no replicates.
3. **sgRNA ranking.** Thirdly, the mean read count and adjusted variance for each sgRNA in each condition are used to parametrise a negative binomial distribution³⁵⁵. Comparing these two distributions for treatment and control conditions generates a p-value for each sgRNA, which represents the likelihood that the read counts for a given sgRNA are drawn from different distributions for treatment and control.
4. **Gene ranking.** Finally, all the sgRNAs for each gene must be combined to generate a single p-value. MAGeCK uses α -RRA, a modified version of an algorithm built for gene set enrichment analysis³⁵⁶. All sgRNAs are ranked by p-value and then each gene is tested against the null assumption that its sgRNA ranks are uniformly distributed. The sgRNAs are randomly assigned to genes and the process is repeated for several permutations to calculate a null score for each gene. Comparing a gene’s actual score to its null score generates p-values and false discovery rates (FDRs). The p-values and FDRs are adjusted using copy-number variation data, if available.

Before scoring gene hits in MAGeCK, I examined the sgRNA read-count distribution for T0, DMSO and VEN samples for each of the biological replicates

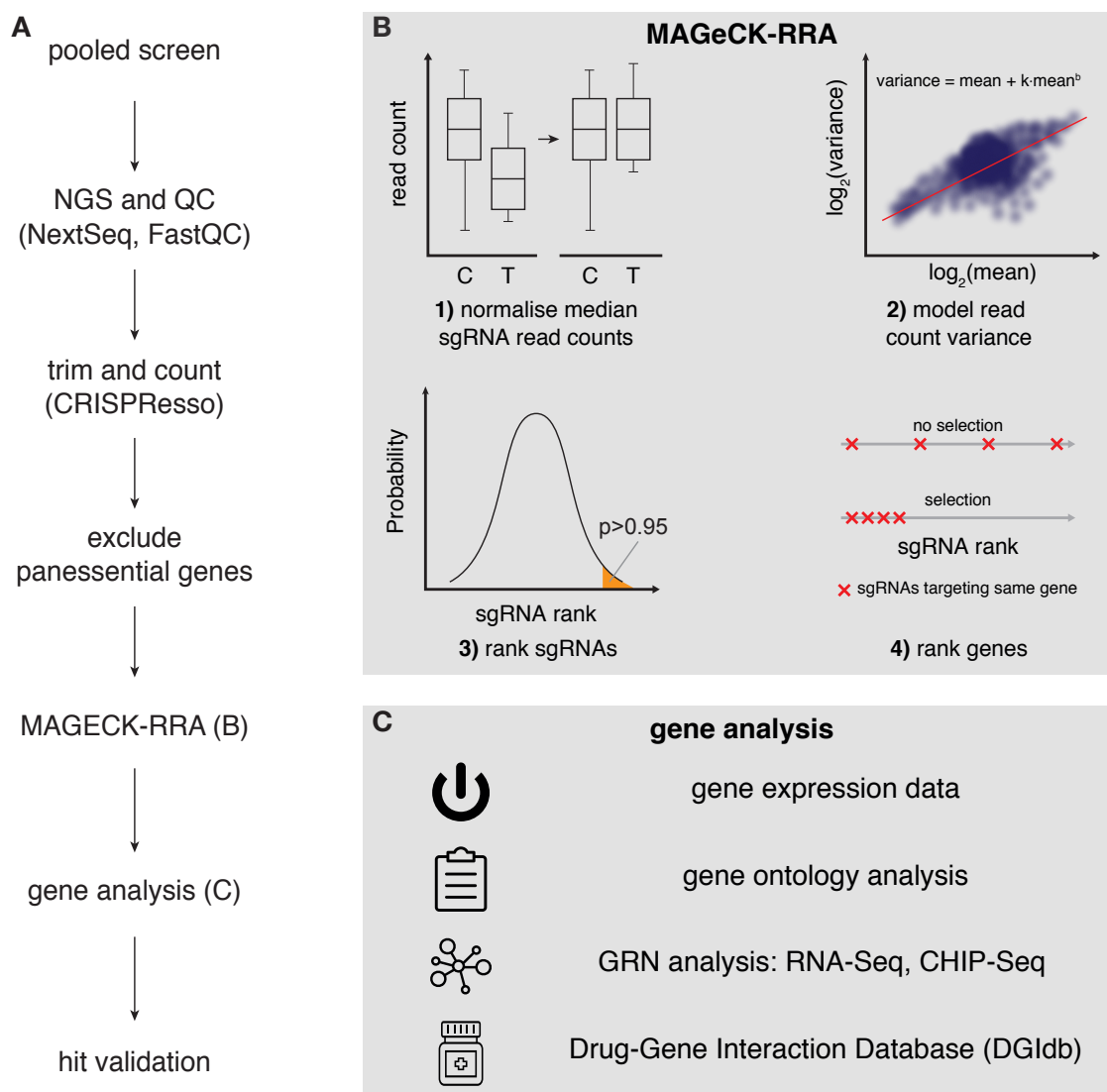


Figure 5.2: Genome-wide screening data analysis. (A) Overview of data analysis pipeline. (B) I used MAGECK-RRA to interpret sgRNA fold-changes and generate hits. See text for details of the MAGECK algorithm. Adapted from Li et al. (2014)²⁷⁷. (C) Positively or negatively selected genes were analysed in four stages.

(Figure 5.3A). As expected, the VEN read count distributions had a longer right tail than T0, indicating some positive selection. Surprisingly, though, the DMSO read count distributions had a lower median read count and longer right tail than either T0 or VEN, indicating that a large fraction of the sgRNAs had been depleted in that condition. Comparing the normalised read counts of individual sgRNAs in each biological replicate, I found that the T0 and VEN replicates were concordant ($R^2 > 0.69$) but the DMSO replicates were not (Figure 5.3B). In fact, returning to

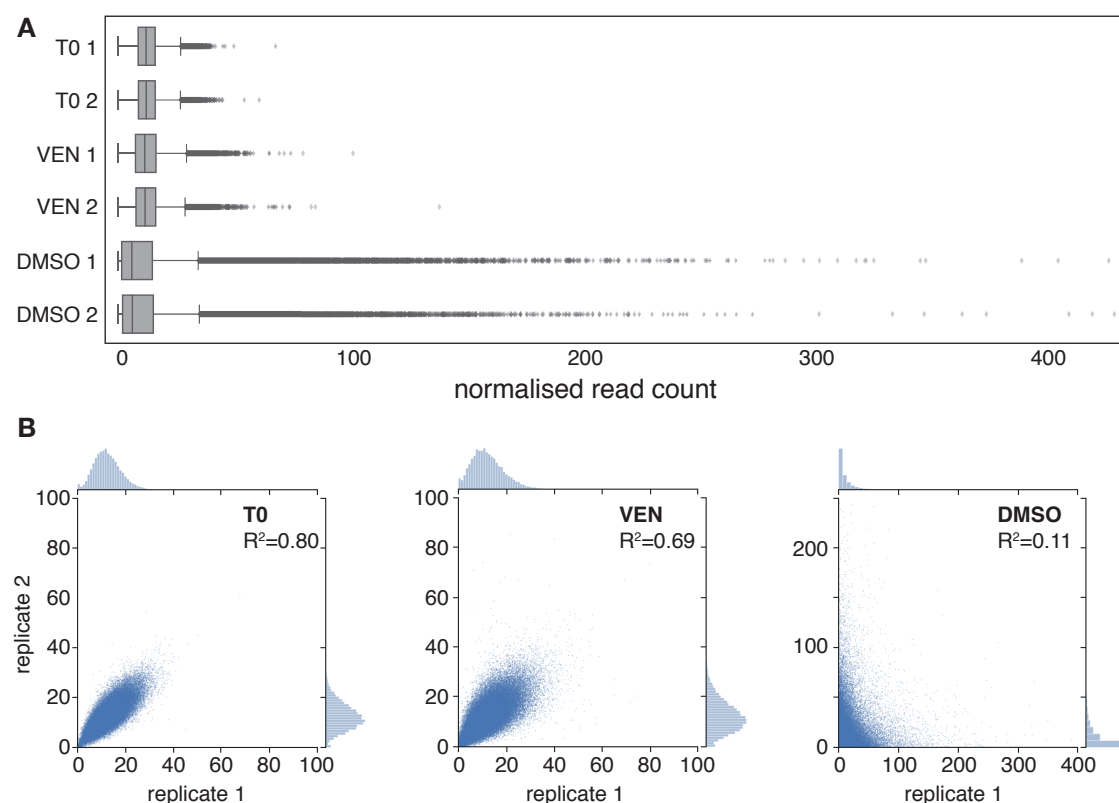


Figure 5.3: Comparing sgRNA read counts from replicate screens (A) Distribution of sgRNA read counts in each condition for two fully independent biological replicates of the genome-wide screen. Tukey boxplots: whiskers show 1.5xIQR; diamonds show outliers. **(B)** Scatter plots to compare normalised read counts for each sgRNA across biological replicates. R^2 : linear regression correlation coefficient.

Table 5.3, 17% of sgRNAs in the library were undetected in the DMSO samples, relative to $<0.6\%$ in the other samples. The most likely explanation was that coverage was insufficient in the DMSO replicates. If many sgRNAs had been lost due to drift, then the remaining sgRNAs in each replicate would represent a random subset of the population and the replicates would not match. If coverage was to blame, the issue must have arisen between day 0, when the replicates matched, and day 18, when the VEN and DMSO samples were harvested and prepared in parallel. In fact, the coverage of the VEN and DMSO samples differed only during the screen itself, where the minimum coverage while passaging the DMSO replicates was 361 cells/sgRNA compared to 1330 cells/sgRNA for VEN (Table 5.3). Accordingly, I concluded that the failure of the DMSO control did not compromise the VEN-selection arm of the screen.

A typical CRISPR drug screen compares the drug selection condition with the control condition at each timepoint. The purpose of the control condition is to exclude sgRNAs which affect cell growth independently of drug selection. The majority of these sgRNAs target so-called essential genes such as *POLR2A* and *TUBB* (chapter 4). Since I was unable to use the DMSO samples to exclude essential genes, I chose instead to compare the VEN condition to T0. I used the literature to identify a conservative set of 640 ‘panessential’ genes to exclude from further analysis. The panessential genes were defined as the intersection of three public sets of essential genes derived from several different cancer cell lines^{248,249,357}.

I analysed the T0 and VEN read counts using MAGeCK-RRA with 10 000 permutations. Using a relatively strict FDR cut-off of 5%, I identified 40% of panessential genes as dropouts (Figure 5.4A, D). I repeated MAGeCK-RRA analysis after eliminating the panessential genes and identified 219 hits, 209 of which were dropout genes (Figure 5.4B). Gene ontology (GO) analysis revealed that the dropouts were highly enriched for genes involved in cell cycle regulation and mitochondrial regulation, such as *PPP1R2* (a serine/threonine phosphatase responsible for the cell stress response), *DNM1L* (a mitochondrial GTPase) and *MUL1* (a mitochondrial E3 ubiquitin ligase) (Figure 5.4C). Interestingly, the dropouts included four components of the 26S proteasome (*PSMB5*, *PSMC1*, *PSMD2* and *SEMI1*). Another proteasome-interacting protein (*PSMC3IP*) was depleted but did not cross the FDR cut-off.

Enriched hits included *BAX* (the immediate effector of mitochondrial apoptosis), *CUL2* (an E3 ubiquitin ligase involved in intrinsic apoptosis), *NOA1* (a mitochondrial GTPase) and the effector caspase *CASP9*, with GO terms relating to apoptosis and mitochondrial processes. The majority of genes were neither enriched nor depleted: for example, *LGMN*, which encodes a lysosomal cysteine protease, had $\text{FDR}=1$ and $\log_2(\text{fold change}) < 1 \times 10^{-4}$.

Apoptotic genes from the miniscreen (chapter 4) were enriched or depleted as expected, but only *BAX* was statistically significant. *BAK* was very weakly enriched ($\text{FDR}=0.96$, $\log_2(\text{FC})=0.02$); *MCL1*, *BCL-X_L* and *BIM* were weakly depleted. *BCL2* was neither enriched nor depleted ($\text{FDR} \leq 1$, $\log_2(\text{fold change}) < 0.01$).

I returned to the sgRNA-level p-values (prior to α -RRA) and compared the strength of sgRNA selection between control sgRNAs and sgRNAs targeting hits. As expected, I found that sgRNAs targeting hits were more strongly selected than controls (Figure 5.4E).

I chose a subset of enriched and dropout hits based on their association with the most heavily enriched GO terms. Analysing mean normalised read counts for these hits confirmed that all four sgRNAs were concordantly enriched or depleted for most hits (Figure 5.5). I observed a similar pattern for some genes such as *PSMC3IP* and *CASP3*, although these did not meet the FDR cut-off, perhaps due to the stochastic nature of MAGECK-RRA. sgRNAs targeting neutral genes such as *LGMN* were not selected significantly during the screen.

I directly validated top hits by individually cloning sgRNA sequences into lentiCRISPRv2. I tested two sgRNA sequences per gene, except for *AAVS1*, *LGMN* and *NOA1* for which I tested only one sgRNA. *AAVS1* and *LGMN* served as silent targeting controls. I used sgRNA sequences from the Brunello library by selecting the highest-scoring sgRNAs for each gene. None of the Brunello sgRNAs had the same sequence as their TKOv3 counterparts. Two TKOv3 sgRNA sequences targeting *BAX* served as positive controls. For each of three independent biological replicates, I transduced 39 different lentivirus preps into THP-1 cells (three technical replicates per lentivirus) at MOI=0.3 and selected them with puromycin following the same schedule as the genome-wide screen (Figure 5.1A). Next I treated them with 20 μ M VEN for 48 hours and measured their viability using the CellTiter-Glo assay (Figure 5.6). I compared viability to the *AAVS1* control. Most dropout and enriched genes performed as expected, confirming that these genes affected VEN resistance.

5.2.2 Finding druggable targets and combination therapies

Dropout hits represent genes which normally contribute to VEN resistance. Targeting these genes alongside VEN treatment could increase sensitivity to VEN, forming the basis of a combination therapy. Most of the human proteome is currently

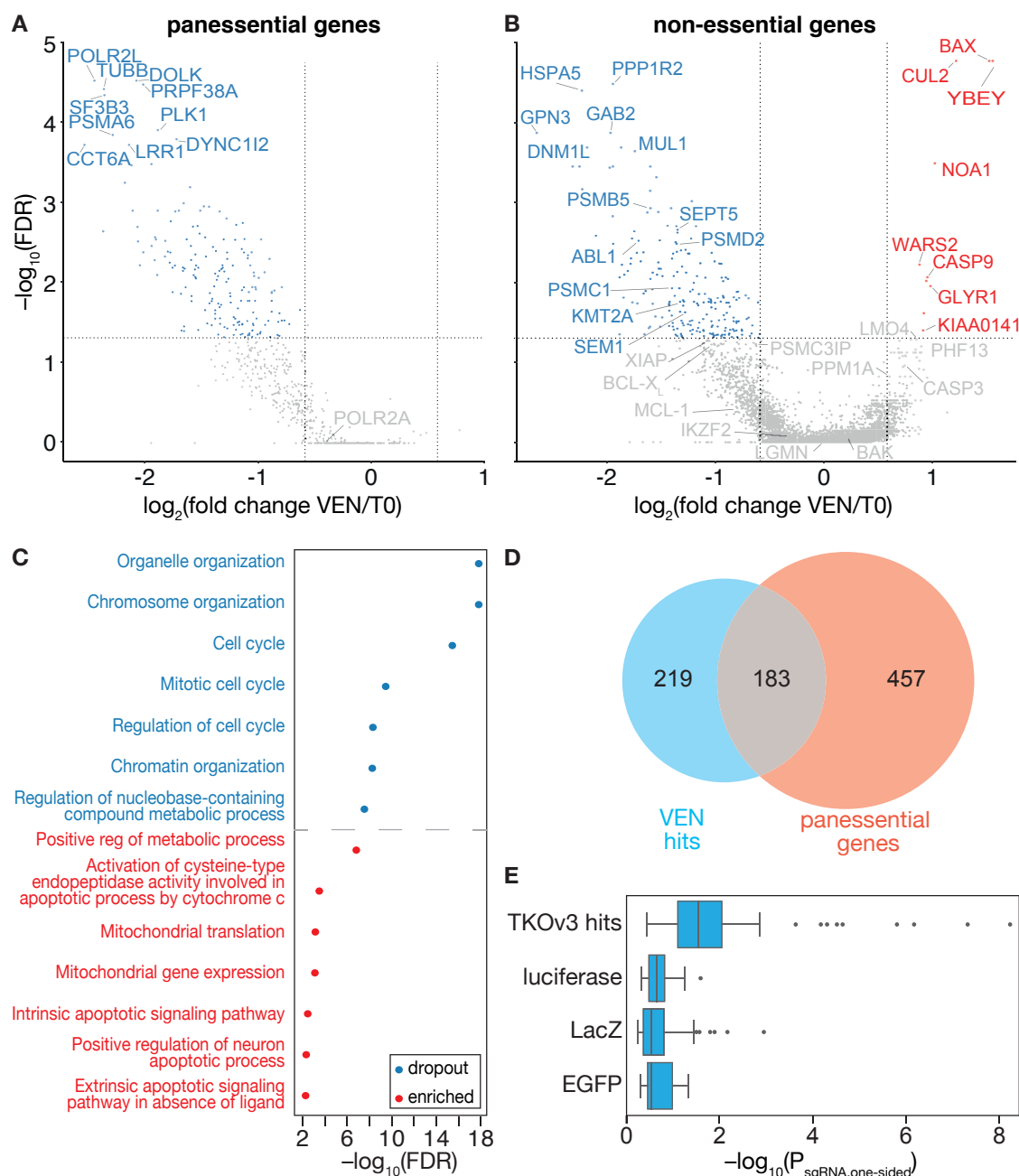


Figure 5.4: Genes affecting VEN resistance in THP-1 cells. (A-B) Volcano plots to show positive (blue) and negative (red) gene-level selection determined using MAGeCK-RRR, comparing VEN (day 18) with day 0. (A) Genes predicted to be panessential according to three genome-wide essentiality screens: the top ten dropouts are labelled. (B) All other protein-coding genes: genes of interest are labelled. $n=2$ biological replicates. FDR: false discovery rate. Cut-offs: $\text{FDR}=5\%$, $\log_2(\text{FC})=1.5$. (C) Enriched gene ontology terms for positively and negatively selected screening hits. (D) Intersection between panessential genes and all hits (positive and negative) in the VEN screen. (E) Comparison of hits and controls at the sgRNA level, based on one-sided P-values (for enrichment or depletion) determined using MAGeCK. Tukey boxplots: whiskers show 1.5xIQR; diamonds show outliers.

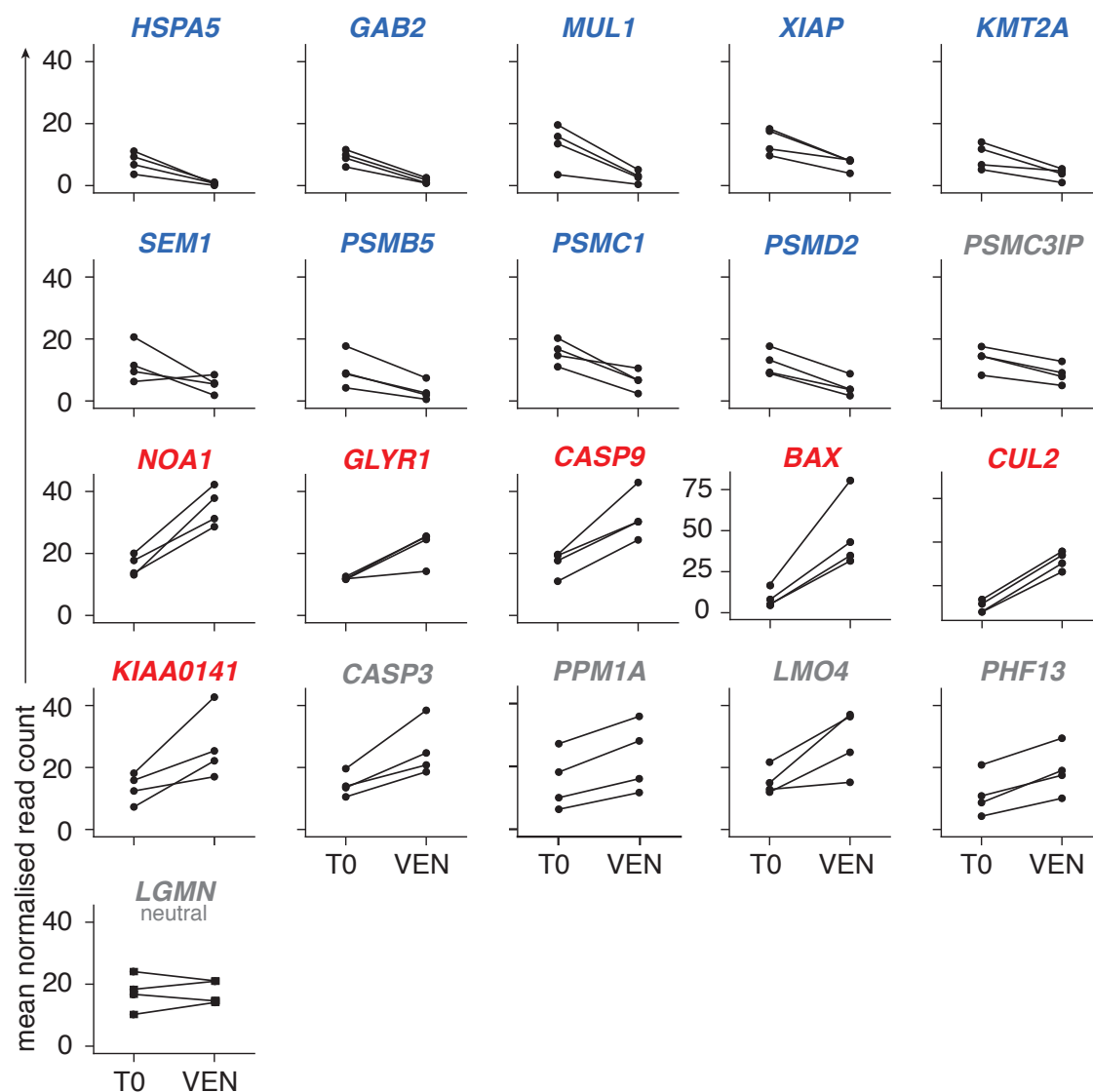


Figure 5.5: sgRNA-level selection for key depleted and enriched genes. Changes in mean normalised read count for dropout hits (blue), enriched hits (red) and other genes which did not meet the FDR cut-off (grey) between day 0 (T0) and day 18 with VEN selection (VEN). Note different scale for *BAX* and *CUL2*. n=2 biological replicates.

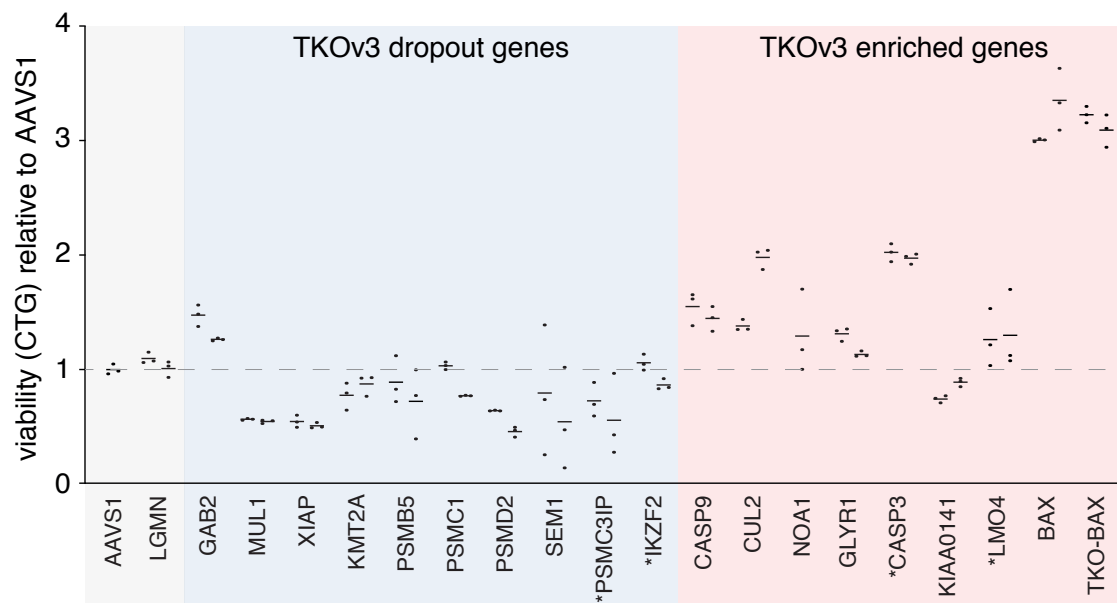


Figure 5.6: Functional validation of screening hits. To validate positive and negative TKOv3 hits, I generated two separate single-sgRNA knockouts of each gene using sgRNAs from the Brunello library³³⁹. I cloned two sgRNAs per gene into lentiCRISPRv2 constructs and individually transduced them into THP-1 cells, then measured their sensitivity to 48-hour treatment with 20 μ M VEN using the CellTiter-Glo assay. Viability was normalised against cells transduced with the silent control sgRNA targeting *AAVS1* from Wang et al.²⁴⁸. The top two *BAX*-targeting sgRNAs from the TKOv3 library were included as positive controls. Genes marked * were selected but did not meet the FDR cut-off in the screen. Only one sgRNA was used for *NOA1*. Dots show biological replicates (n=3); bars show mean.

undruggable³⁵⁸. In fact, <10% of known proteins are associated with an approved or investigational small molecule³⁵⁸. Focusing on the validated dropout hits, I asked which of these could be inhibited using known small molecules. By querying the Drug-Gene Interaction Database (DGIdb), a searchable database of drug-gene interactions, I found that the most druggable dropout hits encoded 26S proteasome components (PSMB5, PSMC1, PSMD2 and SEM1)³⁵⁹.

Next I investigated whether any dropout hits from the screen were targeted by drugs in the DGIdb and associated with the MLLr transcriptional programme. Previous work in the Milne group had established a model gene regulatory network (GRN) for MLL-AF4 leukemia by combining a promoter motif network from THP-1 cells with MLL-AF4 ChIP-seq and RNA-seq [*manuscript in preparation*]. Although THP-1 cells express MLL-AF9 instead of MLL-AF4, it is known that MLL-AF4

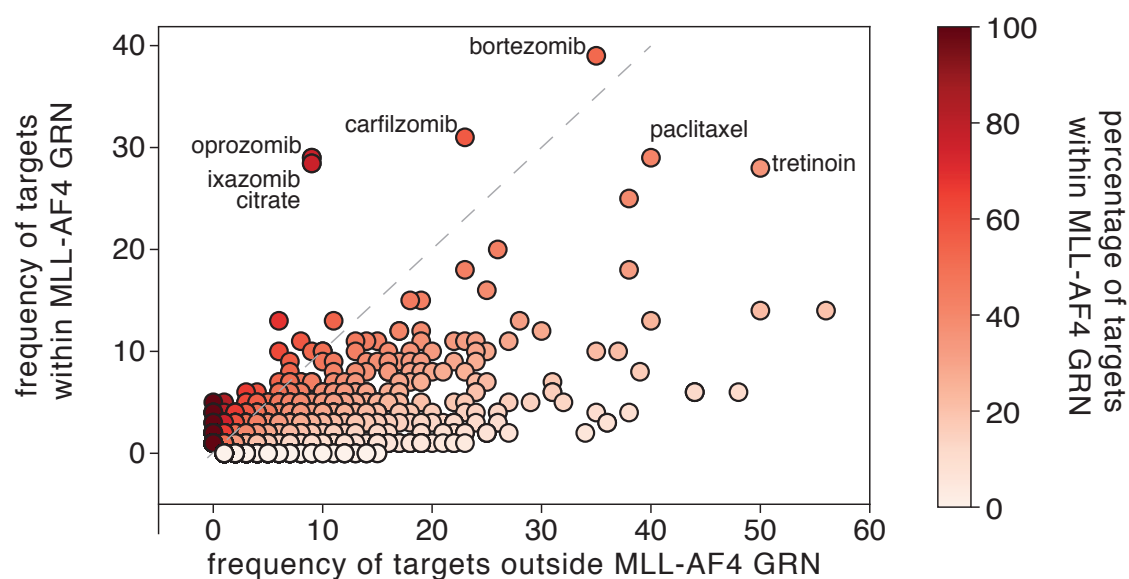


Figure 5.7: Functional analysis of screening hits. (A) Druggability of genes associated with MLL-AF4 leukemia. The MLL-AF4 gene regulatory network (GRN) consists of genes affected by MLL-AF4 knockdown. The Drug-Gene Interaction Database (DGIdb) lists drugs and the genes they target. 6277 drugs target at least one gene within the MLL-AF4 GRN. Dotted line represents drugs with 50% of gene targets within MLL-AF4 GRN. Network analysis provided by Joe Harman.

leukemias are sensitive to VEN and I reasoned that the core VEN sensitivity program may be maintained between different MLLr subtypes²⁰⁰. For each of >6000 small molecules in DGIdb, we compared the number of validated target genes within and outside the MLL-AF4 GRN, reasoning that drugs with greater specificity for the MLL-AF4 GRN would be safer and more effective (Figure 5.7). Four drugs had >25 targets within the MLL-AF4 GRN and >50% specificity for the MLL-AF4 GRN. All four were proteasome inhibitors: bortezomib, carfilzomib, oprozomib and ixazomib citrate. According to DGIdb, one of these drugs, bortezomib (BTZ, Velcade), targeted all four 26S proteasome hits and PSMC3IP. This strongly supported the prospect of a BTZ-VEN combination therapy for VEN-resistant leukemia.

Finally, I used drug combination assays in THP-1 cells to test for synergy between BTZ and VEN. I hypothesised that if BTZ and VEN acted independently, combinations of these two drugs should have an additive effect; if BTZ sensitised cells to VEN, drug combinations should be synergistic. As a positive control for synergy, I used A-1210477, an investigational small-molecule MCL-1 inhibitor known

to synergise with VEN in THP-1 cells³³¹. Firstly I determined the IC_{50} of A-1210477 and BTZ in THP-1 cells: these values matched published data(Figure 5.8A-B)^{327,360}. I treated THP-1 cells with varying combinations of A-1210477 and VEN for 48 hours, then measured cell viability with the CellTiter-Glo assay (Figure 5.8C). Cells treated with VEN–BTZ exhibited more cell death at low concentrations of VEN than the VEN–VEN negative control, but not as much as VEN–A-1210477. I calculated the zero interaction potency (ZIP) score of each drug combination using SynergyFinder to reveal that the VEN–BTZ combination had moderate synergy in THP-1 cells, with synergy score greater than VEN–VEN but less than VEN–A-1210477 (Figure 5.8D)²⁶¹. This indicated that a proteasome inhibitor partially relieved VEN resistance in THP-1 cells, as predicted by the pooled screen.

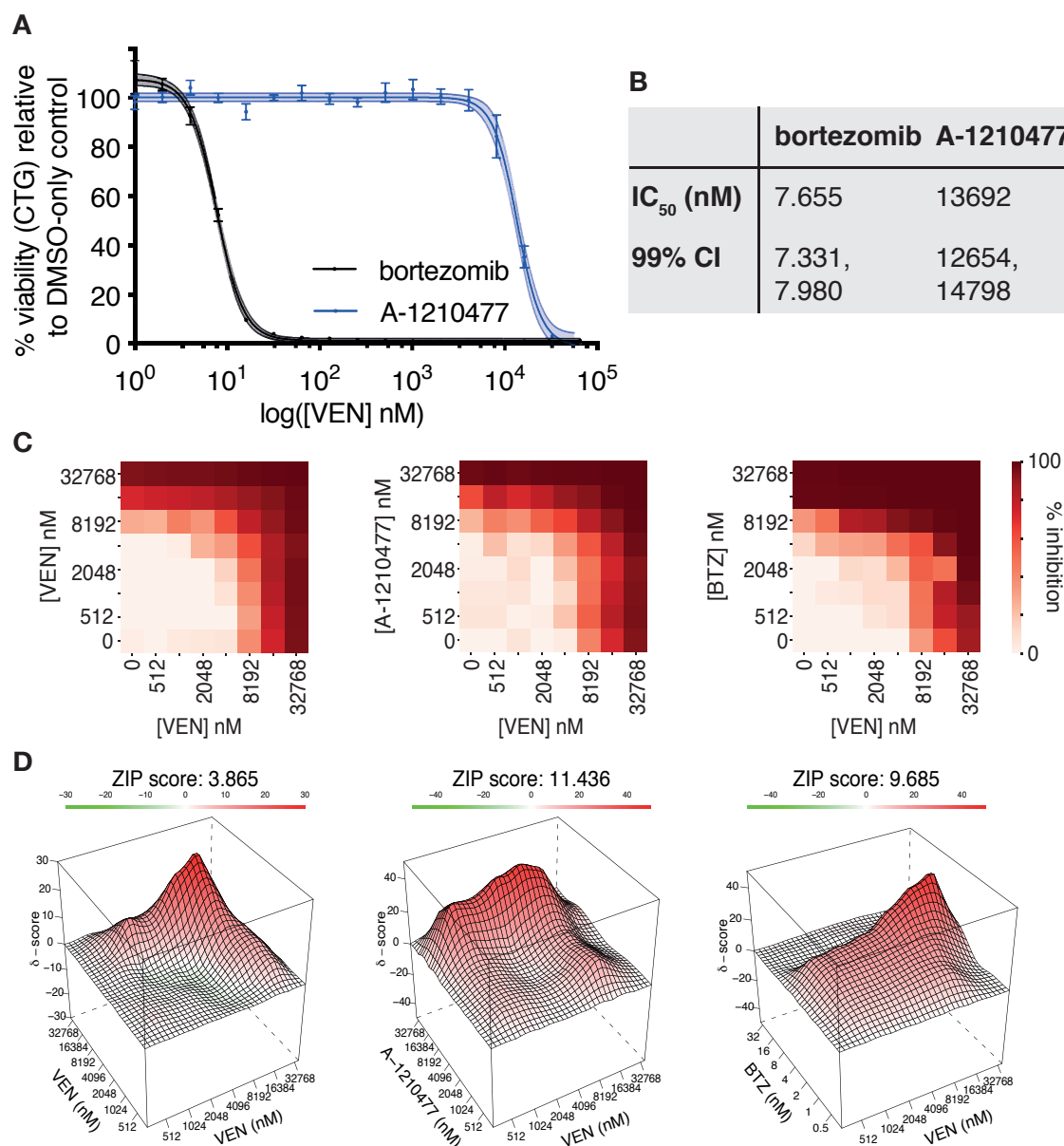


Figure 5.8: Venetoclax-bortezomib combination assays. (A-B) IC_{50} of BTZ and A-1210477 in THP-1 cells. Dots show mean; error bars show SD. Lines show unconstrained four-parameter sigmoidal least-squares regression best fit; shaded region shows 99% confidence interval. For both curves, $R^2 > 0.99$. $n=3$ biological replicates. (C) Drug combination heat maps. Cells were treated with drug combinations (one per square) for 48 hours and analysed with the CellTiter-Glo assay. $n=3$ biological replicates. (D) Drug combination synergy surface plots calculated from heat maps in (C). δ -score: synergy score for a given drug combination. ZIP score: zero interaction potency score. Note different scales in subplots.

5.3 Discussion

In this chapter I show that:

1. Genes outside the canonical intrinsic apoptosis pathway are involved in the response to VEN in THP-1 cells.
2. Resistance-causing genes are associated with cell-cycle regulation; sensitising genes are associated with apoptotic signalling and mitochondrial processes.
3. Knockout of genes encoding proteasome components sensitises THP-1 cells to VEN.
4. Proteasome inhibitor drugs have relatively high specificity for genes involved in MLL-rearranged leukemia.
5. Bortezomib is a small-molecule inhibitor which synergises with VEN in THP-1 cells.

5.3.1 Maintaining sufficient coverage is essential in pooled CRISPR screens

VEN holds great promise for targeted treatment of leukemia but many patients with MLL rearrangements are resistant. Using MLL-AF9 THP-1 cells as a model for VEN resistance, I conducted a TKOv3 genome-wide CRISPR screen with two biological replicates. Cells were cultured for six passages (18 days) in control media (DMSO) or with drug selection (VEN). Comparing read counts of individual sgRNAs between replicates, I found that the day 0 and day 18 VEN replicates were concordant but not the DMSO replicates (Figure 5.3). The observed pattern of random dropout was consistent with a loss of screening coverage. I reasoned that this arose because the mean minimum coverage in the DMSO arm was much lower than the VEN arm, even though DMSO coverage (361 cells/sgRNA) was far above the recommended coverage of 200 cells/sgRNA for the TKOv3 library²⁷⁴. Other protocols, however, recommend a coverage of 500 cells/sgRNA²⁴³.

Since the VEN replicates were concordant and internal controls performed as expected, I set aside the DMSO condition and instead compared VEN to T0 using MAGeCK-RRA (Figure 5.4). Although this strategy is unlikely to have affected the statistical significance of hits, it created the additional challenge of removing essential genes from the dropout hits. I compiled a conservative set of panessential genes by combining three published essentiality screens. None of these screens included THP-1 cells so it is possible that I falsely excluded some true hits from the data. The Yusa essentiality study demonstrated that essential genes may be highly cell type-specific²⁴⁹. All essentiality screens are known to have a high rate of false positives³⁵⁷, as evidenced by the observation that the TKOv3 screen recovered only ~30% of panessential genes.

Given sufficient resources, it would be preferable to repeat the screen at > 1000-fold coverage in both conditions and establish a specific set of essential genes in THP-1 cells. Most published screens do not report concordance of replicates; frequently, published screens have not been replicated at all. Given the differing standards for minimum coverage, it would be interesting to establish how often published screens are affected by similar issues to those described here.

5.3.2 Screen design impacts discovery rates

sgRNA library design is a major barrier to improving the yield and reliability of CRISPR screens. The estimated false negative rate for the Yusa, GeCKOv2, TKOv3 and Brunello libraries is ~20%, suggesting that the sgRNA design rules could be improved³³⁸. Indeed, 87% of target genes are common to all four of these libraries, but only 16% of sgRNA sequences are present in at least two of the four libraries³³⁸. A recent exhaustive comparison of CRISPR-KO libraries suggested that changes to the sgRNA scaffold sequence could also improve library performance³⁶¹. sgRNAs can exhibit off-target activity at sites with more than six mismatches, a fact ignored by all major CRISPR-KO libraries³⁶². It follows that my choice of library could have had a significant effect on screening hits. I favoured the TKOv3 library due to its robust empirical design, but it may have been more reliable to test multiple libraries,

or to apply a less stringent FDR threshold to my hits and then to refine them using a secondary screen and custom library. sgRNA design continues to improve and will lead to more consistent and reliable pooled screens^{289,363}. New sgRNA design tools such as FlashFry should be used for future versions of the TKO library³⁶⁴.

I screened THP-1 cells for six passages at $IC_{50}(\text{VEN})$ ($7.4\ \mu\text{M}$). In a drug screen, it is essential to culture cells for long enough to allow selection to take place through changes in the rate of either cell death or proliferation³⁶⁵. Excessively high drug concentrations may kill cells too quickly or lead to non-specific effects. On the other hand, low drug concentrations may prevent robust selection. One issue is that I determined IC_{50} in a standard 48-hour assay in wild type cells (Figure 4.2). This bears little resemblance to the screen, in which TKOv3-transduced cells were exposed to VEN for multiple passages over 18 days. It may have been more reliable instead to calculate the IC_{50} in a TKOv3-transduced population across multiple passages. Furthermore, selection for VEN-resistant cells meant that the effective IC_{50} of the population increased over time. This may explain why only $\sim 25\%$ of cells were killed by VEN during the final passage of the screen. Thus, it may have been more effective to treat cells at a higher IC_{50} or to use dose escalation. Recent CRISPR drug screens have used drug concentrations ranging from IC_{20} to IC_{70} ^{365,366}. It would also be interesting to repeat the screen in HL-60 cells, in which $IC_{50}(\text{VEN})$ is much lower (Figure 4.2).

The TKOv3 library is specifically optimised for a single-vector format in which the sgRNA and Cas9 are both delivered on lentiCRISPRv2. An alternative strategy is to use Cas9-expressing cell lines. In pilot experiments, I established clonal THP-1 cell lines containing an integrated doxycycline-inducible Cas9 construct but detected a high degree of clonal heterogeneity in Cas9 expression. This presented a risk of confounding the screen so I chose to use lentiCRISPRv2. Nonetheless, a disadvantage of continuous Cas9 expression via lentiCRISPRv2 is that the toxicity of Cas9-induced DSBs has a greater effect. The silent targeting sgRNAs in TKOv3 are, however, an effective control against this effect when combined with copy-number variation (CNV) correction in MAGeCK.

Paracrine effects are another confounding factor²⁴¹. In a highly heterogeneous population, intracellular signalling may rescue a knockout phenotype. This is an intrinsic limitation of pooled CRISPR screening, but may have been more serious in the pooled miniscreen (Figure 4.2), in which the sgRNA pool was very small, than in the genome-wide screen.

5.3.3 Analysis of pooled screens has not been standardised

Previous methods of measuring chemogenetic interactions (based on RNAseq data, for example) depended on high read counts. The low read counts in CRISPR screens make these methods unsuitable. MAGeCK has become the standard analysis tool for pooled CRISPR screens: the majority of recently published screens have used MAGeCK-RRA. This reflects MAGeCK's relative ease of use and analytical advantages.

Despite its popularity, MAGeCK suffers from some known shortcomings. For example, MAGeCK does not implicitly model sgRNA efficiency despite the fact that this can be predicted for each library. Secondly, like EdgeR and DESeq, MAGeCK still uses the negative binomial distribution to model read counts. The negative binomial distribution performs better with RNA-seq data than the Poisson distribution because the negative binomial distribution assumes that the variance is greater than the mean, which is normally true for RNA-seq data where there is a broad range of read lengths^{346,367}. In CRISPR screens, however, most reads are the same length so it may be more appropriate to use a method such as the beta-binomial distribution, which allows variance to be greater or less than the mean³⁶⁷. 'CB²', a recently published pipeline for pooled CRISPR screen analysis which uses the beta-binomial distribution, outperformed MAGeCK-RRA at identifying true hits from public data sets³⁶⁷.

MAGeCK-MLE (maximum likelihood estimate) is a more recent version of MAGeCK which explicitly accounts for differences in sgRNA efficiency and can model

multiple experimental conditions at once. In future screens, I would harvest sgRNAs at several time points and compare them using MAGeCK-MLE instead of -RRA.

Finally, MAGeCK-RRA has the option to account explicitly for gene copy-number variants (CNVs) using reference data, a feature which I implemented in my pipeline. This is essential to avoid false positive hits and a few similar methods have been published^{368–370}. Yet the same correction is possible even in cell lines without reference CNV data because genes with high copy number are clustered into amplified regions which can be identified by mapping gene-level FDRs to the genome. There are three algorithms for this purpose: two (CERES and LDO) use linear mappings^{278,370}; the third uses an unsupervised learning approach³⁷¹. Any of these would improve the versatility of my pipeline.

MAGeCK was designed for essentiality screens rather than chemogenetic interaction screens, although it has been widely used for both. Durocher and coworkers recently developed DrugZ, a Bayesian algorithm specially designed for CRISPR screens with paired samples and drug selection^{365,372}. In direct comparisons, DrugZ returned more dropout hits than MAGeCK at any FDR cut-off value; but more studies are needed to confirm whether DrugZ outperforms other algorithms in conjunction with libraries besides TKOv1 and TKOv3. An effective strategy for overcoming the differences between libraries and analysis pipelines is to run a primary genome-wide screen with a relatively relaxed FDR threshold, and then to refine hits in a subsequent screen with a more stringent cut-off across a broader range of cell types^{241,339}.

BAGEL, a noteworthy alternative to MAGeCK developed alongside the TKO libraries, may be more sensitive than MAGeCK in some contexts but it is less widely used³⁷³. I reanalysed my screen with BAGEL and recovered similar positive and negative hits. Project Score, a large-scale pooled screening effort, uses BAGEL³⁷⁴.

5.3.4 Mitochondrial apoptosis may control VEN resistance in THP-1 cells

In my study, 95% of hits were dropouts, suggesting that there may be more genes in THP-1 cells contributing to VEN resistance than VEN sensitivity. Care should be taken in interpreting gene ontology analysis because many GO terms are redundant and some genes have conflicting annotations³⁷⁵. Due to the small number of hits, I did not consider it reliable to apply pathway analysis. Dropout hits were, nonetheless, associated with a broader range of cellular processes than enriched hits, suggesting that several different mechanisms in THP-1 cells may contribute to VEN resistance (although this may be an artefact of the small number of enriched hits). Surprisingly, *MCL-1* and *BCL-X_L* were only weakly depleted.

Almost all enriched hits were associated with apoptosis or mitochondrial processes. This supports the notion that overall mitochondrial stability – not only mitochondrial apoptosis – mediates the response to VEN in THP-1 cells. Although this hypothesis needs to be tested further, it may provide an important piece in the puzzle of VEN resistance. Perhaps mitochondrial profiling could be used to stratify patients into VEN-resistant and VEN-sensitive groups.

The difference between *BAX* and *BAK*, sometimes considered to have equivalent roles, is striking. *BAX* was the most significantly enriched hit while *BAK* was neutral. This supports growing evidence that *BAX*, which may shuttle between the cytosol and mitochondrial outer membrane, and *BAK*, which is constitutively bound to the mitochondrial membrane, are not functionally redundant^{145,159}. A CRISPR screen using ABT-737 (a BH3-mimetic inhibitor of BCL-2, BCL-X_L and BCL-W) in *Mcl-1* KO mouse embryonic fibroblasts showed that the voltage-dependent MOM channel protein VDAC2 is necessary for *BAX*, but not *BAK*, to initiate MOMP – although the study offered no mechanistic insight³⁷⁶. A difference between *BAX* and *BAK* was also evident in the IC₅₀ assay (Figure 4.7), where *BAX* KO cells were considerably more sensitive to VEN than *BAK* KO cells, although not in the pooled miniscreen (Figure 4.11). This disparity may be due to the greater reliability of using four sgRNAs per gene or a large sgRNA pool.

Hit identification is sensitive to the choice of analysis pipeline. In order to avoid missing important hits, many genome-wide CRISPR screens use a relaxed FDR cut-off for the primary screen and a more stringent cut-off for the secondary screen. Due to practical limitations, I applied a relatively strict FDR cut-off to the primary screen and used GO analysis to identify important hits for individual validation. One disadvantage of this strategy was an increased risk of false negatives, as demonstrated by the analysis in Figure 5.5 which shows several genes of interest such as *CASP3* and *PSMC3IP* which did not meet the FDR threshold despite strong selection of all four sgRNAs.

I used individual sgRNAs from the Brunello library to validate the TKOv3 screen (Figure 5.6). Since it is unlikely that sgRNAs targeting the same gene from different libraries would have the same off-target effects, using sgRNAs from a different library for validation reduced the risk of false positive hits. The validation experiment was shorter, and therefore stricter, than a secondary screen; however, it may have been more reliable to passage cells at least once under VEN selection. *CASP3* and *PSMC3IP*, which were not hits, had a greater effect on viability than some hits, underlining the risk of false negatives due to differences between libraries. Conducting screens with multiple libraries would be an effective strategy to overcome this limitation.

5.3.5 VEN–BTZ combination therapy may overcome VEN resistance in MLLr leukemia

PSMB5, PSMC1, PSMD2, SEM1 and PSMC3IP all appeared to contribute to VEN resistance in THP-1 cells. Independent analysis of the MLL-AF4 GRN showed that four proteasome inhibitor drugs, including bortezomib, are highly specific to targets associated with MLLr leukemia (Figure 5.7).

The proteasome inhibitors are already a promising drug class for the treatment of several hematologic malignancies. The 26S proteasome is a multi-enzyme complex that catalyses the degradation of most intracellular proteins^{377,378}. Misregulation of the proteasome in cancer cells may directly disrupt processes such as the cell cycle,

apoptosis or the DNA damage response; alternatively, cancer cells may depend on the proteasome to degrade the damaged proteins which accumulate as a result of aberrant cellular processes. Early results showed that proteasome inhibition preferentially induces apoptosis in HL-60 cells, in primary CLL lymphocytes and, in a mechanism dependent on NF κ B, in primary AML cells.^{379–381}

Bortezomib (BTZ, Velcade), which was approved for single-agent use in relapsed/refractory multiple myeloma in 2003, is known to affect mitochondrial apoptosis³⁷⁷. In xenograft models, treatment with BTZ overcomes VEN sensitivity in cells expressing BCL-2 and MCL-1 but not BCL-2 and BCL-X_L, irrespective of p53 status³⁸². BTZ was tested *in vitro* in combination with the early BH3 mimetic obataclax over a decade ago¹⁷³. Although the direct mechanism of action is unclear, BTZ is known to inhibit pro-survival NF κ B signalling and promote autophagy³⁸³. It also appears to increase transcription of NOXA, a proapoptotic BH3-only protein which is normally sequestered by MCL-1³⁸⁴. BCL-2/MCL-1 expressing cells are resistant to VEN because MCL-1 can bind propapptotic BIM released from BCL-2 by VEN. If MCL-1 is saturated with NOXA, however, excess BIM leads to BAX/BAK activation and apoptosis. Carfilzomib and the corticosteroid dexamethasone may also be effective combinations with VEN in multiple myeloma because they increase BIM expression^{385,386}.

Early trials for VEN–BTZ combination therapy in R/R multiple myeloma brought promising results³⁸⁷. A phase III trial (BELLINI, NCT02755597) found improved progression-free survival and overall response rate with VEN–BTZ compared to VEN monotherapy but the trial was recently placed on clinical hold due to an increased risk of patient death. BTZ monotherapy has also been the subject of clinical trials in AML³⁸³. A phase III trial of BTZ and the kinase inhibitor sorafenib is ongoing in patients with FLT3/ITD⁺ AML (NCT01371981).

Despite evidence from multiple myeloma and encouraging results for VEN or BTZ monotherapy in MLLr leukemia, there have been no clinical studies of VEN–BTZ combination therapy in MLLr leukemia to date^{4,388}. A drug screen showed that BTZ, carfilzomib and delanzomib were effective in an infant MLLr

ALL model³⁸⁹. However, a phase II trial of BTZ in MLLr pediatric patients recently failed due to a lack of response (NCT02419755).

The genome-wide screen and arrayed KO validation provided evidence that inhibiting each of four components of the 26S proteasome sensitises THP-1 cells to VEN, indicating that a VEN–BTZ combination therapy may be effective in MLLr leukemia. None of these 26S proteasome genes was in the panessential gene set, suggesting that inhibiting them in the absence of VEN does not have a significant effect on cell growth. Taken together, even though the precise relationship between the proteasome and mitochondrial apoptosis pathway in THP-1 cells remains unclear, these results predict synergy between BCL-2 inhibition and proteasome inhibition. If so, the effect of dual inhibition should be greater than the sum of the effects of single inhibition³⁹⁰. This cannot be measured at the genetic level, so I turned to VEN and BTZ to measure inhibition at the protein level. Unfortunately this is an imperfect comparison. Despite its high specificity for the MLLr GRN, BTZ interacts with >70 genes (Figure 5.7) – not only PSMB5, PSMC1, PSMD2 and SEM1. For this reason, individual loss of function of each of these genes had no significant effect on cell proliferation, whereas BTZ treatment is highly cytotoxic. Nonetheless, separate evidence suggests that PSMB5 is among the most critical targets of BTZ in THP-1 cells³⁹¹.

I attempted to separate the overall cytotoxicity of BTZ from its interaction with the BCL-2 pathway by measuring drug synergy. Screening for synergy is experimentally straightforward but analytically complex^{261,390}. There are many competing models for determining drug synergy (and antagonism) but all models attempt to test whether there is a significant difference between the effect of a combination therapy and the effect predicted by a zero-interaction model, which is normally the sum of the individual effects. I used the Zero Interaction Potency (ZIP) model, which combines three older reference models, to show that VEN–BTZ is synergistic in THP-1 cells. Future work should confirm VEN–BTZ synergy in other MLLr cell lines using other protease inhibitors such as oprozomib and carfilzomib.

Instead of using bulk viability as a readout for drug synergy, in future work it may be more reliable to use a single-cell assay such as phospho flow^{392,393}.

A fascinating link between HOXA9/MEIS1 and BTZ hints at a special role for the proteasome in MLLr leukemia⁴. McGinty et al. chemically ubiquitinated histone H2B lysine 120, demonstrating that H2BK120ub recruits DOT1L and promotes DOT1L-mediated H3K79 methylation³⁹⁴ (see also Figure 1.1). The same histone mark promotes H3K4 methylation³⁹⁵. BTZ treatment reduces global H2B ubiquitination⁴. Wang et al. screened a panel of 13 known H2A/B ubiquitinases for their effect on tumour proliferation in an MLL-AF9 mouse model³⁹⁶. They identified the H2BK120 monoubiquitinase RNF20 and showed that RNF20 knockdown reduced proliferation of MOLM-13, THP-1 and MV4-11 cells. RNF20 knockdown specifically reduced H2BK120ub, H3K79me2 and H3K4me3 at *HOXA9* and *MEIS1* loci and reduced the expression of these genes. Thus, it is plausible that inhibition of RNF20 by BTZ prevents MLL fusion protein-mediated upregulation of *HOXA9* and *MEIS1*. If so, it is unclear why *RNF20* did not drop out in the screen. Nonetheless, this is unlikely to be the only effect of BTZ. In a separate study, Liu et al. reported that BTZ treatment led to accumulation of both MLL-AF9 and BID in THP-1 cells, promoting apoptosis; meanwhile in MLL-AF4 cell lines and MLL-AF4 mouse models, BTZ protects MLL-AF4, promoting overexpression of p27 which causes cell-cycle arrest³⁶⁰. In summary, BTZ may synergise with VEN via the anti-apoptotic pathway and also act on separate cancer pathways.

5.3.6 Dropout hits suggest other potential combination therapies for AML/ALL

Several of the dropout hits correspond to known mechanisms of VEN resistance in clinical treatment of AML/ALL. For example, CDK9 overexpression in AML is known to increase expression of XIAP (X-linked inhibitor of apoptosis)³⁹⁷. Both *CDK9* and *XIAP* were dropout hits in the genome-wide screen, although I validated only *XIAP*. CDK9 is one of the elongation-assisting proteins and MLLr cell lines are known to be sensitive to CDK9 inhibition⁵⁴. Several CDK9 inhibitors are being

trialled in AML/ALL and these have been shown to reduce MCL-1 and XIAP expression in cell lines and primary cells. Similarly, pharmacological inhibition of the Ras-homologous GTPase Rac1 increases apoptosis by downregulating XIAP in THP-1 cells (MLL-AF9) but not HL-60 cells (wild-type MLL)³⁹⁸. Rac1 may protect XIAP from proteasomal degradation but it is unclear whether there is an MLL fusion protein-specific mechanism.

Another strong dropout hit, MAPK1 (ERK2), cooperates with ERK1 and the PI3K/mTOR pathway to activate MCL-1 and directly inhibit BIM expression³⁹⁹. Triple combination therapy with (1) VEN, (2) an ERK inhibitor and (3) a dual PI3K/mTOR inhibitor is highly effective at promoting intrinsic apoptosis in THP-1 cells and primary AML patient samples³⁹⁹.

HOXA9 – the *HOX* gene linked to proliferation in MLLr leukemia – is a predictor of VEN sensitivity, especially in MLLr AML/ALL²⁰⁹. This may be because *HOXA9* increases BCL-2 expression. *HOXA9* was weakly enriched in the screen. Two other *HOX* genes which predict VEN sensitivity of AML patient cells, *HOXB7* and *HOXB8*, were more strongly enriched²¹⁰.

The proteasome components were not the only druggable targets suggested by the screen. Two dropout hits which merit further attention are the nuclear exporter XPO1 and the Nedd8 activating enzyme NAE1. The orally bioavailable XPO1 inhibitor KPT-330 (Selinexor, approved for myeloma), which reduces MCL-1 expression, has been demonstrated to synergise with VEN in primary AML samples, including one sample with MLL partial tandem duplication⁴⁰⁰. The NAE1 inhibitor pevonedistat leads to increased expression of NOXA and is effective as monotherapy or in combination with VEN in AML^{401,402}.

5.3.7 CRISPR screens offer new perspectives on VEN sensitivity

Two companion studies describing genome-wide CRISPR screens with VEN selection were published soon after the conclusion of the work described in this chapter. Both emphasise the role of mitochondria in VEN resistance. One study, by

Nechiporuk et al.⁴⁰³, focuses on sensitising genes; the other, by Chen et al.⁴⁰⁴, on resistance-causing genes.

Nechiporuk et al. conducted two separate enrichment screens in MOLM-13 cells using the Yusa and Brunello libraries. They selected with 1 μ M VEN (the IC_{50} in MOLM-13 cells). In the Kosuke screen, they harvested gDNA on both day 7 and day 14; in the Brunello screen, they harvested gDNA on day 14 only. Similar to the pipeline described in this chapter, they used MAGeCK-RRA, although their data are incomplete and they report implausible false discovery rates for both screens. This may be because they did not perform sufficient permutations in their α -RRA analysis. Comparing their results showed that gene ranks and P-values were not concordant between the Kosuke and Brunello screens, apart from key positively selected genes such as *BAX*. Similarly, neither the Kosuke nor the Brunello screen closely matched my own results.

Nechiporuk et al. validated hits individually using new sgRNA sequences. Focusing on p53 as a positively selected hit in their screen, Nechiporuk et al. observed in patient data sets that low or absent *TP53* expression correlated with VEN resistance. They showed that *TP53* mutations in AML cells suppressed the mitochondrial stress response and reduced mitochondrial respiration, but they were unable to detect a corresponding increase in anaerobic respiration in the cytosol. Their study offers an interesting perspective the role of p53 in VEN resistance, a question which my screen could not address since THP-1 cells, unlike MOLM-13, are *TP53*^{-/-}. But as discussed in the next chapter (subsection 6.2.3), there are experimental issues associated with conducting CRISPR-KO screens in *TP53*^{+/+} cell lines.

Chen et al. reported a dropout/enrichment screen in MOLM-13 cells using the Brunello library, with VEN selection escalating from 10 nM to 2 μ M. They harvested gDNA on day 8 and day 16 and analysed the screen by simply averaging the fold change in sgRNA read counts for each gene – a strategy known to be unreliable²⁴¹. Once again, the most significant enriched hit was *BAX*; p53 was also enriched. In validation experiments, sgRNAs targeting p53 increased IC_{50} (VEN) in MOLM-13

cells but not in KASUMI-1 cells (which, like THP-1 cells, are $TP53^{-/-}$). Focusing on mitochondrial proteins in their depleted hits, Chen et al. demonstrated that VEN treatment induces changes in mitochondrial ultrastructure and outer membrane potential. They pinpointed OPA1, a structural protein found in mitochondrial cristae, and CLPB, a mitochondrial ATPase chaperone. In a mouse model, CLPB appeared to increase sensitivity specifically to VEN. Moreover, they isolated CLPB-deficient mitochondria from THP-1 cells and used a cell-free assay to demonstrate greater propensity for cytochrome *c* release and apoptosis. Nonetheless, it is unclear whether this result applies *in vivo* and, if so, why CLPB was not a significant hit in my TKOv3 screen. Currently there are no human CLPB inhibitors.

5.3.8 A high-throughput pipeline for combination therapy discovery

The TKOv3 screen demonstrates the enormous potential of CRISPR screening for discovering combination therapies in oncology and beyond. Better libraries, more reliable analysis platforms and higher throughput will lead to richer discoveries. CRISPR inhibition/activation could also allow for more clinically relevant screens.

The genome-wide screen presented here provides strong justification for exploring the clinical potential of VEN–BTZ combination therapy, as well as providing proof of concept for a powerful discovery pipeline designed for repurposing existing drugs to find effective combination therapies. Mitochondrial profiling and imaging will shed further light on the mechanism for synergy between VEN and BTZ. The screen revealed other potential therapeutic targets such as XIAP and XPO1: future studies should focus on these targets as well.

6

Discussion and future directions

Contents

6.1	A model for non-coding RNA and apoptosis in MLLr leukemia	142
6.2	Pipelines for high-content drug resistance screening in primary cells	151
6.3	Concluding remarks	160

The research described in this thesis had two central goals: (1) to investigate the role of miR-196b and (2) to explore mechanisms for VEN resistance in MLLr leukemia. I accomplished this using versatile CRISPR screening technologies. This strategy has the potential to be applied to other models of treatment resistance in the future.

This thesis makes two main contributions to the field. Firstly, I propose that the putative oncomir miR-196b affects colony-forming capacity in an MLL-AF9 cell line, but does not appear to regulate the expression of *HOXA9*, *MEIS1* or *FAS*, three of its predicted targets. Secondly, I conduct a genome-wide forward genetic screen to study VEN resistance in THP-1 cells. This screen and subsequent *in vitro* assays suggest that many genes beyond the canonical BCL-2 family are involved in VEN resistance and that the proteasome inhibitor BTZ is a promising candidate for VEN combination therapy in MLLr leukemia.

In this concluding chapter, I discuss the overall limitations and implications of the work in this thesis. I attempt to synthesise the results presented in the preceding chapters and I highlight several outstanding scientific questions. I conclude by discussing future pipelines for elucidating the causes and remedies of drug resistance in acute leukemia.

6.1 A model for non-coding RNA and apoptosis in MLLr leukemia

6.1.1 ncRNA in MLLr leukemia: driver or passenger?

In chapter 3, I use dual-sgRNA CRISPR-KO to delete miR-196b and perform a simple reverse genetic screen in THP-1 cells. Dual-sgRNA CRISPR-KO is effective but the technique has some limitations. Deleting the entire mature miRNA sequence probably ensures that the miRNA will not be expressed, although NGS would be necessary to prove this. On the other hand, there is a risk that deleting the entire miRNA could disrupt other functional genetic elements, such as miRNAs in the same cluster or long non-coding RNAs. This risk is even greater for intronic, rather than intergenic, miRNAs. Improved sgRNA design, Cas9 variants with less off-target activity or base editors could all enable more specific miRNA inactivation^{405,406} and avoid the very large deletions that sometimes ensue from Cas9-induced DSBs⁴⁰⁷.

Although my experiments cast doubt on the Chen model for the role of miR-196b in MLL fusion-mediated leukemogenesis, I was unable to propose an alternative model. Ultimately, falsifying the Chen model in THP-1 cells raises more questions than it answers. Many lines of evidence suggest that HOXA9 and miR-196b each contribute to leukemogenesis^{71,84,124}. They are genomically linked and act cooperatively during embryogenesis¹²⁶. Their expression is correlated in the presence of MLL fusion proteins¹²¹. *HOXA9* contains predicted miR-196b response elements and miR-196b expression is controlled by other HOX proteins, if not specifically *HOXA9*¹²⁸. Why, then, has it proven so difficult to establish a model for the role of miR-196b (or any other putative oncogenic miRNA) in the natural history

of MLLr leukemias? The screen by Meyer et al. established *CDKN1B* as a key miR-196b target but it is unclear whether this is the only important target of miR-196b and how this relates to *HOXA9*¹²⁴. Perhaps miR-196b and *HOXA9* are both independently affected by MLL fusion proteins but have no functional link in leukemogenesis. A better understanding of the activity of *HOXA9* during normal and malignant hematopoiesis may help elucidate the role of miR-196b.

I cannot rule out the possibility that miR-196b does affect *HOX* gene expression at other stages of HSC differentiation or in cells containing MLL fusion proteins other than MLL-AF9. Moreover, it is crucial not to overlook the differences between MLL fusion proteins. These are reflected in clinical outcomes: the Interfant-06 trial suggested that infants with MLL-AF9 ALL may have better survival than those with MLL-AF4 or MLL-ENL, although the result was not significant¹³. Schotte et al. found that expression of miR-196b was 500-fold greater in MLLr AML than non-MLLr ALL but miR-196b expression did not correlate with MLL fusion type. This casts doubt on the existence of any specific differences between MLL fusion proteins with respect to miR-196b – or it suggests that there is no direct causal link between miR-196b and MLL fusion proteins⁴⁰⁸.

At the root of this issue is the question whether miRNAs – and, more generally, non-coding RNAs – have a causative role in the initiation or maintenance of leukemia. There is much clinical interest in sequencing the dark transcriptome to stratify leukemia patients. Although the expression of many miRNAs correlates with MLL fusion proteins, the targets of these miRNAs are not known. It seems most likely that each of these miRNAs has manifold subtle effects on a multitude of interconnected targets, such that reverse genetic screens targeting a single miRNA and the measurement of gross phenotypic effects are not sufficiently sensitive¹⁰⁹. If miRNAs do have a driving role in MLLr leukemia, their effects may be combinatorial and they may vary considerably for each MLL fusion protein. Improved ncRNA transcriptomics should allow miRNAs such as miR-196b to be integrated into bioinformatic models of the regulatory networks downstream of MLL fusions⁴⁰⁹.

In summary, many ncRNA species are highly conserved and their expression patterns change consistently during hematopoiesis. Klusmann and coworkers recapitulated the entire hematopoietic hierarchy using t-distributed stochastic neighbour embedding on ncRNA⁴¹⁰. Beyond their potential prognostic value, miR-196b and related miRNAs could be tractable targets for antisense therapy. It remains to be determined whether miRNAs are drivers or passengers in the etiology of MLLr leukemia.

6.1.2 A proposed pipeline for functional identification of oncogenic miRNAs

In the case of MLLr leukemia, many miRNAs other than miR-196b have been functionally linked to leukemogenesis (see subsection 3.3.5). I propose an experimental workflow for discovering these miRNAs and dissecting their function (Figure 6.1). The first step in building a comprehensive model for miRNA-mediated control of leukemogenesis would be to perform an exhaustive search for relevant miRNAs using CRISPR screening (Figure 6.1A). ~6% of sgRNAs in the GeCKOv2 library target miRNAs²⁵⁸ and a single-sgRNA library targeting all known human miRNAs has been described⁴¹¹ but, as discussed in chapter 3, single-sgRNA CRISPR-KO of miRNAs may not be effective depending on the location of the PAM within the miRNA stem-loop. Alternatively, Shirley Liu and coworkers developed a system for paired sgRNA screening in which each lentiviral vector expresses two sgRNAs under the control of separate U6 promoters⁴¹². Each pair contains a unique sgRNA to act as a barcode. Surprisingly, the rate of recombination during viral transduction was relatively low. The same technique could be used to extend the dual-sgRNA miRNA deletion strategy developed in chapter 3 into a pooled screen. Bioinformatic pipelines such as CRISPETa are available for the design of optimised sgRNA pairs⁴¹³.

It is still difficult to validate the predicted targets of a miRNA. For example, it is possible that miR-196b exerts its effects via unknown target genes. The miR-196b pulldown screen recovered the sodium transporter *SLC9A6* which, like *HOXC8* and *HOXA7*, contains four miR-196b target sites in the 3' UTR¹²⁴. As a

second step in the proposed pipeline, I would implement MRE-CRISPR (described in subsection 3.3.3) to survey all known targets of miR-196b (Figure 6.1B) and to understand whether these change in response to differences in MLL fusion or *HOX* gene expression. Bioinformatic packages combining miRNA target prediction and tissue-specific expression data are already available⁴¹⁴, as are packages for network prediction based on CRISPR screening data⁴¹⁵. Work is ongoing in the Milne group to establish the gene regulatory network downstream of MLL fusion proteins. It would be highly instructive to include ncRNA regulatory elements in this network. By combining results for MRE-CRISPR screens on multiple miRNAs with transcriptomic data for top predicted targets, it may be possible to establish a miRNA expression atlas for MLLr leukemogenesis (Figure 6.1C), building on the work of Klusmann and coworkers which determined an MLLr-specific ncRNA signature⁴¹⁰.

6.1.3 Current models for apoptosis in cancer are inadequate

In chapter 4, I use CRISPR-KO to test the effects of specific BCL-2 family proteins on proliferation and VEN resistance. The results of these experiments largely recapitulated previous results about the contribution of BCL-2 family proteins to BH3 mimetic resistance. In particular, they confirmed that MCL-1 and BCL-X_L mediate VEN resistance in MLL-AF9 cells, presumably by sequestering BH3-only proteins that are released when VEN binds to BCL-2. MLLr cells are relatively sensitive to VEN because DOT1L-mediated H3K79 methylation increases BCL-2 expression²⁰⁰, but the results in chapter 4 suggest that MCL-1 and BCL-X_L can still drive resistance. A major issue with CRISPR-KO is that the extent of gene knockout in a mixed population depends on sgRNA activity. Without generating single-cell edited clones, it is difficult to interpret the relative effect of each BCL-2 family protein on VEN resistance from my data. This limits the mechanistic insight that I can gain from these experiments. It would have been interesting to include the other anti-apoptotic proteins, BCL-1 and BFL-1, in these assays. Indeed, a

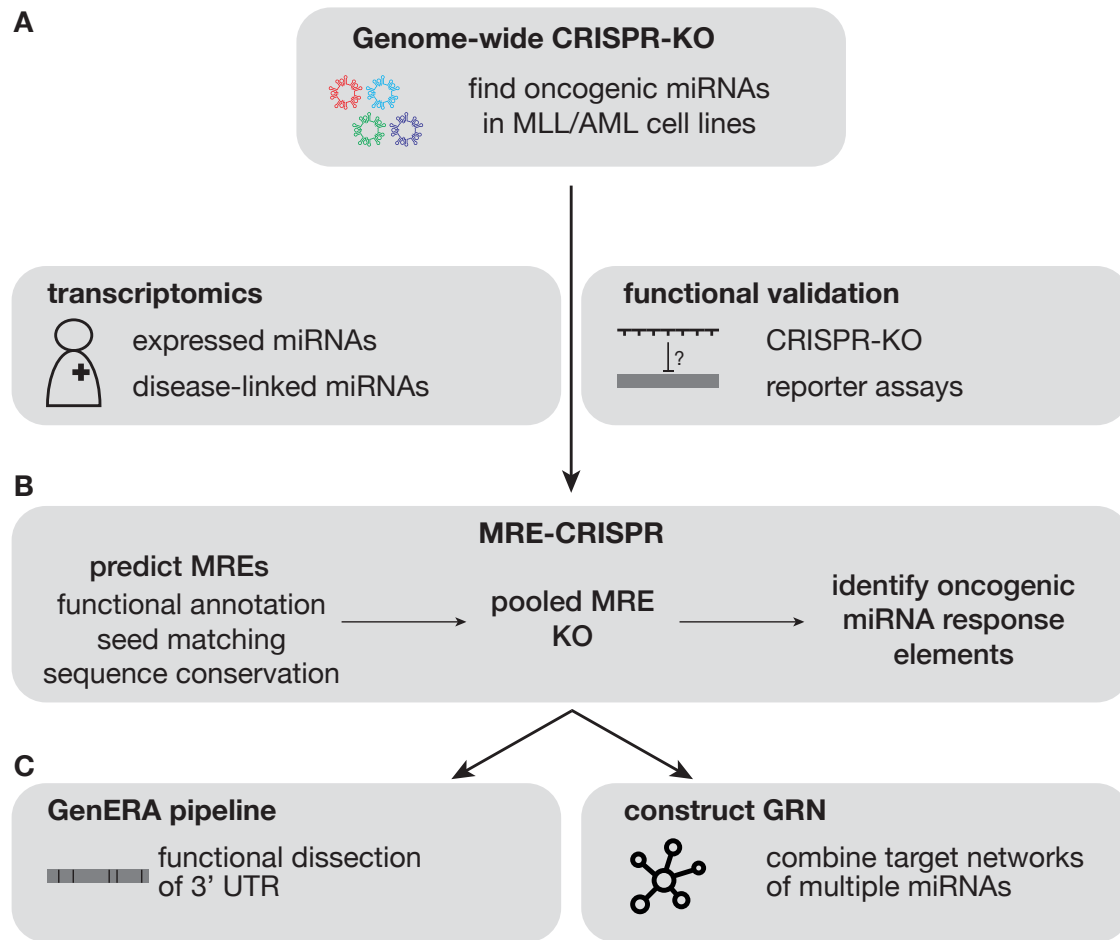


Figure 6.1: A proposed pipeline for oncomir discovery. (A) Pooled CRISPR-KO screening in MLLr cell lines would be used to test all known miRNAs for their effect on viability, as determined by annexin-V immunomagnetic sorting. Candidate miRNAs would be validated by transcriptomic association studies in patient samples and functional assays in primary cells. (B) Next, the MRE-CRISPR pipeline (details in text) would be applied to top candidate miRNAs in order to identify their full oncogenic target networks. (C) Finally, the 3' UTR of top target genes could be dissected using the GenERA pipeline; in parallel, the target networks of multiple oncomirs could be integrated into existing MLLr GRNs.

recent CRISPR screen with combined MCL-1 and BCL-X_L inhibition emphasised the importance of BFL-1 and BCL-w⁴¹⁶.

The pooled screen described in chapter 5 demonstrated that most of the genes controlling VEN resistance are involved in mitochondrial processes or cell-cycle regulation. Since the mechanism of action of VEN is targeted at intrinsic apoptosis, the screen indicated more broadly that resistance to intrinsic apoptosis in MLL-AF9 leukemia depends on mitochondrial stability and cell-cycle signalling. This was

later corroborated by two independent VEN resistance screens^{403,404}. It follows that better models of mitochondrial function in MLLr-transformed cells could lead to new treatments for MLLr leukemia.

Cancer cells are addicted to high proteasome activity, making them highly susceptible to protease inhibitor treatment^{377,378,417,418}. The mechanism of action of the protease inhibitor bortezomib (BTZ) in CLL and AML/ALL is still poorly understood³⁸³. As discussed in subsection 5.3.5, there is tentative evidence that the ubiquitin-dependent 26S proteasome engages in specific regulatory interactions with *HOXA9/MEIS1* and the mitochondrial apoptosis pathway^{360,394}. Both of these interactions need to be understood more fully in order to develop better synergistic combination therapies. In an effective combination therapy, the drugs' efficacy is synergistic but their side effects are additive. BTZ has a very broad spectrum of targets. It would be interesting to target the key resistance-causing proteasome genes (PSMB5, PSMC1, PSMD2 and SEM1) using more specific inhibitors such as antisense oligos. Performing these experiments in HCT116 Octa-KO cells (which lack eight BH3-only proteins²¹²) would eliminate the effects of any direct inhibition of BCL-2 family signalling.

Advances in pro-apoptotic therapy for MLLr leukemia may require a better understanding of mitochondrial dynamics during HSC differentiation. For example, post-translational modifications of BAX – predictably the strongest enriched hit in all three genome-wide VEN screens – may change over time or in response to different MLL fusion proteins¹⁴⁵. VEN sensitivity is known to vary with differentiation state in primary AML cells⁴¹⁹. One outstanding issue is whether mitochondria in malignant cells are specially primed to respond to BH3 mimetics. Khaw et al. reported that healthy B cells and cells from patients with CLL respond similarly to VEN (although healthy B-cell precursors are spared). This contradicts earlier work by Letai, Armstrong and coworkers who suggested that malignant cells are primed for apoptosis^{420,421}. Interestingly, Chen et al. concluded that VEN-resistant AML cells are not primed for apoptosis until knockout of the mitochondrial structural chaperone CLPB⁴⁰⁴. Any effect is likely to differ between cell lines and primary

cells, highlighting the importance of using primary cells to study mitochondrial dynamics. Quantitative modelling of BCL-2 family proteins via dynamic BH3 profiling of patient samples may shed some light on this issue^{145,422}.

6.1.4 New approaches to treating MLLr leukemia

As discussed in this thesis, there are many possible interactions between the signalling pathways associated with MLLr fusion proteins, tumour proliferation and apoptosis. Many questions remain unanswered about the molecular biology and natural history of MLLr leukemia. For example, a GWAS study discovered consistent frameshift insertion mutations in MLL3 (*KMT2C*) which were strongly associated with infant MLLr leukemia⁴²⁵. MLL3 is a homolog of MLL and clearly warrants further study. In another example, a recent preprint suggested that the oncogenic miR-17~92 cluster contributes to survival of HSCs and committed progenitors in a BIM-dependent manner^{426,427}. Interestingly, this matches an observation in a separate study that treating THP-1 cells with a combination of VEN and the CDK9 inhibitor alvocidib leads to a decrease in expression of miR-17~92 miRNAs, although the authors do not propose a mechanism to explain this observation⁴²⁸.

It remains a significant challenge to draw all these threads together into a synthesised model. How important is the HOXA9/MEIS1/FAS/miR-196b axis for intrinsic apoptosis in MLLr cells? Are there several links between these pathways? If so, how many? Which are the most important? These questions are summarised in Figure 6.2.

The pooled CRISPR screen suggested that BTZ, already a successful monotherapy and combination therapy for CLL, may synergise with VEN in MLL-resistant MLL-AF9 leukemia. Much work is needed to validate this result in other cell lines and *in vivo*, but this result offers exciting insight into the role of the proteasome in MLLr leukemia and proves the power of CRISPR screening for discovering drug combinations. As revealed by the analysis of DGIdb data in chapter 5, BTZ has a very broad spectrum of targets. The more specific protease inhibitors carfilzomib and ixazomib are currently undergoing phase I/II trials in R/R pediatric ALL⁴²⁹.

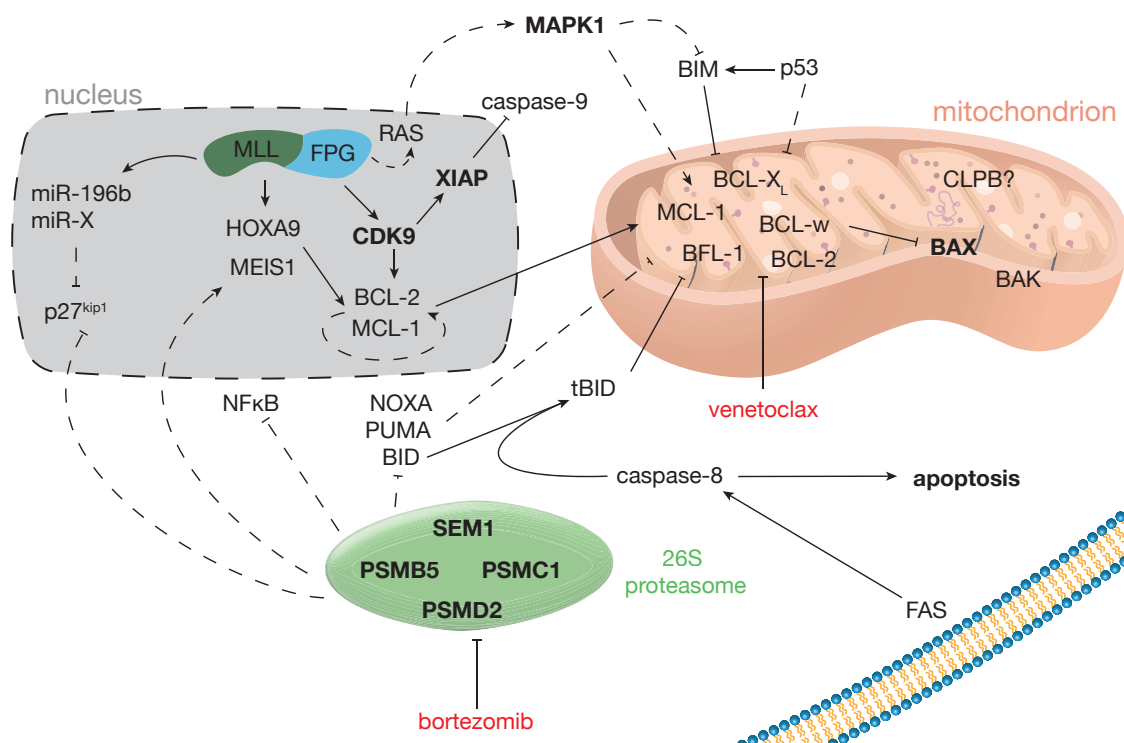


Figure 6.2: MLL fusion proteins and apoptosis. Summary of interactions between MLL fusion proteins, HOXA9/MEIS1/miR-196b and the BCL-2 family. *Solid arrows: experimentally verified interactions. Dashed arrows: indirect interactions. Proteins in bold: validated hits in the genome-wide VEN screen. See main text for details and full references.* Proteins in the nucleus, MOM, cytosol and plasma membrane interact to regulate the response to VEN. MLL-FPG (fusion partner gene) fusions promote expression of the transcription factors HOXA9 and MEIS1²⁷. These upregulate the anti-apoptotic factors BCL-2 and MCL-1, which may also promote their own transcription¹⁶⁰. MLL fusion proteins are associated with overexpression of the RNA pol II activator CDK9, which promotes expression of BCL-2 and MCL-1³⁹⁷. CDK9 also promotes expression of XIAP which inhibits caspase-9. Thirdly, MLL fusion proteins upregulate miR-196b and other miRNAs (miR-X). miR-196b represses the transcription factor p27^{kip1} [124]. MLL fusion proteins may directly upregulate *RAS* isoforms⁴²³; independent gain-of-function mutations in *RAS* are also common⁴²⁴. *RAS* activation upregulates MAPK1 (*ERK2*); MAPK1 may repress MCL-1 and increase BIM expression³⁹⁹. p53 may promote inhibitory post-translational modification of anti-apoptotic proteins²¹². 26S proteasome inhibition indirectly leads to loss of MEIS1⁷³ and accumulation of p27^{kip1} [360], NOXA, PUMA and BID³⁶⁰ and NFkB³⁸³. This may explain why proteasome inhibition synergises with VEN treatment. BID protein is cleaved by caspase-8 in the cytosol to generate active tBID¹⁵⁷. Structural proteins such as CLPB appear to increase resistance to apoptosis⁴⁰⁴ but their interaction partners are unknown.

One of two trials with ixazomib is testing an oral formulation. If ixazomib proves to be effective, it is tempting to envisage an oral VEN/ixazomib combination therapy for pediatric ALL.

Although clinical outcomes for infant ALL have not improved since 1999¹³, other advances in cancer therapy are changing the treatment landscape. Immunotherapies such as blinatumomab⁴³⁰ are likely to be added to future treatment schedules, as are epigenetic drugs such as the DNA methylation inhibitor azacitidine⁴. It will be fascinating to track pivotal trials of VEN in combination with hypomethylating agents. A limited phase Ib study posted encouraging results²⁰⁶.

A recent study showed that combinations of the topoisomerase II inhibitor daunorubicin and inhibitors of MCL-1 or BCL-2 were highly effective in murine MLL-AF9 cells overexpressing *Mcl-1* or *Bcl-2*⁴³¹. In THP-1 cells, the triple combination of daunorubicin, VEN and the investigational BH3-mimetic MCL-1 inhibitor S63845 was highly synergistic. It will be interesting to note whether these results can be replicated in an MLLr AML patient xenograft model⁴³¹. There is preclinical evidence that novel Menin inhibitors are effective in mouse MLLr models as monotherapy⁴³² or in combination with DOT1L inhibitors⁴³³, but no clinical trials have been announced. Although multidrug combination therapies increase the risk of adverse events, some, such as R-CHOP in non-Hodgkin lymphoma, have been highly successful. CRISPR screens with combined BCL-2/MCL-1 inhibition (using separate inhibitors or single-agent navitoclax) could be used to discover similar combinations in MLLr leukemia.

Even as the effect of common MLLr fusion proteins on intrinsic apoptosis becomes better understood, there are many other fusion proteins and cytogenetic abnormalities associated with AML and ALL^{124,203}. It is likely that each of these has its own stable transcriptional signature. More comprehensive gene regulatory networks may suggest combination therapies for even the most treatment-resistant patients. Garnett and coworkers recently reported an impressive systematic analysis of gene fusions in cancer⁴³⁴. They detected fusions in RNA-seq data sets, predicted those fusions' effects on drug sensitivity based on the Genomics of Drug Sensitivity in Cancer database, and scored essential fusion partners using Project Score and

DepMap data. They discovered a small fraction of novel fusions and validated their role as oncogenic drivers.

High-throughput techniques other than CRISPR screening are useful for drug discovery in MLLr leukemia. Henderson and coworkers performed a series of phenotypic screens in PER-485, an infant MLL-AF4 ALL cell line, with a compound library of 34 000 small molecules^{435,436}. They discovered two structurally unrelated hits, CCI-006 and CCI-007, which caused selective cytotoxicity in MLLr cells compared to WT MLL, solid tumour and healthy cell lines. CCI-006 synergised with etoposide, cytarabine and the DOT1L inhibitor EPZ-5676 by triggering the unfolded protein response, leading to apoptosis. Compared to CCI-006-resistant MLLr cells, CCI-006-sensitive MLLr and CALM-AF10 cells exhibited differential expression of components of the mitochondrial electron transport chain, greater sensitivity to a glycolytic inhibitor and reduced expression of MEIS1 and the hypoxia-sensing transcription factor HIF-1 α . A separate study reported that MEIS1 suppresses oxidative phosphorylation in MLL-AF9 mouse models, thereby reducing their sensitivity to pro-apoptotic drugs⁴³⁷. This represents an exciting and novel role for MEIS1 in the maintenance of MLLr leukemia. On the other hand, CCI-007, which also induced apoptosis, reduced transcript levels of *HOXA9*, *MEIS1*, *MYC* and *BCL-2*⁴³⁶. The targets of both CCI-006 and CCI-007 remain unknown.

6.2 Pipelines for high-content drug resistance screening in primary cells

Drug sensitivity screening has made huge progress since early low-throughput haploinsufficiency screens in yeast⁴³⁸. Although, in the last few years, CRISPR screening has become a popular and widely adopted technique, there remain many experimental challenges – including some described first-hand in this thesis, such as library representation. In this section, I provide an overview of these challenges and highlight some exciting new developments in CRISPR screening which could be used to study MLLr leukemia.

6.2.1 Pooled CRISPR screening methodologies are constantly improving

CRISPR screening needs to be optimised for each combination of drug and cell type. The optimisation in chapter 4 was essential for determining correct screening conditions for the genome-wide screen in chapter 5. The CRISPR screens reported by Chen et al.⁴⁰⁴ and Nechiporuk et al.⁴⁰³ show that differences in cell type and drug selection, as well as technical differences such as coverage and the choice of sgRNA library, can have major effects on results. In line with the need to improve cell culture techniques and cell-based assays for drug screening (discussed below), researchers should establish a standardised pipeline for CRISPR drug screening so that results from different studies can be more reliably reported and compared.

The screen I performed demonstrates that results are highly sensitive to the analysis platform used. Pipelines such as MAGeCK are still in their infancy. Published CRISPR screens such as those of Nechiporuk et al.⁴⁰³ and Chen et al.⁴⁰⁴ contain statistical inaccuracies. As discussed in chapter 5, improved modelling of sgRNA distributions and more standardised analysis metrics will make CRISPR screening more reliable.

Dual-sgRNA screening technology (subsection 6.1.1) could be adapted to search for synthetic lethal combinations of target genes without the need for drug selection by perturbing two unrelated coding genes in the same cell. This would present a combinatorial challenge: constructing every possible sgRNA pair from the TKOv3 library would require over 5×10^9 different constructs. Thus it would be necessary to screen a custom subset of target genes. sgRNA pairs could be randomly cloned using an established technique⁴³⁹ although very long next-generation sequencing read length would be required to characterise each pair. A better alternative may be to label each pair with a unique barcode and synthesise a single construct containing the barcode, both sgRNAs and their promoters. Using CRISPRa/i may be preferable in order to avoid the risk of large deletions when both sgRNAs in a pair target nearby loci. Bassik and coworkers described an impressive dual-sgRNA screen for drug combinations in K562 cells⁴⁴⁰. They developed a convenient cloning

strategy and expressed one sgRNA under a human U6 promoter and the other under a mouse U6 promoter for a total of ~490 000 unique dual-sgRNA combinations. As an alternative approach, Doench and coworkers have constructed dual-sgRNA libraries with two different Cas9 orthologues in parallel⁴⁴¹.

In all forward genetic screens, there is a balance between experimental throughput and biological accuracy. CRISPR screening in cell lines favours high throughput in a simple model system at the cost of a higher rate of false positives. Hits must then be validated in a more physiologically relevant model in culture or *in vivo*. Some recent studies, however, have opted to run screens in primary cells or *in vivo*. Examples include two important recent screens in primary T cells^{442,443} and a bone marrow delivery platform which enables pooled CRISPR screens of immune cell lineage specification⁴⁴⁴. An important next step for understanding VEN resistance will be to repeat the VEN screen *in vivo*, such as in a recently published humanised MCL-1 mouse which expresses human MCL-1 instead of its mouse homologue⁴⁴⁵.

Perhaps gene knockout is not physiologically relevant in the context of drug resistance. CRISPRi-mediated gene knockdown may be a more realistic representation of drug activity than CRISPR-KO. Although more complex to conduct and analyse, CRISPRa/i screens – especially using combinatorial sgRNAs – may yield more useful results⁴⁴⁶. Doench and coworkers showed that their CRISPRi library Dolcetto matched the performance of the CRISPR-KO library Brunello in an essentiality screen and out-performed previous CRISPRi screens in a vemurafenib selection screen in melanoma cells³⁶¹.

For similar reasons, it may be useful to view siRNA screening and CRISPR screening as complementary, rather than competing, techniques. Partial knockdown of a target may be a better model for the clinical activity of small-molecule inhibitors than complete knockout. One study showed that, for this reason, some hits are uniquely identified by siRNA or CRISPR screens, alongside many hits common to both techniques⁴⁴⁷. Perhaps each technique measures a different kind of synthetic lethality⁴⁴⁸.

The lack of experimental reproducibility is a well-known issue in high-throughput screening, especially in cancer drug discovery^{449,450}. This may explain some of the disagreement between the miniscreen (chapter 4) and the genome-wide screen (chapter 5). Slight differences in cell culture have significant effects on cell metabolism⁴⁵¹. Stocks of the same common cancer cell lines differ between laboratories and exhibit genetic drift during culture, which has measurable effects on drug responses^{324,325}. Even short-term culture leads to widespread transcriptomic changes in freshly isolated stem cells⁴⁵². Westerhoff and colleagues have proposed a method for proactively counteracting metabolic differences between cell lines during *in vitro* assays, although it would be difficult to scale this approach to a high-throughput CRISPR screening programme⁴⁵³. Nonetheless, researchers must work towards standardising cell culture conditions and carefully reporting them in the literature. Novel ‘physiological’ growth media such as Plasmix may also help to suppress the metabolic artefacts of cell culture⁴⁵⁴.

The work described in this thesis and the comparable studies by Chen et al.⁴⁰⁴ and Nechiporuk et al.⁴⁰³ all implicate mitochondrial metabolism in VEN resistance. This striking concordance between the overall conclusions of three separate studies suggests that mitochondrial metabolism is a robust target. This raises an interesting issue. If the metabolic effects of *in vitro* culture are mediated by mitochondria, then perhaps cell culture has particularly strong confounding effects on pro-apoptotic drug resistance models. This could explain the variation in values of $IC_{50}(VEN)$ reported by different groups (Table 4.1), as well as the inconsistencies between the three VEN CRISPR screens. These inconsistencies are probably also due to differences between cell lines, sgRNA libraries and analysis pipelines.

An analysis of two large-scale pharmacogenetic screening studies, the Cancer Genome Project (CGP) and Cancer Cell line Encyclopedia (CCLE), revealed strong consistency within each study but poor agreement between the two studies, due in particular to a lack of standardised measures for drug responses⁴⁵⁵. Revisiting both studies in the light of more recent results showed that media conditions and

cell seeding density were also major confounding factors⁴⁵⁶. CRISPR screening projects should learn from these lessons.

6.2.2 Automation will enable parallel high-throughput CRISPR screening

CRISPR screening is still a highly laborious technique. There are four strategies to combat this:

1. **Better selection.** Viability screens depend on cell-growth effects to select sgRNAs. Unfortunately, such screens are slow because it may take many passages before differences in growth rate lead to measurable changes in the sgRNA distribution. Moreover, a high rate of cell death across many passages demands a very high number of input cells at the start of the screen so that there are sufficient cells at the end of the screen to maintain coverage. Large-scale viral transduction is expensive and inevitably leads to variation in cell culture conditions. As CRISPR screening develops into a mature technology, I predict that selection techniques such as FACS will become more common than viability screening because physical sorting allows faster screens, is more reliable and allows the experimenter to address more specific phenotypes. Of course, FACS is not without its technical and practical limitations. Moffat, Kelley and coworkers have developed a microfluidic platform for immunomagnetic cell sorting which sorted 1×10^8 cells from a TKOv3 screen in ~1 hour, a rate an order of magnitude faster than most FACS runs⁴⁵⁷.
2. **Smaller libraries.** The smaller the number of sgRNAs in a pooled screening library, the smaller the number of cells that must be screened. Screening fewer cells simplifies transduction, cell culture, selection, DNA harvesting and NGS; further, it facilitates screening in cell types where growth is very slow or cell numbers are limiting. Fundamentally, there are two ways to reduce the total number of sgRNAs in a screening library: (1) use fewer sgRNAs per gene,

which will need better sgRNA design or better data analysis; or (2) design custom sgRNA libraries which do not target the whole genome. Custom libraries demand confident bioinformatic prediction of target genes but, from a practical perspective, it is easier and less expensive for most researchers to purchase a genome-wide library than to synthesise and clone a custom library. Cheaper library synthesis and cloning services may make custom libraries more attainable in the future. New and improved genome-wide CRISPR libraries, such as the H1/H2 libraries which contain 19nt sgRNAs⁴⁵⁸, continue to be published. Over time, this may permit the use of fewer than four sgRNAs per gene. Publicly available libraries such as TKOv3 can be customised by using Cas9 to deplete unwanted sgRNAs from the plasmid pool before generating lentivirus, although this is not a common technique⁴⁵⁹.

3. **Lower coverage.** In a standard pooled CRISPR screen, the integrated sgRNA acts as a barcode to represent the knockout gene in each cell. To account for stochastic effects of transduction, mutagenesis and cell culture, screens are run at high coverage. Cells containing the same sgRNA are indistinguishable. Two groups have described techniques for including a unique molecular identifier (UMI) in every construct in the library^{460,461}. Thus even cells transduced with the same sgRNA can be identified, enabling single-cell lineage tracing. Both studies reported superior sensitivity and robustness with the addition of UMIs for a given screening coverage. Over time, this technique should allow screens to be conducted at lower coverage.
4. **Higher MOI.** Recent work has shown that it is possible to reduce cell culture volumes by performing some whole-genome screens at $\text{MOI} \leq 1$ (L. Crepaldi, personal communication). At this MOI, ~20% of cells are transduced by two sgRNAs per cell. Depending on the phenotype being selected, it is unlikely that both genes in a randomly selected pair will affect the phenotype. Alternatively, Zhu et al.⁴⁶² placed UMIs within the sgRNA scaffold sequence and successfully performed screens at $\text{MOI}=3$ or even $\text{MOI}=10$. Future studies

using this technique will need to develop robust analysis pipelines and address concerns about high toxicity due to Cas9 activity.

5. **Automation.** Many aspects of the CRISPR screening workflow presented in chapter 5 will be accelerated and simplified by automation. Automated cell culture systems will make screening faster and more consistent between replicates, especially for screens involving primary stem cells⁴⁶³. For example, Buchser and coworkers are developing a microfluidic platform for automated arrayed CRISPR screening⁴⁶⁴.

6.2.3 p53: a double-edged sword for CRISPR screening and cancer treatment

Some controversy surrounds the role of the tumour suppressor protein p53 (*TP53*) in AML/ALL and in cancer screens. It has long been disputed whether p53 is an artefact of *in vitro* culture or a prerequisite for immortalisation⁴⁶⁵. p53 mutations are common in cancer cell lines, although the p53 status of some NCI-60 cell lines is still unknown⁴⁶⁶. In the clinic, p53 is mutated in roughly half of solid tumours, but far less frequently in hematological malignancies¹⁹⁹. Given that p53 has a known role in apoptosis, it is a high priority to understand the role of p53 in MLLr leukemia.

p53 status may affect the outcome of VEN treatment in CLL patients. One large study of VEN monotherapy for R/R CLL found that p53 mutant patients were associated with poorer outcomes⁴⁶⁷; another found no association⁴⁶⁸. Two *in vitro* studies also concluded that VEN activity in CLL cells was p53-independent^{193,469}. Similarly, in their systematic search for genetic markers of cancer drug sensitivity, Garnett et al. did not find a link between VEN sensitivity and p53 status⁴⁷⁰. This makes sense because BCL-2 inhibitors act downstream of p53 according to the embedded together model (subsection 1.4.2).

In AML, on the other hand, p53 mutation is associated with poorer responses to VEN⁴⁷¹ and poorer outcomes in general⁴⁷². In HCT116 Octa-KO cells, deletion of both *MCL-1* and *BCL-X_L* abrogates the effect of p53 status on VEN sensitivity, suggesting that MCL-1 and BCL-X_L are regulated by p53-dependent signalling,

perhaps in the cytosol²¹². It is tempting to speculate, therefore, that AML/ALL is sensitive to p53 status due to increased dependence on MCL-1 and BCL-X_L. However, BCL-2 and the BH3-only proteins are also targets of p53 signalling, so it is unclear why CLL is unaffected¹⁵⁶. As discussed in subsection 1.4.5, Konopleva, Andreeff and coworkers have successfully trialled a combination of VEN and an MDM2 inhibitor (which leads to p53 activation) in AML patient-derived xenografts.

Besides its clinical implications, wild-type p53 status may be an issue in CRISPR screens. p53-dependent signalling downstream of Cas9-induced DSBs may cause death or cell-cycle arrest regardless of an sgRNA's target. If so, sgRNAs targeting essential and non-essential genes alike would be depleted in p53^{WT} cell lines. It is known that sgRNAs with more off-target activity or those targeting genes with higher copy number are preferentially depleted in all cell lines but there are several methods to correct for this effect, as discussed in chapter 5. The issue with a p53-specific effect is that it would invalidate screening results from p53^{WT} cell lines, not to mention potentially severe implications for therapeutic applications of CRISPR/Cas9. Haapaniemi et al. ran pooled genome-wide CRISPR-KO screens in p53^{WT} and p53 knockout RPE1 cells and found that *TP53* and related tumour suppressor genes *CDKN1A* and *pRB* were more strongly depleted in p53^{WT} cells. This, they argued, was evidence that p53 activity mediates DSB-induced growth arrest in lentiviral CRISPR screens⁴⁷³. Ihry et al. described a similar phenomenon in human pluripotent stem cells and also found p53-dependent changes in *BCL2*, *BAX* and *FAS* transcription following Cas9-induced DSBs⁴⁷⁴. Taken at face value, these results may explain why successful early CRISPR screens all used p53 mutant cell lines, such as the screens in AML lines by Vassiliou, Yusa and coworkers²⁴⁹.

Moffat and coworkers responded robustly in defence of CRISPR screening in p53^{WT} cell lines, observing that depletion of p53 pathway genes is common to screens in both p53^{WT} and p53 mutant cell lines and even to RNAi screens in which there are no DSBs⁴⁷⁵. Meta-analysis of screens from the DepMap project (discussed below) showed no dependence on p53. Moffat et al. also reiterated that it is important to conduct CRISPR screens in multiple cell types. This also applies to the screen that

I conducted: repeating it in other MLLr cell lines would be an important validation step. Nonetheless, this issue should not affect THP-1 cells because they are p53⁻.

Haapaniemi et al. responded in turn⁴⁷⁶. They argued that CRISPR screens are hampered by p53 activity in RPE1 cells and in primary cells because, in these cell lines, Cas9-mediated editing occurs much more slowly than in non-tumour lines, leading to an enhanced effect of p53-dependent transient cell-cycle arrest. They criticised Moffat and coworkers' use of highly optimised cell lines and scoring metrics for CRISPR screens, implying that these can obscure scientific findings.

Perhaps this controversy is less relevant in CRISPR screening, where Cas9-induced DSBs are a discovery tool, compared to the clinic, where p53-dependent cell cycle arrest could prevent successful gene editing or, worse still, be tumourigenic. Base-editor enzymes which do not induce DSBs could supersede Cas9/CRISPR-KO and avoid this issue^{477,478}. Nonetheless, p53 deserves significant attention, both for its implications for CRISPR/Cas9 editing and its central role in integrating cell stress associated with MLLr fusion proteins and intrinsic apoptosis.

6.2.4 The future of drug resistance screening

There has long been interest in using mutational markers to understand or even predict drug sensitivity. CRISPR screening has become a transformative new tool for this effort^{470,479}. The work presented in this thesis demonstrates the power of CRISPR drug screening for identifying combination therapies and investigating drug resistance pathways. Large-scale CRISPR screening projects will increase the rate of leukemia drug discovery (Figure 6.3). As of mid-2019, two such projects – Project Score (Sanger Institute) and the DepMap project (Broad Institute) – had together performed CRISPR essentiality screens in 371 cell lines across 33 cancer types^{374,434}. Other groups are working on new analysis tools to combine these screening data and use them for target discovery⁴⁸⁰. Future work will bring a deeper understanding of the regulatory networks downstream of MLLr fusion genes and the importance of intrinsic apoptosis as a hallmark of cancer.

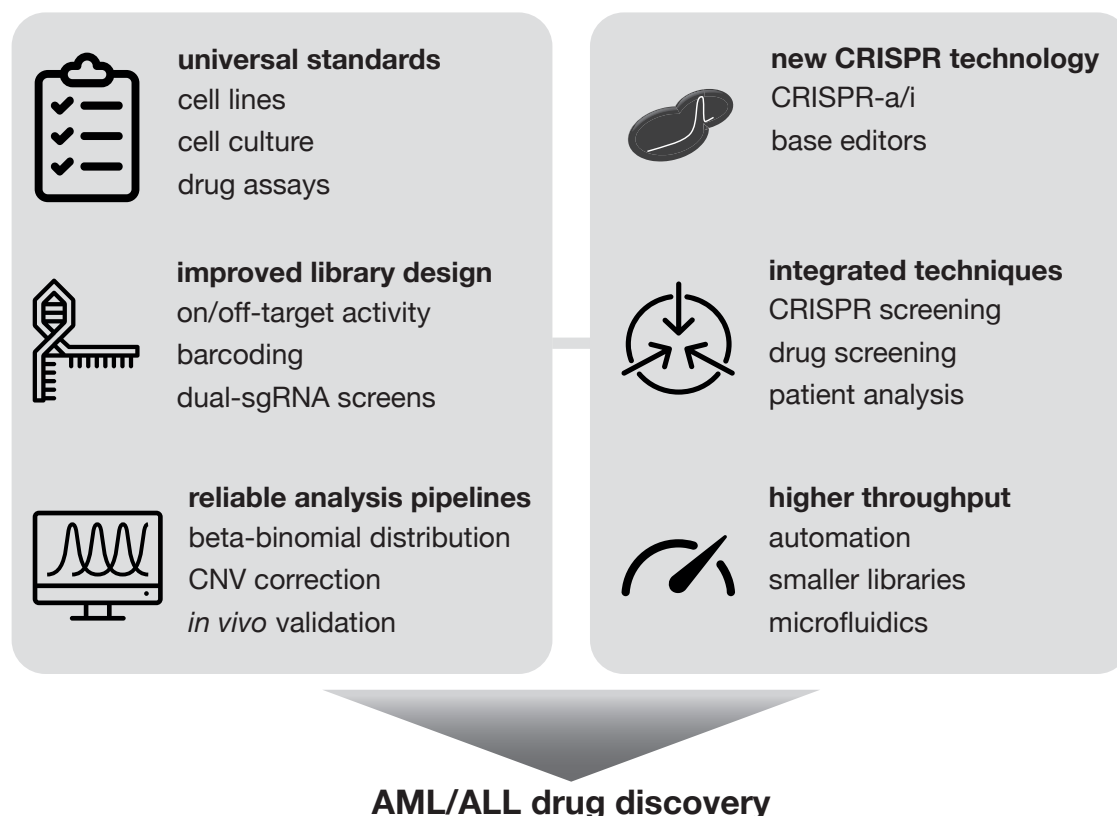


Figure 6.3: High-throughput screening for leukemia drug discovery. Advances in the design of CRISPR screens and other high-throughput assays will create novel opportunities for drug discovery in MLLr leukemia.

6.3 Concluding remarks

In this thesis, I have documented my efforts to combine the biology of MLLr leukemia and the technology of pooled CRISPR screening.

The prognosis for patients with *MLL* rearrangements, especially infants, remains dismal. Even as we explore novel and exciting cancer immunotherapies, there is much that we do not understand about traditional chemotherapy. MLLr leukemia exemplifies the challenges of targeted cancer therapy: the disease is genetically well defined but druggable targets are elusive.

Resistance to targeted therapies for MLLr leukemia reveals the gaps in our understanding of this disease. This thesis shows that a combination of rational genetic analysis and unbiased genome-wide screening can shed some light on a small aspect of how MLL fusion proteins cause leukemia, but it also highlights how poorly

we understand the mechanisms of action of many chemotherapies. Indeed, a recent study demonstrated that the mechanism of action of many clinical drugs is unknown or incorrect⁴⁸¹. CRISPR has greatly enhanced our ability to probe the genome but it cannot single-handedly explain how non-coding RNA contributes to leukemogenesis or how MLLr fusion proteins affect resistance to BH3 mimetic drugs. A combination of precise reverse genetics, high-throughput forward screening and *in silico* modelling will eventually yield answers. Over time, a better understanding of the downstream effects of different MLL fusion proteins should close the gap in therapeutic outcomes between leukemia patients with and without MLL rearrangements.

Appendices



Oligonucleotide sequences

Oligonucleotide	Sequence
D501	AATGATACGGCGACCACCGAGATCTACA CTATAGCCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTTTGTGGAAAGGACGA AACACCG
D502 A	AATGATACGGCGACCACCGAGATCTACA CATAGAGGCACACTCTTTCCCTACACGAC GCTCTTCCGATCTATTGTGGAAAGGACG AAACACCG
D503 GA	AATGATACGGCGACCACCGAGATCTACA CCCTATCCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTGATTGTGGAAAGGAC GAAACACCG
D504 CGA	AATGATACGGCGACCACCGAGATCTACA CGGCTCTGAACACTCTTTCCCTACACGAC GCTCTTCCGATCTCGATTGTGGAAAGGA CGAAACACCG
D505 TCGA	AATGATACGGCGACCACCGAGATCTACA CAGGCGAAGACACTCTTTCCCTACACGAC GCTCTTCCGATCTTCGATTGTGGAAAGG ACGAAACACCG
D506 ATCGA	AATGATACGGCGACCACCGAGATCTACA CTAATCTTAACACTCTTTCCCTACACGAC GCTCTTCCGATCTATCGATTGTGGAAAG GACGAAACACCG

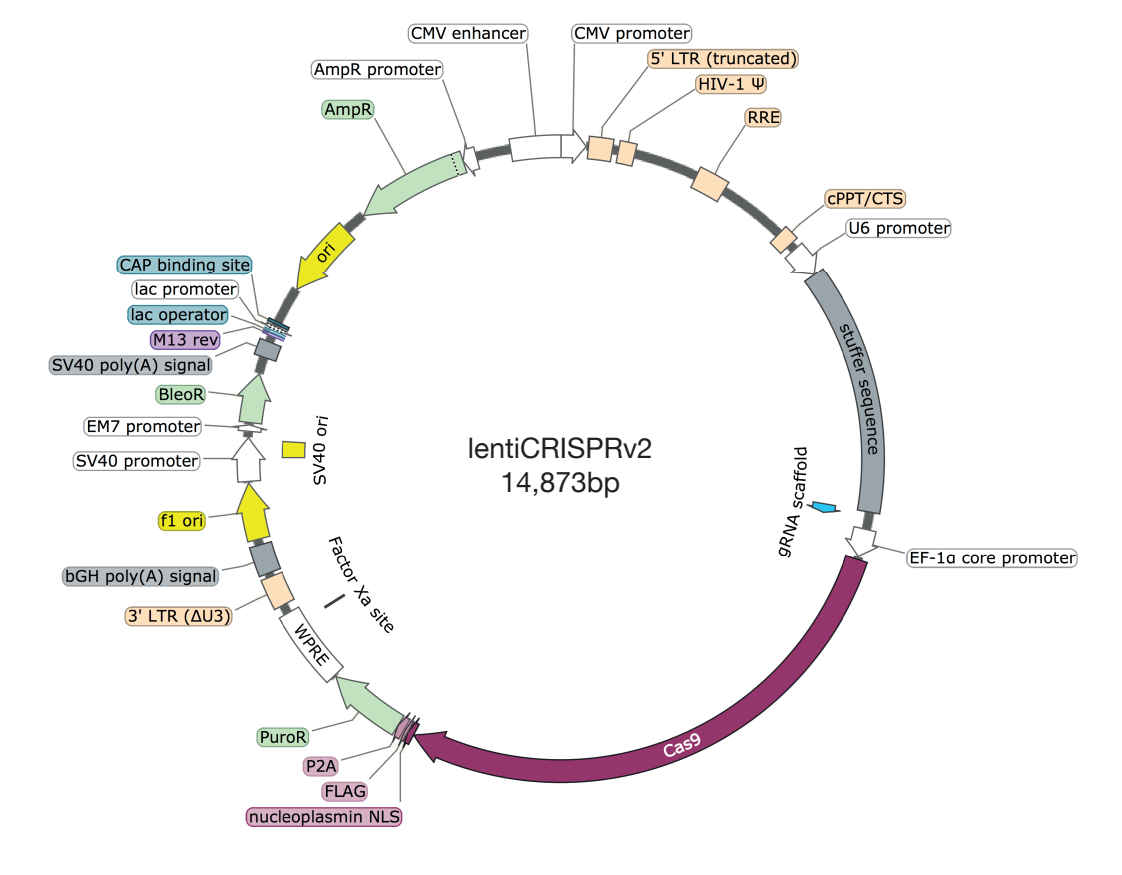
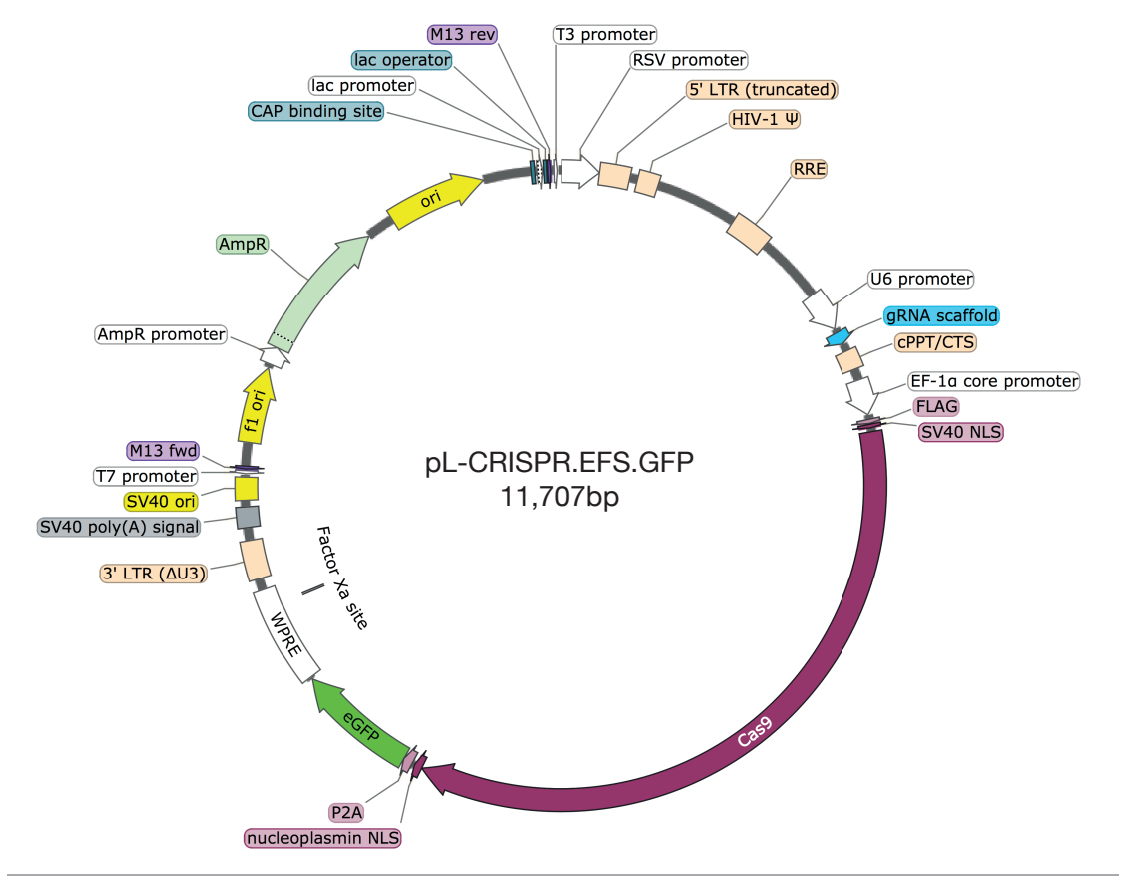
Oligonucleotide	Sequence
D701	CAAGCAGAAGACGGCATAACGAGATCGAG TAATGTGACTGGAGTTCAGACGTGTGCT CTTCCGATCTACTTGCTATTTCTAGCTCT AAAAC
D702	CAAGCAGAAGACGGCATAACGAGATTCTC CGGAGTGACTGGAGTTCAGACGTGTGCT CTTCCGATCTACTTGCTATTTCTAGCTCT AAAAC
D704	CAAGCAGAAGACGGCATAACGAGATGGAA TCTCGTGACTGGAGTTCAGACGTGTGCT CTTCCGATCTACTTGCTATTTCTAGCTCT AAAAC
D705	CAAGCAGAAGACGGCATAACGAGATTTCT GAATGTGACTGGAGTTCAGACGTGTGCT CTTCCGATCTACTTGCTATTTCTAGCTCT AAAAC
D705	CAAGCAGAAGACGGCATAACGAGATTTCT GAATGTGACTGGAGTTCAGACGTGTGCT CTTCCGATCTACTTGCTATTTCTAGCTCT AAAAC
D706	CAAGCAGAAGACGGCATAACGAGATACGA ATTCGTGACTGGAGTTCAGACGTGTGCT CTTCCGATCTACTTGCTATTTCTAGCTCT AAAAC
D706	CAAGCAGAAGACGGCATAACGAGATACGA ATTCGTGACTGGAGTTCAGACGTGTGCT CTTCCGATCTACTTGCTATTTCTAGCTCT AAAAC
D707	CAAGCAGAAGACGGCATAACGAGATAGCT TCAGGTGACTGGAGTTCAGACGTGTGCT CTTCCGATCTACTTGCTATTTCTAGCTCT AAAAC
D707	CAAGCAGAAGACGGCATAACGAGATAGCT TCAGGTGACTGGAGTTCAGACGTGTGCT CTTCCGATCTACTTGCTATTTCTAGCTCT AAAAC
M_0014_GAPDH-qPCR-1_F	AACAGCGACACCCACTCCTC
M_0015_GAPDH-qPCR-1_R	CATACCAGGAAATGAGCTTGACAA
M_0067_MIR196B-KO-1_F	CACCgTCGTGCTGTTCGAGAGAAAGG
M_0068_MIR196B-KO-1_R	AAACCCTTTCTCTCGACAGCACGAC
M_0069_MIR196B-KO-2_F	CACCgCCAGAACTGGTCGGTGATTT
M_0070_MIR196B-KO-2_R	AAACAAATCACCGACCAGTTCTGGC
M_0071_SCRAMBLE-1_F	CACCgGCACTACCAGAGCTAACTCA
M_0072_SCRAMBLE-1_R	AAACTGAGTTAGCTCTGGTAGTGCC

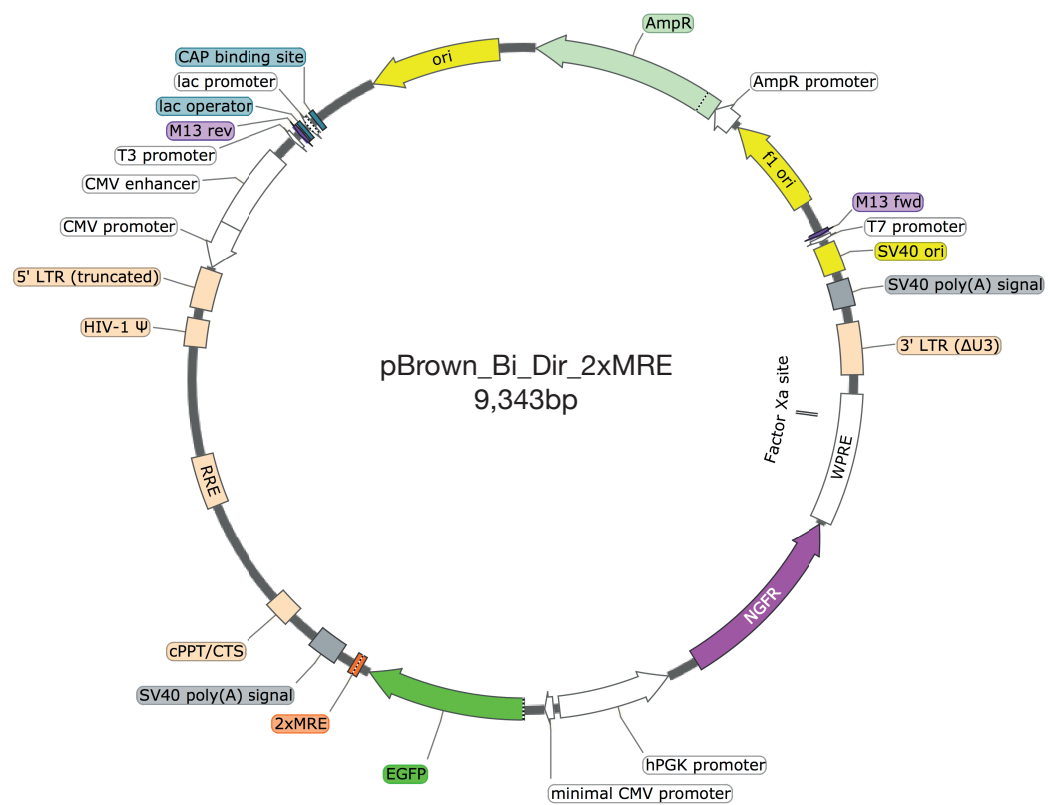
Oligonucleotide	Sequence
M_0107_BAX-KO-1_F	CACCAGCCCAACAGCCGCTCCCGG
M_0108_BAX-KO-1_R	AAACCCGGGAGCGGCTGTTGGGCT
M_0109_BAX-KO-2_F	CACCGATGGTCACGGTCTGCCACG
M_0110_BAX-KO-2_R	AAACCGTGGCAGACCGTGACCATC
M_0111_BAK-KO-1_F	CACCGGAACCTCTGAGTCATAGCGT
M_0112_BAK-KO-1_R	AAACACGCTATGACTCAGAGTTCC
M_0113_BAK-KO-2_F	CACCTGCATCACTGCATTGCCCGG
M_0114_BAK-KO-2_R	AAACCCGGGCAATGCAGTGATGCA
M_0115_BCL2L1-KO-1_F	CACCAATCCTGAAGGCAATCACGG
M_0116_BCL2L1-KO-1_R	AAACCCGTGATTGCCTTCAGGATT
M_0117_BCL2L1-KO-2_F	CACCGGAGGATCGAGACAGCAGGG
M_0118_BCL2L1-KO-2_R	AAACCCCTGCTGTCTCGATCCTCC
M_0119_MCL1-KO-1_F	CACCTCTGGCGTGAGGGTGGACGG
M_0120_MCL1-KO-1_R	AAACCCGTCCACCCTCACGCCAGA
M_0121_MCL1-KO-2_F	CACCGCTTCCGCCAATCACCGCGC
M_0122_MCL1-KO-2_R	AAACGCGCGGTGATTGGCGGAAGC
M_0123_BCL2L1-KO-1_F	CACCAGTAAAGCAAGCGCTGAGGG
M_0124_BCL2L1-KO-1_R	AAACCCCTCAGCGCTTGCTTTACT
M_0125_BCL2L1-KO-2_F	CACCTGTGGAAGAGAACAGGACTG
M_0126_BCL2L1-KO-2_R	AAACCAGTCCTGTTCTCTTCCACA
M_0129_HOXA9-qPCR-4_F	GTCCCACGCTTGACACTCA
M_0130_HOXA9-qPCR-4_R	GGGTTATTGGGATCGATGGG
M_0131_POLR2A-KO-1_F	CACCGGACTGGTTGGAGAATAGGA
M_0132_POLR2A-KO-1_R	AAACTCCTATTCTCCAACCAGTCC
M_0133_POLR2A-KO-2_F	CACCGGTAGGCGAGTACTTGGGAG
M_0134_POLR2A-KO-2_R	AAACCTCCCAAGTACTCGCCTACC
M_0135_TUBB-KO-1_F	CACCTGAGCTGGCCAGGGAAACGG
M_0136_TUBB-KO-1_R	AAACCCGTTTCCCTGGCCAGCTCA
M_0137_TUBB-KO-2_F	CACCTGAGGTATCGGCCGTGGCGG
M_0138_TUBB-KO-2_R	AAACCCGCCACGGCCGATACCTCA
M_0139_AAVS1-KO-1-F	CACCGGGGCCACTAGGGACAGGAT
M_0140_AAVS1-KO-1-R	AAACATCCTGTCCCTAGTGGCCCC
M_0141_sgCTRL-1-F	CACCGGATACTTCTTTCGAACGTTT
M_0142_sgCTRL-1-R	AAACAAACGTTTCGAAGAAGTATCC
M_0143_v2.1-F1 Moffat	GAGGGCCTATTTCCCATGATTC
M_0144_v2.1-R1 Moffat	GTTGCGAAAAAGAACGTTACCGG
M_0145_AAVS1_1_T7_1_F	AGGAGAGAGATGGCTCCAGG
M_0146_AAVS1_1_T7_1_R	CCGGTTAATGTGGCTCTGGT
M_0147_BAK_1_T7_1_F	GGTTACTGGCTCACCTGCAT
M_0148_BAK_1_T7_1_R	CATGACAGAGGAGTGACTGGAG
M_0149_BAK_2_T7_1_F	CCTGTTTGAGAGTGGCATCAAT
M_0150_BAK_2_T7_1_R	AGATATAGGAACCCGAGGAAGG
M_0151_BAX_1_T7_1_F	CCTGGCCATGTTTACAATTTTT
M_0152_BAX_1_T7_1_R	GTGAGGAGGCTTGAGGAGTCT

Oligonucleotide	Sequence
M_0153_BAX_2_T7_1_F	AGTCACTGAAGCGACTGATGTC
M_0154_BAX_2_T7_1_R	GGTTTATTACCCCCTCAAGACC
M_0155_BCL2L1_2_T7_1_F	AGCTGGTGGTTGACTTTCTCTC
M_0156_BCL2L1_2_T7_1_R	GTACCGCAGTTCAAACCTCGTC
M_0157_BCL2L11_1_T7_1_F	TCTGAGTGTGACCGAGAAGGTA
M_0158_BCL2L11_1_T7_1_R	GCAGGGAGGATCTTCTCATAAA
M_0175_BCL2_1_T7_1_F	CACCTGTGGTCCACCTGAC
M_0176_BCL2_1_T7_1_R	CGGTTTCAGGTACTCAGTCATCC
M_0179_BCL2_2_T7_1_F	GTGGCATTCTCTGCTCTCTCTC
M_0180_BCL2_2_T7_1_R	ACCTTTTGCATATTTGTTTGGG
MGJ_0052_MEIS1-qPCR-1_F	GCGCAAAGGTACGACGATCT
MGJ_0053_MEIS1-qPCR-1_R	GGTACTGATGCGAGTGCAGA
MGJ_0054_FAS-qPCR-1_F	GTGACCCTTGCACCAAATGT
MGJ_0055_FAS-qPCR-1_R	CCTTTCTGTGCTTTCTGCATGT

B

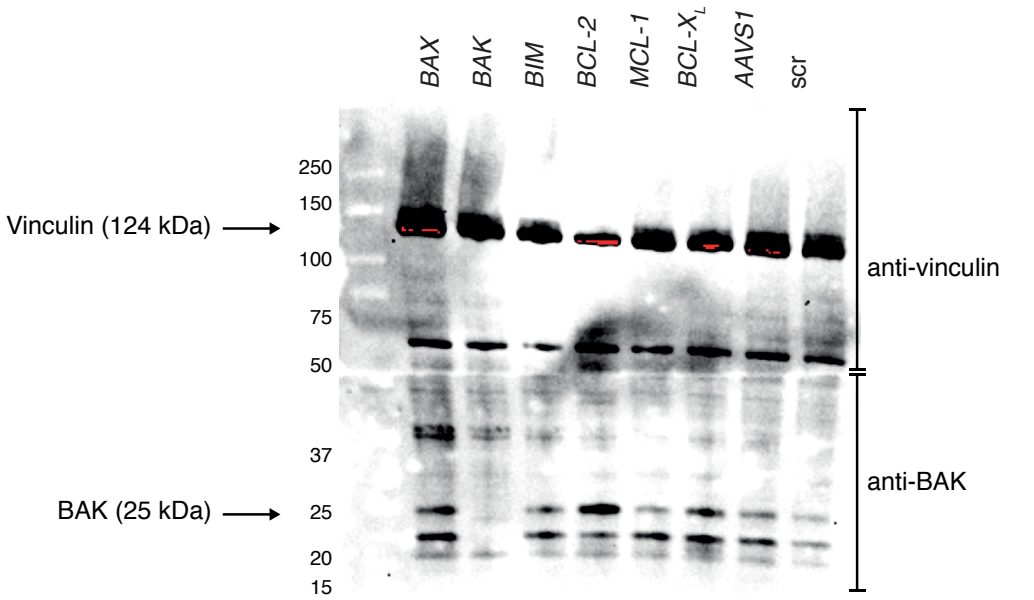
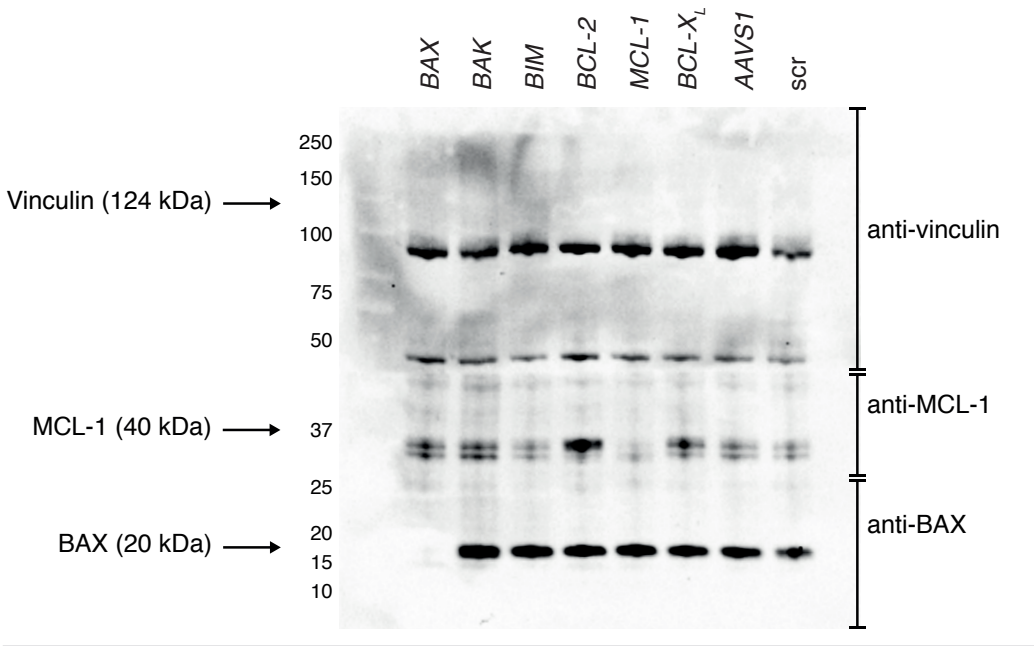
Plasmid maps

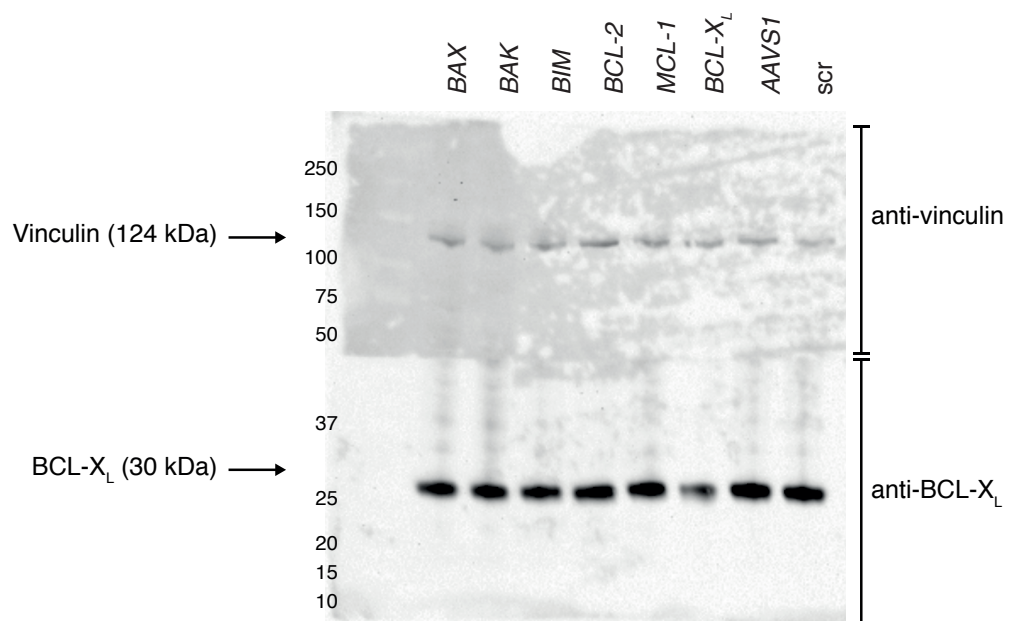
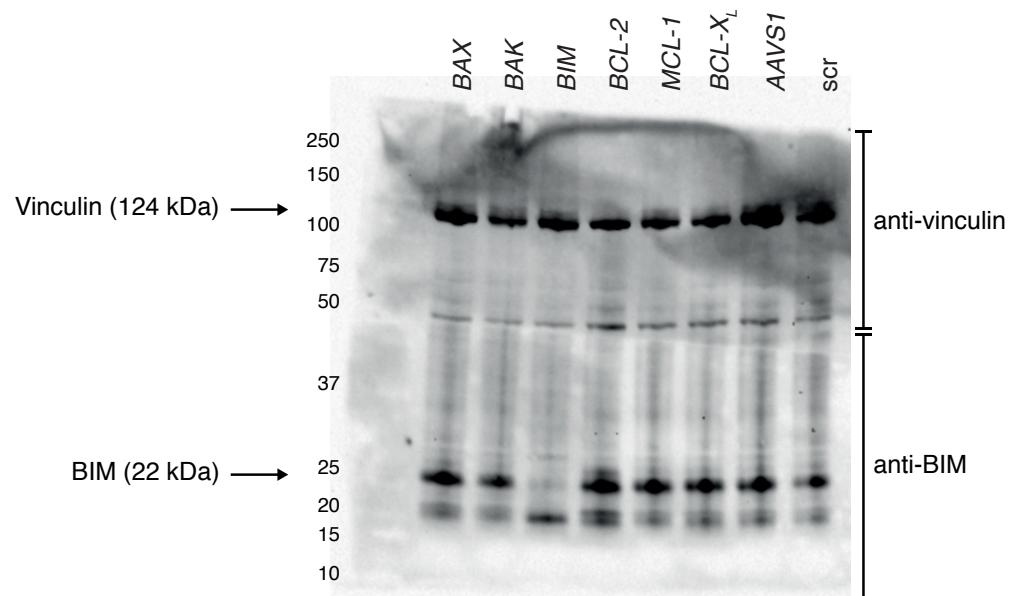




C

Western blots





References

1. Orkin, S. H. & Zon, L. I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631–44 (Feb. 2008).
2. Hunger, S. P. & Mullighan, C. G. Acute Lymphoblastic Leukemia in Children. *N. Engl. J. Med.* **373** (ed Longo, D. L.) 1541–1552 (Oct. 2015).
3. Huret, J., Dessen, P. & Bernheim, A. An Atlas on Chromosomes in Hematological Malignancies. Example: 11q23 and MLL partners. *Leukemia* **15**, 987–989 (June 2001).
4. Winters, A. C. & Bernt, K. M. MLL-Rearranged Leukemias-An Update on Science and Clinical Approaches. *Front. Pediatr.* **5**, 4 (2017).
5. Krivtsov, A. V. & Armstrong, S. A. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat. Rev. Cancer* **7**, 823–833 (Nov. 2007).
6. Cox, M. C. *et al.* Chromosomal Aberration of the 11q23 Locus in Acute Leukemia and Frequency of MLL Gene Translocation Results in 378 Adult Patients. *Am. J. Clin. Pathol.* **122**, 298–306 (Aug. 2004).
7. Super, H. J. *et al.* Rearrangements of the MLL gene in therapy-related acute myeloid leukemia in patients previously treated with agents targeting DNA-topoisomerase II. *Blood* **82**, 3705–11 (Dec. 1993).
8. Andersen, M. K. *et al.* Therapy-related acute lymphoblastic leukaemia with MLL rearrangements following DNA topoisomerase II inhibitors, an increasing problem: report on two new cases and review of the literature since 1992. *Br. J. Haematol.* **114**, 539–543 (Sept. 2001).
9. Schoch, C. *et al.* AML with 11q23/MLL abnormalities as defined by the WHO classification: incidence, partner chromosomes, FAB subtype, age distribution, and prognostic impact in an unselected series of 1897 cytogenetically analyzed AML cases. *Blood* **102**, 2395–402 (Oct. 2003).
10. Krumbholz, M. *et al.* From initiation to eradication: the lifespan of an MLL-rearranged therapy-related paediatric AML. *Bone Marrow Transplant.* **50**, 1382–1384 (Oct. 2015).
11. Arber, D. A. *et al.* The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* **127**, 2391–2405 (May 2016).
12. Pui, C.-H. *et al.* Outcome of treatment in childhood acute lymphoblastic leukaemia with rearrangements of the 11q23 chromosomal region. *Lancet* **359**, 1909–1915 (June 2002).

13. Pieters, R. *et al.* Outcome of Infants Younger Than 1 Year With Acute Lymphoblastic Leukemia Treated With the Interfant-06 Protocol: Results From an International Phase III Randomized Study. *J. Clin. Oncol.* **37**, 2246–2256 (Sept. 2019).
14. Kotecha, R. S., Gottardo, N. G., Kees, U. R. & Cole, C. H. The evolution of clinical trials for infant acute lymphoblastic leukemia. *Blood Cancer J.* **4**, e200–e200 (Apr. 2014).
15. Pieters, R. *et al.* A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. *Lancet* **370**, 240–250 (July 2007).
16. Dreyer, Z. E. *et al.* Intensified chemotherapy without SCT in infant ALL: Results from COG P9407 (Cohort 3). *Pediatr. Blood Cancer* **62**, 419–426 (Mar. 2015).
17. Hilden, J. M. *et al.* Analysis of prognostic factors of acute lymphoblastic leukemia in infants: report on CCG 1953 from the Children's Oncology Group. *Blood* **108**, 441–51 (July 2006).
18. Creutzig, U. *et al.* Randomized trial comparing liposomal daunorubicin with idarubicin as induction for pediatric acute myeloid leukemia: results from Study AML-BFM 2004. *Blood* **122**, 37–43 (July 2013).
19. Balgobind, B. V. *et al.* Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. *Blood* **114**, 2489–96 (Sept. 2009).
20. Alexander, T. B. *et al.* Decreased relapsed rate and treatment-related mortality contribute to improved outcomes for pediatric acute myeloid leukemia in successive clinical trials. *Cancer* **123**, 3791–3798 (Oct. 2017).
21. Krivtsov, A. V., Hoshii, T. & Armstrong, S. A. Mixed-Lineage Leukemia Fusions and Chromatin in Leukemia. *Cold Spring Harb. Perspect. Med.* **7**, a026658 (Nov. 2017).
22. Rao, R. C. & Dou, Y. Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat. Rev. Cancer* **15**, 334–346 (June 2015).
23. Slany, R. K. The molecular mechanics of mixed lineage leukemia. *Oncogene* **35**, 5215–5223 (Oct. 2016).
24. Milne, T. A. *et al.* MLL associates specifically with a subset of transcriptionally active target genes. *Proc. Natl. Acad. Sci.* **102**, 14765–14770 (Oct. 2005).
25. Dou, Y. *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13**, 713–719 (Aug. 2006).
26. Yokoyama, A. & Cleary, M. L. Menin Critically Links MLL Proteins with LEDGF on Cancer-Associated Target Genes. *Cancer Cell* **14**, 36–46 (July 2008).
27. Milne, T. A. *et al.* Multiple Interactions Recruit MLL1 and MLL1 Fusion Proteins to the HOXA9 Locus in Leukemogenesis. *Mol. Cell* **38**, 853–863 (June 2010).
28. Wang, Z. *et al.* Pro Isomerization in MLL1 PHD3-Bromo Cassette Connects H3K4me Readout to CyP33 and HDAC-Mediated Repression. *Cell* **141**, 1183–1194 (June 2010).

29. Milne, T. A. *et al.* MLL Targets SET Domain Methyltransferase Activity to Hox Gene Promoters. *Mol. Cell* **10**, 1107–1117 (Nov. 2002).
30. Yokoyama, A. *et al.* Proteolytically cleaved MLL subunits are susceptible to distinct degradation pathways. *J. Cell Sci.* **124**, 2208–19 (July 2011).
31. Von Bergh, A. *et al.* A DNA probe combination for improved detection of MLL/11q23 breakpoints by double-color interphase-FISH in acute leukemias. *Genes, Chromosom. Cancer* **28**, 14–22 (May 2000).
32. Meyer, C. *et al.* The MLL recombinome of acute leukemias in 2013. *Leukemia* **27**, 2165–2176 (Nov. 2013).
33. Meyer, C. *et al.* The MLL recombinome of acute leukemias in 2017. *Nat. Publ. Gr.* **32**, 273–284 (2017).
34. Greaves, M. F. *et al.* Leukemia in twins: lessons in natural history. *Blood* **102**, 2321–33 (Oct. 2003).
35. Ford, A. M. *et al.* In utero rearrangements in the trithorax-related oncogene in infant leukaemias. *Nature* **363**, 358–360 (May 1993).
36. Gale, K. B. *et al.* Backtracking leukemia to birth: Identification of clonotypic gene fusion sequences in neonatal blood spots. *Proc. Natl. Acad. Sci.* **94**, 13950–13954 (Dec. 1997).
37. Strissel, P. L., Strick, R., Rowley, J. D. & Zeleznik-Le, N. J. An in vivo topoisomerase II cleavage site and a DNase I hypersensitive site colocalize near exon 9 in the MLL breakpoint cluster region. *Blood* **92**, 3793–803 (Nov. 1998).
38. Ng, A., Taylor, G. M., Wynn, R. F. & Eden, O. B. Effects of topoisomerase 2 inhibitors on the MLL gene in children receiving chemotherapy: a prospective study. *Leukemia* **19**, 253–259 (Feb. 2005).
39. Aplan, P., Chervinsky, D., Stanulla, M. & Burhans, W. Site-specific DNA cleavage within the MLL breakpoint cluster region induced by topoisomerase II inhibitors. *Blood* **87** (1996).
40. Slany, R. K., Lavau, C. & Cleary, M. L. The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX. *Mol. Cell. Biol.* **18**, 122–9 (Jan. 1998).
41. Okuda, H. *et al.* MLL fusion proteins link transcriptional coactivators to previously active CpG-rich promoters. *Nucleic Acids Res.* **42**, 4241–56 (Apr. 2014).
42. Stam, R. W. MLL-AF4 driven leukemogenesis: what are we missing? *Cell Res.* **22**, 948–949 (June 2012).
43. Bueno, C. *et al.* A human ESC model for MLL-AF4 leukemic fusion gene reveals an impaired early hematopoietic-endothelial specification. *Cell Res.* **22**, 986–1002 (June 2012).
44. Montes, R. *et al.* Enforced expression of MLL-AF4 fusion in cord blood CD34+ cells enhances the hematopoietic repopulating cell function and clonogenic potential but is not sufficient to initiate leukemia. *Blood* **117**, 4746–58 (May 2011).
45. Basecke, J., Whelan, J. T., Griesinger, F. & Bertrand, F. E. The MLL partial tandem duplication in acute myeloid leukaemia. *Br. J. Haematol.* **135**, 438–449 (Nov. 2006).

46. Lavau, C., Szilvassy, S. J., Slany, R. & Cleary, M. L. Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. *EMBO J.* **16**, 4226–4237 (July 1997).
47. Krivtsov, A. V. *et al.* Transformation from committed progenitor to leukaemia stem cell initiated by MLLAF9. *Nature* **442**, 818–822 (Aug. 2006).
48. Lin, S. *et al.* Instructive Role of MLL-Fusion Proteins Revealed by a Model of t(4;11) Pro-B Acute Lymphoblastic Leukemia. *Cancer Cell* **30**, 737–749 (Nov. 2016).
49. Thomas, M. *et al.* Targeting MLL-AF4 with short interfering RNAs inhibits clonogenicity and engraftment of t(4;11)-positive human leukemic cells. *Blood* **106**, 3559–66 (Nov. 2005).
50. Tamai, H. *et al.* Activated K-Ras protein accelerates human MLL/AF4-induced leukemo-lymphomogenicity in a transgenic mouse model. *Leukemia* **25**, 888–891 (May 2011).
51. Andersson, A. K. *et al.* The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nat. Genet.* **47**, 330–337 (Apr. 2015).
52. Zeisig, D. T. *et al.* The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin. *Oncogene* **24**, 5525–5532 (Aug. 2005).
53. Mueller, D. *et al.* A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood* **110**, 4445–54 (Dec. 2007).
54. Mueller, D. *et al.* Misguided Transcriptional Elongation Causes Mixed Lineage Leukemia. *PLoS Biol.* **7** (ed Zeleznik-Le, N.) e1000249 (Nov. 2009).
55. Bitoun, E., Oliver, P. L. & Davies, K. E. The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Hum. Mol. Genet.* **16**, 92–106 (Jan. 2007).
56. Lin, C. *et al.* AFF4, a component of the ELL/P-TFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol. Cell* **37**, 429–37 (Feb. 2010).
57. Muntean, A. G. & Hess, J. L. The Pathogenesis of Mixed-Lineage Leukemia. *Annu. Rev. Pathol. Mech. Dis.* **7**, 283–301 (Feb. 2012).
58. Kerry, J. *et al.* MLL-AF4 Spreading Identifies Binding Sites that Are Distinct from Super-Enhancers and that Govern Sensitivity to DOT1L Inhibition in Leukemia. *Cell Rep.* **18**, 482–495 (2017).
59. Godfrey, L. *et al.* DOT1L inhibition reveals a distinct subset of enhancers dependent on H3K79 methylation. *Nat. Commun.* **10**, 2803 (Dec. 2019).
60. Biswas, D. *et al.* Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15751–6 (Sept. 2011).
61. Gilan, O. *et al.* Functional interdependence of BRD4 and DOT1L in MLL leukemia. *Nat. Struct. Mol. Biol.* **23**, 673–681 (July 2016).

62. Ballabio, E. & Milne, T. A. Molecular and Epigenetic Mechanisms of MLL in Human Leukemogenesis. *Cancers (Basel)*. **4**, 904–944 (Sept. 2012).
63. Vlaming, H. & van Leeuwen, F. The upstreams and downstreams of H3K79 methylation by DOT1L. *Chromosoma* **125**, 593–605 (Sept. 2016).
64. Jo, S. Y., Granowicz, E. M., Maillard, I., Thomas, D. & Hess, J. L. Requirement for Dot1l in murine postnatal hematopoiesis and leukemogenesis by MLL translocation. *Blood* **117**, 4759–68 (May 2011).
65. Chang, M.-J. *et al.* Histone H3 Lysine 79 Methyltransferase Dot1 Is Required for Immortalization by MLL Oncogenes. *Cancer Res.* **70**, 10234–10242 (Dec. 2010).
66. Maethner, E. *et al.* MLL-ENL Inhibits Polycomb Repressive Complex 1 to Achieve Efficient Transformation of Hematopoietic Cells. *Cell Rep.* **3**, 1553–1566 (May 2013).
67. Marschalek, R. Mechanisms of leukemogenesis by MLL fusion proteins. *Br. J. Haematol.* **152**, 141–154 (Jan. 2011).
68. Bridges, C. B. Current Maps of the Location of the Mutant Genes of *Drosophila Melanogaster*. *Proc. Natl. Acad. Sci.* **7**, 127–132 (Apr. 1921).
69. Lewis, E. B. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570 (Dec. 1978).
70. Shah, N. & Sukumar, S. The Hox genes and their roles in oncogenesis. *Nat. Rev. Cancer* **10**, 361–371 (May 2010).
71. Collins, C. T. & Hess, J. L. Deregulation of the HOXA9/MEIS1 axis in acute leukemia. *Curr. Opin. Hematol.* **23**, 354–361 (July 2016).
72. Li, Z. *et al.* PBX3 is an important cofactor of HOXA9 in leukemogenesis. *Blood* **121**, 1422–1431 (Feb. 2013).
73. Garcia-Cuellar, M.-P., Steger, J., Filler, E., Hetzner, K. & Slany, R. K. Pbx3 and Meis1 cooperate through multiple mechanisms to support Hox-induced murine leukemia. *Haematologica* **100**, 905–13 (July 2015).
74. Sauvageau, G. *et al.* Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12223–7 (Dec. 1994).
75. Pineault, N., Helgason, C. D., Lawrence, H. & Humphries, R. Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. *Exp. Hematol.* **30**, 49–57 (Jan. 2002).
76. Argiropoulos, B. & Humphries, R. K. Hox genes in hematopoiesis and leukemogenesis. *Oncogene* **26**, 6766–6776 (Oct. 2007).
77. Cheng, J. *et al.* Hematopoietic defects in mice lacking the sialomucin CD34. *Blood* **87**, 479–90 (Jan. 1996).
78. Izon, D. J. *et al.* Loss of function of the homeobox gene Hoxa-9 perturbs early T-cell development and induces apoptosis in primitive thymocytes. *Blood* **92**, 383–93 (July 1998).
79. Hisa, T. *et al.* Hematopoietic, angiogenic and eye defects in Meis1 mutant animals. *EMBO J.* **23**, 450–459 (Jan. 2004).

80. Lawrence, H. J. *et al.* Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. *Blood* **106**, 3988–94 (Dec. 2005).
81. Kumar, A. R. *et al.* Hoxa9 influences the phenotype but not the incidence of Mll-AF9 fusion gene leukemia. *Blood* **103**, 1823–1828 (Mar. 2004).
82. Behm, F. G. *et al.* Rearrangement of the MLL gene confers a poor prognosis in childhood acute lymphoblastic leukemia, regardless of presenting age. *Blood* **87**, 2870–7 (Apr. 1996).
83. Armstrong, S. A. *et al.* MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat. Genet.* **30**, 41–47 (Jan. 2002).
84. Popovic, R. *et al.* Regulation of mir-196b by MLL and its overexpression by MLL fusions contributes to immortalization. *Blood* **113**, 3314–22 (Apr. 2009).
85. Kroon, E. *et al.* Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J.* **17**, 3714–3725 (July 1998).
86. Thorsteinsdottir, U., Kroon, E., Jerome, L., Blasi, F. & Sauvageau, G. Defining Roles for HOX and MEIS1 Genes in Induction of Acute Myeloid Leukemia. *Mol. Cell. Biol.* **21**, 224–234 (Jan. 2001).
87. Lawrence, H. J. *et al.* *Frequent co-expression of the HOXA9 and MEIS1 homeobox genes in human myeloid leukemias* tech. rep. (1999), 1993–1999.
88. Abate-Shen, C. Deregulated homeobox gene expression in cancer: cause or consequence? *Nat. Rev. Cancer* **2**, 777–785 (Oct. 2002).
89. Hess, J. L. *et al.* c-Myb is an essential downstream target for homeobox-mediated transformation of hematopoietic cells. *Blood* **108**, 297–304 (July 2006).
90. Dorsam, S. T. *et al.* The transcriptome of the leukemogenic homeoprotein HOXA9 in human hematopoietic cells. *Blood* **103**, 1676–84 (Mar. 2004).
91. Collins, C. T. & Hess, J. L. Role of HOXA9 in leukemia: dysregulation, cofactors and essential targets. *Oncogene* **35**, 1090–1098 (Mar. 2016).
92. Breitinger, C., Maethner, E., Garcia-Cuellar, M.-P. & Slany, R. K. The homeodomain region controls the phenotype of HOX-induced murine leukemia. *Blood* **120**, 4018–4027 (Nov. 2012).
93. Sun, Y. *et al.* HOXA9 Reprograms the Enhancer Landscape to Promote Leukemogenesis. *Cancer Cell* **34**, 643–658.e5 (Oct. 2018).
94. Zhong, X. *et al.* HoxA9 transforms murine myeloid cells by a feedback loop driving expression of key oncogenes and cell cycle control genes. *Blood Adv.* **2**, 3137–3148 (Nov. 2018).
95. Garcia-Cuellar, M.-P., Bttner, C., Bartenhagen, C., Dugas, M. & Slany, R. Leukemogenic MLL-ENL Fusions Induce Alternative Chromatin States to Drive a Functionally Dichotomous Group of Target Genes. *Cell Rep.* **15**, 310–322 (Apr. 2016).
96. Thiel, A. T. *et al.* MLL-AF9-Induced Leukemogenesis Requires Coexpression of the Wild-Type Mll Allele. *Cancer Cell* **17**, 148–159 (Feb. 2010).

97. Milne, T. A., Martin, M. E., Brock, H. W., Slany, R. K. & Hess, J. L. Leukemogenic MLL Fusion Proteins Bind across a Broad Region of the *Hoxa9* Locus, Promoting Transcription and Multiple Histone Modifications. *Cancer Res.* **65**, 11367–11374 (Dec. 2005).
98. Bernt, K. *et al.* MLL-Rearranged Leukemia Is Dependent on Aberrant H3K79 Methylation by DOT1L. *Cancer Cell* **20**, 66–78 (July 2011).
99. Nguyen, A. T., Taranova, O., He, J. & Zhang, Y. DOT1L, the H3K79 methyltransferase, is required for MLL-AF9-mediated leukemogenesis. *Blood* **117**, 6912–22 (June 2011).
100. Daigle, S. R. *et al.* Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood* **122**, 1017–1025 (Aug. 2013).
101. Yokoyama, A. *et al.* The Menin Tumor Suppressor Protein Is an Essential Oncogenic Cofactor for MLL-Associated Leukemogenesis. *Cell* **123**, 207–218 (Oct. 2005).
102. Chen, Y.-X. *et al.* The tumor suppressor menin regulates hematopoiesis and myeloid transformation by influencing Hox gene expression. *Proc. Natl. Acad. Sci.* **103**, 1018–1023 (Jan. 2006).
103. Nakamura, T., Largaespada, D. A., Shaughnessy, J. D., Jenkins, N. A. & Copeland, N. G. Cooperative activation of Hoxa and Pbx1-related genes in murine myeloid leukaemias. *Nat. Genet.* **12**, 149–153 (Feb. 1996).
104. Dik, W. A. *et al.* CALM-AF10+ T-ALL expression profiles are characterized by overexpression of HOXA and BMI1 oncogenes. *Leukemia* **19**, 1948–1957 (Nov. 2005).
105. Soulier, J. *et al.* Blood. *Blood* **94**, 3491–3498 (July 2005).
106. Bissels, U., Bosio, A. & Wagner, W. MicroRNAs are shaping the hematopoietic landscape. *Haematologica* **97**, 160–7 (Feb. 2012).
107. Friedman, R. C., Farh, K. K.-H., Burge, C. B. & Bartel, D. P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **19**, 92–105 (Oct. 2008).
108. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **42**, D68–73 (Jan. 2014).
109. Bartel, D. P. Metazoan MicroRNAs. *Cell* **173**, 20–51 (Mar. 2018).
110. Ramalingam, P. *et al.* Biogenesis of intronic miRNAs located in clusters by independent transcription and alternative splicing. *RNA* **20**, 76–87 (Jan. 2014).
111. Guo, H., Ingolia, N. T., Weissman, J. S. & Bartel, D. P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **466**, 835–40 (Aug. 2010).
112. Djuranovic, S., Nahvi, A. & Green, R. miRNA-Mediated Gene Silencing by Translational Repression Followed by mRNA Deadenylation and Decay. *Science (80-.)*. **336**, 237–240 (Apr. 2012).

113. Tat, T. T., Maroney, P. A., Chamnongpol, S., Collier, J. & Nilsen, T. W. Cotranslational microRNA mediated messenger RNA destabilization. *Elife* **5**, e12880 (Apr. 2016).
114. Schotte, D., Pieters, R. & Den Boer, M. L. MicroRNAs in acute leukemia: from biological players to clinical contributors. *Leukemia* **26**, 1–12 (Jan. 2012).
115. Buza-Vidas, N. *et al.* Dicer is selectively important for the earliest stages of erythroid development. doi:10.1182/blood-2011-10-383653 (2012).
116. Mi, S. *et al.* MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia. *Proc. Natl. Acad. Sci.* **104**, 19971–19976 (Dec. 2007).
117. Chen, J., Odenike, O. & Rowley, J. D. Leukaemogenesis: more than mutant genes. *Nat. Rev. Cancer* **10**, 23–36 (Jan. 2010).
118. Marcucci, G., Mrozek, K., Radmacher, M. D., Garzon, R. & Bloomfield, C. D. The prognostic and functional role of microRNAs in acute myeloid leukemia. *Blood* **117** (2011).
119. Lu, J. *et al.* MicroRNA expression profiles classify human cancers. *Nature* **435**, 834–838 (June 2005).
120. Jongen-Lavrencic, M., Sun, S. M., Dijkstra, M. K., Valk, P. J. M. & Lowenberg, B. MicroRNA expression profiling in relation to the genetic heterogeneity of acute myeloid leukemia. *Blood* **111**, 5078–5085 (May 2008).
121. Li, Z. *et al.* miR-196b directly targets both HOXA9/MEIS1 oncogenes and FAS tumour suppressor in MLL-rearranged leukaemia. *Nat. Commun.* **3**, 688 (Feb. 2012).
122. Schotte, D. *et al.* Expression of miR-196b is not exclusively MLL-driven but is especially linked to activation of HOXA genes in pediatric acute lymphoblastic leukemia. *Haematologica* **95**, 1675–82 (Oct. 2010).
123. Velu, C. S. *et al.* Therapeutic antagonists of microRNAs deplete leukemia-initiating cell activity. *J. Clin. Invest.* **124**, 222–236 (Jan. 2014).
124. Meyer, S. E. *et al.* miR-196b target screen reveals mechanisms maintaining leukemia stemness with therapeutic potential. *J. Exp. Med.* **215**, 2115–2136 (Aug. 2018).
125. Sacar Demirci, M. D., Baumbach, J. & Allmer, J. On the performance of pre-microRNA detection algorithms. *Nat. Commun.* **8**, 330 (Dec. 2017).
126. Yekta, S., Tabin, C. J. & Bartel, D. P. MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nat. Rev. Genet.* **9**, 789–796 (Oct. 2008).
127. Yekta, S., Shih, I.-H. & Bartel, D. P. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**, 594–6 (Apr. 2004).
128. Fantini, S., Salsi, V., Vitobello, A., Rijli, F. M. & Zappavigna, V. MicroRNA-196b is transcribed from an autonomous promoter and is directly regulated by Cdx2 and by posterior Hox proteins during embryogenesis. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1849**, 1066–1080 (Aug. 2015).

129. Erfurth, F. E. *et al.* MLL protects CpG clusters from methylation within the Hoxa9 gene, maintaining transcript expression. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 7517–22 (May 2008).
130. Chen, P. *et al.* miR-9 is an essential oncogenic microRNA specifically overexpressed in mixed lineage leukemia-rearranged leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11511–6 (July 2013).
131. Drexler, H. G., Quentmeier, H. & MacLeod, R. A. F. Malignant hematopoietic cell lines: in vitro models for the study of MLL gene alterations. *Leukemia* **18**, 227–232 (Feb. 2004).
132. Rcker, F. G. *et al.* Molecular profiling reveals myeloid leukemia cell lines to be faithful model systems characterized by distinct genomic aberrations. *Leukemia* **20**, 994–1001 (June 2006).
133. Tsuchiya, S. *et al.* Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *Int. J. Cancer* **26**, 171–176 (Aug. 1980).
134. Odero, M. D., Zeleznik-Le, N. J., Chinwalla, V. & Rowley, J. D. Cytogenetic and molecular analysis of the acute monocytic leukemia cell line THP-1 with an MLL-AF9 translocation. *Genes. Chromosomes Cancer* **29**, 333–8 (Dec. 2000).
135. Quentmeier, H. *et al.* The LL-100 panel: 100 cell lines for blood cancer studies. *Sci. Rep.* **9**, 8218 (June 2019).
136. Sugimoto, K. *et al.* Frequent mutations in the p53 gene in human myeloid leukemia cell lines. *Blood* **79**, 2378–83 (May 1992).
137. Davies, H. *et al.* Mutations of the BRAF gene in human cancer. *Nature* **417**, 949–954 (June 2002).
138. Prange, K. H. M. *et al.* MLL-AF9 and MLL-AF4 oncofusion proteins bind a distinct enhancer repertoire and target the RUNX1 program in 11q23 acute myeloid leukemia. *Oncogene* **36**, 3346–3356 (June 2017).
139. Lin, S., Luo, R. T., Shrestha, M., Thirman, M. J. & Mulloy, J. C. The full transforming capacity of MLL-Af4 is interlinked with lymphoid lineage commitment. *Blood*, blood-2017-04-777185 (June 2017).
140. Bosshart, H. & Heinzelmann, M. THP-1 cells as a model for human monocytes. *Ann. Transl. Med.* **4**, 438 (Nov. 2016).
141. Nagata, S. Apoptosis and Clearance of Apoptotic Cells. *Annu. Rev. Immunol.* **36**, 489–517 (Apr. 2018).
142. Medina, C. B. & Ravichandran, K. S. Do not let death do us part: find-me’ signals in communication between dying cells and the phagocytes. *Cell Death Differ.* **23**, 979–989 (June 2016).
143. Fulda, S. Tumor resistance to apoptosis. *Int. J. Cancer* **124**, 511–515 (Feb. 2009).
144. Hanahan, D. & Weinberg, R. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (Mar. 2011).
145. Kale, J., Osterlund, E. J. & Andrews, D. W. BCL-2 family proteins: changing partners in the dance towards death. *Cell Death Differ.* **25**, 65–80 (2018).

146. Shamas-Din, A., Brahmabhatt, H., Leber, B. & Andrews, D. W. BH3-only proteins: Orchestrators of apoptosis. *Biochim. Biophys. Acta - Mol. Cell Res.* **1813**, 508–520 (Apr. 2011).
147. McArthur, K. *et al.* BAK/BAX macropores facilitate mitochondrial herniation and mtDNA efflux during apoptosis. *Science* (80-.). **359**, eaao6047 (Feb. 2018).
148. Thompson, C. B. *et al.* Apoptosis in the pathogenesis and treatment of disease. *Science* **267**, 1456–62 (Mar. 1995).
149. Lomonosova, E. & Chinnadurai, G. BH3-only proteins in apoptosis and beyond: an overview. *Oncogene* **27**, S2–S19 (Dec. 2008).
150. Roy, S. S. *et al.* Bad Targets the Permeability Transition Pore Independent of Bax or Bak to Switch between Ca²⁺-Dependent Cell Survival and Death. *Mol. Cell* **33**, 377–388 (Feb. 2009).
151. Westphal, D., Kluck, R. M. & Dewson, G. Building blocks of the apoptotic pore: how Bax and Bak are activated and oligomerize during apoptosis. *Cell Death Differ.* **21**, 196–205 (Feb. 2014).
152. Letai, A. *et al.* Distinct BH3 domains either sensitize or activate mitochondrial apoptosis, serving as prototype cancer therapeutics. *Cancer Cell* **2**, 183–192 (Sept. 2002).
153. Ku, B., Liang, C., Jung, J. U. & Oh, B.-H. Evidence that inhibition of BAX activation by BCL-2 involves its tight and preferential interaction with the BH3 domain of BAX. *Cell Res.* **21**, 627–641 (Apr. 2011).
154. Chi, X., Kale, J., Leber, B. & Andrews, D. W. Regulating cell death at, on, and in membranes. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 2100–2113 (Sept. 2014).
155. Shamas-Din, A. *et al.* tBid undergoes multiple conformational changes at the membrane required for Bax activation. *J. Biol. Chem.* **288**, 22111–27 (July 2013).
156. Reed, J. C. Bcl-2-family proteins and hematologic malignancies: history and future prospects. *Blood* **111**, 3322–3330 (Apr. 2008).
157. Fulda, S. Regulation of cell death in cancer-possible implications for immunotherapy. *Front. Oncol.* **3**, 29 (2013).
158. Dai, H. *et al.* Constitutive BAK activation as a determinant of drug sensitivity in malignant lymphohematopoietic cells. *Genes Dev.* **29**, 2140–52 (Oct. 2015).
159. Schellenberg, B. *et al.* Bax exists in a dynamic equilibrium between the cytosol and mitochondria to control apoptotic priming. *Mol. Cell* **49**, 959–71 (Mar. 2013).
160. Brumatti, G. *et al.* HoxA9 regulated Bcl-2 expression mediates survival of myeloid progenitors and the severity of HoxA9-dependent leukemia. *Oncotarget* **4**, 1933–1947 (Nov. 2013).
161. Escudero, S. *et al.* Dynamic Regulation of Long-Chain Fatty Acid Oxidation by a Noncanonical Interaction between the MCL-1 BH3 Helix and VLCAD. *Mol. Cell* **69**, 729–743.e7 (Mar. 2018).
162. Wu, X. *et al.* Extra-mitochondrial prosurvival BCL-2 proteins regulate gene transcription by inhibiting the SUFU tumour suppressor. *Nat. Cell Biol.* **19**, 1226–1236 (Sept. 2017).

163. Oltersdorf, T. *et al.* An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **435**, 677–681 (June 2005).
164. Mason, K. D. *et al.* Programmed Anuclear Cell Death Delimits Platelet Life Span. *Cell* **128**, 1173–1186 (Mar. 2007).
165. Zhang, H. *et al.* Bcl-2 family proteins are essential for platelet survival. *Cell Death Differ.* **14**, 943–951 (May 2007).
166. Cang, S., Iragavarapu, C., Savooji, J., Song, Y. & Liu, D. ABT-199 (venetoclax) and BCL-2 inhibitors in clinical development. *J. Hematol. Oncol.* **8**, 129 (Nov. 2015).
167. Konopleva, M. *et al.* Mechanisms of apoptosis sensitivity and resistance to the BH3 mimetic ABT-737 in acute myeloid leukemia. *Cancer Cell* **10**, 375–388 (Nov. 2006).
168. Del Gaizo Moore, V., Schlis, K. D., Sallan, S. E., Armstrong, S. A. & Letai, A. BCL-2 dependence and ABT-737 sensitivity in acute lymphoblastic leukemia. *Blood* **111**, 2300–2309 (Feb. 2008).
169. Wilson, W. H. *et al.* Navitoclax, a targeted high-affinity inhibitor of BCL-2, in lymphoid malignancies: a phase 1 dose-escalation study of safety, pharmacokinetics, pharmacodynamics, and antitumour activity. *Lancet Oncol.* **11**, 1149–1159 (Dec. 2010).
170. Ackler, S. *et al.* ABT-263 and rapamycin act cooperatively to kill lymphoma cells *in vitro* and *in vivo*. *Mol. Cancer Ther.* **7**, 3265–3274 (Oct. 2008).
171. Suryani, S. *et al.* Cell and Molecular Determinants of In Vivo Efficacy of the BH3 Mimetic ABT-263 against Pediatric Acute Lymphoblastic Leukemia Xenografts. *Clin. Cancer Res.* **20**, 4520–4531 (Sept. 2014).
172. Khaw, S. L. *et al.* Both leukaemic and normal peripheral B lymphoid cells are highly sensitive to the selective pharmacological inhibition of prosurvival Bcl-2 with ABT-199. *Leukemia* **28**, 1207–1215 (June 2014).
173. Nguyen, M. *et al.* Small molecule obatoclax (GX15-070) antagonizes MCL-1 and overcomes MCL-1-mediated resistance to apoptosis. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 19512–7 (Dec. 2007).
174. Schimmer, A. D. *et al.* A multicenter phase I/II study of obatoclax mesylate administered as a 3- or 24-hour infusion in older patients with previously untreated acute myeloid leukemia. *PLoS One* **9**, e108694 (2014).
175. O'Brien, S. *et al.* 5-Year Survival in Patients With Relapsed or Refractory Chronic Lymphocytic Leukemia in a Randomized, Phase III Trial of Fludarabine Plus Cyclophosphamide With or Without Oblimersen. *J. Clin. Oncol.* **27**, 5208–5212 (Nov. 2009).
176. Souers, A. J. *et al.* ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat. Med.* **19**, 202–208 (Feb. 2013).
177. Roberts, A. W. *et al.* Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **374**, 311–322 (Jan. 2016).
178. Croce, C. M. & Reed, J. C. Finally, An Apoptosis-Targeting Therapeutic for Cancer. *Cancer Res.* **76**, 5914–5920 (Oct. 2016).

179. Seymour, J. F. *et al.* VenetoclaxRituximab in Relapsed or Refractory Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **378**, 1107–1120 (Mar. 2018).
180. Fischer, K. *et al.* Venetoclax and obinutuzumab in chronic lymphocytic leukemia. *Blood* **129**, 2702–2705 (May 2017).
181. Jain, N. *et al.* Combined Venetoclax and Ibrutinib for Patients with Previously Untreated High-Risk CLL, and Relapsed/Refractory CLL: A Phase II Trial. *Blood* **130** (2017).
182. Rogers, K. A. *et al.* Initial Results of the Phase 2 Treatment Naïve Cohort in a Phase 1b/2 Study of Obinutuzumab, Ibrutinib, and Venetoclax in Chronic Lymphocytic Leukemia. *Blood* **130** (2017).
183. Valentin, R., Grabow, S. & Davids, M. S. The rise of apoptosis: targeting apoptosis in hematologic malignancies. *Blood*, blood–2018–02–791350 (July 2018).
184. Byrd, J. C. *et al.* The mechanism of tumor cell clearance by rituximab in vivo in patients with B-cell chronic lymphocytic leukemia: evidence of caspase activation and apoptosis induction. *Blood* **99**, 1038–43 (Feb. 2002).
185. Bose, P., Gandhi, V. & Konopleva, M. Pathways and mechanisms of venetoclax resistance. *Leuk. Lymphoma* **58**, 1–17 (2017).
186. McNeer, N. A. *et al.* Genetic mechanisms of primary chemotherapy resistance in pediatric acute myeloid leukemia. *Leukemia* **33**, 1934–1943 (Aug. 2019).
187. Konopleva, M. *et al.* A Phase 2 Study of ABT-199 (GDC-0199) in Patients with Acute Myelogenous Leukemia (AML). *Blood* **124** (2014).
188. Pan, R. *et al.* Selective BCL-2 inhibition by ABT-199 causes on-target cell death in acute myeloid leukemia. *Cancer Discov.* **4**, 362–75 (Mar. 2014).
189. Robertson, L. E., Plunkett, W., McConnell, K., Keating, M. J. & McDonnell, T. J. Bcl-2 expression in chronic lymphocytic leukemia and its correlation with the induction of apoptosis and clinical outcome. *Leukemia* **10**, 456–9 (Mar. 1996).
190. Knight, T., Luedtke, D., Edwards, H., Taub, J. W. & Ge, Y. A delicate balance The BCL-2 family and its role in apoptosis, oncogenesis, and cancer therapeutics. *Biochem. Pharmacol.* **162**, 250–261 (Apr. 2019).
191. Niu, X. *et al.* Binding of Released Bim to Mcl-1 is a Mechanism of Intrinsic Resistance to ABT-199 which can be Overcome by Combination with Daunorubicin or Cytarabine in AML Cells. *Clin. Cancer Res.* **22**, 4440–4451 (Sept. 2016).
192. Vogler, M. *et al.* Concurrent up-regulation of BCL-XL and BCL2A1 induces approximately 1000-fold resistance to ABT-737 in chronic lymphocytic leukemia. *Blood* **113**, 4403–13 (Apr. 2009).
193. Herling, C. D. *et al.* Clonal dynamics towards the development of venetoclax resistance in chronic lymphocytic leukemia. *Nat. Commun.* **9**, 727 (Dec. 2018).
194. Lin, K. H. *et al.* Targeting MCL-1/BCL-XL Forestalls the Acquisition of Resistance to ABT-199 in Acute Myeloid Leukemia. *Sci. Rep.* **6**, 27696 (Sept. 2016).
195. Mali, R. S. *et al.* FLT3-ITD Activation Mediates Resistance to the BCL-2 Selective Antagonist, Venetoclax, in FLT3-ITD Mutant AML Models. *Blood* **130** (2017).

196. Gale, R. E. *et al.* The impact of FLT3 internal tandem duplication mutant level, number, size, and interaction with NPM1 mutations in a large cohort of young adult patients with acute myeloid leukemia. doi:10.1182/blood-2007-08 (2008).
197. Heath, E. M. *et al.* Biological and clinical consequences of NPM1 mutations in AML. *Leukemia* **31**, 798–807 (Apr. 2017).
198. Han, L. *et al.* Concomitant targeting of BCL2 with venetoclax and MAPK signaling with cobimetinib in acute myeloid leukemia models. *Haematologica*, haematol.2018.205534 (May 2019).
199. Pan, R. *et al.* Synthetic Lethality of Combined Bcl-2 Inhibition and p53 Activation in AML: Mechanisms and Superior Antileukemic Efficacy. *Cancer Cell* **32**, 748–760.e6 (Dec. 2017).
200. Benito, J. *et al.* MLL-Rearranged Acute Lymphoblastic Leukemias Activate BCL-2 through H3K79 Methylation and Are Sensitive to the BCL-2-Specific Antagonist ABT-199. *Cell Rep.* **13**, 2715–2727 (Dec. 2015).
201. Khaw, S. L. *et al.* Venetoclax responses of pediatric ALL xenografts reveal sensitivity of MLL-rearranged leukemia. *Blood* **128**, 1382–1395 (Sept. 2016).
202. Stein, E. M. & Tallman, M. S. Emerging therapeutic drugs for AML. *Blood* **127**, 71–8 (Jan. 2016).
203. Steinhilber, D. & Marschalek, R. How to effectively treat acute leukemia patients bearing MLL-rearrangements ? *Biochem. Pharmacol.* **147**, 183–190 (Jan. 2018).
204. Wouters, B. J. & Delwel, R. Epigenetics and approaches to targeted epigenetic therapy in acute myeloid leukemia. *Blood* **127**, 42–52 (Jan. 2016).
205. DiNardo, C. D. *et al.* Safety and preliminary efficacy of venetoclax with decitabine or azacitidine in elderly patients with previously untreated acute myeloid leukaemia: a non-randomised, open-label, phase 1b study. *Lancet Oncol.* **19**, 216–228 (Feb. 2018).
206. DiNardo, C. D. *et al.* Venetoclax combined with decitabine or azacitidine in treatment-naïve, elderly patients with acute myeloid leukemia. *Blood* **133**, 7–17 (2019).
207. Tsao, T. *et al.* Concomitant inhibition of DNA methyltransferase and BCL-2 protein function synergistically induce mitochondrial apoptosis in acute myelogenous leukemia cells. *Ann. Hematol.* **91**, 1861–1870 (Dec. 2012).
208. Hurtz, C. *et al.* Rationale for targeting BCL6 in MLL-rearranged acute lymphoblastic leukemia. *Genes Dev.* **33**, 1265–1279 (Sept. 2019).
209. Kontro, M. *et al.* HOX gene expression predicts response to BCL-2 inhibition in acute myeloid leukemia. *Leukemia* **31**, 301–309 (Feb. 2017).
210. Karjalainen, R. *et al.* Elevated expression of S100A8 and S100A9 correlates with resistance to the BCL-2 inhibitor venetoclax in AML. *Leukemia*, 1 (June 2019).
211. Menges, C. W., Altomare, D. A. & Testa, J. R. FAS-Associated Factor 1 (FAF1): Diverse functions and implications for oncogenesis. *Cell Cycle* **8**, 2528–2534 (Aug. 2009).

212. O'Neill, K. L., Huang, K., Zhang, J., Chen, Y. & Luo, X. Inactivation of prosurvival Bcl-2 proteins activates Bax/Bak through the outer mitochondrial membrane. *Genes Dev.* **30**, 973–88 (Apr. 2016).
213. Greaves, G. *et al.* BH3-only proteins are dispensable for apoptosis induced by pharmacological inhibition of both MCL-1 and BCL-XL. *Cell Death Differ.* **1** (Sept. 2018).
214. Cimmino, A. *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13944–9 (Sept. 2005).
215. Mrquez-Jurado, S. *et al.* Mitochondrial levels determine variability in cell death by modulating apoptotic gene expression. *Nat. Commun.* **9**, 389 (Dec. 2018).
216. Santos, L. C. *et al.* Mitochondrial origins of fractional control in regulated cell death. *Nat. Commun.* **10**, 1313 (Dec. 2019).
217. Kotschy, A. *et al.* The MCL1 inhibitor S63845 is tolerable and effective in diverse cancer models. *Nature* **538**, 477–482 (Oct. 2016).
218. Ramsey, H. E. *et al.* A Novel MCL-1 Inhibitor Combined with Venetoclax Rescues Venetoclax Resistant Acute Myelogenous Leukemia. *Cancer Discov.* CD-18-0140 (Sept. 2018).
219. Frankish, A. *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773 (Jan. 2019).
220. Gurumurthy, C. B. *et al.* CRISPR: a versatile tool for both forward and reverse genetics research. *Hum. Genet.* **135**, 971–976 (Sept. 2016).
221. Patton, E. E. & Zon, L. I. The art and design of genetic screens: zebrafish. *Nat. Rev. Genet.* **2**, 956–966 (Dec. 2001).
222. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–11 (Feb. 1998).
223. Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* **10**, 94–108 (Feb. 2009).
224. Dorsett, Y. & Tuschl, T. siRNAs: applications in functional genomics and potential as therapeutics. *Nat. Rev. Drug Discov.* **3**, 318–329 (Apr. 2004).
225. Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* **16**, 299–311 (May 2015).
226. Adli, M. The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* **9**, 1911 (Dec. 2018).
227. Gaj, T., Gersbach, C. A., Barbas, C. F. & III. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405 (July 2013).
228. Root, D. E., Hacohen, N., Hahn, W. C., Lander, E. S. & Sabatini, D. M. Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat. Methods* **3**, 715–719 (Sept. 2006).
229. Paddison, P. J. *et al.* A resource for large-scale RNA-interference-based screens in mammals. *Nature* **428**, 427–431 (Mar. 2004).

230. Moffat, J. & Sabatini, D. M. Building mammalian signalling pathways with RNAi screens. *Nat. Rev. Mol. Cell Biol.* **7**, 177–187 (Mar. 2006).
231. Evers, B. *et al.* CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat. Biotechnol.* **34**, 631–633 (June 2016).
232. Housden, B. E. & Perrimon, N. Comparing CRISPR and RNAi-based screening technologies. *Nat. Biotechnol.* **34**, 621–623 (June 2016).
233. Shum, K. & Rossi, J. *SiRNA Delivery Methods* (eds Shum, K. & Rossi, J.) doi:10.1007/978-1-4939-3112-5 (Springer New York, New York, NY, 2016).
234. Lino, C. A., Harper, J. C., Carney, J. P. & Timlin, J. A. Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv.* **25**, 1234–1257 (Nov. 2018).
235. Ramezani, A. & Hawley, R. G. in *Curr. Protoc. Mol. Biol.* 1, 16.21.1–16.21.15 (John Wiley & Sons, Inc., Hoboken, NJ, USA, Nov. 2002). doi:10.1002/0471142727.mb1621s60.
236. Naldini, L. *et al.* In Vivo Gene Delivery and Stable Transduction of Nondividing Cells by a Lentiviral Vector. *Science (80-.)*. **272**, 263–267 (Apr. 1996).
237. Leander Johansen, J., Dag, L., Torne, J., Rosenblad, C. & Kusk, P. A New Versatile and Compact Lentiviral Vector. *Mol. Biotechnol.* **29**, 47–56 (2005).
238. Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L. & Trono, D. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat. Biotechnol.* **15**, 871–875 (Sept. 1997).
239. Kabadi, A. M., Ousterout, D. G., Hilton, I. B. & Gersbach, C. A. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. *Nucleic Acids Res.* **42**, e147–e147 (Oct. 2014).
240. Ilg, T. Investigations on the molecular mode of action of the novel immunostimulator ZelNate: Activation of the cGAS-STING pathway in mammalian cells. *Mol. Immunol.* **90**, 182–189 (Oct. 2017).
241. Doench, J. G. Am I ready for CRISPR? A user's guide to genetic screens. *Nat. Rev. Genet.* **19**, 67–80 (Dec. 2017).
242. Parnas, O. *et al.* A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. *Cell* **162**, 675–686 (July 2015).
243. Joung, J. *et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12**, 828–863 (Mar. 2017).
244. Schuster, A. *et al.* RNAi/CRISPR Screens: from a Pool to a Valid Hit. *Trends Biotechnol.* doi:10.1016/J.TIBTECH.2018.08.002 (Aug. 2018).
245. Park, R. J. *et al.* A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors. *Nat. Genet.* **49**, 193–203 (Feb. 2017).
246. Basheer, F. T. & Vassiliou, G. S. Genome-scale drop-out screens to identify cancer cell vulnerabilities in AML. *Curr. Opin. Genet. Dev.* **54**, 83–87 (Feb. 2019).
247. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–4 (Jan. 2014).
248. Wang, T. *et al.* Identification and characterization of essential genes in the human genome. *Science (80-.)*. **350** (2015).

249. Tzelepis, K. *et al.* A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. *Cell Rep.* **17**, 1193–1205 (Oct. 2016).
250. Erb, M. A. *et al.* Transcription control by the ENL YEATS domain in acute leukaemia. *Nature* **543**, 270–274 (Mar. 2017).
251. Wallace, J. *et al.* Genome-Wide CRISPR-Cas9 Screen Identifies MicroRNAs That Regulate Myeloid Leukemia Cell Growth. *PLoS One* **11**, e0153689 (2016).
252. Kurata, M., Yamamoto, K., Moriarity, B. S., Kitagawa, M. & Largaespada, D. A. CRISPR/Cas9 library screening for drug target discovery. *J. Hum. Genet.* **63**, 179–186 (Feb. 2018).
253. Zender, L. *et al.* An Oncogenomics-Based In Vivo RNAi Screen Identifies Tumor Suppressors in Liver Cancer. *Cell* **135**, 852–864 (Nov. 2008).
254. Rudalska, R. *et al.* In vivo RNAi screening identifies a mechanism of sorafenib resistance in liver cancer. *Nat. Med.* **20**, 1138–1146 (Oct. 2014).
255. Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science (80-.)*. **343** (2014).
256. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (Dec. 2014).
257. Heckl, D. *et al.* Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. *Nat. Biotechnol.* **32**, 941–946 (Sept. 2014).
258. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (Aug. 2014).
259. Dmitriev, D. A. & Rakitov, R. A. Decoding of Superimposed Traces Produced by Direct Sequencing of Heterozygous Indels. *PLoS Comput. Biol.* **4** (ed Stormo, G.) e1000113 (July 2008).
260. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168 (Dec. 2014).
261. Ianevski, A., He, L., Aittokallio, T. & Tang, J. SynergyFinder: a web application for analyzing drug combination doseresponse matrix data. *Bioinformatics* **33** (ed Stegle, O.) 2413–2415 (Aug. 2017).
262. Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J.-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**, 1012–9 (June 2014).
263. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* **44**, W272–W276 (July 2016).
264. Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* doi:10.1093/nar/gkz365 (May 2019).
265. Brattelid, T. *et al.* Normalization strategy is critical for the outcome of miRNA expression analyses in the rat heart. *Physiol. Genomics* **43**, 604–610 (May 2011).

- 266. Bustin, S. A. *et al.* The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* **55** (2009).
- 267. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, research0034.1 (June 2002).
- 268. Michaels, Y. S. *et al.* Precise tuning of gene expression levels in mammalian cells. *Nat. Commun.* **10**, 818 (Dec. 2019).
- 269. Campeau, E. *et al.* A Versatile Viral System for Expression and Depletion of Proteins in Mammalian Cells. *PLoS One* **4** (ed Fritz, J. H.) e6529 (Aug. 2009).
- 270. Ellis, E. L. & Delbrck, M. THE GROWTH OF BACTERIOPHAGE. *J. Gen. Physiol.* **22**, 365–84 (Jan. 1939).
- 271. Hart, T. *et al.* Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3 (Bethesda)*. **7**, 2719–2727 (Aug. 2017).
- 272. Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (Oct. 2016).
- 273. Canver, M. C. *et al.* Integrated design, execution, and analysis of arrayed and pooled CRISPR genome-editing experiments. *Nat. Protoc.* **13**, 946–986 (Apr. 2018).
- 274. Aregger, M., Chandrashekar, M., Tong, A. H. Y., Chan, K. & Moffat, J. in *Methods Mol. Biol.* 169–188 (2019). doi:10.1007/978-1-4939-8805-1_15.
- 275. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (May 2011).
- 276. Gillooly, J. F., Hein, A. & Damiani, R. Nuclear DNA Content Varies with Cell Size across Human Cell Types. *Cold Spring Harb. Perspect. Biol.* **7**, a019091 (July 2015).
- 277. Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).
- 278. Meyers, R. M. *et al.* Computational correction of copy number effect improves specificity of CRISPR/Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (Oct. 2017).
- 279. Wang, B. *et al.* Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. *Nat. Protoc.* **14**, 756–780 (Mar. 2019).
- 280. Li, W. *et al.* Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol.* **16**, 281 (Dec. 2015).
- 281. Ge, S. X. & Jung, D. ShinyGO: a graphical enrichment tool for animals and plants. *bioRxiv*, 315150 (May 2018).
- 282. Ge, S. X., Son, E. W. & Yao, R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **19**, 534 (Dec. 2018).
- 283. Flyamer, I. *et al.* Phyla/adjustText. doi:10.5281/ZENODO.1494343 (Nov. 2018).

284. Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **47**, D155–D162 (Jan. 2019).
285. Chang, H. *et al.* CRISPR/cas9, a novel genomic tool to knock down microRNA in vitro and in vivo. *Sci. Rep.* **6**, 22312 (Mar. 2016).
286. Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **17**, 148 (Dec. 2016).
287. Suzuki, Y. & Suzuki, Y. in *Viral Gene Ther.* (InTech, July 2011). doi:10.5772/18155.
288. Canver, M. C. *et al.* Variant-aware saturating mutagenesis using multiple Cas9 nucleases identifies regulatory elements at trait-associated loci. *Nat. Genet.* **49**, 625–634 (Feb. 2017).
289. Allen, F. *et al.* Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat. Biotechnol.* **37**, 64–72 (Nov. 2018).
290. Chen, X. *et al.* Dual sgRNA-directed gene knockout using CRISPR/Cas9 technology in *Caenorhabditis elegans*. *Sci. Rep.* **4**, 7581 (May 2015).
291. Adikusuma, F., Pfitzner, C. & Thomas, P. Q. Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression. *PLoS One* **12** (ed Wang, T. T.) e0187236 (Dec. 2017).
292. Panwar, B., Omenn, G. S. & Guan, Y. miRmine: A Database of Human miRNA Expression Profiles. *Bioinformatics* **44**, btx019 (Jan. 2017).
293. Pereira, C., Clarke, E. & Damen, J. in, 177–208 (Humana Press, 2007). doi:10.1007/978-1-59745-536-7_14.
294. Mullokandov, G. *et al.* High-throughput assessment of microRNA activity and function using microRNA sensor and decoy libraries. *Nat. Methods* **9**, 840–846 (July 2012).
295. Strasser, A., Jost, P. J. & Nagata, S. The many roles of FAS receptor signaling in the immune system. *Immunity* **30**, 180–92 (Feb. 2009).
296. Hong Min, Y. *et al.* Expression of Fas antigen in acute myeloid leukaemia is associated with therapeutic response to chemotherapy. *Br. J. Haematol.* **93**, 928–930 (June 1996).
297. Min, Y. J. *et al.* Prognostic significance of Fas (CD95) and TRAIL receptors (DR4/DR5) expression in acute myelogenous leukemia. *Leuk. Res.* **28**, 359–365 (Apr. 2004).
298. Suminoe, A. *et al.* mRNA expression of apoptosis-associated genes in infant acute lymphoblastic leukemia: low Fas expression is an independent predictor for poor prognosis. *Leukemia* **18**, 365–368 (Feb. 2004).
299. Zhang, Z., Ursin, R., Mahapatra, S. & Gallicano, G. I. CRISPR/CAS9 ablation of individual miRNAs from a miRNA family reveals their individual efficacies for regulating cardiac differentiation. *Mech. Dev.* **150**, 10–20 (Apr. 2018).
300. Raab, N. *et al.* CRISPR/Cas9 Mediated Knockout of MicroRNA744 Improves Antibody Titer of CHO Production Cell Lines. *Biotechnol. J.* **14**, 1800477 (May 2019).

- 301. Barski, A. *et al.* High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* **129**, 823–837 (May 2007).
- 302. Chien, C.-H. *et al.* Identifying transcriptional start sites of human microRNAs based on high-throughput sequencing data. *Nucleic Acids Res.* **39**, 9345–56 (Nov. 2011).
- 303. Ebert, M. S. & Sharp, P. A. MicroRNA sponges: progress and possibilities. *RNA* **16**, 2043–50 (Nov. 2010).
- 304. Krutzfeldt, J. *et al.* Silencing of microRNAs in vivo with ‘antagomirs’. *Nature* **438**, 685–689 (Dec. 2005).
- 305. Cheng, T., Rodrigues, N., Dombkowski, D., Stier, S. & Scadden, D. T. Stem cell repopulation efficiency but not pool size is governed by p27kip1. *Nat. Med.* **6**, 1235–1240 (Nov. 2000).
- 306. Wang, X., Gorospe, M., Huang, Y. & Holbrook, N. J. p27Kip1 overexpression causes apoptotic death of mammalian cells. *Oncogene* **15**, 2991–2997 (Dec. 1997).
- 307. Bassett, A. R. *et al.* Understanding functional miRNA target interactions in vivo by site-specific genome engineering. *Nat. Commun.* **5**, 4640 (Dec. 2014).
- 308. Wu, Q. *et al.* In situ functional dissection of RNA cis-regulatory elements by multiplex CRISPR-Cas9 genome engineering. *Nat. Commun.* **8**, 2109 (Dec. 2017).
- 309. Price, C. & Chen, J. MicroRNAs in cancer biology and therapy: Current status and perspectives. *Genes Dis.* **1**, 53–63 (Sept. 2014).
- 310. Haetscher, N. *et al.* STAT5-regulated microRNA-193b controls haematopoietic stem and progenitor cell expansion by modulating cytokine receptor signalling. *Nat. Commun.* **6**, 8928 (Dec. 2015).
- 311. Bhayadia, R. *et al.* Endogenous Tumor Suppressor microRNA-193b: Therapeutic and Prognostic Value in Acute Myeloid Leukemia. *J. Clin. Oncol.* JCO2017752204 (Feb. 2018).
- 312. Jiang, X. *et al.* miR-22 has a potent anti-tumour role with therapeutic potential in acute myeloid leukaemia. *Nat. Commun.* **7**, 11452 (Sept. 2016).
- 313. He, Y., Jiang, X. & Chen, J. The role of miR-150 in normal and malignant hematopoiesis. *Oncogene* **33**, 3887–3893 (July 2014).
- 314. Li, Z. *et al.* Overexpression and knockout of miR-126 both promote leukemogenesis. *Blood* **126**, 2005–15 (Oct. 2015).
- 315. Chen, P. *et al.* miR-9 is an essential oncogenic microRNA specifically overexpressed in mixed lineage leukemia-rearranged leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11511–6 (July 2013).
- 316. Li, Z. *et al.* Molecular and Cellular Pathobiology PBX3 and MEIS1 Cooperate in Hematopoietic Cells to Drive Acute Myeloid Leukemias Characterized by a Core Transcriptome of the MLL-Rearranged Disease. doi:10.1158/0008-5472.CAN-15-1566 (2016).
- 317. Li, Z. *et al.* PBX3 is an important cofactor of HOXA9 in leukemogenesis. *Blood* **121**, 1422–1431 (Feb. 2013).

318. Pla, A., Zhong, X. & Rayner, S. miRAW: A deep learning-based approach to predict microRNA targets by analyzing whole microRNA transcripts. *PLOS Comput. Biol.* **14** (ed Ioshikhes, I.) e1006185 (July 2018).
319. Tokar, T. *et al.* mirDIP 4.1-integrative database of human microRNA target predictions. *Nucleic Acids Res.* **46**, D360–D370 (2018).
320. Seitz, H. Issues in current microRNA target identification methods. *RNA Biol.* **14**, 831–834 (2017).
321. Helwak, A., Kudla, G., Dudnakova, T. & Tollervey, D. Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding. *Cell* **153**, 654–665 (Apr. 2013).
322. Lenoir, W. F., Lim, T. L. & Hart, T. PICKLES: the database of pooled in-vitro CRISPR knockout library essentiality screens. *Nucleic Acids Res.* **46**, D776–D780 (Jan. 2018).
323. Niu, X. *et al.* Acute myeloid leukemia cells harboring MLL fusion genes or with the acute promyelocytic leukemia phenotype are sensitive to the Bcl-2-selective inhibitor ABT-199. *Leukemia* **28**, 1557–60 (July 2014).
324. Ben-David, U. *et al.* Genetic and transcriptional evolution alters cancer cell line drug response. *Nature* **560**, 325–330 (Aug. 2018).
325. Liu, Y. *et al.* Multi-omic measurements of heterogeneity in HeLa cells across laboratories. *Nat. Biotechnol.* **37**, 314–322 (Mar. 2019).
326. EMA Committee for Medicinal Products for Human Use (CHMP). *Assessment report EMA/CHMP/717199/2018 (Venclyxto)* tech. rep. (2018).
327. Luedtke, D. A. *et al.* Inhibition of Mcl-1 enhances cell death induced by the Bcl-2-selective inhibitor ABT-199 in acute myeloid leukemia cells. *Signal Transduct. Target. Ther.* **2**, 17012 (Apr. 2017).
328. Schwartz, J. *et al.* Synergistic anti-leukemic interactions between ABT-199 and panobinostat in acute myeloid leukemia *ex vivo*. *Am. J. Transl. Res.* **8**, 3893–3902 (2016).
329. Bui, M. H. *et al.* Preclinical Characterization of BET Family Bromodomain Inhibitor ABBV-075 Suggests Combination Therapeutic Strategies. *Cancer Res.* **77**, 2976–2989 (June 2017).
330. Liu, X. *et al.* Editing DNA Methylation in the Mammalian Genome. *Cell* **167**, 233–247.e17 (2016).
331. Levenson, J. D. *et al.* Potent and selective small-molecule MCL-1 inhibitors demonstrate on-target cancer cell killing activity as single agents and in combination with ABT-263 (navitoclax). *Cell Death Dis.* **6**, e1590–e1590 (Jan. 2015).
332. Xiang, W., Yang, C.-Y. & Bai, L. MCL-1 inhibition in cancer treatment. *Onco. Targets. Ther.* **11**, 7301–7314 (2018).
333. Tahir, S. K. *et al.* Potential mechanisms of resistance to venetoclax and strategies to circumvent it. *BMC Cancer* **17**, 399 (Dec. 2017).

- 334. Fresquet, V., Rieger, M., Carolis, C., Garca-Barchino, M. J. & Martinez-Climent, J. A. Acquired mutations in BCL2 family proteins conferring resistance to the BH3 mimetic ABT-199 in lymphoma. doi:10.1182/blood-2014-03-560284 (2014).
- 335. Delbridge, A. R. *et al.* The BH3-only proteins BIM and PUMA are not critical for the reticulocyte apoptosis caused by loss of the pro-survival protein BCL-XL. *Cell Death Dis.* **8**, e2914–e2914 (July 2017).
- 336. Yecies, D., Carlson, N. E., Deng, J. & Letai, A. Acquired resistance to ABT-737 in lymphoma cells that up-regulate MCL-1 and BFL-1. *Blood* **115**, 3304 (2010).
- 337. Esteve-Arenys, A. & Roue, G. BFL-1 expression determines the efficacy of venetoclax in MYC+/BCL2+ double hit lymphoma. *Oncoscience* **5**, 59–61 (Mar. 2018).
- 338. Ong, S. H., Li, Y., Koike-Yusa, H. & Yusa, K. Optimised metrics for CRISPR-KO screens with second-generation gRNA libraries. *Sci. Rep.* **7**, 7384 (Dec. 2017).
- 339. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (Feb. 2016).
- 340. Hart, T. *et al.* High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* **163**, 1515–26 (Dec. 2015).
- 341. Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9mediated gene inactivation. *Nat. Biotechnol.* **32**, 1262–1267 (Dec. 2014).
- 342. Horlbeck, M. A. *et al.* Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife* **5**. doi:10.7554/eLife.12677 (Mar. 2016).
- 343. vanOverbeek, M. *et al.* DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. *Mol. Cell*. doi:10.1016/j.molcel.2016.06.037 (2016).
- 344. Strezoska, a. *et al.* Optimized PCR Conditions and Increased shRNA Fold Representation Improve Reproducibility of Pooled shRNA Screens. *PLoS One* **7** (ed Jeang, K. T.) e42341 (Aug. 2012).
- 345. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (Jan. 2010).
- 346. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (Oct. 2010).
- 347. Luo, B. *et al.* Highly parallel identification of essential genes in cancer cells. *Proc. Natl. Acad. Sci.* **105**, 20380–20385 (Dec. 2008).
- 348. Knig, R. *et al.* A probability-based approach for the analysis of large-scale RNAi screens. *Nat. Methods* **4**, 847–849 (Oct. 2007).
- 349. Winter, J. *et al.* caRnpools: an R package for exploratory data analysis and documentation of pooled CRISPR/Cas9 screens. *Bioinformatics* **32**, 632–634 (Feb. 2016).
- 350. Winter, J. *et al.* CRISPRAnalyzeR: Interactive analysis, annotation and documentation of pooled CRISPR screens. *bioRxiv*, 109967 (Feb. 2017).

351. Spahn, P. N. *et al.* PinAPL-Py: A comprehensive web-application for the analysis of CRISPR/Cas9 screens. *Sci. Rep.* **7**, 15854 (Dec. 2017).
352. Trmbach, D. *et al.* ENCoRE: an efficient software for CRISPR screens identifies new players in extrinsic apoptosis. *BMC Genomics* **18**, 905 (Dec. 2017).
353. You, Q. *et al.* CRISPRMatch: An Automatic Calculation and Visualization Tool for High-throughput CRISPR Genome-editing Data Analysis. *Int. J. Biol. Sci.* **14**, 858–862 (2018).
354. Allen, F. *et al.* JACKS: joint analysis of CRISPR/Cas9 knockout screens. *Genome Res.* **29**, 464–471 (Mar. 2019).
355. Di, Y., Schafer, D. W., Cumbie, J. S. & Chang, J. H. The NBP Negative Binomial Model for Assessing Differential Gene Expression from RNA-Seq. *Stat. Appl. Genet. Mol. Biol.* **10**. doi:10.2202/1544-6115.1637 (Jan. 2011).
356. Kolde, R., Laur, S., Adler, P. & Vilo, J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* **28**, 573–580 (Feb. 2012).
357. Hart, T. *et al.* Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3* **7**, 2719–2727 (Aug. 2017).
358. Oprea, T. I. *et al.* Unexplored therapeutic opportunities in the human genome. *Nat. Rev. Drug Discov.* **17**, 317–332 (May 2018).
359. Cotto, K. C. *et al.* DGIdb 3.0: a redesign and expansion of the drug-gene interaction database. *Nucleic Acids Res.* **46**, D1068–D1073 (Jan. 2018).
360. Liu, H. *et al.* Proteasome inhibitors evoke latent tumor suppression programs in pro-B MLL leukemias through MLL-AF4. *Cancer Cell* **25**, 530–42 (Apr. 2014).
361. Sanson, K. R. *et al.* Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416 (Dec. 2018).
362. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (Feb. 2015).
363. Campenhout, C. V. *et al.* Towards best-practice approaches for CRISPR/Cas9 gene engineering. *bioRxiv*, 469544 (Dec. 2018).
364. McKenna, A. & Shendure, J. FlashFry: a fast and flexible tool for large-scale CRISPR target design. *BMC Biol.* **16**, 74 (Dec. 2018).
365. Zimmermann, M. *et al.* CRISPR screens identify genomic ribonucleotides as a source of PARP-trapping lesions. *Nature* **559**, 285–289 (July 2018).
366. Liu, J. *et al.* A genome-scale CRISPR-Cas9 screening in myeloma cells identifies regulators of immunomodulatory drug sensitivity. *Leukemia*, 1 (July 2018).
367. Jeong, H.-H., Kim, S. Y., Rousseaux, M. W. C., Zoghbi, H. Y. & Liu, Z. Beta-binomial modeling of CRISPR pooled screen data identifies target genes with greater sensitivity and fewer false negatives. *Genome Res.* doi:10.1101/gr.245571.118 (Apr. 2019).
368. Aguirre, A. J. *et al.* Genomic Copy Number Dictates a Gene-Independent Cell Response to CRISPR/Cas9 Targeting. *Cancer Discov.* **6**, 914–29 (Aug. 2016).
369. Munoz, D. M. *et al.* CRISPR Screens Provide a Comprehensive Assessment of Cancer Vulnerabilities but Generate False-Positive Hits for Highly Amplified Genomic Regions. *Cancer Discov.* **6**, 900–13 (Aug. 2016).

- 370. De Weck, A. *et al.* Correction of copy number induced false positives in CRISPR screens. *PLOS Comput. Biol.* **14** (ed Raphael, B. J.) e1006279 (July 2018).
- 371. Iorio, F. *et al.* Unsupervised correction of gene-independent cell responses to CRISPR-Cas9 targeting. *BMC Genomics* **19**, 604 (Dec. 2018).
- 372. Colic, M. *et al.* Identifying chemogenetic interactions from CRISPR screens with drugZ. *Genome Med.* **11**, 52 (Dec. 2019).
- 373. Hart, T. & Moffat, J. BAGEL: a computational framework for identifying essential genes from pooled library screens. *BMC Bioinformatics* **17**, 164 (Dec. 2016).
- 374. Behan, F. M. *et al.* Prioritization of cancer therapeutic targets using CRISPRCas9 screens. *Nature*, 1 (Apr. 2019).
- 375. Khatry, P. & Draghici, S. Ontological analysis of gene expression data: current tools, limitations, and open problems. *Bioinformatics* **21**, 3587–3595 (Sept. 2005).
- 376. Chin, H. S. *et al.* VDAC2 enables BAX to mediate apoptosis and limit tumor development. *Nat. Commun.* **9**, 4976 (Dec. 2018).
- 377. Adams, J. The proteasome: a suitable antineoplastic target. *Nat. Rev. Cancer* **4**, 349–360 (May 2004).
- 378. Navon, A. & Ciechanover, A. The 26 S proteasome: from basic mechanisms to drug targeting. *J. Biol. Chem.* **284**, 33713–8 (Dec. 2009).
- 379. Drexler, H. Activation of the cell death program by inhibition of proteasomefunction. *Proc. Natl. Acad. Sci.* **94**, 855–860 (Feb. 1997).
- 380. Masdehors, P. *et al.* Increased sensitivity of CLL-derived lymphocytes to apoptotic death activation by the proteasome-specific inhibitor lactacystin. *Br. J. Haematol.* **105**, 752–757 (June 1999).
- 381. Guzman, M. L. *et al.* Preferential induction of apoptosis for primary human leukemic stem cells. *Proc. Natl. Acad. Sci.* **99**, 16220–16225 (Dec. 2002).
- 382. Punnoose, E. A. *et al.* Expression Profile of BCL-2, BCL-XL, and MCL-1 Predicts Pharmacological Response to the BCL-2 Selective Antagonist Venetoclax in Multiple Myeloma Models. *Mol. Cancer Ther.* **15**, 1132–1144 (May 2016).
- 383. Csizmar, C. M., Kim, D.-H. & Sachs, Z. The role of the proteasome in AML. *Blood Cancer J.* **6**, e503 (2016).
- 384. Qin, J.-Z. *et al.* Proteasome Inhibitors Trigger NOXA-Mediated Apoptosis in Melanoma and Myeloma Cells. *Cancer Res.* **65**, 6282–6293 (July 2005).
- 385. Matulis, S. M. *et al.* Dexamethasone treatment promotes Bcl-2 dependence in multiple myeloma resulting in sensitivity to venetoclax. *Leukemia* **30**, 1086–1093 (May 2016).
- 386. Rahbari, K. J., Nosrati, J. D., Spektor, T. M. & Berenson, J. R. Venetoclax in Combination With Bortezomib, Dexamethasone, and Daratumumab for Multiple Myeloma. *Clin. Lymphoma. Myeloma Leuk.* **18**, e339–e343 (Sept. 2018).
- 387. Moreau, P. *et al.* Promising efficacy and acceptable safety of venetoclax plus bortezomib and dexamethasone in relapsed/refractory MM. *Blood* **130**, 2392–2400 (Nov. 2017).

388. Grant, S. Rational combination strategies to enhance venetoclax activity and overcome resistance in hematologic malignancies. *Leuk. Lymphoma* **59**, 1292–1299 (June 2018).
389. Cruickshank, M. N. *et al.* Systematic chemical and molecular profiling of MLL-rearranged infant acute lymphoblastic leukemia reveals efficacy of romidepsin. *Leukemia* **31**, 40–50 (Jan. 2017).
390. Yadav, B., Wennerberg, K., Aittokallio, T. & Tang, J. Searching for Drug Synergy in Complex DoseResponse Landscapes Using an Interaction Potency Model. *Comput. Struct. Biotechnol. J.* **13**, 504–513 (Jan. 2015).
391. Oerlemans, R. *et al.* Molecular basis of bortezomib resistance: proteasome subunit beta5 (PSMB5) gene mutation and overexpression of PSMB5 protein. *Blood* **112**, 2489–99 (Sept. 2008).
392. Myhrvold, I. K. *et al.* Single cell profiling of phospho-protein levels in chronic lymphocytic leukemia. *Oncotarget* **9**, 9273–9284 (Feb. 2018).
393. Sknland, S. S. *et al.* An in vitro assay for biomarker discovery and dose prediction applied to ibrutinib plus venetoclax treatment of CLL. *Leukemia*, 1–10 (Aug. 2019).
394. McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G. & Muir, T. W. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **453**, 812–816 (June 2008).
395. Shilatifard, A. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr. Opin. Cell Biol.* **20**, 341–348 (June 2008).
396. Wang, E. *et al.* Histone H2B ubiquitin ligase RNF20 is required for MLL-rearranged leukemia. *Proc. Natl. Acad. Sci.* **110**, 3901–3906 (Mar. 2013).
397. Boffo, S., Damato, A., Alfano, L. & Giordano, A. CDK9 inhibitors in acute myeloid leukemia. *J. Exp. Clin. Cancer Res.* **37**, 36 (Dec. 2018).
398. Hinterleitner, C. *et al.* Rac1 signaling protects monocytic AML cells expressing the MLL-AF9 oncogene from caspase-mediated apoptotic death. *Apoptosis* **18**, 963–979 (Aug. 2013).
399. Su, Y. *et al.* Targeting PI3K, mTOR, ERK, and Bcl-2 signaling network shows superior antileukemic activity against AML ex vivo. *Biochem. Pharmacol.* **148**, 13–26 (Feb. 2018).
400. Luedtke, D. A. *et al.* Inhibition of XPO1 enhances cell death induced by ABT-199 in acute myeloid leukaemia via Mcl-1. *J. Cell. Mol. Med.* **22**, 6099–6111 (Dec. 2018).
401. Knorr, K. L. B. *et al.* MLN4924 induces Noxa upregulation in acute myelogenous leukemia and synergizes with Bcl-2 inhibitors. *Cell Death Differ.* **22**, 2133–2142 (Dec. 2015).
402. Swords, R. T. *et al.* Pevonedistat, a first-in-class NEDD8-activating enzyme inhibitor, combined with azacitidine in patients with AML. *Blood* **131**, 1415–1424 (Mar. 2018).
403. Nechiporuk, T. *et al.* The TP53 Apoptotic Network is a Primary Mediator of Resistance to BCL2 inhibition in AML Cells. *Cancer Discov.* CD–19–0125 (May 2019).

404. Chen, X. *et al.* Targeting mitochondrial structure sensitizes acute myeloid leukemia to Venetoclax treatment. *Cancer Discov.* CD-19-0117 (May 2019).
405. Komor, A. C., Badran, A. H. & Liu, D. R. Editing the Genome Without Double-Stranded DNA Breaks. *ACS Chem. Biol.* **13**, 383–388 (Feb. 2018).
406. Liu, J. *et al.* Pooled library screening with multiplexed Cpf1 library. *Nat. Commun.* **10**, 3144 (Dec. 2019).
407. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (July 2018).
408. Schotte, D. *et al.* Identification of new microRNA genes and aberrant microRNA profiles in childhood acute lymphoblastic leukemia. *Leukemia* **23**, 313–322 (Feb. 2009).
409. Shukla, A. & Huangfu, D. Decoding the noncoding genome via large-scale CRISPR screens. *Curr. Opin. Genet. Dev.* **52**, 70–76 (Oct. 2018).
410. Schwarzer, A. *et al.* The non-coding RNA landscape of human hematopoiesis and leukemia. *Nat. Commun.* **8**, 218 (Dec. 2017).
411. Kurata, J. S. & Lin, R.-J. MicroRNA-focused CRISPR-Cas9 Library Screen Reveals Fitness-Associated miRNAs. *RNA*, rna.066282.118 (May 2018).
412. Zhu, S. *et al.* Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. *Nat. Biotechnol.* **34**, 1279–1286 (Dec. 2016).
413. Pulido-Quetglas, C. *et al.* Scalable Design of Paired CRISPR Guide RNAs for Genomic Deletion. *PLOS Comput. Biol.* **13** (ed Ioshikhes, I.) e1005341 (Mar. 2017).
414. Vila-Casades, M. *et al.* Deciphering microRNA targets in pancreatic cancer using miRComb R package. *Oncotarget* **9**, 6499–6517 (Jan. 2018).
415. Guo, Y. *et al.* Network-Based Combinatorial CRISPR-Cas9 Screens Identify Synergistic Modules in Human Cells. *ACS Synth. Biol.* **8**, 482–490 (Mar. 2019).
416. Soderquist, R. S. *et al.* Systematic mapping of BCL-2 gene dependencies in cancer reveals molecular determinants of BH3 mimetic sensitivity. *Nat. Commun.* **9**, 3513 (Dec. 2018).
417. Voorhees, P. M. & Orłowski, R. Z. The proteasome and proteasome inhibitors in cancer therapy. *Annu. Rev. Pharmacol. Toxicol.* **46**, 189–213 (Feb. 2006).
418. Tsvetkov, P. *et al.* Oncogenic addiction to high 26S proteasome level. *Cell Death Dis.* **9**, 773 (July 2018).
419. Kuusanmki, H. *et al.* Phenotype-based drug screening reveals association between venetoclax response and differentiation stage in acute myeloid leukemia. *Haematologica*, haematol.2018.214882 (July 2019).
420. Certo, M. *et al.* Mitochondria primed by death signals determine cellular addiction to antiapoptotic BCL-2 family members. *Cancer Cell* **9**, 351–365 (May 2006).
421. Sarosiek, K. A. & Letai, A. Directly targeting the mitochondrial pathway of apoptosis for cancer therapy with BH3 mimetics: recent successes, current challenges and future promise. *FEBS J.* **283**, 3523 (2016).

- 422. Montero, J. & Letai, A. Why do BCL-2 inhibitors work and where should we use them in the clinic? *Cell Death Differ.* **25**, 56–64 (Jan. 2018).
- 423. Manara, E. *et al.* MLL-AF6 fusion oncogene sequesters AF6 into the nucleus to trigger RAS activation in myeloid leukemia. *Blood* **124**, 263–272 (July 2014).
- 424. Knight, T. & Irving, J. A. E. Ras/Raf/MEK/ERK Pathway Activation in Childhood Acute Lymphoblastic Leukemia and Its Therapeutic Targeting. *Front. Oncol.* **4**, 160 (2014).
- 425. Valentine, M. C. *et al.* Excess congenital non-synonymous variation in leukemia-associated genes in MLL infant leukemia: a Children’s Oncology Group report. *Leukemia* **28**, 1235–1241 (June 2014).
- 426. Xu, Z. *et al.* BET inhibition represses miR17-92 to drive BIM-initiated apoptosis of normal and transformed hematopoietic cells. *Leukemia* **30**, 1531–1541 (July 2016).
- 427. Brinkmann, K. *et al.* miR17~92 is essential for the survival of hematopoietic stem and progenitor cells by restraining pro-apoptotic BIM. *bioRxiv*, 342071 (June 2018).
- 428. Bogenberger, J. *et al.* Combined venetoclax and alvocidib in acute myeloid leukemia. *Oncotarget* **8**, 107206–107222 (Dec. 2017).
- 429. Kuhlen, M., Klusmann, J.-H. & Hoell, J. I. Molecular Approaches to Treating Pediatric Leukemias. *Front. Pediatr.* **7**, 368 (Sept. 2019).
- 430. Gore, L. *et al.* Survival after blinatumomab treatment in pediatric patients with relapsed/refractory B-cell precursor acute lymphoblastic leukemia. *Blood Cancer J.* **8**, 80 (Sept. 2018).
- 431. Anstee, N. S. *et al.* Impact of elevated anti-apoptotic MCL-1 and BCL-2 on the development and treatment of MLL-AF9 AML in mice. *Cell Death Differ.* **26**, 1316–1331 (July 2019).
- 432. Borkin, D. *et al.* Complexity of Blocking Bivalent ProteinProtein Interactions: Development of a Highly Potent Inhibitor of the MeninMixed-Lineage Leukemia Interaction. *J. Med. Chem.* **61**, 4832–4850 (June 2018).
- 433. Dafflon, C. *et al.* Complementary activities of DOT1L and Menin inhibitors in MLL-rearranged leukemia. *Leukemia* **31**, 1269–1277 (June 2017).
- 434. Picco, G. *et al.* Functional linkage of gene fusions to cancer cell fitness assessed by pharmacological and CRISPR-Cas9 screening. *Nat. Commun.* **10**, 2198 (Dec. 2019).
- 435. Somers, K. *et al.* CCI-007, a novel small molecule with cytotoxic activity against infant leukemia with MLL rearrangements. *Oncotarget* **7**, 46067–46087 (July 2016).
- 436. Somers, K. *et al.* A novel small molecule that kills a subset of MLL-rearranged leukemia cells by inducing mitochondrial dysfunction. *Oncogene*, 1 (Jan. 2019).
- 437. Unnisa, Z. *et al.* Meis1 preserves hematopoietic stem cells in mice by limiting oxidative stress. *Blood* **120**, 4973–81 (Dec. 2012).
- 438. Giaever, G. *et al.* Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nat. Genet.* **21**, 278–283 (Mar. 1999).
- 439. Vidigal, J. A. & Ventura, A. Rapid and efficient one-step generation of paired gRNA CRISPR-Cas9 libraries. *Nat. Commun.* **6**, 8083 (Aug. 2015).

- 440. Han, K. *et al.* Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. *Nat. Biotechnol.* **35**, 463–474 (Mar. 2017).
- 441. Najm, F. J. *et al.* Orthologous CRISPR-Cas9 enzymes for combinatorial genetic screens. *Nat. Biotechnol.* **36**, 179–189 (Dec. 2018).
- 442. Henriksson, J. *et al.* Genome-wide CRISPR Screens in T Helper Cells Reveal Pervasive Crosstalk between Activation and Differentiation. *Cell*. doi:10.1016/j.cell.2018.11.044 (Jan. 2019).
- 443. Dong, M. B. *et al.* Systematic Immunotherapy Target Discovery Using Genome-Scale InVivo CRISPR Screens in CD8T Cells. *Cell* **178**, 1189–1204.e23 (Aug. 2019).
- 444. LaFleur, M. W. *et al.* A CRISPR-Cas9 delivery system for in vivo screening of genes in the immune system. *Nat. Commun.* **10**, 1668 (Dec. 2019).
- 445. Brennan, M. S. *et al.* Humanized Mcl-1 mice enable accurate preclinical evaluation of MCL-1 inhibitors destined for clinical use. *Blood* **132**, 1573–1583 (Oct. 2018).
- 446. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–61 (Oct. 2014).
- 447. Deans, R. M. *et al.* Parallel shRNA and CRISPR-Cas9 screens enable antiviral drug target identification. *Nat. Chem. Biol.* **12**, 361–366 (May 2016).
- 448. Ryan, C. J., Bajrami, I. & Lord, C. J. Synthetic Lethality and Cancer - Penetrance as the Major Barrier. *Trends in cancer* **4**, 671–683 (Oct. 2018).
- 449. Prinz, F., Schlange, T. & Asadullah, K. Believe it or not: how much can we rely on published data on potential drug targets? *Nat. Rev. Drug Discov.* **10**, 712 (Sept. 2011).
- 450. Errington, T. M. *et al.* An open investigation of the reproducibility of cancer biology research. *Elife* **3**. doi:10.7554/eLife.04333 (Dec. 2014).
- 451. Griffiths, J. B. The Effect of Medium Changes on the Growth and Metabolism of the Human Diploid Cell, W1-38. *J. Cell Sci.* **8** (1971).
- 452. Kim, T. *et al.* In situ transcriptome characteristics are lost following culture adaptation of adult cardiac stem cells. *Sci. Rep.* **8**, 12060 (Dec. 2018).
- 453. Wright Muelas, M., Ortega, F., Breitling, R., Bendtsen, C. & Westerhoff, H. V. Rational cell culture optimization enhances experimental reproducibility in cancer cells. *Sci. Rep.* **8**, 3029 (Dec. 2018).
- 454. Vande Voorde, J. *et al.* Improving the metabolic fidelity of cancer models with a physiological cell culture medium. *Sci. Adv.* **5**, eaau7314 (Jan. 2019).
- 455. Haibe-Kains, B. *et al.* Inconsistency in large pharmacogenomic studies. *Nature* **504**, 389–393 (Dec. 2013).
- 456. Haverty, P. M. *et al.* Reproducible pharmacogenomic profiling of cancer cell line panels. *Nature* **533**, 333–337 (May 2016).
- 457. Mair, B. *et al.* High-throughput genome-wide phenotypic screening via immunomagnetic cell sorting. *Nat. Biomed. Eng.* 1–10 (Sept. 2019).
- 458. Chen, C.-H. *et al.* Improved design and analysis of CRISPR knockout screens. *Bioinformatics*. doi:10.1093/bioinformatics/bty450 (June 2018).

459. Kweon, J., Kim, D.-e., Jang, A.-H. & Kim, Y. CRISPR/Cas-based customization of pooled CRISPR libraries. *PLoS One* **13** (ed Maas, S.) e0199473 (June 2018).
460. Michlits, G. *et al.* CRISPR-UMI: single-cell lineage tracing of pooled CRISPRCas9 screens. *Nat. Methods*. doi:10.1038/nmeth.4466 (Oct. 2017).
461. Schmierer, B. *et al.* CRISPR/Cas9 screening using unique molecular identifiers. *Mol. Syst. Biol.* **13**, 945 (Oct. 2017).
462. Zhu, S. *et al.* Guide RNAs with embedded barcodes boost CRISPR-pooled screens. *Genome Biol.* **20**, 20 (Dec. 2019).
463. Daniszewski, M. *et al.* Automated Cell Culture Systems and Their Applications to Human Pluripotent Stem Cell Studies. *SLAS Technol. Transl. Life Sci. Innov.* **23**, 315–325 (Aug. 2018).
464. Buchser, W. Automated CRISPR Screening using the CellRaft Technology.
465. Drexler, H. *et al.* p53 alterations in human leukemia/lymphoma cell lines: in vitro artifact or prerequisite for cell immortalization? *Leukemia* **14**, 198–206 (Jan. 2000).
466. Leroy, B. *et al.* Analysis of TP53 mutation status in human cancer cell lines: a reassessment. *Hum. Mutat.* **35**, 756–65 (June 2014).
467. Mato, A. R. *et al.* Real-world outcomes and management strategies for venetoclax-treated chronic lymphocytic leukemia patients in the United States. *Haematologica* **103**, 1511–1517 (Sept. 2018).
468. Eyre, T. A. *et al.* Efficacy of venetoclax monotherapy in patients with relapsed chronic lymphocytic leukaemia in the post-BCR inhibitor setting: a UK wide analysis. *Br. J. Haematol.* doi:10.1111/bjh.15802 (Feb. 2019).
469. Anderson, M. A. *et al.* The BCL2 selective inhibitor venetoclax induces rapid onset apoptosis of CLL cells in patients via a TP53-independent mechanism. *Blood* **127**, 3215–24 (June 2016).
470. Garnett, M. J. *et al.* Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* **483**, 570–575 (Mar. 2012).
471. Aldoss, I. *et al.* Venetoclax and hypomethylating agents in TP53 mutated acute myeloid leukaemia. *Br. J. Haematol.* bjh.16166 (Aug. 2019).
472. Kuykendall, A., Duployez, N., Boissel, N., Lancet, J. E. & Welch, J. S. Acute Myeloid Leukemia: The Good, the Bad, and the Ugly. *Am. Soc. Clin. Oncol. Educ. book. Am. Soc. Clin. Oncol. Annu. Meet.* **38**, 555–573 (May 2018).
473. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPRCas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* **1** (June 2018).
474. Ihry, R. J. *et al.* p53 inhibits CRISPRCas9 engineering in human pluripotent stem cells. *Nat. Med.* **1** (June 2018).
475. Brown, K. R., Mair, B., Soste, M. & Moffat, J. CRISPR screens are feasible in TP 53 wildtype cells. *Mol. Syst. Biol.* **15**. doi:10.15252/msb.20188679 (Aug. 2019).

- 476. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. Reply to CRISPR screens are feasible in <i>TP53</i> wildtype cells. *Mol. Syst. Biol.* **15**. doi:10.15252/msb.20199059 (Aug. 2019).
- 477. Kuscu, C. *et al.* CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. *Nat. Methods*. doi:10.1038/nmeth.4327 (June 2017).
- 478. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (May 2016).
- 479. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (Mar. 2012).
- 480. Gilvary, C. *et al.* A machine learning approach predicts essential genes and pharmacological targets in cancer. *bioRxiv*, 692277 (July 2019).
- 481. Lin, A. *et al.* Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Sci. Transl. Med.* **11**, eaaw8412 (Sept. 2019).