University Library

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Xu, Zhuofan

Title:
Identifying Novel Strategies to Enhance the Anti-cancer Activity of Venetoclax by Manipulating NOXA Expression

Date:
2021

Persistent Link:
https://hdl.handle.net/11343/288705

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.

Identifying Novel Strategies to Enhance the Anti-cancer Activity of Venetoclax by Manipulating *NOXA* Expression

Zhuofan Xu
0000-0002-6481-1238

Submitted in total fulfilment for the degree of

Master of Research

June 2021

Department of Medial Biology
Faculty of Medicine, Dentistry and Health Sciences
The University of Melbourne

Abstract

Apoptosis is a form of programmed cell death. The intrinsic pathway of apoptosis is governed by the BCL2 family proteins. Targeting BCL2 proteins by small molecules that mimicking the BH3-only proteins to induce apoptosis has proven to be a successful strategy for cancer therapy. Venetoclax, a specific inhibitor of BCL2, has exhibited remarkable efficacy in treating cancers that rely on BCL2 for survival. However, the activity of venetoclax is often limited in other cancers whose survival relies on MCL1, another BCL2 family member. Selective MCL1 inhibitors have been developed and are currently being evaluated in clinical trials. However, the clinical development of these agents has been hampered by toxicity, especially cardiac toxicity. Potentially, another strategy to target MCL1 is by modulating NOXA, a BH3-only protein that selectively binds to MCL1 and mediates its degradation.

I hypothesised that increased *NOXA* expression would prime cancer cells to venetoclax killing and that this would reduce their co-dependence on MCL1. In order to identify new targets to modulate *NOXA* expression, I generated and validated cell lines that report on *NOXA* transcription and then carried out CRISPR-Cas9 genetic screens in those *NOXA* reporter cell lines.

In CRISPR-Cas9 loss-of-function screens focused on epigenetic regulators, I found several genes whose mutation or loss modulated *NOXA* expression, including *CTBP*1, *CHTOP*, *ZMYM3, SPEN*, *HSPA1A*, *KEAP1*, *FOXA1*, *HDAC3* and *SAP30.* Some of these factors have been targeted for cancer therapies, for example *KEAP1* and *HDAC3*, while the others have not yet been recognized for their therapeutic possibilities. Subject to their validation, my results have identified interesting novel mechanisms of *NOXA* regulation, thus providing the rationale basis for the development of new anti-cancer agents.

In CRISPR-Cas9 tiling screens that focused on the *NOXA* promoter region, five *cis*-regulatory elements were identified that contributed to regulation of *NOXA* expression. Among them, a hypermethylated element on the *NOXA* promoter was found to be important for repressing *NOXA* expression across diverse cell lines derived from blood cancers. Disrupting this region

led to *NOXA* induction. Potentially, the findings could provide a rational basis of combining hypomethylating agents with venetoclax in a range of haematological malignancies.

In summary, several potential *NOXA* regulating proteins and DNA elements were discovered by the CRISPR-Cas9 screening approaches. Once validated, these findings should provide new insights into *NOXA* regulation.

Declaration

I declare that this thesis comprises my original work with the following exceptions:

Chapter 4 & 5 − The Miseq analysis (figure 4.5, figure 5.4.B) was performed by Dr. Mark van Delft (WEHI). The bisulfite sequencing analysis (figure 5.7) was performed by Dr. Quentin Gouil (WEHI)

This thesis is less than 30,000 words in length, excluding tables and references.

Zhuofan Xu

June 2021

Publications

Parts of this thesis have been prepared or submitted for publication in the papers listed below:

"Lin VS, **Xu ZF**, Huang DCS, Thijssen R. BH3 Mimetics for the Treatment of B-Cell Malignancies- Insights and Lessons from the Clinic. Cancers (Basel). 2020 Nov 12;12(11):3353. doi: 10.3390/cancers12113353. PMID: 33198338; PMCID: PMC7696913."

Acknowledgements

Firstly, I would like to thank my supervisor Professor David Huang for all of his great guidance and wisdom throughout this project. His mentorship has transformed me into an independent and critical researcher. I would sincerely thank my co-supervisor Dr. Mark van Delft for all the help and encouragement. His great knowledge and insights in science have deeply inspired me. I would like to express great gratitude to my co-supervisor Dr. Mary Ann Anderson for her invaluable advice and warm support. Her encouragement and faith in me helped me getting through many difficulties throughout my study. I would also like to thank my advisory committee, Dr. Gemma Kelly, Dr. Ashley Ng and Dr. Rhys Alan for their valuable suggestions for my projects as well as constructive assurances of my progress.

Many thanks to Dr. Anita Horvath (The University of Melbourne), Professor Li Wu and Dr. Ning Wu (Tsinghua University) and many others who have contributed to the collaboration program between Tsinghua University and the University of Melbourne. To Kerry Ko, Catherine McLean and Michelle Birrell for providing useful information regarding candidature requirements and administrative works. To the China Scholarship Council and the University of Melbourne for funding and scholarship supports.

Many of my colleagues have also provided patient guidance and timely assistance throughout my study. I would like to particularly mention Shuai Huang for patiently guiding me when I first came into the laboratory, and Dr. Andrew Jarratt for the invaluable advice and assistance with completing this thesis. I would like to thank all the members in Huang and Roberts Labs, Dr. Rachel Thijssen, Dr. Christine White, Dr. Jianan Gong, Dr. Mario Djajawi, Angela Georgiou, Chris Riffkin, Farima Moghaddaskho, Victor Lin, Hongke Peng and Mingyuan Gao for their help, support and encouragement. I would also like to thank Dr. Michael Dengler, Dr. Marco Herold, Dr. Zhen Xu, Dr. Quentin Gouil, Dr. Sweta Iyer, Dr. Zilu Wang, Hema Subas-Satish, Melissa Shi and past as well as present members of Blood Cells and Blood Cancer Division for great scientific advice and timely help on reagents and protocols. I also want to thank Dr. Stephen Wilcox from the Genomics lab, Casey Antilla, HuyBenny Nguyen, Athena Chen and all members from FACS lab, who have helped me greatly with my sequencing and FACS analysis.

Last but not least, this thesis would not be possible without the love and support from my family and friends. Thanks to my parents for their unconditional support and understanding. Thanks to my friends in Melbourne and in China for their accompany and encouragement. And finally, thanks to my boyfriend, Jinkai Lin for sharing my joy and tears throughout the journey and for the unwavering faith in me.

Abbreviations

TNF	Tumor necrosis factor
DISC	Death-inducing signalling complex
MOMP	Mitochondrial outer membrane permeabilisation
XIAP	Inhibitor of Apoptosis protein
BH	BCL2 homology
ER	Endoplasmic reticulum
CLL	Chronic lymphocytic leukemia
SLL	Small lymphocytic lymphoma
AML	Acute myeloid leukemia
STAT	Signal transducers and activators of transcription
MM	Multiple myeloma
MEF	Mouse embryonic fibroblasts
FADD	FAS-associated death domain protein
TRADD	TNFR-associated death domain protein
APAF1	Apoptotic protease activating factor 1
CLR5	Cullin-Ring-ligase-5
HDAC	Histone deacetylase
DLBCL	Diffuse large B-cell lymphoma
ALL	Acute lymphoblastic leukaemia
FCS	Fetal calf serum
HRP	Horseradish peroxidase
FACS	Fluorescence-activated cell sorting
CPM	Counts per million
CRISPR	Clustered regularly interspaced short palindromic repeats
sgRNA	Small guide RNA
DSB	Double-strand break
PAM	Protospacer adjacent motif
NHEJ	Nonhomologous end joining

HDR	Homology-directed repair
SSB	Single-strand breaks
NLS	Nuclear localization signal
RNAi	RNA-interference
shRNA	Short hairpin RNA
B/B DKO	BAX/BAK double knock-out
FDR	False discovery rate
CREs	*Cis*-regulatory elements
TSS	Translation start site

Table of Contents

List of figures

List of tables

Chapter 1: Introduction

1.1 Apoptosis: programmed cell death

The balance between cell proliferation and death is fundamental for normal tissue development and health in multicellular organisms. Classic forms of cell death include necrosis, the unregulated lysis of injured cells, and apoptosis, the programmed death of unwanted or damaged cells with well-controlled mechanisms. First described by Kerr, Wyllie and Currie in 1972, apoptosis is characterised by distinct morphological changes including cell shrinkage, fragmentation into apoptotic bodies and rapid phagocytosis by neighbouring cells[2]. Apoptosis is considered an essential component of various processes including embryonic development, immune system development and functioning, normal cell turnover and drug-induced cell death[3-5]. Dysregulation of apoptosis contributes to many human diseases including autoimmune disorders, neurodegenerative diseases and cancers[3,4,6].

To date, two distinct apoptosis pathways have been identified in vertebrates: the death receptor (extrinsic) pathway and the mitochondrial (intrinsic) pathway. The extrinsic pathway enables cells to respond to extracellular death signals. Using signal transduction via transmembrane proteins, the extrinsic pathway involves receptors that are members of the tumor necrosis factor (TNF) receptor super-family[7]. Upon ligand binding, receptors recruit adapter proteins to form the death-inducing signaling complex (DISC), leading to activation of procaspase-8[8]. As the result, activated caspase-8 triggers the downstream execution of apoptosis.

Conversely, the intrinsic pathway is initiated by intracellular cues, such as DNA damage, ischaemia, and oxidative stress. These stress signals alter the balance of interactions between BCL2 proteins, causing certain pro-apoptotic BCL2 family proteins to become activated to disrupt the mitochondrial outer membrane integrity, triggering an event known as mitochondrial outer membrane permeabilization (MOMP).

The BCL2 family proteins consists of members that either promote or inhibit apoptosis. The interplay between the pro- and anti-apoptotic family members regulates apoptosis by

governing MOMP[9]. Recognised as the "point of no return", MOMP causes the release of apoptogenic factors into the cytosol, where they facilitate activation of caspases that mediate apoptotic cell death. These apoptogenic factors include cytochrome c, which triggers apoptotic protease activating factor 1 (APAF1) mediated activation of caspase-9, and SMAC, which prevents the Inhibitor of Apoptosis protein (XIAP) from inhibiting caspase-3/7 [9,10]. (Figure1.1)

1.2 BCL2 family proteins: the major regulators of apoptosis

1.2.1 The discovery of *BCL2*

Commitment to the intrinsic pathway of apoptosis is governed by a complex network of interactions between BCL2 family proteins. The *BCL2* gene, encoding the founding member of this protein family, was initially identified at the break-point region of t(14;18), a recurrent chromosomal translocation in human follicular lymphoma[11,12]. Driven by the immunoglobulin heavy chain gene promoter on chromosome 14, the translocated *BCL2* gene becomes overexpressed. Subsequent studies found that enforced expression of *BCL2* in hematopoietic cells rendered them insensitive to cell death induced by cytokine deprivation[13]. Unlike other oncogenes that cause cell proliferation, overexpression of *BCL2* promoted cell survival by maintaining cells in G0 phase[13]. In addition, mice harbouring the *BCL2* transgene (representing the *BCL2-Ig* fusion gene found at the chromosomal breakpoint of t(14;18)) showed lymphocyte accumulation, but this was not sufficient enough for tumorigenesis[14]. When *BCL2* and another oncogene *MYC* were both overexpressed in cell lines or as transgenes in mice, marked synergy in promoting oncogenic transformation was observed, resulting in rapid cell proliferation and lymphoma development[13,15]. Taken together, these findings uncovered the unique role of BCL2 as a mediator of cell survival that enables nascent neoplastic cells to survive long enough to gain other oncogenic mutations, ultimately facilitating their full malignant transformation. Therefore, evading apoptosis is defined as a hallmark of cancer[16].

Since the identification of BCL2, sixteen additional BCL2 family proteins have been discovered, which share one or more regions of sequence homology, termed BCL2 homology (BH) domains. Based on their structures and function, members of BCL2 family can be divided into three distinct subgroups based on their structural and functional

differences: 1) the pro- survival proteins; 2) the pro-apoptotic proteins and 3) the BH3-only proteins. BCL2 and its closest homologues (BCL-XL, BCLW, MCL1 and BCL2A1), which contain four conserved sequence motifs BH1-4, function to promote cell survival. The pro-apoptotic members, BAX, BAK and less studied BOK share similar structural features of all BH1-3 domains with their pro-survival relatives. Despite this similarity, BAX and BAK promote cell death rather than survival[17]. Lastly, the BH3-only proteins (BIM, BAD, PUMA, NOXA, BIK, HRK, BID and BMF) share with each other and their BCL2 relatives only the small BH3 domain. The BH3-only proteins function as apoptosis initiators, as they engage and inactivate their pro-survival relatives[18,19,20]. (Figure1.2)

1.2.2 Pro-survival proteins

All pro-survival proteins (BCL2, BCL-XL, BCLW, MCL1 and BCL2A1) share similar overall structure, which comprises of 9 α-helices bundled into a globular structure with a hydrophobic groove on its surface[21]. The hydrophobic groove corresponds to the region of the proteins that binds to pro-apoptotic members of the BCL2 family. In particular, pro- survival proteins sequester BAX and BAK via the interaction between the BH3 domains of BAX and BAK and the hydrophobic grooves of pro-survival proteins[20]. The feature that accounts for variation in the functional profiles of the different pro-survival members lies in the hydrophobic groove as well. Subtle changes in the structural topology and electrostatic potential of the binding grooves translate into differences in binding affinity, thus, anti- apoptotic activity.

Most pro-survival proteins in the cell localise to the mitochondrial outer membrane via their short hydrophobic transmembrane domains at C-termini. However, studies have also found BCL2 family proteins in other organelles including the endoplasmic reticulum (ER), Golgi apparatus, the nucleus and the peroxisomes, where their roles for cell fate determination are still being studied[22].

Although some functional redundancy exists between different pro-survival proteins, expression of certain members of the family is important for development and homeostasis in specific tissues[23]. For instance, BCL2 is highly expressed in the renal tissues and in mature lymphocytes. BCL2 knock-out in mice leads to polycystic kidney disease, lymphocytopenia and premature greying[24]. BCLW knock-out causes male sterility in mice, whereas BCL2A1 knock- out leads to accelerated death of granulocytes and mast cells in culture[25]. In the case

of BCL-XL and MCL1, their expression patterns show lower tissue specificity, and they are widely expressed in essential tissues such as brain and heart. Loss of either *MCL1* or *BCL-XL* in mice results in embryonic lethality[26,27]. In particular, loss of one allele of BCL-XL encoding gene (*BCL2L1*) reduces platelet numbers and impairs spermatogenesis[28,29], whereas conditional knock-out of BCL-XL in mice leads to death of hepatocytes and neurons[26,30]. Similarly, conditional knock-out mice models of MCL1 revealed its critical role in diverse tissues, including cardiomyocytes[31], hematopoietic stem cells[32], developing and mature lymphocytes[33,34] and oocytes[35].

1.2.3 Pro-apoptotic proteins

Despite sharing sequence homology with their pro-survival relatives, the pro-apoptotic proteins BAX, BAK and BOK promote rather than inhibit apoptosis.

Pro-apoptotic proteins undergo conformational changes during activation. As previously indicated, the overall structures of inactive BAX and BAK are strikingly similar to their pro- survival counterparts, possessing the same overall fold, surface hydrophobic groove and a transmembrane domain at their C-termini. The BH3 domain of BAX and BAK is inaccessible in their inactive conformation and becomes exposed upon activation, facilitating oligomerization and binding to the pro-survival proteins.

In healthy cells, BAX is predominantly found in the cytosol as an inactive monomer, with its transmembrane domain folded into its hydrophobic groove. Upon activation, BAX undergoes conformational rearrangement, and the transmembrane domain is released from its hydrophobic groove, enabling BAX's translocation to the outer membrane of mitochondria[36]. In contrast to BAX, BAK predominantly locates to the outer mitochondria membrane where it adopts an inactive conformation and associates with VDAC2[37]. Upon activation, BAK changes conformation, dissociates from VDAC2. Released BAX and BAK undergo further conformational changes, exposing their hydrophobic grooves and BH3 domains. The exposed domains bind to each other across different molecules and facilitate the formation of BAX and BAK oligomers[38]. These oligomers are predicted to form pores in the outer mitochondrial membrane and trigger MOMP. BOK, a less well-studied pro-apoptotic protein, is predominantly localised in the ER and Golgi compartment[39]. Its role in apoptosis regulation is largely unknown.

1.2.4 BH3-only proteins

BH3-only proteins are named because they only share a single region of sequence homology with other BCL2 family members, the BH3 domain. The known BH3-only proteins BIM, BAD, PUMA, NOXA, BIK, HRK, BID and BMF interact with their BCL2 family relatives, functioning to initiate apoptosis.

Unlike other BCL2 family proteins that form an α helical bundle, most BH3-only proteins are unstructured prior to engaging pro-survival proteins[40]. Within their BH3 domain, there are four to five hydrophobic residues forming an amphipathic α-helix that allow them to engage their pro-survival relatives[19]. There is also a conserved aspartic acid residue that forms a salt bridge to a conserved arginine on the pro-survival protein's hydrophobic groove. Binding by the BH3-only proteins will block the ability of the pro-survival proteins to bind BAX and BAK, thereby neutralising their pro-survival function[41].

There is evidence that BH3-only proteins could bind to BAX and BAK as well. Short peptides representing the α-helical BH3 domains of BID or BIM were shown to be capable of inducing oligomerisation of BAX and BAK[42]. However, whether this interaction could happen under physiological condition still remains unclear.

BH3-only are recoginised as apoptosis initiator for their distinct roles in response to various cytotoxic stimuli. Different type of cell stress can activate different BH3-only proteins to initiate apoptosis. For example , PUMA and NOXA are activated by DNA damage in a p53- dependent manner[43], whereas BIM is activated following the withdrawal of cytokines such as IL-3[44]. BID is a special member of the BH3-only protein group, linking the death receptor signaling pathway and mitochondrial pathway. BID's cleavage by caspase-8 generates an active form, termed tBID[45].

1.2.5 Regulation of apoptosis by BCL2 family proteins

The "Priming-capture-displacement" model was proposed to describe how the interplay between the three subsets of BCL2 family proteins regulate apoptosis (Figure1.3)[1]. In general, the execution of intrinsic apoptosis is determined by the activity of the pro-apoptotic effectors BAX and BAK. Other members of the BCL2 family including pro-survivals and BH3- only proteins act to regulate BAX and BAK function. Pro-survival proteins bind and capture BAX and BAK to prevent their pro-apoptotic function, while BH3-only proteins bind

to pro-survival proteins and displace BAX and BAK, thus enable BAX and BAK to induce apoptosis. Cells undergo apoptosis once sufficient BAX and BAK are activated and displaced.

In healthy cells, most BAX and BAK are kept in their inactive form, however some become activated, or primed[46]. The primed BAX and BAK are immediately captured by the pro-survival proteins, thus inhibited from further conformational changes. During cellular stress, BH3-only proteins are activated either transcriptionally or post-translationally. Upregulation of BH3- only proteins leads to displacement of the captured BAX and BAK from the pro-survival proteins. When sufficient pro-survival proteins are neutralised, the primed BAX and BAK can freely oligomerise and induce MOMP. The molecular mechanism of BAX and BAK activation remains debatable. As mentioned, evidence supports some BH3-only, for example NOXA, can bind and activate BAX and BAK. Those interactions have been argued as an essential requirement for BAX and BAK conformational changes that lead to oligomerisation on the mitochondrial outer membrane[42]. However, cells lacking all BH3-only proteins as well as p53 and Rb could still undergo BAX/BAK-driven apoptosis[47].

Interestingly, the interactions among BCL2 family proteins exhibit specificity. Their binding preference of the particular protein determined the potency for apoptosis induction (Figure1.4)[41]. For example, BH3-only proteins BIM, PUMA and BID bind with high affinity to all pro-survival proteins, which makes them potent death agonists. Conversely, proteins such as BAD and NOXA only bind preferentially to certain pro-survival proteins. BAD, for example, predominantly binds to BCL2, BCL-XL and BCLW, but poorly to MCL1. NOXA, on the other hand, shows strong binding to MCL1 and BCL2A1, but not the other pro-survival proteins. All pro-survival proteins can bind and sequester BAX but BCL2 and BCLW cannot efficiently sequester BAK.

1.3 BH3 mimetic drugs

Targeting the pro-survival BCL2 family members for cancer therapy has considerable appeal. Overexpression of the pro-survival proteins BCL2, MCL1 and BCL-XL is often observed in various cancer types. While cells overexpressing one or more pro-survival proteins have a clear survival advantage, paradoxically, they are also ‘primed for death’, as inhibition of these upregulated pro-survival proteins can result in the release of large amounts of bound pro- apoptotic proteins, driving the cell towards apoptosis. In other words, the cells become dependent on specific pro-survival proteins for their survival. Such discoveries sparked the

search for small molecules that could mimic the function of BH3-only proteins in binding with high affinity to the hydrophobic grooves on the surfaces of specific pro-survival proteins commonly overexpressed in malignancies, thereby restoring the capacity of cancer cells to undergo apoptosis. (Figure1.5)

BH3 mimetics are small molecule inhibitors that mimic the activity of BH3-only proteins. The first validated BH3 mimetic, ABT-737, was reported in 2005[48]. ABT-737 binds to BCL2, BCL-XL and BCLW with high affinity, thereby mimicking the action of the BH3-only protein BAD. ABT- 737 demonstrated sub-micromolar killing of primary cancer cells and cancer cell lines as well as anti-tumour activity in mouse models[49]. Navitoclax, an orally bioavailable analogue of ABT-737 was successful in reaching clinical trials[50]. Among several tested hematologic malignancies, Navitoclax was most potent against chronic lymphocytic leukemia (CLL) as predicted by the ubiquitously high levels of BCL2 expression observed in this disease[50,51]. However, the clinical development of Navitoclax was curtailed by dose-limiting thrombocytopenia caused by on-target inhibition of BCL-XL, an important protein for platelet survival[29,52].

The limitations of Navitoclax highlighted the need for a compound that selectively targeted BCL2. In 2013, a new compound called venetoclax (ABT-199) was reported. Venetoclax binds to BCL2 with sub-nanomolar affinity but interacts weakly with BCL-XL and BCLW[53]. Venetoclax has shown remarkable efficacy both *in vitro* and *in vivo*. In sparing BCL-XL, venetoclax had limited effect on both mouse and human platelets *in vitro* and *in vivo* but retained its cytotoxic activity on BCL2-dependent tumours[53,54]. Based on its safety and efficacy, venetoclax has been approved by the FDA for use in patients with CLL, small lymphocytic lymphoma (SLL) and acute myeloid leukemia (AML). Venetoclax demonstrated exceptional efficacy in the clinic, igniting enthusiasm for developing additional BH3-mimetic drugs, including specific inhibitors for MCL1 and BCL-XL. (figure1.6)

1.4 MCL1: another important pro-survival protein

Venetoclax has shown remarkable activity in treating blood cancers whose survival relies principally on BCL2. However, besides BCL2, it is known that the pro-survival protein MCL1 is also essential for the growth of diverse types of cancer.

MCL1 was initially discovered as a protein upregulated in human myeloid leukaemia cells when these cells were induced with differentiation factors[55]. As discussed, expression of

MCL1 is important for the survival of many cell types, in particular, haematopoietic stem cells and progenitor cells, and a variety of haematopoietic cells such as neutrophils, T-lymphocytes and plasma B cells[33,56], as well as non-blood cells such as hepatocytes, cardiomyocytes and neurons[57,58]. Cancers derived from these tissues frequently remain dependent on MCL1 for their survival. This includes myeloid leukemias[59], lymphomas, multiple myeloma (MM)[60], hepatocellular carcinoma[61] and certain subtypes of lung or breast cancer[62]. Apart from that, genetic amplification of *MCL1* is found in numerous solid and haematological cancers[63]. Dependency or elevated expression of MCL1 frequently confers resistance to BH3 mimetics navitoclax and venetoclax[62,64,65]. These observations highlight the importance of MCL1 in the process of apoptosis evasion of various type of cancers. It is therefore crucial to eliminate MCL1 activity in order to enhance venetoclax treatment outcome.

1.4.1 Regulation of MCL1

Given its importance for the survival of many cell types, the amount of MCL1 protein in a cell is tightly regulated at multiple levels. At the level of gene transcription, several constitutively active and extracellular signal-activated transcription factors are found to regulate *MCL1* expression, including the signal transducers and activators of transcription (STAT) proteins. Consensus binding sites of several STAT proteins were identified on the *MCL1* promoter[66]. Notable examples include STAT3, which upregulates *MCL1* transcription upon IL-6 treatment, and STAT5, which contributes to the BCR/ABL induced expression of *MCL1*[67]. At the post- transcriptional level, alternative splicing gives rise to two distinct *MCL1* mRNAs encoding *MCL1* and *MCL1s* isoforms with opposing functions[68]. Translationally, *MCL1* was shown to be regulated by micro RNAs through a mir-29b binding site in the 3'-UTR of *MCL1* mRNA. The binding directly inhibits expression of *MCL1*[69].

Finally, at the post-translational level, multiple modes of MCL1 regulation have been defined. Proteasome-dependent degradation is one of the main routes responsible for the rapid turnover of MCL1. Several ubiquitin ligases of MCL1 have been discovered, including, Mule[70], MARCH5[71] and Parkin[72]. Structural study revealed that MULE contains a BH3 domain which is able to bind MCL1 directly[70]. How the other E3 ligases without BH3 domains interact with MCL1 remained unknown until recently when it was shown that NOXA promotes MCL1 ubiquitination by the E3 ligase MARCH5[73,74]. Upon the formation of

MCL1- NOXA complex, several proteins including MARCH5, MTCH2 and UBE2K are recruited and co-operate to ubiquitinate the complex, results in its rapid degradation[74].

1.4.2 Targeting MCL1 for cancer therapies

Given the established role of MCL1 as a critical prosurvival factor in a range of haematological malignancies and a resistance factor to BH3 mimetics targeting BCL2 and BCL-XL, there has been longstanding interest in designing potent therapeutic interventions that abrogate the anti-apoptotic function of MCL1.

Downregulation of MCL1 expression has been reported as one of the mechanisms for the activities of many anticancer agents. For example, MCL1 is transcriptionally regulated by STATs, compounds including Dasatinib and Imatinib have been demonstrated to have inhibitory effects on STAT proteins including STAT3 and STAT5, leading to MCL1 downregulation and apoptosis induction[75,76]. Treatment with cyclin-dependent kinase (CDK) inhibitors resulted in inhibition of RNA polymerase II, which leads to downregulation of *MCL1* mRNA and rapid reduction of short-lived MCL1 protein[77]. However, most of these compounds lack specificity and often affect multiple signalling pathways[78]

Progress in the development of selective MCL1 inhibitors has accelerated over recent years. Several compounds targeting MCL1 have now entered clinical trials, including AZD5991[79], S63845[80], AMG-176 and AMG-397[81, 82]. These new compounds have shown promising efficacy in MCL1-dependent cancers, including MM, AML and B cell lymphoma *in vitro*. The safety and efficacy MCL1 inhibitors have been further evaluated in the clinic. However, a phase 1 trial assessing the safety and tolerability of AMG-397 in patients was placed on hold by the FDA in 2019 because of concerns over potential cardiac toxicity. Indeed, MCL1 has been shown in murine knockout models to play a critical role in diverse tissues. While the tolerability of an MCL1 inhibitor S63845 at doses capable of inhibiting tumour progression has been demonstrated in mice[83], many MCL1 inhibitors have been shown to have weaker affinity for mouse MCL1 than human MCL1, emphasising the importance of clinical trials in establishing whether a therapeutic window can be found for MCL1 inhibitors in patients. As such, elucidating alternative ways to inhibit MCL1 while avoiding the systemic toxicity are sought after to realise the full therapeutic potential of MCL1 down regulation without the dose limiting toxicity.

1.5 NOXA: a BH3-only protein holds great therapeutic potential

Among the many MCL1 regulators, NOXA holds great therapeutic potential. NOXA is a BH3-only protein that selectively binds MCL1 and inhibits its pro-survival function. NOXA binding to MCL1 is an important mechanism through which MCL1 is degraded by the ubiquitin proteasome pathway[84]. Interestingly, other BH3-only proteins, like BIM, bind to and stabilise MCL1, most likely by preventing NOXA or MULE from binding to it, which highlights NOXA's distinct role in mediating MCL1 degradation[70,85,86]. As such, there is a rationale to explore whether NOXA can be harnessed to induce apoptosis in MCL1-dependent cancer cells.

1.5.1 Regulation of NOXA and its binding partners through protein modification

NOXA regulates apoptosis through forming interactions predominantly to MCL1 and BCL2A1[41]. The function of BCL2A1 has been mostly described in the context of lymphocytes after CD40 stimulation, but detailed insight into its interaction with NOXA is still incomplete[87]. Conversely, for MCL1, the rapid turn-over of MCL1-NOXA complex is one of the major routes by which both proteins are regulated. Several E3 ligases were discovered to participate in this process.

In addition to promoting MCL1 ubiquitin modification, NOXA itself is also regulated by the protein ubiquitination system. The Cullin-Ring-ligase-5 (CLR5) E3 ligase is reported to ubiquitinate NOXA, resulting in proteasomal targeting and degradation[88].

NOXA itself and its critical binding partner MCL1 are under tight protein level regulation. This interaction holds great therapeutic potential but remains to be fully elucidated for new targeted therapies.

1.5.2 Transcriptional regulation of *NOXA*

Another potential way to harness NOXA for cancer therapy is through inducing *NOXA* expression by targeting its upstream signalling pathways.

NOXA was first identified as a phorbol-12-myristate-13-acetate (PMA) responsive gene in T-cell acute leukemia cells and in human peripheral blood cells[89]. Consequently, the gene was named PMA-induced protein 1 (*PMAIP1*). *PMAIP1* was then rediscovered and named

NOXA (Latin for damage) in a different approach using mRNA from γ-irradiated mouse embryonic fibroblasts (MEF)[43].

The human *NOXA* locus spans 4.4kb and contains several regulatory regions around the gene body. There are three exons within the human *NOXA* but only exon1 and 3 are the protein encoding regions. Exon 2 is only found in two splice variants that are rapidly degraded[90].

NOXA was originally described as a p53 responsive gene upregulated in DNA damage response triggered by γ-irradiation[43,91,92]. As expected, promoter analysis identified a consensus p53- binding site ~195bp upstream of the transcription start site[93]. However, studies have shown DNA damage induced *NOXA* expression to be cell type specific[43,91]. In some tissues, *NOXA* induction occurred efficiently in the absence of p53, indicating that *NOXA* induction can occur in a p53-independent manner. Additional studies suggested that p73, a member of the p53 family of proteins, can also induce *NOXA* expression independently of p53 upon DNA damage[94]. Consistently, studies have demonstrated that p73 is also able to bind the p53 consensus site in the *NOXA* promoter[93].

In addition to genotoxic stresses, various other cellular stresses can also induce *NOXA* expression, including hypoxia. *NOXA* was reported to be a direct target of the hypoxia inducible factor-alpha subunit (Hif1a), mediating hypoxic cell death in a p53-independent manner[95]. Accordingly, a hypoxia-responsive element (HRE) was identified in the *NOXA* promoter ~1kb upstream of the transcription start site[96]. Endoplasmic reticulum (ER) stress is another type of cell stress that have been described to induce the expression of selective BH3- only proteins including NOXA[97]. ATF4, one of the main transcription regulators of the ER stress response, is suggested to function as a transcription activator of *NOXA* in tumour cells[98]. Expectedly, studies have demonstrated that ATF3 binds to the *NOXA* promoter and recruits ATF4 to this site[99].

NOXA is dynamically regulated by epigenetic modulators. Studies in memory T cells revealed that the polycomb complex protein BMI1 represses *NOXA* gene expression by inducing methylation of histone and by recruiting DNA methyltransferase DNMT1, resulting in CpG methylation[100]. (Figure1.7)

1.5.3 Induction of *NOXA* expression

Elevated *NOXA* expression has been associated to cell death in various hematopoietic and solid cancers, including MM[101], CLL[102], MCL[103] and melanoma[104]. Induction of *NOXA* expression can be achieved by pre-existing or novel therapeutics. Although diverse types of cancer therapies have been shown to increase *NOXA* expression, specific strategies to induce *NOXA* directly have not yet been developed. Current anti-cancer drugs that function in part by inducing *NOXA* can be categorised into three groups: cellular stress inducers; proteasome inhibitors and epigenetic therapeutics.

Drugs that are known to induce *NOXA* expression by provoking cellular stress include chemotherapeutics, DNA damaging and p53-activating agents. Many of these agents also induce expression of another BH3-only protein PUMA, which contributes to their ability to promote apoptosis.

Proteasome inhibitors could also cause *NOXA* induction. Bortezomib is a highly selective inhibitor of ^{26}S proteasome and is approved as treatment for MM and MCL[105]. While its mechanism of action is not yet fully understood, Bortezomib was demonstrated to induce *NOXA* in numerous cancer cell types[106]. One of the main signalling pathways through which bortezomib is proposed to activate *NOXA* transcription is the PERK/ATF4 pathway[107]. Ubiquitylation of NOXA protein has recently been shown to be involved in NOXA regulation as well. Increased ubiquitylation and proteasome degradation of NOXA was shown to be one of the mechanisms of chemotherapy resistance[108]. CLR5 E3 ligase can ubiquitinate NOXA and provoke its degradation[88]. Inhibition of CLR5 neddylation has been shown to control NOXA levels. MLN4924, a neddylation inhibitor, demonstrated potent anti-cancer activity in part by preventing NOXA degradation[109].

Epigenetic drugs such as histone deacetylase (HDAC) inhibitors and DNA methyltransferase inhibitors constitute a novel class of anti-cancer agents. A novel small molecule inhibitor of the polycomb complex protein BMI1, PTC-209, was shown to have potent activity in MM cell lines in part by upregulating *NOXA*[110]. Another component of the polycomb repressive complex, EZH2, has been successfully targeted by small molecule inhibitor GSK126. One of the mechanisms of GSK126 induced apoptosis in melanoma cells is by activating the *NOXA* inducing transcription factor ATF3 and increasing *NOXA* expression[111]. Several HDAC inhibitors were shown to upregulate the transcription of genes encoding NOXA and other

BH3-only proteins (mostly BIM and PUMA). For instance, SNX-275 (Entinostat) was active in AML and shown to upregulate *BIM* and *NOXA*[112], and Panobinostat was shown to synergise with the BCL2 inhibitor S55746 in diffuse large B-cell lymphoma (DLBCL) cell lines and upregulate *NOXA*[113]. In mantle cell lymphoma (MCL), NOXA promoter was shown to be one of the target of hypomethylating agent Decitabine. NOXA promoter is directly demethylated upon Decitabine treatment, at the same time NOXA expression is induced[114]. In a more recent study, hypomethylating agent 5-Azacitidine was demonstrated to prime AML cells for venetoclax killing and induce *NOXA* and *PUMA* via the integrated stress response pathway[115].

1.6 Thesis outline

Due to the close association between NOXA and MCL1 and the distinct role of NOXA in mediating MCL1 degradation, regulating the expression of *NOXA* can be used as an alternative way to counter the pro-survival activity of MCL1. Accumulating evidence suggests that modulating *NOXA* holds great potential for the development of novel cancer therapeutics. However, a clear understanding of how *NOXA* expression is regulated both epigenetically and transcriptionally is still lacking. In this thesis, I hypothesise that there are signalling pathways converging on the *NOXA* locus to control its expression that could be targets for the development of new therapies to effectively reduce the dependency of cancer cells on MCL1.

In order to find such pathways, I first generated cell lines in which changes in *NOXA* transcription could be measured readily. In chapter 3, I created and validated *NOXA* reporter cell lines using CRISPR-Cas9 knock-in approach.

In chapter 4, I investigated the mechanisms of *NOXA* transcriptional regulation by utilising focused sgRNA libraries to perform CRISPR-Cas9 screens. I focused primarily on epigenetic modulators, with the aim to identify epigenetic modulators that could be targeted to increase *NOXA* expression.

In chapter 5, I turned my focus to the *cis*-regulatory elements on *NOXA* promoter. By conducting a *NOXA* promoter tiling screen, I performed CRISPR-Cas9 mediated mutagenesis of the *NOXA* promoter region to identify elements relevant to *NOXA* expression and therefore venetoclax drug resistance.

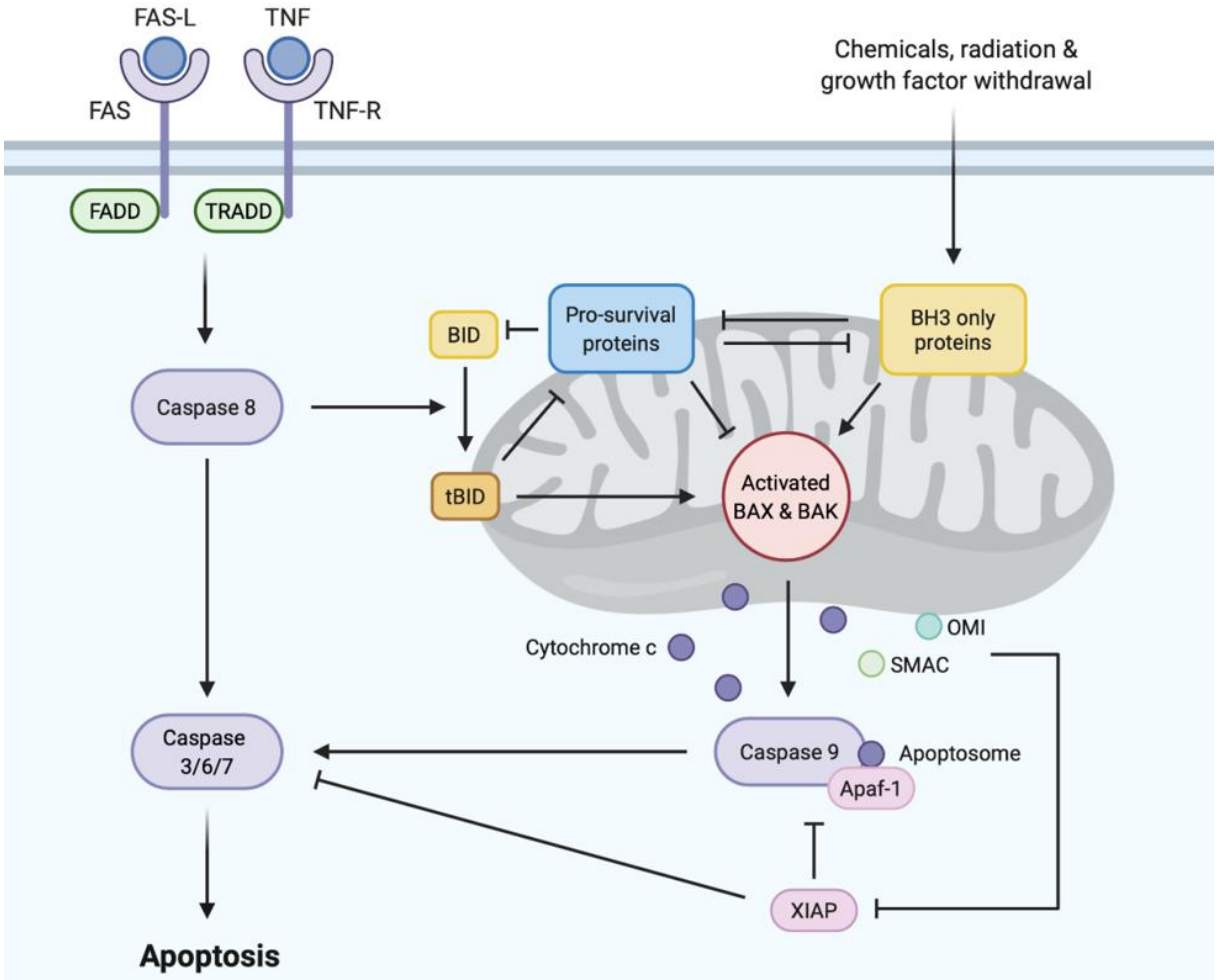

FAS-L
TNF
FAS
TNF-R
Chemicals, radiation &
growth factor withdrawal
FADD
TRADD
BID
Pro-survival
proteins
BH3 only
proteins
Caspase 8
tBID
Activated
BAX & BAK
Cytochrome c
OMI
SMAC
Caspase
3/6/7
Caspase 9
Apoptosome
Apaf-1
XIAP
Apoptosis

Figure 1.1. The extrinsic and intrinsic pathways to apoptosis.

The extrinsic pathway is initiated when death receptor ligands (e.g., FAS-L, TNF) bind to their cognate death receptors (e.g. FAS, TNFR1) on the plasma membrane, resulting in activation of caspase-8 via FAS-associated death domain protein (FADD) with or without TNFR- associated death domain protein (TRADD). The intrinsic pathway is triggered when diverse stress signals activate pro-apoptotic BH3-only proteins, which carry out their pro- apoptotic function by neutralising pro-survival BCL2 family proteins, displace activated BAX and BAK, thus inducing MOMP. MOMP results in the release of a range of apoptogenic factors, including cytochrome c, from the intermembrane space of the mitochondria into the cytoplasm. These factors trigger APAF1 mediated activation of caspase-9, and SMAC, which prevents XIAP from inhibiting its caspase target. The extrinsic and intrinsic pathways converge with initiator caspases (e.g., caspase-8, caspase-9) activating executioner caspases (e.g., caspase-3, caspase-7), which mediate cellular destruction. Caspase-8 also process the BH3-only protein BID, the truncated BID (tBID) can then activate the intrinsic pathway.

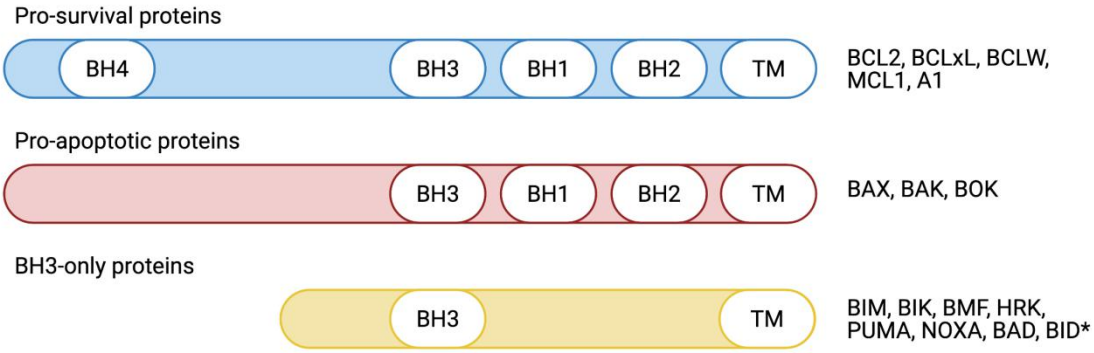

Pro-survival proteins
BH4
BH3
BH1
BH2
TM
BCL2, BCLxL, BCLW, MCL1, A1
Pro-apoptotic proteins
BH3
BH1
BH2
TM
BAX, BAK, BOK
BH3-only proteins
BH3
TM
BIM, BIK, BMF, HRK, PUMA, NOXA, BAD, BID*

Figure 1.2. Classification of the BCL2 family proteins.

BCL2 family proteins are divided into three subgroups: pro-survival proteins, pro-apoptotic proteins and BH3-only proteins. All BCL2 family proteins share regions of sequence homology called the BCL2 homology (BH) domains. The pro-survival and pro-apoptotic proteins share up to four such regions (BH1-BH4), while BH3-only proteins share only the BH3 domain. Most family members, except BID, contain a C-terminal transmembrane (TM) region, which allows them to be anchored on the outer mitochondria membrane, either constitutively or after apoptotic stimuli.

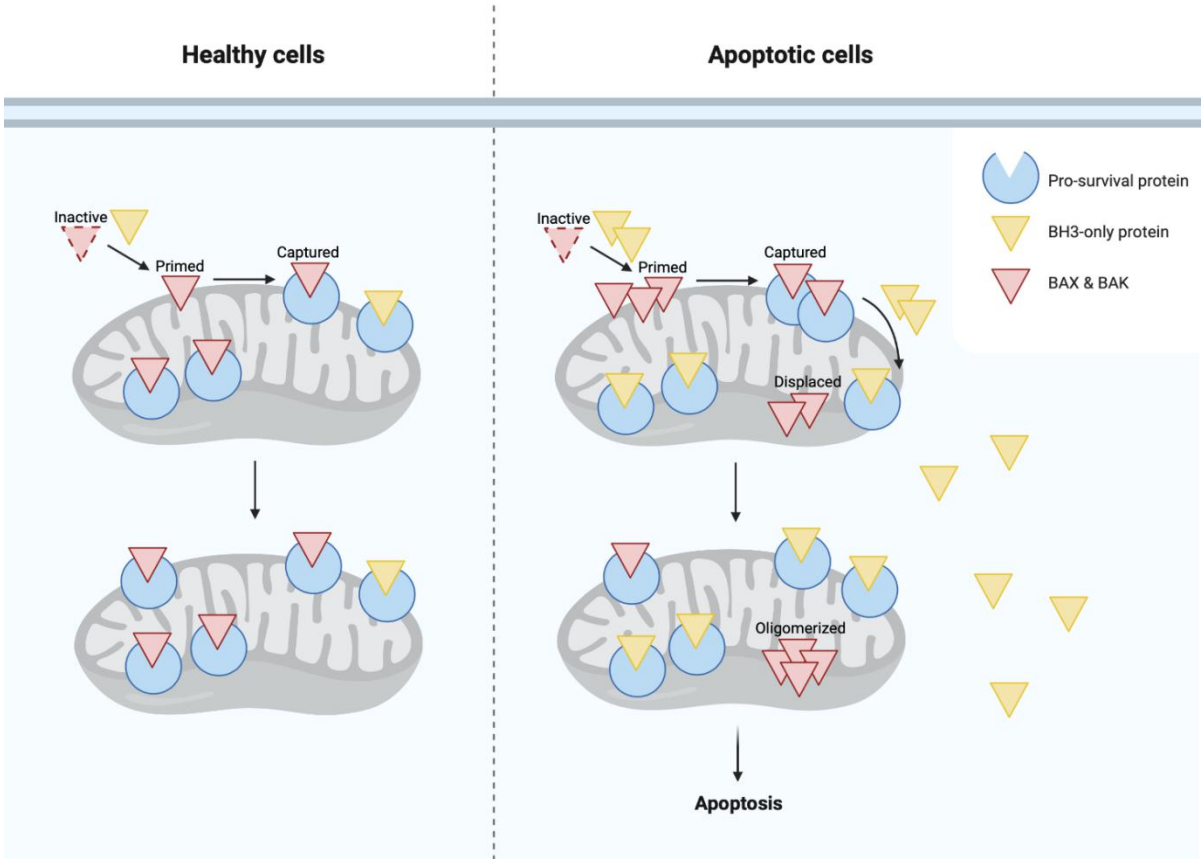

Healthy cells
Apoptotic cells
Inactive
Primed
Captured
Displaced
Oligomerized
Apoptosis
Pro-survival protein
BH3-only protein
BAX & BAK

Figure 1.3. The "Priming-capture-displacement" model[1] of BCL2 family proteins interaction.

In healthy cells, most pro-apoptotic proteins, shown in red, are kept in their inactive forms. They can be primed spontaneously or by existing BH3-only proteins, shown in yellow. However, the primed pro-apoptotic proteins are immediately captured by a pro-survival relative, shown in blue. In healthy cells, there are sufficient pro-survival proteins to sequester the primed BAX and BAK. Upon receiving apoptotic signals, BH3-only proteins are upregulated. As consequence, more BAX and BAK are either primed or displaced from pro-survival proteins by BH3-only proteins. The displaced BAX and BAK can then form dimers and higher oligomers to induce MOMP.

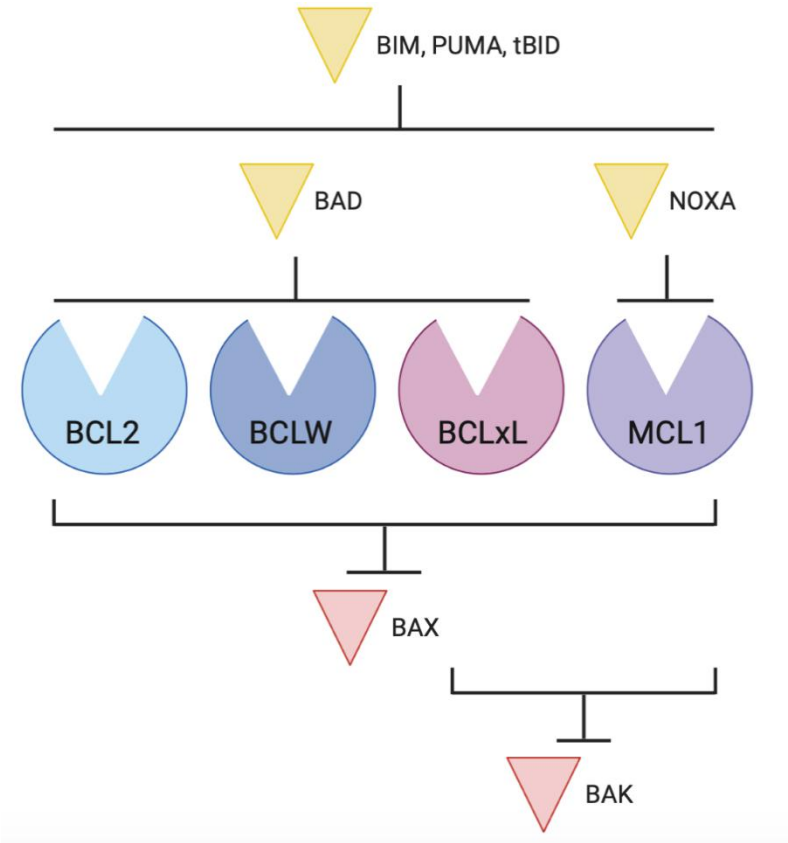

BIM, PUMA, tBID
BAD
NOXA
BCL2
BCLW
BCLxL
MCL1
BAX
BAK

Figure 1.4. Binding selectivity of BH3-only proteins to pro-survival members of BCL2 family. BH3-only protein BIM, PUMA and BID bind with high affinity to all pro-survival proteins, BAD predominantly binds to BCL2, BCL-XL and BCLW, but poorly to MCL1, NOXA binds strongly towards MCL1, but not the other pro-survival proteins. BAX interacts with most pro-survival proteins, while BAK prodominantly interacts with to BCL2 and BCLW.

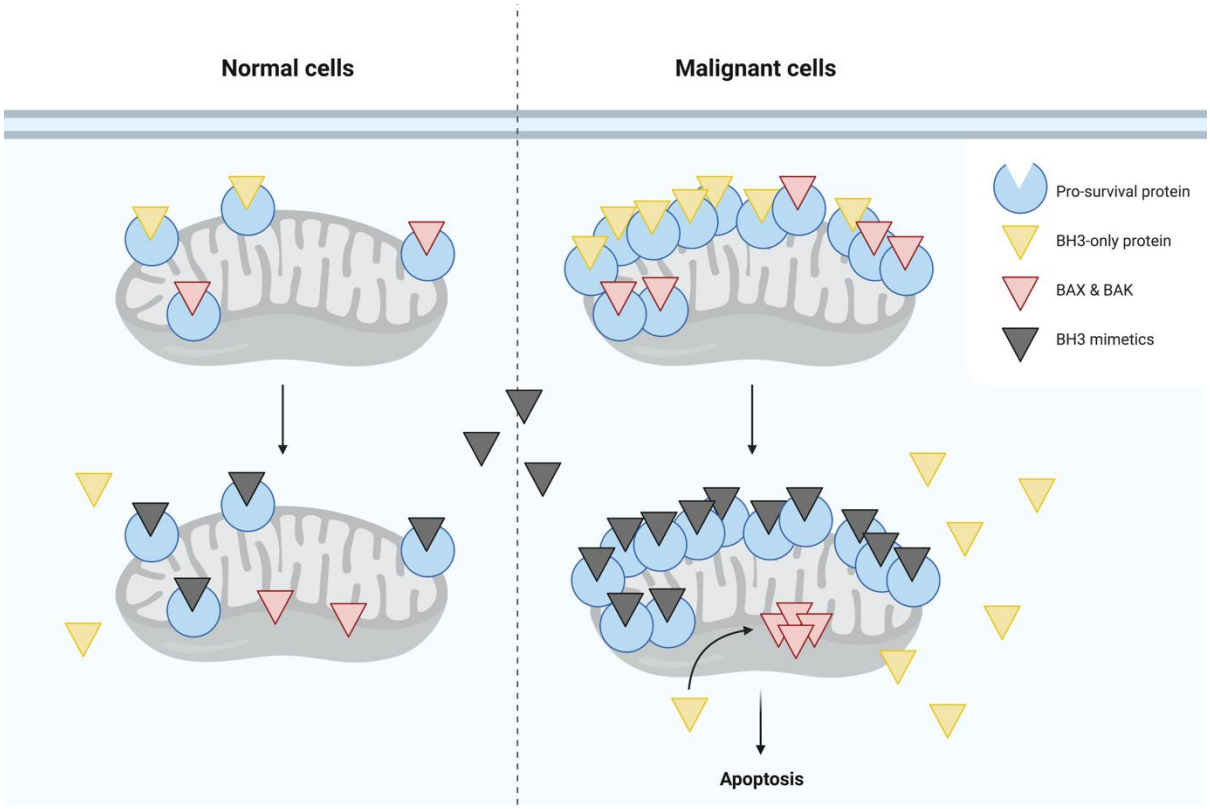

Normal cells
Malignant cells
Pro-survival protein
BH3-only protein
BAX & BAK
BH3 mimetics
Apoptosis

Figure 1.5. Mechanism of action of BH3 mimetics.

Pro-survival proteins, shown in blue, are commonly overexpressed in cancer cells, where they sequester high levels of BH3-only proteins, shown in yellow, and pro-apoptotic proteins BAX and BAK, shown in red, through their BH3 motif to maintain cell survival. These cells are paradoxically "primed" for death, as inhibition of upregulated pro-survival proteins by BH3 mimetics would liberate large quantities of sequestered BH3-only proteins and pro-apoptotic proteins, driving the cell toward apoptosis.

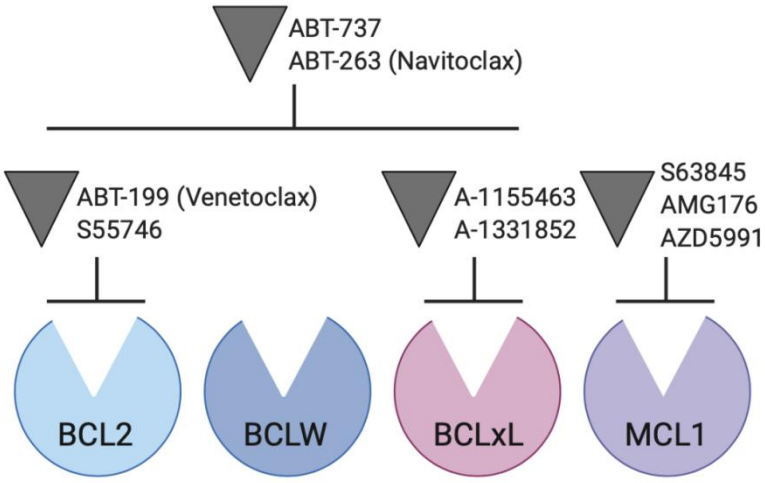

ABT-737
ABT-263 (Navitoclax)
ABT-199 (Venetoclax)
S55746
A-1155463
A-1331852
S63845
AMG176
AZD5991
BCL2
BCLW
BCLxL
MCL1

Figure 1.6. Different BH3 mimetics target different pro-survival proteins.

ABT-737 is the first BH3 mimetic compound, binding to BCL2, BCLW and BCL-XL. An orally bioavailable version called navitoclax (ABT-263) was later developed. A structural analogue of navitoclax known as venetoclax was then developed. Venetoclax binds more tightly to BCL2 in comparison to BCLW and BCL-XL. Similar to venetoclax, S55746 also shows high affinity to BCL2 while sparing BCLW and BCL-XL. Success of venetoclax paved the way to the development of inhibitors that selectively target other pro-survival BCL2 proteins. These include A-1155463 and A-1331852 that selectively bind to BCL-XL, and S63845, AMG176 and AZD5991 that selectively bind to MCL1.

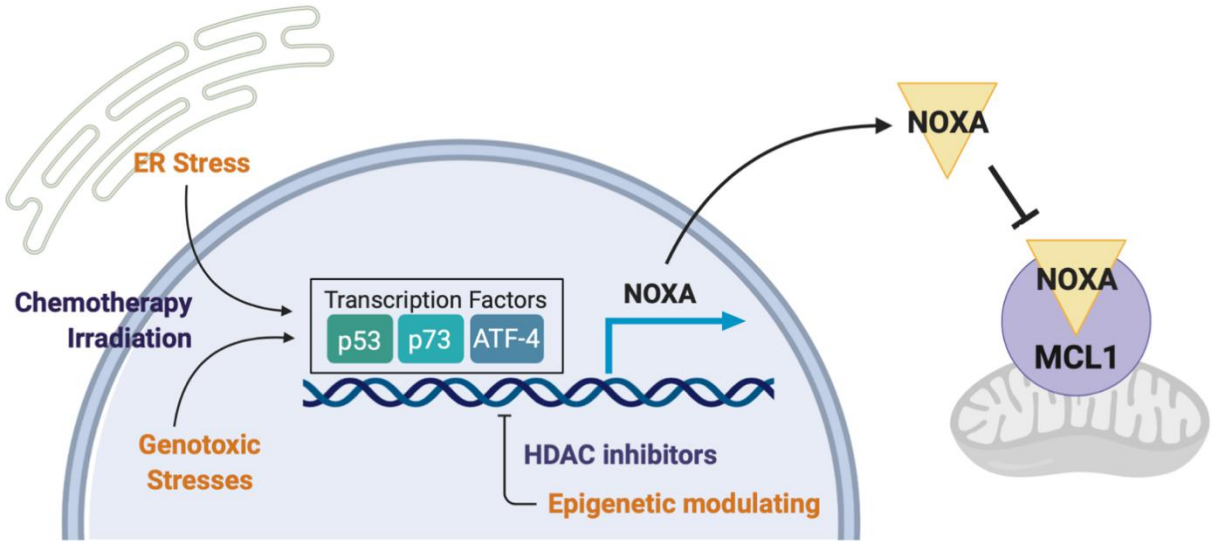

ER Stress
Chemotherapy
Irradiation
Genotoxic
Stresses
Transcription Factors
p53
p73
ATF-4
NOXA
HDAC inhibitors
Epigenetic modulating
NOXA
NOXA
MCL1

Figure 1.7. Known regulators of *NOXA*.

NOXA was initially identified as a p53-responsive gene, upregulated in response to DNA damage triggered by classical chemotherapy and irradiation which leads to p53 activation. Further studies have demonstrated transcription factors other than p53 can also regulate *NOXA*, including p73 and ATF4. Most pathways that induce *NOXA* are activated in response to cellular stress, but recent studies suggest that *NOXA* is also regulated by epigenetic modifications. Several small molecule drugs that modulate histone methylation and acetylation have shown potent impact in killing cancer cells, and in some cases, this has been attributed to increased *NOXA* expression.

Chapter 2: Materials and methods

2.1 Cell lines and cell culture

The human cell lines derived from acute lymphoblastic leukaemia (ALL) RS4;11, multiple myeloma (MM) KMS-12-PE, acute myeloid leukemia (AML) MV4;11, MOLM13, diffuse large B cell lymphoma (DLBCL) OCI-Ly19 and mantle cell lymphoma (MCL) Granta-519 were cultured in RPMI-1640 supplemented with 10% (v/v) fetal calf serum (FCS) at 37°C with 5% CO_2. The human embryonic kidney cell line HEK293T was cultured in DMEM supplemented with 10% (v/v) FCS at 37°C with 10% CO_2. All media contained 1% penicillin and streptomycin. All cell lines were regularly tested for mycoplasma contamination using the MycoAlert mycoplasma detection kit (Lonza #LT07) and were consistently negative for infection.

2.2 Mammalian expression constructs

All constructs used in this thesis are listed in Table 2.1.

The sgRNA expression constructs were generated by inserting the annealed DNA oligos into backbone plasmids[116]. The generated plasmids were amplified in Stbl3 E. coli and extracted by DNA Miniprep, Midiprep or Maxiprep kits (Thermofisher). The generated plasmids were validated by Sanger sequencing using U6 sequencing primer (Table 2.3). The sequences of the DNA oligos used in this thesis are listed in Table 2.2.

2.3 Generation of CRISPR-Cas9 gene editing cell lines

2.3.1 Transient transfection

Transient transfection of RS4;11 cells was done using an Amaxa Nucleofector shuttle apparatus and transfection kit R following the T-016 protocols (Amaxa, Germany). Transient transfection of KMS-12-PE cells also employed an Amaxa Nucleofector shuttle apparatus and transfection kit V following the X-001 protocols (Amaxa, Germany). Briefly, two million cells were pelleted by centrifugation and resuspended in 100 ul of transfection reagent after which 4.0 ug of DNA was added. The DNA contained CRISPR-Cas9 vectors (pX458, pX335 etc.)

with sgRNA expressing constructs. Resuspended cells in media containing DNA were transferred to a nucleo-cuvette module and transfected as per protocol T-016 or X-001 using the Amaxa shuttle instrument. Transfected cells were then transferred to 6-well plates and resuspended in 3.0 ml media. Successfully transfected cells were sorted base on fluorescent protein expression using the BD FACSAria machine.

2.3.2 Lentiviral infection

HEK293T cells were used to generate lentivirus for infection of human cells. Typically, HEK293T packaging cells were plated in a 10-cm dish at a density of 2.5 x 10^6 cells, 24 hours prior to transfection. The next day, 1.5 µg of Cas9 expression plasmid (FU-Cas9-Cherry, FU- Cas9-GFP, Lenti-Cas9-BFP etc.) was mixed with lentiviral packaging plasmids (1 µg pRSV- REV, 1.5 µg pMDLg and 1 µg pVSVG) and transfected into HEK293T using FuGENE6 reagent (Promega).

The culture media was replaced the following day. At 48 hours post-transfection, viral supernatants were filtered from the packaging cells using 0.45 µm syringe filter (Millipore) to remove cell debris.

To infect cells, one million target cells were seeded in 6-well plates prior to transduction. To enhance transduction efficiency, filtered viral supernatants were added to the target cells and centrifuged (2,500 rpm at 20 ^oC for 1 hour) in the presence of Polybrene (Sigma) at a final concentration of 5 µg/ml. 48h after infection, transduced cells were sorted based on fluorescent protein expression using the BD FACSAria machine.

sgRNA vectors (pKLV etc.) were sequentially transduced into cells utilising the same protocol as for the Cas9 vectors. Transduced cells were either sorted based on fluorescent protein expression using the BD FACSAria machine or selected based on puromycin resistance gene expression. The concentration and time of puromycin treatment were determined by puromycin titration experiments. In this thesis, puromycin treatment using 0.5ug/mL for 72h for RS4;11 cells and 1ug/mL for 72h for KMS-12-PE cells was established as being sufficient to select for sgRNA expressing cells.

Gene-editing in cells expressing both Cas9 and sgRNA was validated by PCR, next-generation sequencing, or other biochemical and functional assays.

2.4 PCR analysis of gene editing cell lines

Suspension cells were collected directly from the tissue culture media by centrifugation. For crude DNA extracts, cell pellets were lysed in Direct PCR lysis reagent (Viagen) supplemented with 5% (v/v) Proteinase K (Sigma). Approximately 5 million cells were lysed with 100uL of reagent. The mixture was incubated with the following parameters: 55^{o}C for 2 hours followed by 85^{o}C for 40 minutes. For purified DNA extracts, 5 million cells were lysed with DNeasy Blood and Tissue Kits (Qiagen) following the manufacturer's guidelines.

Unique primers (Table 2.3) to amplify the DNA sequence spanning the sites of CRISPR- mediated insertion were designed using the online tool PrimerBlast https://www.ncbi.nlm.nih.gov/tools/primer-blast/. To amplify DNA, 2 µl of cell lysate from crude extracts or 200 ng of genomic DNA from purified DNA extracts were suspended in 21 µl of GoTaq Green PCR mix containing 2 µl of sequence-specific primer pairs (forward and reverse). The 25 µl reactions were then incubated in an automated thermal cycler with the parameters according to manufacturer's guidelines. The DNA amplicons were then run on a 1% agarose gel with ethidium bromide and exposed under uv light using a GelDoc machine to visualise DNA products at the expected size.

2.5 DNA sequencing of CRISPR-Cas9 gene edited cell lines

2.5.1 Sanger sequencing

Using the same techniques as described in the PCR analysis section, targeted sites of CRISPR- mediated indels were amplified using pairs of sequence-specific primers (Table 2.3). The DNA amplicons were then run on a 1% agarose gel. DNA amplicons were purified using a gel extraction kit (Qiagen), mixed with sanger sequencing primers and subjected to Sanger sequencing (Australian Genome Research Facility (AGRF), Melbourne).

2.5.2 Next generation sequencing

As well as Sanger sequencing, the presence of CRISPR-mediated indels at the target DNA sites were also analysed using next-generation sequencing (Miseq). For this method, approximately 20 base pairs of overhang sequence were added to the 5' ends of both forward and reverse primers (Table 2.3). These overhang sequences were later used for

secondary PCR barcoding of DNA. CRISPR-targeted DNA sites were amplified by PCR with overhang primers. After the first PCR reaction, 2 µl of DNA amplicons were re-suspended in 21 µl of GoTaq Green PCR containing 2 µl of the unique barcoded primer pairs. The resulting 25 µl mixtures were then incubated in the automated thermal cycler. The cycling conditions were described in table 2.4. The size of the barcoded DNA amplicons was expected to be approximately 50 base pairs longer than the non-barcoded amplicons from the first PCR.

To purify DNA for Miseq analysis, 5 µl of DNA amplicons from each second PCR reaction were mixed in a microcentrifuge tube and incubated with AMPure XP beads in a ratio of 1: 1 (v/v) for 5 minutes. Beads containing the barcoded DNA amplicons were separated from non- specific PCR products by placing the microcentrifuge tube into a magnetic rack. These beads were then washed with 80% (v/v) ethanol twice and the supernatants were carefully removed in between washes to ensure that the beads were left undisturbed. Finally, the beads were left to air-dry for 15 minutes. 100 µl of nuclease-free water was then added to elute the attached DNA. The purified barcoded DNA amplicons were then subjected MiSeq analysis.

2.6 Western blot

Suspension cells were collected directly from the tissue culture media by centrifugation. Following harvest, cells were washed once in PBS and centrifuged at 1500 rpm at 4^{o}C for 5 minutes to obtain cell pellets. Cell pellets were lysed with lysis buffer containing: 20 mM Tris pH7.8, 135 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 10% (v/v) glycerol, 1% (v/v) Triton X-100 and 0.5 µg/ml protease inhibitors (Roche Life Science), for 30 minutes on ice. The lysates were centrifuged at 13,000 rpm at 4^{o}C for 5 minutes to remove insoluble materials. Supernatants containing whole cell protein extracts were collected and the Bradford assay was used to quantify protein concentration.

For whole cell protein extracts, 4 x SDS-PAGE sample reducing buffer (0.25M Tris pH 5.8, 1% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 20% (v/v) glycerol and 0.02% (v/v) bromophenol blue) was added to each protein lysate. The mixture was heated for 5 minutes at 95^{o}C, loaded into a 4-12% pre-cast NuPAGE Bis-Tris gel (Invitrogen) and run in 2-(N-morpholino) ethane sulfonic acid (MES), sodium dodecyl sulphate (SDS) running buffer (50 mM MES pH 7.4, 50mM Tris Base, 0.1% (w/v) SDS and1 mM EDTA) at 80V for the first 10 minutes and at

100V for the remaining time. Proteins in the gels were then transferred to nitrocellulose membrane using an iBlot Gel Transfer Device (Invitrogen) at 23V for 7 minutes.

Proteins in the nitrocellulose membrane were detected by chemiluminescent assay. After protein transfer, membranes were incubated in blocking buffer for 1 hour in a rotating tube. Milk blocking buffer (5% (w/v) skim milk powder, 0.1% (v/v) Tween in TBS) were used. Membranes were then incubated with various primary antibodies (1 to 10 µg/ml) (Table 2.5) rotating overnight at 4^{o}C. This was followed by incubation of secondary antibodies for 1 hour at room temperature (Table 2.6). Between antibody incubations, membranes were washed three times with TBS-Tween (0.01% (v/v)) with gentle rocking for 5 minutes each time at room temperature (RT). Finally, proteins on the membrane were detected using Luminata Forte Western Horseradish peroxidase (HRP) substrate (EMD Millipore) and visualised on the ChemiDoc Imager System (BioRad). 2x diluted HRP substrate was used for HSP70, 1x HRP substrate was used for MCL1, NOXA etc.

2.6 Flow cytometry analysis and cell sorting

Suspension cells were collected directly from tissue culture media. Cells were then transferred into fluorescence-activated cell sorting (FACS) tubes and centrifuged at 1500 rpm at 4^{o}C for 5 minutes. For flow cytometry, cells were resuspended in FACS buffer (KDS BSS + 2% FCS), incubated with antibodies or stains and then analysed using either the LSRIIW or Fortessa FACS machines. FACS data were analysed using FlowJo software. For cell sorting, cells were resuspended in cell culture media and sorted by using the Aria Fusion cell sorter, into either FACS tubes or 96-well plates for single clonal cultures. Cells sorted into FACS tubes were centrifuged again and resuspended in cell culture media.

2.7 Quantitative reverse transcription-PCR (qRT-PCR) analysis

Firstly, total RNA was extracted from cultured cells using the RNeasy kit (Qiagen) following the manufacturers guidelines. The RNA concentration was then measured using a NanoDrop (Thermo Scientific) and stored at -80°C.

The isolated total RNA was then reverse transcribed into single strand cDNA using the SuperScript First-Strand Synthesis System kit for RT-PCR (Invitrogen). 1ug of total RNA, 1 µL oligo dT (50 ng/mL) and 1 µL dNTPs (10 mM) were mixed together in a PCR tube and made

up to a total volume of 13 µL with distilled water. The mixture was then incubated at 65°C for 5 min, followed by cooling on ice for 2 min. Next, 4µL of first strand buffer (5x), 1 µL DTT (0.1 M), 1 µL RNasin (40 U/mL) and 1 µL Superscript III (200 U/mL), were added to the sample solution. The mixed solution was then incubated in a thermocycler for cDNA synthesis according to the program following program: 50°C 60 min, 70°C 15 min. The solution containing cDNA was stored at -20°C.

All qRT-PCR analyses were performed using the QuantStudio 12K Real-time PCR machine (Applied Biosystems) with probes purchased from Taqman Gene Expression assays (Applied Biosystems). To set up the 384-well qRT-PCR plates, 2uL of cDNA, 1uL of probe, 10uL of TaqmanTM Universal PCR Mastermix (Applied Biosystems) and 7uL of distilled water were mixed per well. Then the Standard program was used to run the qRT-PCR on the QuantStudio 12K. The expression levels of genes of interest were then standardised to the levels of the reference genes *GAPDH* or *PRMT1* from the same sample. Gene expression levels could then be compared between treated and untreated cells.

2.8 Targeted bisulfite sequencing

Suspension cells were harvested from culture media and DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen) following the manufacturer's guidelines. Next, 200ng purified DNA were aliquoted into 28uL water for bisulfite conversion. The EM-seq conversion module (NEB) was utilised for the conversion in the following steps according to the manufacturer's guidelines: Oxidation of 5-Methylcytosines and 5-Hydroxymethylcytosines by TET2; clean-up of TET2 converted DNA; denaturation of DNA by formamide; deamination of cytosines and clean-up of deaminated DNA. The converted DNA was stored at -20°C.

Bisulfite specific primers were designed by the online tool MethPrimer https://www.urogene.org/methprimer/ (Table 2.3). PCR amplification was performed using the bisulfite specific primers plus the converted DNA and Phusion high-fidelity DNA polymerase (ThermoFisher) following manufacture's guidelines. The amplicons were assessed on a 1% agarose gel containing ethidium bromide.

Amplicons were next indexed by native barcoding. Briefly, 20ng of amplicon DNA from each sample was aliquoted in a total of 25uL of water. An end-prep reaction was set up for each sample by adding 3.5uL of Ultra II End Prep Reaction Buffer and 1.5uL Ultra II End Prep

Enzyme Mix (NEB). Samples were incubated at RT for 10mins, then at 65°C for 10mins and finally on ice for 30secs. 2.5uL of specific NBXX barcode, 20uL of Ultra II Ligation Master Mix and 1uL of Ligation Enhancer (NEB) were added directly to the previous reactions. Samples were incubated at RT for 20mins and then 70°C for 5mins. The barcoded samples were pooled together, followed by bead clean-up using Ampure XP beads. Next, the adapter ligation reaction was set up by adding 2.5uL Adapter Mix II, 10uL Quick Ligation Reaction Buffer (5x) and 5uL Quick T4 DNA ligase (NEB) to 40ng of barcoded sample in 32.5uL water. The sample was incubated at RT for 10mins, then cleaned up by Ampure XP beads to a final volume of 7uL.

The sequencing mix for nanopore sequencing was prepared by mixing 15uL Sequencing Buffer II, 10uL Loading Beads II and 5uL ligated DNA library. The sequencing mix was loaded into a Flongle flow cell (Nanopore Technologies) for sequencing.

2.9 Cell death assay

2.9.1 Annexin V staining and propidium iodide assay

Suspension cells were collected from culture media and centrifuged at 1500 rpm at 4^{o}C for 5 minutes. Both viable and dead cells were then resuspended in 100uL FACS buffer containing APC-conjugated anti Annexin V (BD Biosciences) and 2ug/mL propidium iodide (PI, Sigma- Aldrich). Cell viability was calculated as the percentage of viable cells (APC and PI double negative) using flow cytometry.

2.9.2 CellTiter-Glo assay

Approximately 5 x 10^3 cells were plated per well in an opaque-walled 96-well plate (Greiner). Drugs and DMSO with the desired range of concentrations were added in a total volume of 100 µl of culture media per well. The 96-well plate was then cultured for 24 hours at 37^{o}C and 5% CO2. On the day of analysis, 40 µl of CellTiter-Glo reagent (Promega) was added into each well. The solutions were then mixed for 10 minutes on an orbital shaker in an opaque container. Next, the 96-well plate was analysed using a LUMIstar OPTIMA luminometer (BMG Labtech). Cell viability was calculated as the luminescence normalised to DMSO controls.

2.10 Data analysis

The IC50 (The half maximal inhibitory concentration) values were calculated using Prism 9 (GraphPad Software). The FACS data were analysed using FlowJo software. Next-generation sequencing data from genome-edited cell lines was analysed using the Corol-DNA comparison tool.

A brief summary of the statistical analysis of CRISPR-Cas9 screens is described below:

Raw counts from each sgRNA read were normalised to counts per million (CPM) and transformed into a log2 scale. An offset buffer of 10 CPM was added to each sgRNA read in order to buffer against extreme log2 fold-changes for sgRNAs with low read counts. The frequency of each sgRNA read was averaged across biological replicate samples between the enriched and baseline values. The enriched values were obtained from sorted mCherry-high or mCherry-low cells in the CRISPR-Cas9 screens in *NOXA* reporter cells.

The mean log2 fold-change for each sgRNA read was then calculated as the difference between the mean of sgRNA frequency in enriched and baseline samples. The mean log2 fold- changes were ranked in descending order to identify hits. Minimum hypergeometric p- values were then calculated from the ordered rank of sgRNA for each of the genes represented in the library using an established algorithm[117]. Finally, false discovery rates (FDR) were calculated by adjusting the hypergeometric p-values using the Benjamini-Hochberg procedure.

Table 2.1 List of mammalian expression constructs

Plasmid name	Description
pX335	hCas9 (D10A) and sgRNA transient expression plasmid.
pX335 NOXA-KI sgRNA1+	pX335 with sgRNA that targets *NOXA* (sense strain).
pX335 NOXA-KI sgRNA1-	pX335 with sgRNA that targets *NOXA* (antisense strain).
pX335 NOXA-KI sgRNA2+	pX335 with sgRNA that targets *NOXA* (sense strain).
pX335 NOXA-KI sgRNA2-	pX335 with sgRNA that targets *NOXA* (antisense strain).
pUC57_mCherry-template	Template plasmid that contains *NOXA* homology arms and mCherry template for HDR.
pX458	hCas9 and sgRNA transient expression plasmid that contains GFP marker.
pX458 hsBAX*	pX458 with sgRNA that targets *BAX*.
pX458 hsBAK*	pX458 with sgRNA that targets *BAK*.
FU-Cas9-GFP	hCas9 lentivirus expression plasmid that contains GFP marker.
Lenti-Cas9-BFP	hCas9 lentivirus expression plasmid that contains BFP marker.
pKLV	sgRNA lentivirus expression plasmid that contains sgRNA insertion site, BFP and puromycin resistance markers.
pKLV-TL-sgRNA078	pKLV with sgRNA078 from the *NOXA* tiling library.
pKLV-TL-sgRNA101	pKLV with sgRNA101 from the *NOXA* tiling library.
pKLV-TL-sgRNA267	pKLV with sgRNA267 from the *NOXA* tiling library.
pKLV-TL-sgRNA283	pKLV with sgRNA283 from the *NOXA* tiling library.
pKLV-TL-sgRNA287	pKLV with sgRNA287 from the *NOXA* tiling library.

*These plasmids were generated by Jianan Gong.

pX335 backbone plasmid was a gift from Feng Zhang (Addgene plasmid # 42335). FU- Cas9- GFP plasmid was a gift from Macro Herold. Lenti-Cas9-BFP plasmid was a gift from Roderic Guigo & Rory Johnson (Addgene plasmid # 78547).

Table 2.2 List of DNA oligos

Genes or Numbers	Sequence (5’ to 3’)	Vector
NOXA 1+	CGGCACCGGCGGAGATGCCT	pX335
NOXA 1-	CCCACTCAGCTACAGAGCCC	pX335
NOXA 2+	ACGCTCAACCGAGCCCCGCG	pX335
NOXA 2-	TTCTTCCCAGGCATCTCCGC	pX335
sgRNA078	ACGACAGGCGTGCACCACCA	pKLV
sgRNA101	AAGAGCTGGGATTACAGGCG	pKLV
sgRNA267	CGCTCCCATAACGCCGTCTG	pKLV
sgRNA283	CCCCTGTCCCCGCCCCTGTC	pKLV
sgRNA287	GGCGTCTAGTTTCCCTACGT	pKLV

Table 2.3 List of primer sequences

Name	Direction	Sequence
U6 sequencing primer	Forward	GAGGGCCTATTTCCCATGATT
NOXA reporter_Primer1	Forward	GCAGGACTGTTCGTGTTCAG
	Reverse	TTGAAGCGCATGAACTCCTTG
NOXA reporter_Primer2	Forward	CACGAGTTCGAGATCGAGGG
	Reverse	CTGGGGAAAAACCCCGAAGA
NOXA reporter_Primer3	Forward	CACTGGACAAAAGCGTGGTC
	Reverse	CGCTGGCCTGAAAACTTACG
NOXA reporter_Primer4	Forward	TGAAGCGGCTCTCAGTAACC
	Reverse	AGCAGCTTGAGGTCTTTCGG
NOXA reporter_Primer5	Forward	GTTCATGCGCTTCAAGGTGC
	Reverse	GCCGTACATGAACTGAGGGG
NOXA reporter_NGS_Primer1	Forward	GTGACCTATGAACTCAGGAGTCGCTAAACATCCACAATGGGCG
	Reverse	CTGAGACTTGCACATCGCAGCGGGAAGGGTTTAACCAGGAGG
NOXA reporter_NGS_Primer2	Forward	GTGACCTATGAACTCAGGAGTCTGTGTGGGCCATACTGCTG
	Reverse	CTGAGACTTGCACATCGCAGCAAGAGGGACTTTGTGCTGCT
pKLV_NGS	Forward	GTGACCTATGAACTCAGGAGTCTCACCGTGTGTAGTTGGCAT
	Reverse	CTGAGACTTGCACATCGCAGCAGACCTGGCCCGACCTT
TLseq-1	Forward	GTGACCTATGAACTCAGGAGTCGCACTTCTAGGGAGCCCATGCA
	Reverse	CTGAGACTTGCACATCGCAGCGAGGCGGGTGGTTCGTTTGAGG
TLseq-2	Forward	GTGACCTATGAACTCAGGAGTCCACCCAGGCTGGAATGCAGTGG
	Reverse	CTGAGACTTGCACATCGCAGCCTGATTCCGAACCAGGCAGGCC
TLseq-3	Forward	GTGACCTATGAACTCAGGAGTCAACCTCCGCCTCCCAGGTTCAA

	Reverse	CTGAGACTTGCACATCGCAGCCTGATTCCGAACCAGGCAGGCC
TLseq-4	Forward	GTGACCTATGAACTCAGGAGTCGCACCACCAAGGCCGGCTAATC
	Reverse	CTGAGACTTGCACATCGCAGCCTGATTCCGAACCAGGCAGGCC
TLseq-5	Forward	GTGACCTATGAACTCAGGAGTCCCTCTGGCATCTCCTCTGCGGT
	Reverse	CTGAGACTTGCACATCGCAGCGCGCTCGGCCCCGAAATTACTT
TLseq-6	Forward	GTGACCTATGAACTCAGGAGTCTGCCAAGGCCTCTGGTCTCTCC
	Reverse	CTGAGACTTGCACATCGCAGCTAGGGAAACTAGACGCCCGGCCC
TLseq-7	Forward	GTGACCTATGAACTCAGGAGTCAAGTAATTTCGGGGCCGAGCGC
	Reverse	CTGAGACTTGCACATCGCAGCGGACGCGAGCTGAACACGAACA
TLseq-8	Forward	GTGACCTATGAACTCAGGAGTCTCCGCTCCCATAACGCCGTCTG
	Reverse	CTGAGACTTGCACATCGCAGCGGACGCGAGCTGAACACGAACA
*NOXA*_BSP	Forward	TTATTTAGGTTGGAATGTAGTGG
	Reverse	AACAAAAACTCTTACCCACTATAACCTT

Table 2.4 General PCR cycling condition

Cycle	Step	Temperature (°C)	Time (seconds)	Repeats (x)
Initial Denaturation		95	300	1
Thermal cycler	**Denaturation**	95	30	35
	Annealing	55-60*	30	
	Extension	72	30-180**	
Final extension		72	300	1
*Annealing temperature varies based on the GC contents of primers pair **Extension time varies based on the size of DNA products				

Table 2.5 List of primary antibodies

Antigen	Dilution	Reactivity	Species	Clonal	Supplier
BAK	1:1000	Human	Rat	Mono	WEHI Antibody Services
BAX	1:1000	Human	Rat	Mono	WEHI Antibody Services
BCL2	1:1000	Human	Mouse	Mono	WEHI Antibody Services
BCL-XL	1:1000	Human	Mouse	Mono	WEHI Antibody Services
HSP70	1:4000	Human	Mouse	Mono	WEHI Antibody Services
MCL1	1:1000	Human	Rat	Mono	WEHI Antibody Services
NOXA	1:400	Human	Mouse	Mono	Novus Biologicals

Table 2.6 List of secondary antibodies

Antigen	Dilution	Conjugate	Species	Supplier
Mouse IgG	1:5000	HRP	Goat	WEHI Antibody Services
Rat IgG	1:5000	HRP	Goat	WEHI Antibody Services

Chapter 3: Generation and validation of *NOXA* reporter cell lines

3.1 Introduction

3.1.1 The CRISPR-Cas9 system

CRISPR-Cas9 is a revolutionary technology in modern biology. It enables researchers to conduct precise editing of the genome by deleting, adding or changing part of the DNA sequence, making genetic manipulation much simpler and more precise. This makes it a powerful tool for interrogating genetic pathways important for human cancer. Clustered regularly interspaced short palindromic repeats (CRISPR) and an associated protein Cas together form a bacterial adaptive immune system that uses RNA-guided nucleases to cleave foreign genetic elements[118]. Recently, this system has been developed into powerful genome engineering tools that can be used in eukaryotic cells. CRISPR-Cas9 is one of the best characterised systems, comprising the Cas9 endonuclease and the CRISPR RNA (crRNA) array. The crRNA encodes a 20-bp small guide RNA (sgRNA) and a unique array of repetitive elements which are derived from exogenous DNA targets[119]. The 20-bp sgRNAs direct Cas9 endonuclease to the targeted DNA and the repetitive elements derived from exogenous DNA recruits the Cas9 endonuclease to induce double-strand break (DSB) at the targeted site. To note, in the commonly used CRISPR-Cas9 system derived from *Streptococcus pyogenes*, the target DNA must immediately precede a 5'-NGG protospacer adjacent motif (PAM) site for efficient cutting to happen[120].

The DSBs at the targeted site usually induce DNA damage repair via two alternative major repair pathways. When a DNA repair template is absent, error-prone nonhomologous end joining (NHEJ) happens at higher frequencies. NHEJ repairs DNA by re-ligating two ends of the broken DNA with random insertions or deletions. This has proved useful in generating gene knockouts both *in vitro* and *in vivo*, as insertions or deletions in coding exons can lead to premature stop codons or frameshift mutations. Homology-directed repair (HDR) is another repair pathway, which occurs at lower frequencies than NHEJ under normal

physiological conditions. However, in the presence of exogenous DNA repair template, HDR can be utilised to generate precise modifications in the genome. The repair template can either be double- stranded DNA for integration of sequence of interest into the genome, or single- stranded DNA oligonucleotides for making single-nucleotide edits in the genome. Unlike NHEJ, the efficiency of HDR is generally lower and can vary widely depending on the cell lines, the genomic locus and the repair template[116].

3.1.2 Generation of fluorescent knock-in cells using CRISPR-Cas9 system

Tagging proteins at an endogenous locus with fluorescent markers is a powerful tool for measuring the cell-specific expression of a gene. Harnessing the HDR repair pathway described above has become revolutionary in recent years in achieving this. Researchers are now able to insert such markers into the loci of their genes of interest with relative speed and efficiency, transforming our ability to understand gene regulatory mechanisms. In this thesis, a paired Cas9 D10A nickase method described by *Kock et al.* was used to generate homozygously tagged *NOXA* reporter cell lines, followed by a validation pipeline[121].

Cas9 D10A nickase harbours a D10A mutation that inactivates its RuvC nuclease domain, impairing the cleavage of the target strand. The HNH nuclease domain remains intact and cleaves the non-target strand. As a result, Cas9D10A cut only one DNA strand, generating single-strand break (SSB) that can be repaired via HDR pathway without introducing indels or mutations. The paired Cas9D10A nickase system comprises a mutant Cas9 with a pair of sgRNAs targeting the sense and antisense DNA strands of the gene of interest, introducing targeted DSBs with minimised off-target effects[122]. In the presence of a DNA repair template containing two homology arms of the breakpoint flanking the sequence of a fluorescent marker, the DSBs are repaired by HDR pathways with sequence of the fluorescent marker inserted into the gene of interest. This system is reported to be capable of effectively generating cell lines with endogenous fluorescent tags while minimising off-target effects[121]. (Figure3.1.)

The aim of my study is to uncover regulatory pathways controlling *NOXA* expression that could be targeted to induce *NOXA* expression in malignant cells where excess MCL1 underpins drug resistance. To achieve this goal, I created a cell line in which changes in *NOXA* expression could be measured readily. Therefore, I generated and validated *NOXA* reporter cell lines using the paired Cas9D10A knock-in approach.

3.2 Result

3.2.1 Design of paired gRNAs and donor plasmids for targeting the *NOXA* locus

To generate the *NOXA* reporter cell line, sgRNAs and donor templates were designed to tag *NOXA* with a *mCherry* fluorescent marker at the N terminus. Two pairs of sgRNAs were designed to target the DNA strands near the *NOXA* start codon. For each pair, one sgRNA targeted the antisense DNA strand and one targeted the sense DNA strand, introducing DSB when worked together with Cas9D10A. sgRNA pairs were designed using the online tool Benchling, based on their on-target and off-target scores and proximity to the start codon. (Figure3.2. A) These sgRNA pairs were then inserted into the Cas9/sgRNA expression plasmid containing a Cas9D10A nickase optimised to human codons and a crRNA array with sgRNA insertion site followed by repetitive elements[123].

The donor plasmid served as DNA template contained sequence of *mCherry* fluorescent marker flanked by two 800-bp homology arms from upstream and downstream of the *NOXA* start codon respectively. The *mCherry* marker is inserted immediately after the *NOXA* start codon and in-frame with the remainder of the *NOXA* gene. Additionally, a nuclear localization signal (NLS) was placed at the 5' end of *mCherry* to localise the synthesised fluorescent protein into the nucleus for improved retention in apoptotic cells, and a T2A linker was placed at the 3' end of *mCherry*. In the reporter cells with the modified *mCherry-NOXA* locus, the genes are transcribed as one mRNA message, but the peptides are translated into separate proteins due to the ribosome skipping effect of the 2A self-cleaving peptides[124]. To note, the sgRNA target sites on the designed DNA template were different from the endogenous gene. Synonymous mutations were introduced during design so that the sgRNA would not cut the donor sequence. The DNA template was synthesised and cloned into pUC57-mini. (Figure 3.2. B).

3.2.2 Establishing the system: functionality test of gRNAs

Prior to delivering sgRNAs and DNA template to the cells, a preliminary experiment was conducted using different pairs of sgRNAs to test the efficiency of the designed sgRNAs in targeting the *NOXA* locus.

KMS-12-PE cells were transfected with a GFP expression plasmid in combination with either two paired sgRNA-Cas9D10A plasmids or one Cas9D10A empty vector. The GFP expression plasmid served as selection marker to identify and sort successfully transfected cells, as the Cas9D10A plasmid does not itself contain mammalian selection markers. KMS-12-PE cells were harvested 48h after transfection, GFP-positive cells were sorted, expanded and analysed by Western blot and Next generation sequencing. The efficiency of each pair of gRNAs was evaluated by their ability to introduce indels (deletions, insertions or point mutations) in the targeted DNA region and thereby alter NOXA protein levels.

Western blot results showed reduction of NOXA protein level in both pools of cells transfected with pair1 and pair2 sgRNAs compared to empty vector. The reduction was more obvious with pair2 sgRNAs than pair1. Further analysis with next generation sequencing showed with both pairs, up to 90% of the reads exhibited mutations or deletions at the targeted region of *NOXA*. However no obvious difference in the percentage of indels was seen between pair1 and pair2 sgRNAs. Based on these results, both pairs of sgRNAs were selected for use in constructing reporter cell lines (Figure3.2. C).

3.2.3 Generation and validation of *NOXA-mCherry* reporter cell lines

Human ALL and MM cell lines RS4;11 and KMS-12-PE were chosen for generating NOXA reporter cell lines. Both lines express relatively high level of NOXA protein as previously tested, which make the selection of the transfected cells feasible. RS4;11 cells express wildtype p53, while KMS-12-PE cells express mutant p53. Different pathways might be dominating *NOXA* expression in different cell lines. Additionally, a previous whole-exon sequencing result in our lab has shown that KMS-12-PE line only has one allele of *NOXA.* (Figure 3.3.)

The first step in generating *NOXA* reporter cell lines is fluorescence-activated cell sorting (FACS) of cell clones that stably express the mCherry reporter. The paired gRNA-Cas9D10A plasmids and donor plasmids were introduced into RS4;11 and KMS-12-PE cells by transient transfection. The transfected cells were then cultured for 7 days and the mCherry positive cells were sorted into pools. The prolonged culture time before sorting is done as there might be some cryptic promoter regions in the donor plasmid (for bacterial expression) that can cause background expression without genome integration in the transfected cells. 7 days allows sufficient time for integration and stable expression of the mCherry reporter.

The proportion of mCherry positive cells generated in my experiments was low in the initial sorts (0.9% to 0.7% in RS4;11, 0.1% to 0.2% in KMS-12-PE). These mCherry positive cells were first sorted into pools and cultured for a further 7-10 days to assess whether the reporter expression would remain stable. A large proportion of the cells reverted to be mCherry negative, but some continued to express mCherry and presumably had stably integrated the reporter into their genome (36.1% to 47.9% in RS4;11, 3.6% to 4.2% in KMS-12-PE). Cells that remained mCherry positive were then single cloned into 96-well plates by FACS. (Figure3.4.)

The second step was to identify correctly targeted clones using junction PCR (Figure3.4.A). This method employs primers located within the *mCherry* marker sequence and outside of the homology arms of *NOXA* locus (Primer 1) allowing correct insertion of the *mCherry* marker to be discerned. To determine whether both alleles of *NOXA* were tagged with the *mCherry* marker, primers bind outside of right and left homology arms were employed (Primer 3). All recovered single cell clones were tested by primer 1 and primer 3 and out of about 150 single cell clones, 12 KMS-12-PE clones and 18 RS4;11 clones showed correct insertion of the *mCherry* marker. To further confirm the insertion, the selected clones were tested using 3 additional primer pairs: primer 2 that binds within the *mCherry* sequence and outside of the right homology arms; primer 4 that bind within two homology arms and primer 5 that binds within the *mCherry* sequence. Primer pairs 3 and 4 will generate one PCR product in the case of homozygous reporter integration, and two products in the case of heterozygous integration. Only homozygous clones were taken forward to the next steps. (Figure3.5. A)

Next, the PCR-validated clones were sequenced. PCR was performed with primers binding to the targeted region (primer 1 and primer 2) followed by Sanger sequencing to detect potential indels within the insertion sites of the *mCherry* marker at the *NOXA* locus. The amplicons produced by primer 1 and 2 were sequenced from both directions and aligned to sequences of parental cells and the donor plasmid. Five different RS4;11 cell clones and five KMS-12-PE cell clones (including 4 clones homozygously tagged by *mCherry*, and 1 clone untagged) were aligned to the wildtype sequences and *mCherry* donor plasmid. The untagged clones of both cell lines did not contain *mCherry* sequence as expected, providing a negative control for the experiment. In all the sequenced clones, *mCherry* marker was

inserted into *NOXA* locus without indels or mutations at the integration sites. (Figure3.5. B) These homozygously tagged clones were selected for use in subsequent experiments.

Lastly, the transcriptional responses of the *NOXA* reporter cells to *NOXA* induction were tested. Nutlin3a, an MDM2 inhibitor was used to treat the RS4;11 *NOXA* reporter cells. MDM2 inhibitors increase cellular p53 concentration by inhibiting its sequestration by MDM2, thus causing the transcriptional activation of p53 downstream targets, including *NOXA*. The transcriptional activation of *NOXA* by MDM2 inhibitors was confirmed by qPCR in RS4;11 cells (Figure3.5. C). RS4;11 *NOXA* reporter cells were treated with 20uM of nutlin-3a in combination with 20uM of QVD-OPH for 24h. High concentrations of nutlin-3a is toxic to cells, so the pan-caspase inhibitor QVD-OPH was used to inhibit caspase activation and apoptosis. By FACS analysis, elevated mCherry fluorescence was observed in the reporter clones tested. (Figure3.6. A) However, MDM2 inhibitors only work in *TP53* wild-type cells. KMS-12-PE is a *TP53* mutant cell line, so panobinostat (a pan-HDAC inhibitor that is reported to increase NOXA transcription activity[113]) was used instead to treat KMS-12-PE *NOXA* reporter cells. The mechanism by which the HDAC inhibitor panobinostat increases *NOXA* expression remains unknown, but the effects of panobinostat on *NOXA* expression in KMS-12-PE cells were also confirmed by qPCR (Figure3.6. C). Again, elevated mCherry fluorescence was seen by FACS analysis following 24h of 50nM panobinostat treatment. No cell death inhibitors were added as pilot tests confirmed that panobinostat is not lethal at low concentration in KMS-12-PE cells (Figure3.6. B). In conclusion, the *NOXA* reporter cells exhibited changes in reporter expression that faithfully matched changes in the expression of endogenous *NOXA*.

3.3 Conclusion

In this chapter, I generated and validated *NOXA* reporter cell lines using a CRISPR-Cas9 knock-in approach. Using two pairs of sgRNAs, ~5% of RS4;11 cells and less than 0.1% of KMS-12-PE cells had the *mCherry* marker successfully inserted into genome. Using PCR to evaluate the position of reporter integration in the genome, 18 RS4;11 clones and 12 KMS-12-PE clones were identified with correct insertion. 4 clones of each line were further tested by sequencing and FACS to confirm the DNA sequence fidelity and the reporter functionality. Those clones that completed the validation pipeline: RS4;11 *NOXA* reporter clones 11, 44, 59 and 95 and KMS-12-PE *NOXA* reporter clones 31, 33, 127 and 132 will be

used in further experiments including CRISPR genetic screens and CRISPR tiling screens in the next two chapters of this thesis.

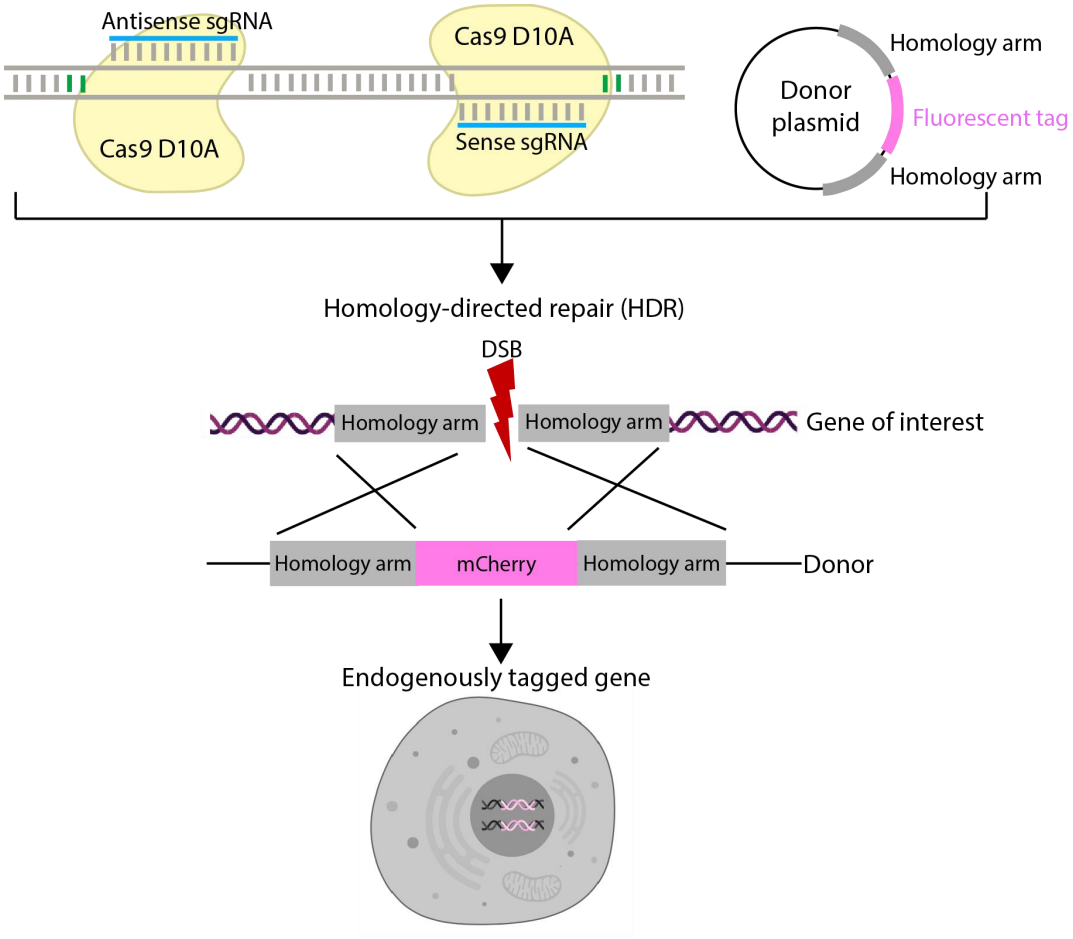

Antisense sgRNA
Cas9 D10A
Cas9 D10A
Sense sgRNA
Donor plasmid
Homology arm
Fluorescent tag
Homology arm
Homology-directed repair (HDR)
DSB
Homology arm
Homology arm
Gene of interest
Homology arm
mCherry
Homology arm
Donor
Endogenously tagged gene

Figure 3.1. The paired CRISPR-Cas9 nickase approach.

Cas9D10A has a mutation in its nuclease domain, which causes it to generate single strand rather than double strand DNA breaks. In order to generate DSBs with this variant of Cas9, simultaneous targeting of both sense and antisense DNA strands are required. The paired Cas9D10A/sgRNA expressing plasmids are transfected together with a donor plasmid containing DNA repair template. The Cas9D10A cutting leads to a DSB at the targeted site, and the subsequent DNA repair occurs through HDR, which leads to the insertion of the donor plasmid sequence (in this case, *mCherry* marker sequence) into the gene of interest.

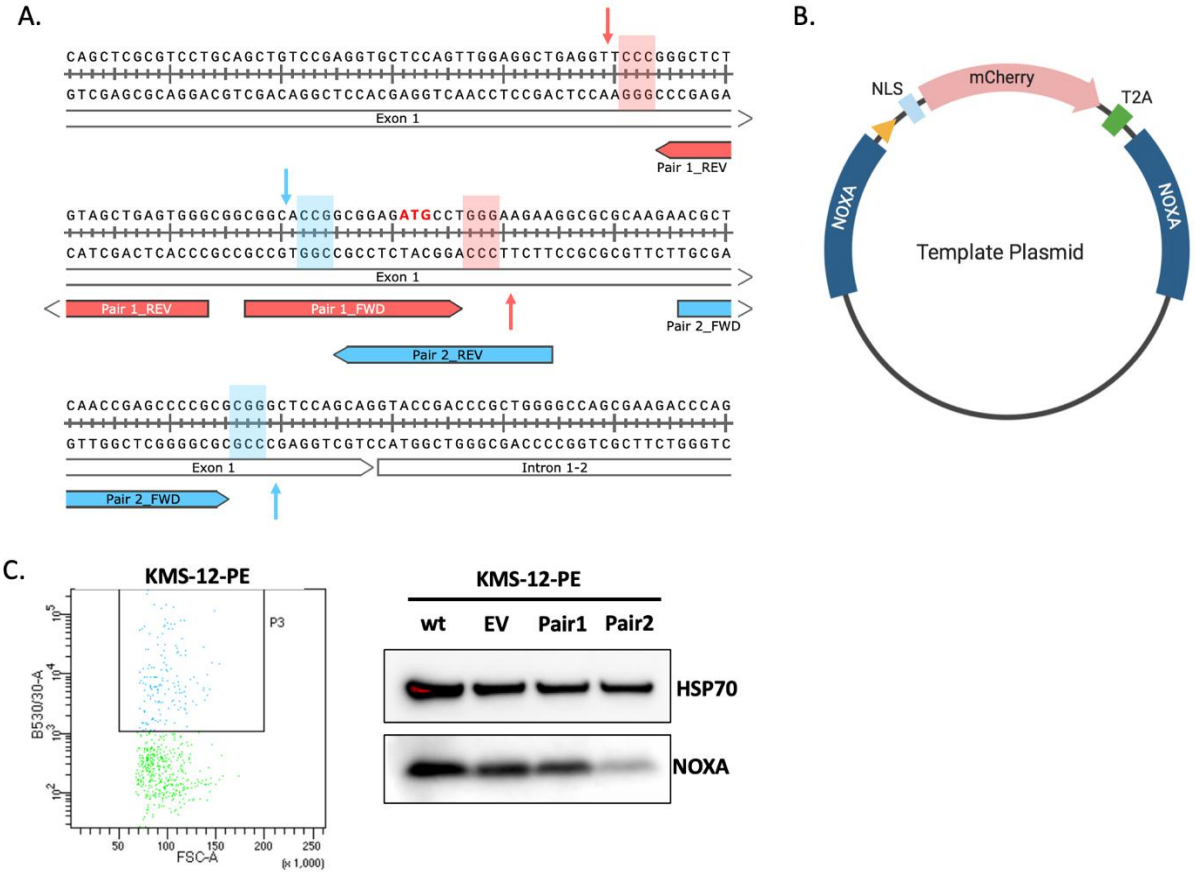

A.
CAGCTCGCGTCCTGCAGCTGTCCGAGGTGCTCCAGTTGGAGGCTGAGGTTCCCGGGCTCT
GTCGAGCGCAGGACGTCGACAGGCTCCACGAGGTCAACCTCCGACTCCAAGGGCCCGAGA
Exon 1
Pair 1_REV
GTAGCTGAGTGGGCGGCGGCACCGGCGGAGATGCCTGGGAAGAAGGCGCGCAAGAACGCT
CATCGACTCACCCGCCGCCGTGGCCGCCTCTACGGACCCTTCTTCCGCGCGTTCTTGCGA
Exon 1
Pair 1_REV
Pair 1_FWD
Pair 2_FWD
Pair 2_REV
CAACCGAGCCCCGCGCGGGCTCCAGCAGGTACCGACCCGCTGGGGCCAGCGAAGACCCAG
GTTGGCTCGGGGCGCGCCCGAGGTCGTCCATGGCTGGGCGACCCCGGTCGCTTCTGGGTC
Exon 1
Intron 1-2
Pair 2_FWD
B.
NLS
mCherry
T2A
NOXA
NOXA
Template Plasmid
C.
KMS-12-PE
P3
B530/30-A
FSC-A
(x 1,000)
KMS-12-PE
wt
EV
Pair1
Pair2
HSP70
NOXA

Figure 3.2. The design of sgRNAs and donor plasmids.

(A) Two pairs of sgRNAs were designed to target the start codon of *NOXA*. Each pair has one sgRNA that covers the start codon for higher homologous recombination efficiency. The shaded red and blue boxes indicate the PAM sites, and the arrows indicate the cut sites of each pair of sgRNAs.

(B) The donor plasmid contains the gene encoding *mCherry* fluorescent protein flanked by two 800-bp homology arms of *NOXA*. The *mCherry* sequence replaces the start codon (indicated as orange triangle) and is in-frame with the endogenous promoter and exons of *NOXA*. A NLS is placed between at the 3' end of mCherry and a T2A linker is placed between *mCherry* and *NOXA*. The donor was synthesised and cloned into pUC57_mini, with no mammalian selection marker contained.

(C) Functionality test of the designed gRNAs were done prior to transfection with donor plasmid. Each pair of sgRNAs were co-transfected with a GFP expression vector. GFP positive cells were sorted (left) and analysed by western blot (right). NOXA protein level was decreased by the indels caused by the sgRNAs.

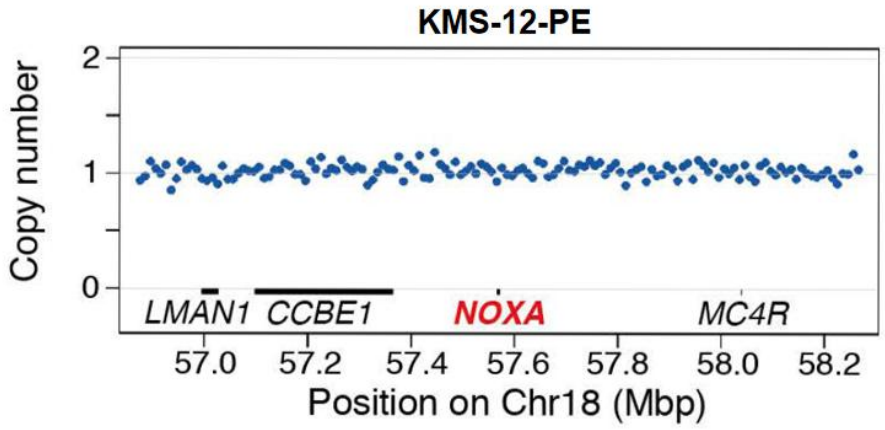

KMS-12-PE
Copy number
2
1
0
LMAN1
CCBE1
NOXA
MC4R
57.0
57.2
57.4
57.6
57.8
58.0
58.2
Position on Chr18 (Mbp)

Figure 3.3. *NOXA* copy number in KMS-12-PE cell line.

NOXA locates on 57.6Mbp on chromosome 18. Whole exome sequencing showed that *NOXA* and its adjacent genes (*LMAN1, CCBE1, MC4R*) only have one copy in the genome pf KMS-12-PE cell line.

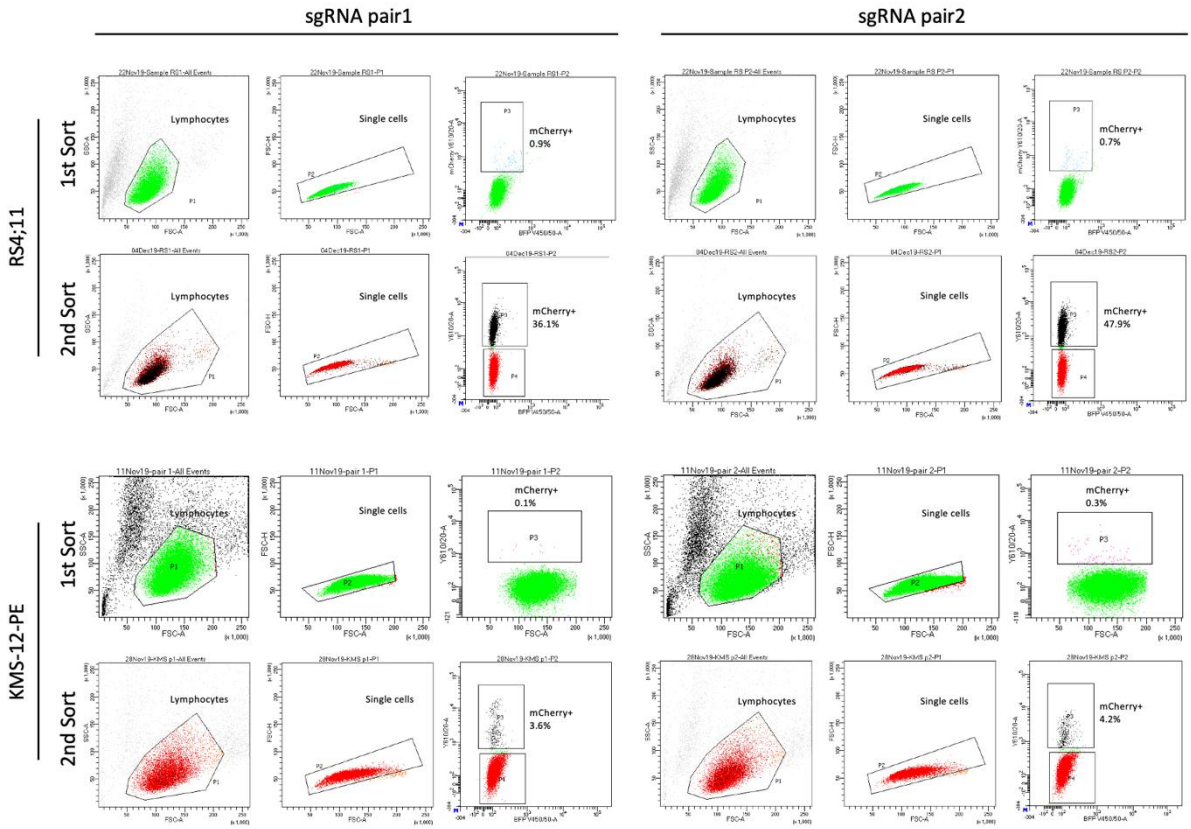
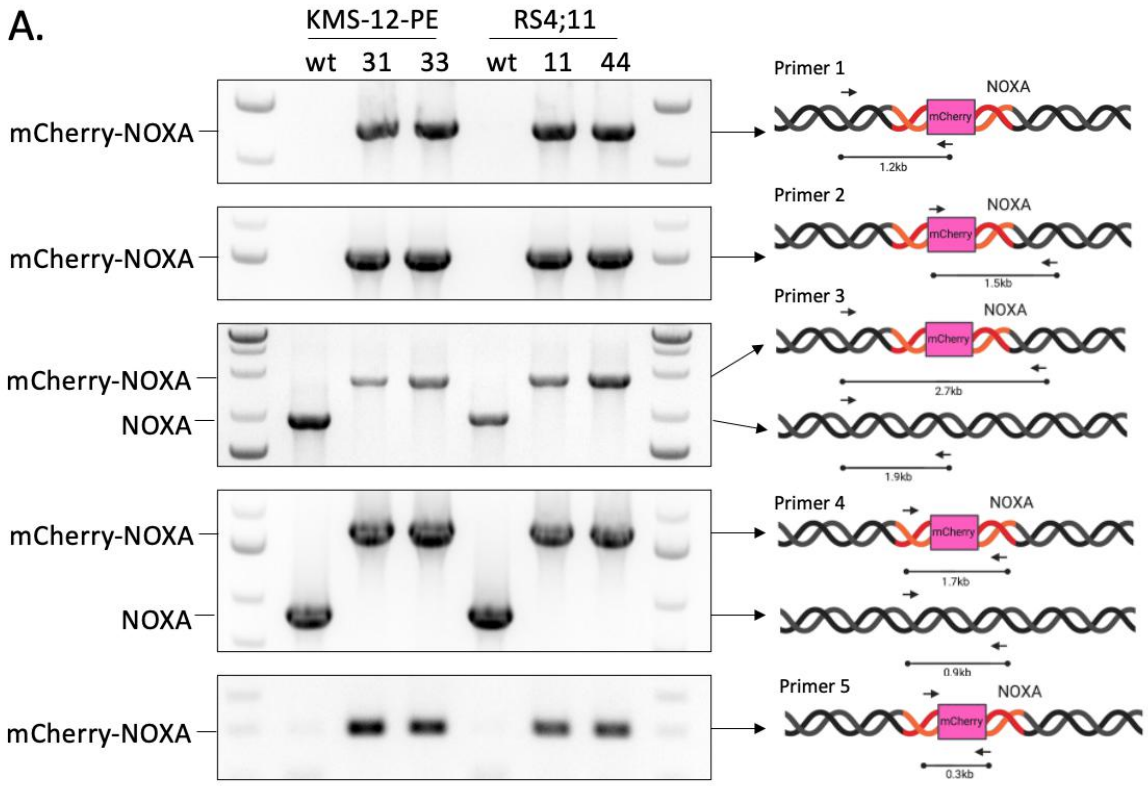

sgRNA pair1
sgRNA pair2
RS4;11
1st Sort
2nd Sort
KMS-12-PE
1st Sort
2nd Sort
Lymphocytes
Single cells
mCherry+
0.9%
mCherry+
0.7%
mCherry+
36.1%
mCherry+
47.9%
mCherry+
0.1%
mCherry+
0.3%
mCherry+
3.6%
mCherry+
4.2%

Figure 3.4. FACS of the transfected cell lines.

RS4;11 and KMS-12-PE cells were transfected with two gRNA-Cas9D10A plasmids and donor plasmid, RS4;11 and KMS-12-PE cells transfected with Cas9D10A empty vector and donor plasmid were used as negative controls to define the background fluorescence. For each sort, cells were gated by their scatter properties for live, single cells, and then by the signal intensity of mCherry. Cells with mCherry level higher than negative controls were selected by two rounds of sorting (indicated as the P3 boxes). The numbers next to the mCherry positive population indicate the percentage of mCherry positive cells. In the 2nd sort, mCherry positive cells were seeded into 96-well plates as single cell clones.



B.

```
ACTGTTCGTGTTCAGCTCGCGTCCTGCAGCTGTCCGAGGTGCTCCAGTTGGAGGCTGAGGTTCACGGGAA
         10        20        30        40        50        60        70
NOXA          ACTGTTCGTGTTCAGCTCGCGTCCTGCAGCTGTCCGAGGTGCTCCAGTTGGAGGCTGAGGTTCACGGGAA 70
clone11       ACTGTTCGTGTTCAGCTCGCGTCCTGCAGCTGTCCGAGGTGCTCCAGTTGGAGGCTGAGGTTCACGGGAA 70
NOXA-mCherry  ACTGTTCGTGTTCAGCTCGCGTCCTGCAGCTGTCCGAGGTGCTCCAGTTGGAGGCTGAGGTTCACGGGAA 70

GCGTCGACGAATGGGCGGCGGCACCGGCGGAGATGACTGCTCCAAAGAAGAAGCGTAAGGTAATGGTGAG
         80        90       100       110       120       130       140
NOXA          GCGTCGACGAATGGGCGGCGGCACCGGCGGAGATG                           140
clone11       GCGTCGACGAATGGGCGGCGGCACCGGCGGAGATGACTGCTCCAAAGAAGAAGCGTAAGGTAATGGTGAG 140
NOXA-mCherry  GCGTCGACGAATGGGCGGCGGCACCGGCGGAGATGACTGCTCCAAAGAAGAAGCGTAAGGTAATGGTGAG 140

CAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCC
         150       160       170       180       190       200       210
NOXA                                                                      210
clone11       CAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCC 210
NOXA-mCherry  CAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCC 210
```

Figure 3.5. Junction PCR and Sanger sequencing of the mCherry positive cells.

(A) PCR was used to test for insertion of *mCherry* into the correct position in the *NOXA* locus. RS4;11 and KMS-12-PE parental cells served as negative controls. Results from two clones (clones 31 and 33 for KMS-12-PE and clones 11 and 44 for RS4;11) were shown. The labels on the left indicate the expected size of bands amplified from a wildtype *NOXA* allele or an allele with correct integration of the *mCherry* marker. The schemes on the right illustrate where the primers bind (black arrows) and the predicted product length. The left and right homology arms are marked with red color.

(B) PCR was carried out with primers spanning the homology arm junctions (Primer 1 and Primer 2), followed by Sanger sequencing. The amplicon sequences were aligned to sequences of parental cells (*NOXA*) and the donor plasmid (*NOXA-mCherry*). The shaded boxes indicate *mCherry* sequence, which was absent from parental RS4;11 and KMS-12-PE cells as expected.

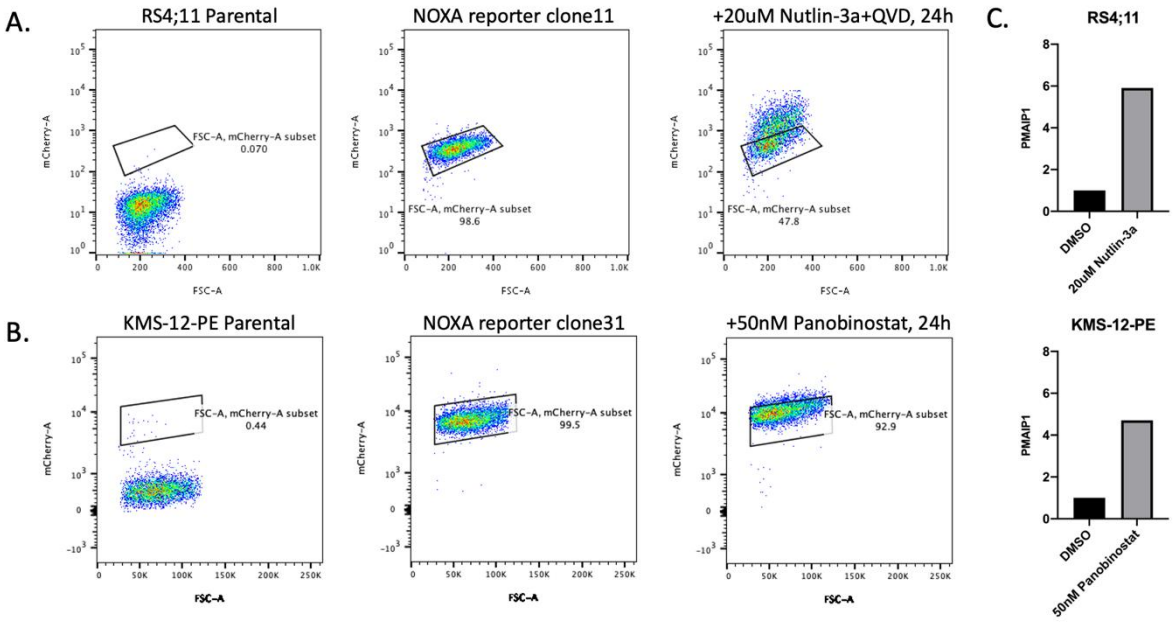

A.
RS4;11 Parental
NOXA reporter clone11
+20uM Nutlin-3a+QVD, 24h
FSC-A, mCherry-A subset
0.070
FSC-A, mCherry-A subset
98.6
FSC-A, mCherry-A subset
47.8
mCherry-A
FSC-A
B.
KMS-12-PE Parental
NOXA reporter clone31
+50nM Panobinostat, 24h
FSC-A, mCherry-A subset
0.44
FSC-A, mCherry-A subset
99.5
FSC-A, mCherry-A subset
92.9
C.
RS4;11
PMAIP1
DMSO
20uM Nutlin-3a
KMS-12-PE
50nM Panobinostat

Figure 3.6. The transcriptional responses of the *NOXA* reporter cells to *NOXA* induction

(A) RS4;11 *NOXA* reporter clone 11 was treated with 20uM nutlin-3a in combination with 20uM QVD for 24h. The mCherry reporter expression was then measured by FACS. Parental cells (left) and untreated cells (middle) were analysed as negative controls. An elevation of cell mCherry level were seen upon treatment.

(B) KMS-12-PE *NOXA* reporter clone 32 was treated with 50nM panobinostat for 24h. Parental cells (left) and untreated cells (middle) were analysed as negative controls. Elevated mCherry reporter expression was observed following treatment with panobinostat.

(C) *NOXA (PMAIP1)* is transcriptionally induced following nutlin-3a and panobinostat treatment. RS4;11 cells were treated with DMSO or 20uM nutlin-3a in combination with 20uM QVD for 24h. KMS-12-PE cells were treated with DMSO or 50nM panobinostat for 24h.

Chapter 4: CRISPR-Cas9 loss-of-function screens to identify epigenetic modulators of *NOXA* expression

4.1 Introduction

High-throughput loss-of-function genetic screens have been widely applied in model systems to link genes to specific phenotypes. For example, RNA-interference (RNAi) screening using short hairpin RNA (shRNA) libraries to downregulate target genes, is a well-established method for forward genetic screening[125]. Recent breakthroughs in genome engineering, especially the CRISPR-Cas9 system, have hugely accelerated the development of new methods to perform genetic loss-of-function studies[116]. Because of its outstanding programmability and efficacy, the CRISPR-Cas9 system is also ideal for performing powerful phenotypic screens. This is in contrast to previous shRNA library screening which may only partially interfere with the targeted mRNAs resulting in varying degrees of functional inhibition of the encoded protein and, potentially, a less severe phenotype compared to that of the knockout phenotype[126]. Several whole-genome CRISPR-Cas9 sgRNA libraries have been constructed for this purpose and have proven to be powerful tools for genome-wide screens *in vitro* and *in vivo*[127-129]. These include the YUSA and Gecko libraries, targeting up to 19,150 mouse or human protein coding genes[129,130], and the more recent Brunello library that was derived from previous libraries with higher on-target specificity and minimum off-target effects[131]. In addition to whole genome sgRNA libraries that allow researchers to unbiasedly associate genes with phenotypes of interests, smaller sgRNA libraries have also been used in studies with more focused interests. These include kinase libraries[132], transcription factor libraries[133] and microRNA libraries[134]. One major advantage of focused libraries is their high coverage of targeted genes, which leads to more robust results and higher feasibility of drop- out screens.

Genetic screens allow researchers to tackle questions relating to a wide variety of phenotypes, including cell proliferation[127], cell migration[135] and cancer cell drug resistance[136]. In addition, the use of genetic screens in cells with fluorescently labelled gene

loci has greatly expanded the application to the study of intracellular processes, including but not limited to DNA repair[137], protein stability[73] and gene expression.

In this chapter, I used three different sgRNA libraries for screens in the *NOXA-mCherry* reporter cell lines that I generated in chapter 3, in order to identify genes involved in the transcriptional regulation of *NOXA* expression. It is hypothesized that modulating *NOXA* expression will regulate the function and the stability of the critical BCL2 family member – MCL1 and will translate into a novel therapeutic approach to treatment of haematologic malignancy.

4.2 Result

4.2.1 sgRNA library creation

To elucidate presently unknown mechanisms of *NOXA* regulation, I utilised three focused sgRNA libraries to perform CRISPR-Cas9 screens. These three libraries were designed respectively to target: 1) transcription factors (TFs library); 2) epigenetic modulators (Epifactors library) and 3) kinases (Kinases library).

The transcriptional regulation of gene expression is a process of high complexity that can operate at the level of chromatin modification, DNA methylation and through the binding of transcription factors. As these processes are generally controlled by transcription factors and epigenetic modifying enzymes, the TFs and Epifactors libraries were chosen for the initial screens. Kinases, on the other hand, play a major role in signal transduction and protein regulation by governing the phosphorylation of transcription factors and epigenetic modulators. In addition, studies have shown that kinases inhibitors can be potent cancer therapeutics, providing a rationale to screen kinases as well. Therefore, the three libraries had the potential to identify direct regulators of *NOXA* expression as well as indirect regulators of any of the known and unknown components of this gene regulatory pathway.

The TFs library was a kind gift from Christopher Vakoc[133]. This library targets the DNA binding domains of 1,427 transcription factor genes and contains a total of 8658 sgRNAs. The kinases library was created by Beinan Wang and targets 3293 kinases and several other drug targets including all of the BCL2 family genes. Altogether, the kinases library is comprised of 14,213 sgRNAs. The Epifactors library was a newly generated library targeting 720 known epigenetic modulator genes[138,139] with 6,047 sgRNAs, including 150 non-

targeting control sgRNAs. Each gene is targeted by 6-8 independent sgRNAs, which were collated from two whole-genome CRISPR libraries: Brunello[131] and Gecko[7]. (Table 4.1)

4.2.2 Optimisation of the CRISPR screening system

To perform the screens, I followed a method previously used in the lab in a genome-wide genetic screen involving selection by the expression of a fluorescent reporter[73]. Specifically, two RS4;11 and KMS-12-PE *NOXA-mCherry* reporter clones were engineered and transduced with one of the three lentiviral sgRNA libraries. The transduced cells were then enriched through FACS according to their mCherry reporter expression level. The design and optimisation of the screening system are discussed below in more detail, as it was critical to the success of the experiment (Figure 4.1):

1) Cell line engineering:

RS4;11 *NOXA* reporter clone 59 and KMS-12-PE *NOXA* reporter clone 132 were validated *NOXA* reporter cell lines as described in chapter 3. To avoid the possible lethality of sgRNA that cause *NOXA* upregulation, *BAX/BAK* double knock-out (B/B DKO) cells were derived from the two *NOXA* reporter clones. BAX and BAK are the major effector proteins of intrinsic apoptosis pathway and so deletion of BAX and BAK prevents cell death caused by high *NOXA* expression. Briefly, the two clones were transiently transfected with the BAX and BAK sgRNA- Cas9 plasmid pX458 and sorted according to their GFP expression. The GFP positive cells were further selected with BH3 mimetics to enrich B/B DKO cells. The selected cells were then subjected to Western blot to confirm deletion of BAX and BAK. (Figure 4.2) The B/B DKO *NOXA* reporter cells were then engineered to constitutively express the Cas9 endonuclease from a lentiviral vector FU-Cas9-GFP.

2) Group design:

For each screen, there were three groups of samples: 1) the screening group; 2) the Cas9 negative group and 3) the baseline control group. Each group contained 2-6 independent samples. The screening group went through several rounds of sorting to enrich cells with the desired phenotype; The Cas9 negative group served as a secondary baseline to assess non- specific fluorescence effects; The baseline control groups record the baseline coverage of the library.

3) Virus titration:

The efficiency of lentiviral infection was assessed prior to performing the screens. Infection efficiency needs to be high enough to ensure full coverage of the targeted genes while minimising the multiplicity of infection so as to reduce the frequency of passenger sgRNAs. For instance, very high rates of infection increase the number of sgRNA per cell, which increases the possibility that non-specific sgRNAs get carried along by phenotype-inducing sgRNAs. Therefore, a high multiplicity of infection can make it more difficult to discern sgRNAs that produce desired phenotypes. Based on our previous experience, it is optimal to keep the overall infection rate around 30% to 40%. Virus titration tests were done with each library preparation in the two cell lines RS4;11 and KMS-12-PE. In brief, 1 million cells were infected with the virus and analysed by FACS 48hr after infection. The infection efficiency was calculated by the percentage of BFP positive cells in the transduced pools. Preliminary experiments demonstrated that the infection efficiency varied from 10% to 60% with different virus concentration across two cell lines. To obtain a 30% infection efficiency, the final virus volume used in the screens varied from 100uL to 2mL per million cells depending on the cell line and viral preparation.

4) sgRNA coverage:

The cell number in each independent sample was determined by the size of the library and the desired coverage of each sgRNA. The coverage can be interpreted as the number of cells expressing each individual sgRNA. In smaller libraries, a higher sgRNA coverage is feasible. Taking the kinase library with 14,213 sgRNAs as an example, in each independent sample, 4.7 million cells were transduced with the sgRNA library with 30% infection rate to reach 100-fold coverage of each guide (this is calculated by 14213/30% x 100 = 4.7 x 10^6 cells).

5) FACS:

Following infection with the sgRNA library, each sample was expanded without splitting to ensure minimum loss of representation. 48 hours after infection, the cells expressing sgRNA were selected for by adding puromycin to the cultures. The puromycin concentrations were determined by cell death assay in parental cell lines. Cells were expanded through transition into larger flasks and the first sort was performed approximately 7d after infection. During the first sort, cells expressing both higher and lower levels of mCherry were seen in the screening groups compared to the baseline control groups. However, the observed changes in mCherry levels were relatively subtle, with only a few cells acquiring significantly higher or lower mCherry levels. This was somewhat expected as the endogenous expression

changes may not necessarily be dramatic. For example, a 2-fold change in expression does not really allow for great separation on a log-scale FACS plot as the variance in the population itself is already larger than this in the untreated populations. Additionally, the percentage of sgRNA in the library expected to produce a phenotype is also low, which means only a small percentage of cells were expected to exhibit a phenotype. To conclude, the observation indicated that multiple rounds of sorting were needed to enrich the desired cells from the rest of the population.

Another problem was that changes of mCherry levels (especially increases) were also observed in the Cas9 negative groups, indicating the Cas9 negative groups were necessary to rule out non-specific effects.

In summary, both screening groups and Cas9 negative groups were sorted into mCherry-high and mCherry-low populations, based on their mCherry level compared to the parental cells. The sorted cells were expanded through transitions into larger format tissue culture flasks and were sorted again once the required cell number was attained. The above process was repeated for 3-4 rounds for the screening groups until an obvious separation in mCherry level between the screening groups and the parental cells were seen. For the Cas9 negative groups, only 1-2 rounds of sorting were repeated.

6) Data collection and analysis

DNA samples were collected at multiple stages throughout the experiment. DNA from the baseline control groups was collected at the time of first sort to record the initial representation of guides in the library. DNA of the screening groups and the Cas9 negative groups was collected at each timepoint of sorting. DNA amplification and indexing were next carried out in a single PCR reaction. The amount of PCR template was calculated to obtain full coverage of each sorted cell sample. For example, if 500,000 cells were sorted in a sample, the DNA template amount for that sample should be 1.8ug (calculated by 500,000 x 3.6pg = 1.8ug, 3.6pg is the average DNA content of a human cell)[140]. The amplified and indexed DNA samples were pooled together and sequenced using the Miseq platform. Sequencing reads could then identify sgRNA representation across the samples.

To analyse the re-occurring sgRNAs across replicate samples and rank targeted genes associated with *NOXA* expression, the algorithm described by Eden et al.[117] was used. This algorithm ranks targeted genes according to their frequency and magnitude of selection in the screening samples as well as by the co-selection of different sgRNAs targeting the same

gene. The top hits for the mCherry high screening samples are potential negative regulators of *NOXA*, whilst the top hits for the mCherry low screening samples are potential positive regulators.

4.2.3 CRISPR genetic screens with the Epifactors library identified novel epigenetic modulators of *NOXA* expression

In the screens with the Epifactors library, RS4;11 and KMS-12-PE *NOXA* reporter B/B DKO cells lines with or without constitutive Cas9 expression were transduced with the library. There were 6 independent samples in the baseline control groups, 2 independent samples in the Cas9 negative groups and 4 independent samples in the screening groups. For each sample, the starting cell number was 3 million. The baseline control groups were harvested at the timepoint of the first sort. The Cas9 negative and screening groups were sorted into mCherry high and mCherry low populations. The Cas9 negative groups were harvested after the first sort, while the screening groups were harvested after 2-4 sorts. Each sample was monitored by FACS at regular time points. (Figure 4.3)

The top hits for the mCherry-high and mCherry-low Epifactors library screens are summarised in Table 4.2. In the RS4;11 screen involving selection for low mCherry expression, only one hit, *CTBP1*, reached significant level (False discovery rate (FDR) ≤ 0.05). In the KMS-12-PE screen for low mCherry expression, two hits, *CHTOP* and *ZMYM3* reached significant FDR. Interestingly, *CHTOP* was also among the most enriched genes in the RS4;11 mCherry low screen, although it was just below the significant FDR threshold in this cell line. These epigenetic modulator genes, enriched by selecting for reduced mCherry levels, are potentially positive regulators of *NOXA* expression. In the screens for high mCherry expression, several genes were enriched in KMS-12-PE cells, while none were evident in RS4;11 cells (FDR ≤ 0.05). The most strongly enriched genes in the KMS-12-PE mCherry-high screen included *SPEN*, *HSPA1A*, *KEAP1*, *FOXA1*, *HDAC3* and *SAP30*. These epigenetic modulator genes are potentially negative regulators of *NOXA* expression (Figure 4.4). In summary, CRISPR genetic screens with the Epifactors library have identified, for the first time, several epigenetic modulators that positively or negatively regulate *NOXA* expression, none of which have been reported to have direct effect on *NOXA* expression before. The known functions of some of the identified genes are summarised here:

CHTOP encodes a small nuclear protein named as Chromatin target of PRMT1. Recent study has reported that this protein binds 5-hydroxymethylcytosine (5hmC) and associates with an arginine methyltransferase complex (methylosome), which promotes methylation of arginine 3 of histone H4 (H4R3) and estrogen-dependent gene activation[141]. The fact that *CHTOP* came up in two independent cell lines is intriguing and strengthens the evidence that it may play a role in the transcriptional regulation of *NOXA*.

SPEN encodes a hormone inducible transcriptional factor. The transcriptional repression effect of SPEN can occur through recruitment of proteins involved in histone deacetylation[142]. SPEN has been reported as a crucial factor in X chromosome inactivation[143] and as a tumor suppressor in hormone-dependent breast cancers[144]. The link between SPEN functions and apoptosis is still unclear and requires further investigation.

KEAP1 is an oxidative stress sensor gene. The coding protein KEAP1 functions as an adaptor for Cullin3 (CUL3)-based E3 ligase complex to regulate proteasomal degradation of NRF2[145]. The KEAP1-NRF2 signalling pathway plays a significant role in protecting cells from exogenous and endogenous stress[146]. This pathway has been reported to regulate oxidative stress induced apoptosis[147]. Interestingly, an increase of both NRF2 and NOXA protein have been observed following treatment with the proteasome inhibitor bortezomib in MCL[114,148], but the direct link between NRF2 and NOXA is yet to be investigated.

HDAC3 belongs to the histone deacetylase superfamily. Histone deacetylases consist of a family of 11 members that are divided into 4 classes. They are regulatory enzymes that catalyse the removal of acetyl groups from histones, a process that is often associated with decreased transcription. Class I HDACs, including HDAC3, have an important role in tumorigenesis and may be candidate targets for many cancer treatments[149]. Several mechanistic studies of HDAC inhibitors indicate that they can induce apoptosis in cancer cells[150]. A recent study reported that the pan-HDAC inhibitor panobinostat induced *NOXA* expression independently of *TP53* status[113]. The screen result may provide a rationale for HDAC inhibitors in cancer therapies. However, the molecular mechanism of HDACs, especially HDAC3 on *NOXA* expression is still to be elucidated.

To conclude, several hypotheses are proposed regarding the potential functions of the identified epigenetic modulators of *NOXA*. The next stage of this study is to knockout the genes of interest using individual sgRNAs to validate these hits. Once confirmed our

longer-term research goal is to investigate the underlying mechanisms by which these genes regulate *NOXA* expression and the therapeutic potential of controlling this action.

4.2.4 CRISPR genetic screens with the other libraries identified control hit but need further optimisation

The methodology and design of screens performed with the Kinases library and the TFs library are the same as the screens with the Epifactors library. However, *NOXA* itself was the only gene that was significantly enriched in these experiments, in both RS4;11 and KMS-12-PE mCherry-low screens with kinases library. No genes were significantly enriched in the mCherry-high screens. One possible explanation for why *NOXA* itself was enriched in the mCherry-low screens is that the sgRNAs targeting *NOXA* coding regions might influence the *mCherry-NOXA* mRNA stability in the *NOXA* reporter cells.

The results suggested that the protocols of CRISPR genetic screens in *NOXA* reporter cells still need further optimisation. Deletion of certain kinases or TFs might have lethal effect in cells. It is possible that better result could be achieved with short-term screens to escape the survival disadvantages caused by deletion of certain genes. Balancing the enrichment of phenotypes with the length of time cells are cultured during the screen is a technical challenge that needs to be carefully optimised in the future.

4.3 Conclusion

CRISPR genetic screens combined with endogenous gene reporter assays have proven to be a powerful tool for studying the regulatory pathways controlling gene expression. In this chapter I used separate libraries of sgRNA targeting transcription factors, kinases and epigenetic regulators to screen *NOXA* reporter cell lines. Whilst I did not uncover any novel kinases regulators or transcription factors of *NOXA*, I did identify several epigenetic regulators that may either positively or negatively regulate *NOXA* expression. Some of these epigenetic modulators (for example, *KEAP1* and *HDAC3*) have previously been associated with *NOXA* expression but lacked direct evidence. Future studies will aim to uncover the mechanisms by which these proteins regulate *NOXA* expression with the ultimate goal of exploiting any therapeutic potential that these proteins may hold. Other epigenetic modulators of *NOXA* identified as a result of my work, have not previously been appreciated

for their potential functions in apoptosis or in *NOXA* regulation. Future studies will aim to elucidate their novel roles in regulating apoptosis and *NOXA* expression in particular.

However, the reporter cell screening system has its limitation in finding hits that have physiologic impact on cell apoptosis. The hits identified might contribute to NOXA basal expression but not inducing apoptosis. Therefore, the hits need to be carefully investigated on their roles in apoptosis and venetoclax sensitivity. Potentially, combining a venetoclax resistant screen with the current screen would give more valid read-out.

Library Name	sgRNAs	Source
Epigenetic factors	6047 guides target 720 genes.	Customed library.
Kinases	14213 guides target 3293 genes.	Customed library.
Transcription factors	8658 guides target 1427 genes.	Addgene #123334.

Table 4.1. List of CRISPR-Cas9 libraries

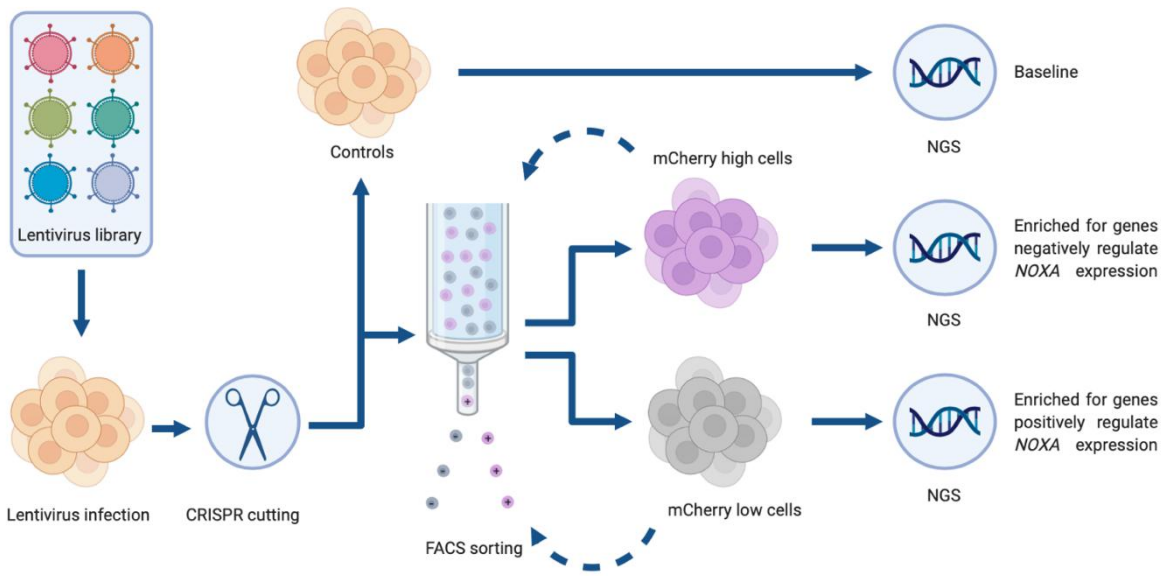

Lentivirus library
Lentivirus infection
CRISPR cutting
Controls
FACS sorting
mCherry high cells
mCherry low cells
Baseline
NGS
NGS
NGS
Enriched for genes negatively regulate NOXA expression
Enriched for genes positively regulate NOXA expression

Figure 4.1. Schematic of CRISPR-Cas9 screens

The *NOXA-mCherry* reporter cells were transduced with the lentiviral sgRNA libraries. Cells in the baseline control groups were harvested without enrichment. Cells in the screening groups with higher or lower mCherry fluorescent level were enriched through several rounds of sorting. DNA samples were collected from the enriched and control samples. Targeted genes in the sorted cells were then identified by next generation sequencing.

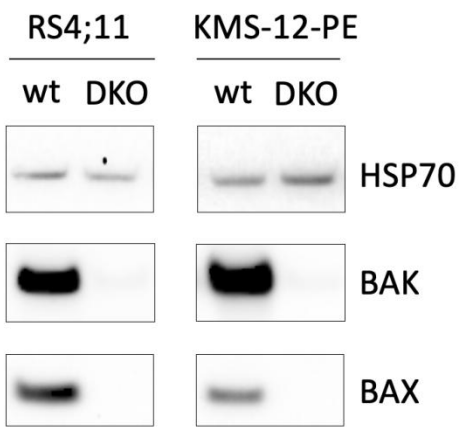

RS4;11
KMS-12-PE
wt
DKO
wt
DKO
HSP70
BAK
BAX

Figure 4.2. Validation of *BAX/BAK* DKO NOXA reporter cells

B/B DKO cells were derived from RS4;11 NOXA reporter clone 59 and KMS-12-PE NOXA reporter clone 132 using CRISPR-Cas9 knock-out approaches. The transfected cells were selected by combination of BH3 mimetics ABT-199, S63845 and A1331852 (1uM, 24h treatment, two times). Compared to parental cells (wt), the DKO cells showed no protein expression of BAX or BAK. The cells were analysed as a polyclonal population.

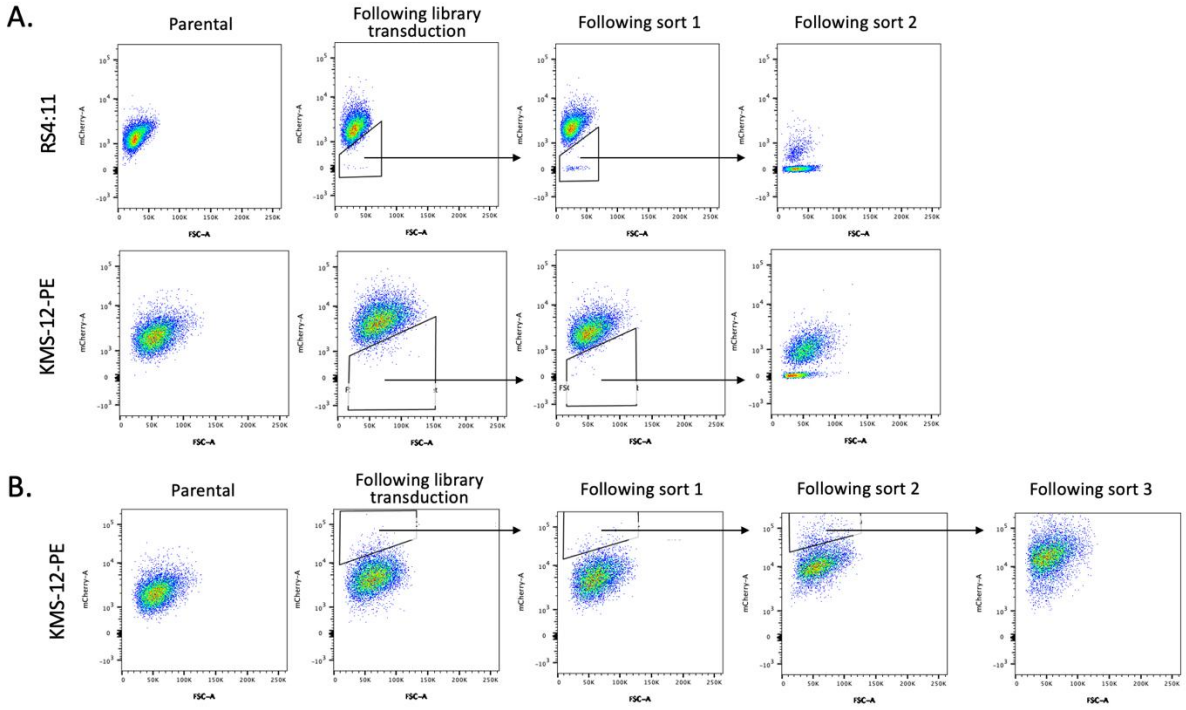

A.
Parental
Following library transduction
Following sort 1
Following sort 2
RS4:11
KMS-12-PE
mCherry-A
FSC-A
B.
Parental
Following library transduction
Following sort 1
Following sort 2
Following sort 3
KMS-12-PE
mCherry-A
FSC-A

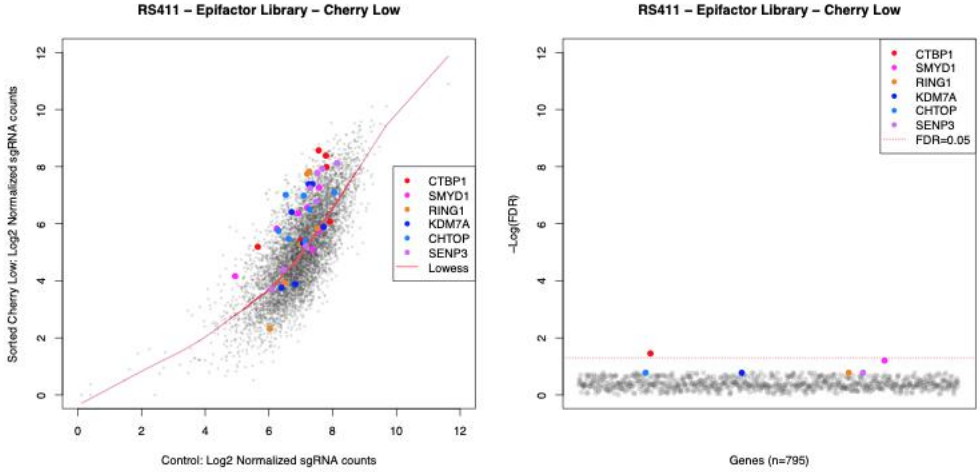
Figure 4.3. Enrichment for cell populations with altered mCherry expression by repeated rounds of cell sorting in the Epifactors library screens

The engineered RS4;11 and KMS-12-PE cells were transduced with the Epifactors library.

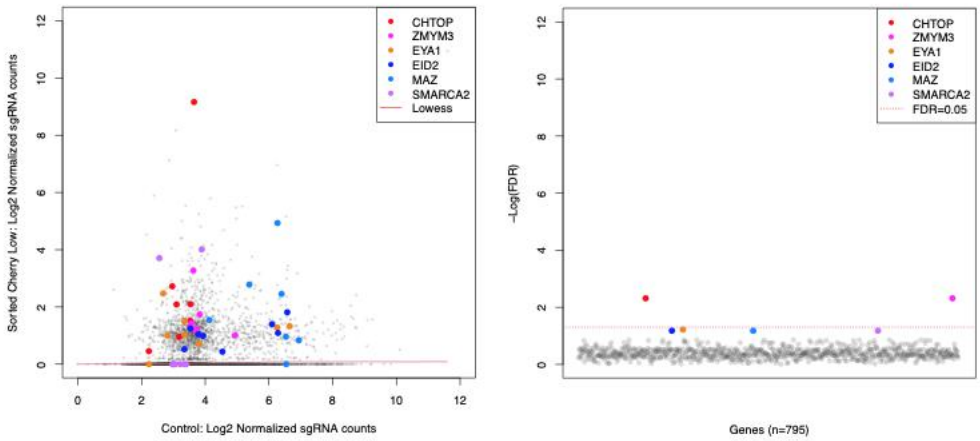
(A) Transduced RS4;11 and KMS12-PE cells expressing low levels of mCherry-low were enriched through two rounds of FACS sorting. The untransduced parental cells were shown on the left. The gates on each plot indicate which cells were sorted.

(B) The mCherry-high cells of transduced KMS-12-PE samples were enriched through three rounds of FACS sorting.

A.


RS411 – Epifactor Library – Cherry Low
Sorted Cherry Low: Log2 Normalized sgRNA counts
Control: Log2 Normalized sgRNA counts
CTBP1
SMYD1
RING1
KDM7A
CHTOP
SENP3
Lowess
RS411 – Epifactor Library – Cherry Low
−Log(FDR)
Genes (n=795)
CTBP1
SMYD1
RING1
KDM7A
CHTOP
SENP3
FDR=0.05

B.


KMS12 – Epifactor Library – Cherry Low
Sorted Cherry Low: Log2 Normalized sgRNA counts
Control: Log2 Normalized sgRNA counts
CHTOP
ZMYM3
EYA1
EID2
MAZ
SMARCA2
Lowess
KMS12 – Epifactor Library – Cherry Low
−Log(FDR)
Genes (n=795)
CHTOP
ZMYM3
EYA1
EID2
MAZ
SMARCA2
FDR=0.05

C.

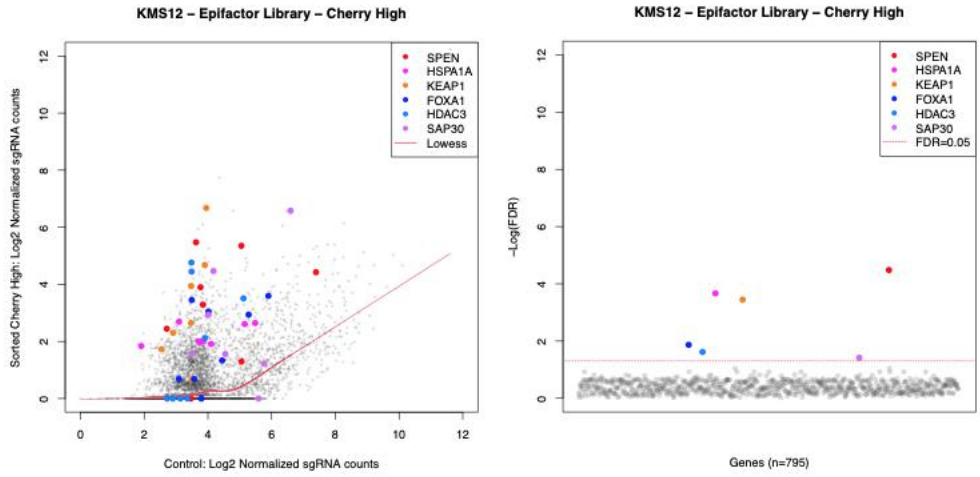

KMS12 – Epifactor Library – Cherry High
Sorted Cherry High: Log2 Normalized sgRNA counts
Control: Log2 Normalized sgRNA counts
SPEN
HSPA1A
KEAP1
FOXA1
HDAC3
SAP30
Lowess
KMS12 – Epifactor Library – Cherry High
−Log(FDR)
Genes (n=795)
SPEN
HSPA1A
KEAP1
FOXA1
HDAC3
SAP30
FDR=0.05

Figure 4.4. Hit FDR and abundance of Epifactors library screens

The FDR and sgRNA abundance highlighting the six most significantly enriched genes in the: (A) RS4;11 mCherry-low; (B) KMS-12-PE mCherry-low; and (C) KMS-12-PE mCherry-high Epifactors screens. The left column shows a scatterplot for the abundance of individual sgRNAs in sorted and unsorted populations. The red line represent a Lowess curve fitted to the median of data. The right column shows the -log(FDR) plot for each targeted gene. The red dotted line indicates the significance threshold (FDR=0.05).

Screen	Gene	Protein	Function	Modification	Target
RS4;11 mCherry-low	CTBP1	C-terminal binding protein 1	Chromatin remodelling	#	Chromatin
KMS-12-PE mCherry-low	CHTOP	Chromatin target of PRMT1	#	#	DNA
	ZMYM3	Zinc finger, MYM-type 3	Histone modification erase cofactor	Histone acetylation	DNA
KMS-12-PE mCherry-high	SPEN	SPEN family transcriptional repressor	Histone modification erase cofactor, TF	Histone acetylation, TF activator, TF repressor	Histone
	HSPA1A	Heat shock 70kDa protein 1A	Histone modification write cofactor	Histone methylation, Histone acetylation	Histone
	KEAP1	Kelch-like ECH-associated protein 1	Chromatin remodelling	#	Chromatin
	FOXA1	Forkhead box A1	Chromatin remodelling, TF	#	Chromatin, DNA
	HDAC3	Histone deacetylase 3	Histone modification erase	Histone acetylation	Histone
	SAP30	Sin3A-associated protein, 30kDa	Histone modification erase cofactor	Histone acetylation	Histone

Table 4.2. Hit summary of Epifactors library screens

Hits from each Epifactors library screen are summarized. The hits from each screen were ranked by their significance level. The protein coded by each gene and their functions described in the HEMD database are listed[138].

Chapter 5: *NOXA* promoter tiling screens identify *cis*- regulatory elements regulating *NOXA* expression

5.1 Introduction

Cis-regulatory elements (CREs) are regions of noncoding DNA that regulate the transcription of adjacent genes. They are vital components of gene regulatory networks and comprise a large part of the human genome. Indeed, Genome-Wide Association Studies (GWAS) have identified many noncoding variants that are associated with human diseases[151]. CREs can be categorized based on their function. Promoters, for example, are DNA sequences to which proteins bind to initiate transcription of a downstream gene. Promoters are usually 100-1000 base pairs in length, located near the transcription start sites of genes, upstream on the DNA. Promoters consist of binding sites for RNA polymerase, general transcription factors (e.g. TATA box) and cell-type specific transcription factors[152]. Enhancers are another important class of CREs that influence the rate of transcription of genes on the same chromosome. Enhancers can be found adjacent to, far away from or even within the introns of the gene they regulate. They act most often by looping through long distances to come into close physical proximity with the promoters of their target genes[152].

Despite their importance, the identification of CREs has been largely confined to predictions of putative DNA regions, identified by a combination of molecular hallmarks such as consensus sequences for transcription factor binding sites, epigenetic modifications (e.g H3K4me1) and chromatin accessibility, and their evolutionary conservation across organisms[153]. *In vitro* studies that have used CRE DNA fragments in luminescence reporter assays have made it possible to study the role of isolated CREs and their effect on reporter gene expression[154]. However, these methods lack the local chromatin context and broader regulatory interactions of native loci and so reporter cell lines and transgenic animal models must be made for more accurate functional studies of CREs of interest.

With the advent of CRISPR-Cas9 genome engineering technology it is now possible to precisely edit the human genome and, as such, high-throughput loss-of-function genetic screens have been applied successfully to associate genes with specific phenotypes.

CRISPR- Cas9 genome editing has likewise been successfully employed to study CREs in both low throughput[155,156] and high throughput[157-159] experimental systems. For example, Rajagopal et al. carried out a high-throughput CRISPR screen that mapped the effects of genomic variation on the expression of murine embryonic stem cell (mESC)-specific genes[160]. They mapped elements that are required for gene expression by utilising a gRNA library that tiled the enhancers of mESC-specific genes *Tdgf1* and *Zfp42* and measured gene expression with GFP reporters knocked into the endogenous loci. Sanjana et al. described a similar approach utilizing a pooled gRNA library to target large genomic regions (100kb upstream and downstream) surrounding the genes *NF1, NF2* and *CUL3*, which are involved in BRAF inhibitor resistance in melanoma, to identify functional elements relevant to cancer drug resistance[159].

NOXA, a cellular stress induced BH3-only gene, is reported to be a direct target of various transcription factors and epigenetic modulators. Numerous binding sites have been identified on the *NOXA* promoter by traditional luminescence reporter assays, including for p53, p73 and SP1 [93]. (figure5.1) However, few studies have been done to evaluate the function of these proposed elements in their native context. That is, no study to date has linked putative *NOXA* CREs to changes in the gene's expression, in an unbiased way.

In this chapter I conducted a *NOXA* promoter tiling screen, performing high-throughput CRISPR-Cas9 mediated mutagenesis of the *NOXA* promoter region to identify elements relevant to *NOXA* expression and therefore venetoclax drug resistance in blood cancer cell lines.

5.2 Result

5.2.1 Tiling sgRNA library creation

In the tiling sgRNA library, sgRNAs targeting all possible PAM sites within the human *NOXA* promoter were pooled. The pooled sgRNA library had 298 sgRNAs targeting the ~2kb *NOXA* promoter, with an average 6.7bp gap between adjacent cleavages (range: 1bp to 95bp). The targeted promoter region covered the position -2182 to -182 relative to the *NOXA* translation start site (TSS). (figure5.1) 150 non-targeting sgRNAs were included as negative controls. The library was cloned to a lentiviral vector pKLV.

5.2.2 *NOXA* promoter tiling screens

The *NOXA* promoter tiling screens were conducted using two different approaches: mCherry reporter sorting and enrichment following venetoclax treatment. The experimental methodology of the screens was similar to the CRISPR genetic screens in the RS4;11 and KMS- 12-PE *NOXA* reporter cells in chapter 4. The overall schematic for experimental design is depicted in Figure5.2. For each individual screen, two groups of samples were included: the screening group and the baseline control group. 4-6 independent samples were included in each group.

The efficiency of lentiviral infection was optimized prior to performing the screens as described in chapter 4, targeting 30% cell infection in each sample. The cell number in each independent sample was determined by the size of the library and the desired coverage of each guide, in this case, 700-fold. 1 million cells were utilised in each sample as calculated by 448/30% x 700 = 1.0 x 10^6 cells. Following infection with the sgRNA library, each independent sample was expanded without splitting. 48h after infection, the library expressing cells were selected using puromycin.

In the *NOXA* promoter screens performed by sorting on mCherry reporter levels, RS4;11 and KMS-12-PE *NOXA* reporter B/B DKO cells lines with constitutive Cas9 expression were transduced with the tiling library. The first sort was performed 7d after infection in the mCherry reporter samples. The mCherry low populations were defined relative to the expression of mCherry reporter in untransduced cells. The mCherry low cells in the screening groups were sorted two times before harvest. Each sample was monitored by FACS at regular time points. (Figure5.3)

In the screens performed using venetoclax treatment, RS4;11 and KMS-12-PE parental cells with constitutive Cas9 expression were transduced with the tiling library. Venetoclax was added for a 24h treatment once the samples reached required cell number of 10 million. The concentration of venetoclax was determined based on the IC50 of *NOXA* knock-out cells. 1.1uM of venetoclax was applied to RS4;11 cells for 24h, while 0.3uM of venetoclax was applied to KMS-12-PE cells for 24h. Cells were recovered without splitting in drug-free media for up to 7 days after venetoclax treatment.

DNA from the baseline control groups was collected at the time of the first sort or before the addition of venetoclax to record the baseline coverage of the library. DNA from the

mCherry sorting screening groups was collected after the last sort. DNA from the venetoclax treated samples was collected following their recovery. The DNA amplification and indexing were carried out as described in chapter 4. The amplified and indexed DNA were pooled together and sequenced using the Illumina Miseq DNA sequencer. Sequencing reads were then used to quantitate sgRNA representation across the samples.

To analyse the data, raw counts of each sgRNA were normalised and transformed into Log2 scale to generate heatmaps to visualise the enrichment or depletion of each sgRNAs in the library (Figure 5.4 A). *t* statistic scores were calculated for each sgRNA by comparing its representation in the screening and baseline groups across replicate samples. The *t* statistic score for each sgRNA was then mapped to its predicted position of genomic cleavage for visualisation (Figure 5.4 B).

5.2.3 Five regions were identified by *NOXA* promoter tiling screens

The heatmaps and *t* statistic visualisation results are shown in figure5.4. According to the *t* statistic scores, an upper baseline was drawn by the 97.5 percentile for non-targeting control guides and a lower baseline was drawn by the 2.5 percentile for non-targeting control guides. sgRNAs with *t*-scores higher than the upper baseline were considered enriched, while sgRNAs with *t*-scores lower than the lower baseline were considered depleted.

In the two cell lines screened by the two different approaches, five clusters of sgRNAs were either enriched or depleted as hits. These were: 1) A depleted region covering base pairs -1532 to -1283, named depleted 1 region; 2) A depleted region covering base pairs -663 to -433, named depleted 2 region; 3) An enriched region covering base pairs -333 to -283, named enriched 1 region; 4) A depleted region covering base pairs -283 to -253, named depleted 3 region; 5) An enriched region covering base pairs -253 to -193, named as enriched 2 region.

The interpretation of these enriched regions in the screens where low *NOXA* expression and venetoclax resistance were selection criteria, is that the enriched sgRNAs are predicted to cause CRISPR-Cas9 mediated down-regulation of *NOXA* expression, suggesting that the targeted regions of these enriched sgRNAs are NOXA inducing elements. On the contrary, the depleted sgRNAs are predicted to induce CRISPR-Cas9 mediated up regulation of *NOXA*

expression, suggesting that the targeted regions of these depleted sgRNAs are *NOXA* repressing elements.

To validate the findings from the *NOXA* promoter tiling screens, the top-scoring sgRNA from each identified region was individually cloned and transduced into RS4;11 and KMS-12-PE *NOXA* reporter B/B DKO cells, including sgRNA078 for depleted 1 region, sgRNA101 for depleted 2 region, sgRNA267 for enriched 1 region, sgRNA283 for depleted 3 region and sgRNA287 for enriched 2 region. The transduced cells were selected with puromycin then analyzed by FACS. Consistent with the screen results, sgRNAs targeting the depleted regions resulted in upregulation of *NOXA* expression (represented as increase of mCherry fluorescence) while sgRNAs targeting the enriched regions resulted in down regulation of *NOXA* expression (represented as decrease of mCherry fluorescence). (Figure5.5) However, the effects of individual sgRNAs were only observed in a relatively small proportion of the transduced cells. One possible explanation for the observation is that some of the indels generated by the sgRNA were not large enough to induce phenotypic changes.

To characterize the indels induced by individual sgRNAs and map their precise location relative to known transcription factor binding sites on the *NOXA* promoter, the *NOXA* promoter regions from sgRNA transduced cells were sequenced by Miseq. Several pairs of primers were designed to amplify the targeted regions for sequencing. However, due to the presence of repetitive elements in the *NOXA* promoter, some primers failed to produce specific products and will require further optimisation. To date, the depleted 2 region and the enriched 3 region have been successfully amplified by TLseq-8 primers without non-specific amplicons. The amplicons were subjected to Illumina Miseq sequencing, and the results are discussed below in detail.

DNA was extracted from RS4;11 *NOXA* reporter B/B DKO cells expressing Cas9 and either sgRNA283 or sgRNA287 and from RS4;11 parental cells. Each was amplified by TLseq-8 primers and subjected to Illumina Miseq sequencing. The sequences of the amplicons were then mapped to the reference *NOXA* sequence. The actual indels generated by sgRNA283 and sgRNA287 were mapped to published transcription factor binding sites respectively. (Figure5.6) Indels induced by sgRNA287 covered 38bp from position -274 to -237, which didn't match to any known *NOXA* transcription factor binding sites. Indels induced by sgRNA283 covered 25bp from position -288 to -264, which were close to the binding site of SP1, a known *NOXA* repressor[93]. To note, the 5' end of the indels induced by sgRNA283 was

close to the TLseq-8 forward primer binding sites, so there might be larger indels generated by sgRNA283 that were not captured by the current primer set.

DNA was also extracted from RS4;11 *NOXA* reporter B/B DKO cells expressing Cas9 and sgRNA078. DNA from these and parental RS4;11 cells were amplified using the TLseq-1 primers and subjected to Illumina Miseq sequencing. The sequences of the amplicons were then mapped to the reference *NOXA* sequence. The specificity of the TLseq-1 primers was lower than the TLseq-8. For example, in the parental samples, only 40% of the reads matched to reference sequence. The remaining 60% of reads were filtered out as they mapped to other positions in the genome. In the transduced samples, no indels were detected. The sequencing was repeated with two additional sgRNAs targeting depleted 1 region and, again, no indels were detected. (Data not shown) There are two hypotheses for these observations: Firstly, the actual indels induced by sgRNAs targeting depleted 1 region may have been larger than the sequencing amplicons (200bp), thus the indels were not captured by the TLseq-1 primers. Secondly, the phenotypes may have been induced by the sgRNA-Cas9 steric effect instead of genome editing effect, thus no indels were detected. The first hypothesis can be tested by utilising primer pairs located further upstream and downstream of depleted 1 region and comparing the amplicons' size between sgRNA transduced samples and parental sample. The second hypothesis can be tested by utilising the dCas9, a mutant from of Cas9 whose endonuclease activity is removed. If endonuclease activity is not required, then these guides should still produce the phenotype when co-expressed with dCas9. The validations of the depleted 1 region are still ongoing. Primers for sgRNA101 for depleted 2 region and sgRNA267 for enriched 1 region are still undergoing optimisation. The enriched 1 region is especially interesting because the predicted cutting site for sgRNA 267 is close to a known p53/p73 binding site[93].

To conclude, I have identified five putative *NOXA* CREs in the *NOX*A promoter region using a targeted CRISPR-Cas9 tiling screen. Importantly, the screen results were reproducible by different approaches and the effects of knocking out the five putative CREs were validated by individual sgRNA transductions. Indels induced by sgRNA283 from depleted 3 region and sgRNA 287 from enriched 2 region were determined by Miseq. Alignment to known transcription factor binding sites suggested that depleted 3 region may involve alterations to an SP1 binding site. No indels by sgRNAs from depleted 1 region were detected by current primer sets. The indels of sgRNAs from other hit regions are still to be determined.

5.2.4 Depleted 1 region on *NOXA* promoter is hypermethylated

DNA methylation is a process by which methyl groups are added to a DNA molecule. When located within a promoter, DNA methylation typically functions to repress gene transcription. In mammalian genomes, DNA methylation is almost exclusively found in CpG dinucleotides, with the cytosines on both strands being methylated[161]. The depleted 1 region identified by the *NOXA* promoter tiling screen does not contain any published transcription factor binding sites within its sequence, so its role in the transcriptional regulation of *NOXA* is particularly interesting. Also, depleted 1 region is a 240bp, GC-rich region located 1.5kb upstream of the *NOXA* transcription start site and bisulfite sequencing data from our lab has shown that it is hypermethylated in the KMS-12-PE cell line. These data raise the possibility that depleted 1 region and its effect on *NOXA* transcription may be regulated by the presence or absence of methylation.

Targeted bisulfite sequencing was carried out with bisulfite specific primers that amplify depleted 1 region. DNA of several blood cancer cell lines were extracted and subjected to bisulfite conversion and sequencing, including KMS-12-PE, RS4;11, MV4;11, MOLM13, OCI- Ly19 and Granta-519. Results showed that most GC dinucleotides were highly methylated in the depleted 1 region, and the methylation status was conserved across all tested blood cancer cell lines. (figure5.7)

The fact that depleted 1 region is hypermethylated provides a potential reason for its impact on *NOXA* expression. It is possible that the depleted 1 region represses *NOXA* expression by either physically impeding the binding of transcription factors to the gene or recruiting additional proteins that affect chromatin structure of the locus. The Cas9-sgRNA targeting of the depleted 1 region might supress these repressive functions and thereby lead to *NOXA* induction and sensitisation to venetoclax. In future investigations we could assess whether the methylation status of the depleted 1 region influences *NOXA* induction by either using hypomethylating agents or molecular engineering tools (e.g. CRISPR-Cas9-TET1).

5.3 Conclusion

CREs are important components of genetic regulatory networks that comprise a large part of the human genome. CRISPR-Cas9 genome editing allowed me to carry out large scale

functional association studies of CREs. In this chapter, I developed a CRISPR promoter tiling library using ~300 sgRNAs targeting 2kb of *NOXA* promoter. The promoter tiling screens proved to be a powerful method to identify new *NOXA* CREs with both positive and negative control over gene expression. The *NOXA* promoter tiling screen I carried out identified five regions involved in *NOXA* transcriptional regulation. The engineered mutations within the identified regions potentially altered transcription factor occupancy and local epigenetic environments, implicating these regions in *NOXA* regulation and venetoclax resistance. Future work will focus on demonstrating how these identified regions regulate *NOXA* expression, and how these findings could be translated into novel targeted therapies for blood cancers.

Leshchenko et al. demonstrated that, in mantle cell lymphoma lines, *NOXA* expression was induced by hypomethylating agent decitabine, and a hypermethylated region on *NOXA* promoter was demethylated after decitabine treatment[114]. Consistent to their findings, we showed that depleted 1 region is hypermethylated in several blood cancer cell lines, disruption of depleted 1 region induced *NOXA* expression. By venetoclax resistant screening, we also demonstrated the importance of this region in venetoclax resistance. In the next stage of this study, we will focus on the association between the methylation status of this region and *NOXA* induction by using hypomethylating agents and molecular engineering tools, we will also investigate whether this association could be trasnformed into therapeutic drug combination.

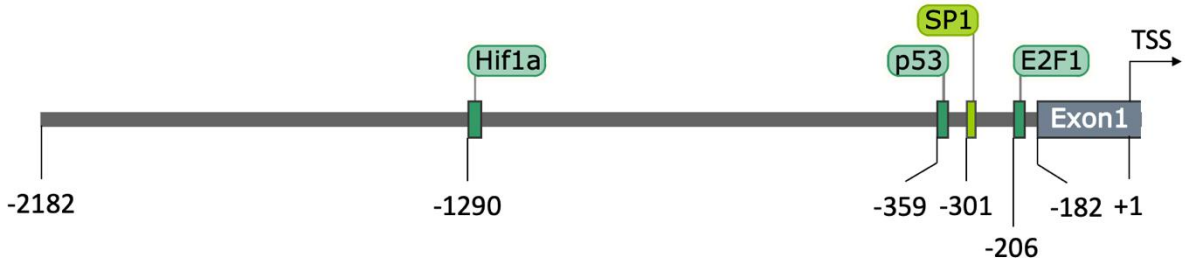

SP1
TSS
Hif1a
p53
E2F1
Exon1
-2182
-1290
-359
-301
-182
+1
-206

Figure 5.1. *NOXA* promoter and consensus binding sites.

Scheme of the *NOXA* promoter with the consensus binding sites of several known transcription factors, including E2F1, SP1, p53 and Hif1a. Numbering starts at the first nucleotide upstream from the TSS.

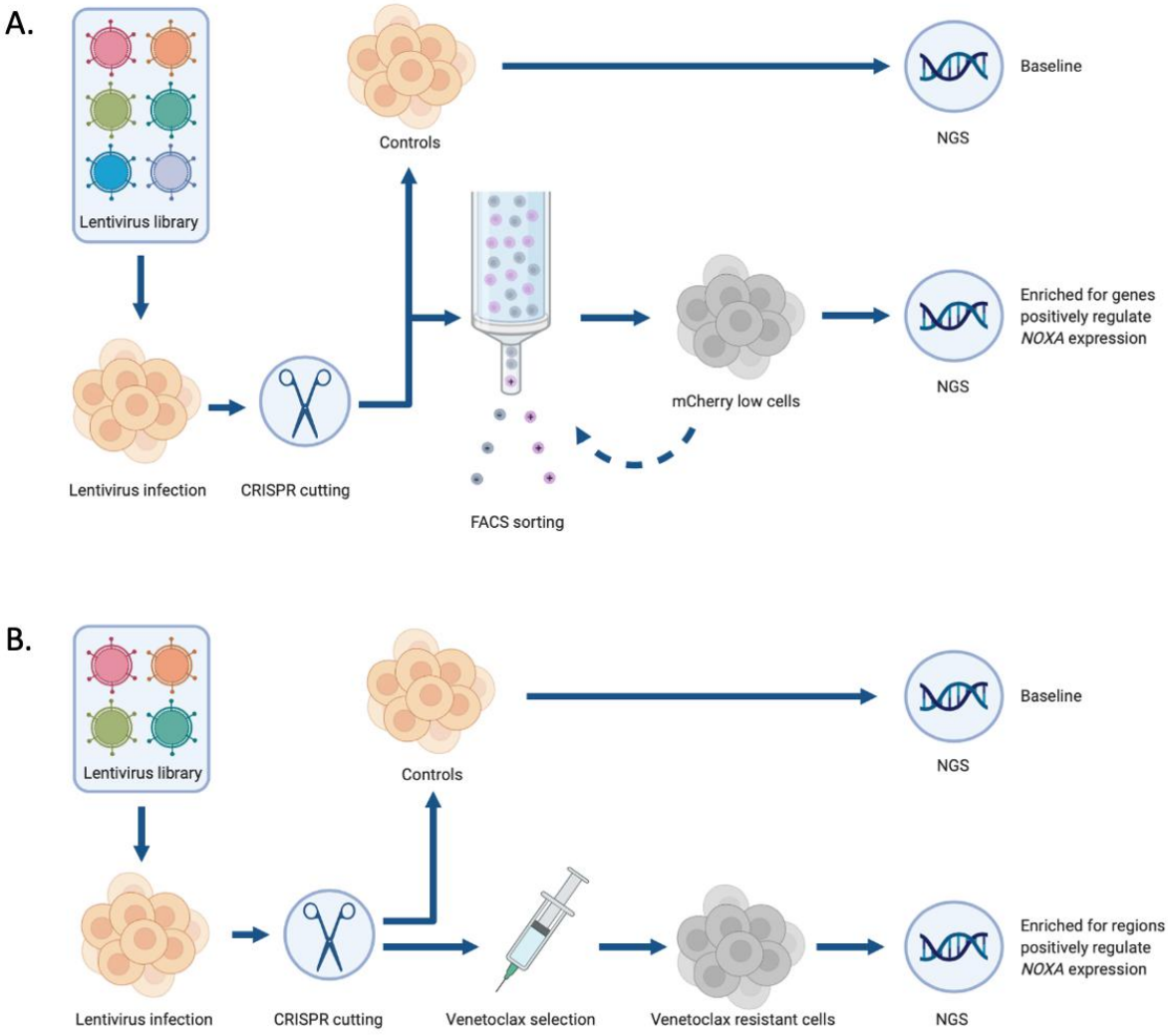

A.
Lentivirus library
Controls
Baseline
NGS
Lentivirus infection
CRISPR cutting
FACS sorting
mCherry low cells
NGS
Enriched for genes positively regulate NOXA expression
B.
Lentivirus library
Controls
Baseline
NGS
Lentivirus infection
CRISPR cutting
Venetoclax selection
Venetoclax resistant cells
NGS
Enriched for regions positively regulate NOXA expression

Figure 5.2. Schematic of *NOXA* promoter tiling screens

(A) mCherry reporter sorting approach: The engineered RS4;11 and KMS-12-PE cells were transduced with *NOXA* promoter tiling libraries. Cells in the baseline control were harvested without enrichment. Cells in the screening groups with lower mCherry fluorescent level were enriched through several rounds of sorting. Targeted *NOXA* promoter regions in the sorted cells were then identified by next generation sequencing.

(B) Venetoclax treatment approach: Cells in the screening groups were enriched by venetoclax treatment, cells that became resistant to venetoclax were enriched.

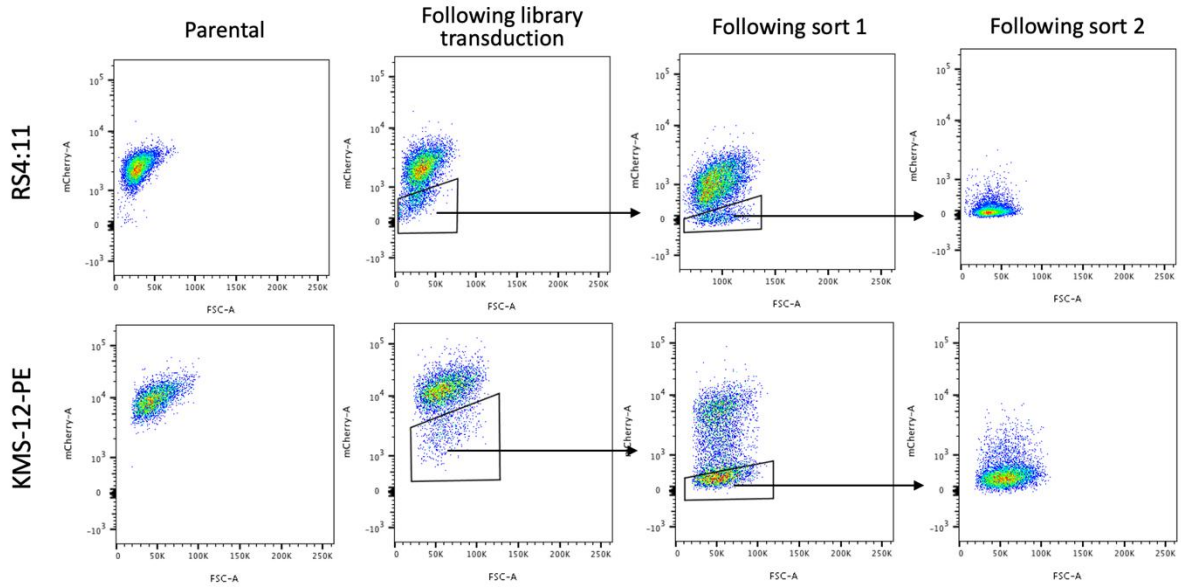
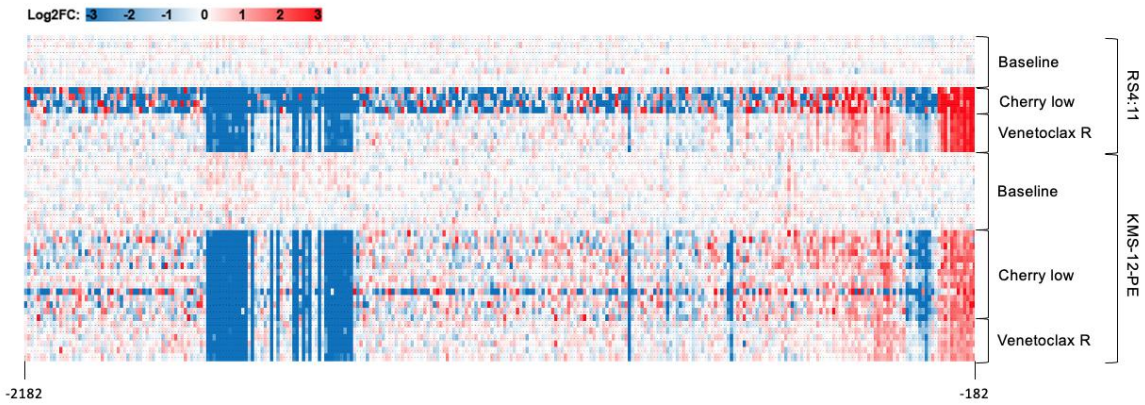

Parental
Following library transduction
Following sort 1
Following sort 2
RS4:11
KMS-12-PE
mCherry-A
FSC-A

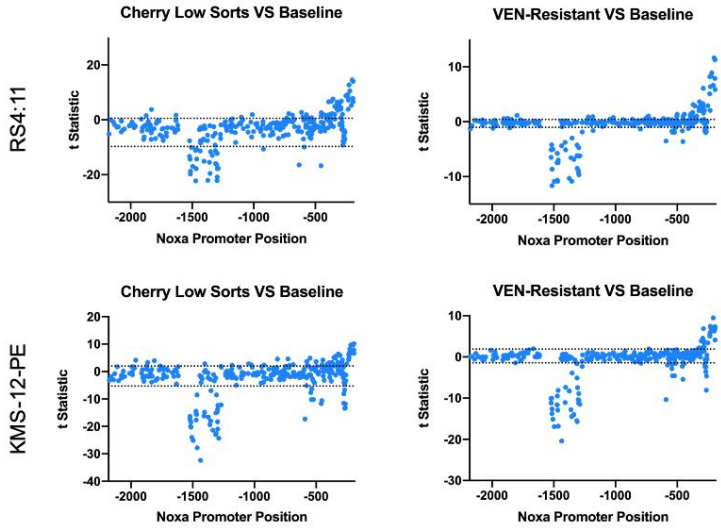
Figure 5.3. Enrichment by mCherry reporter sorting in *NOXA* promoter tiling screens

The engineered RS4;11 and KMS-12-PE cells were transduced with the *NOXA* promoter tiling library. The mCherry low cells of transduced RS;11 and KMS-12-PE samples were enriched through two rounds of sorting. The mCherry fluorescent level of each state of samples were shown. The untransduced parental cells were shown on the left. The gates on each plot indicate which cells were sorted.

A.


Log2FC: -3 -2 -1 0 1 2 3
Baseline
Cherry low
Venetoclax R
Baseline
Cherry low
Venetoclax R
RS4:11
KMS-12-PE
-2182
-182

B.


Cherry Low Sorts VS Baseline
VEN-Resistant VS Baseline
RS4:11
KMS-12-PE
t Statistic
Noxa Promoter Position
-2000
-1500
-1000
-500

C.

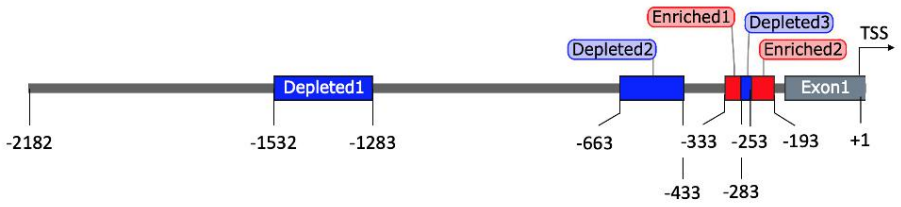

Enriched1
Depleted3
Depleted2
Enriched2
TSS
Depleted1
Exon1
-2182
-1532
-1283
-663
-333
-253
-193
+1
-433
-283

Figure 5.4. Five regions were identified by *NOXA* promoter tiling screens

(A) Heatmap result of the tiling screen. Each column represents one sgRNA in the library. The sgRNAs were arranged by their distance to *NOXA* TSS. Each row represents one sample in the screens. The sample types and cell lines are indicated on the right. sgRNAs are colored by their normalised Log2FC.

(B) *t*-score result of the tiling screen. *t*-score of each sgRNA was calculated and plotted against sgRNA position on *NOXA* promoter. sgRNAs with *t*-score higher than the upper baseline were considered enriched, sgRNAs with *t*-score lower than the lower baseline were considered depleted.

(C) Five regions were named accordingly to their depletion or enrichment effect.

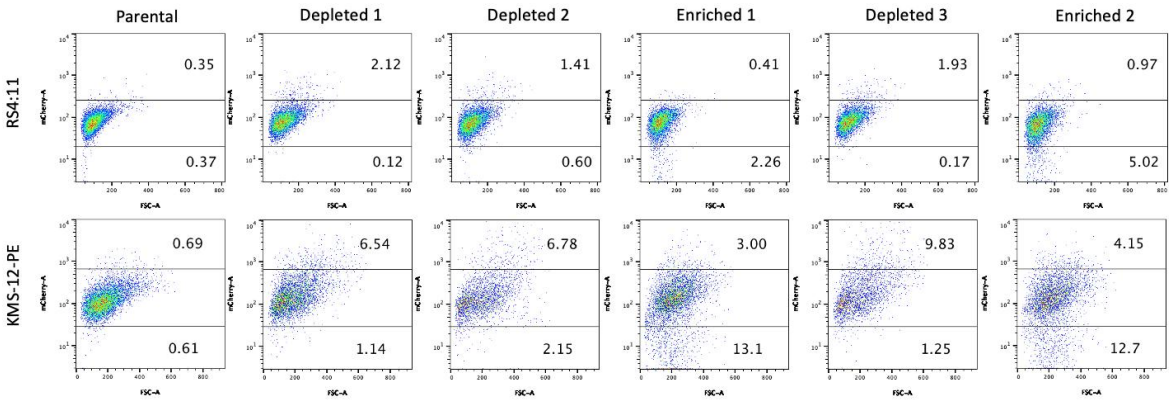

Parental
Depleted 1
Depleted 2
Enriched 1
Depleted 3
Enriched 2
RS4:11
KMS-12-PE
0.35
0.37
2.12
0.12
1.41
0.60
0.41
2.26
1.93
0.17
0.97
5.02
0.69
0.61
6.54
1.14
6.78
2.15
3.00
13.1
9.83
1.25
4.15
12.7
mCherry-A
FSC-A

Figure 5.5. Single sgRNA validation of the identified regions.

The top-scoring sgRNAs from each identified region were individually cloned and transduced into RS4;11 and KMS-12-PE *NOXA* reporter B/B DKO cells. The transduced cells were selected with puromycin then analyzed by FACS. The cells were gated into three proportions (Top gate: high mCherry, middle gate: parental mCherry and bottom gate: low mCherry). sgRNAs targeting the depleted regions resulted in upregulation of *NOXA* expression (high mCherry) sgRNAs targeting the enriched regions resulted in down regulation of *NOXA* expression (low mCherry). The proportion of high mCherry and low mCherry cells were indicated.

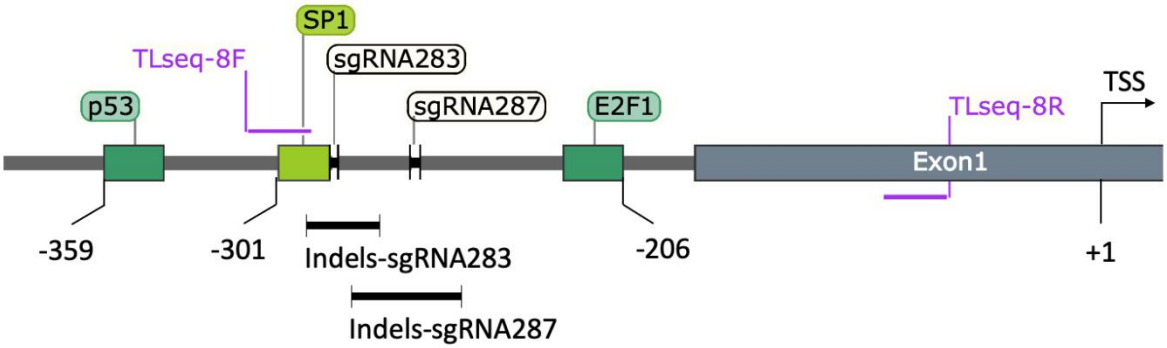

SP1
TLseq-8F
sgRNA283
p53
sgRNA287
E2F1
TLseq-8R
TSS
Exon1
-359
-301
Indels-sgRNA283
-206
+1
Indels-sgRNA287

Figure 5.6. Indels generated by sgRNA283 and sgRNA287.

The actual indels generated by sgRNA283 (depleted 3) and sgRNA287 (enriched 2) were mapped to the consensus transcription factor binding sites. The PAM sites of sgRNA283 and sgRNA287 were marked as well. Indels induced by sgRNA283 covered 25bp from position -288 to -264. Indels induced by sgRNA287 covered 38bp from position -274 to -237.

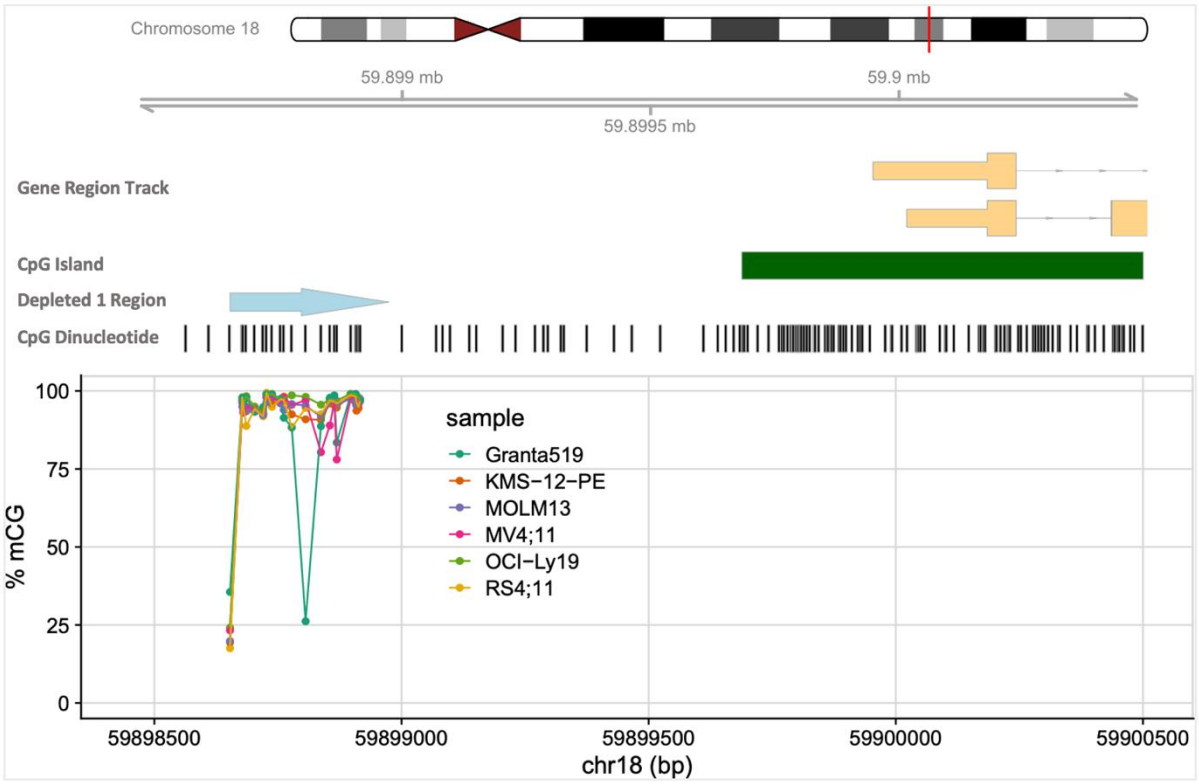

Chromosome 18
59.899 mb
59.9 mb
59.8995 mb
Gene Region Track
CpG Island
Depleted 1 Region
CpG Dinucleotide
100
75
50
25
0
% mCG
sample
Granta519
KMS−12−PE
MOLM13
MV4;11
OCI−Ly19
RS4;11
59898500
59899000
59899500
59900000
59900500
chr18 (bp)

Figure 5.7. Depleted 1 region is hypermethylated in various blood cancer cell lines.

NOXA locates on the chromosome 18q. The shown genome region includes full length of *NOXA* promoter and part of *NOXA* gene body. The exon and protein coding region are marked with yellow boxes, the predicted CpG island are marked with green box. The CpG dinucleotides were indicated as black bars. The depleted 1 region is marked with blue arrow. In the depleted 1 region, most GC dinucleotides were highly methylated. The methylation status was conserved across the tested blood cancer cell lines.

Chapter 6: Conclusion and prospective

It is well established that many cancer cells evade apoptosis to survive. Using BH3 mimetic drugs to target members of the BCL2 family that govern the intrinsic apoptosis pathway has been shown to re-engage apoptosis through direct inhibition of pro-survival proteins. Venetoclax is a BCL2-specific inhibitor that has had great success in treating patients with CLL. However, the potency of venetoclax is often limited in other types of blood cancers due to their co-dependency on MCL1 and other pro-survival proteins.

The aim of my study was to identify new strategies to enhance the action of venetoclax in cancers where MCL1 limits its action. In this thesis, I focused on the BCL2 family member NOXA, a BH3-only protein that binds specifically to MCL1 and mediates its degradation. I hypothesised modulation of *NOXA* expression could reduce cancer cells' dependency on MCL1 and therefore prime cancer cells to venetoclax-induced killing.

Despite holding great therapeutic potential, little is known about the control of *NOXA* gene expression, limiting our ability to target these processes for increased expression and ultimately greater protein production to inhibit MCL1. The work I have presented in this thesis has greatly expanded our understanding of both the transcriptional and epigenetic regulation of *NOXA* and provides a rationale for combining novel agents with venetoclax in cancer therapies.

In order to study the pathways that modulate *NOXA* expression, I generated and validated two *NOXA* reporter cell lines that faithfully report *NOXA* transcription. In Chapter 4, I used these cell lines to perform CRISPR-Cas9 loss-of-function screens and identified multiple genes whose loss-of-function increased *NOXA* expression, including *SPEN*, *HSPA1A*, *KEAP1*, *FOXA1*, *HDAC3* and *SAP30*. KEAP1 is an E3 ligase adapter protein for CUL3 that regulates proteasomal degradation of NRF2, which is a critical transcription factor in the cell stress pathway. This gene and its relation to *NOXA* expression is of particular interest as no published work has linked KEAP1/NRF2 to *NOXA* transcriptional regulation before. In addition, HDAC3 is a member of the histone deacetylase family. Histone deacetylation is a well-studied mechanism of transcriptional silencing and it is known that several pan-HDAC inhibitors are able to increase *NOXA* expression in cancer cells. However, the specific

mechanism by which this happens is still to be determined and if HDAC3 is particularly important, this could help with designing therapies that are targeted more specifically. Apart from *KEAP1* and *HDAC3*, none of the other genes identified in my screens have has been associated with *NOXA* regulation before. Future validation studies of single gene knockouts are needed to confirm their importance. Once confirmed, the screen results may provide novel mechanisms of *NOXA* regulation, and rationale for new cancer therapies.

To further integrate the transcriptional regulation of *NOXA*, in Chapter 5 I carried out *NOXA* promoter tiling screens to identify CREs involved in the gene's function. Utilising a tiling sgRNA library that targeted 2kb of the *NOXA* promoter, I found five distinct CREs that regulate *NOXA* transcription. CRISPR-Cas9 induced genome editing of these regions led to either an increase (depleted 1,2 and 3 regions) or decrease (enriched 1 and 2 regions) in *NOXA* expression, and respectively, either an increase or decrease in sensitivity to venetoclax treatment. These identified regions might be binding sites of *NOXA* transcription factors or important elements for *NOXA* epigenetic regulation. Mapping these CREs relative to known *NOXA* transcription factor binding sites is one way to explore this idea. Enriched 1 region, whose deletion resulted in decreased *NOXA* expression, was in close proximity to a known P53 binding site. Likewise, depleted 3 region, whose deletion was associated with increased *NOXA* expression, overlapped a known transcriptional repressor, SP1, binding site. Excitingly, none of the other three regions that I identified map to published regulatory elements of *NOXA*. Depleted 1 region is a hypermethylated sequence that represses *NOXA* expression and, interestingly, upon further investigation I found that this region is also hypermethylated across several blood cancer cell lines, indicating that the methylation state of this region may have a crucial role in the epigenetic regulation of *NOXA* expression. Future investigations will focus on elucidating how these elements regulate *NOXA* expression and whether the regulation can be tweaked by known or novel agents.

To conclude, my work has shown that CRISPR-Cas9 genetic loss-of-function screens and promoter tiling screens combined with endogenous reporter assay are powerful tools in studying the *trans* and *cis* regulation pathways of *NOXA*. I have uncovered a role for six genes and five DNA elements in the *NOXA* promoter region which modulate expression in the *NOXA*-reporter cell lines that I created. The next stage of this study will be to investigate the mechanisms of action of the individual hits identified in the screen in order to further understand their roles in *NOXA* regulation and their therapeutic potential. Furthermore, the

associations of these novel *trans* and *cis* regulators will provide a more detailed and comprehensive understanding of the complex network of signalling pathways acting upstream of *NOXA* and, specifically, establishing how the CREs I have identified act to integrate these inputs via the binding of TFs will be of great importance.

There are limitations of the experiment design and could be optimised in future studies. First, the generated *NOXA* reporter lines are transcription reporter system that cannot represent the actual protein level in cells. The role of protein degradation in *NOXA* regulation is not represented in the *NOXA* reporter system. Second, the reporter screening method might generate "non-functional" hits that increase *NOXA* expression but have little impact on venetoclax sensitivity. However, this could be easily validated by individual knock-out studies. Finally, in addition to the 2-kb *NOXA* promoter region that I studied, it is possible that other CREs exist within the gene body of *NOXA* itself, further upstream of the promoter or indeed downstream of the gene. The *NOXA* tiling screen library could be expanded to search for these additional elements to further expand our understanding of the regulation of this critical gene.

Furthermore, the methodology of CRISPR screens coupled with gene-specific reporter assays could be applied to the study of the transcriptional regulation of other BCL2 family proteins, which would provide critical information in elucidating resistance mechanisms to BH3 mimetic drugs, and potentially uncover more therapeutic targets to induce the apoptotic pathway in human diseases where its initiation is stalled. Future studies may also use the *NOXA* reporter cell lines, and potentially other BCL2 family protein reporter cell lines to perform small molecule screens for new drug discoveries.

Bibliography

1 Strasser, A., Cory, S. & Adams, J. M. Deciphering the rules of programmed cell death to improve therapy of cancer and other diseases. *EMBO J* **30**, 3667-3683, doi:10.1038/emboj.2011.307 (2011).

2 Kerr, J. F., Wyllie, A. H. & Currie, A. R. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer* **26**, 239-257, doi:10.1038/bjc.1972.33 (1972).

3 Renehan, A. G., Bach, S. P. & Potten, C. S. The relevance of apoptosis for cellular homeostasis and tumorigenesis in the intestine. *Can J Gastroenterol* **15**, 166-176, doi:10.1155/2001/164727 (2001).

4 Nijhawan, D., Honarpour, N. & Wang, X. Apoptosis in neural development and disease. *Annu Rev Neurosci* **23**, 73-87, doi:10.1146/annurev.neuro.23.1.73 (2000).

5 Greenhalgh, D. G. The role of apoptosis in wound healing. *Int J Biochem Cell Biol* **30**, 1019-1030, doi:10.1016/s1357-2725(98)00058-2 (1998).

6 Kerr, J. F., Winterford, C. M. & Harmon, B. V. Apoptosis. Its significance in cancer and cancer therapy. *Cancer* **73**, 2013-2026, doi:10.1002/1097-0142(19940415)73:8<2013::aid-cncr2820730802>3.0.co;2-j (1994).

7 Locksley, R. M., Killeen, N. & Lenardo, M. J. The TNF and TNF receptor superfamilies: integrating mammalian biology. *Cell* **104**, 487-501, doi:10.1016/s0092-8674(01)00237-9 (2001).

8 Kischkel, F. C. *et al.* Cytotoxicity-dependent APO-1 (Fas/CD95)-associated proteins form a death-inducing signaling complex (DISC) with the receptor. *EMBO J* **14**, 5579-5588 (1995).

9 Jost, P. J. *et al.* XIAP discriminates between type I and type II FAS-induced apoptosis. *Nature* **460**, 1035-1039, doi:10.1038/nature08229 (2009).

10 Green, D. R. & Kroemer, G. The pathophysiology of mitochondrial cell death. *Science* **305**, 626-629, doi:10.1126/science.1099320 (2004).

11 Bakhshi, A. *et al.* Cloning the chromosomal breakpoint of t(14;18) human lymphomas: clustering around JH on chromosome 14 and near a transcriptional unit on 18. *Cell* **41**, 899-906, doi:10.1016/s0092-8674(85)80070-2 (1985).

12 Tsujimoto, Y., Cossman, J., Jaffe, E. & Croce, C. M. Involvement of the bcl-2 gene in human follicular lymphoma. *Science* **228**, 1440-1443, doi:10.1126/science.3874430 (1985).

13 Vaux, D. L., Cory, S. & Adams, J. M. Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. *Nature* **335**, 440-442, doi:10.1038/335440a0 (1988).

14 McDonnell, T. J. *et al.* bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation. *Cell* **57**, 79-88, doi:10.1016/0092-8674(89)90174-8 (1989).

15 Strasser, A., Harris, A. W., Bath, M. L. & Cory, S. Novel primitive lymphoid tumours induced in transgenic mice by cooperation between myc and bcl-2. *Nature* **348**, 331-333, doi:10.1038/348331a0 (1990).

16 Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57-70, doi:10.1016/s0092-8674(00)81683-9 (2000).
17 Lindsten, T. *et al.* The combined functions of proapoptotic Bcl-2 family members bak and bax are essential for normal development of multiple tissues. *Mol Cell* **6**, 1389-1399, doi:10.1016/s1097-2765(00)00136-2 (2000).
18 Puthalakath, H. & Strasser, A. Keeping killers on a tight leash: transcriptional and post-translational control of the pro-apoptotic activity of BH3-only proteins. *Cell Death Differ* **9**, 505-512, doi:10.1038/sj.cdd.4400998 (2002).
19 Day, C. L. *et al.* Solution structure of prosurvival Mcl-1 and characterization of its binding by proapoptotic BH3-only ligands. *J Biol Chem* **280**, 4738-4744, doi:10.1074/jbc.M411434200 (2005).
20 Sattler, M. *et al.* Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis. *Science* **275**, 983-986, doi:10.1126/science.275.5302.983 (1997).
21 Petros, A. M. *et al.* Solution structure of the antiapoptotic protein bcl-2. *Proc Natl Acad Sci U S A* **98**, 3012-3017, doi:10.1073/pnas.041619798 (2001).
22 Popgeorgiev, N., Jabbour, L. & Gillet, G. Subcellular Localization and Dynamics of the Bcl-2 Family of Proteins. *Front Cell Dev Biol* **6**, 13, doi:10.3389/fcell.2018.00013 (2018).
23 Sorenson, C. M. Bcl-2 family members and disease. *Biochim Biophys Acta* **1644**, 169-177, doi:10.1016/j.bbamcr.2003.08.010 (2004).
24 Veis, D. J., Sorenson, C. M., Shutter, J. R. & Korsmeyer, S. J. Bcl-2-deficient mice demonstrate fulminant lymphoid apoptosis, polycystic kidneys, and hypopigmented hair. *Cell* **75**, 229-240, doi:10.1016/0092-8674(93)80065-m (1993).
25 Hamasaki, A. *et al.* Accelerated neutrophil apoptosis in mice lacking A1-a, a subtype of the bcl-2-related A1 gene. *J Exp Med* **188**, 1985-1992, doi:10.1084/jem.188.11.1985 (1998).
26 Motoyama, N. *et al.* Massive cell death of immature hematopoietic cells and neurons in Bcl-x-deficient mice. *Science* **267**, 1506-1510, doi:10.1126/science.7878471 (1995).
27 Rinkenberger, J. L., Horning, S., Klocke, B., Roth, K. & Korsmeyer, S. J. Mcl-1 deficiency results in peri-implantation embryonic lethality. *Genes Dev* **14**, 23-27 (2000).
28 Rucker, E. B., 3rd *et al.* Bcl-x and Bax regulate mouse primordial germ cell survival and apoptosis during embryogenesis. *Mol Endocrinol* **14**, 1038-1052, doi:10.1210/mend.14.7.0465 (2000).
29 Mason, K. D. *et al.* Programmed anuclear cell death delimits platelet life span. *Cell* **128**, 1173-1186, doi:10.1016/j.cell.2007.01.037 (2007).
30 Takehara, T. *et al.* Hepatocyte-specific disruption of Bcl-xL leads to continuous hepatocyte apoptosis and liver fibrotic responses. *Gastroenterology* **127**, 1189-1197, doi:10.1053/j.gastro.2004.07.019 (2004).
31 Wang, X. *et al.* Deletion of MCL-1 causes lethal cardiac failure and mitochondrial dysfunction. *Genes Dev* **27**, 1351-1364, doi:10.1101/gad.215855.113 (2013).
32 Opferman, J. T. *et al.* Obligate role of anti-apoptotic MCL-1 in the survival of hematopoietic stem cells. *Science* **307**, 1101-1104, doi:10.1126/science.1106114 (2005).
33 Vikstrom, I. *et al.* Mcl-1 is essential for germinal center formation and B cell memory. *Science* **330**, 1095-1099, doi:10.1126/science.1191793 (2010).

34 Peperzak, V. *et al.* Mcl-1 is essential for the survival of plasma cells. *Nat Immunol* **14**, 290-297, doi:10.1038/ni.2527 (2013).
35 Omari, S. *et al.* Mcl-1 is a key regulator of the ovarian reserve. *Cell Death Dis* **6**, e1755, doi:10.1038/cddis.2015.95 (2015).
36 Wolter, K. G. *et al.* Movement of Bax from the cytosol to mitochondria during apoptosis. *J Cell Biol* **139**, 1281-1292, doi:10.1083/jcb.139.5.1281 (1997).
37 Cheng, E. H., Sheiko, T. V., Fisher, J. K., Craigen, W. J. & Korsmeyer, S. J. VDAC2 inhibits BAK activation and mitochondrial apoptosis. *Science* **301**, 513-517, doi:10.1126/science.1083995 (2003).
38 Dewson, G. *et al.* To trigger apoptosis, Bak exposes its BH3 domain and homodimerizes via BH3:groove interactions. *Mol Cell* **30**, 369-380, doi:10.1016/j.molcel.2008.04.005 (2008).
39 Echeverry, N. *et al.* Intracellular localization of the BCL-2 family member BOK and functional implications. *Cell Death Differ* **20**, 785-799, doi:10.1038/cdd.2013.10 (2013).
40 Hinds, M. G. *et al.* Bim, Bad and Bmf: intrinsically unstructured BH3-only proteins that undergo a localized conformational change upon binding to prosurvival Bcl-2 targets. *Cell Death Differ* **14**, 128-136, doi:10.1038/sj.cdd.4401934 (2007).
41 Chen, L. *et al.* Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. *Mol Cell* **17**, 393-403, doi:10.1016/j.molcel.2004.12.030 (2005).
42 Letai, A. *et al.* Distinct BH3 domains either sensitize or activate mitochondrial apoptosis, serving as prototype cancer therapeutics. *Cancer Cell* **2**, 183-192, doi:10.1016/s1535-6108(02)00127-7 (2002).
43 Oda, E. *et al.* Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis. *Science* **288**, 1053-1058, doi:10.1126/science.288.5468.1053 (2000).
44 Dijkers, P. F., Medema, R. H., Lammers, J. W., Koenderman, L. & Coffer, P. J. Expression of the pro-apoptotic Bcl-2 family member Bim is regulated by the forkhead transcription factor FKHR-L1. *Curr Biol* **10**, 1201-1204, doi:10.1016/s0960-9822(00)00728-4 (2000).
45 Scaffidi, C. *et al.* Two CD95 (APO-1/Fas) signaling pathways. *EMBO J* **17**, 1675-1687, doi:10.1093/emboj/17.6.1675 (1998).
46 Willis, S. N. *et al.* Apoptosis initiated when BH3 ligands engage multiple Bcl-2 homologs, not Bax or Bak. *Science* **315**, 856-859, doi:10.1126/science.1133289 (2007).
47 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**, 973-988, doi:10.1101/gad.276725.115 (2016).
48 Oltersdorf, T. *et al.* An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **435**, 677-681, doi:10.1038/nature03579 (2005).
49 Del Gaizo Moore, V. *et al.* Chronic lymphocytic leukemia requires BCL2 to sequester prodeath BIM, explaining sensitivity to BCL2 antagonist ABT-737. *J Clin Invest* **117**, 112-121, doi:10.1172/JCI28281 (2007).
50 Tse, C. *et al.* ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Res* **68**, 3421-3428, doi:10.1158/0008-5472.CAN-07-5836 (2008).

51 Mason, K. D. *et al.* The BH3 mimetic compound, ABT-737, synergizes with a range of cytotoxic chemotherapy agents in chronic lymphocytic leukemia. *Leukemia* **23**, 2034-2041, doi:10.1038/leu.2009.151 (2009).
52 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, doi:10.1016/S1470-2045(10)70261-8 (2010).
53 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, doi:10.1038/nm.3048 (2013).
54 Vogler, M., Dinsdale, D., Dyer, M. J. & Cohen, G. M. ABT-199 selectively inhibits BCL2 but not BCL2L1 and efficiently induces apoptosis of chronic lymphocytic leukaemic cells but not platelets. *Br J Haematol* **163**, 139-142, doi:10.1111/bjh.12457 (2013).
55 Kozopas, K. M., Yang, T., Buchan, H. L., Zhou, P. & Craig, R. W. MCL1, a gene expressed in programmed myeloid cell differentiation, has sequence similarity to BCL2. *Proc Natl Acad Sci U S A* **90**, 3516-3520, doi:10.1073/pnas.90.8.3516 (1993).
56 Dzhagalov, I., St John, A. & He, Y. W. The antiapoptotic protein Mcl-1 is essential for the survival of neutrophils but not macrophages. *Blood* **109**, 1620-1626, doi:10.1182/blood-2006-03-013771 (2007).
57 Hikita, H. *et al.* Mcl-1 and Bcl-xL cooperatively maintain integrity of hepatocytes in developing and adult murine liver. *Hepatology* **50**, 1217-1226, doi:10.1002/hep.23126 (2009).
58 Arbour, N. *et al.* Mcl-1 is a key regulator of apoptosis during CNS development and after DNA damage. *J Neurosci* **28**, 6068-6078, doi:10.1523/JNEUROSCI.4940-07.2008 (2008).
59 Glaser, S. P. *et al.* Anti-apoptotic Mcl-1 is essential for the development and sustained growth of acute myeloid leukemia. *Genes Dev* **26**, 120-125, doi:10.1101/gad.182980.111 (2012).
60 Derenne, S. *et al.* Antisense strategy shows that Mcl-1 rather than Bcl-2 or Bcl-x(L) is an essential survival protein of human myeloma cells. *Blood* **100**, 194-199, doi:10.1182/blood.v100.1.194 (2002).
61 Fleischer, B. *et al.* Mcl-1 is an anti-apoptotic factor for human hepatocellular carcinoma. *Int J Oncol* **28**, 25-32 (2006).
62 Song, L., Coppola, D., Livingston, S., Cress, D. & Haura, E. B. Mcl-1 regulates survival and sensitivity to diverse apoptotic stimuli in human non-small cell lung cancer cells. *Cancer Biol Ther* **4**, 267-276, doi:10.4161/cbt.4.3.1496 (2005).
63 Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899-905, doi:10.1038/nature08822 (2010).
64 van Delft, M. F. *et al.* The BH3 mimetic ABT-737 targets selective Bcl-2 proteins and efficiently induces apoptosis via Bak/Bax if Mcl-1 is neutralized. *Cancer Cell* **10**, 389-399, doi:10.1016/j.ccr.2006.08.027 (2006).
65 Bojarczuk, K. *et al.* BCR signaling inhibitors differ in their ability to overcome Mcl-1-mediated resistance of CLL B cells to ABT-199. *Blood* **127**, 3192-3201, doi:10.1182/blood-2015-10-675009 (2016).
66 Isomoto, H. *et al.* Interleukin 6 upregulates myeloid cell leukemia-1 expression through a STAT3 pathway in cholangiocarcinoma cells. *Hepatology* **42**, 1329-1338, doi:10.1002/hep.20966 (2005).

67 Senichkin, V. V., Streletskaia, A. Y., Gorbunova, A. S., Zhivotovsky, B. & Kopeina, G. S. Saga of Mcl-1: regulation from transcription to degradation. *Cell Death Differ* **27**, 405-419, doi:10.1038/s41418-019-0486-3 (2020).
68 Cui, J. & Placzek, W. J. Post-Transcriptional Regulation of Anti-Apoptotic BCL2 Family Members. *Int J Mol Sci* **19**, doi:10.3390/ijms19010308 (2018).
69 Mott, J. L., Kobayashi, S., Bronk, S. F. & Gores, G. J. mir-29 regulates Mcl-1 protein expression and apoptosis. *Oncogene* **26**, 6133-6140, doi:10.1038/sj.onc.1210436 (2007).
70 Zhong, Q., Gao, W., Du, F. & Wang, X. Mule/ARF-BP1, a BH3-only E3 ubiquitin ligase, catalyzes the polyubiquitination of Mcl-1 and regulates apoptosis. *Cell* **121**, 1085-1095, doi:10.1016/j.cell.2005.06.009 (2005).
71 Subramanian, A., Andronache, A., Li, Y. C. & Wade, M. Inhibition of MARCH5 ubiquitin ligase abrogates MCL1-dependent resistance to BH3 mimetics via NOXA. *Oncotarget* **7**, 15986-16002, doi:10.18632/oncotarget.7558 (2016).
72 Carroll, R. G., Hollville, E. & Martin, S. J. Parkin sensitizes toward apoptosis induced by mitochondrial depolarization through promoting degradation of Mcl-1. *Cell Rep* **9**, 1538-1553, doi:10.1016/j.celrep.2014.10.046 (2014).
73 Djajawi, T. M. *et al.* MARCH5 requires MTCH2 to coordinate proteasomal turnover of the MCL1:NOXA complex. *Cell Death Differ* **27**, 2484-2499, doi:10.1038/s41418-020-0517-0 (2020).
74 Haschka, M. D. *et al.* MARCH5-dependent degradation of MCL1/NOXA complexes defines susceptibility to antimitotic drug treatment. *Cell Death Differ* **27**, 2297-2312, doi:10.1038/s41418-020-0503-6 (2020).
75 Nam, S. *et al.* Dasatinib (BMS-354825) inhibits Stat5 signaling associated with apoptosis in chronic myelogenous leukemia cells. *Mol Cancer Ther* **6**, 1400-1405, doi:10.1158/1535-7163.MCT-06-0446 (2007).
76 Mow, B. M. *et al.* Effects of the Bcr/abl kinase inhibitors STI571 and adaphostin (NSC 680410) on chronic myelogenous leukemia cells in vitro. *Blood* **99**, 664-671, doi:10.1182/blood.v99.2.664 (2002).
77 MacCallum, D. E. *et al.* Seliciclib (CYC202, R-Roscovitine) induces cell death in multiple myeloma cells by inhibition of RNA polymerase II-dependent transcription and down-regulation of Mcl-1. *Cancer Res* **65**, 5399-5407, doi:10.1158/0008-5472.CAN-05-0233 (2005).
78 Akgul, C. Mcl-1 is a potential therapeutic target in multiple types of cancer. *Cell Mol Life Sci* **66**, 1326-1336, doi:10.1007/s00018-008-8637-6 (2009).
79 Tron, A. E. *et al.* Discovery of Mcl-1-specific inhibitor AZD5991 and preclinical activity in multiple myeloma and acute myeloid leukemia. *Nature Communications* **9**, 5341, doi:10.1038/s41467-018-07551-w (2018).
80 Kotschy, A. *et al.* The MCL1 inhibitor S63845 is tolerable and effective in diverse cancer models. *Nature* **538**, 477-482, doi:10.1038/nature19830 (2016).
81 Caenepeel, S. *et al.* AMG 176, a Selective MCL1 Inhibitor, Is Effective in Hematologic Cancer Models Alone and in Combination with Established Therapies. *Cancer Discov* **8**, 1582-1597, doi:10.1158/2159-8290.Cd-18-0387 (2018).
82 Hird, A. W. & Tron, A. E. Recent advances in the development of Mcl-1 inhibitors for cancer therapy. *Pharmacol Ther* **198**, 59-67, doi:10.1016/j.pharmthera.2019.02.007 (2019).

83 Kotschy, A. *et al.* The MCL1 inhibitor S63845 is tolerable and effective in diverse cancer models. *Nature* **538**, 477-482, doi:10.1038/nature19830 (2016).
84 Czabotar, P. E. *et al.* Structural insights into the degradation of Mcl-1 induced by BH3 domains. *Proc Natl Acad Sci U S A* **104**, 6217-6222, doi:10.1073/pnas.0701297104 (2007).
85 Song, T. *et al.* Deactivation of Mcl-1 by Dual-Function Small-Molecule Inhibitors Targeting the Bcl-2 Homology 3 Domain and Facilitating Mcl-1 Ubiquitination. *Angew Chem Int Ed Engl* **55**, 14250-14256, doi:10.1002/anie.201606543 (2016).
86 Gomez-Bougie, P. *et al.* Noxa controls Mule-dependent Mcl-1 ubiquitination through the regulation of the Mcl-1/USP9X interaction. *Biochem Biophys Res Commun* **413**, 460-464, doi:10.1016/j.bbrc.2011.08.118 (2011).
87 Vogler, M. BCL2A1: the underdog in the BCL2 family. *Cell Death Differ* **19**, 67-74, doi:10.1038/cdd.2011.158 (2012).
88 Zhou, W. *et al.* Neddylation E2 UBE2F Promotes the Survival of Lung Cancer Cells by Activating CRL5 to Degrade NOXA via the K11 Linkage. *Clin Cancer Res* **23**, 1104-1116, doi:10.1158/1078-0432.CCR-16-1585 (2017).
89 Hijikata, M., Kato, N., Sato, T., Kagami, Y. & Shimotohno, K. Molecular cloning and characterization of a cDNA for a novel phorbol-12-myristate-13-acetate-responsive gene that is highly expressed in an adult T-cell leukemia cell line. *J Virol* **64**, 4632-4639, doi:10.1128/JVI.64.10.4632-4639.1990 (1990).
90 Wang, Z. & Sun, Y. Identification and characterization of two splicing variants of human Noxa. *Anticancer Res* **28**, 1667-1674 (2008).
91 Fei, P., Bernhard, E. J. & El-Deiry, W. S. Tissue-specific induction of p53 targets in vivo. *Cancer Res* **62**, 7316-7327 (2002).
92 Villunger, A. *et al.* p53- and drug-induced apoptotic responses mediated by BH3-only proteins puma and noxa. *Science* **302**, 1036-1038, doi:10.1126/science.1090072 (2003).
93 Grande, L. *et al.* Transcription factors Sp1 and p73 control the expression of the proapoptotic protein NOXA in the response of testicular embryonal carcinoma cells to cisplatin. *J Biol Chem* **287**, 26495-26505, doi:10.1074/jbc.M112.376319 (2012).
94 Flinterman, M. *et al.* E1A activates transcription of p73 and Noxa to induce apoptosis. *J Biol Chem* **280**, 5945-5959, doi:10.1074/jbc.M406661200 (2005).
95 Carmeliet, P. *et al.* Role of HIF-1alpha in hypoxia-mediated apoptosis, cell proliferation and tumour angiogenesis. *Nature* **394**, 485-490, doi:10.1038/28867 (1998).
96 Kim, J. Y., Ahn, H. J., Ryu, J. H., Suk, K. & Park, J. H. BH3-only protein Noxa is a mediator of hypoxic cell death induced by hypoxia-inducible factor 1alpha. *J Exp Med* **199**, 113-124, doi:10.1084/jem.20030613 (2004).
97 Toth, A. *et al.* Endoplasmic reticulum stress as a novel therapeutic target in heart diseases. *Cardiovasc Hematol Disord Drug Targets* **7**, 205-218, doi:10.2174/187152907781745260 (2007).
98 Puthalakath, H. *et al.* ER stress triggers apoptosis by activating BH3-only protein Bim. *Cell* **129**, 1337-1349, doi:10.1016/j.cell.2007.04.027 (2007).
99 Sharma, K. *et al.* p53-independent Noxa induction by cisplatin is regulated by ATF3/ATF4 in head and neck squamous cell carcinoma cells. *Mol Oncol* **12**, 788-798, doi:10.1002/1878-0261.12172 (2018).

100 Yamashita, M. *et al.* Bmi1 regulates memory CD4 T cell survival via repression of the Noxa gene. *J Exp Med* **205**, 1109-1120, doi:10.1084/jem.20072000 (2008).
101 Gomez-Bougie, P. *et al.* Noxa up-regulation and Mcl-1 cleavage are associated to apoptosis induction by bortezomib in multiple myeloma. *Cancer Res* **67**, 5418-5424, doi:10.1158/0008-5472.CAN-06-4322 (2007).
102 Lamothe, B., Wierda, W. G., Keating, M. J. & Gandhi, V. Carfilzomib Triggers Cell Death in Chronic Lymphocytic Leukemia by Inducing Proapoptotic and Endoplasmic Reticulum Stress Responses. *Clin Cancer Res* **22**, 4712-4726, doi:10.1158/1078-0432.CCR-15-2522 (2016).
103 Perez-Galan, P. *et al.* The proteasome inhibitor bortezomib induces apoptosis in mantle-cell lymphoma through generation of ROS and Noxa activation independent of p53 status. *Blood* **107**, 257-264, doi:10.1182/blood-2005-05-2091 (2006).
104 Qin, J. Z. *et al.* Proteasome inhibitors trigger NOXA-mediated apoptosis in melanoma and myeloma cells. *Cancer Res* **65**, 6282-6293, doi:10.1158/0008-5472.CAN-05-0676 (2005).
105 Albert, M. C., Brinkmann, K. & Kashkar, H. Noxa and cancer therapy: Tuning up the mitochondrial death machinery in response to chemotherapy. *Mol Cell Oncol* **1**, e29906, doi:10.4161/mco.29906 (2014).
106 Fribley, A. M. *et al.* Proteasome inhibitor PS-341 induces apoptosis in cisplatin-resistant squamous cell carcinoma cells by induction of Noxa. *J Biol Chem* **281**, 31440-31447, doi:10.1074/jbc.M604356200 (2006).
107 Narita, T. *et al.* Lower expression of activating transcription factors 3 and 4 correlates with shorter progression-free survival in multiple myeloma patients receiving bortezomib plus dexamethasone therapy. *Blood Cancer J* **5**, e373, doi:10.1038/bcj.2015.98 (2015).
108 Brinkmann, K. *et al.* Ubiquitin C-terminal hydrolase-L1 potentiates cancer chemosensitivity by stabilizing NOXA. *Cell Rep* **3**, 881-891, doi:10.1016/j.celrep.2013.02.014 (2013).
109 Wang, Y. *et al.* Targeting protein neddylation with an NEDD8-activating enzyme inhibitor MLN4924 induced apoptosis or senescence in human lymphoma cells. *Cancer Biol Ther* **16**, 420-429, doi:10.1080/15384047.2014.1003003 (2015).
110 Bolomsky, A., Schlangen, K., Schreiner, W., Zojer, N. & Ludwig, H. Targeting of BMI-1 with PTC-209 shows potent anti-myeloma activity and impairs the tumour microenvironment. *J Hematol Oncol* **9**, 17, doi:10.1186/s13045-016-0247-4 (2016).
111 Tiffen, J. C. *et al.* Targeting activating mutations of EZH2 leads to potent cell growth inhibition in human melanoma by derepression of tumor suppressor genes. *Oncotarget* **6**, 27023-27036, doi:10.18632/oncotarget.4809 (2015).
112 Zhou, L. *et al.* HDAC inhibition by SNDX-275 (Entinostat) restores expression of silenced leukemia-associated transcription factors Nur77 and Nor1 and of key pro-apoptotic proteins in AML. *Leukemia* **27**, 1358-1368, doi:10.1038/leu.2012.366 (2013).
113 Liu, Y. *et al.* NOXA genetic amplification or pharmacologic induction primes lymphoma cells to BCL2 inhibitor-induced cell death. *Proc Natl Acad Sci U S A* **115**, 12034-12039, doi:10.1073/pnas.1806928115 (2018).
114 Leshchenko, V. V. *et al.* Harnessing Noxa demethylation to overcome Bortezomib resistance in mantle cell lymphoma. *Oncotarget* **6**, 27332-27342, doi:10.18632/oncotarget.2903 (2015).

115 Jin, S. *et al.* 5-Azacitidine Induces NOXA to Prime AML Cells for Venetoclax-mediated Apoptosis. *Clin Cancer Res*, doi:10.1158/1078-0432.CCR-19-1900 (2020).
116 Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308, doi:10.1038/nprot.2013.143 (2013).
117 Eden, E., Lipson, D., Yogev, S. & Yakhini, Z. Discovering motifs in ranked lists of DNA sequences. *PLoS Comput Biol* **3**, e39, doi:10.1371/journal.pcbi.0030039 (2007).
118 Deveau, H., Garneau, J. E. & Moineau, S. CRISPR/Cas system and its role in phage-bacteria interactions. *Annu Rev Microbiol* **64**, 475-493, doi:10.1146/annurev.micro.112408.134123 (2010).
119 Brouns, S. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960-964, doi:10.1126/science.1159689 (2008).
120 Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821, doi:10.1126/science.1225829 (2012).
121 Koch, B. *et al.* Generation and validation of homozygous fluorescent knock-in cells using CRISPR-Cas9 genome editing. *Nat Protoc* **13**, 1465-1487, doi:10.1038/nprot.2018.042 (2018).
122 Trevino, A. E. & Zhang, F. Genome editing using Cas9 nickases. *Methods Enzymol* **546**, 161-174, doi:10.1016/B978-0-12-801185-0.00008-8 (2014).
123 Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823, doi:10.1126/science.1231143 (2013).
124 Szymczak-Workman, A. L., Vignali, K. M. & Vignali, D. A. Design and construction of 2A peptide-linked multicistronic vectors. *Cold Spring Harb Protoc* **2012**, 199-204, doi:10.1101/pdb.ip067876 (2012).
125 Silva, J., Chang, K., Hannon, G. J. & Rivas, F. V. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. *Oncogene* **23**, 8401-8409, doi:10.1038/sj.onc.1208176 (2004).
126 Schuster, A. *et al.* RNAi/CRISPR Screens: from a Pool to a Valid Hit. *Trends Biotechnol* **37**, 38-55, doi:10.1016/j.tibtech.2018.08.002 (2019).
127 Chen, S. *et al.* Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* **160**, 1246-1260, doi:10.1016/j.cell.2015.02.038 (2015).
128 Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet* **16**, 299-311, doi:10.1038/nrg3899 (2015).
129 Zhou, Y. *et al.* High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509**, 487-491, doi:10.1038/nature13166 (2014).
130 Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* **11**, 783-784, doi:10.1038/nmeth.3047 (2014).
131 Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184-191, doi:10.1038/nbt.3437 (2016).
132 Tarumoto, Y. *et al.* LKB1, Salt-Inducible Kinases, and MEF2C Are Linked Dependencies in Acute Myeloid Leukemia. *Mol Cell* **69**, 1017-1027 e1016, doi:10.1016/j.molcel.2018.02.011 (2018).
133 Huang, Y. H. *et al.* POU2F3 is a master regulator of a tuft cell-like variant of small cell lung cancer. *Genes Dev* **32**, 915-928, doi:10.1101/gad.314815.118 (2018).
134 Kurata, J. S. & Lin, R. J. MicroRNA-focused CRISPR-Cas9 library screen reveals fitness-associated miRNAs. *RNA* **24**, 966-981, doi:10.1261/rna.066282.118 (2018).

135 Prolo, L. M. *et al.* Targeted genomic CRISPR-Cas9 screen identifies MAP4K4 as essential for glioblastoma invasion. *Sci Rep* **9**, 14020, doi:10.1038/s41598-019-50160-w (2019).
136 Koike-Yusa, H., Li, Y., Tan, E. P., Velasco-Herrera Mdel, C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol* **32**, 267-273, doi:10.1038/nbt.2800 (2014).
137 Roidos, P. *et al.* A scalable CRISPR/Cas9-based fluorescent reporter assay to study DNA double-strand break repair choice. *Nat Commun* **11**, 4077, doi:10.1038/s41467-020-17962-3 (2020).
138 Huang, Z. *et al.* HEMD: an integrated tool of human epigenetic enzymes and chemical modulators for therapeutics. *PLoS One* **7**, e39917, doi:10.1371/journal.pone.0039917 (2012).
139 Medvedeva, Y. A. *et al.* EpiFactors: a comprehensive database of human epigenetic factors and complexes. *Database (Oxford)* **2015**, bav067, doi:10.1093/database/bav067 (2015).
140 Soltis, D. E., Soltis, P. S., Bennett, M. D. & Leitch, I. J. Evolution of genome size in the angiosperms. *Am J Bot* **90**, 1596-1603, doi:10.3732/ajb.90.11.1596 (2003).
141 van Dijk, T. B. *et al.* Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases. *Mol Cell Biol* **30**, 260-272, doi:10.1128/MCB.00645-09 (2010).
142 Shi, Y. *et al.* Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev* **15**, 1140-1151, doi:10.1101/gad.871201 (2001).
143 Dossin, F. *et al.* SPEN integrates transcriptional and epigenetic control of X-inactivation. *Nature* **578**, 455-460, doi:10.1038/s41586-020-1974-9 (2020).
144 Legare, S. *et al.* The Estrogen Receptor Cofactor SPEN Functions as a Tumor Suppressor and Candidate Biomarker of Drug Responsiveness in Hormone-Dependent Breast Cancers. *Cancer Res* **75**, 4351-4363, doi:10.1158/0008-5472.CAN-14-3475 (2015).
145 Kobayashi, A. *et al.* Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2. *Mol Cell Biol* **24**, 7130-7139, doi:10.1128/MCB.24.16.7130-7139.2004 (2004).
146 Kensler, T. W., Wakabayashi, N. & Biswal, S. Cell survival responses to environmental stresses via the Keap1-Nrf2-ARE pathway. *Annu Rev Pharmacol Toxicol* **47**, 89-116, doi:10.1146/annurev.pharmtox.46.120604.141046 (2007).
147 He, J. *et al.* KEAP1/NRF2 axis regulates H2O2-induced apoptosis of pancreatic beta-cells. *Gene* **691**, 8-17, doi:10.1016/j.gene.2018.11.100 (2019).
148 Weniger, M. A. *et al.* Treatment-induced oxidative stress and cellular antioxidant capacity determine response to bortezomib in mantle cell lymphoma. *Clin Cancer Res* **17**, 5101-5112, doi:10.1158/1078-0432.CCR-10-3367 (2011).
149 Lakshmaiah, K. C., Jacob, L. A., Aparna, S., Lokanatha, D. & Saldanha, S. C. Epigenetic therapy of cancer with histone deacetylase inhibitors. *J Cancer Res Ther* **10**, 469-478, doi:10.4103/0973-1482.137937 (2014).
150 Zhang, J. & Zhong, Q. Histone deacetylase inhibitors and cell death. *Cell Mol Life Sci* **71**, 3885-3901, doi:10.1007/s00018-014-1656-6 (2014).
151 Maurano, M. T. *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190-1195, doi:10.1126/science.1222794 (2012).

152 Wittkopp, P. J. & Kalay, G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet* **13**, 59-69, doi:10.1038/nrg3095 (2011).
153 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).
154 Melnikov, A. *et al.* Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat Biotechnol* **30**, 271-277, doi:10.1038/nbt.2137 (2012).
155 Bauer, D. E. *et al.* An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* **342**, 253-257, doi:10.1126/science.1242088 (2013).
156 Mansour, M. R. *et al.* Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science* **346**, 1373-1377, doi:10.1126/science.1259037 (2014).
157 Canver, M. C. *et al.* BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* **527**, 192-197, doi:10.1038/nature15521 (2015).
158 Korkmaz, G. *et al.* Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat Biotechnol* **34**, 192-198, doi:10.1038/nbt.3450 (2016).
159 Sanjana, N. E. *et al.* High-resolution interrogation of functional elements in the noncoding genome. *Science* **353**, 1545-1549, doi:10.1126/science.aaf7613 (2016).
160 Rajagopal, N. *et al.* High-throughput mapping of regulatory DNA. *Nat Biotechnol* **34**, 167-174, doi:10.1038/nbt.3468 (2016).
161 Zemach, A., McDaniel, I. E., Silva, P. & Zilberman, D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**, 916-919, doi:10.1126/science.1186366 (2010).