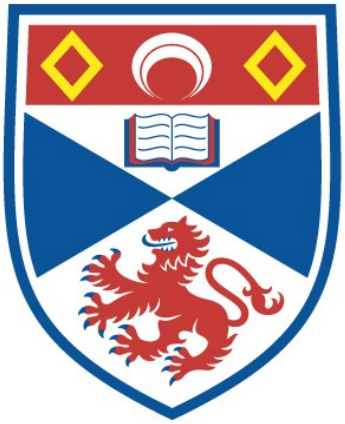
Investigating resistance mechanisms to CB-839 in renal cell carcinoma using a whole-genome CRISPR/Cas9 approach

Aleksandra Maria Raczka

A thesis submitted for the degree of PhD
at the
University of St Andrews



2021

Abstract

Renal cell carcinoma (RCC) represents around 3% of all malignancies worldwide, with its most common type, clear cell RCC (ccRCC), making up around 70-80% of all RCC. RCC is characterised by significant tumour heterogeneity and inherent (in 25-30% cases) or acquired resistance to available chemotherapeutics. Glutamine addiction is a potential new therapeutic target for RCC as it has been recently shown not only to rely on glutamine for energy generation and maintenance of redox homeostasis, but also *de novo* pyrimidine synthesis. CB-839 is a small, orally administered reversible inhibitor of human kidney-type glutaminase (GLS). CB-839 is currently in early phase clinical trials and shows promising results.

Given the significant incidence of resistance to previously approved therapies, this thesis describes a genome-scale CRISPR/Cas9 screen approach in a cell culture model of ccRCC (786-0 cell line) to identify candidate genes (hits), which when knocked down confer resistance to CB-839. Next generation sequencing data analysis of drug-selected sgRNA library representation from two timepoints was performed using the MAGeCKFlute bioinformatics workflow. To validate generated hits, single-gene knockout cell lines were created using several independent sgRNAs per gene of interest.

Knockout of FOXC1, which appeared as a strong hit, could not be achieved using multiple approaches. Additionally, knockout of RPLP2, gene, a hit in both timepoints, could not be confirmed, likely due to it being an essential gene. FOXO3 and ALDH18A1 knockout pools were confirmed using Western blotting, however, their resistance to CB-839 compared to wild type 786-0 was not significantly different. Upon further inspection of expression databases, the hits selected for validation did not show significant difference in expression or impact on survival between normal and kidney

cancer tissues. However, a CB-839 resistant cell line (10-fold more resistant) was created from 786-0 cells, suggesting that resistance factors do potentially exist.

While the screen aligned with metrics usually associated with good quality control, no candidate genes were validated when subsequently assessed individually. This study highlights the challenges in performing and validating CRISPR/Cas9 screens and describes the issues surrounding false-positive candidates.

Declaration

Candidate's declaration

I, Aleksandra Maria Raczka, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 35,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2017.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 05 May 2021 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date 05 May 2021 Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Aleksandra Maria Raczka, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

Embargo on all of print copy for a period of 3 years on the following ground(s):

- Publication would preclude future publication

Supporting statement for printed embargo request

My research is being prepared for publication and time is needed to make these journal submissions.

Electronic copy

Embargo on all of electronic copy for a period of 3 years on the following ground(s):

- Publication would preclude future publication

Supporting statement for electronic embargo request

My research is being prepared for publication and time is needed to make these journal submissions.

Title and Abstract

- I agree to the title and abstract being published.

Date 05 May 2021 Signature of candidate

Date 05 May 2021 Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Aleksandra Maria Raczka, understand that by declaring that I have original research data or digital outputs, I should make every effort in meeting the University's and research funders' requirements on the deposit and sharing of research data or research digital outputs.

Date 05 May 2021 Signature of candidate

Permission for publication of underpinning research data or digital outputs

We understand that for any original research data or digital outputs which are deposited, we are giving permission for them to be made available for use in accordance with the requirements of the University and research funders, for the time being in force.

We also understand that the title and the description will be published, and that the underpinning research data or digital outputs will be electronically accessible for use in accordance with the license specified at the point of deposit, unless exempt by award of an embargo as requested below.

The following is an agreed request by candidate and supervisor regarding the publication of underpinning research data or digital outputs:

Embargo on all of electronic files for a period of 3 years on the following ground(s):

- Publication would preclude future publication

Supporting statement for embargo request

My research is being prepared for publication and time is needed to make these journal submissions.

Date 05 May 2021 Signature of candidate

Date 05 May 2021 Signature of supervisor

Acknowledgments

General acknowledgements

I would like to thank my supervisor, Dr Paul Reynolds, for his guidance and all the help during this PhD. I would also like to thank Melville Trust for the Care and Cure of Cancer for the studentship that allowed me to conduct my research.

I would also like to extend my thanks to my 'work wife', Christie, this would not have been the same without you, my souvlaki. I also thank other PhD students and researchers within the School of Medicine for their continued support in the lab and all those cakes and banter on the sofas.

To my best friends, Ola and Patrycja, thank you for putting up with me and being the best support I could ask for despite the distance. To all the great people I have met since my first day of uni and kept in touch throughout the years, thank you. I am also very grateful for my Covid 'bubble' and late night wine on the landing. And finally, to my family, for trying so hard not to celebrate too much over seven years ago when I boarded that plane from Poland to the UK. I could not be more grateful for everyone's support over the years that have led to the submission of this PhD thesis.

Funding

This work was supported by the Melville Trust for the Care and Cure of Cancer

(grant numbers M00108.0001/TZH/JGC and M00109.0001/TZH/MHR).

Research Data/Digital Outputs access statement

Research data underpinning this thesis are available at

https://doi.org/10.17630/077598ff-ddca-4f6f-9ed0-5d2f2144c9f6.

Table of Contents

List of figures

List of tables

Abbreviations

αKG	alpha-ketoglutarate
ab	antibody
ALK-1	Activin receptor-like kinase 1
AML	acute myeloid leukaemia
ARNT	aryl hydrocarbon receptor nuclear translocator
AS	active surveillance
ATP	adenosine triphosphate
BAP1	BRCA-associated protein-1
BCA	Bicinchoninic Acid
BIRC5	baculoviral inhibitor of apoptosis repeat-containing 5 gene
BPTES	Bis-2-(5-phenylacetamido-1,3,4-thiadiazol-2-yl)ethyl sulfide
BRCA	Breast cancer susceptibility gene
Cas9	CRISPR associated protein 9
cc	clear cell
CDS	coding sequence
chsA	chalcone synthetase gene
CT	computed tomography
CTLA-4	cytotoxic T lymphocytes antigen 4
COSMIC	Catalogue of Somatic Mutations in Cancer
CRISPR	clustered regularly interspaced short palindromic repeats
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DON	6-diazo-5-oxo-l-norleucine
ds	double stranded
DTT	Dithiothreitol
ECM	extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
ENU	N-ethyl-N-nitrosourea
ESC	Embryonic stem cells

FBS	foetal bovine serum
FH	fumarate hydrase gene
FLCN	folliculin gene
GAC	glutaminase C
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GeCKO library	Genome-scale CRISPR Knock Out library
GIST	Gastrointestinal stromal tumours
GF	growth factor
GLS	glutaminase
GSH	glutathione
HCC	hepatocellular carcinoma
HDR	homology-directed repair
HIF-1α/2α	hypoxia-inducible factor 1α/2α protein
ITGA7	integrin α7
ITH	intratumour heterogeneity
KGA	kidney-type glutaminase
LOH	loss of heterozygosity
MAGeCK	Model-based Analysis of Genome-wide CRISPR/Cas9 Knockout
MOI	multiplicity of infection
mTOR	mammalian target of rapamycin protein/pathway
NCI-60	National Cancer Institute 60
NHEJ	non-homologous end joining
OS	overall survival
PAM	protospacer adjacent motif
PBRM1	Polybromo 1 gene
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline with Tween20
PCR	polymerase chain reaction
PC	pyruvate carboxylase
PDGF	platelet-derived growth factor
PDGFR-β	platelet-derived growth factor receptor β

PDH	pyruvate dehydrogenase
PD-1	programmed death receptor
PD-L1	programmed death receptor ligand 1
PFS	progression-free survival
PFT	pericyte-fibroblast transition
PN	partial nephrectomy
PTEN	phosphatase and tensin homolog protein
PTGS	post-transcriptional gene silencing
PVDF	Polyvinylidene fluoride (membrane)
qPCR	quantitative PCR
RA	radiofrequency ablation
RCC	renal cell carcinoma
RECIST	Response Evaluation Criteria in Solid Tumours
RN	radical nephrectomy
RNA	ribonucleic acid
RNAi	RNA interference
RPMI	Roswell Park Memorial Institute medium
RT-gPCR	reverse transcriptase- qPCR
SBRT	stereotactic body radiation therapy
SDS	sodium dodecyl sulphate
shRNA	short hairpin RNA
sgRNA	single guide RNA
siRNA	small interference RNA
SLC1A5	Solute Carrier Family 1, member 5
SRB	Sulforhodamine B
SRM	small renal mass
ss	single stranded
STR	short tandem repeats
T7E1	T7 endonuclease 1
TALEN	Transcription activator-like effector nucleases
TCA	trichloroacetic acid

TF	transcription factor
TGI	tumour growth inhibition
TKI	tyrosine kinase inhibitor
TME	tumour microenvironment
TNBC	triple negative breast cancer
TSC	Tuberous Sclerosis Complex gene
UTR	untranslated region
VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptor
VHL	von Hippel Lindau gene
WS	working solution
ZFN	zinc-finger nuclease

1. Introduction

1.1. Renal Cell Carcinoma

1.1.1. Overview

Renal cell carcinoma (RCC) is a collective term used for very heterogenous and highly vascularised cancers that arise from the renal tubular epithelium. It accounts for approximately 90% of all kidney cancers and is the most lethal genitourinary cancer. Incidence, prevalence and mortality rates vary across world regions and individual countries. The lifetime risk of developing RCC ranges from 1.3 to 1.8% in Europe and North America, with highest worldwide estimated incidence recorded in North America [1]. It is more commonly diagnosed in men (6th most common cancer, accounts for 5% of all cases) than in women (10th most common, 3% of all cases) [2]. Not only is RCC strongly correlated with gender, but also age, with higher incidence in older population. Worldwide, it is the cause of more than 140,000 deaths annually, making it the 13th most common cause of cancer-related deaths [1]. According to Cancer Research UK, the incidence rates have risen by more than 85% since the early 1990s. This trend can be seen worldwide. This could be partially explained by increase in abdominal imaging leading to more common incidental findings. The majority of detected lesions are small; however, locally advanced disease is still being diagnosed in a significant proportion of patients. Moreover, up to 17% of patients present with distant metastases at the time of diagnosis [1]. Although the 10-year survival in the UK has increased from 23 to 50% in the last 40 years, it remains a significant global burden.

1.1.2. Pathophysiology and genetics

There are more than ten histological subtypes of RCC, of which clear cell (ccRCC) and papillary (pRCC; type 1 and 2) are the most common. Lesser prevalent types include

chromophobe and collecting duct carcinomas. Most renal carcinomas arise due to sporadic mutations, with inherited forms accounting for 3-5% of all cases [3]. Familial RCC is usually inherited in autosomal dominant manner. Clear cell and papillary subtypes are most commonly associated with mutations in genes located on short arm of chromosome 3 and include *von Hippel Lindau (VHL)* and *c-Met* genes [4, 5]. There are ten hereditary syndromes and 12 genes which have been found to be associated with increased risk of RCC so far, but it is likely that this list is incomplete [6]. Patients suffering from von Hippel Lindau disease or Birt-Hogg-Dubé syndrome develop tumours at an earlier age than sporadic RCC patients. Familial RCC tumours are also more often bilateral and multifocal [7-9].

ccRCC alone accounts for 70-80% of all RCC cases and is characterised by clear cytoplasm due to accumulation of lipids and glycogen. The most commonly mutated gene is *VHL* gene located on chromosome 3p25 [4, 5]. Frameshift, nonsense and missense mutations, often concurrent with loss of heterozygosity (LOH) and thus causing biallelic inactivation of *VHL*, are found in up to 80% of sporadic ccRCC. Moreover, promoter methylation events, such as hypermethylation, are present in up to of 19% of patients. *VHL* encodes a tumour-suppressor protein, which plays a pivotal role in hypoxia pathway via its E3 ubiquitin ligase activity. In normoxia, functional VHL protein associates with elongin C, elongin B, cullin-2 and Rbx1 to form a complex that targets various proteins for polyubiquitination followed by degradation by proteasomes. One of the targets for this complex is hypoxia-inducible factor 1α and 2α (HIF-1α, HIF-2α) [10]. Non-functional VHL leads to stabilisation and accumulation of HIF-1α, which can interact with the aryl hydrocarbon receptor nuclear translocator (ARNT) forming a heterodimer. Such formed complex translocates to the nucleus, where HIF-1α can act as a transcription factor (TF) for hypoxia-inducible genes, such as those involved in angiogenesis. It remains elusive whether HIF-1α and HIF-2α have similar or different functions in

tumorigenesis. Alterations of genes involved in chromatin regulation are also commonly observed in ccRCC patients. Genes such as Polybromo 1 (*PBRM1*) and BRCA-associated protein-1 (*BAP1*) are also located on chromosome 3p25 and inactivating mutations in these genes are often observed in higher grade tumours. Other genetic aberrations causing ccRCC include mutations in mTOR (mammalian target of rapamycin) pathway proteins, such as PTEN (phosphatase and tensin homolog), Akt, PIK3CA and mTOR [11-13].

The second most common histological subtype is papillary RCC with incidence of 13-15%. It can grow as one tumour or multiple tumours, unilaterally or bilaterally. It is distinguished from other RCC subtypes by its papillary, tubular, or tubopapillary morphology. The cytoplasm of pRCC cells ranges from clear, basophilic to eosinophilic [14, 15]. Of the two types, type 1 is more common and less aggressive [16]. Type 1 tumours are also characterised by low-grade nuclei and a single layer of cells on top of papillae. In contrast, type 2 tumours have generally poorer prognosis (overall survival rate 89% for type 1 and 55% for type 2; disease-free survival rate 92% for type 1 and 44% for type 2), high-grade nuclei and more than one layer of cells on papillae [15, 17]. Most commonly detected pRCC are sporadic, however, inherited mutations have also been reported. Chromosome 7 and 17 trisomy as well as absence of chromosome Y in males are the main known underlying causes of sporadic pRCC [16, 18, 19]. Mutations in *MET* proto-oncogene are commonly found in hereditary and sporadic pRCC and are characteristic of type 1 pRCC [20]. Predisposition to the aggressive form of the disease is associated with hereditary leiomyomatosis and renal cell cancer syndrome [21]. Such tumours arise due to germline mutations in fumarate hydratase (*FH*) enzyme, which is involved in Krebs cycle. Accumulation of fumarate has been shown to affect oxidative stress pathways as well as its response pathways, such as the NRF2-ARE pathway [22].

Rare forms of RCC, chromophobe and collecting duct RCC, account for 4-5% and 1% of all cases, respectively. Chromophobe RCC is often found in patients with Birt-Hogg-Dubé syndrome caused by mutation in the folliculin (*FLCN*) gene [23]. These neoplasms originate in distal renal tubular epithelium of the cortical collecting duct. They are less aggressive than other RCCs and tumours can grow significantly before metastasising to other organs [24]. Collecting duct RCCs, on the other hand, are very aggressive and characterised by their non-responsiveness to drugs [25]. It remains elusive what causes this type of RCC but tumours arise from epithelial cells of Bellini's ducts of the nephron.

1.1.3. Environmental risk factors

Several environmental risk factors have been identified to increase the probability of developing RCC. Cigarette smoking is one of the most extensively studied risk factors for RCC. A meta-analysis of 24 studies concluded that smoking has a modest, yet significant impact on RCC development (overall risk ratio for the development of RCC associated with ever smoking was 1.38; 95% CI=1.27-1.50 as compared to never smokers) [26]. Moreover, smoking cessation appears to reduce the risk. In a study published in 2012, researchers at the University of Texas found that 25.4% patients with RCC had a history of diabetes mellitus, and the incidence was higher in female patients [27]. A study of 1,297 patients in Canada found that obesity plays and important role in development of RCC in both sexes [28]. Arterial hypertension has been studied for association with RCC, however it is not clear, whether it is a risk factor or a consequence of RCC [29]. Certain drugs, such as acetaminophen, or exposure to chemicals including asbestos, have also been implicated to increase the risk of RCC [30, 31].

1.1.4. Diagnosis

Classic clinical presentation of RCC includes flank pain, flank mass and haematuria. However, only 5-10% patients present with these symptoms [32]. Other symptoms are often nonspecific, such as anorexia, tiredness or unexplained weight loss. Incidental findings have been on a rise and can be attributed to the development of various imaging techniques, such as ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI) [33].

Renal tumour core biopsy is also used in the diagnostic process, especially before ablative therapies and in patients with metastatic disease [33]. Because RCCs are often asymptomatic, patients may present with distant lesions in brain, bones, ovaries, testes, lung and liver at the time of diagnosis [34].

1.1.5. Staging

TNM (tumour, node, metastasis) is often used to assess the stage of kidney tumours (Table 1.1). Alternatively, or complementary, Stages I-IV are used to describe tumours (Table 1.2).

Table 1.1: TNM staging of RCC.

Union for International Cancer Control (UICC) tumour, node, metastasis (TNM) staging of RCC, 8th edition

T—primary tumour	
TX Primary tumour cannot be assessed	
T0 No evidence of primary tumour	
T1 Tumour 7 cm or less in greatest dimension, limited to the kidney	
T1a Tumour 4 cm or less	
T1b Tumour more than 4 cm but not more than 7 cm	
T2 Tumour more than 7 cm in greatest dimension, limited to the kidney	
T2a Tumour more than 7 cm but not more than 10 cm	
T2b Tumour more than 10 cm, limited to the kidney	
T3 Tumour extends into major veins or perinephric tissues but not into the ipsilateral adrenal gland and not beyond Gerota fascia	
T3a Tumour extends into the renal vein or its segmental branches, or tumour invades the pelvicalyceal system or tumour invades perirenal and/or renal sinus fat (peripelvic) fat but not beyond Gerota fascia	
T3b Tumour extends into vena cava below diaphragm	
T3c Tumour extends into vena cava above the diaphragm or invades the wall of the vena cava	
T4 Tumour invades beyond Gerota fascia (including contiguous extension into the ipsilateral adrenal gland)	
N—regional lymph nodes	**M—distant metastasis**
NX Regional lymph nodes cannot be assessed	M0 No distant metastasis
N0 No regional lymph node metastasis	M1 Distant metastasis
N1 Metastasis in regional lymph node(s)	

Table 1.2: Pathological TNM (pTNM) classification of RCC.

(UICC, 8th edition)

Stage	pTNM pathological classification
I	T1 N0 M0
II	T2 N0 M0
III	T3 N0 M0
	T1, T2, T3 N1 M0
IV	T4 Any N M0
	Any T Any N M1

1.2. Management and treatment options

1.2.1. Active surveillance

Development of abdominal imaging techniques in the last couple of decades has increased incidental finding of localised renal masses. It is estimated that small renal masses (SRMs, <4cm in dimension) represented 48-66% of all renal masses detected between 1988 and 2003 [35]. The vast majority of SRMs has been shown to be benign and lack the ability to metastasise [36]. Moreover, SRMs are often detected in patients above the age of 65 with significant comorbidities [37]. Therefore, active surveillance (AS), which refers to initial systematic monitoring of mass size, is a common management technique in patients with asymptomatic SRMs. Patients who progress may be offered various treatment options.

1.2.2. Minimally invasive techniques

Tumour ablation, including radiofrequency ablation (RA) and cryoablation, is often a preferred treatment option for older patients with small tumours and various comorbidities [38]. Moreover, patients with familial RCC often suffer from decrease in renal function following repeated partial nephrectomies making ablation a valid choice in this setting [39].

1.2.3. Surgery

Partial nephrectomy (PN), also referred to as nephron-sparing surgery, allows for removal of tumour mass without the need to remove the normal renal parenchyma. This can be achieved during open, laparoscopic or robot-assisted surgery and the technique of choice depends on tumour characteristics as well as surgeon's expertise [40]. PN was

first described in the 19th century, but has not been widely used until around two decades ago. Increase in PN, especially for SRMs, is the effect of studies demonstrating that it has equal oncologic outcomes to radical nephrectomy (RN) in T1a tumours, with the benefit of sparing part of the kidney [41]. It is now a gold-standard treatment option for SRMs. Preservation of parenchyma is preferred because it reduces the risk of cardiovascular complications following surgery, which is often observed in patients who underwent removal of the whole organ [42, 43]. RN also carries risk of developing grade 3 chronic kidney disease. PN is oncologically effective in highly selected patients with T2 tumours, where the mass is bigger than 7 cm, but confined to kidney, and T3bN0M0 tumours [44]. However, because PN does not remove the whole organ, this can also increase the risk of local recurrence. If PN cannot be performed on SRMs or when tumour is bigger, laparoscopic RN remains a gold-standard for tumour removal [45].

1.2.4. Radiotherapy and chemotherapy

RCC has been traditionally viewed as intrinsically resistant to radiotherapy and chemotherapy, however the molecular basis of this remains elusive. Resistance was verified *in vitro*, where RCC cell lines were shown to be among the most radioresistant cell types [46]. However, radiotherapy is sometimes used in palliative setting in metastatic disease. In this setting it helps reduce pain associated with secondary tumours in brain and bones. Radioresistance of primary tumours is now debated with new techniques, such as stereotactic body radiation therapy (SBRT), showing promising results in patients with primary and metastatic RCC [47, 48].

1.2.5. Targeted therapy for advanced and metastatic RCC

Before introduction of targeted therapies, cytokines, such as interferon-α (IFN-α) and interleukin-2 (IL-2), were the standard of care for metastatic RCC for over 20 years [49]. Although a subset of patients achieved complete response, which lasted for decades, significant toxicities and lack of predictive biomarkers to best select patients, which would benefit from cytokine therapy, lead to restriction of their use over time [50]. In 2007, small molecule TKI, Sunitinib, was shown to be superior over IFN-α and became the new standard of care, starting the era of targeted therapies [51]. Schematic pathways for targeted therapies used in ccRCC are shown on Figure 1.1 [52].

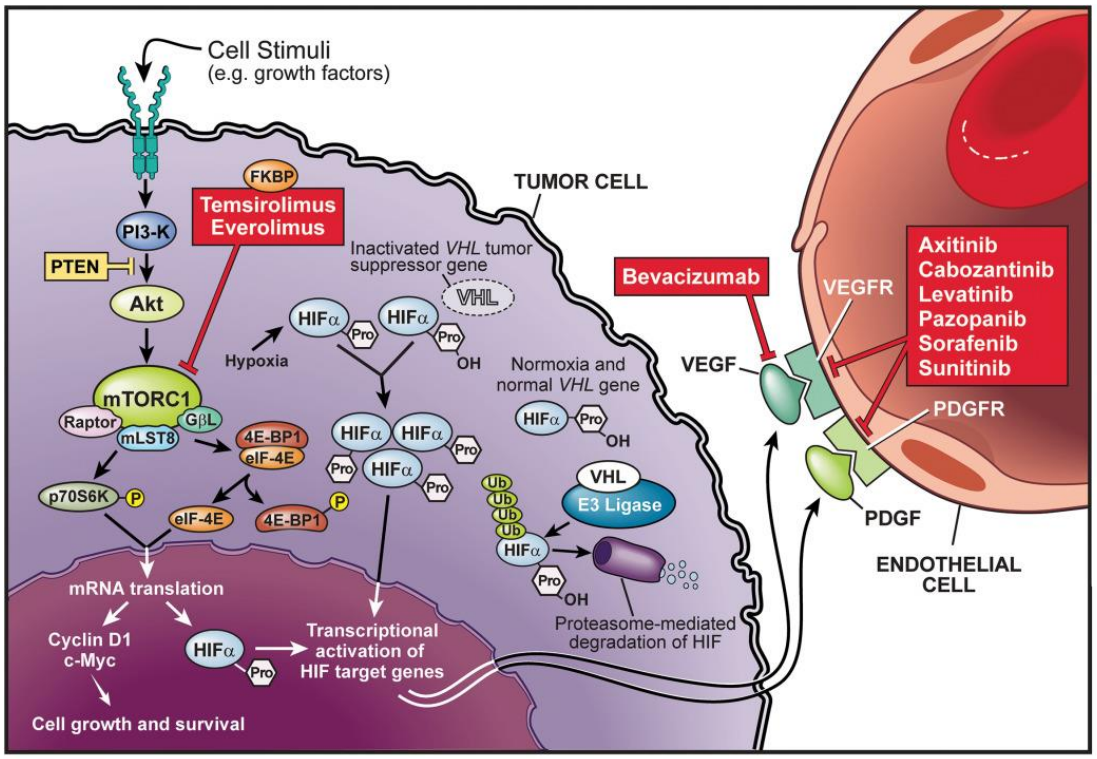


Figure 1.1: Pathways for targeted therapies in ccRCC.

Inactivation of the *VHL* gene leads to accumulation of HIF1α protein, which in turn causes overexpression of VEGF and PDGF. Additionally, HIF activity may also be increased via mTOR pathway. Understanding the molecular basis of ccRCC and the role of HIF in disease pathogenesis has led to the development of targeted therapies. These include mTOR inhibitors as well as therapies against various growth factors involved in the development and progression of ccRCC.

4E-BP1, 4E binding protein-1; AKT, protein kinase B; FKBP, forkhead binding protein; eIF-4E, eukaryotic initiation factor-4 subunit E; FGF, fibroblast growth factor; HIF, hypoxia-inducible factor; IL-8, interleukin-8; mLST8, mammalian lethal with SEC13 protein 8; mTORC1, mammalian target of rapamycin complex 1; P70S6K, P70S6 kinase; PDGFR, platelet-derived growth factor receptor; P, phosphorous; PI3K, phosphoinositide 3-kinase; Pro, proline; PTEN, phosphatase and tensin homolog; Ub, ubiquitin; VEGFR, vascular endothelial growth factor receptor; VHL, von Hippel-Lindau.

Figure taken from [52]

1.2.5.1. Anti-angiogenic therapy

Inactive VHL causes abnormal accumulation of transcription factors HIF-1α and HIF-2α in normoxia. This in turn leads to enhanced expression of various growth factors (GF) involved in angiogenesis and tumour development and progression. Affected pathways

include RAF/MEK/ERK, vascular endothelial GF (VEGF) and platelet-derived GF (PDGF) pathways.

Sunitinib, sold under the name of Sutent, is one of the first-line treatment options for advanced RCC, as well as being used to treat imatinib-resistant gastrointestinal stromal tumours (GISTs) [53]. It is an orally administered, small molecule multi-kinase inhibitor, which exerts anti-proliferative effects on tumour cells with constitutively active receptor tyrosine kinases (RTKs), including VEGFR and PDGFR. Sunitinib causes mostly reversible side effects including nausea, vomiting and bleeding events, with rarer skin toxicity and hypothyroidism [54]. Pazopanib (Votrient) has a similar mode of action to sunitinib, but was shown to have a slightly better toxicity profile [55]. Sorafenib is a small molecule inhibitor of VEGF and related receptors and also inhibits an intracellular signaling enzyme, raf kinase [56]. Axitinib is another small molecule drug, which inhibits VEGFR and PDGFR [57]. Bavacizumab is a monoclonal antibody that binds and neutralizes circulating VEGF protein [58].

1.2.5.2. mTOR inhibitors

Mammalian target of rapamycin (mTOR) is a highly conserved serine/threonine kinase that plays an important role in cell growth, proliferation and survival [59]. Its role in normal cells and in pathogenesis was possible to study thanks to the discovery of its inhibitor, rapamycin, in the soil bacteria in 1975 [60]. Rapamycin was first described to have anti-fungal properties, but it was later found to also be an immunosuppressant and thus often used to prevent rejection of transplanted organs [61, 62]. It works by diffusing into cells, where it forms a complex with FK506-binding protein (FKB-12), which then binds and inhibits mTOR in a highly specific manner, without inhibiting any other kinases [63].

mTOR pathway plays a role in various kidney diseases, such as acute kidney injury, autosomal dominant polycystic kidney disease and chronic kidney disease [64]. Phosphatase and tensin homolog (PTEN) is a tumour suppressor and has reduced activity in up to a third of RCCs leading to dysregulation of various Akt-dependent pathways, including the mTOR pathway [65] . Patients with tuberous sclerosis have increased predisposition to developing RCC due to mutation in *TSC1* or *TSC2* (Tuberous sclerosis 1 or 2) genes [66]. Loss-of-function mutations in *PTEN* and *TSC* result in upregulation and constitutive activation of mTOR which in turn leads to upregulation of HIF-1α. Everolimus and Temsirolimus are two mTOR inhibitors approved by the FDA and their mechanism of action is illustrated in (Figure 1.1).

1.2.5.3. Immunotherapy

Immune checkpoint inhibitors are the most recent examples of targeted therapies. Increasing understanding of tumour microenvironment and immunosuppression caused by tumour cells has led to the development of agents leading to increased anti-cancer immune response [67, 68]. Programmed death receptor 1/programmed death receptor ligand 1 (PD-1/PD-L1) and cytotoxic T lymphocytes antigen 4 (CTLA-4) are constituents of pathways, which are often upregulated in RCC and other tumour types resulting in dampened immune response [69]. Inhibition of PD-1 interaction with its ligand by Nivolumab was shown to enhance T-cell response [70]. Similarly, CTLA-4 inhibition with an antibody, such as Ipilimumab, promotes maturation and activation of T-cells, which are specific to tumour cells [71].

1.2.5.4. Combination therapy

Until recently, serial administration of single agents remained the standard of care for ccRCC. Even though this approach produces improved overall survival (OS), lack of long-lasting response and resistance occurring within 6-15 months remain a significant burden in mRCC patients. Combination therapy has been introduced to improve treatment efficacy in a wider patient population. Novel combinations have been made possible due to better understanding of tumour biology, including heterogeneity and resistance mechanisms.

Combination therapies including TKI plus IFN-α, mTOR inhibitor plus immune-checkpoint inhibitor or two checkpoint inhibitors, are a standard of care for first-, second- and third-line therapies (Table 1.3). Additional combination therapies, such as Axitinib (TKI) plus Evelumab (PD-L1 inhibitor), are approved by the FDA and European Commission as optional therapies. There is a wide range of combinations in clinical trials and they show promising results. This shows a switch of paradigm from single agent treatment to administration of two agents at the same time or one followed closely by another to treat advanced ccRCC.

Table 1.3: Current first-, second- and third-line treatment of ccRCC.

Adapted from ESMO Guidelines Feb 2019.

Nivo- Nivolumab, Ipi- Ipilimumab, Cabo- Cabozantinib

Drug	Target	Standard line of therapy	Optional therapy
TKIs			2nd- any TKI if 1st Nivo+Ipi; 3rd- another TKI if 1st Nivo+Ipi and 2nd TKI
Sunitinib	PDGFR, VEGFR, c-KIT, FLT3	1st- good risk	1st- intermediate, poor risk
Bevacizumab	VEGFR	1st- good	1st- intermediate
Pazopanib	PDGFR, VEGFR, c-KIT, FLT3	1st- good	1st- intermediate, poor
Axitinib	PDGFR, VEGFR, c-KIT	2nd if 1st TKI	3rd if 1st TKI and 2nd Nivo/Cabo
Cabozantinib	VEGFR, MET, RET	2nd if 1st TKI; 3rd if 1st TKI and 2nd Nivolumab /TKI	
Tivozanib	PDGFR, VEGFR, c-KIT	1st- good	1st- intermediate
mTOR inhibitors			
Everolimus	mTORC1		2nd if 1st TKI; 3rd if 1st TKI and 2nd Nivo/Cabo/TKI; 3rd if 1st Nivo+Ipi and 2nd TKI
Temsirolimus	mTOR		1st- poor
Immunotherapy			
Nivolumab	PD-1	2nd if 1st TKI; 3rd if 1st TKI and 2nd Cabozantinib /TKI	
IL-2	Cytokine therapy		1st (high dose)- good
Combination therapy			
Bevacizumab + low IFN	VEGF + cytokine therapy	1st- good	
Nivo + Ipi	PD-1 + CTLA-4	1st- intermediate, poor	
Lenvatinib+ Everolimus	VEGFR, FGFR + mTORC1		2nd if 1st TKI/Nivo+Ipi

1.3. Drug resistance

1.3.1. Tumour heterogeneity

In his 1976 publication Peter Nowell established evolutionary theory of cancer. Nowell stated that cancer is driven by stepwise process of somatic cell mutations accompanied by sequential subclonal selection [72]. Modern cancer biology further validated this model, in which cancer is governed by complex processes, Darwinian selection and adaptability [73]. For a long time now, the two types of tumour heterogeneity, intratumoral (ITH) and intertumoral heterogeneity, have been used by pathologists to describe tumours. High heterogeneity in RCC was shown almost three decades ago and is considered the primary cause of resistance to drug therapy [74].

ITH at the genetic level refers to the presence of cell clones which are genetically different and reside in different subpopulations within the same tumour. The key to different response to chemotherapeutics and chances of resistance in ccRCC is *VHL* mutation rate and chromosome copy number. To assess ITH in ccRCC patients, Gerlinger *et al* assessed tumour biopsies from four patients and reported that 63-69% mutations had not been detected in all biopsied regions [75]. In another study, of 10 biopsied patients, all showed ITH and 73-75% driver mutations were subclonal. Moreover, the only ubiquitous mutations were VHL mutations and loss of chromosome 3p [12]. In a paired primary vs metastatic tumour biopsies, Kim *et al.* reported that drug-target pathways varied between primary and metastatic sites as well as among cells within a single site [76].

Intertumour heterogeneity refers to differences in tumours between patients. Worse survival in black populations when compared to white populations used to be attributed

to racial differences in access to and quality of health care received, increased number of comorbidities. Various socioeconomic factors, such as insurance coverage, competing demands (eg. childcare) and fear of loss of income due to hospitalisation also impact the decision whether to undergo nephrectomy. However, genetic analysis showed that black populations carry lower frequencies of VHL mutations and higher frequencies of ccB subtype of ccRCC, both of which are associated with poorer prognosis [74].

1.3.2. Intrinsic and acquired resistance

Disease progression despite treatment is defined in RECIST 1.1 (Response Evaluation Criteria in Solid Tumours) guidelines as 'at least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. (Note: the appearance of one or more new lesions is also considered progression.)' [77].

Resistance to targeted therapeutics is classified into two groups; primary, or intrinsic and secondary, or acquired. Intrinsic resistance refers to lack of efficacy of therapeutic agent(s) from the start of therapy. It can be caused by the presence of inherited or evolutionary selected resistant clones ahead of therapy. Acquired resistance is defined as initial response followed by tumour growth despite patient still receiving the same therapy. Although precise resistance mechanisms in RCC remain elusive, various pre-clinical and clinical studies have identified several underlying causes (Figure 1.2) [78, 79].

1.3.3. Compensatory angiogenic pathways

Revascularisation of kidney tumours as resistance mechanism to anti-VEGFR therapy is an emerging concept with supporting evidence seen in several clinical studies [80-86]. In one such study, increased tumour size after 4 months of therapy was positively correlated with increased tumour blood flow [80]. Resistance to anti-VEGFR agents often develops while the drug is actively inhibiting its target. One of the mechanisms by which it may arise, is upregulation of VEGFR-1 and VEGFR-3 pathways, since most anti-VEGFR molecules target only VEGFR-2. Targeting multiple VEGFRs is therefore a plausible mechanism to avoid resistance [81, 82]. An example of multi-targeted TKI is axitinib, which was shown to have antitumour activity in patients who had previously developed resistance to sorafenib [83]. Other possible mechanisms involve deregulation of non-VEGFR angiogenic pathways. Increased FGRF-1 signalling was shown to positively correlate with reduced PFS in patients receiving sorafenib and was suggested to be an intrinsic resistance mechanism [84]. Involvement of activating-like kinase 1 (ALK-1), which is a pro-angiogenic receptor in TGFβ family, and angiopoietin-1/2, glycoproteins involved in angiogenesis, have also been suggested as potential compensatory mechanisms in RCC [85, 86].

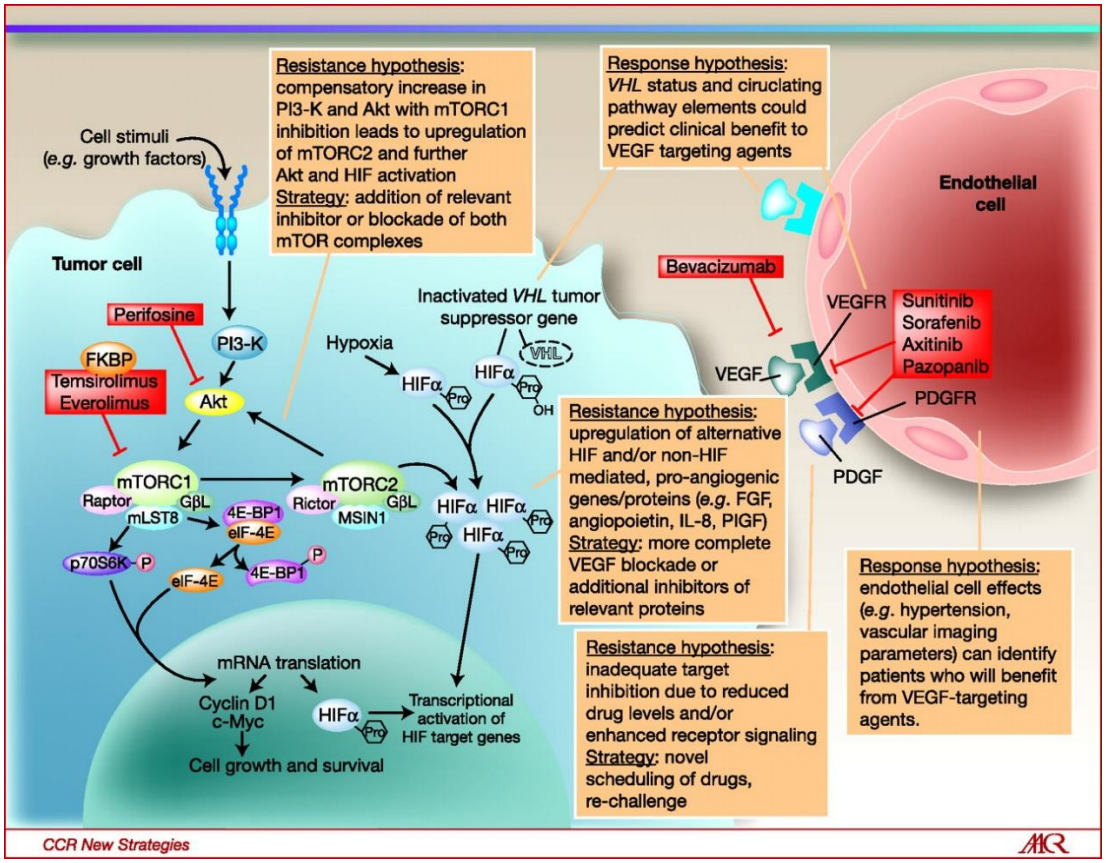


Figure 1.2: Potential mechanisms of resistance to targeted therapies in ccRCC.

Resistance to targeted therapy in metastatic RCC is thought to be related to *VHL* status and upregulation of HIF, upregulation of alternative proteins in setting of persistent VEGF or mTOR blockade, non-anti-angiogenic mechanisms, such as immunostimulatory mechanisms, re-emergence of VEGF-driven vasculature, among other mechanisms.

Figure taken from [79]

1.3.4. Compensatory mTOR signalling

Rapamycin analogues (rapalogs), temsirolimus and everolimus, are a first generation of mTOR inhibitors in RCC. Despite their effectiveness in targeting mTORC1, they are not effective inhibitors of mTORC2 [87]. mTORC2 activation is normally inhibited by a feedback loop, but if mTORC1/S6K is inhibited, the negative feedback is lost. As a result, mTORC1 becomes activated and leads to increased expression of HIF-2α via AKT. This in turn causes tumour growth [88, 89]. HIF-2α is argued to be the more significant form of HIF in RCC and its dependency on mTORC2 activation suggests that inhibition of this

pathway is a rational mechanism to prevent resistance [88]. Dual inhibitors of mTORC1 and mTORC2 are currently being evaluated in pre-clinical studies, however, they show varying results, some being inferior to rapalogs [90-92]. Other possible mechanisms of resistance to mTOR inhibitors include overexpression of survivin gene, baculoviral inhibitor of apoptosis repeat-containing 5 (*BIRC5),* augumented PI3K activity with concomitant *KRAS* mutations, elevated integrin α7 (ITGA7) expression and reduced affinity of rapalogs to FKBP12 [93-96].

1.3.5. Increased tumour invasion

One of the adaptation mechanisms to reduced formation of new blood vessels into tumour, is invasion into normal tissue to benefit from normal vasculature to sustain growth and metabolic demands [97]. Another mechanism of increased tumour invasion is upregulation of two RTKs, MET and AXL, by sunitinib [98, 99]. They play various roles in tumorigenesis in several cancers and are often overexpressed in RCC causing tumour progression, invasion and metastasis [100]. In a study by Zhou *et al*, the group administered cabozantinib to mouse xenografts previously treated with sunitinib and observed rescue of acquired resistance to sunitinib. In their *in vitro* study, addition of cabozantinib prevented pro-metastatic behaviour of cells in culture [99].

1.3.6. Tumour microenvironment

The tumour microenvironment (TME) consists of various cellular and non-cellular components surrounding tumours. It is made of immune cells, fibroblasts, blood vessels and extracellular matrix amongst others. It has been shown that TME plays a pivotal role

in every step of tumorigenesis and that the tumour cells closely interact with their TME [101].

Xian-De *et al.* reported that anti-angiogenic therapy causes increased immunosuppression in patients, whose primary tumours were treated with sunitinib. They demonstrated increased CD4+ and CD8+ T-lymphocyte infiltration of metastatic sites compared to other tissues and untreated control. This was positively correlated with infiltration of immunosuppressive Tregs and upregulation of PD-L1. As a result, those patients exhibited reduced OS (8.0% vs 22.3% ($P<0.01$)) and PFS (8.4% vs 20.3% ($P<0.01$)). This suggests that infiltrating Tregs and upregulation of PD-L1 are plausible biomarkers for anti-angiogenic therapy response [102].

Pericytes, or mural cells, are perivascular cells characterised by relatively large nucleus and small amount of cytoplasm. They wrap around blood vessels and have a supportive role preventing vessels from leaking and dilation. Moreover, it was demonstrated that pericytes retain some of the progenitor characteristics and can differentiate into other cell types, such as osteoblasts, phagocytes and macrophages [103, 104]. One of their significant features is expression of several surface markers and response to growth factors. One of the best characterised pathways that involve endothelial cells and pericytes is PDGF BB-PDGFRβ signaling pathway. PDGF-BB is expressed by endothelial cells and activates PDGFRβ, what leads to pericyte recruitment to newly formed blood vessels. Pericyte-fibroblast transition (PFT) was recently shown to be involved in vascular development and angiogenesis. PFT is a process by which pericytes, which detached from the vessel, differentiate into stromal fibroblasts, which contribute to cancer cell invasion and metastasis [105]. Increased pericyte coverage of mature vessels was shown to be an unfavourable prognostic factor for patients with ccRCC [106]. Conversely, increased coverage of immature blood vessels was positively

correlated with better overall response to sunitinib and lower occurrence of metastases [107].

1.3.7. Lysosomal sequestration of TKIs

Lysosomes are intracellular organelles and sites of lipid, nucleic acid and protein degradation. They are also involved in other cellular processes, such as apoptosis and autophagy [108]. Hydrophobicity of sunitinib allows it to easily diffuse through membranes, however, as a weak base, it becomes protonated in acidic environment of lysosomes and loses its ability to cross membranes. Gotnik *et al* showed that sunitinib accumulates in lysosomes, with tenfold higher concentrations in sunitinib-resistant cells in comparison to their sensitive counterparts [109].

1.3.8. Adaptation to hypoxia conditions

All cells can sense changing levels of oxygen, from normoxia (normal oxygen levels), through hypoxia (low oxygen levels) to anoxia (extreme hypoxia or lack of oxygen) [110]. Due to uncontrolled proliferation, cancer cells use oxygen more rapidly than healthy cells. Reduced oxygen availability is now recognised as one of the most common features of almost all solid tumours [111]. In order to sustain their rapid growth, cancer cells adapted to low oxygen and nutrient levels by production of angiogenic factors, lower consumption of oxygen as a result of metabolic shift and reduction of apoptotic potential. Adaptation to hypoxia and anoxia is believed to be one of the main mechanisms that allows for survival and selection of more invasive and therapy-resistant cells. Although the exact mechanisms of this adaptation are not fully understood, several processes have been implicated. These include, but are not limited to direct effect of hypoxia on resistance by

unavailability of oxygen, which is required by some cytotoxic agents to reach their maximum effectiveness, alteration of metabolism, thus reducing drug cytotoxicity and increased genetic instability allowing for faster development of resistant cells [112, 113].

Regulation of HIF is one of the major processes by which cells adapt to hypoxia. In hypoxic tumour cells, HIF expression can lead to HIF-induced autophagy [114]. Autophagy is an evolutionarily conserved process, which occurs in all cells at basal level and functions to recycle old or damaged cellular compartments [115]. Nutrient starvation, metabolic stress and hypoxia lead to increased autophagy, which is thought to have a cytoprotective role in cancer due to recycling of ATP and cellular components to maintain biosynthesis and cell survival. HIF-1α was shown to be involved in induction of autophagy by induction of BNIP3 and BNIP3L, which in turn interfere with formation of Bcl-Beclin1 complex [114]. Anoxia causes HIF-independent autophagy by activation of AMP kinase and thus inhibition of mTOR pathway [116].

1.4. Glutamine metabolism

1.4.1. Normal kidney and other tissues

Glutamine is the most abundant amino acid found in the human body. It accounts for 20% of all amino acids found in circulation and is involved in the largest number of metabolic processes of all amino acids [117]. Its metabolism was first described by Hans Krebs in 1935, where he noted that glutamine is found across various tissues and species, what suggested its central metabolic role [118]. Glutamine is a precursor for various molecules, such as nucleotides, where it serves as nitrogen donor in three and two enzymatic processes for purine and pyrimidine synthesis, respectively [119]. It is also important for production of amino acids involved in proliferation of immune

cells [120]. Additionally, it is a part of the pathway leading to production of glutathione, which has a protective role against oxidative stress [121]. It is viewed as conditionally essential amino acid due to its extensive use during disease by various organs [122].

Although synthesised by most tissues, skeletal muscles release the largest amounts of glutamine into plasma, followed by lungs and adipose tissue. Kidneys and gut contribute to the most glutamine uptake [120, 123]. Glutamine is an important substrate for gluconeogenesis, synthesis of glucose from non-glucose substrates, which occurs in liver and kidney [124]. These two organs are the exclusive sites of gluconeogenesis due to presence of complete set of gluconeogenic enzymes, including glucose-6-phosphate, which is absent in other organs. Studies of humans following overnight fast showed that gluconeogenesis accounts for 50% of all glucose released into circulation [125]. Moreover, gluconeogenic capacity of the kidney is significantly larger than that of the liver with approximately 25% of all glucose in the circulation being released by the kidney [126]. Kidneys also take up around 20% of all available glucose [127].

1.4.2. Cancer cells

Metabolic reprogramming is now widely recognised as one of the hallmarks of cancer [128]. Due to fast growth, tumour cells switch to alternative metabolic pathways to increase energy and nutrient production. The Warburg effect, which is enhanced glycolysis, is an important mechanism allowing cells to meet their enhanced bioenergetics demands. Unlike normal cells, cancer cells are largely dependent on glutamine, which is a non-essential amino acid. It can be equally important as glucose for survival of tumour cells [121]. It is converted into glutamate by the mitochondrial enzyme glutaminase, which was found to be upregulated in some, but not all, tumours,

which rely on glutamine for their survival [129, 130]. Moreover, glutamine levels are often found to be depleted in tumour environment caused by its uptake by cancer cells. This in turn impacts T cell growth as glutamine is an important nutrient required in this process [131].

1.4.3. Glutamine addiction in ccRCC

Glutaminase is a mitochondrial enzyme, which hydrolises glutamine to glutamate, making it a key enzyme in glutamine metabolism [132]. It has two isozymes, GLS1 and GLS2. GLS1 is expressed as two splice variants, kidney-type glutaminase (KGA) and glutaminase C (GAC), both widely expressed across all normal tissues [132]. Additionally, GLS1 has been found to be often upregulated in tumour tissues in comparison to their normal counterparts, however, this is not the case in ccRCC tumours [130, 133, 134]. GLS2 is mostly expressed in liver, brain and pancreas and is not elevated in ccRCC [132]. GLS1 and GLS2 are thought to play oncogenic or anti-oncogenic role, depending on the cancer type [130].

Despite no significant increase in glutaminase expression in ccRCC, these tumours have higher levels of glutamine and glutamate compared to normal tissue [135]. Moreover, they show increased expression of proteins involved in glutamine import into the cells, such as Solute Carrier Family 1, member 5 (SLC1A5), which is associated with poorer prognosis in ccRCC patients [136]. Addiction to glutamine seems to be a consequence of ccRCC tumours increased need for glutamate.

It was recently shown that RCC deficient of *VHL* gene not only rely on glutamine for energy generation and maintenance of redox homeostasis, but also for *de novo*

pyrimidine synthesis [137]. The group also tested response to GLS1 inhibition in two cell lines and showed that UMRC2 cell line, which has wild-type *VHL*, can adapt to these conditions by increasing oxidation of glucose in Krebs cycle. On the other hand, UMRC2 cells, which lack *VHL*, are unable to adapt because they lack compensatory mechanisms like in *VHL*+/+ cell line [137]. Stabilisation of HIF1α leads to reprogramming of cellular metabolism and increased glucose uptake due to upregulation of glucose transporters and related enzymes, such as GLUT1, hexokinase and aldolase [138]. Despite upregulation of glycolysis, increased levels of HIF1α inhibit pyruvate dehydrogenase (PDH) and pyruvate carboxylase (PC) activity, leading to reduced glucose-derived carbon to enter the TCA cycle [139]. It causes the TCA cycle to shift from using glucose to using glutamine as its main substrate, as shown by *in vitro* as well as *in vivo* experiments. Therefore, *VHL*-/- tumours rely on glycolysis for energy generation, but glutamate-derived citrate and malate for lipogenesis and oxaloacetate for synthesis of nucleotides [140-142]. Activation of HIF is both indispensable and sufficient for increased utilisation and addiction to glutamine. It achieves it by limiting glucose's ability to fuel the TCA cycle and reprogramming glutamine metabolism to serve as a substrate for production of macromolecules crucial to sustain the rapid growth and proliferation of tumour cells [141].

1.5. Glutaminase inhibitors

1.5.1. DON, BPTES, C.968

Small molecule GLS inhibitors have been studied since the 1950s and suggested to have antitumour activity [143, 144]. Early compounds included azaserine, acivicin and 6-diazo-5-oxo-l-norleucine (DON) [143, 144]. Of the above, DON was first isolated from *Streptomyces* cultures and has been the most extensively studied. Early Phase I clinical studies were carried out at low daily dosing in patients with refractory solid tumours, such as breast and lung tumours, Hodgkin's Lymphoma, testicular cancer and paediatric leukaemia [145-148]. Phase II studies, which ran from the 1980s to the early 2000s, involved high intermittent dosing of DON for patients in advanced colorectal carcinomas, sarcomas and several other solid tumours [149, 150]. Despite promising initial studies, where DON was used at low doses, use of DON in the clinic had to be halted due to its limited chemotherapeutic activity and severe toxicities observed in Phase II studies [150, 151]. The most common adverse events included gastrointestinal toxicities and inhibition of cell cycle in rapidly proliferating intestinal epithelial cells. This is thought to be due to DON's mechanism of action- it binds irreversibly not only to the active site of the GLS, but also other glutamine-utilising enzymes. Thus, DON also inhibits glutamine binding to glutamine amidotransferases, which are involved in synthesis of various crucial macromolecules, such as purines and pyrimidines, coenzymes and several amino acids [152, 153]. However, recent developments suggest that DON could potentially be used at low doses with greater success than in the past. For example, it has been successfully used as prodrug in combination with another drug in patients with paediatric acute leukaemia and several brain tumours due to its CNS activity [154]. DON has also been of interest in non-cancer-related medical settings, such as in acute pain management and showing promising results [155].

Bis-2-(5-phenylacetamido-1,2,4-thiadizol-2-yl)ethyl sulfide 3 (BPTES) is another example of GLS inhibitor. Unlike DON, it inhibits GLS in a highly specific and reversible manner. It binds to both, free enzyme and enzyme in complex with its substrate, and forms stable, yet inactive tetramer complex [156]. Its high specificity for GLS means that it is less likely than DON to inhibit other molecules, which use glutamine as their substrate. Moreover, its inhibitory potential is higher that DON (Ki in μM vs mM) [157]. Despite its ability to slow proliferation of malignant cells *in vitro* and in mouse models, BPTES shows moderate potency to inhibit GLS, low solubility and poor metabolic stability, limiting its use in the clinic [158]. More recently identified allosteric inhibitor of GLS, compound 968, inhibits glutaminolysis in transformed cells in a highly specific manner and does not impact normal cells [159]. Compound 968 has so far only been tested *in vitro* and shown to have cytotoxic effects in models of lymphoma, glioblastoma, lung cancer and breast cancer [160, 161].

1.5.2. CB-839 (Telaglenastat)

CB-839 is a small, orally administered molecule, which inhibits human glutaminase in an allosteric and reversible manner. It exerts a dual action by inhibiting tumour cells and activating T cells [162] (Figure 1.4). It was shown to downregulate mTOR signalling in RCC cell lines and RMPI8226 myeloma cells, as shown by a decrease in proteins such as phospho-mTOR and phospho-S6. The same study found a reduction in oncogenic proteins c-Myc and c-Kit and an increase in programmed cell death proteins, such as cleaved-PARP [163]. It also inhibits glutaminase in triple negative breast cancer (TNBC) cell lines, which are highly sensitive to glutaminolysis inhibition [164, 165]. CB-839 exhibits high specificity towards GLS1, demonstrated by suppression of ATP production, biosynthesis and maintenance of redox balance, in which glutamine acts as an

intermediate [164]. CB-839 also has antitumor activity *in vivo*, as shown by two studies of xenograft models of TNBC [164]. In a patient-derived xenograft model, administration of 200 mg/kg of CB-839 twice a day resulted in 61% tumour growth suppression when compared to untreated control. The HER2+ basal-like cell line JIMT-1 was used in a cell-line based xenograft model. Two treatment regimens were evaluated: CB-839 on its own or in combination with the TNBC standard therapy paclitaxel. Administration of paclitaxel alone led to tumour reduction followed by its expansion, whereas CB-839 alone or in combination with paclitaxel reduced tumour growth with no regrowth over time [164].

The activity of CB-839 was tested in a range of kidney cancer cell lines. Exposure to 1μM CB-839 for 72h resulted in varying degrees of death of ccRCC lines 786-0, 769-P and A498. CB-839 did not have anti-proliferative effect on Caki-2 and ACHN cells, which are RCC lines expressing wild-type (WT) *VHL*. Interestingly, the drug affected growth of Caki-1 cells, which express WT *VHL*. CB-839 did not seem to affect growth of six non-RCC cell lines, such as G402 and JMU-RTK2 (Figure 1.3) [166].

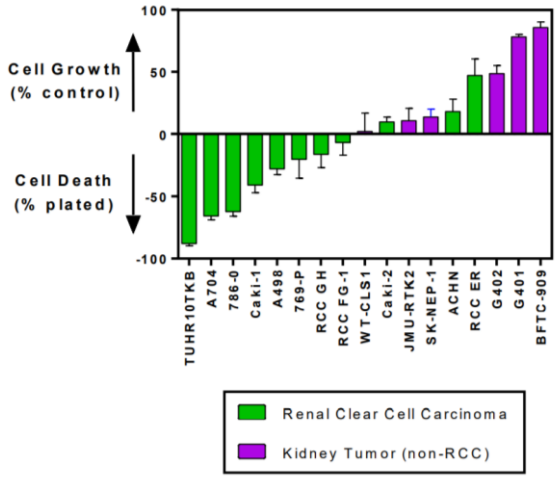


Figure 1.3: Relative cell growth or cell death across a panel of kidney tumour-derived cell lines following a 72 h treatment with CB-839 (1 µM).

Several renal cancer cell lines, including A704 and 786-0 are sensitive to treatment with 1 µM CB-839 over the period of 72 h. Other commonly used models of clear cell RCC, such as Caki-2 and ACHN are intrinsically more resistant to the same treatment conditions.

Figure taken from [166]

Safety and efficacy of CB-839 alone or in combination with Everolimus, Cabozantinib or Nivolumab is currently being evaluated in Phase I and II clinical trials (ClinicalTrials.gov Identifier: NCT02071862, NCT02771626) [167]. This multicentre, open-label, dose escalation study includes patients presenting with metastatic or locally advanced tumours. To date, the study evaluating monotherapy showed that continuous administration of the drug is characterized by an acceptable safety profile. Moreover, it significantly inhibits glutaminase and shows promising signs of clinical activity in various tumours, including RCC [168, 169]. Pre-clinical studies of Caki-1 xenografts showed that monotherapy with CB-839 or everolimus, resulted in tumour growth inhibition (TGI) of 51% and 52%, respectively. However, the combination of the two gave TGI of 85% [169]. So far it has been shown that the combination is well tolerated in ccRCC and papillary carcinoma patients and CB-839 does not increase the toxicities, compared to everolimus

alone. The most common AEs were nausea and fatigue [169]. Early pre-clinical studies using the Caki-1 cell line and Caki-1 xenografts showed superiority of cabozantinib plus CB-839 over monotherapy when measuring its anti-proliferative effect. The combination reduced signal transduction more than cabozantinib alone by reducing pERK and pAKT proteins. The two drugs also synergize to inhibit TCA cycle activity as measured by oxygen consumption rate. In xenograft models, where tumours of approximately 400 mm^3 were treated with DMSO, CB-839, cabozantinib or a combination of the drugs, monotherapy resulted in slower tumour growth, which was enhanced in combination models [170].

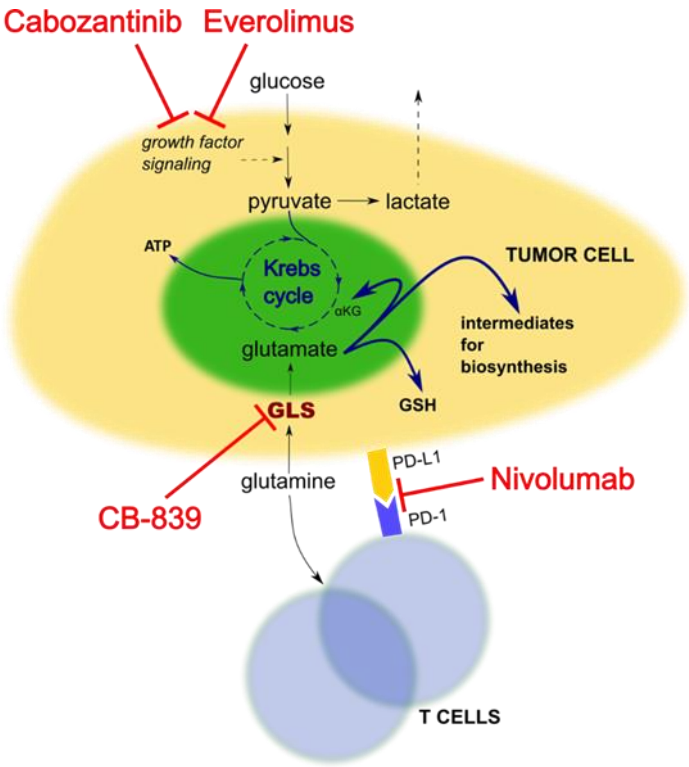


Figure 1.4: Pathways inhibited by CB-839, nivolumab, cabozantinib and everolimus.

CB-839 is a small, orally administered molecule, which inhibits human glutaminase in an allosteric and reversible manner. It exerts a dual action by inhibiting glutamine-addicted tumour cells by inhibiting GLS, first enzyme in glutamine metabolism, and activating T cells which require glutamine to exert their function. Cabozantinib and Everolimus are multi-targeted TKIs and therefore limit tumour growth. Nivolumab is an example of immunotherapy, it works by blocking the interaction between PD-1 and its ligand (PD-L1), therefore inducing PD-1 pathway-mediated inhibition of the immune response, including the anti-tumour immune response, resulting in decreased tumour growth.

GLS- glutaminase, GSH- glutathione, ATP- adenosine triphosphate, αKG- alpha-ketoglutarate, PD-1- programmed death-receptor 1, PD-L1- programmed death-ligand 1. Figure taken from [162]

1.6. Cancer drug screening

1.6.1. Cell-based drug screening

The first high-throughput cancer screening technology was developed in 1990 and consists of 60 cell lines from nine cancer types. National Cancer Institute 60 (NCI 60) platforms allows for screening across 6 leukaemia-derived cell lines, 8 melanoma, 7 colon, 6 brain, 7 ovary, 6 breast, 2 prostate and 6 kidney cancer cell lines [171, 172]. At the time when NCI60 was introduced, most anti-tumour therapies involved use of non-specific cytotoxic agents, which had high (up to 70%) response rates across multiple tumour types. Therefore, use of this limited number of cell lines was sufficient to capture such frequencies of drug response. However, over the years, as targeted therapy was becoming the gold standard of anti-cancer therapy, limitations of NCI60 platform became more apparent. Targeted therapies have lower adverse effect rates than cytotoxic agents, but their activity varies significantly between patients even with the same histological subtypes. Such low (1-10%) frequency response is not feasible to be captured by a limited number of cell lines accurately. Despite this limitation, NCI60 platform has paved way for many cell-based screening technologies and as of 2010 was used to screen approximately 100,000 pure compounds and 50,000 crude extracts [173].

Other screening platforms include JFCR-39 platform, with a subset of NCI60 cell lines and added gastric cell lines, due to gastric cancers' prevalence in Japanese population [174]. CMT1000 platform uses 1,200 cell lines and by 2009 was used to screen 127 candidate and known anticancer agents to assess their cytostatic and cytotoxic efficiencies. Moreover, it allows to detect cell line-drug pairings, which are not necessarily obvious otherwise. For example, comparison of 14 different TKIs' activity across 500 cell lines revealed unforeseen similarities between TKIs with distinct targets.

The Cancer Cell Line Project at the Sanger Institute aims at describing mutation profiles of more than 1,000 cell lines used in cancer research. The database of annotated exome sequences, copy number variants, gene expression profiles, RNASeq and CpG methylation patters is regularly updated and open to public (Catalogue of Somatic Mutations in Cancer (COSMIC) database (https://cancer.sanger.ac.uk/cell_lines/download)).

Advantages of using tumour-derived immortalised cell lines are plentiful. They provide unlimited source of various omics data, including genomics, transcriptomics and proteomics. Additionally, they are a well-controlled system with well-established culture conditions. Moreover, many phenotypes, such as cytotoxicity, growth rate, gene expression patterns and metabolism can be easily measured using well established assays. Lastly, molecular data is often widely available (eg. COSMIC).

Cell-based, or two-dimensional (2D) screening also has several limitations [175-179]. Some cells grow too slow or are difficult to grow, such as prostate cancer cell lines, therefore they are likely to be underrepresented. Culture conditions do not recapitulate *in vivo* conditions perfectly. Examples include non-physiological oxygen levels, composition of growth media and use of foetal bovine serums [179]. Moreover, due to lack of cells found in tumour microenvironment, it is impossible to measure pharmacokinetic effects as well as immune system-drug interactions [178]. Lastly, culture introduces new mutations and gene expression profiles can vary from those found in primary tissues [176, 177].

Despite these limitations, using tissue-specific cell line panels, cultured tumour-derived cells were showed to be credible genetic surrogates of tumours *in vivo* due to very similar genetic landscapes between cell lines and tumours from which they originated.

Moreover, other approaches, such as xenografts and genetically engineered mouse models are a low throughput screening methods in comparison to often high throughput of cell-based screening platforms.

1.6.1. Mutagenesis methods and cell based loss-of-function screens

Zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) use nucleases to induce DSBs, which are then repaired either by HDR or NHEJ [180, 181]. ZFNs are fusion proteins that comprise of site-specific DNA-binding proteins and are attached to the endonuclease domain of the *FokI* restriction enzyme of bacterial origin. The DNA-binding domains are based on zinc-finger domains found all over the genome and engineered to recognise 3-4 bp DNA sequence. In order to cleave the site of interest, ZFNs are designed to recognise sequences on each site of the targeted region. Binding to the DNA strands induces dimerization of the FokI nuclease, which cleaves DNA on both strands and results in 5' overhangs [180]. ZFN-mediated genome editing is not a very straightforward tool due to the significant difficulties surrounding design and assembly of highly effective zinc-finger domains [182]. Off-target effects are also prevalent [183, 184]. Similarly to ZNFs, TALENs comprise of endonuclease FokI fused to a DNA-binding domain. Here, the DNA-binding domain is made of highly conserved repeats derived from transcription activator-like effectors based on bacterial elements used to interfere with host plants' gene expression [185]. ZNFs and TALENs induce DSBs, which are repaired usually by the error-prone NHEJ. Alternatively, sister chromatid is used as a template in HDR, however, this is only possible in late S or G2 phases of the cell cycle. TALENs have several advantages over ZNFs, including feasibility of design and their repeats being longer, even to target whole genome [186, 187]. However, the issue of off-target effects remains, but this has recently been shown to be at low rates [188]. Additionally, due to their larger size, TALENs are

more difficult to deliver, with cDNA encoding it being approximately 3 kb in comparison to 1 kb for a ZFN. More examples of mutagenesis methods are listed in Table 1.4.

Table 1.4: Examples of mutagenesis methods used in cancer drug discovery.

Method	Type of mutagenesis		Reference
Zinc-finger nucleases (ZFN)	Targeted	*In vitro* and various organisms *in vivo*	[189]
Transcription-activator like effector nucleases (TALEN)	Targeted	*In vitro* and various organisms *in vivo*	[189]
PiggyBac transposon	Insertional mutagenesis	Mouse cell lines	[190]
N-ethyl-N-nitrosourea (ENU)	Point mutations	Mice and other model organisms	[191]
Gene trapping	Random mutagenesis	Mouse ES cells	[192]

RNA interference (RNAi) is a method widely used for silencing of gene expression across plethora of organisms. Similarly to CRISPR/Cas, RNAi is an endogenous mechanism, which functions to degrade foreign viral genetic material. In mammals, the defensive role of the RNAi has been overtaken by evolution of highly specialised immune system. The scientific community first came across RNAi in plants in the 1990s, when group led by R. Jorgensen attempted to upregulate chalcone synthetase (*chsA*), gene encoding enzyme involved in production of a pigment in petunia flowers. To their surprise, some of transgenic flowers lost activity of both endogenous and exogenous copies of the gene, making flowers lose pigmentation as a result of reduced mRNA levels, but not reduced gene expression. Jorgensen called this phenomenon co-suppression to account for effects on endo- and exogenous mRNA [193, 194]. Several other laboratories observed similar events in other plants and the term co-suppression was changed to post-transcriptional gene silencing (PTGS) [195, 196]. Initially only studied in plants, laboratories worldwide started reporting similar phenomena in other organisms, including fungi, such *Neurospora crassa*, where it was called quelling [197].

The mechanism of action of PTSG by RNA was first reported by Fire *et al* in 1998 [198]. The group observed that introduction of a single strand of RNA, although sufficient to induce silencing, is not as effective as double-stranded RNA (dsRNA) in a model organism, *Caenorhabditis elegans*. They noticed that dsRNA has a highly specific effect, which persisted in animals, which were injected with it, as well as their progeny [198]. The same group also showed that interference occurs when the worms are soaked in a solution containing dsRNA or fed *E.coli* expressing dsRNA [199]. Since then, PTGS, now known as RNA interference, was shown to be highly specific in many invertebrate as well as vertebrate. Over the next decade, many groups worked on deciphering exact mechanisms of RNAi and adapting it for the use as a gene editing technique in research (Figure 1.5).

Having extensively studied RNAi in model organisms and using it as a research tool, it became apparent that it can be harboured to aid cancer drug screening *in vitro*. Nowadays, high throughput screens are performed using vector-based siRNAs or short hairpin (sh)RNAs, which are available as libraries covering the whole human genome. This method of screening is based on chemically synthesised oligonucleotides and has been used to identify genes involved in various physiological processes as well as those involved in response to drug therapy [200]. Examples of *in vitro* RNAi-based screens are shown in Table 1.5.

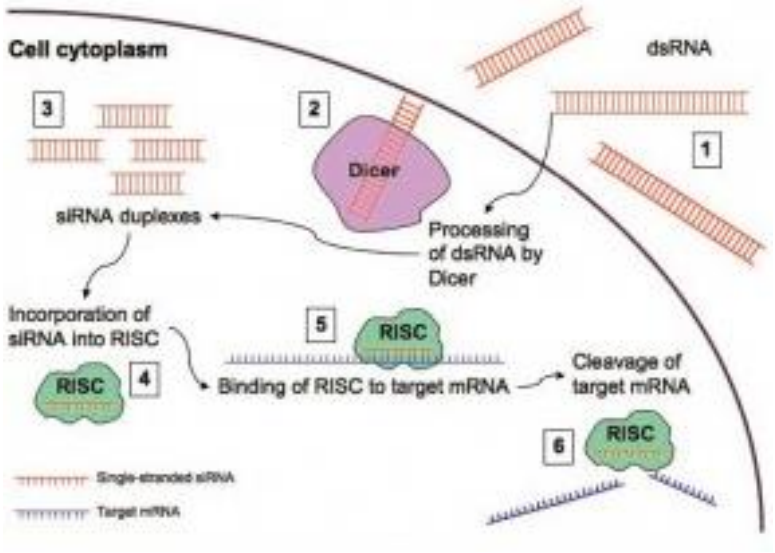


Figure 1.5: Schematic representation of the steps involved in RNA interference.

dsRNA enters cell (1) and is cleaved by enzyme Dicer (2) to produce siRNA (3). siRNA duplexes are then incorporated into multiprotein complex, RISC (RNA-induced silencing complex) (4). RISC then binds to the complementary target mRNA transcript (5). In the final step Argonaute 2, one of the proteins which makes up RISC, activates and cleaves the targer mRNA (6).

Table 1.5: Examples of RNAi-based screens in various tumours.

Tumour type	Setting	RNAi targets	Candidate genes	Reference
TNBC	Sensitivity to Paclitaxel	'Gene set from overlay of the druggable genome and a collection of genomically deregulated gene transcripts'	*PPM1D, CENPF, BCL2L1, FRAP1*	[201]
Colorectal cancer	Essential genes and synergistic combinations with Trametinib	siKinome SMARTpool library (720 genes) & custom SMARTpool siRNA library (95 genes commonly mutated in colorectal cancer)	Essential genes: *PLK1, AURKA, WEE1, SF3B;* Combinations*:* *siPINK1, siCRIM1, siPIK3CA, siPIM1, siHUNK, siBRAF, siCDKN2D*	[202]
ccRCC	Survival regulators	Whole-genome RNAi screen	*MCT4*	[203]
Lymphoma	Sensitivity to Doxorubicin	MiR-30-based shRNA library targeting 1000 gene set (≈2,300 shRNAs)	*Top2A*	[204]

The use of RNAi in high throughput screening has some major drawbacks [200, 205]. Despite early research showing high specificity, it soon became apparent that RNAi often induces off-target effects, which are classified as non-specific and sequence-specific off-target effects [205]. One of the reasons for such events is that the interferon response is triggered by the length and type of siRNAs (ss or ds), certain sequence motifs and cell type [206, 207]. Disturbance of tissue homeostasis as a result of delivery of transfection has also been shown to induce non-specific off-target effects [208]. Sequence-specific off-target effects were first described by Jackson *et al*, but only fully appreciated when RNAi became a plausible technique used in research and screening [209]. It was later discovered that the main reason for sequence-specific off-target effects are 3′ untranslated region (UTR) matches with the seed region of the antisense strand of the siRNA rather than overall sequence [210]. It is not possible to design siRNAs that do not have seed regions matching 3' UTRs found in the transcriptome, however the effects can be reduced by siRNA redundancy, siRNA pooling and chemical modifications, 2′ O-methyl ribosyl modifications [211]. However, there is currently not a perfect solution to the off-target effects of RNAi. Despite the limitations listed above, pooled shRNA screens has aided identification of essential genes in normal and cancer cells, as well as drug targets and gene mutations modulating drug response by inducing perturbations and making some cells more fit than the other to survival under specific experimental conditions. Upon completion of the experiment, DNA from the whole population of cells is isolated and quantified using NGS. Genes essential for survival will be underrepresented, whereas genes aiding survival will appear over-represented in the sequenced populations. [201-204].

1.6.2. CRISPR/Cas9 technology

The discovery of a bacterial defence mechanism against viral infection in *E.coli* in 1987 opened a new era of mammalian genome engineering and drug discovery [212, 213]. Bacteria evolved CRISPR (clustered regularly interspaced palindromic repeat)/Cas9 (CRISPR-associated protein 9) system as a memory of past infections that uses antisense RNA to avoid subsequent viral attacks in three steps [214-217]. First, spacer sequence of foreign viral or plasmid DNA is inserted into CRISPR. Then precursor CRISPR RNAs (pre-crRNAs) are transcribed and mature into mature crRNAs with repeat sequence and virus-targeting spacer sequence. Finally, Cas proteins cleave foreign nucleic acids. This occurs at sites complementary to crRNA in complex with Cas. There are three types of CRISPR/Cas systems adapted for use in mammalian cells, with type II requiring a single endonuclease and thus being the most commonly used as the other two types require ribonucleoprotein complexes. The elements of type II system (CRISPR/Cas9) together with its mechanism of action are shown on Figure 1.6.

The CRISPR/Cas9 system is versatile, simple and modular. The fact that the only requirement for target recognition is the presence of a protospacer adjacent motif (PAM), comprising of 3 nucleotides, NGG, means that almost all genes can be edited. This system has also found its application as a high throughput genetic screening tool to determine novel targets for anti-cancer drug [218]. Two screens can be performed using CRISPR libraries: loss-of-function (LoF) screen and gain-of-function (GoF) screen [219, 220]. LOF screens aim to discover the function of target genes by reducing or ablating the original genes, mRNA or proteins. In GOF screens, genes are overexpressed so more mRNA and protein is produced to observe the consequences of gains in function. Positive, as well as negative selection assays have been deployed to screen large

libraries of sgRNAs. In a positive screen, Cas-9 dependent gene modification results in engineering of resistant cells, which survive upon drug treatment, whereas cells without modification die. Negative selection works by creating sensitive cells, which die and cells without modification survive the assay [221, 222].

CRISPR/Cas 9 technology has its limitations. Several studies have showed that Cas9 can bind to sites in the genome which are not the sequences meant for cleavage and off-target events have been found at rates as high as 50% [223]. More than three mismatches between target sequences and 20 nt of sgRNA can induce off-target cleavage by the nuclease. Off-target effects can induce large deletions and genomic rearrangements or lead to lethal mutations. Off-target effects can be minimised by careful design of sgRNAs. Moreover, CRISPR/Cas9 editing can trigger a p53-mediated DNA damage response and cell cycle arrest [224]. p53 can be temporary inhibited to improve efficiency of genome editing, however it leaves cells vulnerable to the introduction of chromosomal rearrangements and other types of mutations.

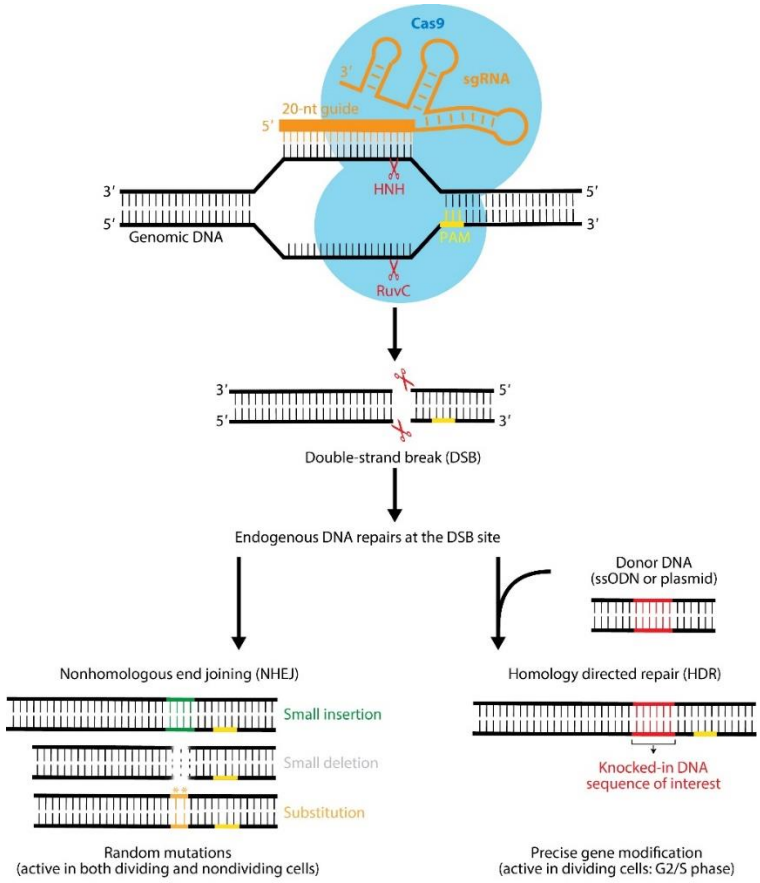


Figure 1.6: CRISPR/Cas9 system.

The system comprises of three components: Cas9 with two endonuclease domains, HNH and RuvC, which introduce double stranded (ds) breaks into DNA; crRNA with sequence complementary to target DNA sequence, and tracrRNA, trans-activating crRNA driving maturation of crRNA [225, 226]. The two RNAs can be engineered into a sgRNA (single guide RNA) [227]. To induce ds breaks in a sequence-specific manner, co-expression of Cas9 and sgRNA is sufficient in mammalian cells. sgRNA in complex with Cas9 recognises a 20 nucleotide complementary sequence, which is upstream of a protospacer adjacent motif (PAM), comprising of 3 nucleotides, NGG [214]. Binding of Cas9 to PAM is required for target recognition and later unwinding of dsDNA for crRNA to invade the target sequence [228]. Strand invasion initiates nuclease activity of HNH and RuvC domains of Cas9, which result in formation of blunt ends on DNA 3nt upstream of PAM [229]. These can be repaired by two mechanisms in mammalian cells. Non-homologous end joining (NHEJ) is an error-prone mechanism, which introduces small insertions or deletions, which cause frameshift mutations and as a result truncated and non-functional protein is expressed [230]. Alternatively, transcribed mRNA is recognised and sent to be degraded. Breaks can also be repaired by error-free homology-directed repair (HDR), but this requires the presence of a homologous DNA template [217, 228, 229].

1.6.3. *In vitro* genome-wide CRISPR/Cas9 screens

CRISPRR/Cas9 screening aims at finding a small number of genes involved in a specific physiological process within larger selection of genes (arrayed screening) or even whole genome (whole-genome screening). This method has been successfully applied to identify candidate genes that impact pluripotency, virus-host interactions and drug resistance among others [231-233]. It allows researchers to perform a large-scale loss-of-function screen with a broad hypothesis. Genes or other genetic sequences identified by CRISPR screening are referred to as 'hits' or 'candidates' and require further validation using various molecular biology methods.

In vitro delivery of sgRNAs and Cas9 can be achieved using a one-plasmid or a two-plasmid system. Plasmids are delivered to recipient cells using lentiviral vectors with optimised viral titre (low multiplicity of infection MOI=0.3) to ensure that each cells receives a single construct. Successfully infected cells are then selected, for example by using puromycin, due to puromycin-resistance gene also being encoded on plasmids. Next steps include collection of baseline DNA before treatment with a drug and subsequent drug treatment of remaining cells for a period of time. DNA is collected at various time points depending on the experimental design, purified, amplified and sequenced to determine what sgRNAs are present in surviving cells [221]. Data is then processed using MAGeCK (Model-based Analysis of Genome-wide CRISPR/Cas9 Knockout), current analysis standard, or other analysis methods, such as BAGEL (Bayesian Analysis of Gene EssentiaLity) [234, 235].

The efficacy of various sgRNAs targeting the same gene is not the same. Various technical and biological factors influence sgRNA efficacy. Therefore, researchers

developed estimation algorithms, which help to design the most efficient sgRNAs [236]. However, predictive ability of such algorithms remain limited and it is recommended to use five or more sgRNAs per gene. Some aspects of sgRNA design include favourable selection of cytosine over thymine in PAM sequence, adenoside preferred in the middle of the sequence, cytosine disfavoured at position 3 and guanine favoured over cytosine at position 20 [236].

Results of one of the first successful screen results were published in 2014 by the researchers at the Broad Institute [221]. They used pooled libraries consisting of sgRNAs targeting 18,080 genes found in human genome. Not only did they find genes which had been previously known to confirm resistance to vemurafenib in melanoma, but also identified novel hits. They also observed that knock out efficiency of individual sgRNAs targeting the same gene, was very consistent [221]. Other examples of successful application of CRISPR screening to identify genes that confer resistance to chemotherapeutics involve resistance to imatinib in GIST, sorafenib in HCC and interferon in melanoma [237-239] (Table 1.6). Successful implementation of this methodology also includes identification of human pluripotency-specific genes, genes essential for oxidative phosphorylation and potential antiviral target for Ebola virus [231, 240, 241].

Table 1.6: Examples of CRISPR screens using GeCKO(v2) library.

GIST- gastrointestinal stromal tumours; HCC- hepatocellular carcinoma, AML- acute myeloid leukemia; HNSCC- head and neck squamous cell carcinoma

Tumour type	Inhibitor	Library used	Candidate genes	Reference
Melanoma	Vemurafenib	GeCKO	Previously known: *NF1, MED12*, novel hits: *NF2, CUL3, TADA2B, TADA1*	[221]
GIST	Imatinib	GeCKO	*TP53, SOC6, DBP, NR3C1, TCF12, ZNF12, ZFP36, ACYP1, DRD1*	[237]
HCC	Sorafenib	GeCKOv2A	*PHGDH*	[238]
Melanoma	Interferon	GeCKOv2	*JAK1*	[239]
AML	Quizartinib	GeCKO	*SPRY3, GSK3*	[242]
HNSCC	radiotherapy	GeCKOv2	*STING*	[243]

1.6.4. GeCKOv2 library

GeCKOv2 is a human CRISPR whole-genome knockout library created by researchers at the Broad Institute. It is an improved version of their previous library, GeCKO [221, 244]. Modification were made to obtain approximately tenfold higher-titer virus (lentiCRISPRv2, previously lentiCRISPRv1) include ‘removal of one of the nuclear localization signals, human-codon optimization of the remaining nuclear localization signal and P2A bicistronic linker sequences, and repositioning of the U6-driven sgRNA cassette’ [244]. The improved library also has lower levels of potential off-targets and additionally targets approximately 1,000 genes, not targeted by sgRNAs in GeCKO, bringing the total to 19,050 targets. The library consists of two half-libraries, A and B, each containing three sgRNAs targeting each gene. These half libraries can be used individually or be combined to create library with six sgRNAs per gene. Additionally, GeCKOv2 library contains 1,000 non-targeting sgRNAs, which act as control and 7,456 sgRNAs, which target 1,864 miRNAs (4 sgRNAs per miRNA). It is available as a one vector system, where singe sgRNA, Cas9 and puromycin-resistance gene are encoded

on a single backbone, or two-vector system, where sgRNA and puromycin-resistance gene are separate to Cas9-containing vector. In addition to human library, GeCKOv2 is also available as mouse knockout library [221, 244]. Specifications of human GeCKOv2 library can be found in Table 1.7.

Table 1.7: Details of human GeCKOv2 library.

Genes targeted	19,050 (6 sgRNAs per gene)
miRNA targeted	1,864 (4 sgRNAs per miRNA)
Control (non-targeting) sgRNAs	1,000
Viral plasmid vector	Single and dual vector: lentiCRISPRv2 and lentiGuide-Puro
Total sgRNA constructs	123,411

1.6.5. Next generation sequencing (NGS)

Sequencing of the human genome has made a tremendous contribution to our understanding of the evolution, origins of disease and impact of the interplay between genes and environment on heredity and complex diseases. Completion of the Human Genome Project would not have been possible if it had not been for a sequencing method developed by Frederic Sanger and colleagues in 1977 [245]. This method, commonly referred to as Sanger sequencing, was first used to determine nucleotide sequence of bacteriophage ϕX174. Over the next 40 years, Sanger sequencing was the most widely used sequencing method. Various improvements of this method and automation of the process allowed for sequencing of whole genomes of various organisms, including *Haemophilus influenzae* and *Drosophila melanogaster* [246, 247]. However, up until the late 20th century, available technologies were not advanced enough to make significant progress towards completion decoding of the human genome. It was becoming clear that the human genome is extremely complex. The first full drafts of the human genome were

delivered independently by Celera Genomics and International Human Genome Sequencing Consortium in 2001 [248, 249]. It took almost a year of work and over two decades since Sanger first published his method.

Today, entire human genome can be sequenced within one day. Moreover, its cost has dropped approximately 50,000-times by mid-2000s in comparison to early 2000s. As of 2016, Veritas Genomics quoted the price of to be USD$1,000. Such significant reduction of time and costs were possible thanks to Next Generation Sequencing (NGS), also known as massively parallel or deep sequencing. It is a collective term for technologies that allow to sequence multiple long (much bigger that 1 million base pairs) DNA or RNA fragments simultaneously. Top NGS companies include Illumina, Ion Torrent, PacBio, Oxford Nanopore Technologies and BGI Genomics. They use different strategies, but have in common three general steps: library preparation, amplification and sequencing (Figure 1.7). NGS has allowed to decode sequences of human mitochondrial DNA, epigenome, transcriptome and exome. In addition to being a tool used in research, NGS is also used in clinic, for example for diagnosis or to guide decisions on treatment by studying small mutations, such as small nucleotide polymorphisms and larger aberrations, such as copy number variation.

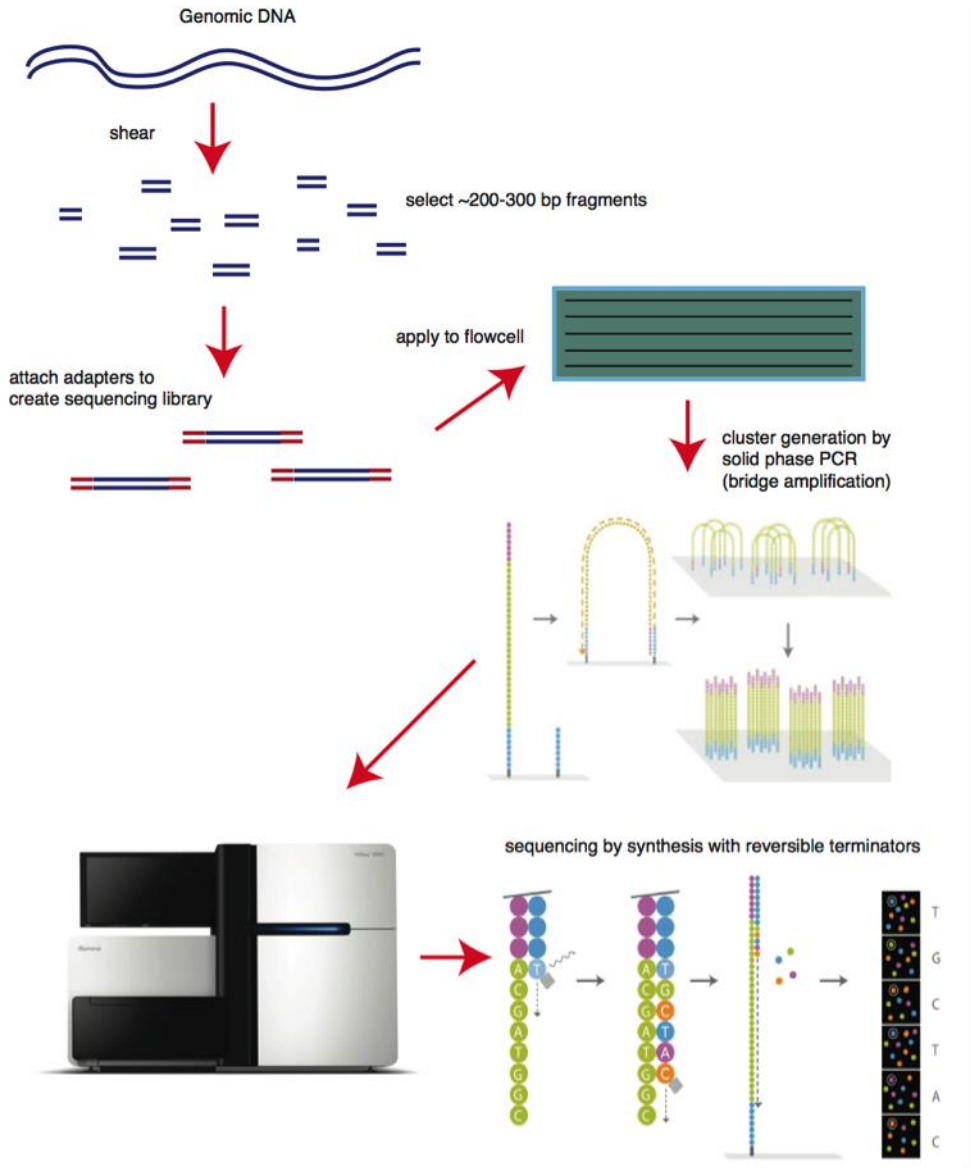


Figure 1.7: Overview of sequencing by synthesis by Illumina.

(Adapted from www.bitesizebio.com)

Sequencing by synthesis is a widely used NGS technology. The process comprises of four steps: library preparation, cluster generation, sequencing and alignment and data analysis. During library preparation, genomic DNA is fragmented, and adapters are added to both ends of each fragment. Such prepared library is then applied to the flow cell and each fragment is hybridised to the surface, following by fragment amplification using bridge amplification. Such formed clonal clusters are then sequenced using fluorescently labelled nucleotides. Each base is identified using emission wavelength and intensity. This is repeated as many times as there are bases on the flow cell. Reads are then aligned to the reference sequence, therefore making it possible to identify differences between the two using bioinformatic software.

1.6.6. Data analysis using MAGECK algorithm

Data analysis is perhaps the most challenging aspect of CRISPR screens. Studies are often carried out with few replicates and therefore it is crucial to employ a proper statistical model to detect statistically significant differences between treatment and control samples. It is important that the selected data analysis model takes into consideration that different sgRNAs targeting the same gene often have varying efficiencies. Robust normalisation methods are also very important to minimise false-positive and false-negative hits.

Currently there is little guidance and consensus when it comes to analysis of pooled CRISPR screens [250]. There are several algorithms, which can be useful in this setting. For example, there are several algorithms commonly used to analyse differential RNASeq expression. EdgeR, DESeq and baySeq are able to evaluate statistical significance of hits in CRISPR screens, but only at the sgRNA level [251-253]. The RNAi Gene Enrichment Ranking (RIGER) and Redundant siRNA Activity (RSA) algorithms were originally designed for siRNA or shRNA screens and can also be used in CRISPR screens [254, 255].

One of the algorithms designed specifically to analyse data from CRISPR screens is MAGeCK [234, 256]. It works by using median-normalisation of read counts from different samples to account for library size read count distribution. It then uses negative binomial model to establish whether sgRNA abundance is significantly different between samples. sgRNAs are then ranked using p-values. Positively or negatively selected genes are then identified using either modified robust rank aggregation (RRA) or maximum likelihood estimation (MLE) algorithms. MLE method is preferred to RRE in

presence of multiple screen conditions, whereas RRE can only compare samples between two conditions [256]. MAGeCK was shown to have high sensitivity and lower false discovery rate when compared to edgeR, DESeq, baySeq, RIGER and RSA algorithms [234].

1.6.7. Other approaches

1.6.7.1. Screens in 3D organoids

Currently, testing of a compound for its utility in oncology, starts with 2D cultures, followed by animal testing and clinical studies. However, many drugs that enter clinical trials, fail Phase III clinical trials, main reasons being lack of efficacy or unaccepted toxicities [257]. Failure of a compound in Phase III trial not only is associated with high cost of that particular stage of testing, but it also carries a collective cost of all previous testing stages.

One of the drawbacks of 2D cultures in cancer studies is the lack of the extracellular matrix (ECM) which has long been known to play a key role in tumour initiation, progression and metastasis [258, 259]. To overcome this limitation, 3D culture systems have been developed. Such cultures comprise of tumour cells, which form aggregates, or organoids, on, inside the matrix or in suspension. This allows to mimic more closely conditions that cells grow in in human body, including cells at various stages of development, rapidly proliferating cells, apoptotic cells and necrotic cells among others [260]. 3D systems also allow for co-culture of tumour cells with other types of tumour microenvironment cells, such as fibroblasts [261].

Despite higher complexity and closer recapitulation of *in vivo* tumour and its environment, 3D cultures have several limitations when it comes to their utility in high-throughput drug

screening. The use of immortalized cells is preferred in both 2D and 3D cultures due to their ease of culture. However, this also means that those cells are not very heterogenous and may not represent the tumour accurately. Co-culture of cells also creates several aspects to consider, including the ratio of various cell types and impact of co-culture on growth rates. Moreover, co-culture introduces higher complexity, which is not desirable in high-throughput screening [262]. Similarly to cells cultured in a monolayer, 3D cultures also require growth media, which alter gene expression profiles. Last but not least, 3D culture models still lack vasculature of tumours, which prevents recapitulation of oxygenation, nutrient absorption and metastasis [263].

1.6.7.2. *In vivo* screens

To date, *in vivo* CRISPR/Cas9 screens have mostly focused on identifying genes involved in cancer initiation and progression in mouse models [264]. They aided identification of novel tumour suppressors, oncogenes, regulators of immunotherapy response and synthetic lethal genes [265-269]. *In vivo* screening is still in its infancy and the main outstanding questions are whether drug response in the engineered mouse mimics that of genetically-matched patient and whether it is possible to faithfully mimic complex interactions in mouse models so as they are useful for prediction of response in the clinic [270].

1.6.7.3. Alternative approaches to genomic screens

Changes in gene expression in tumours when compared to normal tissues, can lead to significant changes in metabolome [271]. The metabolome of a cell is made up of complex biochemical pathways and numerous types of molecules, including amino acids, sugars and lipids among others. Comparison of metabolomic profiles between

cancerous and healthy tissues can help identify changes that promote carcinogenesis. For example, recent advancements in this field have led to the better understanding of predictive risk of developing and resistance mechanisms in breast cancer due to the dysregulated glucose metabolism, fatty acid synthesis and glutaminolysis [272, 273].

Another study, which focused on resistance to glutaminolysis inhibition in RCC showed that accelerated lipid catabolism and autophagy are cancer survival mechanisms [274]. Here Halama *et al*. used compound C.968 to inhibit glutamine metabolism and found that it suppressed cell proliferation but did not cause cell death. They found that accelerated lipid catabolism was activated as a compensatory mechanism to glutaminolysis inhibition together with oxidative stress, which then triggered autophagy. Moreover, dual inhibition of glutaminolysis and beta oxidation or autophagy led to cell death. They used chloroquine to inhibit autophagy and since it is an FDA-approved drug in other settings, they suggested that combination of glutaminolysis-inhibiting drugs and chloroquine can be a viable option for patients presenting with glutamine-addicted tumours.

Pancreatic duct adenocarcinoma (PDAC) is another example of a tumour with altered metabolic profile. *In vitro* studies of GLS inhibition in pancreatic cancer show promising results, however, compensatory metabolic pathways allow cells to survive *in vitro* and *in vivo*. Biancur *et al*. used both metabolomics and proteomic approaches to study this adaptation [275]. They showed that mice treated with CB-839 had no significant improvement in survival and that they actually developed secondary pancreatic tumours. They used uniformly labelled 13C-labelled Gln tracing to identify the itinerary of Gln-derived carbons in control versus CB-839 treated conditions. They found that labelled Glu species decreased but that there was an increase of unlabelled species as well as several downstream metabolites suggesting activation of alternative pathways. The

group also performed metabolomics profiling of PDAC cells and found enrichment in oxidation of branched chain fatty acid metabolism. This was also true for CB-839-treated tumours. To further understand these adaptations, the group also performed proteomic analysis and identified upregulation of proteins involved in oxidative stress response, fatty acid and lipid metabolism, glycolysis, amino acid metabolism and lysosomal processes.

1.6.8. Drug-resistant cancer cell lines

One of the well-established approaches to study drug resistance *in vitro* is development of chemotherapy and targeted therapy drug-resistant cell lines. This approach has been used since as early as 1970s when Biedler and Riehm described Chinese hamster cells resistant to Actinomycin D with cross-resistance to several other drugs, including vinblastine and vincristine [276]. They used a stepwise dose escalation approach, which resulted in 2500-fold more resistant cells when compared to parental cells. Resistant models have also been made in *in vivo* mouse models and include resistance to methotrexate, 8-azaguanine and terephthalanilide [277].

Establishment of a drug-resistant cell line usually takes 3-18 months and there is no specific protocol, although two approaches are generally followed. First, called pulse-selection method, involves relatively low drug concentration recovery of cells in drug-free media between treatments [278]. Examples of cells created using this method are platinum or taxane-resistant lung cancer cell lines (H1299, H460) and taemozaolamide-resistant melanoma cell line (Malme-3M) [279]. Second approach employs continuous treatment with the drug and escalation of doses over time. On of the cell lines created using this method is lapatinib-resistant breast cancer line (HCC1954).

Fold resistance is calculated by dividing IC50 value of the resistant ell line by IC50 of parental cell line. It is crucial to determine a clinically relevant level of resistance. It is generally done by comparison of cell lines established from cancer patients before and after chemotherapy. Most established post-treatment cell lines show two to five-fold increase in resistance to the drug that a patient had been treated with based on comparison of IC50 values. Some cell lines showed higher, 8-12-fold increase in resistance. Many clinically-relevant models use pulsed treatment method to mimic treatment patients receive in the clinic. However, this approach has its disadvantages, mainly potential of creating unstable resistance and very low levels of resistance which are hard to study. Laboratory models often use higher doses which are escalated over time and continues treatment with the agent. In this case, levels of resistance are often higher than in clinical models and therefore molecular aberrations associated with the resistance are easier to study [279].

1.6.9. Validation studies and technologies

Validation of hits generated from the whole-genome screen is a crucial step in order to confirm that a knockout of a given gene is in fact responsible for the observed phenotype. First step is often an assay similar to the screen prior to any other downstream steps. This consists of creating CRISPR knockout cell pools, where all cells harbour the same gene knockout.

Following transfection knockout is validated using various techniques, such as qPCR to detect transcript level changes or Western blotting to confirm that the protein expression is abolished. However, Western blotting requires specific antibodies and depending on the target, it may be difficult to find good quality ones. The use of downstream gene

ontology or biological pathway analyses can complement target validation and provide biological context to hits. Another option is confirmation of Cas-dependent modifications using T7 endonuclease 1 (T7E1), which is a structure-selective enzyme that detects structural deformities in heteroduplexed DNA [280]. A less commonly and more costly approach is fragment analysis via automated capillary electrophoresis (ACE) [281]. This approach allows to detect even single nucleotide changes. High resolution melting (HRM) genotyping is another alternative to detect variations in nucleic acid sequences [282].

It is important to remember that not all cells in the pool will harbour the desired mutation and that in some cells gene ablation will be incomplete. If appropriate template is available, cells may use HDR to repair lesions and therefore regain functional copies of the gene. Therefore it is beneficial to create single-cell clones with confirmed knockout. However, this process is lengthy and laborious. Another reason for validation of hits is the possibility that the observed phenotype resulting from knockout appears only in the cell line the original screen was performed in. Therefore, it is important to replicate a result in multiple cell lines.

Following hit validation *in vitro* using primary screen method and potentially another method, such as RNAi, it may be useful to extend into models of higher complexity, such as organoids or mouse models. This is particularly important in screens aimed at the identification of resistance or sensitisation mechanisms for drugs.

1.7. Thesis aims

Resistance to therapies is a common occurrence in cancer treatment, including ccRCC treatment. The overall aim of this thesis was to study genes which are potentially involved in resistance to a novel, yet to be approved for the use in clinic, glutaminase inhibitor CB-839. The technology used to achieve this goal was a genome-scale CRISPR/Cas9 screen.

The first aim was to optimise conditions of the whole-genome CRISPR screen in a selected human RCC cell line. Parameters requiring optimisation included 786-0 cell line selection and seeding density, puromycin concentration for the selection of transfected cells, PCR cycling conditions of isolated genomic DNA and selection of quality control methods. Optimisation of various conditions at a smaller scale is a crucial step ahead of performing a whole-genome screen. Failure to do so might result in poor quality data. All steps involved in the optimisation process are described in Chapter 3.

The second aim was to perform the whole-genome CRISPR screen using parameters established previously. Chapter 4 describes all experimental procedures performed in order to study what genes when knocked out using the GeCKOv2 library, confer resistance to CB-839 in 786-0 cells.

The third and final aim was to validate candidate genes obtained from the screen. This was done by creating 786-0 cell pools with single gene knockouts to study the effect of loss of function of this particular gene. Several genes were selected for independent validation based on data analysis using MAGeCK algorithm. Experimental proceedings regarding hit validation are described in Chapter 5.

2. Materials and methods

2.1. Cell culture

2.1.1. Recovery and freezing cells

In order to culture previously frozen cells, 1ml of cells in freeze media (90% FBS, 10% DMSO) was recovered from cell archive, where cells were kept in liquid nitrogen. They were then quickly thawed in water bath at 37°C. Cell suspension was then transferred to a 50 ml tube containing 9 ml of appropriate media and centrifuged in the Heraeus LAbofuge 400 R centrifuge (Thermofisher Scientific) at 1,200 rpm for 5 minutes at room temperature. Supernatant was then discarded and cell pellet re-suspended in 10 ml of media and transferred into a new T75 or T175 flask with 15-40 ml media. Cells were then placed in Heracell™ 150i incubator at 37°C and 5% CO_2. Subsequent freezing of cells was performed by re-suspending cell pellet in freeze media at 1-2 x 10^6 cells/ml in cryovial and stored in -80°C, followed by long-term storage in liquid nitrogen.

2.1.2. Cell lines and culture conditions

786-0 cells were cultured in RPMI 1640, GlutaMAX™ medium supplemented with 10% (v/v) FBS and 1% (v/v) Pen/Strep. HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Gibco™) supplemented with 10% (v/v) FBS and 1% (v/v) Pen/Strep. All cells were cultured in Heracell™ 150i incubator at 37°C and 5% CO_2. 786-0 cell line was authenticated using short tandem repeats (STR) and compared against STR profiles from Cellosaurus website (Supplementary Data 1). All cells were regularly checked for Mycoplasma contamination.

2.1.3. Sub-culturing of cells

Cells were sub-cultured at 80-90% confluence unless stated otherwise, by removing media, washing with phosphate buffered saline (PBS) and adding 5 ml of TrypLE Express reagent (Gibco™). Flasks were then placed back into an incubator and after 5-10 min cells were examined under light microscope for detachment from the flask. 5 ml of appropriate media was then added to the flask, cell suspension transferred to a 50 ml tube and centrifuged at 1,200 rpm for 5 minutes at room temperature (Thermo Scientific Heraeus Labofuge 400 R). Supernatant was then removed and cell pellet re-suspended in fresh media, counted using Neubauer Haemocytometer and the required number of cells was plated in an appropriate cell culture dish for subsequent experiments.

2.1.4. Preparation of drug solutions

5 mg of CB-839 (Cambridge Bioscience) was dissolved in 8.748 ml DMSO to give a final concentration of 1 mM. The solution was then sterile-filtered using 0.22 µm filter and aliquoted into sterile 1.5 ml Eppendorf tubes. 100 µg staurosporine (Cambridge Bioscience) was dissolved in 214 µl DMSO to a final concentration of 1 mM. 25 mg puromycin dihydrochloride (Sigma) was dissolved in dH_2O to a final concentration of 10 mg/ml, sterile filtered as above and aliquoted. All drug solutions were stored in -20°C freezer.

2.1.5. Cytotoxicity assay

2.1.5.1. Cell treatment

Cells were seeded in 200 µl of RPMI (see section 2.1.2.) at constant density in 96-well plates, with perimeter wells filled with 200 µl of PBS only. Plates were then incubated overnight to allow for cell attachment. The following day, media was removed from all wells and replaced in fresh media (control) or media with increasing drug concentration (0-1 µM) in replicates of 6. Cells were then incubated for 72 h with CB-839. In order to create a CB-839-resistant cell line (786-0-R), 786-0 cells were continuously treated with 0.7 µM of CB-839.

2.1.5.2. Sulforhodamine B (SRB) assay

After treatment time elapsed, cells were fixed by adding 50 µl of 25% trichloroacetic acid (TCA) (Sigma, T4885-500g) to each well containing media. Plates were then left at 4°C and after 60 min all solution was discarded. Each plate was then washed ten times with tap water and dried overnight in the oven at 50°C. The following day, 50 µl of 0.4% (w/v) SRB solution (Sigma) was added and evenly distributed in each well. Plates were left for 30 min at room temperature. Plates were then washed 4 times with 1% v/v glacial acetic acid and drying step was repeated as above (or for a minimum of 3 hours). 150 µl of 10 mM Tris buffer (pH 10.5) was then added to each well and left on the rocker for 1 hour at room temperature to dissolve SRB dye.

2.1.5.3. Cell survival analysis

Absorbance of dissolved SRB dye at 540 nm was measured using Biohit BP800 Platereader as a surrogate of cell density.

2.2. Protein detection and analysis

2.2.1. Preparation of protein lysates

Cells were seeded onto 10 or 15 cm tissue culture dishes and incubated until 80-90% confluent or the required period of time for experiments requiring collection of lysates at specific time points. Growth medium was then discarded and dishes washed twice with 10-25 ml ice cold PBS. For apoptosis markers, floating cells were also collected, centrifuged and washed with ice cold PBS. Attached cells were then lysed using cell lysis buffer (New England Biolabs; cat no. 9803). 10x buffer (Table 2.1) was supplemented with 100 µl PhosSTOP™ phosphatase inhibitor cocktail (Roche; cat no. 04906845001) and cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail tablet (Roche; cat no. 11836170001) and diluted 1 in 10 with deionised water. Depending on dish size and confluency of cells, 50-200 µl of buffer was added to the dish and cells were scraped using Corning® cell lifter (Sigma, cat no. CLS3008). Dish was then sat on ice at an approximately 45 degree angle so that the cell suspension collected at the bottom of the dish and left for 10-15 min for cells to lyse. Lysates were then collected into sterile Eppendorf tubes and sonicated for 2 min (in 10 sec on, 20 sec off cycles) (Qsonica, Q125 Sonicator). Tubes were then centrifuged in a pre-cooled Hereaus Fresco microcentrifuge (13,000 x g for 6 min at 4ºC). Supernatant was then transferred to a new Eppendorf and kept in -80°C until ready to use for Western Blotting.

2.2.2. Isolation of the mitochondrial fraction (ALDH18A1)

Cells were cultured and protein isolated as described in Section 2.2.1. and mitochondrial fraction isolated using Mitochondria Isolation Kit for Cultured Cells (Thermo Scientific™, cat.no. 89874) as per manufacturer's protocol. Briefly, 800 µl of Reagent A was added to the sonicated sample and after 2 min of incubation on ice, 10 µl of Reagent B was added and tubes were kept on ice for 5 min, vortexing every minute. 800 µl of Reagent C was then added to the tubes and centrifuged in a pre-cooled Hereaus Fresco microcentrifuge (700 x g for 10 min at 4ºC). supernatant was then collected into new tubes and centrifuged again (12,000 x g for 15 min at 4ºC). Supernatant was then removed and pellet (mitochondrial fraction) was washed with 500 µl of Reagent C and centrifuged as in previous step. Mitochondrial pellet was kept on ice before downstream processing.

2.2.3. Protein quantification

Concentration of protein lysates was determined using the Bicinchoninic Acid (BCA) Assay (Pierce™ BCA Protein Assay Kit; Thermo Scientific™, cat no. 23227). The assay was performed in borosilicate tubes (Fisher; cat no.14-961-26). 2 mg/ml albumin was diluted in distilled water to final concentrations of 0-2 mg/ml to prepare an 8-point standard curve. Samples were diluted 1 in 10 (5 µl of sample in 45 µl distilled water) and 1 ml BCA solution (1 part of Copper Sulphate solution in 50 parts Bicinchoninic Acid solution) was added to each standard and sample tubes and thoroughly mixed by vortexing. Tubes were then incubated in a water bath at 50°C for 60 min. 200 µl of each standard and sample was then aliquoted in duplicates in flat-bottomed 96-well plate and absorbance read at 540 nm using Biohit BP800 Platereader. A standard curve was generated based on the absorbance values of the known concentrations of BSA (bovine

serum albumin) (0.0 to 2.0 mg/mL). Concentration of diluted (1:10) protein lysates was assayed by comparison to the standard curve. Concentrations were calculated by comparison to the standard curve values of albumin All samples were then diluted with lysis buffer and 5x loading buffer (Table 2.1). Before first use, samples were denatured on a heating block at 95°C for 5 min (Techne, Dri-Block® DB100/2TC).

2.2.4. Western Blotting

2.2.4.1. Buffers

Table 2.1: Buffers used for Western blotting analysis.

Buffer	Components
1x lysis buffer	20 mM Tris-HCl (pH 7.5) 150 mM NaCl 1 mM Na_2EDTA 1 mM EGTA 1% Triton 2.5 mM sodium pyrophosphate 1 mM β-glycerophosphate 1 mM Na_3VO_4 1 µg/ml leupeptin
5x loading buffer	12.5 ml Trizma base (2M, pH6.8 stock) 5 g SDS (sodium dodecyl sulphate) 50 ml Glycerol 25 ml Mercapthoethanol 2.5 ml Bromophenol Blue (saturated) Distilled H_2O to 100 ml
10x running buffer (pH range of 8.3 +/- 0.3)	30 g Trizma base (250 mM) 144 g Glycine (1.92 M) 10 g SDS (1% v/v) Distilled H_2O to 1 l
1x transfer buffer (pH range of 8.3 +/- 0.3)	6.06 g Trizma base 28.8 g Glycine Distilled H_2O to 2 l

2.2.4.2. Polyacrylamide gel electrophoresis

Tris-glycine polyacrylamide gels were prepared as per Table 2.2 and placed in a Biorad mini gel tank filled with 1x running buffer. 8 µl of Chamaleon Duo pre-stained protein ladder and 20 µl of previously prepared protein samples were loaded onto gel. Proteins

were resolved first at 85 V for 15 min following by increased voltage of 140 V for up to 45 min.

Table 2.2: Composition of gels used for Western blotting analysis.

Component	10% Separating gel	10% Stacking gel
dH_2O	4.9 ml	3.55 ml
30% Acrylamide/Bis	3.3 ml	825 µl
2 M Tris-HCl pH 8.8	1.8 ml	-
1 M Tris-HCl pH 6.8	-	625 µl
10% APS (Ammonium persulphate)	100 µl	50 µl
TEMED (Tetramethylethylenediamine)	10 µl	5 µl
Total	10 ml	5 ml

2.2.4.3. Protein transfer onto polyvinylidene fluoride (PVDF) membranes

Following gel electrophoresis, proteins were transferred onto methanol-activated PVDF membrane. Briefly, PVDF membrane was submerged in methanol for 10 minutes and then washed with water, then in the following order: sponge, Whatman filter paper, Tris-glycine gel, PVDF membrane, Whatman filter paper, sponge, they were assembled into a ‘sandwich’ cassette (Biorad) and then placed into a transfer tank (Biorad), containing cold transfer buffer and a cooling block.Transfer was run at 225 mA for 90 min. Membranes were then incubated with Odyssey® Blocking Buffer (LI-COR Biosciences; cat no. 927-4000) diluted 50:50 with PBS on the rocker at room temperature for 1h. This step reduces subsequent non-specific binding of antibodies and background signal.

2.2.4.4. Probing blots with antibodies and blot analysis

Blocked membrane was then probed with primary antibody (ab) against protein of interest as well as anti-β-actin ab to serve as a loading control at the same time. All primary antibodies were diluted 1/1,000 in Odyssey® Blocking Buffer diluted 50:50 with PBS. Only anti-β-actin ab was diluted 1/10,000. Blots with all primary ab solutions were incubated on the rocker at 4°C overnight. The following morning, membranes were washed with 0.1% PBS-Tween (PBS-T; 3 x, on the rocker for 5 min) following by incubation with appropriate secondary ab (prepared as above, all diluted 1/10,000). After 45 min of incubation on the rocker at room temperature, blots were washed with PBS-T as in previous step, followed by 3x washes with PBS to remove residual Tween. Membranes were then placed overnight in the oven at 50°C to dry in order to enhance the signal. The following day blots were scanned using Licor Odyssey Scanner and band intensity analysed using Image Studio™ Lite software (LI-COR Biosciences).

2.2.4.5. Western blotting of proteins <20 kDa (RPLP2)

In order to detect proteins smaller than 20 kDa, several changes to the above describe methods had to be made. Stacking and separating gel components for 2 gels are shown in Table 2.3. Proteins were resolved for 15 min at 80 V followed by 30 min at 150 V. Transfer onto 0.2 μM PVDF transfer membrane was done in the presence of 20% methanol in transfer buffer (added to reagents shown in Table 2.1) and in cold room (4°C) for 2h at 100 V. Blots were then processed as described in Section 2.2.4.4.

Table 2.3: Composition of gels used for Western blotting analysis of proteins < 20 kDa.

Component	15% Separating gel	3.6% Stacking gel
dH_2O	2.3 ml	4 ml
30% Acrylamide/Bis	10 ml	9 ml
1 M Tris-HCl pH 8.85	7.5 ml	-
0.375 M Tris-HCl pH 6.8	-	2.5 ml
10% SDS	200 µl	0.85 µl
10% APS (Ammonium persulphate)	50 µl	25 µl
TEMED (Tetramethylethylenediamine)	50 µl	25 µl
Total	10 ml	5 ml

2.3. DNA preparation for Next Generation Sequencing

2.3.1. Lentivirus production from HEK293T cells

$8x10^6$ HEK293T cells were seeded onto 2 x 10 cm dishes ($4x10^6$ cells per dish) and incubated for 24 h. Human GeCKOv2 library was obtained from Addgene (#1000000048) as two plasmid DNA half-library reagents, A and B. Dr Reynolds had previously amplified the plasmid DNA into half-library A and B working stocks. HEK293T cells were then transiently transfected with the two half libraries (one dish with A, another with B) together with pVSV.g and psPAX2 packaging plasmids. Ratios of reagents are shown in Table 2.4. Produced viral particles were collected after 48 h and 72 h, pooled from both dishes to use as full library and sterile filtered using a 0.45 µm filter. Polybrene at 8 µg/ml was then added to the pooled virus to increase efficiency of infection of recipient cells.

Table 2.4: Volumes of reagents to produce lentiviral particles from HEK293T cells.

Library	No of 10 cm plates	pVSVs (pSD11) plasmid (µl)	psPAX2 plasmid (µl)	DNA (µl)	Mirus LT1 (µl)	Opti-MEM (ml)
LentiCRISPRv2 half-library A	1	2	3	4	27	1.5
LentiCRISPRv2 half-library B	1	2	3	4	27	1.5

2.3.2. Transduction of GeCKO library in 786-0 cells

Total of $13x10^7$ recipient cells (786-0, p54) were plated evenly in 20x T175 flasks and after 24h infected with virus to obtain MOI=0.3 (multiplicity of infection). This step theoretically prevented introduction of more than 1 sgRNAs to any given cell and therefore disruption of a single gene. It was based on calculations and probabilities, which make introduction of more than 1 sgRNA cell very unlikely. After 48h, cells were selected with puromycin (10µl of 10mg/ml stock solution added to 50ml media to obtain the final concentration of 2µg/ml). Puromycin concentration required to kill all cells without construct was determined by incubating cells with varying concentrations (0-2 µg/ml) for 7 days. After completion of selection, all cells were pooled and baseline DNA was collected (6 pellets of $6x10^7$ cells each) by trypsinising and washing with cold PBS. Cell pellets was then stored in -80°C freezer. $5.2x10^7$ cells were then split equally for two treatment conditions: drug (0.7 µM CB-839; Cayman Chemicals) or vehicle control (DMSO) and surplus discarded. After 16 days (4 cycles of 72h treatment cycles) cells from all flasks for each condition were pooled. DNA was collected as above. Remaining live cells that were not harvested for DNA, were plated and treated again. On day 21 of treatment (after the 5^{th} treatment cycle), DNA was collected again as described above.

2.3.3. Genomic DNA purification

Previously frozen cell pellet ($6x10^7$ cells) was re-suspended in 6 ml PBS, transferred to 50 ml Nalgene tubes (2 ml of cell suspension per tube) and purified using QIAGEN Blood & Cell Culture DNA Midi Kit (#13343). Briefly, cells were centrifuged (4°C, 15 min at 1,300 x g) using a JA25-50 rotor in the Beckman-Coulter Avanti J26-XP centrifuge and PBS discarded followed by re-suspension in 1 ml ice-cold Buffer C1 and 3ml distilled water per tube. Cells were centrifuged again (as described above) following pellet re-suspension in 5ml Buffer G2 to lyse nuclei and denature various proteins including histones. Next, 95 µl of Proteinase K (QIAGEN) was added and suspension incubated at 50°C for 60 min. In the meantime, QIAGEN Genomic tip 100/G was equilibrated with 4 ml of Buffer QBT and suspension applied to the tip to move through it by gravity flow. QIAGEN Genomic-tip was then washed with 2 x 7.5 ml of Buffer QC followed by eluting of genomic DNA with 5ml of Buffer QF and precipitated with 3.5 ml isopropanol into clean Nalgene tubes. Samples were then centrifuged (4°C, 30 min at 13,000 rpm) and pellet re-suspended in 2 ml of 70% ethanol, centrifuged again (4°C, 10 min at 13,000 rpm) and remaining pellet was re-suspended in 100 µl of MilliQ water. Tubes were then placed in 4°C for the DNA to dissolve overnight.

2.3.4. DNA quantification using QubiT fluorometer 2.0

Concentration of purified DNA was assessed using Qubit® dsDNA BR Assay Kit (Invitrogen; cat no. Q32850). Working solution (WS) was prepared by diluting 1 part of dsDNA BR Reagent in 199 parts of dsDNA BR Buffer. 190 µl of WS was aliquoted into each of the two 0.6 ml QubiT assay tubes and 198 µl aliquots were added to sample tubes. 10 µl of Standard A or Standard B were then added to the two tubes and 2 µl of

each sample was added to the remaining tubes. All tubes were briefly vortexed and incubated at room temperature for 2-3 minutes and read using QubiT fluorometer 2.0.

2.3.5. sgRNA amplification and purification

In order to achieve 300X coverage of GeCKO library, volume containing 260µg genomic DNA was calculated for each sample. 26 reactions containing 10 µg DNA each were prepared and amplified using G-Storm GS4 Multi Block Thermal Cycler. 30x PCR (polymerase chain reaction) mix containing Herculase II Fusion DNA Polymerase, 5x Herculase buffer, DMSO and dNTPs mix (Agilent; cat no. 600677) and custom oligos (Eurofins Genomics; synthesis scale 0.05 µM, purified using HPLC) as shown in Table 2.5. Amplification was performed using conditions shown in Table 2.6 and 10µl of PCR products from each reactions were next run on 2% agarose gel to confirm amplification of a 290 bp target and check for purity of samples. 100 µl PCR products were then purified using 180 µl Agentcourt AMPure beads (Beckman Coulter, cat no. A63880) as per manufacturer's instructions.

After optimisation of PCR conditions, 5 µl of each product from PCR1 was then used to run PCR2 (in duplicates) (Tables 2.7 and 2.8) with unique combination of forward and reverse primers for barcoding each sample (Table 2.9, also Appendix). Products were inspected for purity and presence of 350 bp fragment on 2% agarose gel. PCR2 products were purified and quantified using QubiT as above.

Table 2.5: PCR1 reagents per reaction.

Reagent	Volume (µl)
dH_2O	(up to 100 µl)
5x Herculase buffer	20
dNTPs Mix (25mM each)	1
DNA template	X µl (10 µg)
Forward adaptor primer	2.5
Reverse adaptor primer	2.5
Herculase DNA Pol	1
DMSO	3
Total	100

Table 2.6: PCR1 cycling conditions.

Step	Temperature	Time	Number of cycles
Initial denaturation	95°C	2 min	1 cycle
Denaturation	95°C	15 s	
Annealing	51°C	20 s	24 cycles
Elongation	72°C	30 s	
Final extension	72°C	3 min	1 cycle

Table 2.7: PCR2 reagents per reaction.

Reagent	Volume (µl)
dH_2O	65
5x Herculase buffer	20
dNTPs Mix (25mM each)	1
PCR1 product	5
Forward primer	2.5
Reverse primer	2.5
Herculase DNA Pol	1
DMSO	3
Total	100

Table 2.8: PCR2 cycling conditions.

Step	Temperature	Time	Number of cycles
Initial denaturation	95°C	2 min	1 cycle
Denaturation	95°C	15 s	
Annealing	68°C	20 s	18 cycles
Elongation	72°C	30 s	
Final extension	72°C	3 min	1 cycle

Table 2.9: Illumina primer combinations used for barcoding of individual libraries.

Sample/library	Forward primer	Reverse primer
Baseline 1	F1	R2
Baseline 2	F4	R4
DMSO D16 1	F1	R1
DMSO D16 2	F2	R1
CB-839 D16 1	F2	R3
CB-839 D16 2	F2	R4
DMSO D21 1	F3	R2
DMSO D21 2	F3	R1
CB-839 D21 1	F4	R1
CB-839 D21 2	F4	R3

2.3.6. Quality control using Bioanalyzer system

To confirm presence of a single main product, PCR products were analysed using Agilent 2100 Bioanalyzer system according to manufacturer's protocol. The High Sensitivity DNA Assay (Agilent, cat. no 5067-4626) allows for quality control of NGS samples following as few as four PCR cycles. Briefly, all reagents were equilibrated to room temperature for 30 min. Gel-Dye Mix was prepared by adding 15 µl High Sensitivity DNA dye concentrate to a vial containing High Sensitivity DNA gel matrix and mixed by vortexing. It was then transferred to a spin filter and centrifuged for 15 min at 2240 x g. Such prepared mix was then kept at 4°C protected from light and used within 6 weeks. Gel-dye mix was then added to specified wells on a chip placed in a priming station, followed by addition of marker and ladder used to assess sizes of DNA fragments. AMPure purified PCR2 products were then loaded at pg/µl concentration range and chip run on 2100 Bioanalyzer. Results were analysed using 2100 Expert software which produced electropherograms.

2.3.7. Quality control using KAPA Library Quantification kit

Samples were then subjected to qPCR analysis using KAPA Library Quantification (KAPA Biosystems, cat. no KK4824). Upon first use of the kit, 1ml of 10x Primer Premix was added the bottle containing 5 ml of 2x KAPA SYBR® FAST qPCR Master Mix and thoroughly mixed using vortex mixer. DNA samples were diluted 1:1,000 – 1:100,000 using 10mM Tris-HCl (pH 8.5, room temperature) to a final concentration of 1 – 24 ng/ µl. Three replicates of the following reaction were prepared as per Table 2.10: six DNA Strands, no-template control (with 4µl of MilliQ water in place of DNA), 10 library preparations. Tubes were then transferred to the Rota-Gene qPCR cycler and reactions

run as described in Table 2.11. After the end of final cycle, standard curve created using DNA Strands (Table 2.12) was analysed and Melt Curve produced. A single peak consisting of clustered samples is sufficient to conclude that Illumina machine will only bind desired fragments and all other fragments, which were amplified in a non-specific manner by PCR, will be lost and not included in sequencing pool. All samples were then diluted to 6.8 nM concentration and pooled together for sequencing at Edinburgh Genomics facilities at the University of Edinburgh, Edinburgh, Scotland, UK.

Table 2.10: KAPA mastermix reagents.

Reagent	Volume (µl)
KAPA SYBR FAST qPCR Master Mix (2X) + Primer Premix (10X)	12
MilliQ water (PCR-grade water)	2
DNA	4
Total	20

Table 2.11: Cycling conditions for qPCR using KAPA Library Quantification kit.

Step	Temperature	Time	Number of cycles
Initial denaturation	95°C	5 min	1 cycle
Denaturation	95°C	30 s	
Annealing/ Extension/ Data			35 cycles
Acquisition	60°C	45 s	
Melt Curve Analysis	65 - 95°C		

Table 2.12: Concentration of DNA Strands used to create standard curve.

Known standard	Concentration (pM)
DNA Strand 1	20
DNA Strand 2	2
DNA Strand 3	0.2
DNA Strand 4	0.02
DNA Strand 5	0.002
DNA Strand 6	0.0002

2.4. Generation of single-knockout stable cell lines

2.4.1. sgRNA oligo design

Guide sequence targeting genes of interest were selected at random from 6 sgRNAs available from Zhang laboratory directory for GeCKOv2 library. This was due to MLE algorithm producing gene, but not sgRNA ranking. Forward oligo was design by the addition of ‘CACCG’ sequence to the beginning of the sgRNA sequence. Reverse oligo was designed by the addition of ‘AAAC’ sequence to the beginning and ‘C’ after the last base of the sgRNA (Table 2.13). 50 nM, 5’-phosphorylated and desalted oligos were ordered from Invitrogen. Oligos were resuspended at 1 nmol/µl in MilliQ water and stored at -20°C.

To use for cloning, oligos were first diluted 1:10 in MilliQ water and forward and corresponding reverse oligos were annealed in an annealing reaction according to Table 2.14. The reaction was performed using heat block at 95°C for 5 min and then tubes were left in the block until the temperature reached 30°C (around 2.5-3 h).

Table 2.13: Primer sequences used to amplify specific guide sequences (sgRNA).

Gene	sgRNA ID	Sequence	Forward primer	Reverse primer
FOXO3	HGLibB_17900	CCTGCCATATC AGTCAGCCG	CACCGCCTGCC ATATCAGTCAG CCG	AAACCGGCTGA CTGATATGGCA GGC
FOXC1	HGLibA_17821	CGGACCGGGC GCGTTGCCGT	CACCGCCGGAC CGGGCGCGTTG CCGT	AAACACGGCAA CGCGCCCGGTC CGC
FOXC1	HGLibB_17799	GGGCGGCTACA CCGCCATGC	CACCGGGGCG GCTACACCGCC ATGC	AAACGCATGGC GGTGTAGCCGC CCC
RPLP2	HGLibA_49301	CGGCCGCTACC TTGTTGAGC	CACCGCGGCCG CTACCTTGTTGA GC	AAACGCTCAAC AAGGTAGCGGC CGC
RPLP2	HGLibA_49302	GTATCGAGGCG GACGACGAC	CACCGGTATCG AGGCGGACGAC GAC	AAACGTCGTCG TCCGCCTCGAT ACC
DIAPH3	HGLibB_13206	GAATTTAAATGA AGATAAAA	CACCGGAATTT AAATGAAGATAA AA	AAACTTTTATCT TCATTTAAATTC
ALDH18A1	HGLibA_01539	GGGCAGCACAG ATGCTGTAC	CACCGGGGCAG CACAGATGCTG TAC	AAACGTACAGT CTGTGCTGCCC C

Table 2.14: Annealing solution reagents.

Reagent	Volume (µl)
0.1 nM forward oligo	1
0.1 nM reverse oligo	1
10X NEB Buffer 2	5
dH_2O	43
Total	50

2.4.2. LentiCRISPRv2 plasmid digestion

LentiCRISPRv2 was a gift from Feng Zhang (Addgene plasmid #52961; http://n2t.net/addgene:52961; RRID:Addgene_52961). It is a 3rd generation lentiviral backbone (10,000 bp in size without insert) and a one-vector system. In addition to the *Cas9* gene insert (4200 bp), it also contains puromycin resistance gene insert (600 bp). It also contains a 2 kb filler region, which is removed by digestion with BsmBI enzyme (Fermentas/ Life Technologies, FD0454). In addition, this plasmid was dephosphorylated before a desired guide sequence can be cloned in. Reagents and volumes used are listed in Table 2.15. Additionally, one tube containing all reagents listed in the table, minus the enzyme, was used as negative control. Tubes containing reaction mix were then incubated at 37°C for 30 min using heating block.

Table 2.15: Reaction mix used to remove the filler and dephosphorylate the LentiCRISPRv2 vector.

Reagent	Volume (µl)
1 µg/µl pLentiCRISPRv2	2
FastDigest Esp3I (*BsmBI*) enzyme	1
10X FastDigest buffer	2
Recombinant shrimp alkaline phosphatase (rSAP) (NEB #M0371S)	0.5
100 mM freshly prepared 1,4-Dithiothreitol (DTT) (Roche, #10197777001)	0.2
dH_2O	14.3
Total	20

2.4.3. Plasmid and sgRNA ligation

sgRNA sequence was then cloned into the enzyme-digested and dephosphorylated vector using the Rapid DNA Ligation kit (5x DNA dilution buffer, 2x ligation buffer, ligase) (Roche #11635379001). First, all ligated oligos were diluted 1:10, 1:100 and 1:1,000 and used in separate annealing reactions to optimise oligo concentration needed for later

steps. Reagents and volumes used are listed in Table 2.16. Additionally, a tube containing all reagents except for any sgRNA oligos was used as negative control. Tubes with mixes were then incubated at 16°C overnight in thermocycler.

Table 2.16: Reaction mix used ligate pLentiCRISPRv2 and guide sequence.

Reagent	Volume (µl)
dH_2O	5
Annealed oligos	1
digested gel purified pLentiCRISPRv2	2 (approximately 50 ng)
5x DNA dilution buffer	2
2x ligation buffer	10
ligase	1
Total	21

2.4.4. STBL3 bacteria transformation

The next day, after the ligation reaction was completed, plasmids containing inserts were used to transform One Shot™ Stbl3™ Chemically Competent *E.coli* bacteria (Thermofisher, #C737303). Briefly, 50 µl aliquots of STBL3 bacteria were thawed on ice before the addition of the previously prepared ligation mix and incubation for 20 min on ice following by heat shock at 42°C for 30 sec and another incubation on ice for 10 min. Transformed bacteria were then added to LB broth (without any antibiotics) and incubated shaking at 37°C for 60 min. 200 µl of each mix was then plated on 10 cm Petri dishes containing set agar and ampicillin. Plates were incubated at 37°C overnight. The next day, single colonies were picked (3 colonies per plate) and grown in individual tubes containing 5 ml LB and 100 µg/ml carbenicillin (disodium salt) (Fisher Scientific, #10396833). Liquid cultures were grown shaking at 37°C for up to 16 hours.

2.4.5. DNA mini-prep

Following amplification in STBL3 bacteria, plasmid DNA was purified using the Monarch® Plasmid DNA Miniprep Kit (NEB #T1010) according to the manufacturer's protocol. Briefly, 3ml of bacterial culture was pelleted in a microcentrifuge for 30 sec at 16,000 x g and supernatant discarded. All subsequent centrifugation steps were performed at 16,000 x g. Pellet was then resuspended by vortexing in 200 µl of Plasmid Resuspension Buffer until no clumps were observed. Bacteria was then lysed by adding 200 µl of Plasmid Lysis Buffer and inverting the tube 5 times until colour changed to dark pink and the solution was clear and viscous. Tubes were then incubated for 1 min at room temperature. The lysate was then neutralised by adding 400 µl of Plasmid Neutralization Buffer and gentle inversion of tubes until colour was uniformly yellow and a precipitate formed. Tubes were then incubated for 2min at room temperature. To clarify the lysate, tubes were then centrifuged for 2 min. Supernatant was then transferred to a DNA spin column, centrifuged for 1 min and flow-through discarded. Spin column was then re-inserted into a collection tube and 200 µl of Plasmid Wash Buffer 1 was added to remove RNA, protein and endotoxin. Tubes were then incubated for 5 min at room temperature and centrifuged for 1min. After the flow-through was discarded, 400 µl of Plasmid Wash Buffer 2 was added to the column and centrifuged for 1 min. The column was then transferred to a clean 1.5 ml collection tube and 30µl of DNA Elution Buffer was added to the centre of the column matrix. Samples were incubated for 1 min at room temperature, following by centrifugation for 1min. Concentration and purity of eluted plasmid DNA was then measured using Nanodrop 2000 spectrophotometer, which uses UV-visible spectrum to measure fluorescence of a sample in order to quantify and assess purity of DNA, RNA or protein samples .

2.4.6. Insert confirmation by the NdeI/ EcoRI double digest

In order to confirm guide sequence insert, purified plasmid DNA was then treated with EcoRI and NdeI restriction enzymes. Reaction reagents and volumes are listed in Table 2.17. Empty vector was used as control. Reaction mixes were then incubated for 1 h at 37°C in heat block followed by gel electrophoresis on 2% (w/v) agarose gel. Following electrophoresis, gel was incubated in 50 ml TAE buffer and 2 µl Ethidium Bromide to enhance the visibility of the smallest band. Three bands were observed: 10,374 bp, 2,299 bp and 154/174 bp. 154 bp band indicated lack of insert, whereas 20 bp shift to 174 bp band indicated presence of the guide sequence insert.

Table 2.17: Double digest reaction mix reagents and volumes to confirm guide sequence insert into pLentiCRISPRv2.

Reagent	Volume (µl) for 1x mix
EcoRI enzyme	1
NdeI enzyme	1
10x CutSmart Buffer	3
Purified plasmid DNA	Equivalent of 1 µg
dH_2O	Up to 30
Total	30

2.4.7. Insert sequence verification using Sanger sequencing

After confirmation of the presence of an insert, it was then sequenced at DNA Sequencing and Services at the University of Dundee. Samples were prepared at a concentration of 25-30 µg/µl in a final volume of 20 µl (500-600 µg plasmid DNA in total). Inserts were sequenced using LKO.1 5' (5'GACTATCATATGCTTACCGT3') primer provided by the sequencing facility.

2.4.8. Glycerol stock preparation

After confirmation of the correct sequence by Sanger sequencing, glycerol stocks were prepared by mixing of 250 µl of 80% (v/v) sterile glycerol and 750 µl liquid culture. Stocks were stored at 80°C.

2.4.9. Maxiprep

In order to achieve sufficient plasmid DNA for transfection, 200 µl of liquid culture was grown in a conical flask containing 200 ml LB and 200 µl 10 mg/ml carbenicillin. Cultures were incubated overnight at 37°C with shaking. The following day, bacteria were harvested by centrifugation at 6,000 x g for 15 min at 4°C in 500 ml bottles (Beckman-Coulter Avanti J26-XP centrifuge with JA10 rotor). Supernatant was then discarded and pellet processed using EndoFree Plasmid Maxi Kit (QIAGEN, #13262) according to manufacturer's protocol. Briefly, pelleted cells were completely re-suspended by pipetting and vortexing in 10 ml of lysis (P1) buffer. Next, 10 ml of Buffer P2 was added and sealed tube inverted six times until the solution appeared viscous. Lysate was then incubated for 5 min at room temperature. After 5 min, 10 ml of chilled Buffer P3 was added to enhance precipitation, sealed tube inverted six times and poured into QIAfilter Maxi Cartridge. Lysate was then incubated in the cartridge closed with a nozzle for 10 min. This step allows for separation of genomic DNA, proteins and detergent, which form precipitate on top of the clear solution, from the rest of the lysate and prevents clogging of the cartridge. Plunger was then inserted and approximately 25 ml of the cell lysate filtered into a 50 ml tube. 2.5 ml of Buffer ER was then added to the lysate, tube sealed and inverted 10 times and incubated on ice for 30 min.

In the meantime, QIAGEN-tip 500 was equilibrated by applying 10 ml of Buffer QBT. Lysate was then applied to the tip and left to pass the column using gravity flow. The column was then washed two times with 30 ml Buffer QC to remove most contaminants and plasmid DNA with 15 ml of Buffer QN into a new 50 ml tube. 10.5 ml of isopropanol was added to the lysate, mixed by inverting and solution immediately centrifuged at ≥15,000 x g for 30 min at 4°C. Supernatant was then carefully poured out and remaining pellet washed with 5 ml of 70% endotoxin-free ethanol by mixing and then immediate centrifugation at ≥15,000 x g for 10 min. Supernatant was then removed using a pipette and pellet air-dried for 5 – 10 min before leaving it to dissolve in 200 – 300 µl of Buffer TE. The following day, concentration and purity of the eluted plasmid DNA was measured using Nanodrop 2000 spectrophotometer, diluted to a final concentration of 1 µg/µl in MilliQ water and stored at -20°C.

2.4.10. Transfection

786-0 cells were used to create stable single-gene knockout cells and used in further phenotypic studies due to their use in the screening experiment. Transfection process was carried as described previously for transfection with the GeCKOv2 library, the only difference being use of only one plasmid DNA with a guide sequence of interest. Following transfection, cells were selected with 2 mg/ml puromycin for 7 days and immediately frozen and stored in liquid nitrogen.

2.5. RNA interference

786-0 cells were seeded onto 6-well plate at a density of 1.2 x 10^5 cells/well in 2.5 ml of complete culture medium and left to attach overnight. After 24h cells were transfected

by using Lipofectamine™ RNAiMAX Transfection Reagent (ThermoFisher, cat.no 13778075) with Silencer® Select Negative Control #1 siRNA (ThermoFisher, cat.no 4390843) or one of two siRNAs targeting FOXC1 (ThermoFisher, cat. no 4392420; s5226 and s5227) at a final concentration of 5 or 10 nM according to the manufacturer's instructions. The following day cells were trypsynised and plated for SRB assay as described in Sections 2.1.5.1 and 2.1.5.2. Alternatively, cells were transfected again and plated for SRB assay the following day.

2.6. Statistics and data handling

Reads generated by Illumina sequencing generated four pools, which were separated into 10 samples based on inline barcodes at the beginning of read one. The 'process_radtags' tool from the 'Stacks' package (version 2.0b) was used to demultiplex the reads with the parameter ' –disable_rad_check'. Reads were trimmed using cutadapt (version 1.9) with the parameters: ' –discard-untrimmed -m 10'. Trimmed reads were aligned to the sequences (provided by you) of the guide sgRNAs in the library using bowtie (version 1.2.2). Only the first read in each pair was mapped, read two was discarded. The 'featureCounts' tools from subread (version 1.5.1)5 was used to count the reads aligning to the sgRNA sequences in each sample. Parameter ' -s 1' was used specify reads should be align to the forward strand.

Version 0.5.3 of the MAGeCK tool was used for data analysis. Batch correction and data QC was performed using the MAGeCK-VISPR pipeline. Testing for selection was performed on the batch corrected data using the 'mle' function of the MAGeCK-VISPR pipeline with the following parameters '–norm-method control nonessential_sgRNA.txt' Where 'nonessential_sgRNA.txt' is a list of all the sgRNAs in the library provided by you

that correspond to genes deemed to be non-essential. The test was performed to compare drug (CB-839) treated to control (DMSO treated) samples, and was performed once for each timepoint. MaGeCKFlute was used to normalise the beta scores produced by 'MAGeCK mle' and compare the drug treated samples to the control samples at each time point. MaGeCKFlute was also used to look for GO terms and Kegg pathways enriched among the drug selected genes. MaGeCKFlute was run once for each timepoint.

MAGeCK mle calculates beta scores for each gene in each condition (compared to the baseline). The higher the beta score, the more positively selected the gene. MaGeCKFlute identifies genes under positive or negative selection by comparing the beta scores calculated by MAGeCK mle. The genes were put in to six groups (groups 1-4 are not mutually exclusive with A and B) (Table 2.18).

Table 2.18: Gene grouping based on beta scores used by MLE algorithm.

Group	Beta score significance
Group 1	Negative selection in control, no selection in drug treated.
Group 2	No selection in control, positive selection in drug treated.
Group 3	Positive selection in control, no selection in drug treated.
Group 4	No selection in control, negative selection in drug treated.
Group A	Higher beta score (more positive selection) in the drug treated samples than the control samples.
Group B	Higher beta score (more positive selection) in the control treated samples than the drug treated samples.

For Western Blot quantification, protocol available at Licor website was used (https://www.licor.com/documents/7bd2dev0rfjofad7dr1zf85wip4g4it3).

For SRB assay, OD values obtained as described in Sections 2.1.5.2 and 2.1.5.3, were adjusted using blank OD subtraction and averaged for each column corresponding to 6 replicate wells treated with the same drug concentration. Standard deviation for adjusted OD values was then calculated. Percentage survival was calculated by dividing adjusted

OD value for a specific drug condition by adjusted OD value for control (no drug treatment). Such obtained values were then used for GraphPad calculations. First concentration values were transformed using log[10]. Then, to analyse data 'log(inhibitor) vs.response (three parameters)' option was selected for IC50 quantification.

3. Optimisation of the CRISPR/Cas9 screen parameters

3.1. Introduction

Pooled genetic screens allow for simultaneous testing of thousands of individual genetic perturbations in a single experiment [283]. They are more cost-effective than arrayed libraries for the discovery of novel regulators of biological processes and allow for a more unbiased approach. Pooled libraries, such as GeCKOv2, available to the academic community via Addgene, remove the need to create a new well represented library, which covers the whole genome and other time-consuming library validation steps, associated with creating a new library. On the other hand, commercial libraries can incur higher initial cost, but after initial amplification, they can be used for multiple screens. It is important to choose the right library for a given application. More sgRNAs targeting single gene allow for higher statistical certainty for hits, especially when hits have weaker initial phenotypic effect. Fewer sgRNAs, such as 2 to 4 per single gene, have the advantage of requirement of fewer cells in the initial phase of the experiment. Despite the chance to test more cell lines at a smaller scale than one cell line with 6 sgRNAs per single gene, the costs associated with either are comparable [283]. GeCKOv2 library was chosen here due to its superiority over other screening methods and libraries (see sections 1.6.3. – 1.6.4.). Regardless, there are several practical considerations that have to be addressed before the screening process could be initiated. Optimisation of

parameters, such as choosing the right cell line and drug concentration, here CB-839, optimal puromycin concentration, viral titre and various quality control methods, are necessary to enhance hit discovery and minimise false positives.

Puromycin dihydrochloride is an antibiotic, which works by inhibiting protein synthesis and kills mammalian cells fast and at low concentration in culture. It is an antibiotic and an analogue of aminoacyl-tRNAs and causes premature termination of polypeptide synthesis [284]. Puromycin concentration of 1 to 10 µg/ml is sufficient to kill all mammalian cells lacking puromycin resistance gene in the dish within 4 to 7 days. Cell line-specific concentrations for selection with puromycin can be found in the literature, however it is necessary to determine to correct value empirically and close in time to the screen. This is because drifts in conditions between laboratories and even within one laboratory are common. It is usually added to the growth media after trypsinisation and before cells attach to the bottom of a new vessel. Selective media have to be replaced every few days in order to remove dead cells and ensure constant antibiotic availability. There are other antibiotics, such as hygromycin and geneticin, commonly used in the laboratory practice, however the choice is dictated by the application. LentiCRISPRv2 used in the GeCKOv2 screen encodes, amongst other elements, puromycin resistance gene, which is expressed in stably transfected 786-0 cells (see section 1.6.4.). Using too little puromycin will results in survival of uninfected and uninformative cells. On the other hand, too much antibiotic will also kill successfully infected cells and is likely to have non-specific and unpredictable consequences on gene expression of those that survive. Thus, by applying selection using puromycin at concentration experimentally determined to kill all cells that were not exposed to the plasmid, only cells that contain LentiCRISPRv2 plasmid survive.

Next step is to calculate the number of viral particles required to infect cells. Based on the Poisson distribution, 20-60% infection efficiency is sufficient and required to integrate the majority of target cells with a single sgRNA per cell [283]. This reduces the initial number of target cells, as well as cells, which will not be successfully infected. Very high infection rates are not desirable, due to nonspecific effects of multiple sgRNAs integrated in a single cell (see section 1.6.3.). Low MOI of 0.3-0.5 is recommended what translates to infection of 25-39% of cells exposed to virus [285]. Additionally, virus production method is a recognised source of inaccuracies in fold-change measurements of sgRNAs following sequencing of samples. Excess plasmid delivered to packaging cells (here 293T cells) and transferred with the viral supernatant to recipient cells dilutes over time, however, it can remain unchanged in early samples used as a baseline (the initial screen reference) [286]. Plasmid DNA may be co-purified with the genomic DNA and amplified in PCR leading to large differences in template abundance between baseline and DNA from cells at a later stage of the screen. Therefore, it is crucial to establish viral titre, but also use appropriate DNA purification kits and methods.

Isolation of genomic DNA from cells is a critical step and its quality affects all downstream sample preparations, including PCR and finally NGS data and its analysis. Many DNA purification kits are available, but they vary in yield and quality of genomic DNA. Use of Endonuclease-free genomic DNA purification kit allows for rapid purification of high yield and quality genomic DNA. It also enhances the effect of MOI on template accuracy by not co-purifying plasmid DNA that might have remained in excess in some samples. It is also of high importance to ensure that there are no PCR inhibitors present in any samples, as they are not detected by UV absorbance or fluorescence-based DNA quantification methods and severely impact PCR performance [283]. Additional purification and quality control methods, including AMPure bead purification were applied

to obtain highest quality material. Optimisation of quality control methods is described in Chapter 4.

PCR is a well-documented source of bias in measuring abundance of multi-template populations [287, 288]. Therefore, it is important to carefully optimise cycling conditions, such as annealing temperature and number of PCR cycles to obtain accurate and representative amplification of such populations [289]. Additionally, compatibility of primer pairs with templates (DNA collected at baseline, DMSO-treated and CB-839-treated samples separately giving 10 templates) had to be tested to ensure successful amplification of all templates in PCR2, where unique barcodes are added to PCR libraries.

3.2. Choosing the right cell line and validating its characteristics

Research on immortalised cultured cell lines has had an enormous impact on understanding of various aspects of cancer. Despite using monolayers of single type of cells as a surrogate for far more complicated systems, for years basic research has been successfully translated into clinical use. In terms of research into ccRCC, most commonly used cell lines include 786-0, 769-P, Caki-1, ACHN and several other. In order to choose the most appropriate cell line, various databases were used, such as Pubmed, ATCC and COSMIC with focus on their *VHL* status, as well as their reported use in experiments involving CB-839. Other determinants included other genetic aberrations and easiness of culture. Comparison of 5 ccRCC cell lines is summarised in Table 3.1. ACHN, although widely used to study ccRCC due to similarity in gene expression, is by some considered to be more representative of the papillary type given its chromosomal aberration profile, characteristic c-met polymorphism and WT *VHL* status [290-292]. Caki-1 cells appeared to be a good candidate based on the above mentioned selection criteria, however, upon

culture, they grew in clumps and were difficult to count. The main reason for choosing 786-0 cells over 769-P was their faster doubling time. After these considerations, 786-0 was chosen as the cell line for library screening.

First set of experiments involved confirmation that the cells being used are in fact 786-0 cells followed by establishment of their doubling time. Analysis of short tandem repeats (STRs) confirmed that the cell line used was derived from the same source as the Cellosaurus reference cell line as profiles matched in 97% (see Appendix). Misidentification and cross-contamination of human cell lines has plagued biomedical research for decades [293]. It continues to cause incorrect conclusions, problems with reproducibility and has been highlighted by the use of HeLa cell line, which is one of the most widely used cell lines [294]. STR profiling is a widely used and recommended method used to unambiguously authenticate human cell lines [295]. Other available methods, such as karyotyping, immunotyping and isoenzyme analysis do not allow to confirm human cell line identity with the same confidence as STR profiling.

Calculation of cell doubling time is crucial for any planned experiments, including assessment of viral particles needed to infect cells for screening purposes. This also prevents cells from reaching 100% confluence in a very short time and frequent exposure to stress caused by trypsinisation. Literature search using PubMed website returned doubling time values of 45h and 25h for 786-0 cells [296, 297]. Based on known seeding density and cell count after estimated two doublings, we found that our cell line needed approximately 26.5h to double.

Table 3.1: Comparison of cell lines widely used to study ccRCC.

Mutation status for each cell line obtained from COSMIC database [298].

Cell line / Feature	786-0	769-P	A704	Caki-1	A498
Origin	Primary [299]	Primary [300]	Primary [301]	Metastatic (skin) [302]	Primary [303]
Karyotype	Male; hyper-3n; Y chromosome in 60% cells [299]	Female; large numbers of 4n-, 6n-, and higher-ploid cells [300]	Male; 2n, hyper-2n, hyper-3n, hyper-4n [301]	Male; aneuploidy, 3n range, no Y chromosome [302]	Female [303]
Tumorigenic in nude mice	Yes [299, 304]	Yes [300, 305] / No [304]	No [301]	Yes [302]	Yes [303]
VHL	Mutated, 1bp del in codon 104 (STOP codon 158) [306]	Methylated promoter [306]; no mRNA	Methylated promoter (no protein) [306]	WT *VHL*	4bp del (STOP 157) [306] / no mutation [307]
PBRM1	WT [308]	WT [309]	Mutated [308]	WT	WT [309]
NF2	WT	WT	WT	WT	WT
SETD2	WT	WT	Mutated	Mutated	Mutated
mTOR	WT	WT	WT		
PTEN	Mutated	WT	WT		
Tp53	Mutated	WT	Heterozygous mutation	WT	Mutated; low expression
3p & 14q loss	Yes	Yes	Yes	Yes	No
VEGF	High levels	High levels		High levels	High levels
HIF1α protein	No (truncated mRNA) [310]	No (& no mRNA) [306]	Yes [311]	Full length mRNA [310]	No [306, 310]
HIF2α protein	Yes [310]	No [306]		Yes	Yes
Other	Expresses vimentin	*BAP1* mutated (expressed)		*Met* mutated	

3.3. Optimal seeding density and CB-839 concentration selection

SRB assay allowed to measure the effect of a range of drug concentration on growth of 786-0 cells. First measurements for lower seeding density ($7.5x10^3$ cells/ml; $1.5x10^3$ cells/well) showed that 700 nM CB-839 is sufficient to inhibit cell growth by 50% in comparison to control cells with no drug added. This, however would translate to 32x T175 flasks for screening purposes. Therefore, higher seeding densities were tested next: that would allow for plating of $2.6x10^7$ cells across 4, 6, 8 and 10 flasks (Table 3.2, Figure 3.1).

Seeding density of $3.25x10^6$ cells per T175 flask and CB-839 concentration of 0.7 μM was then chosen for further experiments. Although IC_{50} value was calculated at 0.53 μM, the reason for choosing higher concentration is to ensure selection pressure that would result in significant findings during sequencing step. We aim to select only cells that are resistant to CB-839. To confirm that observed changes are due to the action of CB-839, cells were grown in parallel in media only, with DMSO (vehicle control) and CB-839 (Figure 3.2). CB-839 (dissolved in DMSO) inhibits growth and affects morphology of 786-0 cells, which appear more elongated suggesting stress. DMSO alone impacts cell growth and morphology, although to the lesser extent than CB-839. This confirms inhibitory potential of CB-839 on 786-0 cells and highlight the importance of using cells cultured in media supplemented with DMSO as a control in screening process.

Table 3.2: Seeding densities for cell growth assay (pilot experiment).

No. of T175 flasks (seeding surface 175 cm^2)	Cells per flask (total of 2.6x107 cells)	Cells per well (96-well plate seeding surface 0.32 cm^2)
4	6.5×10^6	1.2×10^4
6	4.3×10^6	7.8×10^3
8	3.25×10^6	5.9×10^3
10	2.6×10^6	4.75×10^3

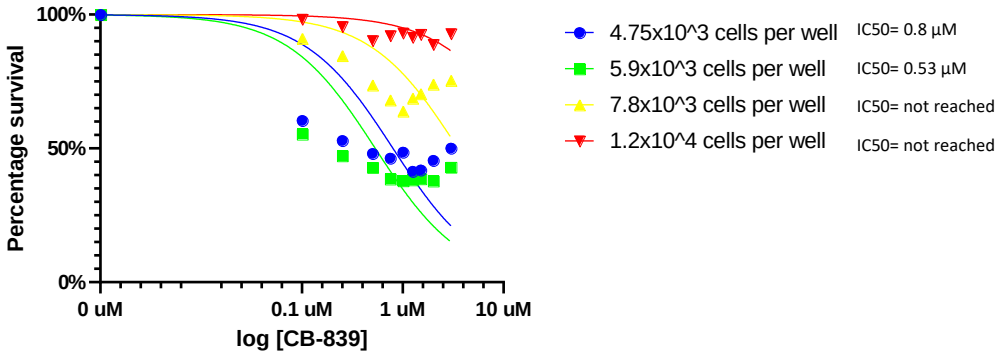


Figure 3.1: Effect of a range of concentration of CB-839 on growth of 786-0 cells seeded at different densities.

CB-839 concentration of 3 µm was not sufficient to inhibit growth by 50% of cells seeded at either of the two higher densities. IC50 values were obtained to the two lower seeding densities and therefore used for further experiments. All conditions were repeated in six wells and tested simultaneously. Experiment was not repeated to obtain biological replicates.OD values were first calculated after blank subtraction, then averaged and standard deviation used to obtain percentage survival.

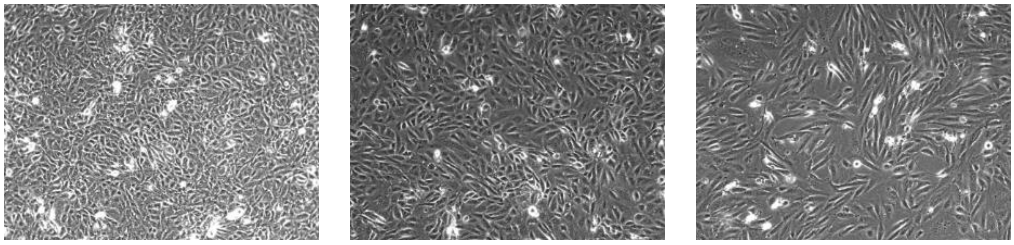


Figure 3.2: Morphology of 786-0 cells cultured in different media conditions.

786-0 cell were cultured in RPMI (a); DMSO (b); 0.7μM CB-839 (c). Initial seeding density was the same for all culture conditions. Although these results are not quantitative, it can be observed that treatment with 0.7μM CB-839 negatively impacts cell proliferation (fewer cells when compared to cells cultured in RMPI only or RPMI with DMSO). DMSO is a solvent for CB-839 and here the same volumes of DMSO were used in DMSO only and CB-839-treated cells.

3.4. Confirmation of glutamine requirements for the growth of 786-0 cells

Additionally, it was important to confirm that proliferation of 786-0 cells is dependent on glutamine. To do that, equal number of 786-0 cells was seeded in four different media conditions: RPMI supplemented with GlutaMAX^{TM} (a more stable alternative to L-glutamine) with or without CB-839 and RPMI without L-glutamine with or without CB-839. It can be concluded from Figure 3.3 that 786-0 in fact rely on glutamine for their growth. Moreover, lack of glutamine results in stronger growth inhibition than that caused by treatment with 0.7 μM CB-839 for 72h. Addition of CB-839 to media lacking glutamine does not cause stronger inhibitory effects on 786-0 cells.

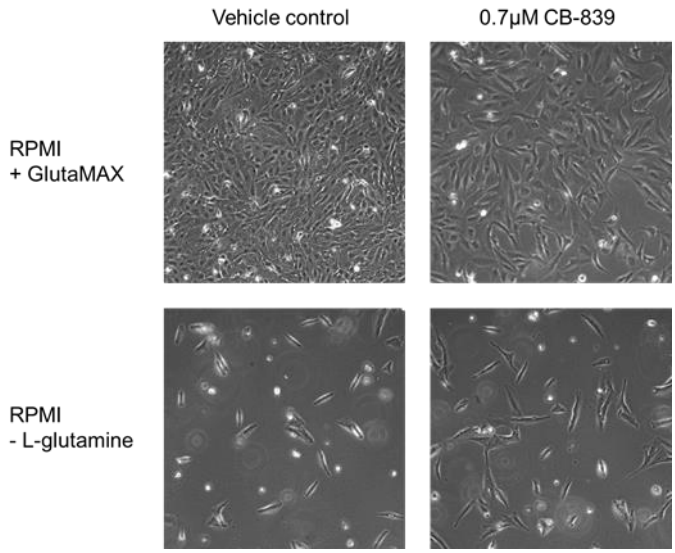


Figure 3.3: Effect of glutamine presence in growth medium on growth of 786-0 cells.

786-0 cells require addition of glutamine in growth media for optimal growth. Culture with 0.7μM CB-839 and glutamine does not negatively impact cell growth to the same extent as culture without glutamine with or without 0.7μM CB-839.

3.5. Pilot experiment

3.5.1. Puromycin titration

In order to determine minimum concentration of puromycin that kills all 786-0 cells, fixed number of cells was plated on a 6-well plate and each well was exposed to different puromycin concentrations, ranging from 0 to 2 µg/ml. Media with puromycin was replaced every 2 to 3 days and images taken under light microscope (x20 magnification) on day 7 of selection. This experiment was repeated three times and 2 µg/ml of puromycin was found to kill all 786-0 cells within 7 days (Figure 3.4).

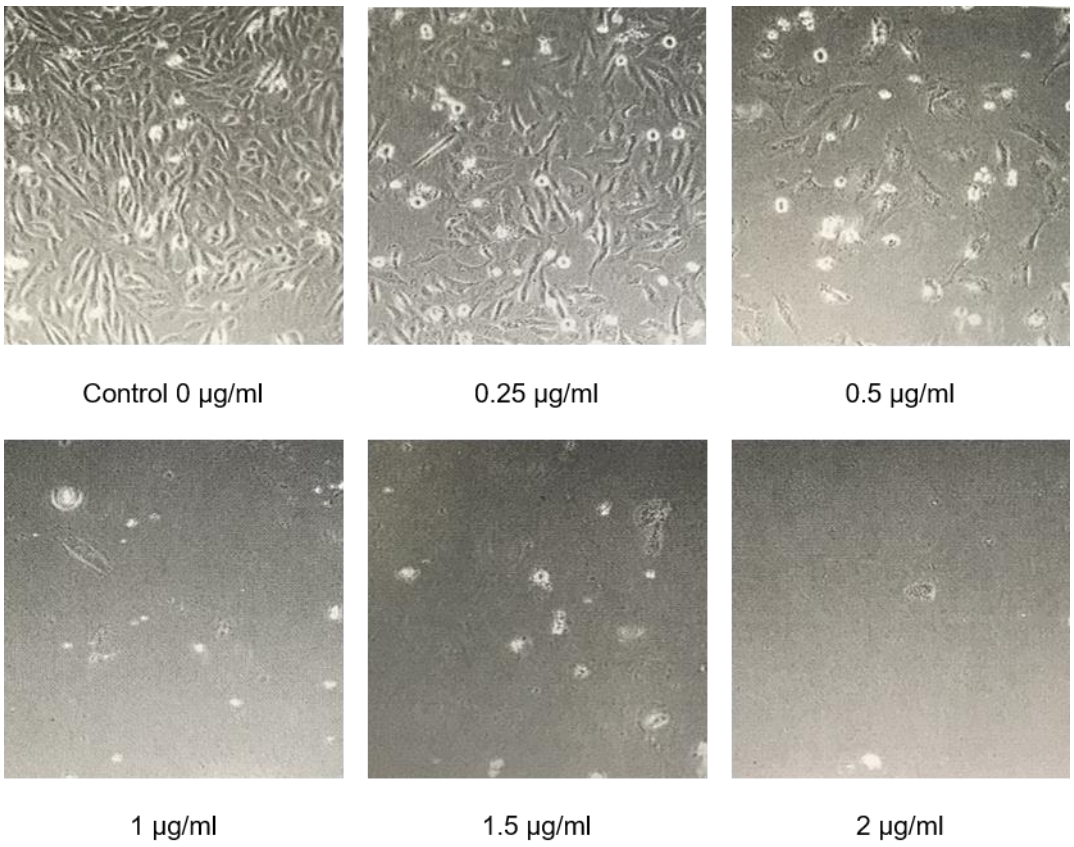


Figure 3.4: Puromycin titration experiment.

786-0 cell were seeded in a 6-well plate (1 x 10^4 cells per well in 2 ml media) and treated with increasing concentrations of puromycin (0 – 2 µg/ml).

3.5.2. Lentivirus titration

To ensure that most cells receive only one copy of sgRNA, number of viral particles required to obtain MOI of 0.3 was calculated. Given that approximately $1.7x10^7$ cells cover whole growth area of T175 flask, I calculated the number of cells that equal 40% confluency at $6.8x10^6$ cells. The reason for choosing 40% of area to be covered is to ensure that cells do not overgrow during time required for infection with viral particles. Since $1.3x10^8$ cells is the minimum number to ensure 300x library coverage, I plated cells across 20x T175 flasks, what gave $6.5x10^6$ cells in each. Bearing in mind the above mentioned doubling time of 786-0 cells, we assumed that after 24h between cell plating and infection, 60% of cells will divide giving $1.1x10^6$ cells per flask. Viral titre had been previously calculated under similar experimental conditions in the host laboratory by Awa Sarr at $4x10^7$ viral particles/ml. Therefore, $3.3x10^6$ particles or 82 µl of suspension containing half library A and B, each, will be used for $1.1x10^7$ recipient cells. It had previously been shown that 15ml is sufficient volume to cover the cells, we then performed viral solution dilution up to 1/200 which resulted in 16.4ml of media containing required number of particles (libraries A and B) per flask.

3.5.3. Optimisation of PCR1 & PCR2 cycling conditions

Baseline DNA from $6x10^7$ cells used in pilot experiment was purified to obtain 300x library coverage or approximately 396 µg of DNA as previously recorded by Reynolds laboratory. However, only 199.7 µg of DNA in 298 µl was collected after taking into account that 2 µl had been used in quantification process using QubiT fluorometer. This could have been due to one of the DNA purification columns being blocked as a result of overloading. Second attempt of purification produced higher yield of 297 µg DNA but this was still lower than expected.

Purified baseline DNA from pilot experiment was then used to optimise conditions for PCR1 and PCR2. Temperatures used to denature and extend PCR product did not require optimisation and were 95 °C and 72 °C, respectively and used in both PCR reactions. In first instance, samples containing 1 or 10 µg DNA as a template were amplified using conditions previously determined by Awa Sarr (24 cycles, annealing temperature (T_A) of 50 °C). Gel electrophoresis (2% agarose gel; 80V, 45 min) confirmed amplification of desired fragment (290 bp) in samples with lower initial DNA quantity, but not for higher value. It was important to optimise these conditions as 10 µg DNA was going to be used to amplify fragments from samples obtained from the complete screen. Thus, gradient PCR, with T_A between 59.3 and 60.2 °C and 20 amplification cycles, was run with 10 µg template DNA. None of the conditions resulted in amplification of the correct size DNA fragment (Figure 3.5). Another PCR gradient was run, with T_A ranging from 50.4 to 50.2 °C and 22 cycles. Sample amplified at 50.4 °C showed faint 290 bp band and that amplified at 51.9 °C showed slightly stronger band. Final conditions with T_A = 51 °C and 22 cycles were confirmed as correct ones by running 10 µg DNA in triplicate (Figure 3.6).

Following optimisation of cycling conditions to be used in first PCR, amplified products were used to optimise second PCR conditions. In first instance, 5 µl of PCR1 product was run using several selected primer pairs using T_A = 60 °C and 20 cycles. Correct size band was observed for all primer pairs, however multiple non-specific bands were also present (Figure 3.7). This suggested that either T_A or number of cycles, or both, are too high. Gradient PCR with T_A ranging from 45.1 to 60.3 °C and 18 cycles returned similar results to those described above. Lowering cycle number to 16 while keeping the same temperature gradient only slightly reduced the intensity of non-specific bands observed on agarose gel. Finally, gradient PCR with T_A ranging from 60.1 to 68.2 °C and 16 cycles

resulted in selection of optimal conditions of T_A = 60 °C and 16 cycles, which were confirmed by running PCR2 with various primer combinations in duplicates (Figure 3.8).

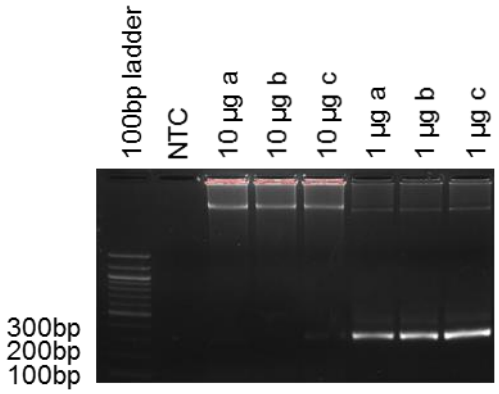


Figure 3.5: First step in PCR1 cycling conditions optimisation.

290bp band was observed only for samples with 1 µg template DNA and not 10 µg template DNA after 24 amplification cycles and T_A = 50 °C. PCR products were run on a 2% agarose gel with ethidium bromide at 80V for 45 minutes and visualised using BioRad Gel Doc™ XR+ Imaging System.

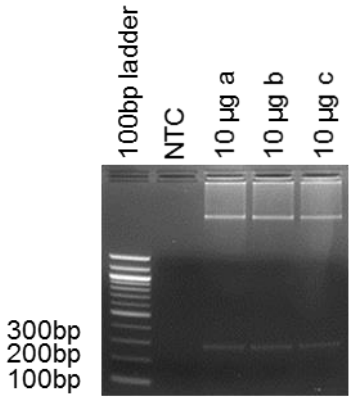


Figure 3.6: Optimised PCR1 cycling conditions using pilot DNA- annealing temperature and number of cycles.

10µg of DNA was loaded on a gel in triplicate, run for 22 cycles with T_A = 51 °C and correct band of 290bp observed for each replicate. PCR products were visualised by electrophoresis using 2% agarose gel with ethidium bromide.

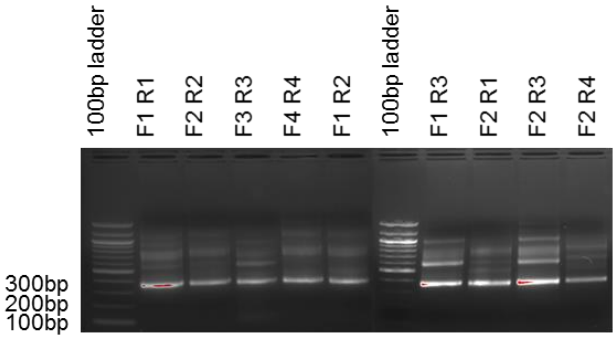


Figure 3.7: First step in PCR2 cycling conditions optimisation.

20 cycles and T_A = 60 °C allowed for all tested primer pairs amplified the correct band size (approximately 350 bp), however many non-specific bands were also observed. Removal of non-specific product can be done by lowering T_A and/or reducing the number of cycles. PCR products were visualised by electrophoresis using 2% agarose gel with ethidium bromide.

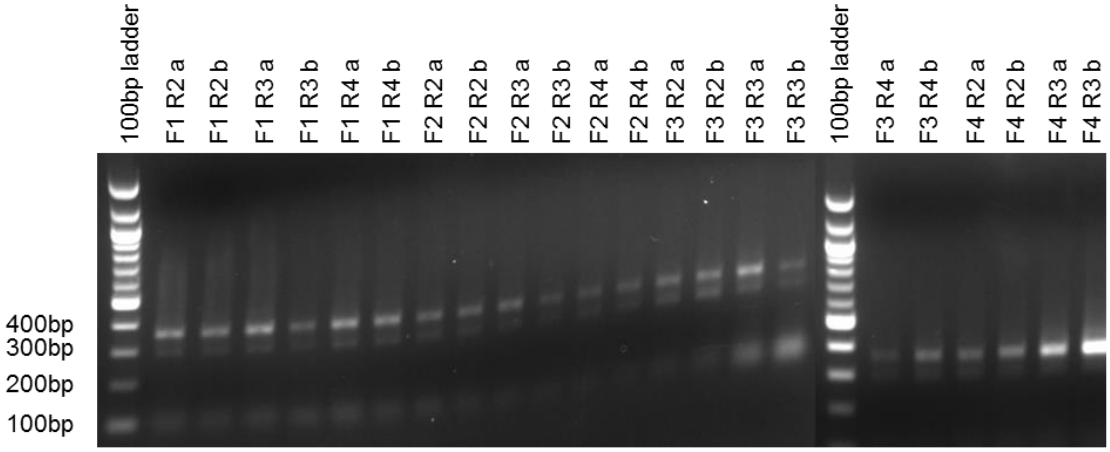


Figure 3.8: Optimised PCR2 cycling conditions using pilot DNA.

Each sample was run for 16 cycles with T_A = 68°C. Band of correct size (approximately 350 bp) was observed for all samples run with unique primer combination in duplicates. PCR products were visualised by electrophoresis using 2% agarose gel with ethidium bromide.

3.6. Summary of optimised conditions of GeCKOv2 screen

Table 3.3: Summary of optimised conditions of GeCKOv2 screen.

Parameter	Selection
Cell line	786-0
Puromycin concentration	2 µg/ ml
MOI	0.3
Viral titre	$3.3x10^6$ particles containing half library A and B, each
Recipient cell number	1.3×10^8
Recipient cell seeding density	3.25×10^6 cells per T175 flask
CB-839 concentration	0.7 µM
PCR1 cycling conditions	22 cycles, T_A = 51 °C
PCR2 cycling conditions	16 cycles, T_A = 68 °C

3.7. Discussion

Whole-genome knockout CRISPR/Cas9 screening requires careful selection and optimisation of several parameters before embarking on performing such complex experiment. Due to the multitude of steps, which have to be carried out in a specified order and within specified timeframe, there are multiple sources of errors, which can affect the success of candidate gene discovery and emergence of false positive hits. In order to minimise risks of errors, this first results chapter described experiments that were carried out and their outcomes.

There are many cell lines available for researchers whose focus is ccRCC. However, these cell lines are characterised by differences in their source of origin (primary or metastatic site), *VHL* status, other mutational profile and how they behave in culture in terms of doubling time and feasibility of infection. Taken into consideration all of these parameters, 786-0 cells were chosen to be used in the screen. They carry homozygous mutation in the *VHL* gene, which is characteristic for ccRCC cases found in the clinic, as well as several other mutations (Table 3.1). Moreover, they have been extensively used in ccRCC research and were found to be easily cultured, with 26.5h doubling time.

Selection of appropriate drug concentration used in selection is crucial for significant hit discovery. Too little drug will result in survival of most cells regardless of what gene is knocked down. On the other hand, very high inhibitory doses will kill most cells and result in false negative hits. Therefore, we decided to first establish drug concentration equal to IC_{50} and then use a higher concentration of 0.7 μM to aid hit discovery and minimise chance of both false positive and false negative hits. This was first calculated using SRB assay performed on 96-well plate scale and then translated to dimensions correct for

T175 flasks which will be used in the screening process. CB-839 concentration would be used at 0.7 μM for cells seeded at 3.25 x 10^6 cells per T175 flask.

As CRISPRLentiv2 plasmid encodes puromycin resistance gene alongside other elements, this feature is used to select successfully infected cells. All non-infected 786-0 cells were found to be killed by 2 μg/ ml puromycin. Next step was to determine to correct number of viral particles needed to infect each cell with no more than 1 sgRNA. Although it is impossible to be certain that a single sgRNA was incorporated into a single cell genome as infection process cannot be controlled to this extent, using MOI of 0.3 is generally recommended.

Amplification of desired fragments of DNA is another important step in preparation of samples (libraries) for sequencing. DNA has to be isolated and purified using reliable and tested kit, such as Endo-free kit from QIAGEN. DNA is then quantified and it was found that QubiT 2.0 fluorometer, which is based on dyes that specifically bind dsDNA, is more reliable method than using Nanodrop, which is based on UV spectroscopy [312]. However, Nanodrop is effective in detecting contamination in the sample and therefore it was used for this purpose, following the use of QubiT 2.0 fluorometer for accurate DNA quantification. Two-step PCR ensures that only the region of interest, with additional barcode, is amplified. It is important to determine the lowest number of cycles required to amplify template in PCR1, because higher values may affect PCR2 and result in non-specific products and higher rate of false positives [313]. PCR1 was optimised to be run using annealing temperature of 51 and 22 cycles. PCR2 was optimised to be run using annealing temperature of 68 and 16 cycles.

Taken together, optimisation of the above parameters was necessary and sufficient to perform the CRISPR/Cas9 screen in 786-0 cells using GeCKOv2 library. Optimisation of quality control methods, such as AMPure bead DNA purification, Bioanalyzer chip electrophoresis and KAPA kit qPCR will be discussed in the next chapter.

4. Whole-genome CRISPR/Cas9 knock-out in 786-0 cell line

4.1. Introduction

ccRCC is characterised by high resistance to currently available therapy (see section 1.3). Either intrinsic, or present from the beginning of the treatment, and acquired resistance, which arises after initial response to drugs, are a major obstacle in successfully eradicating tumours in patients. Despite significant progress in our understanding of mechanisms of resistance, patients with advanced tumours still require multiple lines of treatment and administration of drug combinations to combat their tumours. It is therefore of great importance to create drug combinations that are most likely to succeed as treatment and to minimise adverse effects of multiple drugs that patients are exposed to in order to find the one that works in their particular case.

Mutations in specific genes or pathways are an important clue as to why a certain drug is ineffective. However, certain inactivating mutations may not be obvious to cause resistance and whole-genome *in vitro* CRISPR/Cas9 screens allow to look at essentially every single gene and its impact on drug resistance (see section 1.6.3). This approach requires no previous knowledge and is a tool used to discover new druggable genes and pathways, which may be exploited to inhibit tumour growth and metastasis.

The gold standard therapeutics used to treat advanced ccRCC are drugs targeting pathways involved in angiogenesis, mTOR inhibitors and, more recently, immune checkpoint inhibitors. CB-839 (Telaglenastat), is a small molecule inhibitor of glutaminase, enzyme involved in first step of glutamine metabolism. While it was not the first glutaminase inhibitor to be studied, it is the only one that has shown favourable safety profile and is currently in Phase I/II clinical trials alone or in combination with previously approved therapies. It is being tested in multiple settings, including ccRCC,

but little is known about the potential resistance to CB-839. However, given high frequency of acquired resistance to other classes of drugs in ccRCC, it is reasonable to assume that it will also arise upon CB-839 administration.

The aim of the experiments described in this chapter was to create a whole-genome CRISPR/Cas9 knockout library in 786-0 cells, expose them to CB-839 and assess which genes when knocked out confer resistance to the drug. This approach has been shown to successfully identify genes and pathways, which are involved in resistance to various classes of drugs in various cancer types (see section 1.6.3) [221, 237]. There are various pooled libraries available for researchers to choose from, with GeCKOv2 being one of the most commonly used ones (see section 1.6.4). As described in Chapter 3, there are several steps, for which optimisation is crucial to minimise bias and enhance discovery of hits. In addition to library preparation, this chapter also describes optimisation of quality control steps necessary for robust data sequencing and analysis.

Sequencing of such immense number of genes requires special facilities, which routinely handle large number of samples. Samples were sequenced and data analysed by Dr Frances Turner at Edinburgh Genomics facilities, which is part of the University of Edinburgh, Scotland UK. Edinburgh Genomics works with collaborators across industry, government and academia, and offers sequencing using various platforms, including massively parallel sequencing using Illumina technology, which was used in this project. In addition to outsourcing sequencing of samples, Edinburgh Genomics also analysed the results using MAGeCK algorithm (see section 1.6.6) and generated all the figures included in the data analysis section of this chapter.

4.2. Experimental design and timeline

In order to produce robust data, the screen was performed in two biological replicates. This means that all steps starting from virus production, all the way to final (day 21 of treatment) DNA collection, were performed twice and independently from one another. Following completion of the above steps, DNA samples were processed to minimise impact of varying conditions that could potentially affect quality of samples, as well as contamination and other aspects of samples that may be influenced by their processing at time points, which were far apart.

Figure 4.1 shows timeline and workflow of the screening process. 786-0 cells were infected with virus particles combined from two half libraries, A and B. Lentiviral particles carried a plasmid with a single sgRNA insert, as well as sequence encoding Cas9. Recipient cells were then selected with puromycin to ensure only those with the plasmid survive. Next, baseline DNA, or DNA isolated from cells that were not exposed to CB-839 or DMSO, was collected. Remaining cells were treated with 0.7 μM CB-839 or equal volume DMSO (same v/v % concentration as that used in samples treated with CB-839, see section 2.1.4). DMSO is a solvent for the powder CB-839 and can have effects on cells itself, therefore it was important to distinguish these effects from those caused by the drug. DNA was collected again after 16 days of treatment, with remaining cells treated for further 5 days before final DNA collection. DNA was then purified, sequences corresponding to sgRNAs amplified and sequenced. Data produced by NGS was then analysed using MAGeCK algorithm to identify genes, which when knocked out confer resistance to CB-839.

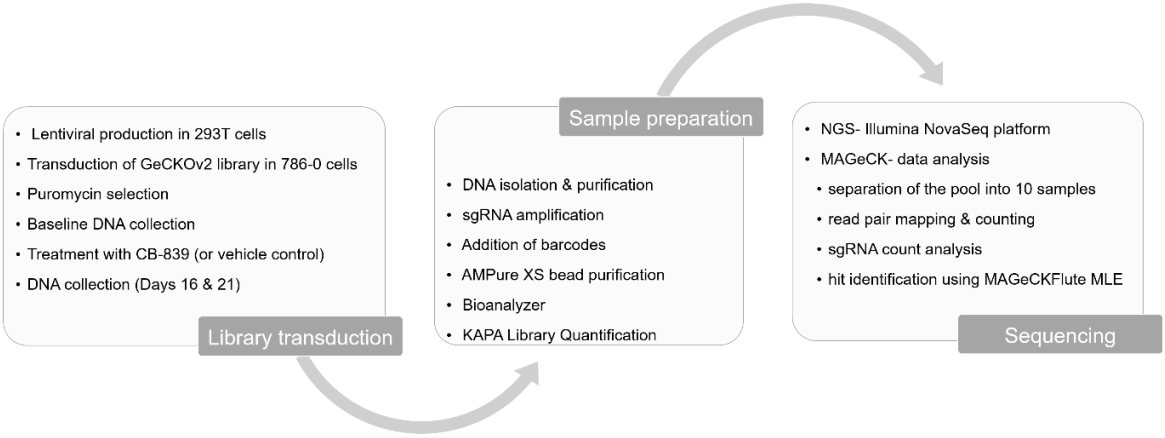


Figure 4.1: CRISPR/Cas9 screen workflow.

4.3. Transduction of GeCKOv2 library in 786-0 cells and selection

Due to the large number of cells, as well as their rapid growth, following transfection (see section 3.4.2), puromycin selection had to be performed in batches. 2 mg/ml puromycin dose was established experimentally (see section 3.4.1). First, all cells were trypsinised and seeded across 20 T175 flasks and treated with puromycin for 2 days. Puromycin works best on unattached, floating cells, but also its effects are often first seen after 3 to 4 days. Therefore, cells were still growing at a rapid rate and after two days were close to reach confluence. To avoid splitting cells into unmanageable number of large flasks, cells from 15 flasks were collected and frozen in -80 °C in three 50 ml falcons. Remaining cells from 5 flasks were plated across 20 new flasks and puromycin selection was continued for further 5 days without cells reaching confluence. Puromycin-selected cells were then frozen and the same process repeated for each of the previously frozen cells, one falcon at a time. After all cells were exposed to puromycin for 7 days, all cells were pooled and counted. To minimise the loss of sgRNA representation in the library, it was crucial to collect a minimum of 6 x 10^7 cells for baseline DNA. Thanks to the large numbers of cells, 4 tubes containing 6 x 10^7 cells, each, were collected and frozen before

future processing and back up cells in case of problems with, for example, clogging of DNA purification columns. Of the remaining cells, 10.4×10^7 were frozen in liquid nitrogen as a backup and 5.2×10^7 were plated across 16 T175 flasks for treatment (8 flasks for CB-839, 8 flasks for vehicle control).

The next day, after overnight attachment, cells were treated with the drug or vehicle control, as described above. After 72 h of treatment (1st treatment cycle), all cells from a given condition were collected and counted and again, 2.6×10^7 cells were plated for each condition and treated the same day (2nd treatment cycle). Remaining cells discarded. Collection, counting and plating was repeated again after 72h, but cells were left to attach overnight before treatment (3rd and 4th treatment cycles). After completion of the 4th treatment cycle, cells were trypsinised and plated across 10 flasks per condition and left for 2 days to allow enough cells to grow before DNA collection on day 16 of treatment. As expected, fewer cells were present after treatment with CB-839 in comparison to DMSO only treatment, however for each biological replicate experiment, sufficient cell pellets were frozen. The same day, remaining cells were plated and treated (5th treatment cycle) and after 72 h trypsinised and plated as described following 4th treatment cycle. On day 21 of treatment, cells were again collected and counted and pellets frozen for DNA processing. The number of cells collected in each of the two biological replicates after puromycin selection, Day 16 and 21 of treatment, together with number of cells pelleted for DNA for each timepoint can be found in Table 4.1.

Table 4.1: Number of cells and cell pellets generated for each timepoint of the screen.

	Replicate 1	Replicate 2
Baseline	47.5 x 10^7	41.1 x 10^7
Day 16 CB-839		
- Total	- 11 x 10^7	- 17 x 10^7
- Frozen pellet	- 8.4 x 10^7	- 14.4 x 10^7
Day 16 DMSO		
- Total	- 21.3 x 10^7	- 31.2 x 10^7
- Frozen pellet	- 18.7 x 10^7	- 24 x 10^7
Day 21 CB-839		
- Total / frozen pellet	- 14.3 x 10^7	- 13.6 x 10^7
Day 21 DMSO		
- Total / frozen pellet	- 24.1 x 10^7	- 24 x 10^7

4.4. Preparation of DNA samples

4.4.1. Genomic DNA purification

After completion of both biological replicates, genomic DNA from all samples was isolated using Blood & Cell Culture DNA Midi Kit as outlined in section 2.3.3. In order to minimise extraction variations between biological repeats, corresponding samples from both biological replicates were purified in parallel, for example Baseline 1 and Baseline 2. During elution of DNA using DNA columns, in some cases columns became blocked causing slow elution and very low concentration of eluted DNA as a result. However, as mentioned in section 4.3., all pellets were frozen in multiples of two or more and therefore more pellet could be purified to account for loss of DNA during purification step. Final concentrations of genomic DNA, together with additional information required for further processing can be found in Table 4.2.

Table 4.2: DNA quantity purified from each of 10 cell pellets.

Sample ID with number 1 refer to pellets from 1st biological replicate of the CRISPR/Cas9 screen, 2 – 2nd screen.

sample ID	QubiT conc µg/ml	Nanodrop 260/280	volume µl	totalDNA µg	vol (µl) for 10 µg	vol (µl) for 260 µg
BASELINE 1	1800	1.89	580	1044	6	144.4
BASELINE 2	810	1.92	580	469.8	12	321.0
DMSO D16 1	1000	1.88	290	290	10	260.0
DMSO D16 2	1260	1.9	580	730.8	8	206.3
CB-839 D16 1	930	1.9	390	362.7	11	279.6
CB-839 D16 2	780	1.89	390	304.2	13	333.3
DMSO D21 1	1520	1.9	290	440.8	7	171.1
DMSO D21 2	1220	1.89	580	707.6	8	213.1
CB-839 D21 1	1720	1.9	290	498.8	6	151.2
CB-839 D21 2	950	1.89	290	275.5	11	273.7

4.4.2. DNA amplification

Yields of genomic DNA were assessed by QubiT 2.0 fluorometer and were within the range 275.5 to 1044 µg for all samples. Additionally, Nanodrop spectrophotometer was used to assess sample purity and 260/280 ratio for all samples was between 1.89 and 1.92, which is expected for DNA (Table 4.2). Regions of the genome corresponding to sgRNAs and their flanking sequences were then amplified using conditions optimised and described in section 3.4.3. However, cycling conditions for both PCR steps had to be re-optimised after testing them with 10 µg template from each of 10 samples. PCR1 amplified some, but not all templates (Figure 4.2a). In order to account for this, number of cycles was increased, but annealing temperature was not changed and therefore final cycling conditions were: 24 cycles and T_A = 51 °C (Figure 4.2b). Number of cycles also had to be increased for PCR2, due to absence of bands after amplification of aforementioned PCR1 products (Figure 4.3a). Re-optimised cycling conditions for PCR2 were: 18 cycles and T_A = 68 °C (Figure 4.3b).

Following re-optimisation of cycling conditions, a total of 260 µg of template DNA from each sample was amplified in 13 separate reactions for each sample. DNA isolated from corresponding timepoints was amplified at the same time. As shown in Figures 4.4 and 4.5, all band of expected size (290 bp) was present for reactions containing DNA corresponding to Baseline 1, Baseline 2 and DMSO D16 2. However, several reactions showed contamination of NTC (band of the same size as expected for samples containing template DNA) (Figures 4.6 and 4.8). However, due to insufficient template DNA to repeat PCR1, these samples were used (CB-839 D16 1 and 2, CB-839 D21 1 and 2) (Figures 4.6, 4.8). Where sufficient template DNA was available, additional reactions were run and used as replacement in pools of amplicons (DMSO D16 1, CB-839 D16 2, CB-839 D21 1) (Figures 4.5a, 4.6b, 4.8a).

The next step was to test unique primer combinations for their ability to add barcodes and amplify PCR1 products (Figure 4.9). Despite all primer combinations producing band of correct size (approximately 365 bp) in this initial test, several primers had to be swapped later due to lack of 365 bp peak on Bioanalyzer traces. Illumina primer combinations used in PCR2 can be found in Materials and Methods chapter, Table 2.8.

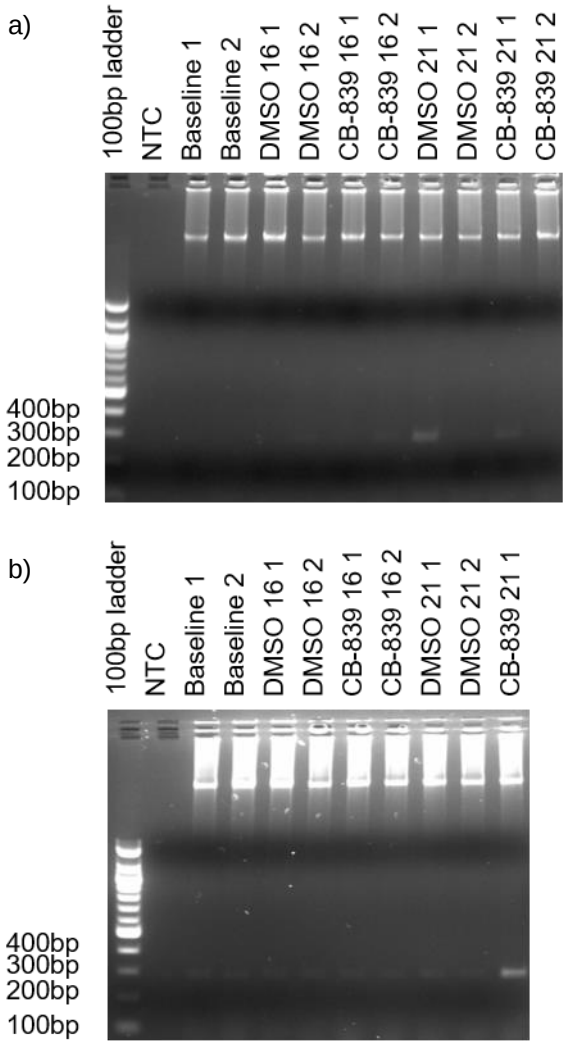


Figure 4.2: PCR1 test of samples from the screen.

PCR products visualised by electrophoresis using 2% agarose gel with ethidium bromide, amplified using PCR1 primers and DNA templates using previously optimised cycling conditions, 22 cycles, T_A = 51 °C (a) and re-optimised cycling conditions, 24 cycles, T_A = 51 °C (b). NTC- no-template control; marker- 100 bp ladder.

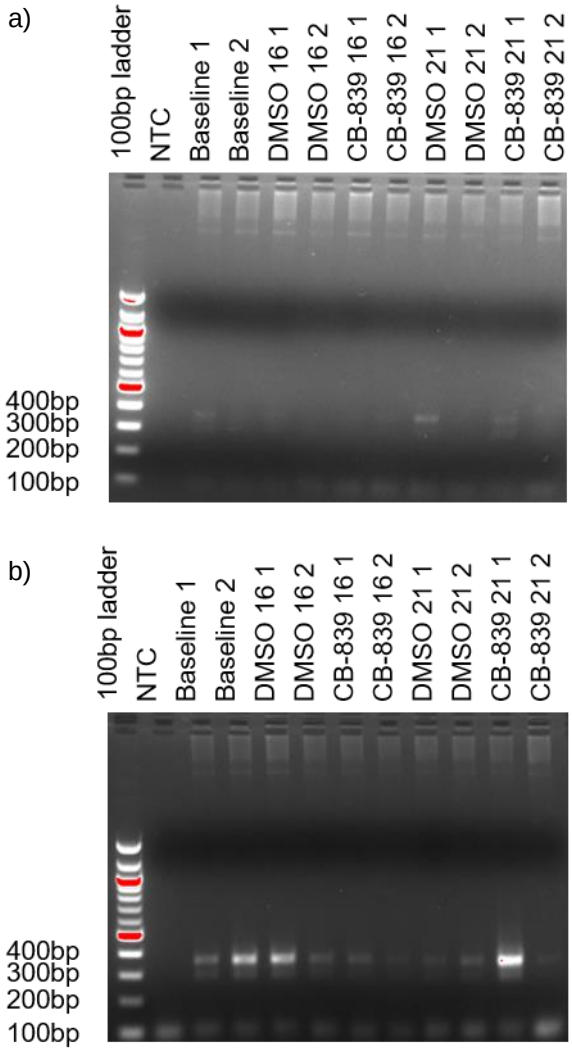


Figure 4.3: PCR2 test of samples from the screen.

PCR products visualised by electrophoresis using 2% agarose gel with ethidium bromide, amplified using PCR2 primers and DNA amplified in PCR1 using previously optimised cycling conditions. 16 cycles, T_A = 68 °C (a) and re-optimised cycling conditions, 18 cycles, T_A = 68 °C (b).

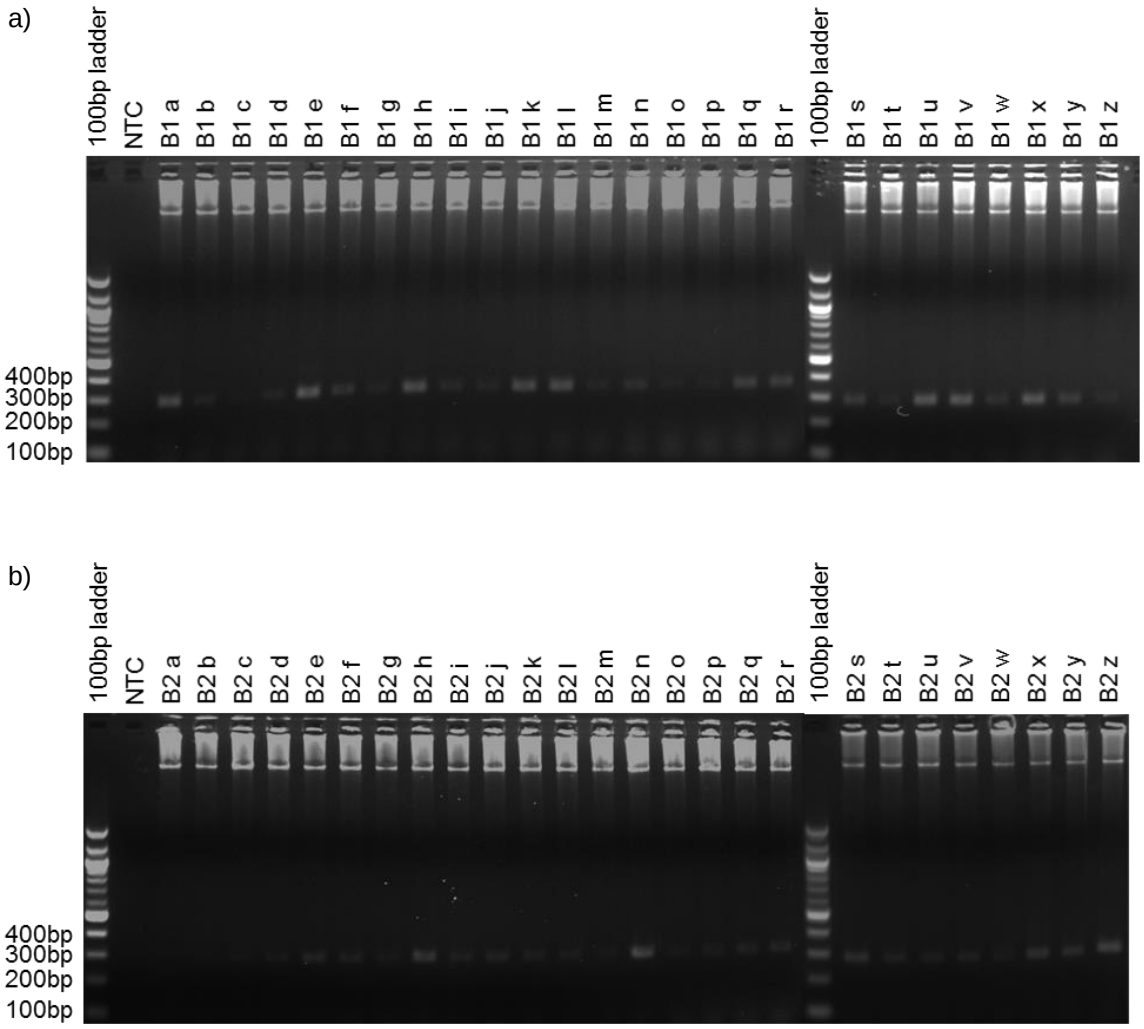


Figure 4.4: PCR1 of Baseline 1 and Baseline 2 samples from the screen.

PCR products visualised by electrophoresis using 2% agarose gel with ethidium bromide, amplified using PCR1 primers and DNA templates - Baseline day 21, replicate 1 (a) and 2 (b).

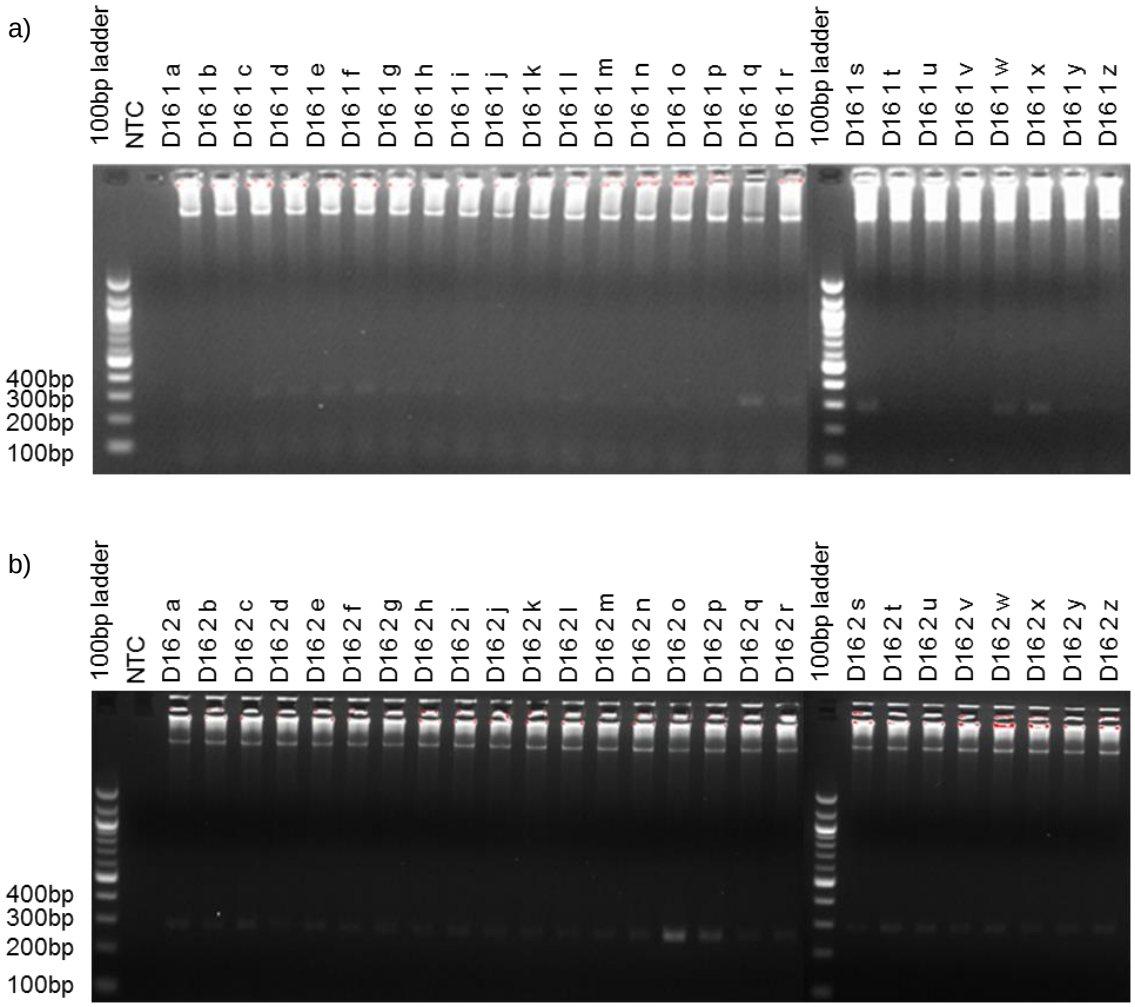


Figure 4.5: PCR1 of DMSO D16 1 and 2 samples from the screen.

PCR products visualised by electrophoresis using 2% agarose gel with ethidium bromide, amplified using PCR1 primers and DNA templates - DMSO day 16, replicate 1 (a) and 2 (b).

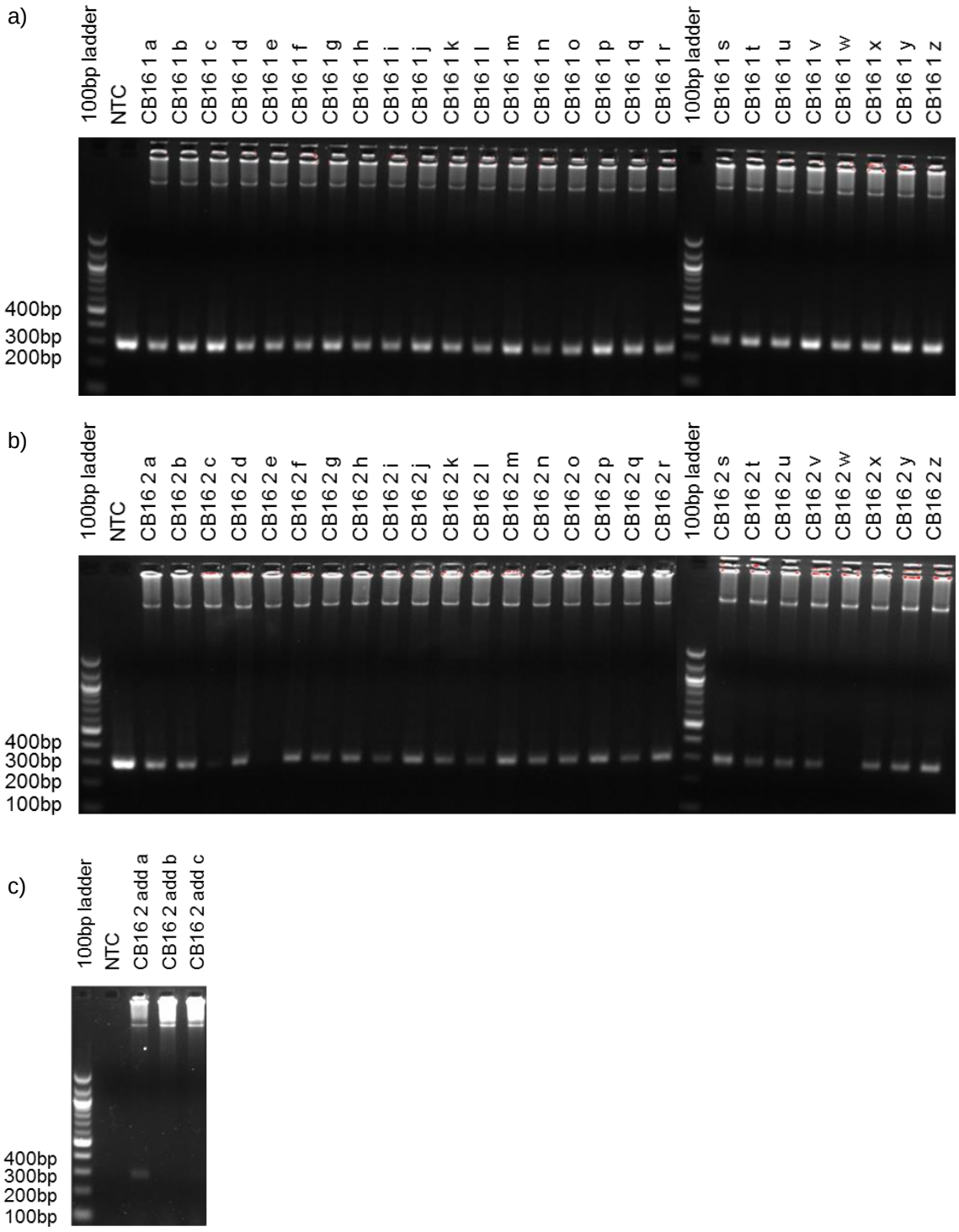


Figure 4.6: PCR1 of CB-839 D16 1 and 2 samples from the screen.

PCR products visualised by electrophoresis using 2% agarose gel with ethidium bromide, amplified using PCR1 primers and DNA templates – CB-839 day 16, replicate 1 (a) and 2 (b,c).

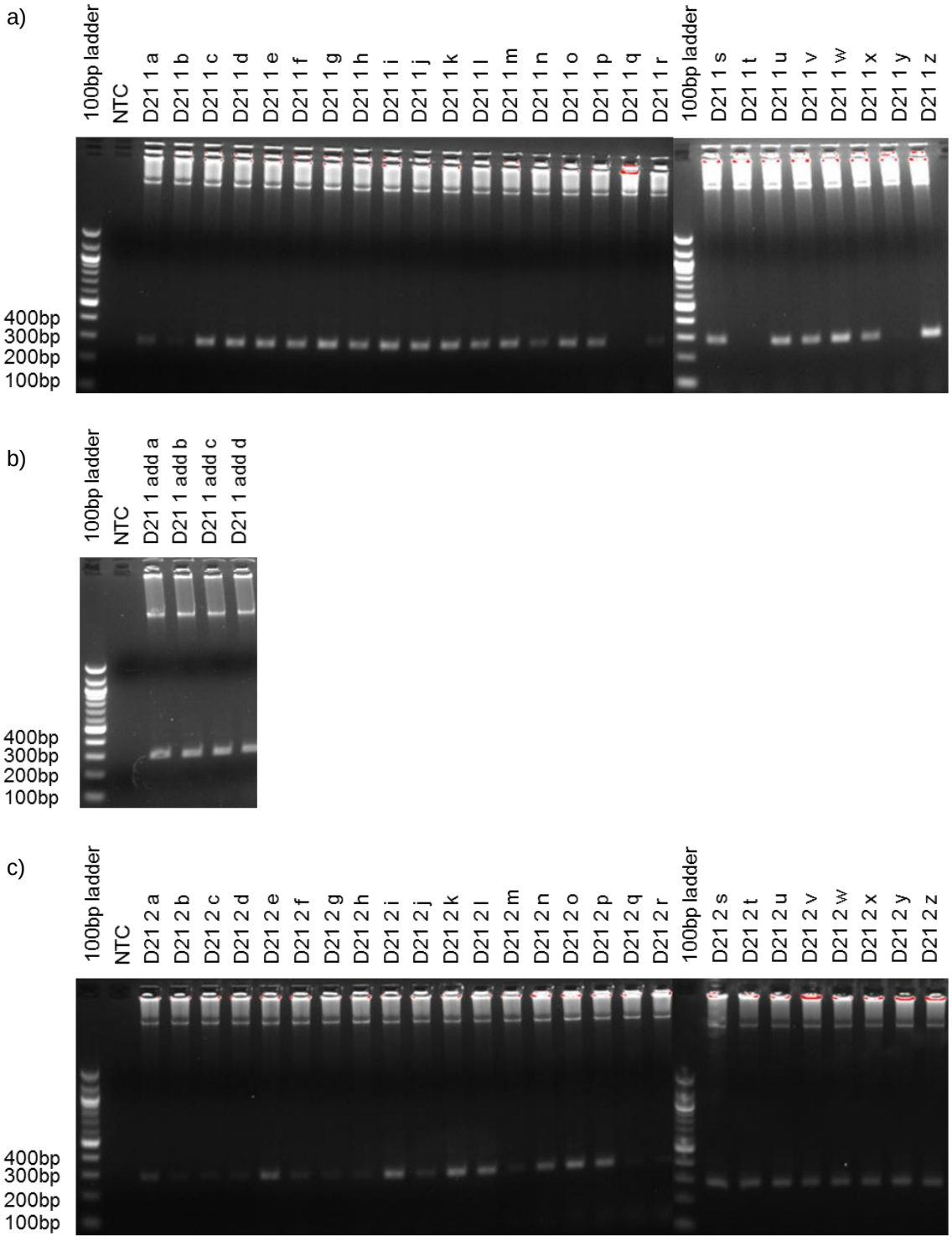


Figure 4.7: PCR1 of DMSO D21 1 and 2 samples from the screen.

PCR products visualised by electrophoresis using 2% agarose gel with ethidium bromide, amplified using PCR1 primers and DNA templates – DMSO day 21, replicate 1 (a,b) and 2 (c).

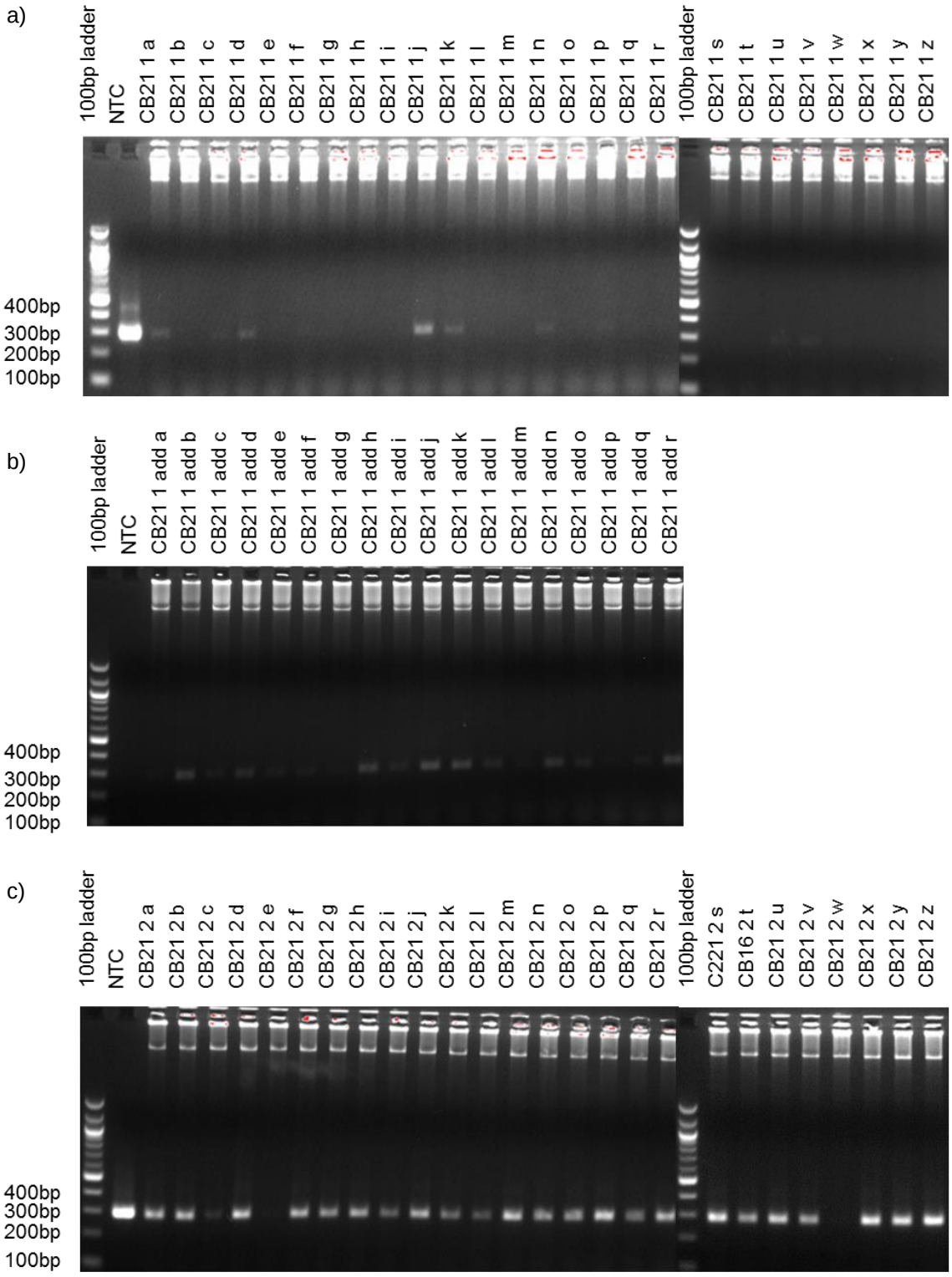


Figure 4.8: PCR1 of CB-839 D21 1 and 2 samples from the screen.

PCR products visualised by electrophoresis using 2% agarose gel with ethidium bromide, amplified using PCR1 primers and DNA templates – CB-839 day 21, replicate 1 (a, b) and 2 (c).

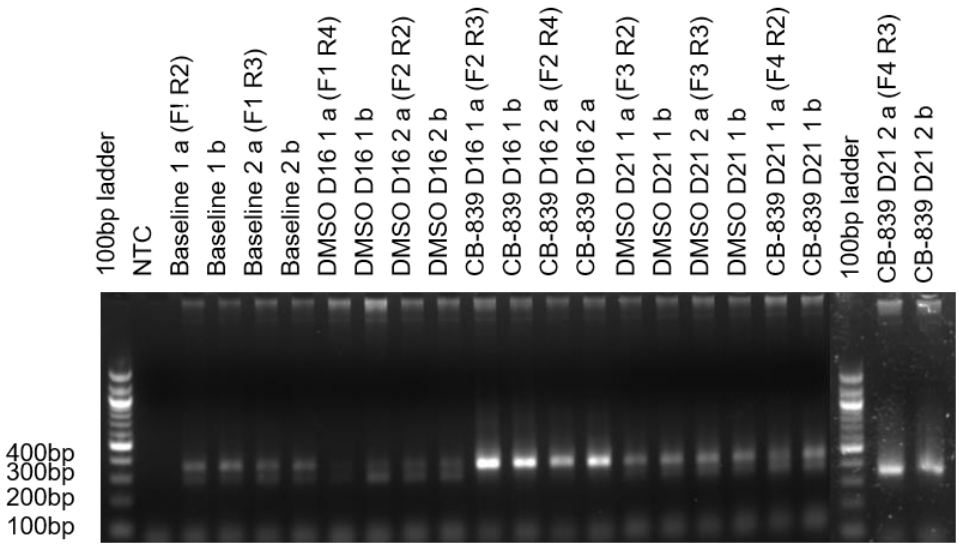


Figure 4.9: Test of PCR2 primer combinations.

Visualised by electrophoresis using 2% agarose gel with ethidium bromide. Additional ~100 bp bands could be explained by the formation of primer dimers or genomic DNA, which was carried, but not amplified in the previous run of PCR (PCR 1). Quality control checks described in the later part of this thesis showed that the DNA corresponding to these bands is cleaned by the bead purification and only DNA corresponding to the right size bands will bind to the Illumina flowcell for sequencing.

4.5. Quality control

4.5.1. AMPure XP bead purification

In order to ensure that the PCR libraries only contained sequences corresponding to sgRNAs, adapter and flanking sequences, a series of quality control methods were used. First, as mentioned in previous section, PCR products were purified using AMPure beads. This kit is recommended as a PCR cleanup kit for NGS applications due to high yield and purity of DNA, as well as relatively short time required to perform it. AMPure beads work by binding to DNA fragments larger than 100 bp and therefore removes any leftover primer dimers, which may have been left in the mix alongside amplicons as seen on Figures in section 4.4.2. Moreover, it removes unincorporated primers, dNTPs, salts and other contaminants. Figure 4.10 shows gel electrophoresis of PCR2 products after their purification with AMPure beads. Primer dimers observed after PCR2 were successfully removed as shown by the absence of bands smaller than 100 bp. Additional band of approximately 290 bp was present in all samples. It could potentially impact the sequencing, but could also be non-specific and not interfere with downstream processes. It was analysed further by other methods.

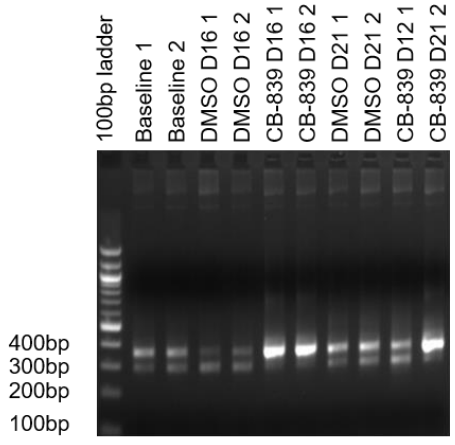


Figure 4.10: PCR2 products after purification with AMPure beads Visualised by electrophoresis using 2% agarose gel with ethidium bromide. Additional ~100 bp bands could be explained by the formation of primer dimers or genomic DNA, which was carried, but not amplified in the previous run of PCR (PCR 1).

Quality control checks described in the later part of this thesis showed that the DNA corresponding to these bands is cleaned by the bead purification and only DNA corresponding to the right size bands will bind to the Illumina flowcell for sequencing.

4.5.2. Bioanalyzer

To confirm presence of a single main product, PCR products were analysed using Agilent 2100 Bioanalyzer system, which is a chip-based capillary electrophoresis machine to analyse NGS samples following as few as four PCR cycles. This step had to be optimised and re-optimised several times due to several factors- too high concentration of DNA, which resulted in absence of a trace that could be correctly interpreted, multiple peaks pointing to the presence of multiple DNA fragments of varying sizes and mistakes in loading of samples and the marker. Moreover, as mentioned in section 4.4.2., some primer combinations used in PCR2 caused amplification of non-specific products resulting in additional peaks as shown on Figure 4.11. After careful troubleshooting, all 10 PCR library samples passed quality control using Bioanalyzer, with a single peak of correct (approximately 365 bp) size and absence of additional peaks (Figure 4.12). 365 bp is an approximate size for all the created libraries, however, it is not the absolute exact size due to varying lengths of flanking sequences added alongside barcode sequences (staggers of two to eight base pairs depending on the choice of Illumina primer combination, see Table 2.8, also Appendix).

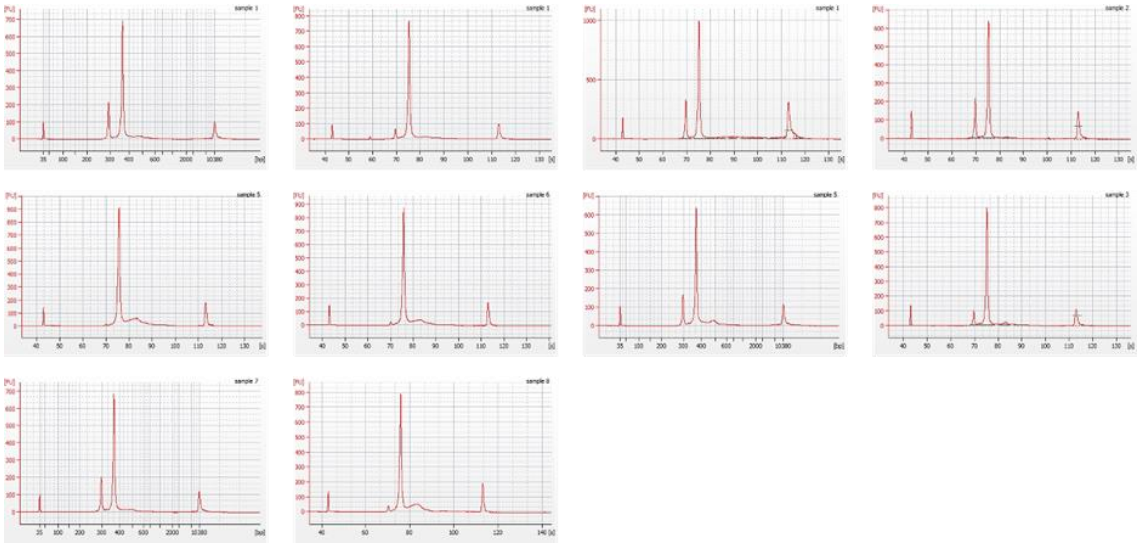
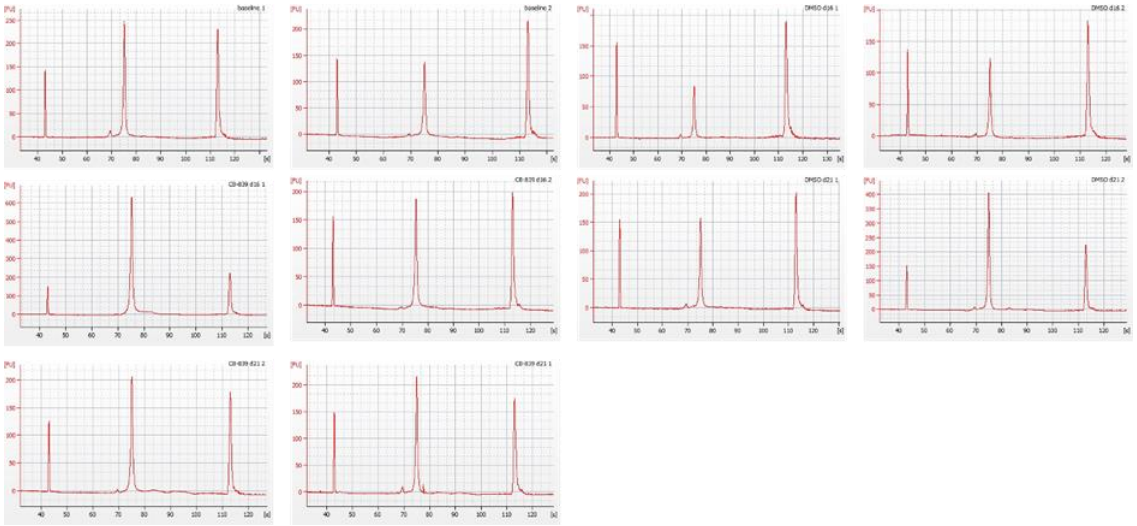


Figure 4.11: Quality control of purified samples using Bioanalyzer.

Peak of around 365 bp corresponds to the PCR library products. Additional peaks shown on electropherograms of samples 1,2,3,4,7,8 and 9, indicate additional, non-specific DNA fragments. Y- axis corresponds to fluorescence intensity (FU), x- axis corresponds to DNA size (bp).



.

Figure 4.12: Bioanalyzer traces after double purification of PCR products.

Lower DNA concentrations than in first run were used and therefore less fluorescence was absorbed, making the peak corresponding to the DNA fragments (PCR library) smaller in relation to upper and lower marker peaks than observed in previous figure. Y- axis corresponds to fluorescence intensity (FU), x- axis corresponds to DNA size (bp).

4.5.3. KAPA library quantification

qPCR-based KAPA Library Quantification kit is a high quality kit widely used for libraries before their sequencing using Illumina or Ion Torrent platforms. Six pre-diluted DNA Standards in concentrations ranging from 0.0002 to 20 pM and appropriately diluted NGS libraries (PCR product pools of sgRNA libraries) are amplified using platform-specific qPCR primers that target adapter sequences added to sgRNA sequences and thus quantifies all and only sequencing-competent fragments. Illumina-specific sequencing primer sequences are as follows: Primer P1: 5′-AAT GAT ACG GCG ACC ACC GA-3′; Primer P2: 5′-CAA GCA GAA GAC GGC ATA CGA-3′. Therefore, this system allows to confirm that non-specific sequences, which are found in the PCR mix, but lack adapter sequences added by PCR2, will not be detected and sequenced by an Illumina machine. The average cycle quantification (Cq) value for each DNA Standard was then plotted against its known concentration to generate a standard curve. The standard curve is used to convert the average Cq values for diluted libraries to concentration, from which the working concentration of each library is calculated.

As shown on Figure 4.13, all standards included in the kit clustered together and all PCR libraries clustered together into another peak, which means that each library consists of a single product in terms of size. No contamination was present as shown by lack of amplification of NTC samples. Therefore, despite additional bands as seen after AMPure bead purification of amplicons from PCR2, these are unlikely to affect sequencing and generated data since they are not amplified using KAPA Library Quantification kit.

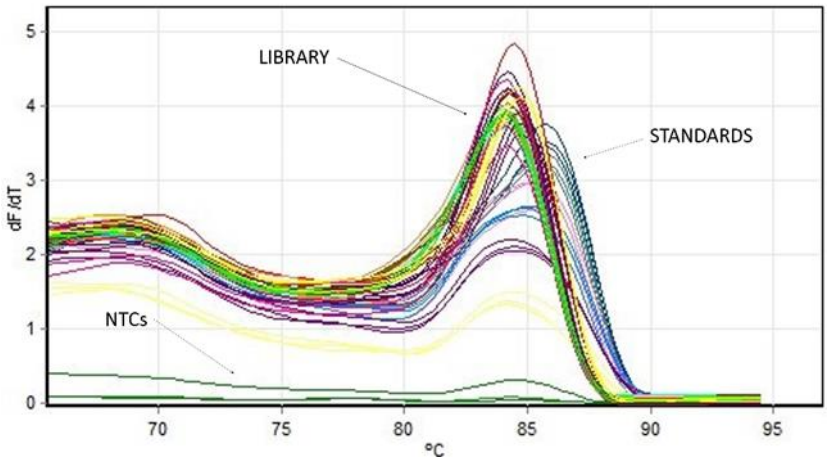


Figure 4.13: Melting curve generated by qPCR of 10 samples and standards from KAPA kit.

4.6. Next Generation Sequencing and data analysis

4.6.1. Raw data generation

Following the completion of all quality control steps described above, ten CRISPR sgRNA libraries corresponding to DNA collected throughout the screen, were prepared for next generation sequencing. First, each of ten samples were diluted to a concentration of 1.65 ng/µl in 10mM Tris buffer. 10 µl of each sample was then combined, giving the final volume of 100 µl, with estimated molarity of 6.9 nM of each library in the pool of 10 libraries.

Libraries were sequenced at Edinburgh Genomics facilities using NovaSeq 50 PE to yield approximately 375M paired reads. Following generation of sequencing data, four pools were separated into 10 samples based on inline barcodes at the beginning of read one. Reads were also trimmed to remove sequences flanking sgRNA, using cutadapt (version 1.9) and aligned to the sequences from GeCKO v2 reference library of the guide sgRNAs in the library using bowtie (version 1.2.2) [314, 315]. Only the first read in each pair was mapped, read two was discarded. Summary of read pair mapping and counting is shown in Table 4.4.

Table 4.3: Summary of read pair mapping and counting.

Read mapping	
Total read pairs	59.3 – 117 M
Reads uniquely assigned to sgRNA	54.3 – 107 M

4.6.2. sgRNA count analysis

In order to analyse generated data, MAGeCK pipeline was used by Frances Turner at Edinburgh Genomics (see section 1.3.3.). Batch correction and data QC was performed using the MAGeCK-VISPR pipeline. Testing for selection was performed on the batch corrected data using the MLE function of the MAGeCK-VISPR pipeline. To do that, a list of all the sgRNAs in the library that correspond to genes deemed to be non-essential was used. The test was performed to compare drug (CB-839) treated to control (DMSO treated) samples, and was performed once for each timepoint.

MaGeCKFlute was used to normalise the beta scores produced by MAGeCK MLE and compare the drug treated samples to the control samples at each time point [316]. MAGeCK MLE calculates beta scores for each gene in each condition (compared to the baseline). The higher the beta score, the more positively selected the gene. MaGeCKFlute identifies genes under positive or negative selection by comparing the beta scores calculated by MAGeCK MLE. The genes were put in to six groups (groups 1-4 are not mutually exclusive with A and B) (Table 4.5). MaGeCKFlute was also used to look for GO terms and Kegg pathways enriched among the drug selected genes. MaGeCKFlute was run once for each timepoint.

Table 4.4: Summary of numbers of genes found to be under different selection after normalisation with the essential genes in the drug treated versus control samples at each time point.

Group	Day 16	Day 21	
Group 1	159	302	Negative selection in control, no selection in drug treated.
Group 2	93	114	No selection in control, positive selection in drug treated.
Group 3	99	92	Positive selection in control, no selection in drug treated.
Group 4	500	555	No selection in control, negative selection in drug treated.
Group A	2,153	2,414	Higher beta score (more positive selection) in the drug treated samples than the control samples.
Group B	2,981	3,710	Higher beta score (more positive selection) in the control treated samples than the drug treated samples.

Important early step in data analysis was generation of plots that allowed to assess whether the complexity of library was not lost. Maintenance of library representation and complexity is required for accurate quantification and interpretation of generated sequencing data. As shown on Figure 4.14, less than 6% of library was below 100 reads, which is to be expected and consistent with original publication on this library. The aim was 300 reads per sgRNA and in both samples, Baseline 1 and 2, this was true for the vast majority of sgRNAs. Therefore his meant that the complexity of the library was preserved.

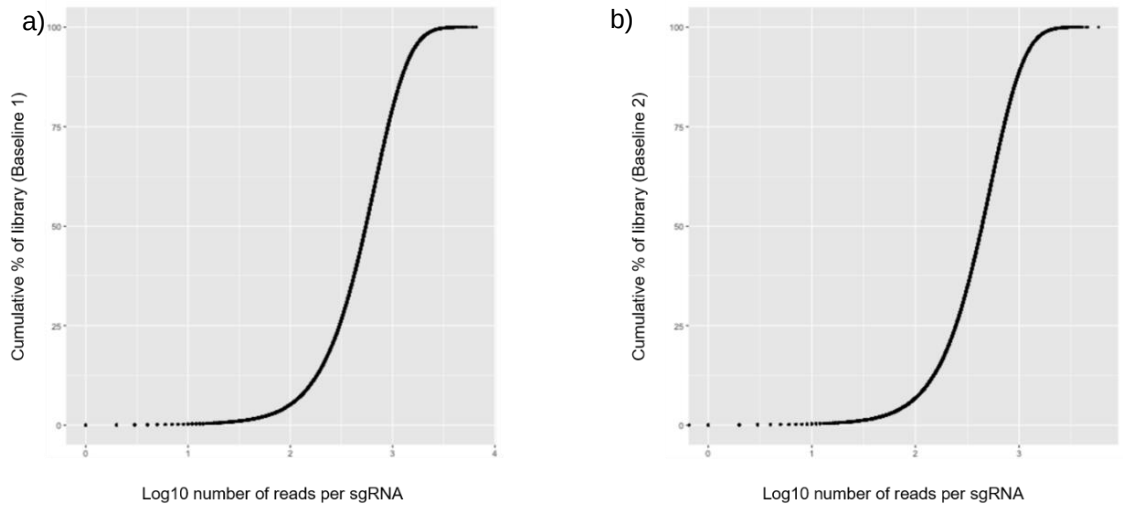


Figure 4.14: The cumulative portion of the libraries accounted for by sgRNAs with increasing read counts.

Baseline sample 1 (a) and Baseline sample 2 (b).

In order to confirm that certain genes when knocked out give cells growth advantage in the presence of CB-839, a boxplot comparing normalised read counts per sgRNA in each sample was generated (Figure 4.15). It is clear from the graph that the counts are comparable between corresponding samples from both biological replicates, but differ significantly across various timepoints and treatment groups. CB-839 was shown to impact sgRNA count when compared to Baseline and DMSO-treated samples. In conclusion, CB-839 has selective properties.

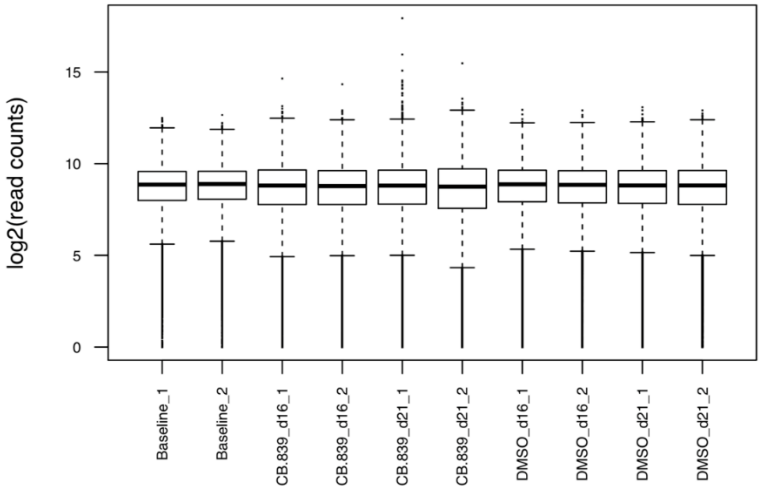


Figure 4.15: Boxplots of the normalized read counts per sgRNA in each sample.

To further confirm selection pressure due to CB-839 treatment, normalised sgRNA counts from CB-839 samples were plotted against those from DMSO only (Figure 4.16). As expected, knockout of most genes had no significant effect on cells' increased resistance or sensitivity, this can be concluded by most genes clustering along the straight line and within 5 % from the line. In other words, this shows no selection. sgRNAs, which are shown as single dots further up from clustered sgRNAs, and are therefore enriched, represent genes, which when knocked out confer resistance to CB-839. sgRNAs, which are shown as single dots further up from clustered sgRNAs, and are therefore depleted, represent genes, which when knocked out increase sensitivity to CB-839. Stronger selection was observed in Day 21 compared to Day 16 samples, as seen by larger number of single dots representing sgRNAs further away from those that cluster together.

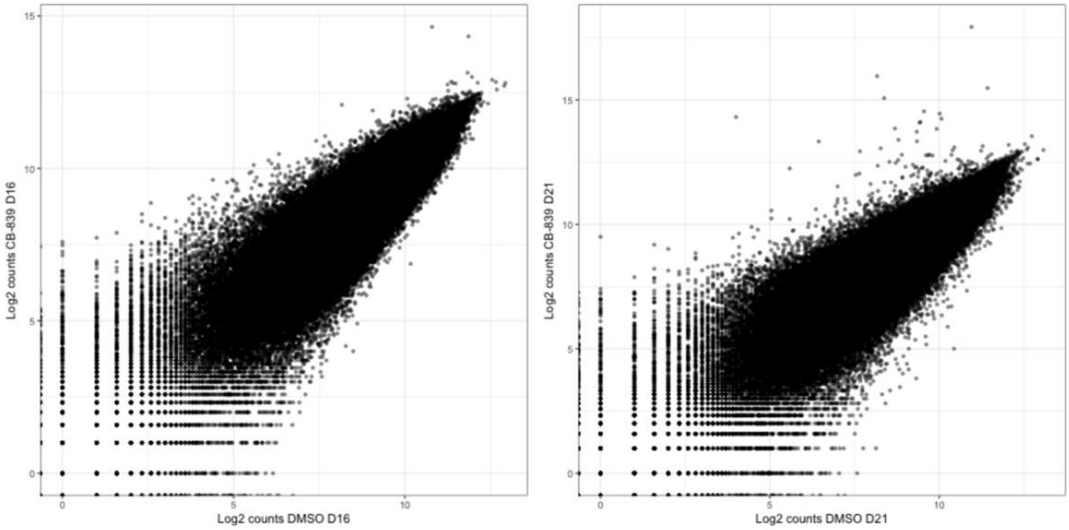


Figure 4.16: Scatter plot of log2 of normalised sgRNA counts in DMSO treated compared to CB-839 treated samples.

Day 16 (a) and day 21 (b).

4.6.3. Hit identification

In order for the hits that are related to the CB-839 treatment, to be identified accurately, FluteMLE assigns genes to four groups (Figure 4.17). Group 1 (green): Negative selection in control, no selection in drug treated- genes located in the pathways targeted by the treatment; Group 2 (orange): No selection in control, positive selection in drug treated-genes whose loss confers treatment resistance; Group 3 (blue): Positive selection in control, no selection in drug treated- regulators of cell proliferation in general, or regulators of the treatment target; Group 4 (purple): No selection in control, negative selection in drug-treated (genes conferring synthetic lethality in combination with the drug treatment). Assignment of a gene to its group is based on independent analysis of each of 6 sgRNAs targeting single gene and then all 6 sgRNAs were combined to produce ranking of genes. sgRNAs were normalised to essential genes. Knockout efficiency of each sgRNA as compared across all sgRNAs targeting the same gene are not available using MLE analysis. Additional file containing gene rank was used alongside data from

the above-mentioned scatterplots to compile the list of genes, which when knocked out confer resistance to CB-839 based on MLE analysis and can be found in Table 4.6.

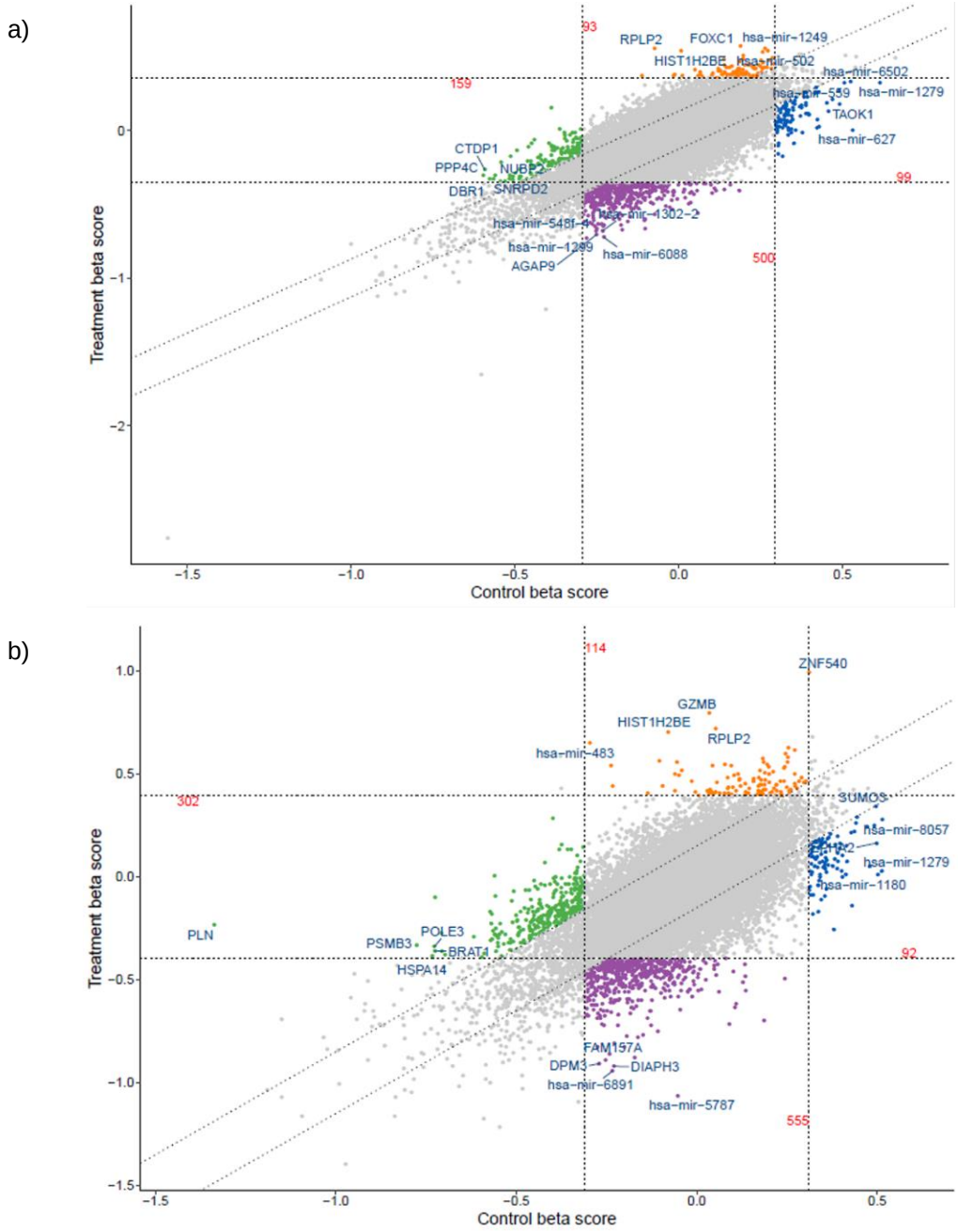


Figure 4.17: Scatter plots of essential gene normalised beta scores in DMSO treated compared to CB-839 treated samples.

Day 16 (a) and 21 (b), highlighting genes belonging to groups 1-4. Group 1 (green): Negative selection in control, no selection in drug treated- genes located in the pathways targeted by the treatment; Group 2 (orange): No selection in control, positive selection in drug treated genes whose loss confers treatment resistance; Group 3 (blue): Positive selection in control, no selection in drug treated- regulators of cell proliferation in general, or regulators of the treatment target; Group 4 (purple): No selection in control, negative selection in drug treated (genes conferring synthetic lethality in combination with the drug treatment).

Table 4.5: Ranking of top genes, which when knocked out confer resistance to CB-839, based on MLE algorithm.

Rank	Day 16	Day 21
1.	FOXC1	ZNF540
2.	hsa-mir-1249	GZMB
3.	RPLP2	RPLP2
4.	hsa-mir-502	HIST1H2BE
5.	HIST1H2BE	hsa-mir-483
6.	GLTSCR1	ALDH18A1
7.	ZDHHC12	FAM162A
8.	BIRC8	FOXO3
9.	CHRNA3	BIRC8
10.	KRTAP9-4	CCDC146

4.7. Discussion

Chapter 4 describes all steps required for the generation of the whole-genome CRISPR/Cas9 screen. It reports experimental procedures starting from transduction of the library in 786-0 cells, through application of selection pressure by treatment with CB-839, preparation of PCR libraries for sequencing, including several quality control steps, all the way to data generation and analysis using MAGeCK pipeline. It also details additional optimisation steps in addition to those described in Chapter 3. The screen, which aimed to identify genes, which when knocked out confer resistance to CB-839, was performed in two biological replicates. In order to accurately identify treatment-related hits, DNA was collected from cells at three timepoints. First timepoint after transduction of library and selection of cells, with incorporated plasmid carrying Cas9 and sgRNA sequences but before CB-839 treatment and serving as baseline. Second one after 16 days of CB-839 or vehicle control treatment. Final genomic material was collected after 21 days of CB-839 or vehicle control treatment.

Isolation of DNA from cells and its careful preparation for sequencing is a crucial step in any CRISPR.Cas9 screen. It is a complex process and requires not only use of highly-specialised DNA purification methods, including high quality kits for isolation of genomic DNA from cells. Moreover, quality of isolated DNA has to be assessed using several methods, such as spectroscopy that allows to confirm lack of impurities in the sample, but is not sensitive or reliable method for DNA quantification of samples for NGS. Measurement of fluorescence with intercalating dyes, such as that using QubiT fluorometer, is a preferred method used to quantify DNA in this setting. Functional test of PCR performance is another crucial step that ensures absence of impurities, such as PCR inhibitors, which may sometimes co-purify with DNA. All of the above mentioned methods were used and confirmed high quality DNA samples.

Despite the optimisation of the cycling conditions using pilot experiment DNA for both rounds of PCR as described in Chapter 3, those had to be re-optimised. Ten PCR libraries, each with specific barcoded added during second round of PCR, were subject to stringent quality control. Amplicon purification using AMPure beads was used to remove primer dimers, unincorporated primers, dNTPs, salts and other contaminants. This process did not remove PCR fragments, which appeared to correspond to amplicons from first round of PCR but lacked barcodes and therefore appeared smaller than expected PCR2 products. This, however, was shown not to interfere with the sequencing results as seen by Bioanalyzer tracer, were one strong peak was present, indicating presence of a single product (PCR library made of equal-size fragments made of sgRNA, barcode and flanking sequence). Additionally, qPCR-based library quantification with KAPA kit, which serves to mimic Illumina-based library amplification and therefore does not detect fragments with no attached barcodes, showed that all PCR libraries clustered together into a single peak, which means that each library consists of a single product in terms of size.

Following library preparation and completion of quality control steps, all ten PCR libraries were sequenced and data analysed at Edinburgh Genomics facilities. Data generated using MAGeCK pipeline showed conserved complexity and representation of library in both libraries corresponding to genomic DNA collected at baseline. Further analysis of sgRNA counts showed that CB-839 had an impact on sgRNA count when compared to Baseline and DMSO-treated samples, confirming selective properties of CB-839. Moreover, this selection pressure increases with the length of treatment (16 vs. 21 days of treatment).

FluteMLE workflow was then applied to rank all sgRNAs and then use this ranking to rank genes and assign hits into four groups. The aim of this screen was to identify hits involved in increased resistance to CB-839, however synthetic lethality was also part of the screen. Hits included in the two remaining groups have a higher chance of being selected as false positives and therefore were not analysed as part of validation process described in the next Chapter.

Knockout screens using pooled CRISPR/Cas9 libraries, such as GeCKOv2, require no previous knowledge of pathways affected by the drug used for selection and allow for an unbiased approach. Little is known about the mechanism of action of CB 839 other than its inhibitory role in the first step of glutamine metabolism that it achieves by allosterically blocking glutaminase. This drug has not been extensively tested in terms of its effect on other enzymes and pathways and there are not many published studies describing potential resistance mechanisms to it in vitro and to date there has been no whole-genome knockout screens described in the literature. Moreover, due to it being in clinical trials and not yet being an available treatment in ccRCC, datasets generated as a result of experimental proceedings described here, cannot be compared with any in vivo data. Additionally, CB-839 is a first in its class drug that has acceptable safety and toxicity

profile and therefore data presented in this chapter cannot be compared to that generated for other glutaminase inhibitors.

Despite the lack of data available generated in similar settings, the screen results show that a) CB-839 effectiveness on killing human kidney cancer-derived cells is affected by inactivating mutations in various genes, and b) the effect of these genes is either more prominent in earlier or later phases of exposure to the drug or knockout of the same gene confers resistance to CB-839 as tested on day 16 and day 21 (eg. RPLP2). Many of the identified hits are not well-characterised in literature.

The next chapter will focus on methods for selection and identification of the most viable and promising hits whose effect on conferring resistance to CB-839 will be validated using various *in vitro* techniques as the screening resulted in the robust set of data.

5. Validation

5.1. Introduction

Among their multitude of applications, large-scale CRISPR/Cas9 screens aided identification of genes involved in drug resistance. This technique has been used to identify targets involved in cellular proliferation, metastasis and drug resistance in a large variety of cancer cell lines [242, 317, 318]. Following sequencing, large sets of data are produced which then have to be analysed using bioinformatics approaches. Algorithms, which are often carefully designed for the specific library used to generate data, produce a list of ranked candidate genes. However, given the volume of data, the output can be noisy and lead to false positive and false negative hits.

Screens are usually performed with multiple sgRNAs targeting each gene of interest. In early years of the CRISPR/Cas9 technology, sgRNAs were designed based on a small number of rules and with little knowledge about sgRNA activity. It was then shown that some sgRNAs may be inactive and non-specific [319]. However, the presence of inactive guides in the library affect screen outcomes by reducing library coverage and creating false negatives. In recent years, multiple rules have been added to minimise non-functional sgRNAs and reduces off-target effects. Despite their on-target activity, each sgRNA varies in terms of their knockout efficiencies and using several sgRNAs increases the chance for the successful gene knockout. However, improved libraries still have a false negative rate between 10 and 20% [320]. False positive results also arise as a result of gene-independent DNA damage phenotypes following DSBs introduced by Cas9 nuclease in the amplified regions of the genome [321]. Cell line ploidy and gene copy number variation should also be taken into consideration. The higher the copy of the gene, the harder it is to knockdown or knockout all copies, so that the gene will most likely still be expressed from the unaltered copies. This in turn causes false negative

results. Moreover, DNA sequence and sequence accessibility may affect its editing [322]. Additionally, some DNA regions can be harder for nuclease to access due to the DNA packaging and chromatin state as well as epigenetic modifications. This will also affect editing efficiency and therefore sequencing results.

Therefore, it is important to validate that the perturbation identified by the screen in fact causes the phenotype of interest rather than it being an off-target effect. The first step in hit validation is recapitulation of a gene knockout by creating a cell line containing knockout pools with a single sgRNA. Then, knockout is validated by confirming protein loss by western blotting. After technical validation, the next step is to confirm cell phenotype resulting from gene knockdown and this is usually done by drug assays similar to those used in the original screen. As the observed phenotype may be specific to a cell line that the screen was performed in, it is important to recapitulate the findings in at least one more cell line.

5.2. Selection of hits for validation

Following the screen using GeCKOv2 library, MAGeCK algorithm generated gene rankings and the focus of the validation process were genes, which when knocked out confer resistance to CB-839. Full gene ranking was generated using the algorithm and beta scores in CB-839 samples (separately for day 16 and day 21) were arranged in descending order. This meant that the genes with the highest beta score were the ones that were enriched (positively selected) in treatment group when compared to the control group. One gene was also selected from group 4 (purple: the group of genes whose loss, in combination with CB-839 treatment, is implicated in synthetic lethality; i.e. no selection in control, negative selection upon drug treatment). Criteria for selection of hits for validation were as follows: PubMed search presence in one or both samples (d16 and

d21), association with formation of progression of RCC or other tumours, constituents of metabolic or apoptotic pathways, metastasis, EMT.

FOXC1 belongs to the superfamily of forkhead box (Fox) proteins, which are often associated with tumorigenesis and cancer progression [323]. Although its role in cancer has not been fully elucidated, it is thought to play a role in aberrant cell proliferation, cancer stem cell maintenance, angiogenesis and metastasis. Increased levels of FOXC1 mRNA and protein have been shown to be associated with poor prognosis in a variety of cancers [324-327]. Moreover, FOXC1 inhibition was shown to reduce cell migration and proliferation in NSCLC [328]. The findings of the role of aberrant expression of FOXC1 in renal cancer has not been published at the time of writing this thesis.

FOXO3 also belongs to the Fox family of transcriptional regulators [323]. It is a tumour suppressor as its downregulation is often observed in cancer development. Loss of FOXO3 was found to induce EMT and promote metastasis *in vitro* (including in 786-0 cells) *and in vivo* [329]. Moreover, it promotes Akt phosphorylation in response to NVP-BEZ235, which is a dual inhibitor of PI3K/mTOR pathways [330]. Upregulation of PI3K/Akt is a validated therapeutic target in RCC. Inhibition of FOXO3 makes cells more susceptible to NVP-BEZ235-mediated cell growth inhibition *in vitro* and tumour shrinkage *in vivo*. Therefore, FOXO3 is a potential therapeutic agent, which could synergise with PI3K/mTOR inhibition as a strategy to treat RCC. FOXO3 also could be used as a biomarker for cancer progression as its overexpression correlates with cancer cell proliferation [331].

RPLP2 gene encodes 60S acidic ribosomal protein P2 and forms a complex with RPLP1 in the large subunit of the ribosome, therefore playing an important role in the elongation step of protein synthesis. Its expression was found to be upregulated in various

gynaecologic tumours and pancreatic cancer cell lines [332, 333]. It is also associated with poor survival of patients with lung adenocarcinomas with bone metastases [334].

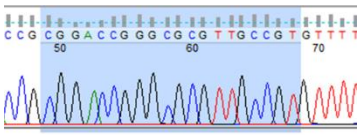
The aldehyde dehydrogenase 18A1 gene (ALDH18A1) encodes delta 1-pyrroline-5-carboxylate synthase (P5CS), which is involved in the biosynthesis of proline, ornithine and arginine [335]. Alongside PYCR1, ALDH18A1 is a key enzyme in the conversion of glutamate to proline [336]. In a recent study, which concluded that glutamine to proline conversion is associated with response to glutaminase inhibition in breast cancer, both ALDH18A1 mRNA and protein levels were significantly upregulated in CB-839 responding xenograft models vs the non-responding models [337]. ALDH18A1 expression was also found to be upregulated in HCC models and to promote HCC proliferation [338].

Diaphanous-related formin 3 (DIAPH3, mDia2) is a member of diaphanous subfamily of proteins, which are involved in actin remodelling and cell motility and adhesion [339]. It was shown that depletion of DIAPH3 was associated with cell shape instability and gain in malignant properties in prostate and breast cancer cells [340]. Moreover, DIAPH3 promoted the growth, migration and metastasis of hepatocellular carcinoma cells [341]. However, inhibition of DIAPH3 caused amoeboid properties, increased invasion and promoted metastasis in mice [342].

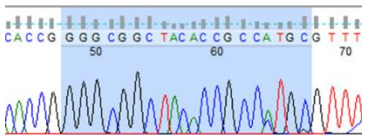
5.3. Creating and assessing knockdowns and their effects on cell viability

The first step in hit validation was to integrate sgRNAs into pLentiCRISPRv2 vector in order to generate stable single knockout cell lines. As mentioned earlier (section 1.6.8.), validation of most screens usually starts by directly replicating conditions (including technology) used in the whole-genome screen and then focuses on further validation using other techniques, such as siRNA. However, siRNA could be used first because it allows for faster confirmation of the effect of the gene knockout in comparison to creating stable knockouts. As mentioned before, MAGeCK MLE algorithm generates gene ranking, but does not generate sgRNA ranking for these genes. Therefore, one or more sgRNAs were selected at random from the six available sequences that are within GECKOv2 library (see section 2.4.1.). The selected sgRNAs were inserted into pLentiCRISPRv2 backbone as described in section 2.4 and inserts confirmed using Sanger sequencing. Sequence chromatograms confirming the presence of sgRNA inserts in the backbone are shown on Figure 5.1. Two individual knockdown plasmids containing sgRNAs targeting different parts of the coding sequence were used for FOXC1 as well as RPLP2 knockdown experiments. For each of the remaining candidates (FOXO3, ALDH18A1 and DIAPH3), a single sgRNA was selected from the library and used to produce knockdown cell pools. As the selected analysis method does not produce sgRNA ranking, sgRNAs were chosen at random.

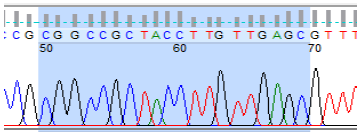
(a) Sequence of sgFOXC1(b) insert



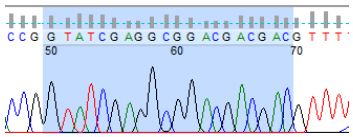
(b) Sequence of sgFOXC1(c) insert



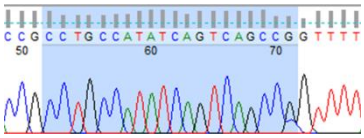
(c) Sequence of sgRPLP2(b) insert



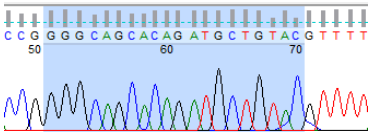
(d) Sequence of sgRPLP2(c) insert



(e) Sequence of sgFOXO3 insert



(f) Sequence of sgALDH18A1 insert



g) Sequence of sgDIAPH3 insert

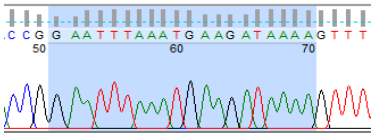


Figure 5.1: Chromatograms showing sgRNA sequence inserts in pLentiCRISPRv2 vector.

Highlighted (in blue) is sgRNA sequence selected from GeCKOv2 library for the following genes: (a,b) FOXC1, (c,d) RPLP2, (e) FOXO3, (f) ALDH18A1 and (g) DIAPH3.

Following the confirmation of successful construction of sgRNA-pLentiCRISPRv2 plasmids, these constructs were then used to generate a single gene knockout cell pools. Transduction of plasmids in 786-0 cells is described in section 2.4.10. Briefly, HEK293T cells were transfected with a plasmid containing Cas9, sgRNA and puromycin-resistance gene inserts as well as packaging plasmids. Produced virus was then used to infect 786-0 cells to generate knockout cell pools. Alongside cells transduced with a plasmid containing guide sequence targeting one of the genes of interests, second pool of cells was transduced with a plasmid containing scrambled sequence (sgScr) to act as a negative control during protein depletion detection. Cells, which were not infected were removed using puromycin selection.

Next, cells were harvested to confirm gene knockout by analysing protein depletion using Western Blotting. This method is one of the most commonly used techniques to assess knockdown efficiency. Protein levels in knockdown cell lines were compared to scramble-transduced cell lines, which served as a control. To ensure validity and consistency of results, all Western Blot samples were run in triplicates (technical repeats) per blot. Each protein was tested three times with protein collected at three different points (biological repeats).

In order to determine whether knockdown of a specific gene impacts cell survival when treated with CB-839, cells were seeded on 96-well plates and after 24h treated with increasing doses of CB-839. Then after 72h cells were fixed and density of such fixed cells was assessed using SRB assay, which is a colorimetric assay. Plates were then scanned using platereader. Dose-response curves were then created using log(inhibitor) vs. response (three parameters) analysis using GraphPad Prism. IC_{50} values were also calculated. This allowed for comparison of cell survival between scrambled and knockout cells and across selected drug concentrations. Additionally, untreated control (cells not

transfected with any version of sgRNA) could have been used to serve as an additional control for the cells transfected with any sgRNA.

5.3.1. DIAPH3

In order to study the potential synthetic lethal effect of DIAPH3 knockout when combined with the treatment with CB-839, 786-0 cells were transduced with the lentiCRISPRv2-sgDIAPH3 construct and designated as 786-0-sgDIAPH3. One sgRNA was chosen out of six from the GeCKOv2 library. As shown on Figure 5.2, knockdown could not be confirmed and therefore synthetic lethality avenue was not progressed any further.

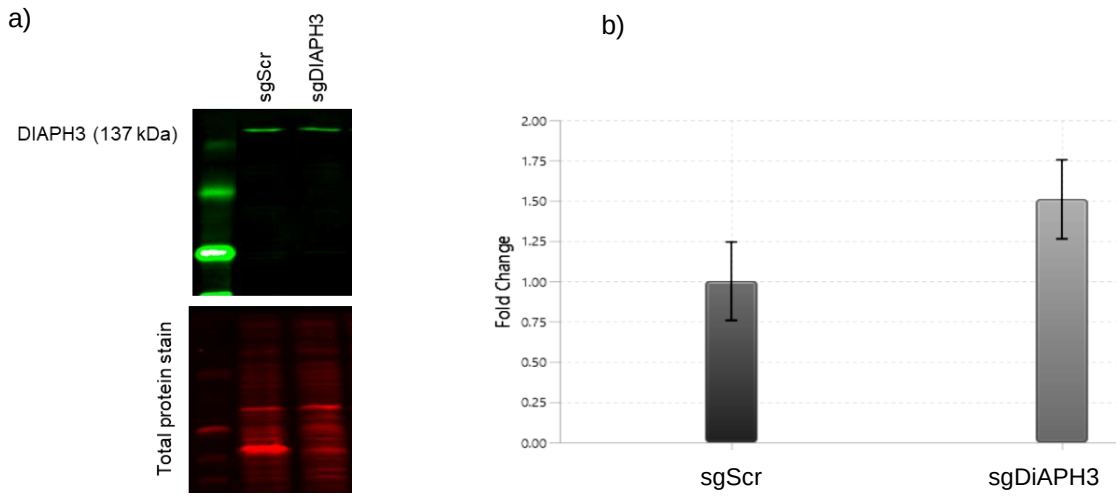


Figure 5.2: DIAPH3 protein expression in 786-0 cells after transduction with lentiCRISPRv2-sgDIAPH3.

(a) Western Blot of DIAPH3 protein in protein lysates collected from 786-0-sgScr and -sgDIAPH3 cells. Total protein staining was used for normalisation. (b) Fold change analysis calculated using Empiria Studio software and error bars indicate standard deviation for each replicate group (n=3).

5.3.2. FOXC1

In order to study the effect of FOXC1 knockdown on cell survival upon treatment with CB-839, 786-0 cell line, which had been used to generate whole-genome knockout data, as well as A498, selected as a viable second cell line to confirm gene knockout effect on cell survival while treated with CB-839, were used. The four knockout cell lines were designated as follows: 786-0-sgFOXC1b, 786-0-sgFOXC1c, A498-sgFOXC1b and A498-sgFOXC1c. Knockdown efficiency was analysed for all four cell lines using Western Blotting with β-actin (housekeeping gene) serving as internal loading control (Figure 5.3). Protein depletion could not be confirmed for any of the four cell lines and it was concluded that *FOXC1* knockdown was not achieved.

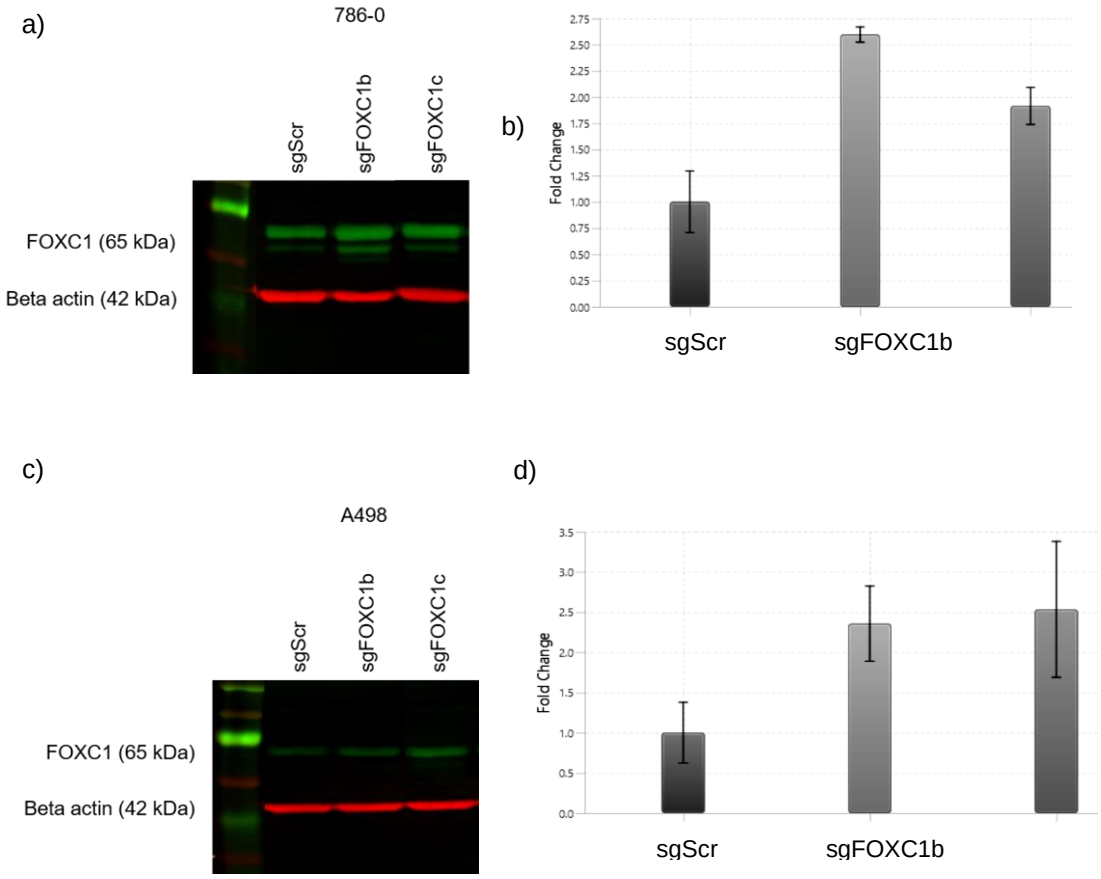


Figure 5.3: FOXC1 protein expressions in 786-0 and A498 cells after transduction with lentiCRISPRv2-sgFOXC1.

(a) Western Blot of FOXC1 protein levels in protein lysates collected from 786-0-sgScr, -sgFOXC1b and -sgFOXC1c cells. Total protein staining was used for normalisation. (c) Western Blot of FOXC1 protein levels in protein lysates collected from A498-sgScr, -sgFOXC1b and -sgFOXC1c cells. Total protein staining was used for normalisation. (b,d) Fold change analysis (vs –sgScr) was calculated using Empiria Studio software (n=3).

Following unsuccessful attempts at using three independent constructs to knock down FOXC1 in 786-0 cells, 786-0 FOXC1 KO cell line was engineered by Synthego using CRISPR technology (see Appendix – guide sequences, PCR and sequencing primers). The supplied report quoted editing efficiency after expansion at 97%. Edited cell pool was delivered together with matched passage (p2) WT cells. WT and KO cells were thawed and seeded as per Synthego's Quick Start Guide and protein isolated for analysis using Western Blotting. As shown on Figure 5.4, FOXC1 knockout could not be confirmed. Next, Synthego attempted to target different part of the exon, however they could not confirm the knockout. This was likely due to the high GC content, which is further discussed in section 5.5.

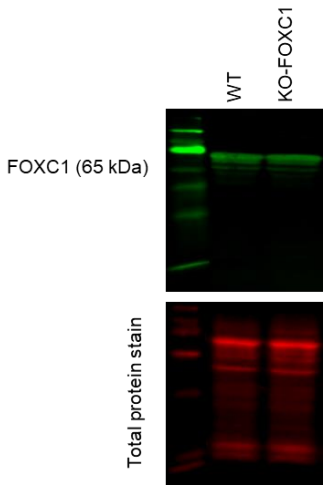


Figure 5.4: FOXC1 protein expression in 786-0 (WT) and FOXC1 KO cells.

Total protein staining was used for normalisation.

Following unsuccessful attempts to knockout FOXC1 in 786-0 and A498 using CRISPR/Cas9 approaches, RNAi technology was then tested. Unlike CRISPR/Cas9, RNAi does not affect DNA, but works by silencing mRNA. Additionally, it causes transient silencing and its effects are best studied in a short period (up to 10 days) after transfection. Two independent siRNAs were selected for this experiment and their effect on protein expression studied using Western Blotting. First, 786-0 cells were transfected with 5 nM siRNA, which effects are shown in Figure 5.5 a,b. It was observed that one of the siRNAs, siFOXC1b, had had some effect on reduction of FOXC1 protein expression (20% lower than control). siFOXC1a did not cause any change in expression levels. In order to test whether increased concentration of the two siRNAs have a stronger effect on the protein expression, 786-0 cells were transfected with 10 nM in the next instance. Again, siFOXC1a had no effect on FOXC1 expression. siFOXC1b caused a more prominent decrease in FOXC1 expression (31% lower than control) when compared to its effect using lower siRNA concentration (Figure 5.5 c,d). To further evaluate whether it is possible to reduce FOXC1 expression even more, 786-0 cells were transfected with 10 nM siFOXC1b twice, second transfection 24h after the first one. As shown in Figure 5.5 e,f, this had similar effect on protein depletion as single siRNA transfection (33% lower than control).

Next, in order to analyse the effect of the highest achieved knockdown of FOXC1, cells were treated with CB-839 and data analysed as previously described. Due to the transient nature of siRNAs, replicates of cytotoxicity assays could not be performed for each transfection. In hindsight, cells could have been transfected with siRNA at the same concentration again to provide statistically significant data at n=3. Figure 5.6 shows that cells with reduced levels of FOXC1 protein (IC_{50}= 507 nM) do not have growth advantage over cells with normal FOXC1 levels (IC_{50}= 569 nM).

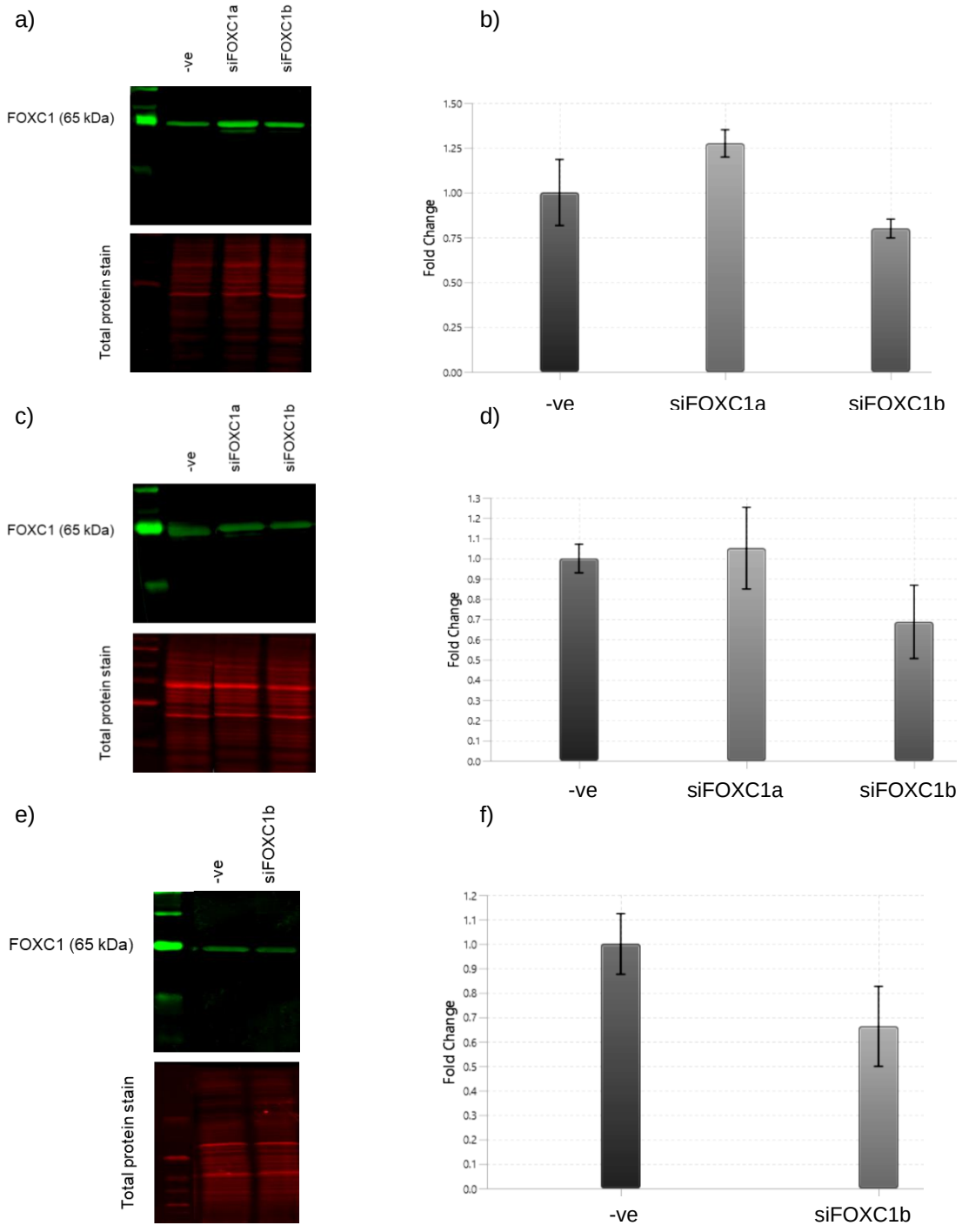


Figure 5.5: FOXC1 protein expressions in 786-0 cells after FOXC1 siRNA transfection detected by Western blot.

(a, d, f) Western Blot of FOXC1 protein levels in protein lysates collected from 786-0 negative control group (-ve), siFOXC1a and siFOXC1b cells. Total protein staining was used for normalisation. (b, c, d) Fold change analysis calculated using Empiria Studio software(n=3).

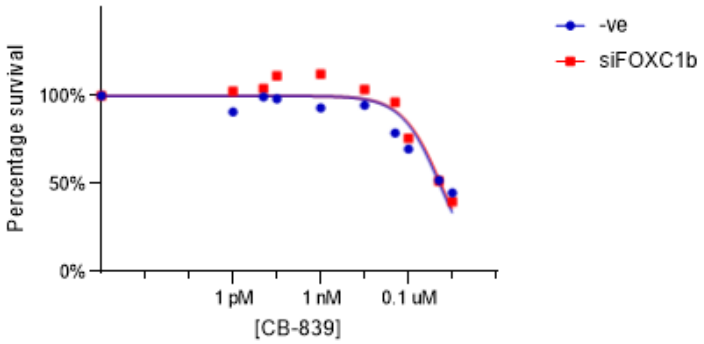


Figure 5.6: Dose-response curves showing control and siRNA (siFOXC1) knockdown cell survival after 72h treatment with increasing doses of CB-839 (0-1 μM).

Data generated using log(inhibitor) vs. response (three parameters) option in GraphPad 8 software, IC50 for 786-0-sgScr= 5.07 $x10^{-8}$M, IC50 786-0-siFOXC1 5.57 $x10^{-8}$M.

5.3.3. RPLP2

To evaluate wheteher RPLP2 knockdown increases resistance of cells to CB-839, 786-0 and A498 parental cell lines were used. The four knockout cell lines were designated as follows: 786-0-sg RPLP2b, 786-0-sg RPLP2c, A498-sg RPLP2b and A498-sg RPLP2c. Knockdown efficiency was analysed for all four cell lines using Western Blotting with total protein stain serving as loading control (Figure 5.7). Protein depletion could not be confirmed for any of the four cell lines and therefore *RPLP2* knockdown was not achieved.

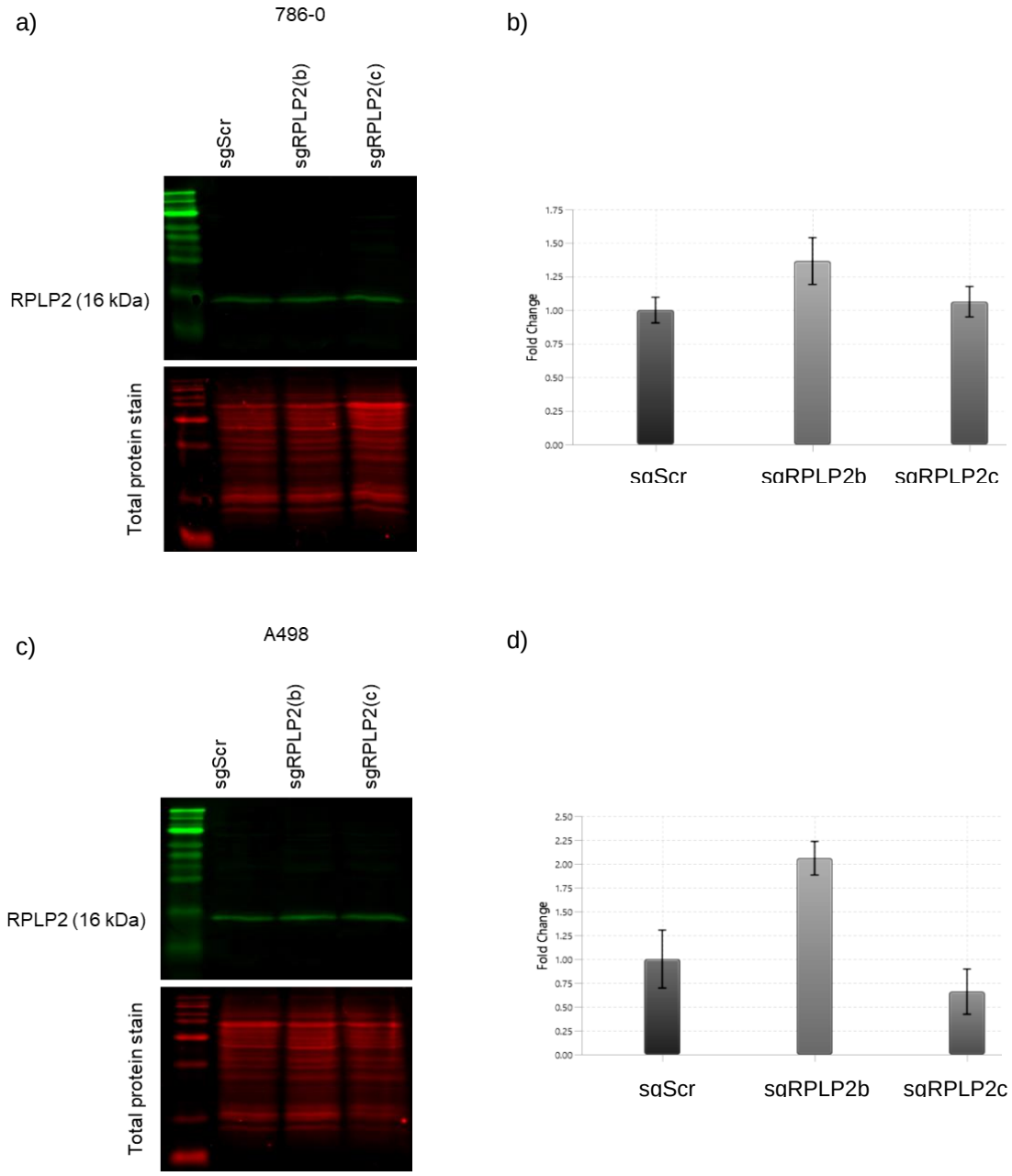


Figure 5.7: RPLP2 protein expressions in 786-0 and A498 cells after transduction with lentiCRISPRv2-sgRPLP2.

(a) Western Blot of RPLP2 protein levels in protein lysates collected from 786-0-sgScr, -sgRPLP2b and -sgRPLP2c cells. Total protein staining was used for normalisation. (c) Western Blot of RPLP2 protein levels in protein lysates collected from A498-sgScr, -sgRPLP2b and -sgRPLP2c cells. Total protein staining was used for normalisation. (b,d) Fold change analysis (vs –sgScr) was calculated using Empiria Studio software.

5.3.4. FOXO3

Next, 786-0 cells were transduced with the lentiCRISPRv2-sgFOXO3 construct and designated as 786-0-sgFOXO3. One sgRNA was chosen out of six from the GeCKOv2 library. Using total protein stain as loading control, it was confirmed that the expression of FOXO3 was lower in 786-0-sgFOXO3 when compared to the expression levels in 786-0-sgScr cells (Figure 5.8). Knockdown cells expressed 68% protein depletion as compared to the control. However, no difference between IC_{50} was observed between 786-0-sgScr cells (IC_{50}= 205 nM) and 786-0-FOXO3 cells (IC_{50}= 226 nM) ($p>0.1$)(Figure 5.9).

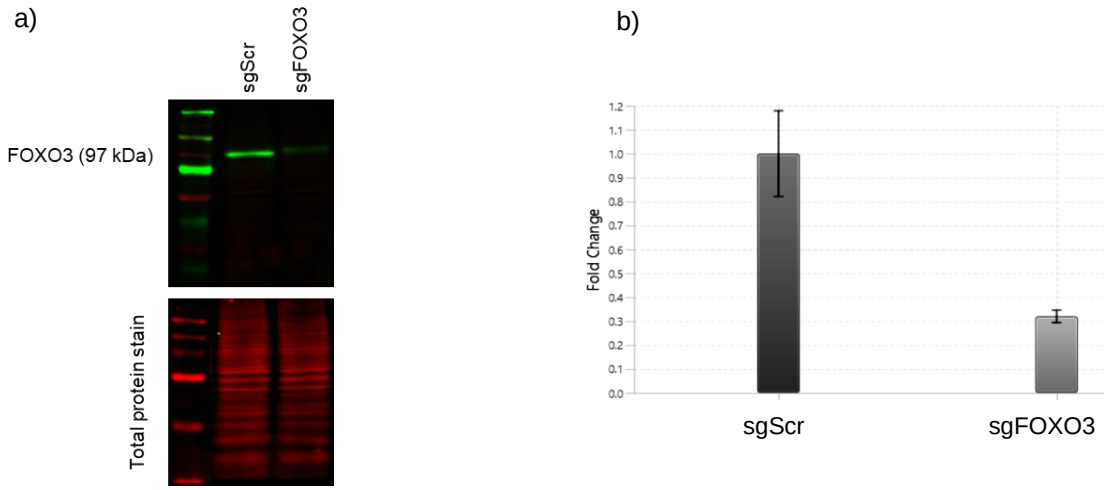


Figure 5.8: FOXO3 protein expressions in 786-0 cells after transduction with lentiCRISPRv2-sgFOXO3.

(a) Western Blot of FOXO3 protein levels in protein lysates collected from 786-0-sgScr and -sgFOXO3 cells. Total protein staining was used for normalisation. (b) Fold change analysis calculated using Empiria Studio software (n=3)

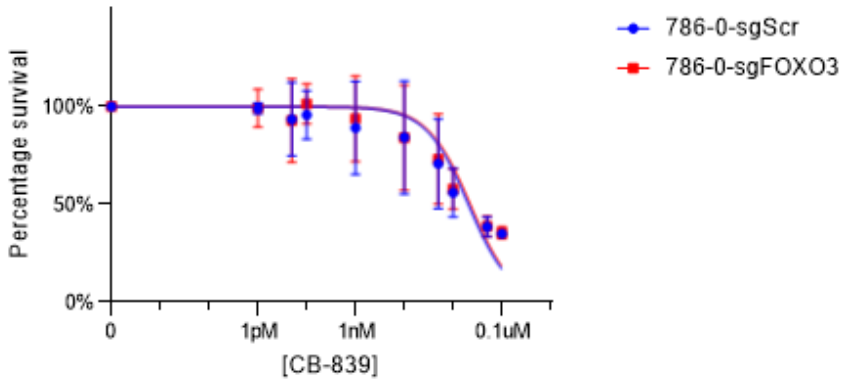


Figure 5.9: Dose-response curves showing control and FOXO3 CRISPR/Cas9 knockdown cell survival after 72h treatment with increasing doses of CB-839 (0-1 μM).

Data generated using log(inhibitor) vs. response (three parameters) option in GraphPad 8 software, IC50 for 786-0-sgScr= 2.05×10^{-8}M, IC50 786-0-sgALDH18A1= 2.26×10^{-8}M.

5.3.5. ALDH18A1

Lastly, 786-0 cells were transduced with the lentiCRISPRv2-sgALDH18A1 construct and designated as 786-0-sgALDH18A1. Since ALDH18A1 protein is located in mitochondria, mitochondrial fraction was isolated from cells using the Mitochondria Isolation Kit for Cultured Cells (see section 2.2.2.). Western Blotting was used to confirm that ALDH18A1 gene was knocked down successfully in 786-0-sgALDH18A1 cells, resulting in 59% protein depletion when compared to the scrambled control (Figure 5.10). However, no difference to CB-839 sensitivity was detected between 786-0-sgScr cells (IC_{50}= 264 nM) and 786-0-sgALDH18A1 cells (IC_{50}= 245 nM) (Figure 5.11).

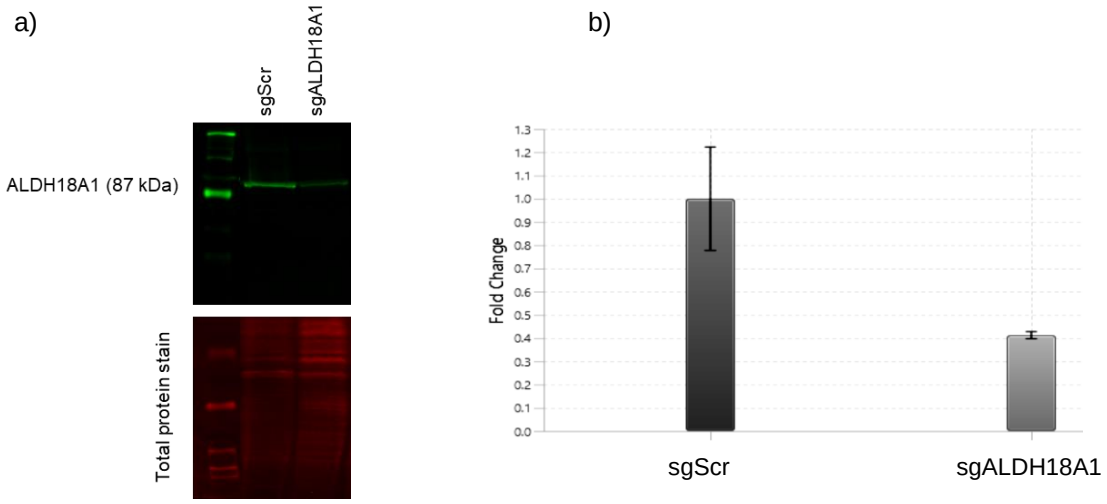


Figure 5.10: ALDH18A1 protein expressions in 786-0 cells after transduction with lentiCRISPRv2-sgALDH18A1.

(a) Western Blot of ALDH18A1 protein levels in protein lysates collected from 786-0-sgScr and -sgALDH18A1 cells. Total protein staining was used for normalisation. (b) Fold change analysis calculated using Empiria Studio software (n=3).

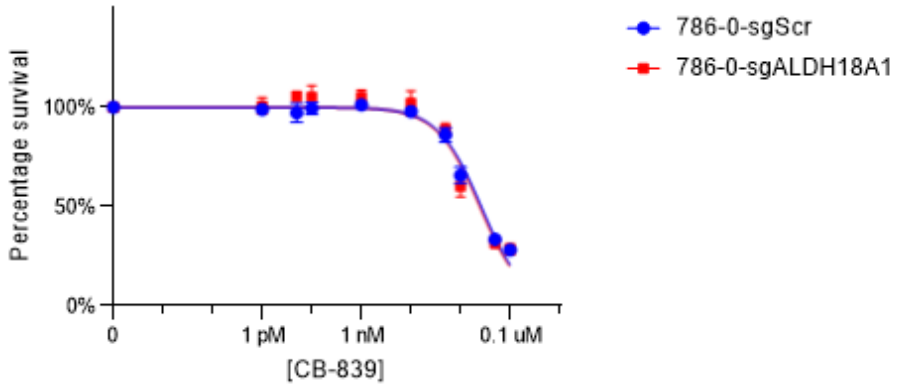


Figure 5.11: Dose-response curves showing control and ALDH18A1 CRISPR/Cas9 knockdown cell survival after 72h treatment with increasing doses of CB-839 (0-1 µM).

Data generated using log(inhibitor) vs. response (three parameters) option in GraphPad 8 software, IC50 for 786-0-sgScr= 2.64×10^{-8}M, IC50 786-0-sgALDH18A1= 2.45×10^{-8}M.

5.4. Generation of a resistant cell line

In addition to valisation of the whole-genome CRISR/Cas9 screen, a resistant cell line was generated during the course of this project. This experiment was not progressed

further due to the limitations caused by the Covid-19 pandemic (time constraint and limited lab space use to accommodate social distancing). However, based on background reading about creating drug resistant cell lines, resistance levels of 786-0-R cell line generated by acute exposure to 0.7 uM of CB-839 in comparison to paternal cell line 786-0 are significant and could be used as an in vitro model of resistance to CB-839. Figure 5.12 shows comparison of resistance between the two cell lines.

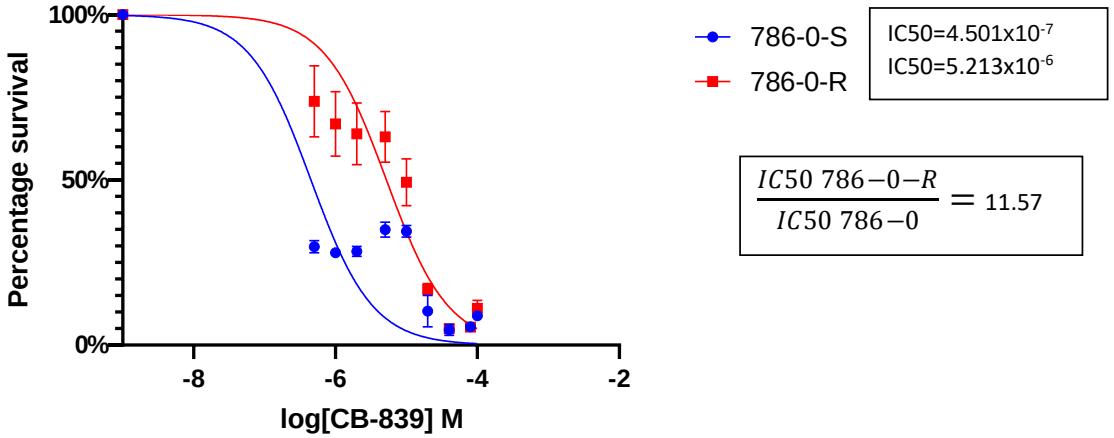


Figure 5.12: Effect of CB-839 on parental cell line (786-0) vs cells continuously exposed to the drug for 4 months (786-0-R).

5.5. Discussion

Genome-wide LOF screens using libraries containing sgRNAs have proven to be a powerful tool to identify novel gene and protein functions by knocking out genes across a population of cells and applying selective pressure followed by identification of genes that are either enriched or depleted in comparison to starting and control populations. Although CRISPR/Cas9 screens combined with NGS technology are now more widely accessible for researchers, every step of the process requires careful optimisation and ultimately, perhaps the most challenging, data analysis and selection of genes for validation. Validation is an essential part of the screen as it aims to confirm that the hits produced are not false positive or false negative results.

Work described in this chapter aimed at confirming that individual knockouts of RPLP2, FOXC1, FOXO3 and ALDH18A1 confer resistance to CB-839. Moreover, one gene, DIAPH3, was selected for validation as it appeared that its knockout is synthetic lethal when combined with cell treatment with CB-839 (since it appeared in the 'purple' groupof genes, Figure 4.17). For each candidate gene single sgRNA was used to create knockout pools followed by confirmation of each knockout using Western blotting to show protein depletion. Cell pools with confirmed knockouts were then treated with increasing doses of CB-839 alongside WT 786-0 cells to test whether knockout of a given gene is in fact responsible for increased resistance to CB-839.

One of the top hits, which appeared on the list for genes, which when knocked out, confer resistance in library pools on Day 16 and Day 21, was RPLP2. It was selected for validation due to it being the only candidate appearing on both days, as well as involvement in protein synthesis and implication in several types of tumours. Attempts at creating 786-0 RPLP2 knockout pools were unsuccessful. Upon further reading, RPLP2

appeared to be an essential gene and therefore cell, which lack this gene are not viable. Therefore, it calls for careful consideration whether this data analysis was performed correctly, and validity of all other hits has to be questioned. This will be further discussed in the next chapter. Moreover, after looking back at MAGeCK data in more detail, it became apparent that one of the RPLP2-targeting sgRNAs in the GeCKOv2 library, HGLibB_41960, has a significantly higher count than any other sgRNA in the library in all four CB-839 treatment samples (day 16 replicate 1 and 2, day 21 replicate 1 and 2). Additionally, none of the other sgRNAs targeting RPLP2 has such high count when compared to all other sgRNAs in the library or across treatment conditions (see Appendix). HGLibB_41960 sgRNA is likely an outlier which failed to be removed by the algorithm and therefore caused RPLP2 to be a false positive hit. What is interesting and unexplained, is the fact that this sgRNA was present at such high count in independent samples. This is an example of false positive and requires further investigation into MAGeCK algorithm and data analysis as it should not have been called as a hit by the algorithm that normalises hits to essential genes. This example highlights the difficulty of data analysis of CRISPR screens.

FOXC1 was another candidate chosen for validation due to its presence high in the ranking of genes as well as its implication in a multiple tumour types. Although majority of literature points at FOXC1 upregulation being involved in worse prognosis, it would be interesting to see whether the opposite, downregulation or knockout of FOXC1, caused problems in CB-839 setting [317-320]. Despite multiple attempts at using different sgRNAs from GeCKOv2 library, as well as other knockout and knockdown approaches, 786-0 FOXC1 cell pools could not be achieved as demonstrated by lack of reduction in protein levels in comparison to protein lysates collected form WT cells. Using siRNA approach showed decrease in FOXC1 protein as compared to WT cells, but it did not increase cells' resistance to CB-839. It is possible that cells require even lower levels of

the protein to observe any effect on CB-839 resistance in 786-0 cells. As two out of 6 sgRNAs failed to produce knockout, it is also possible that FOXC1 is a false positive and again further investigation of how MAGeCK algorithm called FOXC1 as a hit, could shine more light on FOXC1 as a candidate gene.

Lastly, knockout of DIAPH3, which ranked high in the synthetic lethality group of genes (purple group, Figure 4.17), could not be confirmed.

Next high ranked candidate, which, according to the screen data, when knocked out, confers resistance to CB-839, was ALDH18A1. It was selected for validation studies due to its involvement in glutamine metabolism, specifically in conversion of glutamate to proline. Although knockout of ALDH18A1 by single sgRNA was confirmed using Western blotting, no significant increase in resistance to CB-839 when compared to WT cells was observed in SRB assay.

FOXO3 was another candidate selected for validation due to its tumour suppressive role and its downregulation is often observed in cancer development. Similarly to ALDH18A1, reduced protein expression was confirmed by Western blotting but FOXO3 knockout cells did not show increased resistance to CB-839 when compared to WT cells. ALDH18A1 and FOXO3 are either false positive hits or observed reduction in protein levels was not sufficient to cause significant changes to increase resistance to CB-839.

None of the selected candidates could be validated in the cell line, in which the original whole-genome screen was performed. It is possible that some lower ranking candidates would validate. However, the fact that the one of highest ranking gene on both days, Day 16 and Day 21, was an essential gene, which should not have been selected by the algorithm, poses the question whether data was analysed accurately and whether

MAGeCK was the right algorithm to use in this setting. This and other aspects of the screen design and optimisation are discussed in detail in Chapter 6, General discussion.

6. General discussion and future work

6.1. General discussion

Current drug therapy options for patients with ccRCC are anti-angiogenic therapies, mTOR inhibitors and immunotherapies (see Section 1.2.5.). However, most patients display intrinsic or acquired resistance to targeted therapies. High heterogeneity in RCC was shown almost three decades ago and is considered the primary cause of resistance [75, 343, 344]. While this heterogeneity presents a challenge in identifying resistance mechanisms, some of the resistance mechanisms are understood and include compensatory angiogenic pathways, compensatory mTOR signalling, increased tumour evasion, various aspects of tumour microenvironment, lysosomal sequestration of TKIs and adaptation to hypoxia conditions (see Section 1.3.). Therefore, finding new resistance mechanisms may be tractable using appropriate model systems.

Glutaminase is a mitochondrial enzyme, which hydrolises glutamine to glutamate, making it a key enzyme in glutamine metabolism [132]. Despite no significant increase in glutaminase expression in ccRCC, these tumours were shown to have higher levels of glutamine and glutamate compared to normal tissue. It was recently shown that RCC deficient of *VHL* gene not only rely on glutamine for energy generation and maintenance of redox homeostasis, but also for *de novo* pyrimidine synthesis [137].

CB-839 is a member of glutaminase inhibitors class of drugs, and first such drug to be in Phase II clinical trials for various solid and haematological tumours [168, 170, 345, 346]. Given high drug resistance in RCC, it is reasonable to expect emerging resistance to CB-839 in some patients. The inherent plasticity of cellular metabolism and the high degree of metabolic heterogeneity in human tumours pose great challenges for developing effective patient stratification and drug combination strategies [347]. Several

groups have published work that is aimed at predicting sensitivity to CB-839 *in vitro*, but none of them have looked at ccRCC [275, 348]. Additionally, these groups applied methods, such as proteomic profiling and metabolomic studies, but to date, whole-genome CRISPR-Cas9 screening with CB-839 has not been published in any setting.

Genome-scale genetic screens are powerful, yet complex tools to identify genomic perturbations responsible for a specific phenotype. Careful selection of a multitude of parameters is required at every step of the screening procedure. Viral titration, drug concentration and optimisation of sgRNA cassette amplification is described in Chapter 3 of this thesis. This general discussion will focus on additional choices and selection process made throughout the experimental workflow described in this thesis. It will aim to identify steps that might have had an impact on data quality and results that lead to identification of false positive hits. It will also consider whether any alternative choices might have improved the outcomes and led to identification of hits validated as those involved in resistance to CB-839 in RCC. This discussion will take into consideration data available at the point when the experiments were conducted as well as any new research published more recently.

Firstly, an important consideration is the choice of library. Here, the choice between arrayed or pooled library was influenced by factors such as drug class, information on drug action and any previous research aiming at understanding resistance to CB-839. Due to CB-839 being in the experimental phase and clinical trials and not yet approved for routine use in patients, there is no real-life data on potential mechanisms of resistance. To date, only a handful of *in vitro* studies aimed at studying resistance to CB-839 [349, 350]. Moreover, none of the published research focused on RCC. Therefore, a pooled screen was chosen as it allows interrogation of the entire genome. Moreover, it is more cost- and time-effective than using several arrayed libraries. Additionally,

pooled libraries have generally higher genome editing efficiency and transfection efficiency than arrayed libraries [351, 352].

Several validated whole-genome CRISPR/Cas9 libraries are commercially available. As of 2016 GeCKOv2 library and one created by the Sabatini/Lander lab were the most common choices in resistance studies [352]. GeCKOv2 library had been previously successfully used in the Reynolds lab to identify resistance mechanisms to Acelarin in a pancreatic cancer model and therefore was selected for the experiments described in this thesis [353]. However, in recent years, several groups have developed potentially superior libraries. The first iteration of the library used here, GeCKOv1, comprised of 3-4 sgRNAs targeting each gene in the genome [221]. The Zhang group designed the library focusing specifically on 5' constitutively expressed exons to be targeted by sgRNAs, but did not incorporate additional sgRNA design rules to increase cleavage efficiency and decrease off-target effects. Although GeCKOv2 comprises of 6 sgRNAs per gene distributed over 3-4 constitutively expressed exons and has an improved viral titre compared to GeCKOv1, it is an example of first generation library, designed before the development of any sgRNA rules [244]. Despite these limitations, in their proof-of-concept study, the Zhang group showed high level of consistency between independent sgRNAs targeting the same gene and a high rate of hit confirmation [221].

There are no generally accepted design principles for sgRNA libraries and published experiments vary substantially in both design and performance [354, 355]. In recent years the design of sgRNAs and CRISPR libraries was vastly improved based on various rules derived from comparing the nucleotide composition of active and non-active guides [319, 356]. Moreover, training models and deep learning algorithms have been used to identify predictive features of efficiently editing sgRNAs [357-359]. Despite many significant improvements to library design, many factors that influence sgRNA efficiency

are still unknown. Potential additional considerations include DNA accessibility and presence of functional protein domains [322, 360, 361]. It is also important to acknowledge that the CRISPR/Cas9 technology has its limitations as described in Section 1.6.2. These include relatively high frequency of off-target effects and DNA-damage toxicity leading to apoptosis rather than gene editing. Genome editing using CRISPR can also trigger p53 signalling pathway and so cell cycle arrest or apoptosis. Therefore, using this technology can result in aberrant or misleading results.

An example of a second generation library is the Toronto Knock Out (TKO) library, which was designed to only utilize effective targeting sgRNAs with 0 or 1 genomic off target sites [319]. Additionally, Brunello library designed by the Doench laboratory incorporated more detailed sgRNA design rules to maximize cleavage efficiency by Cas9 while decreasing off-target effects [358]. In a 2018 paper, Sanson *et al.* compared GeCKOv1, GeCKOv2, Brunello and TKOv3 libraries and performed 14 screens across 3 cell lines [355]. They found that Brunello outperformed all other libraries at both the sgRNA and gene level. TKOv3 library was the next best performer and GeCKOv2 outperformed GeCKOv1. Interestingly, even with 1 sgRNA per gene Brunello outperformed GeCKOv2 with 6 sgRNAs per gene. These findings show that GeCKOv2 has an inferior screening efficiency compared to the next-generation libraries. Had this data been available at the time of library selection for the study described in this thesis, Brunello library, which was first described in 2016, could have been chosen instead. On the other hand, GeCKOv2 is still a widely-used library. A 2020 study by Lau *et al.* used GeCKOv2 to identify cancer multi-drug resistance genes [233]. They screened cells exposed to 27 drugs, including various classes of drugs, such as DNA damage-inducing drugs and drugs affecting replication, DNA repair, cell cycle and metabolism. They identified known drug targets and relevant pathways as well as gained novel insights into the drugs' mechanisms of actions, including finding ten multi-drug resistance genes. This and other recent studies

suggest that GeCKOv2 is a valid choice for whole-genome resistance screens across multiple cell lines and drugs [238, 362, 363].

The next crucial step is selection of a cell line [352]. In some cases, it may be beneficial to perform the screen in multiple cell lines to avoid specific caveats, such as the genetic background of the cell line or transduction efficiency. Moreover, the choice of a cell line may dictate what genes and how many are identified from the screen. Cell line ploidy is an important consideration. Most normal human cells are diploid, meaning that they contain two copies of each gene. However, cancer cells are characterised by abnormal number of some or all chromosomes and may contain amplified genomic regions due to the inherent genomic instability [364]. Performing a screen in a diploid cell line may produce more robust data with less noise than a screen performed in a hyperdiploid or hypertriploid cell line. This is due to the increased difficulty in knocking out each copy of a gene in a higher ploidy cells. Screens performed in haploid cells may be even more sensitive and produces even higher quality data [365]. Several cancer cell lines, such as KBM7 and HAP1 cells of leukaemic origin, have a near-haploid genotype [233].

However, haploid or near-haploid cell lines are not available as established cell lines to study RCC. Available primary ccRCC cell lines are hypertriploid (786-0) and tetraploid (A498). 769-P cell line has an even less desirable genotype because it contains large numbers of tetra-, hexa-, and higher-ploid cells (www.atcc.org). Although A704 cell line is diploid to hyperdiploid, it is also hypertriploid to hypertetraploid with abnormalities including breaks, dicentrics and endoreduplication. Therefore, given ploidy and taking into consideration other characteristics shown in Table 3.1. 786-0 cell line was chosen as the most viable model of ccRCC here. Moreover, 786-0 was also one of the cell lines used in the Project Achilles, which is a large-scale study aimed at identifying and cataloging gene essentiality across hundreds of genomically characterized cancer cell

lines. Additionally, 786-0 transduction efficiency was confirmed empirically making it a suitable candidate for the screen (see Figure 3.4).

Having selected the appropriate cell line, it was then important to expose cells to varying drug concentrations to establish the correct dose to be used in the screen. There is no consensus on selecting the optimal drug concentration for screening purposes. Choosing low dose could result in not enough selection pressure to observe significant changes in sgRNA populations. A dose too high would kill the majority of cells and could prevent survival of the sufficient number of cells and thus hinder library coverage in treatment groups. Therefore, drug concentration required to kill slightly more than 50% of the wild-type cells was selected here in order to see the desired effect. Number of cells collected at each timepoint confirmed this, with CB-839-treated cells being always half of the count of DMSO-treated cells (data not shown). In hindsight, a higher concentration of the drug could have caused a more prominent effect and could have been used here with sufficient number of cells still being available for collection.

sgRNA library was transduced at an MOI = 0.3 to ensure that most cells receive at most one genetic perturbation. Transducing at higher MOI's may confound screening results when several genes are being targeted per cell and increase false-positive discovery rate [352]. HEK293T cell line was used here for lentiviral production. This is a common choice, however HEK293FT, cell line derived from HEK293T cells, is often recommended as it is fast growing, highly transfectable cell line designed specifically for lentiviral production [285]. Because the transduction potency depends on both the viral preparation and the cell type that is to be transduced, it was established in pilot experiments that HEK293T produced sufficient quantity of virus after surveying recipient cell survival following transduction and puromycin selection.

The next step considers DNA collection and processing. An important change to the original protocol described by Zhang group is the timeline for collection of samples in treatment groups. In this project, DNA from treated cells was harvested at days 16 and 21 to allow for sufficient selection due to the drug action and recovery of cells to maintain the desired library coverage. Similar timelines had been previously successfully used to identify candidate genes, including using this library previously in Reynolds lab. Yau and Rana recommend two methods for DNA isolation [366]. First one is the salt precipitation method described by Zhang lab and the other is extraction of gDNA from frozen cell pellets using QIAamp Blood midi kit according to the manufacturer's protocol [367]. In the study described here, second option was selected.

sgRNA cassettes were then amplified using empirically tested cycling conditions. PCR is known to be a source of bias in measuring abundance of multi-template populations [287, 368]. To capture accurate and representative amplification of multi-template populations, it is crucial to carefully optimise cycle number and reaction conditions [369]. Therefore, different cycling conditions were tested to ensure that all templates would amplify using the exact same conditions and that no additional products would be amplified. Additionally, variation in template abundance between screen and control populations results in inaccuracies in fold-change measurements [370]. A common source of variation is delivery of excess plasmid, which although dilutes over time, may remain in baseline samples [371]. This may lead to the excess plasmid co-purification with cellular genomic DNA and also act as a PCR template. Moreover, excess plasmid may lead to enhanced recovery of GC-rich templates in such samples relative to their recovery from later time-point samples, due to a difference in effective number of PCR cycles [370]. This may cause apparent depletion of GC-rich samples. As described above, low MOI was used to avoid excess plasmid being delivered to recipient cells. Additionally, bacterial nuclease Benzonase may be used to degrade contaminating

plasmid DNA from viral supernatant prior to transduction and eliminate GC bias, but this was not used in this study [372]. Another consideration is sufficient template used for amplification, to keep library representation, which was ensured in this study by following Zhang group recommendations.

Herculase II Polymerase is one of the recommended reagents used in library preparation and 18 to 20 cycles should be optimal using this polymerase [366]. However, in this study 22 cycles of PCR1 were used after previously testing fewer cycles that resulted in lack of visible bands in some samples. Following PCR1, 18 cycles were sufficient for PCR2, as per recommendations. Corresponding samples from biological replicates were processed at the same time to avoid cross-contamination between samples collected from one biological replicate at different timepoints. Additional bands or large molecular weight smears were not observed on agarose gels, suggesting that only one product of the correct size was present. This was also confirmed using Bioanalyzer, an additional quality control measure. Another quality control measure was library quantification using qPCR-based KAPA Library Quantification Kit. All of the steps described confirmed high quality and purity and therefore met criteria necessary for NGS of the samples.

Perhaps the most challenging aspect of genome-wide CRISPR/Cas9 screens is data analysis following NGS. Computational challenges include methods for evaluation of data quality, identification of genes and pathways including their statistical significance and visualisation of results [256]. Several algorithms are used for analysis of screening methods, some originating from RNAi screens and other designed specifically for CRISPR screen data. In this study, data analysis was outsourced to Edinburgh Genomics, where NGS was also performed. Algorithm selected here was MAGeCK-MLE described in section 1.6.6. Unlike an alternative version of MAGeCK, MAGeCK-RRE, this algorithm was designed specifically for screens with multiple conditions, such as

baseline, drug-treated and vehicle control screens and therefore was preferentially chosen here. Moreover, Li *et al*. state that most other algorithms do not take into account varying DNA cutting efficiencies associated with individual sgRNAs, but MAGeCK-MLE was designed to overcome this [256]. MAGeCK was shown to have higher sensitivity and generate fewer false positive hits than other algorithms available at that time [234]. However, in a study published in 2020, Bodapati *et al*. demonstrate that MLE does not perform as well as other available algorithms when guides have a decreased efficiency [250]. They also suggest that RRA is a better choice overall as it performed the most consistently across multiple datasets they analysed. They conclude that MLE could be a suitable option when multiple screens are available on multiple cell lines to identify both common hit genes and cell-type specific hits.

There are several quality control measures suggested for data analysed using MAGeCK pipeline. High quality data was generated based on information such as mapped reads expressed as 'coverage (500) x number of sgRNAs', Gini index (at most 0.1 for Baseline samples), mapped read percentage (above 65%) and other parameters shown in Chapter 4.

After more in-depth analysis of individual sgRNA effect for some genes, which were called as hits, MAGeCK's ability to take into account size effects of various sgRNAs targeting the same gene, has to be questioned. For example, in 'all.count_normalized.txt' output file, HGLibB_41960 sgRNA targeting RPLP2 has a significantly higher count than any other sgRNA in the library in all four CB-839 treatment samples (day 16 replicate 1 and 2, day 21 replicate 1 and 2). It does not have an abnormal (low or high) count in baseline or vehicle control samples. Additionally, none of the other sgRNAs targeting RPLP2 has such high count when compared to all other sgRNAs in the library or across treatment conditions (see Appendix). HGLibB_41960 sgRNA is likely an outlier which

failed to be removed by the algorithm and therefore caused RPLP2 to be a false positive hit. What is interesting and unexplained, is the fact that this sgRNA was present at such high count in independent samples. Moreover, according to DepMap (www.depmap.org), RPLP2 is an essential gene and therefore cells lacking RPLP2 are likely not viable. This could explain why knockout of RPLP2 could not be confirmed on Western Blot.

Another potentially worrying aspect of the data is the presence of non-targeting controls in the top ranking genes when arranged according to decreasing 'CB-839 beta' score. Top ranking genes should be the ones which when knocked out confer resistance to CB-839 and therefore non-targeting controls should not be present here (see Appendix).

This screen generated a list of hits, of which the majority has not been studied in any biological setting, not to mention the specific aspect of RCC and glutamine metabolism. This led to a significant challenge when selecting hits for validation. FOXC1 was the only hit with literature available on its effects in various cancer types. However, most data suggest that FOXC1 overexpression is associated with poor prognosis, rather than FOXC1 downregulation, which is what data described in this thesis suggest [373]. FOXC1 was still considered and interesting candidate for validation as there was no available information on its effect in RCC. Therefore, it was the main focus in validation studies described in Chapter 5. Despite using two individual sgRNAs to knock out FOXC1 and additional custom knockout outsourced to a commercial company Synthego (www.synthego.com), knockout of FOXC1 could not be confirmed using any of these approaches. The siRNA approach led to reduced protein levels, however this did not cause any detectable effect in cell viability when compared to wild-type cells exposed to the same CB-839 concentrations. One possible explanation for this is threshold effect. Despite reduced FOXC1 protein expression, it could still be too high to see any

phenotypic change in response to drug exposure. A more efficient protein level reduction or gene inactivation may be necessary to observe desired effects.

FOXO3 and ALDH18A1 were validated using only one sgRNA per gene. Significant protein depletion was observed for both candidates. However, these proved to be two false positive hits, as cell viability remained unchanged when exposed to CB-839.

Public databases, such as UACLAN, offer a wide range of information about genes and targets of interest and can be used to compare expression profiles in tumour compared to normal tissue [374]. According to the TCGA database on UACLAN, FOXC1 expression is lower in primary tumour samples (median 7.014 transcripts per million) compared to the normal tissue (median 35.287 transcripts per million) but expression levels do not affect patient survival. Expression of FOXO3 is comparable in normal vs tumour samples (median 20.874 vs 20.19 transcripts per million) but higher expression is correlated with better survival probability. Expression of ALDH18A1 is no different (median normal 32.196 vs tumour 39.225 transcripts per million) and lower expression correlates with better survival probability.RPLP2 is overexpressed in primary tumour (median 2,194.654 transcripts per million) when compared to normal tissue (median 1,444.235 transcripts per million) and lower expression in tumour suggests better survival probability. DIAPH3 has low and comparable expression in both normal (median 2.544 transcripts per million) and primary tumours (median 2.118 transcripts per million) and has no effect on survival. Expression levels of hit genes in primary tumour compared to normal tissues do not suggest that any of the hits are strongly implicated in ccRCC further supporting the conclusion of false positive hits.

All validation was done using sgRNAs present in the original library but chosen at random for the validation purposes. This is because the selected algorithm generated a ranking

of genes, but not sgRNAs and therefore it was not possible to find which of the 6 sgRNAs targeting each gene was the most effective. This random selection could have also affected validation studies. Moreover, siRNAs would have given results faster and would allow for testing of larger number of hits in the same time rather than focusing on a small subset.

To summarise, the work described in this thesis generated a high-quality data as defined by recommended quality control checks, such as recommended library representation (<6% of library was below 100 reads, which was expected and consistent with original publication on this library), high percentage of mapped reads, absence of sgRNAs with zero reads and a low Gini index (a high Gini index suggests that the sgRNA read count is distributed very heterogeneously across the target genes). However, this study has significant limitations. First, CB-839 is a first in class glutaminase inhibitor that entered clinical trials and therefore there is no long-term patient-derived data. This limits what results described here could be compared to. Additionally, although CRISPR screens for drug resistance are on the rise, screens that only produce false positive or false negative hits are not being published. CRISPR screens suffer from some similar caveats as siRNA and shRNA screens, but there is still a lack of substantial knowledge about off-target effects, false-positive hits and other aspects of screening. Data analysis is considered the most difficult aspect of CRISPR screens. Although detailed protocols are available, there are not many scientists who specialise in this type of analysis. A significant limitation of this study is limited experience of scientists at Edinburgh Genomics with this type of data analysis.

6.2. Future work

The list of potential candidates was generated using MAGeCK-MLE pipeline. One potential direction could be to reanalyse the data using MAGeCK-RRE to compare it with that of MLE in terms of hit selection. Further analysis of MLE data could be interesting in terms of finding out the reasons behind the selection of RPLP2. Additionally, it would be interesting to validate hits in the 'purple group'- genes, which when knocked out cause synthetic lethality with CB-839. FOXC1 validation using additional sgRNAs or alternative methods could lead to the desired gene knockdown threshold needed to observe a changed phenotype. Work described in this thesis focused on an attempt to validate individual genes, which when knocked out, confer resistance to CB-839. However, gene ranking was not the only information available following NGS data analysis. Several pathways, such as arginine and proline metabolism, N-glycan biosynthesis and Hedgehog signalling pathway (data not shown). It could be interesting to focus on these pathways and perform an arrayed CRISPR screen to further analyse these data. Additionally, alternative approaches could be used to study resistance to CB-839 in RCC setting, such as metabolomics and proteomic studies.

In addition to the whole-genome CRISPR/Cas9 screen, a CB-839-resistant cell line, 786-0-R, was created. As mentioned in Section 1.2.8. of the Introduction, resistant cell lines are often used to study resistance mechanisms to chemotherapeutics and targeted therapies. Various protocols and resistance fold-increases in comparison to the parental cell line can be found in literature. Cell line described here had an over 11-fold higher IC50 when compared to the parental cell line thus suggesting that it could be a relevant model of resistance. However, due to the COVID-19 pandemic and time constraint, this approach was not studied any further. One of the approaches which could be explored in the future is comparison of gene expression between parental and resistant cell line.

As there are multiple resistance mechanisms for each drug, it would be beneficial to create another resistant cell line, derived for example from 769-P or A498 cell lines, and compare the data. The fact that cells with over 11-fold more resistance can be created, suggests that there are cellular factors that do confer resistance to CB-839.

References

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA: A Cancer Journal for Clinicians. 2014;64(1):9-29. doi: 10.3322/caac.21208.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA: A Cancer Journal for Clinicians. 2018;68(1):7-30. doi: 10.3322/caac.21442.
3. Maher ER. Hereditary renal cell carcinoma syndromes: diagnosis, surveillance and management. World journal of urology. 2018;36(12):1891-8. Epub 04/21. doi: 10.1007/s00345-018-2288-5. PubMed PMID: 29680948.
4. Gnarra JR, Tory K, Weng Y, Schmidt L, Wei MH, Li H, et al. Mutations of the VHL tumour suppressor gene in renal carcinoma. Nature genetics. 1994;7(1):85-90. Epub 1994/05/01. doi: 10.1038/ng0594-85. PubMed PMID: 7915601.
5. Clifford SC, Prowse AH, Affara NA, Buys CH, Maher ER. Inactivation of the von Hippel-Lindau (VHL) tumour suppressor gene and allelic losses at chromosome arm 3p in primary renal cell carcinoma: evidence for a VHL-independent pathway in clear cell renal tumourigenesis. Genes, chromosomes & cancer. 1998;22(3):200-9. Epub 1998/06/13. PubMed PMID: 9624531.
6. Haas NB, Nathanson KL. Hereditary kidney cancer syndromes. Advances in chronic kidney disease. 2014;21(1):81-90. doi: 10.1053/j.ackd.2013.10.001. PubMed PMID: 24359990.
7. Zbar B, Tory K, Merino M, Schmidt L, Glenn G, Choyke P, et al. Hereditary papillary renal cell carcinoma. J Urol. 1994;151(3):561-6. Epub 1994/03/01. doi: 10.1016/s0022-5347(17)35015-2. PubMed PMID: 8308957.
8. Benusiglio PR, Giraud S, Deveaux S, Méjean A, Correas J-M, Joly D, et al. Renal cell tumour characteristics in patients with the Birt-Hogg-Dubé cancer susceptibility syndrome: a retrospective, multicentre study. Orphanet Journal of Rare Diseases. 2014;9(1):163. doi: 10.1186/s13023-014-0163-z.
9. Neumann HPH, Bender BU, Berger DP, Laubenberger J, Schultze-Seemann W, Wetterauer U, et al. PREVALENCE, MORPHOLOGY AND BIOLOGY OF RENAL CELL CARCINOMA IN VON HIPPEL-LINDAU DISEASE COMPARED TO SPORADIC RENAL CELL CARCINOMA. The Journal of Urology. 1998;160(4):1248-54. doi: https://doi.org/10.1016/S0022-5347(01)62509-6.
10. Iliopoulos O, Levy AP, Jiang C, Kaelin WG, Goldberg MA. Negative regulation of hypoxia-inducible genes by the von Hippel-Lindau protein. Proceedings of the National Academy of Sciences. 1996;93(20):10595. doi: 10.1073/pnas.93.20.10595.
11. Network CGAR. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature. 2013;499(7456):43.
12. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nature genetics. 2014;46(3):225.
13. Guo G, Gui Y, Gao S, Tang A, Hu X, Huang Y, et al. Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma. Nature genetics. 2012;44(1):17-9. doi: 10.1038/ng.1014.
14. Mancilla-Jimenez R, Stanley RJ, Blath RA. Papillary renal cell carcinoma. A clinical, radiologic, and pathologic study of 34 cases. Cancer. 1976;38(6):2469-80. doi: 10.1002/1097-0142(197612)38:6<2469::AID-CNCR2820380636>3.0.CO;2-R.
15. Delahunt B, Eble JN, McCredie MRE, Bethwaite PB, Stewart JH, Bilous AM. Morphologic typing of papillary renal cell carcinoma: Comparison of growth kinetics and patient survival in 66 cases. Human Pathology. 2001;32(6):590-5. doi: https://doi.org/10.1053/hupa.2001.24984.
16. Jiang F, Richter J, Schraml P, Bubendorf L, Gasser T, Sauter G, et al. Chromosomal imbalances in papillary renal cell carcinoma: genetic differences between histological subtypes.

The American journal of pathology. 1998;153(5):1467-73. doi: 10.1016/S0002-9440(10)65734-3. PubMed PMID: 9811338.
17. Delahunt B, Eble JN, McCredie MR, Bethwaite PB, Stewart JH, Bilous AM. Morphologic typing of papillary renal cell carcinoma: comparison of growth kinetics and patient survival in 66 cases. Hum Pathol. 2001;32(6):590-5. Epub 2001/06/30. doi: 10.1053/hupa.2001.24984. PubMed PMID: 11431713.
18. Zhao WP, Gnarra JR, Liu S, Knutsen T, Linehan WM, Whang-Peng J. Renal cell carcinoma: Cytogenetic analysis of tumors and cell lines. Cancer Genetics and Cytogenetics. 1995;82(2):128-39. doi: https://doi.org/10.1016/0165-4608(95)00024-J.
19. Corless CL, Aburatani H, Fletcher JA, Housman DE, Amin MB, Weinberg DS. Papillary renal cell carcinoma: quantitation of chromosomes 7 and 17 by FISH, analysis of chromosome 3p for LOH, and DNA ploidy. Diagnostic molecular pathology : the American journal of surgical pathology, part B. 1996;5(1):53-64. Epub 1996/03/01. PubMed PMID: 8919546.
20. Lubensky IA, Schmidt L, Zhuang Z, Weirich G, Pack S, Zambrano N, et al. Hereditary and sporadic papillary renal carcinomas with c-met mutations share a distinct morphological phenotype. The American journal of pathology. 1999;155(2):517-26. doi: 10.1016/S0002-9440(10)65147-4. PubMed PMID: 10433944.
21. Alam NA, Rowan AJ, Wortham NC, Pollard PJ, Mitchell M, Tyrer JP, et al. Genetic and functional analyses of FH mutations in multiple cutaneous and uterine leiomyomatosis, hereditary leiomyomatosis and renal cancer, and fumarate hydratase deficiency. Human molecular genetics. 2003;12(11):1241-52. Epub 2003/05/23. PubMed PMID: 12761039.
22. Ooi A, Wong JC, Petillo D, Roossien D, Perrier-Trudova V, Whitten D, et al. An antioxidant response phenotype shared between hereditary and sporadic type 2 papillary renal cell carcinoma. Cancer cell. 2011;20(4):511-23. Epub 2011/10/22. doi: 10.1016/j.ccr.2011.08.024. PubMed PMID: 22014576.
23. Leter EM, Koopmans AK, Gille JJ, van Os TA, Vittoz GG, David EF, et al. Birt-Hogg-Dube syndrome: clinical and genetic studies of 20 families. The Journal of investigative dermatology. 2008;128(1):45-9. Epub 2007/07/06. doi: 10.1038/sj.jid.5700959. PubMed PMID: 17611575.
24. Vera-Badillo FE, Conde E, Duran I. Chromophobe renal cell carcinoma: A review of an uncommon entity. International Journal of Urology. 2012;19(10):894-900. doi: 10.1111/j.1442-2042.2012.03079.x.
25. Verdorfer I, Culig Z, Hobisch A, Bartsch G, Hittmair A, Duba HC, et al. Characterisation of a collecting duct carcinoma by cytogenetic analysis and comparative genomic hybridisation. International journal of oncology. 1998;13(3):461-4. Epub 1998/07/31. PubMed PMID: 9683779.
26. Hunt JD, van der Hel OL, McMillan GP, Boffetta P, Brennan P. Renal cell carcinoma in relation to cigarette smoking: Meta-analysis of 24 studies. International Journal of Cancer. 2005;114(1):101-8. doi: 10.1002/ijc.20618.
27. Habib SL, Prihoda TJ, Luna M, Werner SA. Diabetes and risk of renal cell carcinoma. Journal of Cancer. 2011;3:42-8. doi: 10.7150/jca.3718. PubMed PMID: 22232697.
28. Hu J, Mao Y, White K. Overweight and obesity in adults and risk of renal cell carcinoma in Canada. Sozial- und Praventivmedizin. 2003;48(3):178-85. Epub 2003/08/02. PubMed PMID: 12891869.
29. Stojanovic M, Goldner B, Ivkovic D. Renal cell carcinoma and arterial hypertension. Clinical and experimental nephrology. 2009;13(4):295-9. Epub 2009/02/03. doi: 10.1007/s10157-008-0122-x. PubMed PMID: 19184271.
30. Pesch B, Haerting J, Ranft U, Klimpel A, Oelschlägel B, Schill W. Occupational risk factors for renal cell carcinoma: agent-specific results from a case-control study in Germany. International Journal of Epidemiology. 2000;29(6):1014-24. doi: 10.1093/ije/29.6.1014.

31. Gago-Dominguez M, Yuan JM, Castelao JE, Ross RK, Yu MC. Regular use of analgesics is a risk factor for renal cell carcinoma. British Journal of Cancer. 1999;81(3):542-8. doi: 10.1038/sj.bjc.6690728. PubMed PMID: PMC2362920.
32. Gibbons RP, Monte JE, Correa RJ, Jr., Mason JT. Manifestations of renal cell carcinoma. Urology. 1976;8(3):201-6. Epub 1976/09/01. doi: 10.1016/0090-4295(76)90366-6. PubMed PMID: 788291.
33. Escudier B, Porta C, Schmidinger M, Rioux-Leclercq N, Bex A, Khoo V, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up[†]. Annals of Oncology. 2019;30(5):706-20. doi: 10.1093/annonc/mdz056.
34. Lee CT, Katz J, Fearn PA, Russo P, editors. Mode of presentation of renal cell carcinoma provides prognostic information. Urologic Oncology: Seminars and Original Investigations; 2002: Elsevier.
35. Lee Cheryl T, Katz J, Shi W, Thaler Howard T, Reuter Victor E, Russo P. SURGICAL MANAGEMENT OF RENAL TUMORS 4 CM. OR LESS IN A CONTEMPORARY COHORT. Journal of Urology. 2000;163(3):730-6. doi: 10.1016/S0022-5347(05)67793-2.
36. Tsui KH, Shvarts O, Smith RB, Figlin RA, deKernion JB, Belldegrun A. Prognostic indicators for renal cell carcinoma: a multivariate analysis of 643 patients using the revised 1997 TNM staging criteria. J Urol. 2000;163(4):1090-5; quiz 295. Epub 2000/03/29. PubMed PMID: 10737472.
37. Hollenbeck BK, Miller DC, Hollingsworth JM, Daignault S. Rising Incidence of Small Renal Masses: A Need to Reassess Treatment Effect. JNCI: Journal of the National Cancer Institute. 2006;98(18):1331-4. doi: 10.1093/jnci/djj362.
38. Quivy A, Daste A, Harbaoui A, Duc S, Bernhard J-C, Gross-Goupil M, et al. Optimal management of renal cell carcinoma in the elderly: a review. Clinical interventions in aging. 2013;8:433-42. Epub 04/19. doi: 10.2147/CIA.S30765. PubMed PMID: 23626463.
39. Steinbach F, Novick AC, Zincke H, Miller DP, Williams RD, Lund G, et al. Treatment of renal cell carcinoma in von Hippel-Lindau disease: a multicenter study. J Urol. 1995;153(6):1812-6. Epub 1995/06/01. PubMed PMID: 7752324.
40. Choi SY, Jung H, You D, Jeong IG, Song C, Hong B, et al. Robot-assisted partial nephrectomy is associated with early recovery of renal function: Comparison of open, laparoscopic, and robot-assisted partial nephrectomy using DTPA renal scintigraphy. Journal of surgical oncology. 2019. Epub 2019/03/02. doi: 10.1002/jso.25429. PubMed PMID: 30820951.
41. Lee CT, Katz J, Shi W, Thaler HT, Reuter VE, Russo P. Surgical management of renal tumors 4 cm. or less in a contemporary cohort. J Urol. 2000;163(3):730-6. Epub 2000/02/25. PubMed PMID: 10687966.
42. Huang WC, Elkin EB, Levey AS, Jang TL, Russo P. Partial nephrectomy versus radical nephrectomy in patients with small renal tumors--is there a difference in mortality and cardiovascular outcomes? The Journal of urology. 2009;181(1):55-62. Epub 11/13. doi: 10.1016/j.juro.2008.09.017. PubMed PMID: 19012918.
43. Provet J, Tessler A, Brown J, Golimbu M, Bosniak M, Morales P. Partial nephrectomy for renal cell carcinoma: indications, results and implications. J Urol. 1991;145(3):472-6. Epub 1991/03/01. PubMed PMID: 1997691.
44. Margulis V, Tamboli P, Jacobsohn KM, Swanson DA, Wood CG. Oncological efficacy and safety of nephron-sparing surgery for selected patients with locally advanced renal cell carcinoma. BJU Int. 2007;100(6):1235-9. Epub 2007/11/06. doi: 10.1111/j.1464-410X.2007.07225.x. PubMed PMID: 17979923.
45. Asimakopoulos AD, Miano R, Annino F, Micali S, Spera E, Iorio B, et al. Robotic radical nephrectomy for renal cell carcinoma: a systematic review. BMC urology. 2014;14:75. Epub

2014/09/23. doi: 10.1186/1471-2490-14-75. PubMed PMID: 25234265; PubMed Central PMCID: PMCPMC4171399.
46. Deschavanne PJ, Fertil B. A review of human cell radiosensitivity in vitro. International journal of radiation oncology, biology, physics. 1996;34(1):251-66. Epub 1996/01/01. PubMed PMID: 12118559.
47. Teh B, Bloch C, Galli-Guevara M, Doh L, Richardson S, Chiang S, et al. The treatment of primary and metastatic renal cell carcinoma (RCC) with image-guided stereotactic body radiation therapy (SBRT). Biomedical imaging and intervention journal. 2007;3(1):e6. Epub 2007/01/01. doi: 10.2349/biij.3.1.e6. PubMed PMID: 21614267; PubMed Central PMCID: PMCPMC3097653.
48. Siva S, Kothari G, Muacevic A, Louie AV, Slotman BJ, Teh BS, et al. Radiotherapy for renal cell carcinoma: renaissance of an overlooked approach. Nature reviews Urology. 2017;14(9):549-63. Epub 2017/06/21. doi: 10.1038/nrurol.2017.87. PubMed PMID: 28631740.
49. Fyfe G, Fisher RI, Rosenberg SA, Sznol M, Parkinson DR, Louie AC. Results of treatment of 255 patients with metastatic renal cell carcinoma who received high-dose recombinant interleukin-2 therapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1995;13(3):688-96. Epub 1995/03/01. doi: 10.1200/jco.1995.13.3.688. PubMed PMID: 7884429.
50. Koneru R, Hotte SJ. Role of cytokine therapy for renal cell carcinoma in the era of targeted agents. Curr Oncol. 2009;16 Suppl 1(Suppl 1):S40-S4. doi: 10.3747/co.v16i0.417. PubMed PMID: 19478896.
51. Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Rixe O, et al. Sunitinib versus Interferon Alfa in Metastatic Renal-Cell Carcinoma. New England Journal of Medicine. 2007;356(2):115-24. doi: 10.1056/NEJMoa065044.
52. Barata PC, Rini BI. Treatment of renal cell carcinoma: Current status and future directions. CA: A Cancer Journal for Clinicians. 2017;67(6):507-24. doi: https://doi.org/10.3322/caac.21411.
53. Mendel DB, Laird AD, Xin X, Louie SG, Christensen JG, Li G, et al. In vivo antitumor activity of SU11248, a novel tyrosine kinase inhibitor targeting vascular endothelial growth factor and platelet-derived growth factor receptors: determination of a pharmacokinetic/pharmacodynamic relationship. Clin Cancer Res. 2003;9(1):327-37. Epub 2003/01/23. PubMed PMID: 12538485.
54. Papaetis GS, Syrigos KN. Sunitinib. BioDrugs. 2009;23(6):377-89. doi: 10.2165/11318860-000000000-00000.
55. Motzer RJ, Hutson TE, Cella D, Reeves J, Hawkins R, Guo J, et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. N Engl J Med. 2013;369(8):722-31. doi: 10.1056/NEJMoa1303989. PubMed PMID: 23964934.
56. Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Staehler M, et al. Sorafenib for Treatment of Renal Cell Carcinoma: Final Efficacy and Safety Results of the Phase III Treatment Approaches in Renal Cancer Global Evaluation Trial. Journal of Clinical Oncology. 2009;27(20):3312-8. doi: 10.1200/JCO.2008.19.5511.
57. Escudier B, Gore M. Axitinib for the Management of Metastatic Renal Cell Carcinoma. Drugs in R&D. 2011;11(2):113-26. doi: 10.2165/11591240-000000000-00000. PubMed PMID: PMC3585900.
58. Ferrara N, Hillan KJ, Novotny W. Bevacizumab (Avastin), a humanized anti-VEGF monoclonal antibody for cancer therapy. Biochem Biophys Res Commun. 2005;333(2):328-35. doi: 10.1016/j.bbrc.2005.05.132. PubMed PMID: 15961063.
59. Fingar DC, Blenis J. Target of rapamycin (TOR): an integrator of nutrient and growth factor signals and coordinator of cell growth and cell cycle progression. Oncogene. 2004;23:3151. doi: 10.1038/sj.onc.1207542.

60. Li J, Kim SG, Blenis J. Rapamycin: one drug, many effects. Cell metabolism. 2014;19(3):373-9. Epub 02/06. doi: 10.1016/j.cmet.2014.01.001. PubMed PMID: 24508508.
61. Eng CP, Sehgal SN, Vézina C. Activity of rapamycin (AY-22,989) against transplanted tumors. The Journal of antibiotics. 1984;37(10):1231-7. Epub 1984/10/01. doi: 10.7164/antibiotics.37.1231. PubMed PMID: 6501094.
62. Martel R, Klicius J, Galet S. Inhibition of the immune response by rapamycin, a new antifungal antibiotic. Canadian journal of physiology and pharmacology. 1977;55(1):48-51.
63. Ballou LM, Lin RZ. Rapamycin and mTOR kinase inhibitors. J Chem Biol. 2008;1(1-4):27-36. Epub 05/15. doi: 10.1007/s12154-008-0003-5. PubMed PMID: 19568796.
64. Fantus D, Rogers NM, Grahammer F, Huber TB, Thomson AW. Roles of mTOR complexes in the kidney: implications for renal disease and transplantation. Nature Reviews Nephrology. 2016;12:587. doi: 10.1038/nrneph.2016.108.
65. Shin Lee J, Seok Kim H, Bok Kim Y, Cheol Lee M, Soo Park C. Expression of PTEN in renal cell carcinoma and its relation to tumor behavior and growth. Journal of surgical oncology. 2003;84(3):166-72. Epub 2003/11/05. doi: 10.1002/jso.10302. PubMed PMID: 14598361.
66. Yang P, Cornejo KM, Sadow PM, Cheng L, Wang M, Xiao Y, et al. Renal cell carcinoma in tuberous sclerosis complex. The American journal of surgical pathology. 2014;38(7):895-909. doi: 10.1097/PAS.0000000000000237. PubMed PMID: 24832166.
67. Santoni M, Massari F, Di Nunno V, Conti A, Cimadamore A, Scarpelli M, et al. Immunotherapy in renal cell carcinoma: latest evidence and clinical implications. Drugs Context. 2018;7:212528-. doi: 10.7573/dic.212528. PubMed PMID: 29899754.
68. Farkona S, Diamandis EP, Blasutig IM. Cancer immunotherapy: the beginning of the end of cancer? BMC medicine. 2016;14:73-. doi: 10.1186/s12916-016-0623-5. PubMed PMID: 27151159.
69. Zou W, Wolchok JD, Chen L. PD-L1 (B7-H1) and PD-1 pathway blockade for cancer therapy: Mechanisms, response biomarkers, and combinations. Sci Transl Med. 2016;8(328):328rv4-rv4. doi: 10.1126/scitranslmed.aad7118. PubMed PMID: 26936508.
70. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, Activity, and Immune Correlates of Anti–PD-1 Antibody in Cancer. New England Journal of Medicine. 2012;366(26):2443-54. doi: 10.1056/NEJMoa1200690.
71. O'Day SJ, Hamid O, Urba WJ. Targeting cytotoxic T-lymphocyte antigen-4 (CTLA-4): a novel strategy for the treatment of melanoma and other malignancies. Cancer. 2007;110(12):2614-27. Epub 2007/11/16. doi: 10.1002/cncr.23086. PubMed PMID: 18000991.
72. Nowell PC. The clonal evolution of tumor cell populations. Science. 1976;194(4260):23-8.
73. Pepper JW, Scott Findlay C, Kassen R, Spencer SL, Maley CC. Cancer research meets evolutionary biology. Evol Appl. 2009;2(1):62-70. doi: 10.1111/j.1752-4571.2008.00063.x. PubMed PMID: 25567847.
74. Krishnan B, Rose TL, Kardos J, Milowsky MI, Kim WY. Intrinsic genomic differences between African American and white patients with clear cell renal cell carcinoma. JAMA oncology. 2016;2(5):664-7.
75. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. New England journal of medicine. 2012;366(10):883-92.
76. Kim K-T, Lee HW, Lee H-O, Song HJ, Shin S, Kim H, et al. Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. Genome biology. 2016;17(1):80.
77. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47. doi: 10.1016/j.ejca.2008.10.026. PubMed PMID: 19097774.

78. Makhov P, Joshi S, Ghatalia P, Kutikov A, Uzzo RG, Kolenko VM. Resistance to Systemic Therapies in Clear Cell Renal Cell Carcinoma: Mechanisms and Management Strategies. Molecular cancer therapeutics. 2018;17(7):1355-64. doi: 10.1158/1535-7163.MCT-17-1299. PubMed PMID: 29967214.
79. Rini BI. New strategies in kidney cancer: therapeutic advances through understanding the molecular basis of response and resistance. Clin Cancer Res. 2010;16(5):1348-54. Epub 2010/02/25. doi: 10.1158/1078-0432.ccr-09-2273. PubMed PMID: 20179240.
80. de Bazelaire C, Alsop DC, George D, Pedrosa I, Wang Y, Michaelson MD, et al. Magnetic resonance imaging–measured blood flow change after antiangiogenic therapy with PTK787/ZK 222584 correlates with clinical outcome in metastatic renal cell carcinoma. Clinical Cancer Research. 2008;14(17):5548-54.
81. Shibuya M. Vascular Endothelial Growth Factor (VEGF) and Its Receptor (VEGFR) Signaling in Angiogenesis: A Crucial Target for Anti- and Pro-Angiogenic Therapies. Genes Cancer. 2011;2(12):1097-105. doi: 10.1177/1947601911423031. PubMed PMID: 22866201.
82. Tammela T, Zarkada G, Wallgard E, Murtomäki A, Suchting S, Wirzenius M, et al. Blocking VEGFR-3 suppresses angiogenic sprouting and vascular network formation. Nature. 2008;454(7204):656-60. Epub 06/25. doi: 10.1038/nature07083. PubMed PMID: 18594512.
83. Rini BI, Wilding G, Hudes G, Stadler WM, Kim S, Tarazi J, et al. Phase II Study of Axitinib in Sorafenib-Refractory Metastatic Renal Cell Carcinoma. Journal of Clinical Oncology. 2009;27(27):4462-8. doi: 10.1200/JCO.2008.21.7034.
84. Ho TH, Liu X-D, Huang Y, Warneke CL, Johnson MM, Hoang A, et al. The impact of FGFR1 and FRS2α expression on sorafenib treatment in metastatic renal cell carcinoma. BMC cancer. 2015;15:304-. doi: 10.1186/s12885-015-1302-1. PubMed PMID: 25900027.
85. Wang X, Bullock AJ, Zhang L, Wei L, Yu D, Mahagaokar K, et al. The role of angiopoietins as potential therapeutic targets in renal cell carcinoma. Translational oncology. 2014;7(2):188-95. Epub 03/04. doi: 10.1016/j.tranon.2014.02.003. PubMed PMID: 24704536.
86. Bhatt RS, Atkins MB. Molecular pathways: can activin-like kinase pathway inhibition enhance the limited efficacy of VEGF inhibitors? Clinical cancer research : an official journal of the American Association for Cancer Research. 2014;20(11):2838-45. Epub 04/08. doi: 10.1158/1078-0432.CCR-13-2788. PubMed PMID: 24714770.
87. Harada K, Miyake H, Kumano M, Fujisawa M. Acquired resistance to temsirolimus in human renal cell carcinoma cells is mediated by the constitutive activation of signal transduction pathways through mTORC2. British Journal of Cancer. 2013;109(9):2389-95. doi: 10.1038/bjc.2013.602.
88. Toschi A, Lee E, Gadir N, Ohh M, Foster DA. Differential dependence of hypoxia-inducible factors 1 alpha and 2 alpha on mTORC1 and mTORC2. The Journal of biological chemistry. 2008;283(50):34495-9. Epub 10/22. doi: 10.1074/jbc.C800170200. PubMed PMID: 18945681.
89. Sun S-Y, Rosenberg LM, Wang X, Zhou Z, Yue P, Fu H, et al. Activation of Akt and eIF4E survival pathways by rapamycin-mediated mammalian target of rapamycin inhibition. Cancer research. 2005;65(16):7052-8.
90. Pan X-d, Gu D-h, Mao J-H, Zhu H, Chen X, Zheng B, et al. Concurrent inhibition of mTORC1 and mTORC2 by WYE-687 inhibits renal cell carcinoma cell growth in vitro and in vivo. PLOS ONE. 2017;12(3):e0172555. doi: 10.1371/journal.pone.0172555.
91. Zhang H, Berel D, Wang Y, Li P, Bhowmick NA, Figlin RA, et al. A comparison of Ku0063794, a dual mTORC1 and mTORC2 inhibitor, and temsirolimus in preclinical renal cell carcinoma models. PloS one. 2013;8(1):e54918-e. Epub 01/22. doi: 10.1371/journal.pone.0054918. PubMed PMID: 23349989.
92. Kauffman EC, Lang M, Rais-Bahrami S, Gupta GN, Wei D, Yang Y, et al. Preclinical efficacy of dual mTORC1/2 inhibitor AZD8055 in renal cell carcinoma harboring a TFE3 gene fusion. BMC Cancer. 2019;19(1):917. doi: 10.1186/s12885-019-6096-0.

93. Carew JS, Espitia CM, Zhao W, Mita MM, Mita AC, Nawrocki ST. Targeting Survivin Inhibits Renal Cell Carcinoma Progression and Enhances the Activity of Temsirolimus. Molecular Cancer Therapeutics. 2015;14(6):1404. doi: 10.1158/1535-7163.MCT-14-1036.
94. Di Nicolantonio F, Arena S, Tabernero J, Grosso S, Molinari F, Macarulla T, et al. Deregulation of the PI3K and KRAS signaling pathways in human cancer cells determines their response to everolimus. The Journal of clinical investigation. 2010;120(8):2858-66. Epub 07/26. doi: 10.1172/JCI37539. PubMed PMID: 20664172.
95. Fruman DA, Wood MA, Gjertson CK, Katz HR, Burakoff SJ, Bierer BE. FK506 binding protein 12 mediates sensitivity to both FK506 and rapamycin in murine mast cells. European journal of immunology. 1995;25(2):563-71.
96. Engl T, Rutz J, Maxeiner S, Fanguen S, Juengel E, Koschade S, et al. Acquired resistance to temsirolimus is associated with integrin α7 driven chemotactic activity of renal cell carcinoma in vitro. Oncotarget. 2018;9(27):18747-59. doi: 10.18632/oncotarget.24650. PubMed PMID: 29721158.
97. Bergers G, Hanahan D. Modes of resistance to anti-angiogenic therapy. Nature reviews Cancer. 2008;8(8):592-603. doi: 10.1038/nrc2442. PubMed PMID: 18650835.
98. van der Mijn JC, Broxterman HJ, Knol JC, Piersma SR, De Haas RR, Dekker H, et al. Sunitinib activates Axl signaling in renal cell cancer. International journal of cancer. 2016;138(12):3002-10. Epub 03/01. doi: 10.1002/ijc.30022. PubMed PMID: 26815723.
99. Zhou L, Liu XD, Sun M, Zhang X, German P, Bai S, et al. Targeting MET and AXL overcomes resistance to sunitinib therapy in renal cell carcinoma. Oncogene. 2016;35(21):2687-97. doi: 10.1038/onc.2015.343.
100. Rankin EB, Fuh KC, Castellini L, Viswanathan K, Finger EC, Diep AN, et al. Direct regulation of GAS6/AXL signaling by HIF promotes renal metastasis through SRC and MET. Proceedings of the National Academy of Sciences. 2014;111(37):13373-8.
101. Wang M, Zhao J, Zhang L, Wei F, Lian Y, Wu Y, et al. Role of tumor microenvironment in tumorigenesis. Journal of Cancer. 2017;8(5):761-73. doi: 10.7150/jca.17648. PubMed PMID: 28382138.
102. Liu X-D, Hoang A, Zhou L, Kalra S, Yetil A, Sun M, et al. Resistance to Antiangiogenic Therapy Is Associated with an Immunosuppressive Tumor Microenvironment in Metastatic Renal Cell Carcinoma. Cancer Immunology Research. 2015;3(9):1017. doi: 10.1158/2326-6066.CIR-14-0244.
103. Brighton CT, Lorich DG, Kupcha R, Reilly TM, Jones AR, Woodbury RA, 2nd. The pericyte as a possible osteoblast progenitor cell. Clin Orthop Relat Res. 1992;(275):287-99. PubMed PMID: 1735227.
104. Balabanov R, Washington R, Wagnerova J, Dore-Duffy P. CNS microvascular pericytes express macrophage-like function, cell surface integrin αM, and macrophage marker ED-2. Microvascular research. 1996;52(2):127-42.
105. Hosaka K, Yang Y, Seki T, Fischer C, Dubey O, Fredlund E, et al. Pericyte–fibroblast transition promotes tumor growth and metastasis. Proceedings of the National Academy of Sciences. 2016;113(38):E5618. doi: 10.1073/pnas.1608384113.
106. Cao Y, Zhang Z-L, Zhou M, Elson P, Rini B, Aydin H, et al. Pericyte coverage of differentiated vessels inside tumor vasculature is an independent unfavorable prognostic factor for patients with clear cell renal cell carcinoma. Cancer. 2013;119(2):313-24. Epub 07/18. doi: 10.1002/cncr.27746. PubMed PMID: 22811049.
107. Kim BH, Sohn JC, Ha JY, Park CH, Choe MS, Jung HR, et al. Relationships between the effect of sunitinib and immature blood vessels in metastatic renal cell cancer. Urol Int. 2015;94(2):137-43. Epub 08/16. doi: 10.1159/000363773. PubMed PMID: 25138147.
108. Xu H, Ren D. Lysosomal Physiology. Annual Review of Physiology. 2015;77(1):57-80. doi: 10.1146/annurev-physiol-021014-071649.

109. Gotink KJ, Broxterman HJ, Labots M, de Haas RR, Dekker H, Honeywell RJ, et al. Lysosomal Sequestration of Sunitinib: A Novel Mechanism of Drug Resistance. Clinical Cancer Research. 2011;17(23):7337. doi: 10.1158/1078-0432.CCR-11-1667.
110. Al Tameemi W, Dale TP, Al-Jumaily RMK, Forsyth NR. Hypoxia-Modified Cancer Cell Metabolism. Frontiers in Cell and Developmental Biology. 2019;7(4). doi: 10.3389/fcell.2019.00004.
111. Ruan K, Song G, Ouyang G. Role of hypoxia in the hallmarks of human cancer. Journal of Cellular Biochemistry. 2009;107(6):1053-62. doi: 10.1002/jcb.22214.
112. Teicher BA. Hypoxia and drug resistance. Cancer and Metastasis Reviews. 1994;13(2):139-68.
113. Teicher BA, Lazo JS, Sartorelli AC. Classification of antineoplastic agents by their selective toxicities toward oxygenated and hypoxic tumor cells. Cancer research. 1981;41(1):73-81.
114. Bellot G, Garcia-Medina R, Gounon P, Chiche J, Roux D, Pouysségur J, et al. Hypoxia-Induced Autophagy Is Mediated through Hypoxia-Inducible Factor Induction of BNIP3 and BNIP3L via Their BH3 Domains. Molecular and Cellular Biology. 2009;29(10):2570. doi: 10.1128/MCB.00166-09.
115. Glick D, Barth S, Macleod KF. Autophagy: cellular and molecular mechanisms. J Pathol. 2010;221(1):3-12. doi: 10.1002/path.2697. PubMed PMID: 20225336.
116. Brugarolas J, Lei K, Hurley RL, Manning BD, Reiling JH, Hafen E, et al. Regulation of mTOR function in response to hypoxia by REDD1 and the TSC1/TSC2 tumor suppressor complex. Genes & development. 2004;18(23):2893-904.
117. Fürst P. Intracellular muscle free amino acids--their measurement and function. The Proceedings of the Nutrition Society. 1983;42(3):451-62. Epub 1983/09/01. doi: 10.1079/pns19830052. PubMed PMID: 6361773.
118. Krebs HA. Metabolism of amino-acids: The synthesis of glutamine from glutamic acid and ammonia, and the enzymic hydrolysis of glutamine in animal tissues. Biochem J. 1935;29(8):1951-69. Epub 1935/08/01. doi: 10.1042/bj0291951. PubMed PMID: 16745865; PubMed Central PMCID: PMCPMC1266709.
119. Moffatt BA, Ashihara H. Purine and pyrimidine nucleotide synthesis and metabolism. Arabidopsis Book. 2002;1:e0018-e. Epub 04/04. doi: 10.1199/tab.0018. PubMed PMID: 22303196.
120. Cruzat V, Macedo Rogero M, Noel Keane K, Curi R, Newsholme P. Glutamine: Metabolism and Immune Function, Supplementation and Clinical Translation. Nutrients. 2018;10(11):1564. doi: 10.3390/nu10111564. PubMed PMID: 30360490.
121. Vander Heiden M, Cantley L, Thompson C. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. Science (New York, NY). 2009;324(5930):1029-33. doi: 10.1126/science.1160809. PubMed PMID: PMC2849637.
122. Lacey JM, Wilmore DW. Is glutamine a conditionally essential amino acid? Nutrition reviews. 1990;48(8):297-309. Epub 1990/08/01. doi: 10.1111/j.1753-4887.1990.tb02967.x. PubMed PMID: 2080048.
123. Kuhn KS, Schuhmann K, Stehle P, Darmaun D, Fürst P. Determination of glutamine in muscle protein facilitates accurate assessment of proteolysis and de novo synthesis–derived endogenous glutamine production. The American Journal of Clinical Nutrition. 1999;70(4):484-9. doi: 10.1093/ajcn/70.4.484.
124. Exton JH. Gluconeogenesis. Metabolism: clinical and experimental. 1972;21(10):945-90. Epub 1972/10/01. doi: 10.1016/0026-0495(72)90028-5. PubMed PMID: 4342011.
125. Landau BR, Wahren J, Chandramouli V, Schumann WC, Ekberg K, Kalhan SC. Contributions of gluconeogenesis to glucose production in the fasted state. The Journal of clinical investigation. 1996;98(2):378-85. doi: 10.1172/JCI118803. PubMed PMID: 8755648.

126. Krebs HA. RENAL GLUCONEOGENESIS. Advances in enzyme regulation. 1963;1:385-400. Epub 1963/01/01. doi: 10.1016/0065-2571(63)90034-7. PubMed PMID: 14190368.
127. Stumvoll M, Meyer C, Perriello G, Kreider M, Welle S, Gerich J. Human kidney and liver gluconeogenesis: evidence for organ substrate selectivity. The American journal of physiology. 1998;274(5):E817-26. Epub 1998/06/05. doi: 10.1152/ajpendo.1998.274.5.E817. PubMed PMID: 9612239.
128. Hanahan D, Weinberg Robert A. Hallmarks of Cancer: The Next Generation. Cell. 2011;144(5):646-74. doi: https://doi.org/10.1016/j.cell.2011.02.013.
129. Xiang L, Mou J, Shao B, Wei Y, Liang H, Takano N, et al. Glutaminase 1 expression in colorectal cancer cells is induced by hypoxia and required for tumor growth, invasion, and metastatic colonization. Cell Death & Disease. 2019;10(2):40. doi: 10.1038/s41419-018-1291-5.
130. Saha SK, Islam SMR, Abdullah-Al-Wadud M, Islam S, Ali F, Park KS. Multiomics Analysis Reveals that GLS and GLS2 Differentially Modulate the Clinical Outcomes of Cancer. J Clin Med. 2019;8(3):355. doi: 10.3390/jcm8030355. PubMed PMID: 30871151.
131. Gross M, JC, JE, 2, JJ, 1, et al. Glutaminase Inhibition With CB-839 Enhances Anti-Tumor Activity of PD-1 and PD-L1 Antibodies by Overcoming a Metabolic Checkpoint Blocking T Cell Activation 2016 [20/01/2018]. Available from: https://www.calithera.com/wp-content/uploads/2016/04/04.2016-2016-AACR-Poster-Gross-et-al-FINAL.pdf.
132. Katt WP, Lukey MJ, Cerione RA. A tale of two glutaminases: homologous enzymes with distinct roles in tumorigenesis. Future Med Chem. 2017;9(2):223-43. Epub 01/23. doi: 10.4155/fmc-2016-0190. PubMed PMID: 28111979.
133. Huang F, Zhang Q, Ma H, Lv Q, Zhang T. Expression of glutaminase is upregulated in colorectal cancer and of clinical significance. Int J Clin Exp Pathol. 2014;7(3):1093-100. PubMed PMID: 24696726.
134. Pan T, Gao L, Wu G, Shen G, Xie S, Wen H, et al. Elevated expression of glutaminase confers glucose utilization via glutaminolysis in prostate cancer. Biochemical and Biophysical Research Communications. 2015;456(1):452-8. doi: https://doi.org/10.1016/j.bbrc.2014.11.105.
135. Martino JJ, Wall BA, Mastrantoni E, Wilimczyk BJ, La Cava SN, Degenhardt K, et al. Metabotropic glutamate receptor 1 (Grm1) is an oncogene in epithelial cells. Oncogene. 2013;32(37):4366-76. doi: 10.1038/onc.2012.471.
136. Liu Y, Yang L, An H, Chang Y, Zhang W, Zhu Y, et al. High expression of Solute Carrier Family 1, member 5 (SLC1A5) is associated with poor prognosis in clear-cell renal cell carcinoma. Sci Rep. 2015;5:16954. Epub 2015/11/26. doi: 10.1038/srep16954. PubMed PMID: 26599282; PubMed Central PMCID: PMCPMC4657035.
137. Okazaki A, Gameiro PA, Christodoulou D, Laviollette L, Schneider M, Chaves F, et al. Glutaminase and poly(ADP-ribose) polymerase inhibitors suppress pyrimidine synthesis and VHL-deficient renal cancers. The Journal of Clinical Investigation. 2017;127(5):1631-45. doi: 10.1172/JCI87800. PubMed PMID: PMC5409089.
138. Iyer NV, Kotch LE, Agani F, Leung SW, Laughner E, Wenger RH, et al. Cellular and developmental control of O2 homeostasis by hypoxia-inducible factor 1 alpha. Genes & development. 1998;12(2):149-62. doi: 10.1101/gad.12.2.149. PubMed PMID: 9436976.
139. Kim JW, Tchernyshyov I, Semenza GL, Dang CV. HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia. Cell Metab. 2006;3(3):177-85. Epub 2006/03/07. doi: 10.1016/j.cmet.2006.02.002. PubMed PMID: 16517405.
140. Gameiro PA, Yang J, Metelo AM, Pérez-Carro R, Baker R, Wang Z, et al. In vivo HIF-mediated reductive carboxylation is regulated by citrate levels and sensitizes VHL-deficient cells to glutamine deprivation. Cell Metab. 2013;17(3):372-85. Epub 2013/03/12. doi:

10.1016/j.cmet.2013.02.002. PubMed PMID: 23473032; PubMed Central PMCID: PMCPMC4003458.
141. Wise DR, Ward PS, Shay JE, Cross JR, Gruber JJ, Sachdeva UM, et al. Hypoxia promotes isocitrate dehydrogenase-dependent carboxylation of α-ketoglutarate to citrate to support cell growth and viability. Proc Natl Acad Sci U S A. 2011;108(49):19611-6. Epub 2011/11/23. doi: 10.1073/pnas.1117773108. PubMed PMID: 22106302; PubMed Central PMCID: PMCPMC3241793.
142. Le A, Lane AN, Hamaker M, Bose S, Gouw A, Barbi J, et al. Glucose-independent glutamine metabolism via TCA cycling for proliferation and survival in B cells. Cell Metab. 2012;15(1):110-21. Epub 2012/01/10. doi: 10.1016/j.cmet.2011.12.009. PubMed PMID: 22225880; PubMed Central PMCID: PMCPMC3345194.
143. Lukey MJ, Wilson KF, Cerione RA. Therapeutic strategies impacting cancer cell glutamine metabolism. Future Med Chem. 2013;5(14):1685-700. doi: 10.4155/fmc.13.130. PubMed PMID: 24047273.
144. Ahluwalia GS, Grem JL, Hao Z, Cooney DA. Metabolism and action of amino acid analog anti-cancer agents. Pharmacology & Therapeutics. 1990;46(2):243-71. doi: https://doi.org/10.1016/0163-7258(90)90094-I.
145. Magill GB, Myers WP, Reilly HC, Putnam RC, Magill JW, Sykes MP, et al. Pharmacological and initial therapeutic observations on 6-diazo-5-oxo-1-norleucine (DON) in human neoplastic disease. Cancer. 1957;10(6):1138-50. Epub 1957/11/01. doi: 10.1002/1097-0142(195711/12)10:6<1138::aid-cncr2820100608>3.0.co;2-k. PubMed PMID: 13489662.
146. Veterans Administration Cancer Chemotherapy Study G, Krantz S, Rivers S, Dwight RW, Corbus HF, Wolf J, et al. A Clinical Study of the Comparative Effect of Nitrogen Mustard and DON in Patients With Bronchogenic Carcinoma, Hodgkin's Disease, Lymphosarcoma, and Melanoma2. JNCI: Journal of the National Cancer Institute. 1959;22(2):433-9. doi: 10.1093/jnci/22.2.433.
147. Sullivan MP, Nelson JA, Feldman S, Van Nguyen B. Pharmacokinetic and phase I study of intravenous DON (6-diazo-5-oxo-L-norleucine) in children. Cancer chemotherapy and pharmacology. 1988;21(1):78-84. Epub 1988/01/01. doi: 10.1007/bf00262746. PubMed PMID: 3342470.
148. Rahman A, Smith FP, Luc PT, Woolley PV. Phase I study and clinical pharmacology of 6-diazo-5-oxo-L-norleucine (DON). Investigational new drugs. 1985;3(4):369-74. Epub 1985/01/01. doi: 10.1007/bf00170760. PubMed PMID: 4086244.
149. Earhart RH, Amato DJ, Chang AY, Borden EC, Shiraki M, Dowd ME, et al. Phase II trial of 6-diazo-5-oxo-L-norleucine versus aclacinomycin-A in advanced sarcomas and mesotheliomas. Investigational new drugs. 1990;8(1):113-9. Epub 1990/02/01. doi: 10.1007/bf00216936. PubMed PMID: 2188926.
150. Lynch G, Kemeny N, Casper E. Phase II evaluation of DON (6-diazo-5-oxo-L-norleucine) in patients with advanced colorectal carcinoma. American journal of clinical oncology. 1982;5(5):541-3. Epub 1982/10/01. PubMed PMID: 7180833.
151. Mueller C, Al-Batran S, Jaeger E, Schmidt B, Bausch M, Unger C, et al. A phase IIa study of PEGylated glutaminase (PEG-PGA) plus 6-diazo-5-oxo-L-norleucine (DON) in patients with advanced refractory solid tumors. Journal of Clinical Oncology. 2008;26(15_suppl):2533-. doi: 10.1200/jco.2008.26.15_suppl.2533.
152. Hartman SC. THE INTERACTION OF 6-DIAZO-5-OXO-L-NORLEUCINE WITH PHOSPHORIBOSYL PYROPHOSPHATE AMIDOTRANSFERASE. J Biol Chem. 1963;238:3036-47. Epub 1963/09/01. PubMed PMID: 14081921.
153. Levenberg B, Melnick I, Buchanan JM. Biosynthesis of the purines. XV. The effect of aza-L-serine and 6-diazo-5-oxo-L-norleucine on inosinic acid biosynthesis de novo. J Biol Chem. 1957;225(1):163-76. Epub 1957/03/01. PubMed PMID: 13416227.

154. Lemberg KM, Vornov JJ, Rais R, Slusher BS. We're Not "DON" Yet: Optimal Dosing and Prodrug Delivery of 6-Diazo-5-oxo-L-norleucine. Molecular cancer therapeutics. 2018;17(9):1824-32. doi: 10.1158/1535-7163.MCT-17-1148. PubMed PMID: 30181331.
155. Crosby HA, Ihnat M, Miller KE. Evaluating the Toxicity of the Analgesic Glutaminase Inhibitor 6-Diazo-5-Oxo-L-Norleucine in vitro and on Rat Dermal Skin Fibroblasts. MOJ Toxicol. 2015;1(1):00005. Epub 03/25. doi: 10.15406/mojt.2015.01.00005. PubMed PMID: 29750203.
156. DeLaBarre B, Gross S, Fang C, Gao Y, Jha A, Jiang F, et al. Full-Length Human Glutaminase in Complex with an Allosteric Inhibitor. Biochemistry. 2011;50(50):10764-70. doi: 10.1021/bi201613d.
157. Robinson M, McBryant S, Tsukamoto T, Rojas C, Ferraris D, Hamilton S, et al. Novel mechanism of inhibition of rat kidney-type glutaminase by bis-2-(5-phenylacetamido-1,2,4-thiadiazol-2-yl)ethyl sulfide (BPTES). The Biochemical Journal. 2007;406(3):407-14. Epub 08/29. doi: 10.1042/BJ20070039. PubMed PMID: 17581113.
158. Shukla K, Ferraris D, Thomas A, Stathis M, Duvall B, Delahanty G, et al. Design, synthesis, and pharmacological evaluation of bis-2-(5-phenylacetamido-1,2,4-thiadiazol-2-yl)ethyl sulfide 3 (BPTES) analogs as glutaminase inhibitors. Journal of medicinal chemistry. 2012;55(23):10551-63. Epub 11/30. doi: 10.1021/jm301191p. PubMed PMID: 23151085.
159. Stalnecker CA, Ulrich SM, Li Y, Ramachandran S, McBrayer MK, DeBerardinis RJ, et al. Mechanism by which a recently discovered allosteric inhibitor blocks glutamine metabolism in transformed cells. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(2):394-9. Epub 12/29. doi: 10.1073/pnas.1414056112. PubMed PMID: 25548170.
160. Simpson NE, Tryndyak VP, Beland FA, Pogribny IP. An in vitro investigation of metabolically sensitive biomarkers in breast cancer progression. Breast cancer research and treatment. 2012;133(3):959-68.
161. Katt WP, Ramachandran S, Erickson JW, Cerione RA. Dibenzophenanthridines as inhibitors of glutaminase C and cancer cell proliferation. Molecular cancer therapeutics. 2012;11(6):1269-78.
162. Matre P, Velez J, Jacamo R, Qi Y, Su X, Cai T, et al. Inhibiting glutaminase in acute myeloid leukemia: metabolic dependency of selected AML subtypes. Oncotarget. 2016;7(48):79722-35. doi: 10.18632/oncotarget.12944. PubMed PMID: PMC5340236.
163. Parlati F, Gross M, Janes J, Lewis E, MacKinnon A, Rodriguez M, et al. Glutaminase Inhibitor CB-839 Synergizes with Pomalidomide in Preclinical Multiple Myeloma Models. Blood. 2014;124(21):4720.
164. Gross MI, Demo SD, Dennison JB, Chen L, Chernov-Rogan T, Goyal B, et al. Antitumor Activity of the Glutaminase Inhibitor CB-839 in Triple-Negative Breast Cancer. Molecular Cancer Therapeutics. 2014;13(4):890.
165. Lampa M, Arlt H, He T, Ospina B, Reeves J, Zhang B, et al. Glutaminase is essential for the growth of triple-negative breast cancer cells with a deregulated glutamine metabolism pathway and its suppression synergizes with mTOR inhibition. PLOS ONE. 2017;12(9):e0185092. doi: 10.1371/journal.pone.0185092.
166. Parlati F. CB-839, a selective glutaminase inhibitor, synergizes with signaling pathway inhibitors to produce an anti-tumor effect in cell lines and tumor xenografts [11 Dec 2018]. Available from: https://www.calithera.com/wp-content/uploads/2017/12/AACR-2015-final.pdf.
167. Study of the Glutaminase Inhibitor CB-839 in Solid Tumors [11 Dec 2018]. Available from: https://clinicaltrials.gov/ct2/show/NCT02071862.
168. Harding J, Telli M, Munster P, Le M, Molineaux C, Bennett M, et al. Safety and tolerability of increasing doses of CB-839, a first-in-class, orally administered small molecule inhibitor of

glutaminase, in solid tumors. Journal of Clinical Oncology. 2015;33(15_suppl):2512-. doi: 10.1200/jco.2015.33.15_suppl.2512.
169. Meric-Bernstam F, Tannir N, Mier J, DeMichele A, Telli M, Fan A, et al. Phase 1 study of CB-839, a small molecule inhibitor of glutaminase (GLS), alone and in combination with everolimus (E) in patients (pts) with renal cell cancer (RCC). Journal of Clinical Oncology. 2016;34(15_suppl):4568-. doi: 10.1200/JCO.2016.34.15_suppl.4568.
170. Emberley E, Bennett M, Chen J, Gross M, Huang T, Li W, et al. CB-839, a Selective Glutaminase Inhibitor, has Anti-Tumor Activity in Renal Cell Carcinoma and Synergizes with Cabozantinib and Everolimus [11 Dec 2018]. Available from: https://www.calithera.com/wp-content/uploads/2017/12/03.2017-Keystone-poster-Emberley-2017.pdf.
171. Boyd MR, Paull KD. Some practical considerations and applications of the national cancer institute in vitro anticancer drug discovery screen. Drug Development Research. 1995;34(2):91-109. doi: 10.1002/ddr.430340203.
172. Shoemaker RH. The NCI60 human tumour cell line anticancer drug screen. Nature Reviews Cancer. 2006;6(10):813-23. doi: 10.1038/nrc1951.
173. Holbeck SL, Collins JM, Doroshow JH. Analysis of Food and Drug Administration-approved anticancer agents in the NCI60 panel of human tumor cell lines. Molecular cancer therapeutics. 2010;9(5):1451-60. Epub 05/04. doi: 10.1158/1535-7163.MCT-10-0106. PubMed PMID: 20442306.
174. Kong D, Yamori T. JFCR39, a panel of 39 human cancer cell lines, and its application in the discovery and development of anticancer drugs. Bioorganic & medicinal chemistry. 2012;20(6):1947-51. Epub 2012/02/18. doi: 10.1016/j.bmc.2012.01.017. PubMed PMID: 22336246.
175. Haibe-Kains B, El-Hachem N, Birkbak NJ, Jin AC, Beck AH, Aerts HJ, et al. Inconsistency in large pharmacogenomic studies. Nature. 2013;504(7480):389-93.
176. Domcke S, Sinha R, Levine DA, Sander C, Schultz N. Evaluating cell lines as tumour models by comparison of genomic profiles. Nat Commun. 2013;4:2126. Epub 2013/07/11. doi: 10.1038/ncomms3126. PubMed PMID: 23839242; PubMed Central PMCID: PMCPMC3715866.
177. Gillet JP, Calcagno AM, Varma S, Marino M, Green LJ, Vora MI, et al. Redefining the relevance of established cancer cell lines to the study of mechanisms of clinical anti-cancer drug resistance. Proc Natl Acad Sci U S A. 2011;108(46):18708-13. Epub 2011/11/10. doi: 10.1073/pnas.1111840108. PubMed PMID: 22068913; PubMed Central PMCID: PMCPMC3219108.
178. Bhadriraju K, Chen CS. Engineering cellular microenvironments to improve cell-based drug testing. Drug discovery today. 2002;7(11):612-20.
179. Langdon SP. Basic principles of cancer cell culture. Cancer Cell Culture: Springer; 2004. p. 3-15.
180. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. Nature Reviews Genetics. 2010;11(9):636-46.
181. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol. 2013;14(1):49-55. Epub 11/21. doi: 10.1038/nrm3486. PubMed PMID: 23169466.
182. Ramirez CL, Foley JE, Wright DA, Müller-Lerch F, Rahman SH, Cornu TI, et al. Unexpected failure rates for modular assembly of engineered zinc fingers. Nature methods. 2008;5(5):374-5.
183. Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppel C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nature biotechnology. 2011;29(9):816.
184. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. Nature methods. 2011;8(9):765.

185. Boch J, Bonas U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annual review of phytopathology. 2010;48:419-36.
186. Kim Y, Kweon J, Kim A, Chon JK, Yoo JY, Kim HJ, et al. A library of TAL effector nucleases spanning the human genome. Nature biotechnology. 2013;31(3):251.
187. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol. 2012;30(5):460-5. Epub 2012/04/10. doi: 10.1038/nbt.2170. PubMed PMID: 22484455; PubMed Central PMCID: PMCPMC3558947.
188. Smith C, Gore A, Yan W, Abalde-Atristain L, Li Z, He C, et al. Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. Cell stem cell. 2014;15(1):12-3.
189. Gaj T, Gersbach CA, Barbas CF, III. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends in Biotechnology. 2013;31(7):397-405. doi: 10.1016/j.tibtech.2013.04.004.
190. Rad R, Rad L, Wang W, Cadinanos J, Vassiliou G, Rice S, et al. PiggyBac transposon mutagenesis: a tool for cancer gene discovery in mice. Science (New York, NY). 2010;330(6007):1104-7. Epub 10/14. doi: 10.1126/science.1193004. PubMed PMID: 20947725.
191. Balling R. ENU mutagenesis: analyzing gene function in mice. Annual review of genomics and human genetics. 2001;2(1):463-92.
192. Morris SM, Mhyre AJ, Carmack SS, Myers CH, Burns C, Ye W, et al. A modified gene trap approach for improved high-throughput cancer drug discovery. Oncogene. 2018;37(31):4226-38. doi: 10.1038/s41388-018-0274-4.
193. Napoli C, Lemieux C, Jorgensen R. Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. The plant cell. 1990;2(4):279-89.
194. Van Blokland R, Van der Geest N, Mol J, Kooter J. Transgene-mediated suppression of chalcone synthase expression in Petunia hybrida results from an increase in RNA turnover. The Plant Journal. 1994;6(6):861-77.
195. Ingelbrecht I, Van Houdt H, Van Montagu M, Depicker A. Posttranscriptional silencing of reporter transgenes in tobacco correlates with DNA methylation. Proceedings of the National Academy of Sciences. 1994;91(22):10502-6.
196. Van der Krol AR, Mur LA, Beld M, Mol J, Stuitje AR. Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. The Plant Cell. 1990;2(4):291-9.
197. Cogoni C, Irelan JT, Schumacher M, Schmidhauser TJ, Selker EU, Macino G. Transgene silencing of the al-1 gene in vegetative cells of Neurospora is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. The EMBO journal. 1996;15(12):3153-63.
198. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998;391(6669):806-11. doi: 10.1038/35888.
199. Timmons L, Fire A. Specific interference by ingested dsRNA. Nature. 1998;395(6705):854-.
200. Aagaard L, Rossi JJ. RNAi therapeutics: principles, prospects and challenges. Adv Drug Deliv Rev. 2007;59(2-3):75-86. Epub 03/16. doi: 10.1016/j.addr.2007.03.005. PubMed PMID: 17449137.
201. Bauer JA, Ye F, Marshall CB, Lehmann BD, Pendleton CS, Shyr Y, et al. RNA interference (RNAi) screening approach identifies agents that enhance paclitaxel activity in breast cancer cells. Breast Cancer Research. 2010;12(3):R41. doi: 10.1186/bcr2595.

202. Williams SP, Barthorpe AS, Lightfoot H, Garnett MJ, McDermott U. High-throughput RNAi screen for essential genes and drug synergistic combinations in colorectal cancer. Scientific Data. 2017;4(1):170139. doi: 10.1038/sdata.2017.139.
203. Gerlinger M, Santos CR, Spencer-Dene B, Martinez P, Endesfelder D, Burrell RA, et al. Genome-wide RNA interference analysis of renal carcinoma survival regulators identifies MCT4 as a Warburg effect metabolic target. J Pathol. 2012;227(2):146-56. doi: 10.1002/path.4006. PubMed PMID: 22362593.
204. Burgess DJ, Doles J, Zender L, Xue W, Ma B, McCombie WR, et al. Topoisomerase levels determine chemotherapy response in vitro and in vivo. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(26):9053-8. Epub 06/23. doi: 10.1073/pnas.0803513105. PubMed PMID: 18574145.
205. Jackson AL, Linsley PS. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. Nature Reviews Drug Discovery. 2010;9(1):57-67. doi: 10.1038/nrd3010.
206. Hornung V, Guenthner-Biller M, Bourquin C, Ablasser A, Schlee M, Uematsu S, et al. Sequence-specific potent induction of IFN-α by short interfering RNA in plasmacytoid dendritic cells through TLR7. Nature medicine. 2005;11(3):263-70.
207. Reynolds A, Anderson EM, Vermeulen A, Fedorov Y, Robinson K, Leake D, et al. Induction of the interferon response by siRNA is cell type–and duplex length–dependent. Rna. 2006;12(6):988-93.
208. Jacobsen LB, Calvin SA, Lobenhofer EK. Transcriptional effects of transfection: the potential for misinterpretation of gene expression data generated from transiently transfected cells. Biotechniques. 2009;47(1):617-24.
209. Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, et al. Expression profiling reveals off-target gene regulation by RNAi. Nature biotechnology. 2003;21(6):635-7.
210. Birmingham A, Anderson EM, Reynolds A, Ilsley-Tyree D, Leake D, Fedorov Y, et al. 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. Nature methods. 2006;3(3):199-204.
211. Jackson AL, Burchard J, Leake D, Reynolds A, Schelter J, Guo J, et al. Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. Rna. 2006;12(7):1197-205.
212. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. Journal of Bacteriology. 1987;169(12):5429-33. PubMed PMID: PMC213968.
213. Mojica FJM, Díez-Villaseñor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Molecular Microbiology. 2000;36(1):244-6. doi: 10.1046/j.1365-2958.2000.01838.x.
214. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology. 2005;151(8):2551-61. doi: doi:10.1099/mic.0.28048-0.
215. Tang T-H, Bachellerie J-P, Rozhdestvensky T, Bortolin M-L, Huber H, Drungowski M, et al. Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus. Proceedings of the National Academy of Sciences. 2002;99(11):7536-41. doi: 10.1073/pnas.112047299.
216. Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic Acids Research. 2011;39(21):9275-82. doi: 10.1093/nar/gkr606. PubMed PMID: PMC3241640.
217. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National

Academy of Sciences of the United States of America. 2012;109(39):E2579-E86. doi: 10.1073/pnas.1208507109. PubMed PMID: PMC3465414.
218. Shi J, Wang E, Milazzo JP, Wang Z, Kinney JB, Vakoc CR. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nature biotechnology. 2015;33(6):661-7. doi: 10.1038/nbt.3235. PubMed PMID: PMC4529991.
219. Liu J-q, Li T. CRISPR-Cas9-mediated loss-of-function screens. Frontiers in Life Science. 2019;12(1):1-13. doi: 10.1080/21553769.2019.1670739.
220. Stark GR, Gudkov AV. Forward genetics in mammalian cells: functional approaches to gene discovery. Human molecular genetics. 1999;8(10):1925-38.
221. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science. 2014;343(6166):84-7. Epub 2013/12/18. doi: 10.1126/science.1247005. PubMed PMID: 24336571; PubMed Central PMCID: PMCPMC4089965.
222. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR/Cas9 system. Science (New York, NY). 2014;343(6166):80-4. doi: 10.1126/science.1246981. PubMed PMID: PMC3972032.
223. Zhang X-H, Tee LY, Wang X-G, Huang Q-S, Yang S-H. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. Molecular Therapy - Nucleic Acids. 2015;4:e264. doi: https://doi.org/10.1038/mtna.2015.37.
224. Haapaniemi E, Botla S, Persson J, Schmierer B, Taipale J. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. Nature Medicine. 2018;24(7):927-30. doi: 10.1038/s41591-018-0049-z.
225. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biology Direct. 2006;1:7-. doi: 10.1186/1745-6150-1-7. PubMed PMID: PMC1462988.
226. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011;471(7340):602-7. doi: 10.1038/nature09886. PubMed PMID: PMC3070239.
227. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 2012;337(6096):816.
228. Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature. 2014;513(7519):569-73. doi: 10.1038/nature13579. PubMed PMID: PMC4176945.
229. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2014;507(7490):62-7. doi: 10.1038/nature13011. PubMed PMID: PMC4106473.
230. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-Guided Human Genome Engineering via Cas9. Science (New York, NY). 2013;339(6121):823-6. doi: 10.1126/science.1232033. PubMed PMID: PMC3712628.
231. Ihry RJ, Salick MR, Ho DJ, Sondey M, Kommineni S, Paula S, et al. Genome-Scale CRISPR Screens Identify Human Pluripotency-Specific Genes. Cell Rep. 2019;27(2):616-30.e6. doi: 10.1016/j.celrep.2019.03.043. PubMed PMID: 30970262.
232. Gebre M, Nomburg JL, Gewurz BE. CRISPR-Cas9 Genetic Analysis of Virus-Host Interactions. Viruses. 2018;10(2). Epub 2018/02/02. doi: 10.3390/v10020055. PubMed PMID: 29385696; PubMed Central PMCID: PMCPMC5850362.
233. Lau M-T, Ghazanfar S, Parkin A, Chou A, Rouaen JR, Littleboy JB, et al. Systematic functional identification of cancer multi-drug resistance genes. Genome biology. 2020;21(1):27-. doi: 10.1186/s13059-020-1940-8. PubMed PMID: 32028983.

234. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biology. 2014;15(12):554. doi: 10.1186/s13059-014-0554-4.
235. Hart T, Moffat J. BAGEL: a computational framework for identifying essential genes from pooled library screens. BMC Bioinformatics. 2016;17(1):164. doi: 10.1186/s12859-016-1015-8.
236. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nature biotechnology. 2014;32(12):1262-7. Epub 09/03. doi: 10.1038/nbt.3026. PubMed PMID: 25184501.
237. Cao J, Wei J, Yang P, Zhang T, Chen Z, He F, et al. Genome-scale CRISPR-Cas9 knockout screening in gastrointestinal stromal tumor with Imatinib resistance. Molecular Cancer. 2018;17(1):121. doi: 10.1186/s12943-018-0865-2.
238. Wei L, Lee D, Law C-T, Zhang MS, Shen J, Chin DW-C, et al. Genome-wide CRISPR/Cas9 library screening identified PHGDH as a critical driver for Sorafenib resistance in HCC. Nature Communications. 2019;10(1):4681. doi: 10.1038/s41467-019-12606-7.
239. Han P, Dai Q, Fan L, Lin H, Zhang X, Li F, et al. Genome-Wide CRISPR Screening Identifies JAK1 Deficiency as a Mechanism of T-Cell Resistance. Front Immunol. 2019;10:251-. doi: 10.3389/fimmu.2019.00251. PubMed PMID: 30837996.
240. Arroyo JD, Jourdain AA, Calvo SE, Ballarano CA, Doench JG, Root DE, et al. A Genome-wide CRISPR Death Screen Identifies Genes Essential for Oxidative Phosphorylation. Cell metabolism. 2016;24(6):875-85. Epub 09/22. doi: 10.1016/j.cmet.2016.08.017. PubMed PMID: 27667664.
241. Flint M, Chatterjee P, Lin DL, McMullan LK, Shrivastava-Ranjan P, Bergeron É, et al. A genome-wide CRISPR screen identifies N-acetylglucosamine-1-phosphate transferase as a potential antiviral target for Ebola virus. Nature Communications. 2019;10(1):285. doi: 10.1038/s41467-018-08135-4.
242. Hou P, Wu C, Wang Y, Qi R, Bhavanasi D, Zuo Z, et al. A Genome-Wide CRISPR Screen Identifies Genes Critical for Resistance to FLT3 Inhibitor AC220. Cancer research. 2017;77(16):4402-13. Epub 06/16. doi: 10.1158/0008-5472.CAN-16-1627. PubMed PMID: 28625976.
243. Hayman TJ, Baro M, Cui W, Contessa JN. Whole Genome CRISPR-Cas9 Screen Identifies STING as a Determinant of Intrinsic Radiosensitivity in Head and Neck Squamous Cell Carcinoma. International Journal of Radiation Oncology • Biology • Physics. 2019;105(1):S77-S8. doi: 10.1016/j.ijrobp.2019.06.538.
244. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nature Methods. 2014;11(8):783-4. doi: 10.1038/nmeth.3047.
245. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America. 1977;74(12):5463-7. doi: 10.1073/pnas.74.12.5463. PubMed PMID: 271968.
246. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science (New York, NY). 1995;269(5223):496-512. doi: 10.1126/science.7542800. PubMed PMID: 7542800.
247. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, et al. The genome sequence of Drosophila melanogaster. Science (New York, NY). 2000;287(5461):2185-95. doi: 10.1126/science.287.5461.2185. PubMed PMID: 10731132.
248. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The Sequence of the Human Genome. Science. 2001;291(5507):1304. doi: 10.1126/science.1058040.
249. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860-921. Epub 2001/03/10. doi: 10.1038/35057062. PubMed PMID: 11237011.

250. Bodapati S, Daley TP, Lin X, Zou J, Qi LS. A benchmark of algorithms for the analysis of pooled CRISPR screens. Genome Biology. 2020;21(1):62. doi: 10.1186/s13059-020-01972-x.
251. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England). 2010;26(1):139-40. Epub 2009/11/11. doi: 10.1093/bioinformatics/btp616. PubMed PMID: 19910308.
252. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biology. 2010;11(10):R106. doi: 10.1186/gb-2010-11-10-r106.
253. Hardcastle TJ, Kelly KA. baySeq: Empirical Bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics. 2010;11(1):422. doi: 10.1186/1471-2105-11-422.
254. König R, Chiang CY, Tu BP, Yan SF, DeJesus PD, Romero A, et al. A probability-based approach for the analysis of large-scale RNAi screens. Nat Methods. 2007;4(10):847-9. Epub 2007/09/11. doi: 10.1038/nmeth1089. PubMed PMID: 17828270.
255. Luo B, Cheung HW, Subramanian A, Sharifnia T, Okamoto M, Yang X, et al. Highly parallel identification of essential genes in cancer cells. Proceedings of the National Academy of Sciences. 2008;105(51):20380. doi: 10.1073/pnas.0810485105.
256. Li W, Köster J, Xu H, Chen C-H, Xiao T, Liu JS, et al. Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. Genome Biology. 2015;16(1):281. doi: 10.1186/s13059-015-0843-6.
257. Hwang TJ, Carpenter D, Lauffenburger JC, Wang B, Franklin JM, Kesselheim AS. Failure of investigational drugs in late-stage clinical development and publication of trial results. JAMA internal medicine. 2016;176(12):1826-33.
258. Bissell M, Kenny P, Radisky DC, editors. Microenvironmental regulators of tissue structure and function also regulate tumor induction and progression: the role of extracellular matrix and its degrading enzymes. Cold Spring Harbor symposia on quantitative biology; 2005: Cold Spring Harbor Laboratory Press.
259. Wirtz D, Konstantopoulos K, Searson PC. The physics of cancer: the role of physical interactions and mechanical forces in metastasis. Nature Reviews Cancer. 2011;11(7):512-22.
260. Luca AC, Mersch S, Deenen R, Schmidt S, Messner I, Schäfer K-L, et al. Impact of the 3D microenvironment on phenotype, gene expression, and EGFR inhibition of colorectal cancer cell lines. PloS one. 2013;8(3):e59689-e. Epub 03/26. doi: 10.1371/journal.pone.0059689. PubMed PMID: 23555746.
261. Wang X, Sun L, Maffini MV, Soto A, Sonnenschein C, Kaplan DL. A complex 3D human tissue culture system based on mammary stromal cells and silk scaffolds for modeling breast morphogenesis and function. Biomaterials. 2010;31(14):3920-9.
262. Goers L, Freemont P, Polizzi KM. Co-culture systems and technologies: taking synthetic biology to the next level. Journal of the Royal Society, Interface. 2014;11(96). Epub 2014/05/16. doi: 10.1098/rsif.2014.0065. PubMed PMID: 24829281; PubMed Central PMCID: PMCPMC4032528.
263. Luca AC, Mersch S, Deenen R, Schmidt S, Messner I, Schäfer K-L, et al. Impact of the 3D microenvironment on phenotype, gene expression, and EGFR inhibition of colorectal cancer cell lines. PloS one. 2013;8(3).
264. Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, et al. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. Cell. 2015;160(6):1246-60. Epub 2015/03/10. doi: 10.1016/j.cell.2015.02.038. PubMed PMID: 25748654; PubMed Central PMCID: PMCPMC4380877.
265. Katigbak A, Cencic R, Robert F, Sénécha P, Scuoppo C, Pelletier J. A CRISPR/Cas9 functional screen identifies rare tumor suppressors. Scientific reports. 2016;6(1):1-8.

266. Song C-Q, Li Y, Mou H, Moore J, Park A, Pomyen Y, et al. Genome-wide CRISPR screen identifies regulators of mitogen-activated protein kinase as suppressors of liver tumors in mice. Gastroenterology. 2017;152(5):1161-73. e1.
267. Kodama M, Kodama T, Newberg JY, Katayama H, Kobayashi M, Hanash SM, et al. In vivo loss-of-function screens identify KPNB1 as a new druggable oncogene in epithelial ovarian cancer. Proceedings of the National Academy of Sciences. 2017;114(35):E7301-E10.
268. Yau EH, Kummetha IR, Lichinchi G, Tang R, Zhang Y, Rana TM. Genome-wide CRISPR screen for essential cell growth mediators in mutant KRAS colorectal cancers. Cancer research. 2017;77(22):6330-9.
269. Manguso RT, Pope HW, Zimmer MD, Brown FD, Yates KB, Miller BC, et al. In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. Nature. 2017;547(7664):413-8.
270. Chow RD, Chen S. Cancer CRISPR Screens In Vivo. Trends in cancer. 2018;4(5):349-58. Epub 03/30. doi: 10.1016/j.trecan.2018.03.002. PubMed PMID: 29709259.
271. Shajahan-Haq AN, Cheema MS, Clarke R. Application of metabolomics in drug resistant breast cancer research. Metabolites. 2015;5(1):100-18. doi: 10.3390/metabo5010100. PubMed PMID: 25693144.
272. Lécuyer L, Victor Bala A, Deschasaux M, Bouchemal N, Nawfal Triba M, Vasson M-P, et al. NMR metabolomic signatures reveal predictive plasma metabolites associated with long-term risk of developing breast cancer. International Journal of Epidemiology. 2018;47(2):484-94. doi: 10.1093/ije/dyx271.
273. Cardoso MR, Santos JC, Ribeiro ML, Talarico MCR, Viana LR, Derchain SFM. A Metabolomic Approach to Predict Breast Cancer Behavior and Chemotherapy Response. International journal of molecular sciences. 2018;19(2):617. doi: 10.3390/ijms19020617. PubMed PMID: 29466297.
274. Halama A, Kulinski M, Dib SS, Zaghlool SB, Siveen KS, Iskandarani A, et al. Accelerated lipid catabolism and autophagy are cancer survival mechanisms under inhibited glutaminolysis. bioRxiv. 2018:230433. doi: 10.1101/230433.
275. Biancur DE, Paulo JA, Małachowska B, Quiles Del Rey M, Sousa CM, Wang X, et al. Compensatory metabolic networks in pancreatic cancers upon perturbation of glutamine metabolism. Nature communications. 2017;8:15965-. doi: 10.1038/ncomms15965. PubMed PMID: 28671190.
276. Biedler JL, Riehm H. Cellular resistance to actinomycin D in Chinese hamster cells in vitro: cross-resistance, radioautographic, and cytogenetic studies. Cancer Res. 1970;30(4):1174-84. Epub 1970/04/01. PubMed PMID: 5533992.
277. Roberts D, Wodinsky I, Hall TC. Studies on folic reductase. 3. The level of enzyme activity and response to methotrexate of transplantable mouse tumors. Cancer Res. 1965;25(11):1899-903. Epub 1965/12/01. PubMed PMID: 5856519.
278. Breen L, Murphy L, Keenan J, Clynes M. Development of taxane resistance in a panel of human lung cancer cell lines. Toxicology in vitro : an international journal published in association with BIBRA. 2008;22(5):1234-41. Epub 2008/06/03. doi: 10.1016/j.tiv.2008.04.005. PubMed PMID: 18514476.
279. McDermott M, Eustace AJ, Busschots S, Breen L, Crown J, Clynes M, et al. In vitro Development of Chemotherapy and Targeted Therapy Drug-Resistant Cancer Cell Lines: A Practical Guide with Case Studies. Front Oncol. 2014;4:40-. doi: 10.3389/fonc.2014.00040. PubMed PMID: 24639951.
280. Mashal RD, Koontz J, Sklar J. Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. Nature genetics. 1995;9(2):177-83.
281. Hendel A, Kildebeck EJ, Fine EJ, Clark J, Punjya N, Sebastiano V, et al. Quantifying genome-editing outcomes at endogenous loci with SMRT sequencing. Cell Rep. 2014;7(1):293-305. Epub 2014/03/27. doi: 10.1016/j.celrep.2014.02.040. PubMed PMID: 24685129.

282. Wang K, Mei DY, Liu QN, Qiao XH, Ruan WM, Huang T, et al. Research of methods to detect genomic mutations induced by CRISPR/Cas systems. Journal of biotechnology. 2015;214:128-32.
283. Doench JG. Am I ready for CRISPR? A user's guide to genetic screens. Nature Reviews Genetics. 2018;19(2):67-80. doi: 10.1038/nrg.2017.97.
284. Nathans D. PUROMYCIN INHIBITION OF PROTEIN SYNTHESIS: INCORPORATION OF PUROMYCIN INTO PEPTIDE CHAINS. Proceedings of the National Academy of Sciences. 1964;51(4):585.
285. Joung J, Konermann S, Gootenberg JS, Abudayyeh OO, Platt RJ, Brigham MD, et al. Genome-scale CRISPR-Cas9 Knockout and Transcriptional Activation Screening. Nature protocols. 2017;12(4):828-63. doi: 10.1038/nprot.2017.016. PubMed PMID: PMC5526071.
286. Segura MM, Mangion M, Gaillet B, Garnier A. New developments in lentiviral vector design, production and purification. Expert opinion on biological therapy. 2013;13(7):987-1011. Epub 2013/04/18. doi: 10.1517/14712598.2013.779249. PubMed PMID: 23590247.
287. Kanagawa T. Bias and artifacts in multitemplate polymerase chain reactions (PCR). Journal of bioscience and bioengineering. 2003;96(4):317-23.
288. Kalle E, Kubista M, Rensing C. Multi-template polymerase chain reaction. Biomolecular detection and quantification. 2014;2:11-29. Epub 2014/12/04. doi: 10.1016/j.bdq.2014.11.002. PubMed PMID: 27896140; PubMed Central PMCID: PMCPMC5121205.
289. Strezoska Ž, Licon A, Haimes J, Spayd KJ, Patel KM, Sullivan K, et al. Optimized PCR conditions and increased shRNA fold representation improve reproducibility of pooled shRNA screens. PloS one. 2012;7(8):e42341-e. Epub 08/01. doi: 10.1371/journal.pone.0042341. PubMed PMID: 22870320.
290. Hakimi AA, Chevinsky M, Hsieh JJ, Sander C, Sinha R. Mp23-11 genomic comparison of renal cell carcinoma cell lines to human tumors. The Journal of Urology. 2014.
291. Schmidt L, Junker K, Nakaigawa N, Kinjerski T, Weirich G, Miller M, et al. Novel mutations of the MET proto-oncogene in papillary renal carcinomas. Oncogene. 1999;18(14):2343-50.
292. Furge KA, Dykema K, Petillo D, Westphal M, Zhang Z, Kort EJ, et al. Combining differential expression, chromosomal and pathway analyses for the molecular characterization of renal cell carcinoma. Canadian Urological Association Journal. 2007;1(2 Suppl):S21.
293. Ioannidis JP. Why most published research findings are false. PLos med. 2005;2(8):e124.
294. Horbach SPJM, Halffman W. The ghosts of HeLa: How cell line misidentification contaminates the scientific literature. PLOS ONE. 2017;12(10):e0186281. doi: 10.1371/journal.pone.0186281.
295. Barallon R, Bauer SR, Butler J, Capes-Davis A, Dirks WG, Elmore E, et al. Recommendation of short tandem repeat profiling for authenticating human cell lines, stem cells, and tissues. In Vitro Cellular & Developmental Biology - Animal. 2010;46(9):727-32. doi: 10.1007/s11626-010-9333-z.
296. Williams RD, Elliott AY, Stein N, Fraley EE. In vitro cultivation of human renal cell cancer. In Vitro. 1978;14(9):779-86. doi: 10.1007/BF02617972.
297. Cowley GS, Weir BA, Vazquez F, Tamayo P, Scott JA, Rusin S, et al. Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies. Scientific Data. 2014;1:140035. doi: 10.1038/sdata.2014.35. PubMed PMID: PMC4432652.
298. COSMIC database. Available from: https://cancer.sanger.ac.uk/cell_lines.
299. 786-0 (ATCC). Available from: https://www.lgcstandards-atcc.org/products/all/CRL-1932.aspx.
300. 769-P (ATCC). Available from: https://www.lgcstandards-atcc.org/products/all/CRL-1933.aspx.

301. A704 (ATCC). Available from: https://www.lgcstandards-atcc.org/products/all/HTB-45.aspx.
302. Caki-1 (ATCC). Available from: https://www.lgcstandards-atcc.org/en/Global/Products/5/6/8/1/HTB-46.aspx.
303. A498 (ATCC). Available from: https://www.lgcstandards-atcc.org/products/all/HTB-44.aspx.
304. Kozlowski JM, Fidler IJ, Campbell D, Xu ZL, Kaighn ME, Hart IR. Metastatic behavior of human tumor cell lines grown in the nude mouse. Cancer Res. 1984;44(8):3522-9. Epub 1984/08/01. PubMed PMID: 6744277.
305. Huang B, Huang YJ, Yao ZJ, Chen X, Guo SJ, Mao XP, et al. Cancer stem cell-like side population cells in clear cell renal cell carcinoma cell line 769P. PloS one. 2013;8(7).
306. Shinojima T, Oya M, Takayanagi A, Mizuno R, Shimizu N, Murai M. Renal cancer cells lacking hypoxia inducible factor (HIF)-1alpha expression maintain vascular endothelial growth factor expression through HIF-2alpha. Carcinogenesis. 2007;28(3):529-36. Epub 2006/08/22. doi: 10.1093/carcin/bgl143. PubMed PMID: 16920734.
307. Ashida S, Nishimori I, Tanimura M, Onishi S, Shuin T. Effects of von Hippel-Lindau gene mutation and methylation status on expression of transmembrane carbonic anhydrases in renal cell carcinoma. Journal of cancer research and clinical oncology. 2002;128(10):561-8.
308. Varela I, Tarpey P, Raine K, Huang D, Ong CK, Stephens P, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature. 2011;469(7331):539-42. Epub 01/19. doi: 10.1038/nature09639. PubMed PMID: 21248752.
309. Pawłowski R, Mühl SM, Sulser T, Krek W, Moch H, Schraml P. Loss of PBRM1 expression is associated with renal cell carcinoma progression. International Journal of Cancer. 2013;132(2):E11-E7. doi: 10.1002/ijc.27822.
310. Shinojima T, Oya M, Takayanagi A, Mizuno R, Shimizu N, Murai M. Renal cancer cells lacking hypoxia inducible factor (HIF)-1α expression maintain vascular endothelial growth factor expression through HIF-2α. Carcinogenesis. 2007;28(3):529-36. doi: 10.1093/carcin/bgl143.
311. Murakami A, Wang L, Kalhorn S, Schraml P, Rathmell WK, Tan AC, et al. Context-dependent role for chromatin remodeling component PBRM1/BAF180 in clear cell renal cell carcinoma. Oncogenesis. 2017;6(1):e287-e. doi: 10.1038/oncsis.2016.89. PubMed PMID: 28092369.
312. Simbolo M, Gottardi M, Corbo V, Fassan M, Mafficini A, Malpeli G, et al. DNA Qualification Workflow for Next Generation Sequencing of Histopathological Samples. PLOS ONE. 2013;8(6):e62692. doi: 10.1371/journal.pone.0062692.
313. Sims D, Mendes-Pereira AM, Frankum J, Burgess D, Cerone MA, Lombardelli C, et al. High-throughput RNA interference screening using pooled shRNA libraries and next generation sequencing. Genome Biol. 2011;12(10):R104. Epub 2011/10/25. doi: 10.1186/gb-2011-12-10-r104. PubMed PMID: 22018332; PubMed Central PMCID: PMCPMC3333774.
314. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011. 2011;17(1):3. Epub 2011-08-02. doi: 10.14806/ej.17.1.200.
315. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25. Epub 2009/03/06. doi: 10.1186/gb-2009-10-3-r25. PubMed PMID: 19261174; PubMed Central PMCID: PMCPMC2690996.
316. Wang B, Wang M, Zhang W, Xiao T, Chen C-H, Wu A, et al. Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. Nature Protocols. 2019;14(3):756-80. doi: 10.1038/s41596-018-0113-7.
317. Hinze L, Pfirrmann M, Karim S, Degar J, McGuckin C, Vinjamur D, et al. Synthetic lethality of Wnt pathway activation and asparaginase in drug-resistant acute leukemias. Cancer cell. 2019;35(4):664-76. e7.

318. Wang C, Vegna S, Jin H, Benedict B, Lieftink C, Ramirez C, et al. Inducing and exploiting vulnerabilities for the treatment of liver cancer. Nature. 2019;574(7777):268-72.
319. Hart T, Tong AHY, Chan K, Van Leeuwen J, Seetharaman A, Aregger M, et al. Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. G3 (Bethesda). 2017;7(8):2719-27. doi: 10.1534/g3.117.041277. PubMed PMID: 28655737.
320. Ong SH, Li Y, Koike-Yusa H, Yusa K. Optimised metrics for CRISPR-KO screens with second-generation gRNA libraries. Scientific Reports. 2017;7(1):7384. doi: 10.1038/s41598-017-07827-z.
321. Munoz DM, Cassiani PJ, Li L, Billy E, Korn JM, Jones MD, et al. CRISPR Screens Provide a Comprehensive Assessment of Cancer Vulnerabilities but Generate False-Positive Hits for Highly Amplified Genomic Regions. Cancer Discov. 2016;6(8):900-13. Epub 2016/06/05. doi: 10.1158/2159-8290.Cd-16-0178. PubMed PMID: 27260157.
322. Daer RM, Cutts JP, Brafman DA, Haynes KA. The Impact of Chromatin Dynamics on Cas9-Mediated Genome Editing in Human Cells. ACS Synthetic Biology. 2017;6(3):428-38. doi: 10.1021/acssynbio.5b00299.
323. Myatt SS, Lam EWF. The emerging roles of forkhead box (Fox) proteins in cancer. Nature Reviews Cancer. 2007;7(11):847-59. doi: 10.1038/nrc2223.
324. Pan F, Yao J, Chen Y, Zhou C, Geng P, Mao H, et al. A novel long non-coding RNA FOXCUT and mRNA FOXC1 pair promote progression and predict poor prognosis in esophageal squamous cell carcinoma. Int J Clin Exp Pathol. 2014;7(6):2838.
325. Van Der Heul-Nieuwenhuijsen L, Dits NF, Jenster G. Gene expression of forkhead transcription factors in the normal and diseased human prostate. BJU international. 2009;103(11):1574-80.
326. Wang LY, Li LS, Yang Z. Correlation of FOXC1 protein with clinicopathological features in serous ovarian tumors. Oncology letters. 2016;11(2):933-8.
327. Xia L, Huang W, Tian D, Zhu H, Qi X, Chen Z, et al. Overexpression of forkhead box C1 promotes tumor metastasis and indicates poor prognosis in hepatocellular carcinoma. Hepatology. 2013;57(2):610-24.
328. Chen S, Jiao S, Jia Y, Li Y. Effects of targeted silencing of FOXC1 gene on proliferation and in vitro migration of human non-small-cell lung carcinoma cells. American Journal of Translational Research. 2016;8(8):3309.
329. Ni D, Ma X, Li H-Z, Gao Y, Li X-T, Zhang Y, et al. Downregulation of FOXO3a Promotes Tumor Metastasis and Is Associated with Metastasis-Free Survival of Patients with Clear Cell Renal Cell Carcinoma. Clinical Cancer Research. 2014;20(7):1779. doi: 10.1158/1078-0432.CCR-13-1687.
330. Lin A, Piao H-L, Zhuang L, Sarbassov DD, Ma L, Gan B. FoxO transcription factors promote AKT Ser473 phosphorylation and renal tumor growth in response to pharmacologic inhibition of the PI3K-AKT pathway. Cancer research. 2014;74(6):1682-93. Epub 2014/01/21. doi: 10.1158/0008-5472.CAN-13-1729. PubMed PMID: 24448243.
331. Cho EC, Kuo ML, Liu X, Yang L, Hsieh YC, Wang J, et al. Tumor suppressor FOXO3 regulates ribonucleotide reductase subunit RRM2B and impacts on survival of cancer patients. Oncotarget. 2014;5(13):4834-44. Epub 2014/06/21. doi: 10.18632/oncotarget.2044. PubMed PMID: 24947616; PubMed Central PMCID: PMCPMC4148103.
332. Gardner-Thorpe J, Ito H, Ashley SW, Whang EE. Differential display of expressed genes in pancreatic cancer cells. Biochem Biophys Res Commun. 2002;293(1):391-5. Epub 2002/06/11. doi: 10.1016/s0006-291x(02)00237-1. PubMed PMID: 12054612.
333. Artero-Castro A, Castellvi J, García A, Hernández J, Ramón y Cajal S, Lleonart ME. Expression of the ribosomal proteins Rplp0, Rplp1, and Rplp2 in gynecologic tumors. Hum Pathol. 2011;42(2):194-203. Epub 2010/11/03. doi: 10.1016/j.humpath.2010.04.020. PubMed PMID: 21040949.

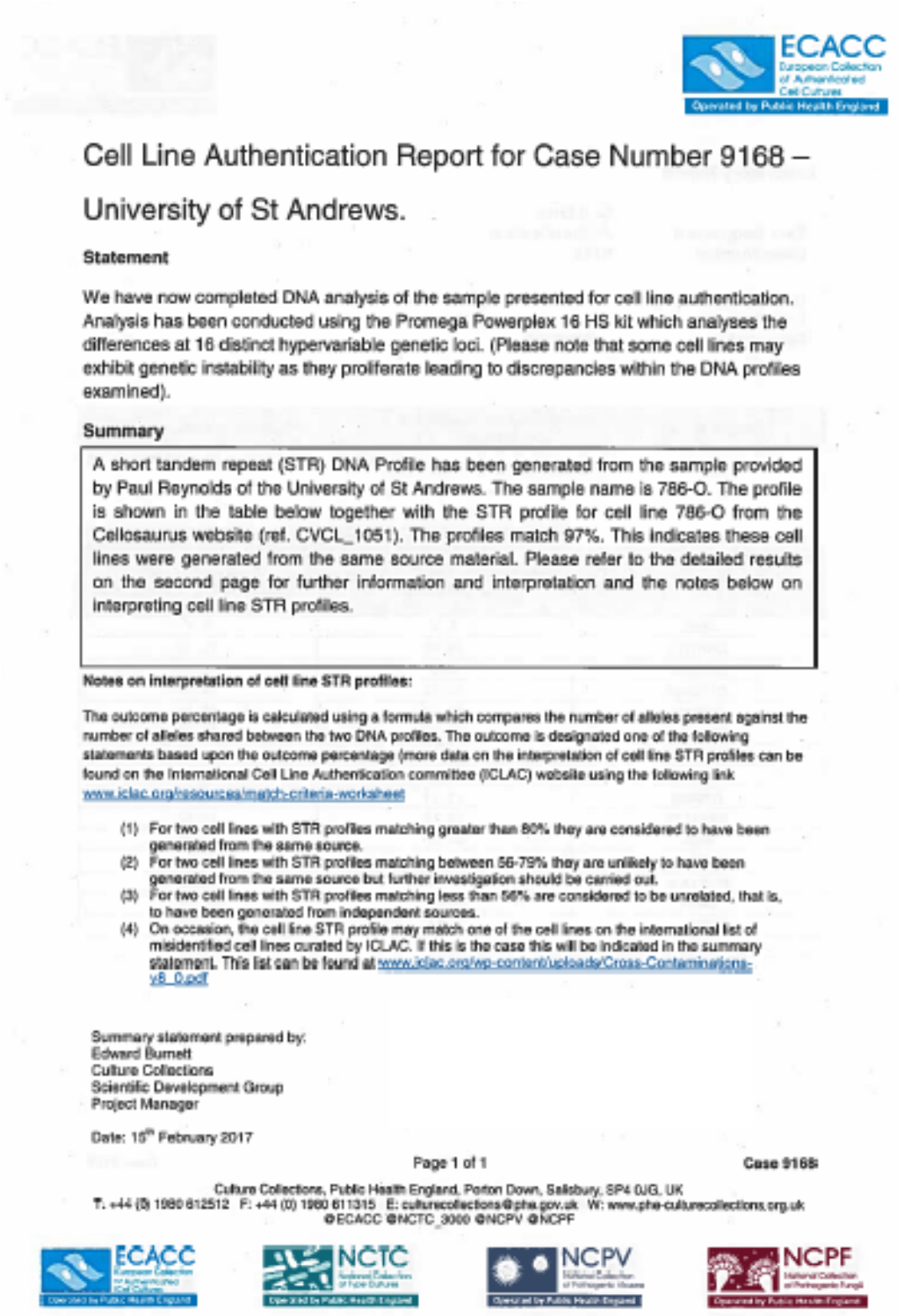
334. Yang M, Sun Y, Sun J, Wang Z, Zhou Y, Yao G, et al. Differentially expressed and survival-related proteins of lung adenocarcinoma with bone metastasis. Cancer medicine. 2018;7(4):1081-92. Epub 2018/03/10. doi: 10.1002/cam4.1363. PubMed PMID: 29522283; PubMed Central PMCID: PMCPMC5911611.
335. Aral B, Schlenzig J-S, Liu G, Kamoun P. Database cloning human delta 1-pyrroline-5-carboxylate synthetase (P5CS) cDNA: a bifunctional enzyme catalyzing the first 2 steps in proline biosynthesis. Comptes rendus de l'Academie des sciences Serie III, Sciences de la vie. 1996;319(3):171.
336. Tang L, Zeng J, Geng P, Fang C, Wang Y, Sun M, et al. Global Metabolic Profiling Identifies a Pivotal Role of Proline and Hydroxyproline Metabolism in Supporting Hypoxic Response in Hepatocellular Carcinoma. Clin Cancer Res. 2018;24(2):474-85. Epub 2017/11/01. doi: 10.1158/1078-0432.Ccr-17-1707. PubMed PMID: 29084919.
337. Grinde MT, Hilmarsdottir B, Tunset HM, Henriksen IM, Kim J, Haugen MH, et al. Glutamine to proline conversion is associated with response to glutaminase inhibition in breast cancer. Breast Cancer Res. 2019;21(1):61-. doi: 10.1186/s13058-019-1141-0. PubMed PMID: 31088535.
338. Ding Z, Ericksen RE, Escande-Beillard N, Lee QY, Loh A, Denil S, et al. Metabolic pathway analyses identify proline biosynthesis pathway as a promoter of liver tumorigenesis. Journal of hepatology. 2020;72(4):725-35. Epub 2019/11/15. doi: 10.1016/j.jhep.2019.10.026. PubMed PMID: 31726117.
339. Katoh M, Katoh M. Identification and characterization of human DIAPH3 gene in silico. International journal of molecular medicine. 2004;13(3):473-8.
340. Reis-Sobreiro M, Chen JF, Novitskaya T, You S, Morley S, Steadman K, et al. Emerin Deregulation Links Nuclear Shape Instability to Metastatic Potential. Cancer Res. 2018;78(21):6086-97. Epub 2018/08/30. doi: 10.1158/0008-5472.Can-18-0608. PubMed PMID: 30154147.
341. Dong L, Li Z, Xue L, Li G, Zhang C, Cai Z, et al. DIAPH3 promoted the growth, migration and metastasis of hepatocellular carcinoma cells by activating beta-catenin/TCF signaling. Molecular and Cellular Biochemistry. 2018;438(1-2):183-90.
342. Hager MH, Morley S, Bielenberg DR, Gao S, Morello M, Holcomb IN, et al. DIAPH3 governs the cellular transition to the amoeboid tumour phenotype. EMBO molecular medicine. 2012;4(8):743-60.
343. Beksac AT, Paulucci DJ, Blum KA, Yadav SS, Sfakianos JP, Badani KK. Heterogeneity in renal cell carcinoma. Urologic Oncology: Seminars and Original Investigations. 2017;35(8):507-15. doi: https://doi.org/10.1016/j.urolonc.2017.05.006.
344. Shanmugasundaram K, Block K. Renal Carcinogenesis, Tumor Heterogeneity, and Reactive Oxygen Species: Tactics Evolved. Antioxidants & Redox Signaling. 2016;25(12):685-701. doi: 10.1089/ars.2015.6569. PubMed PMID: PMC5069729.
345. Tannir N, Motzer R, Agarwal N, Liu P-Y, Whiting S, O'Keeffe B, et al. CANTATA: A randomized phase 2 study of CB-839 in combination with cabozantinib vs. placebo with cabozantinib in patients with advanced/metastatic renal cell carcinoma. Journal of Clinical Oncology. 2018;36(15_suppl):TPS4601-TPS. doi: 10.1200/JCO.2018.36.15_suppl.TPS4601.
346. Gross M, Budczies J, Demo S, Julie Janes J, E L, Parlati F, et al. Antitumor Activity of the Glutaminase Inhibitor, CB-839, in Triple-Negative Breast Cancer. 2013.
347. Lukey MJ, Katt WP, Cerione RA. Targeting Therapy Resistance: When Glutamine Catabolism Becomes Essential. Cancer cell. 2018;33(5):795-7. doi: https://doi.org/10.1016/j.ccell.2018.04.009.
348. MacKinnon AL, Bennett M, Gross M, Janes J, Li W, Rodriquez M, et al. Metabolomic, Proteomic and Genomic Profiling Identifies Biomarakers of Sensitivity to Glutaminase Inhibitor

CB-839 in Multiple Myeloma. Blood. 2015;126(23):1802-. doi: 10.1182/blood.V126.23.1802.1802.
349. Reis LMD, Adamoski D, Ornitz Oliveira Souza R, Rodrigues Ascenção CF, Sousa de Oliveira KR, Corrêa-da-Silva F, et al. Dual inhibition of glutaminase and carnitine palmitoyltransferase decreases growth and migration of glutaminase inhibition-resistant triple-negative breast cancer cells. J Biol Chem. 2019;294(24):9342-57. Epub 2019/05/02. doi: 10.1074/jbc.RA119.008180. PubMed PMID: 31040181; PubMed Central PMCID: PMCPMC6579458.
350. Singleton DC, Dechaume A-L, Murray PM, Katt WP, Baguley BC, Leung EY. Pyruvate anaplerosis is a mechanism of resistance to pharmacological glutaminase inhibition in triple-receptor negative breast cancer. BMC Cancer. 2020;20(1):470. doi: 10.1186/s12885-020-06885-3.
351. Agrotis A, Ketteler R. A new age in functional genomics using CRISPR/Cas9 in arrayed library screening. Frontiers in Genetics. 2015;6(300). doi: 10.3389/fgene.2015.00300.
352. Miles LA, Garippa RJ, Poirier JT. Design, execution, and analysis of pooled in vitro CRISPR/Cas9 screens. FEBS J. 2016;283(17):3170-80. Epub 2016/06/03. doi: 10.1111/febs.13770. PubMed PMID: 27250066.
353. Sarr A, Bré J, Um IH, Chan TH, Mullen P, Harrison DJ, et al. Genome-scale CRISPR/Cas9 screen determines factors modulating sensitivity to ProTide NUC-1031. Scientific reports. 2019;9(1):7643-. doi: 10.1038/s41598-019-44089-3. PubMed PMID: 31113993.
354. Henkel L, Rauscher B, Schmitt B, Winter J, Boutros M. Genome-scale CRISPR screening at high sensitivity with an empirically designed sgRNA library. BMC Biology. 2020;18(1):174. doi: 10.1186/s12915-020-00905-1.
355. Sanson KR, Hanna RE, Hegde M, Donovan KF, Strand C, Sullender ME, et al. Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. Nature Communications. 2018;9(1):5416. doi: 10.1038/s41467-018-07901-8.
356. Tzelepis K, Koike-Yusa H, De Braekeleer E, Li Y, Metzakopian E, Dovey OM, et al. A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. Cell Rep. 2016;17(4):1193-205. Epub 2016/10/21. doi: 10.1016/j.celrep.2016.09.079. PubMed PMID: 27760321; PubMed Central PMCID: PMCPMC5081405.
357. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. Nature Biotechnology. 2014;32(12):1262-7. doi: 10.1038/nbt.3026.
358. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nature Biotechnology. 2016;34(2):184-91. doi: 10.1038/nbt.3437.
359. Wang D, Zhang C, Wang B, Li B, Wang Q, Liu D, et al. Optimized CRISPR guide RNA design for two high-fidelity Cas9 variants by deep learning. Nature Communications. 2019;10(1):4284. doi: 10.1038/s41467-019-12281-8.
360. Isaac RS, Jiang F, Doudna JA, Lim WA, Narlikar GJ, Almeida R. Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. eLife. 2016;5:e13450. doi: 10.7554/eLife.13450.
361. Horlbeck MA, Gilbert LA, Villalta JE, Adamson B, Pak RA, Chen Y, et al. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. eLife. 2016;5:e19760. doi: 10.7554/eLife.19760.
362. Ouyang Q, Liu Y, Tan J, Li J, Yang D, Zeng F, et al. Loss of ZNF587B and SULF1 contributed to cisplatin resistance in ovarian cancer cell lines based on Genome-scale CRISPR/Cas9 screening. Am J Cancer Res. 2019;9(5):988-98. PubMed PMID: 31218106.
363. Zhuang X, Veltri DP, Long EO. Genome-Wide CRISPR Screen Reveals Cancer Cell Resistance to NK Cells Induced by NK-Derived IFN-γ. Front Immunol. 2019;10:2879-. doi: 10.3389/fimmu.2019.02879. PubMed PMID: 31921143.

364. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458(7239):719-24. Epub 2009/04/11. doi: 10.1038/nature07943. PubMed PMID: 19360079; PubMed Central PMCID: PMCPMC2821689.
365. Carette JE, Guimaraes CP, Wuethrich I, Blomen VA, Varadarajan M, Sun C, et al. Global gene disruption in human cells to assign genes to phenotypes by deep sequencing. Nat Biotechnol. 2011;29(6):542-6. Epub 2011/05/31. doi: 10.1038/nbt.1857. PubMed PMID: 21623355; PubMed Central PMCID: PMCPMC3111863.
366. Yau EH, Rana TM. Next-Generation Sequencing of Genome-Wide CRISPR Screens. Methods Mol Biol. 2018;1712:203-16. doi: 10.1007/978-1-4939-7514-3_13. PubMed PMID: 29224076.
367. Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, et al. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. Cell. 2015;160(6):1246-60. Epub 2015/03/05. doi: 10.1016/j.cell.2015.02.038. PubMed PMID: 25748654.
368. Kalle E, Kubista M, Rensing C. Multi-template polymerase chain reaction. Biomolecular detection and quantification. 2014;2:11-29.
369. Strezoska Ž, Licon A, Haimes J, Spayd KJ, Patel KM, Sullivan K, et al. Optimized PCR conditions and increased shRNA fold representation improve reproducibility of pooled shRNA screens. PLoS One. 2012;7(8):e42341.
370. Sack LM, Davoli T, Xu Q, Li MZ, Elledge SJ. Sources of Error in Mammalian Genetic Screens. G3 (Bethesda). 2016;6(9):2781-90. doi: 10.1534/g3.116.030973. PubMed PMID: 27402361.
371. Segura MM, Mangion M, Gaillet B, Garnier A. New developments in lentiviral vector design, production and purification. Expert opinion on biological therapy. 2013;13(7):987-1011.
372. Sastry L, Xu Y, Cooper R, Pollok K, Cornetta K. Evaluation of plasmid DNA removal from lentiviral vectors by benzonase treatment. Human gene therapy. 2004;15(2):221-6.
373. Yang Z, Jiang S, Cheng Y, Li T, Hu W, Ma Z, et al. FOXC1 in cancer development and therapy: deciphering its emerging and divergent roles. Ther Adv Med Oncol. 2017;9(12):797-816. Epub 2017/11/23. doi: 10.1177/1758834017742576. PubMed PMID: 29449899.
374. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia (New York, NY). 2017;19(8):649-58. Epub 2017/07/22. doi: 10.1016/j.neo.2017.05.002. PubMed PMID: 28732212; PubMed Central PMCID: PMCPMC5516091.

Appendix

Cell Line Authentication Report of 786-0 cell line.



ECACC
European Collection of Authenticated Cell Cultures
Operated by Public Health England

Cell Line Authentication Report for Case Number 9168 – University of St Andrews.

Statement

We have now completed DNA analysis of the sample presented for cell line authentication. Analysis has been conducted using the Promega Powerplex 16 HS kit which analyses the differences at 16 distinct hypervariable genetic loci. (Please note that some cell lines may exhibit genetic instability as they proliferate leading to discrepancies within the DNA profiles examined).

Summary

A short tandem repeat (STR) DNA Profile has been generated from the sample provided by Paul Reynolds of the University of St Andrews. The sample name is 786-O. The profile is shown in the table below together with the STR profile for cell line 786-O from the Cellosaurus website (ref. CVCL_1051). The profiles match 97%. This indicates these cell lines were generated from the same source material. Please refer to the detailed results on the second page for further information and interpretation and the notes below on interpreting cell line STR profiles.

Notes on interpretation of cell line STR profiles:

The outcome percentage is calculated using a formula which compares the number of alleles present against the number of alleles shared between the two DNA profiles. The outcome is designated one of the following statements based upon the outcome percentage (more data on the interpretation of cell line STR profiles can be found on the International Cell Line Authentication committee (ICLAC) website using the following link www.iclac.org/resources/match-criteria-worksheet

(1) For two cell lines with STR profiles matching greater than 80% they are considered to have been generated from the same source.
(2) For two cell lines with STR profiles matching between 56-79% they are unlikely to have been generated from the same source but further investigation should be carried out.
(3) For two cell lines with STR profiles matching less than 56% are considered to be unrelated, that is, to have been generated from independent sources.
(4) On occasion, the cell line STR profile may match one of the cell lines on the international list of misidentified cell lines curated by ICLAC. If this is the case this will be indicated in the summary statement. This list can be found at www.iclac.org/wp-content/uploads/Cross-Contaminations-v8_0.pdf

Summary statement prepared by:
Edward Burnett
Culture Collections
Scientific Development Group
Project Manager

Date: 15th February 2017

Page 1 of 1 Case 9168

Culture Collections, Public Health England, Porton Down, Salisbury, SP4 0JG, UK
T. +44 (0) 1980 612512 F: +44 (0) 1980 611315 E: culturecollections@phe.gov.uk W: www.phe-culturecollections.org.uk
@ECACC @NCTC_3000 @NCPV @NCPF

ECACC NCTC NCPV NCPF

ECACC
European Collection of Authenticated Cell Cultures
Operated by Public Health England

Laboratory Report

Test Requested	**Cell Line Authentication**
Case Number	**9168**
Date Sample Tested	**13/2/2018**
Date Sample Reported	**15/2/2018**

Sample Name	Sample/Comparison Profile Source	Sample Reference
786-O	Cellosaurus 786-O CVCL_1051	SC_01_01014790_9168_LW6000056-18

Table of Allelic Data	Genotypes	
STR Locus	**Test Sample – 786-O – University of St Andrews**	**Database Sample – Cellosaurus 786-O CVCL_1051**
AMEL	X, Y	X, Y
CSF1PO	10, 10	10, 10
D13S317	8, 8	8, 8
D16S539	12, 12	12, 12
D18S51	13, 14	13, 14
D21S11	29, 30	29, 30
D3S1358	16, 16	16, 16
D5S818	9, 9	9, 9
D7S820	11, 11	11, 12
D8S1179	13, 13	13, 13
FGA	24, 24	24, 24
PENTA D	9, 12	9, 12
PENTA E	7, 16	7, 16
THO1	6, 9.3	6, 9.3
TPOX	8, 11	8, 11
vWA	15, 17	15, 17

Culture Collections, Public Health England, Porton Down, Salisbury, SP4 0JG, UK
T: +44 (0) 1980 612512 F: +44 (0) 1980 611315 E: culturecollections@phe.gov.uk W: www.phe-culturecollections.org.uk
@ECACC @NCTC_3000 @NCPV @NCPF

ECACC Operated by Public Health England · NCTC Operated by Public Health England · NCPV Operated by Public Health England · NCPF Operated by Public Health England

Supplementary Table – Illumina primers for PCR2 of GeCKOv2 library

Stagger

Barcode

Primer ID	Sequence
Illumina F1	5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG CTCTTCCGATCTT**AAGTAGAG**TCTTGTGGAAAGGACGAAACACCG-3'
Illumina F2	5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG CTCTTCCGATCTAT**ACACGATC**TCTTGTGGAAAGGACGAAACACCG-3'
Illumina F3	5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG CTCTTCCGATCTGAT**CGCGCGGT**TCTTGTGGAAAGGACGAAACACCG-3’
Illumina F4	5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG CTCTTCCGATCTCGAT**CATGATCG**TCTTGTGGAAAGGACGAAACACCG-3'
Illumina R1	5'-CAAGCAGAAGACGGCATACGAGAT**AAGTAGAG**GTGACTGGAGTTCA GACGTGTGCTCTTCCGATCTTTCTACTATTCTTTCCCCTGCACTGT-3'
Illumina R2	5'-CAAGCAGAAGACGGCATACGAGAT**ACACGATC**GTGACTGGAGTTCA GACGTGTGCTCTTCCGATCTATTCTACTATTCTTTCCCCTGCACTGT-3’
Illumina R3	5'-CAAGCAGAAGACGGCATACGAGAT**CGCGCGGTGTG**ACTGGAGTTCA GACGTGTGCTCTTCCGATCTGATTCTACTATTCTTTCCCCTGCACTGT- 3
Illumina R4	5'-CAAGCAGAAGACGGCATACGAGAT**CATGATCG**GTGACTGGAGTTCA GACGTGTGCTCTTCCGATCTCGATTCTACTATTCTTTCCCCTGCACTGT-3'

Supplementary Table - QC data generated by MAGeCK.

Label	Reads	Mapped	Percentage	TotalsgRNAs	Zerocounts	GiniIndex
Baseline_1	82745487	82745487	1	123411	4	0.07909
Baseline_2	64086987	64086987	1	123411	7	0.07872
CB.839_d16_1	69297349	69297349	1	123411	40	0.1006
CB.839_d16_2	55031582	55031582	1	123411	18	0.101
CB.839_d21_1	73080673	73080673	1	123411	23	0.0967
CB.839_d21_2	97179980	97179980	1	123411	18	0.1104
DMSO_d16_1	55794332	55794332	1	123411	9	0.0957
DMSO_d16_2	81391945	81391945	1	123411	14	0.09296
DMSO_d21_1	83053588	83053588	1	123411	13	0.09165
DMSO_d21_2	76759635	76759635	1	123411	10	0.09866

Supplementary Table – Excerpt from 'all.count_normalized.txt' file generated by MAGeCK for all sgRNAs targeting RPLP2, FOXC1, FOXO3 and ALDH18A1 genes.

sgRNA	Gene	Baseline_1	Baseline_2	CB.839_d16_1	CB.839_d16_2	CB.839_d21_1	CB.839_d21_2	DMSO_d16_1	DMSO_d16_2	DMSO_d21_1	DMSO_d21_2
HGLibA_49301	RPLP2	116.45542	105.868183	55.40694305	42.32905156	49.93485295	13.56064781	18.22013999	74.69970757	104.1872387	102.8616846
HGLibA_49302	RPLP2	324.91558	212.919396	183.2673633	160.8645702	216.1262085	322.2023168	96.59379945	173.8125816	89.77107439	155.0914015
HGLibA_49303	RPLP2	797.88192	968.568283	738.1519161	729.3943204	826.3400197	754.1674372	470.5293987	686.4983136	552.7225126	566.2858799
HGLibB_41960	RPLP2	1733.7324	3358.34029	25531.84211	20602.08501	249425.5945	45457.00652	1773.913334	3710.190718	1984.696167	2751.498044
HGLibB_41961	RPLP2	638.45145	608.703977	923.342687	1071.919665	822.0554956	743.0409601	1133.193801	1014.209555	1307.518337	920.6899358
HGLibB_41962	RPLP2	90.204805	92.317934	352.483287	163.7125555	65.41366749	131.4115998	52.3697755	332.802292	94.26256469	355.473041
HGLibA_17821	FOXC1	1260.0727	965.101399	2715.444623	2343.911886	1236.405022	1702.826362	1111.575966	1821.222424	1366.881824	2201.602474
HGLibA_17822	FOXC1	440.19037	319.648246	559.7535688	470.3645994	441.6204435	715.429373	397.2674391	427.6977062	356.0847562	663.5237268
HGLibA_17823	FOXC1	141.44387	158.922609	1674.564786	974.2862294	770.6058225	1232.259001	232.1611552	915.6558346	515.1987074	1142.345728
HGLibB_17798	FOXC1	457.65535	605.859439	563.3151957	692.1775725	787.372619	541.9899489	600.7632992	647.8433801	792.3996396	429.5780593
HGLibB_17799	FOXC1	328.39491	239.678921	325.4232751	352.8723877	385.4957958	300.4894771	187.705061	382.5147014	462.0111571	447.8097407
HGLibB_17800	FOXC1	320.65188	206.377286	221.2403768	241.9766285	73.57268705	80.77460395	147.9255219	218.3945062	244.2199986	177.0904617
HGLibB_17900	FOXO3	136.99386	231.53514	306.6200705	254.5625136	286.7188608	296.7736806	137.070574	241.0433467	244.3641016	236.910175
HGLibB_17901	FOXO3	849.60598	964.644696	1346.062042	1011.999143	1623.0285	1764.635292	835.1470634	1189.008387	968.8296983	1552.596138
HGLibB_17902	FOXO3	346.46103	243.034246	590.7203105	325.5057631	709.9113482	988.8206513	107.1023536	467.4178916	108.3262125	617.2294351
HGLibA_17923	FOXO3	913.21084	690.485968	1213.577827	1058.408822	908.1833108	972.3692877	639.3997679	1145.31713	725.8215108	1249.131112
HGLibA_17924	FOXO3	434.92448	357.731898	492.5893451	566.0956587	801.0097962	400.9226015	287.9678863	495.445565	427.8326641	400.0991687
HGLibA_17925	FOXO3	147.13847	195.473593	417.0935747	283.8815196	348.6165823	473.3978008	265.6221085	329.5036192	166.2277749	370.460461
HGLibB_01535	ALDH18A1	168.26281	204.744259	186.1635471	187.9476904	381.3528246	139.8893716	175.8827529	213.6496887	109.6781458	144.2715973
HGLibB_01536	ALDH18A1	166.15558	166.670088	285.8342171	130.905107	352.9327047	254.683785	63.5743341	270.3927393	37.26263872	399.5571826
HGLibB_01537	ALDH18A1	338.03359	316.219805	782.9513477	532.8526553	711.2992322	778.83776	459.366583	801.2756154	318.3053863	900.604187
HGLibA_01537	ALDH18A1	250.3282	374.816191	241.1667387	232.5754114	267.3511662	277.9053832	206.8857782	216.486556	339.6401073	269.1230944
HGLibA_01538	ALDH18A1	573.69013	516.645026	1292.368086	1049.156282	1356.007231	1707.519588	575.3818573	1041.879195	620.1926052	1280.836227
HGLibA_01539	ALDH18A1	539.32421	480.746265	1299.790375	953.7352415	1264.15924	1431.904208	377.579016	1040.843821	384.0996744	1041.671042

Supplementary Table - CB-839 vs DMSO D16 data generated by MAGeCK. Genes are organised according to decreasing beta-scores in CB-839 samples on day 16.

Gene	sgRNA	DMSO|beta	DMSO|z	DMSO|p-value	DMSO|fdr	DMSO|wald-p-value	DMSO|wald-fdr	CB-839|beta	CB-839|z	CB-839|p-value	CB-839|fdr	CB-839|wald-p-value	CB-839|wald-fdr
hsa-mir-4679-2	1	0.61208	0.61638	0.035546	0.36182	0.53765	0.85337	0.89807	0.90753	0.014784	0.28052	0.36413	0.66168
NonTargetingControlGuideForHuman_0575	2	0.56681	1.0217	0.0071184	0.18182	0.30691	0.73151	0.88313	1.592	0.00036505	0.039823	0.11138	0.38392
NonTargetingControlGuideForHuman_0183	2	0.28265	0.62666	0.15054	0.62457	0.53088	0.85131	0.81996	1.8182	0.0011408	0.074627	0.069038	0.31092
hsa-mir-4679-1	1	0.41357	0.48666	0.14976	0.62457	0.6265	0.89004	0.79853	0.94324	0.022815	0.34339	0.34556	0.64778
NonTargetingControlGuideForHuman_0426	2	0.64288	1.1896	0.0031029	0.11972	0.23422	0.67861	0.77694	1.4381	0.0018709	0.096471	0.1504	0.44064
NonTargetingControlGuideForHuman_0418	2	0.54377	1.2289	0.0088068	0.19815	0.21911	0.66515	0.72272	1.6341	0.0039699	0.15876	0.10225	0.36957
NonTargetingControlGuideForHuman_0250	2	0.47083	1.1871	0.020945	0.3	0.2352	0.67919	0.70094	1.7673	0.0048825	0.17657	0.077179	0.32553
hsa-mir-5582	2	0.84878	1.3269	0.00022815	0.023585	0.18456	0.63128	0.69943	1.0922	0.0049281	0.17734	0.27475	0.58195
NonTargetingControlGuideForHuman_0598	2	0.55613	1.2128	0.0078941	0.18893	0.22521	0.67105	0.68698	1.4985	0.005932	0.19006	0.13399	0.41688
NonTargetingControlGuideForHuman_0205	2	0.68259	1.4434	0.0016883	0.083333	0.14891	0.5995	0.68364	1.4461	0.0060233	0.19242	0.14815	0.43698
NonTargetingControlGuideForHuman_0531	2	0.37828	1.0221	0.058362	0.44225	0.30674	0.73151	0.68046	1.8388	0.0060689	0.19303	0.065946	0.30449
NonTargetingControlGuideForHuman_0895	2	0.48089	0.90297	0.018389	0.28471	0.36654	0.77093	0.66121	1.2434	0.0073466	0.2124	0.21374	0.51837
NonTargetingControlGuideForHuman_0374	2	0.23737	0.68061	0.22651	0.69524	0.49612	0.83775	0.6498	1.8633	0.0087155	0.22248	0.062422	0.29674
NonTargetingControlGuideForHuman_0225	2	0.19408	0.62175	0.32288	0.77081	0.5341	0.85238	0.64599	2.0697	0.0090805	0.22665	0.038476	0.23975
NonTargetingControlGuideForHuman_0244	2	0.1771	0.56334	0.36605	0.79654	0.5732	0.86873	0.64597	2.0549	0.0090805	0.22665	0.039891	0.24298
NonTargetingControlGuideForHuman_0217	2	0.56251	1.3505	0.0073922	0.18664	0.17687	0.62486	0.64494	1.5487	0.0091262	0.22671	0.12146	0.39889
NonTargetingControlGuideForHuman_0965	2	0.67837	1.4982	0.0019165	0.090517	0.13409	0.58044	0.64094	1.416	0.0098563	0.23529	0.15679	0.44827
NonTargetingControlGuideForHuman_0227	2	0.24808	0.77784	0.20566	0.67693	0.43667	0.80852	0.63379	1.988	0.010586	0.24192	0.046813	0.25992
NonTargetingControlGuideForHuman_0954	2	0.58587	1.4876	0.0059776	0.16522	0.13686	0.5828	0.62264	1.5811	0.011636	0.24976	0.11385	0.38683
TSC2	6	0.35569	1.7171	0.002601	0.10899	0.085961	0.50988	0.61895	2.9891	0	0	0.002798	0.079858
NonTargetingControlGuideForHuman_0342	2	0.26177	0.70755	0.18166	0.65406	0.47923	0.8286	0.61636	1.6667	0.012183	0.2572	0.095566	0.35826
NonTargetingControlGuideForHuman_0044	2	0.10193	0.27964	0.61045	0.90362	0.77975	0.94037	0.60834	1.6692	0.013233	0.26927	0.095077	0.3577
NonTargetingControlGuideForHuman_0440	2	0.55027	1.4698	0.0082592	0.19421	0.14162	0.58942	0.59781	1.5969	0.015286	0.28414	0.11029	0.38239
NonTargetingControlGuideForHuman_0616	2	0.30994	0.97793	0.1175	0.581	0.32811	0.74151	0.5935	1.8731	0.016062	0.29139	0.061057	0.29379
FOXC1	6	0.18852	0.81583	0.099019	0.54736	0.4146	0.79799	0.57076	2.4734	9.13E-05	0.018349	0.013384	0.15543
NonTargetingControlGuideForHuman_0305	2	0.14208	0.39988	0.46945	0.84669	0.68925	0.91495	0.56722	1.5974	0.021492	0.33381	0.11017	0.38224
NonTargetingControlGuideForHuman_0551	2	0.32703	1.141	0.098198	0.54548	0.25386	0.69379	0.56347	1.966	0.022085	0.33888	0.049294	0.26517
NonTargetingControlGuideForHuman_0783	2	0.38083	1.1746	0.0564	0.43708	0.24015	0.68361	0.56085	1.7321	0.022724	0.34274	0.083258	0.33658
NonTargetingControlGuideForHuman_0888	2	0.57102	1.5784	0.0067534	0.17412	0.11447	0.5549	0.55978	1.5477	0.023089	0.3449	0.12169	0.39941
NonTargetingControlGuideForHuman_0419	2	0.59783	1.3981	0.0052019	0.1551	0.16208	0.61132	0.55765	1.3064	0.023546	0.34747	0.19141	0.49199
hsa-mir-1249	4	0.26282	0.93019	0.062697	0.45633	0.35227	0.75902	0.55456	1.9662	0.0018709	0.096471	0.049278	0.26517
NonTargetingControlGuideForHuman_0190	2	0.27075	0.96977	0.16797	0.64095	0.33216	0.7448	0.55324	1.9819	0.024686	0.35616	0.04749	0.26162
RPLP2	6	-0.074395	-0.23496	0.56057	0.88515	0.81424	0.9513	0.55075	1.7585	0.00013689	0.02521	0.078654	0.32808
NonTargetingControlGuideForHuman_0144	2	0.32658	1.196	0.098654	0.54663	0.2317	0.67667	0.5467	2.0023	0.026694	0.36816	0.045257	0.25633
hsa-mir-4264	4	0.48803	1.2785	0.00091262	0.057307	0.20108	0.64879	0.54526	1.4267	0.0021446	0.10609	0.15366	0.44444
hsa-mir-502	4	0.27207	1.1157	0.053525	0.42936	0.26455	0.70451	0.53526	2.1958	0.0026466	0.12159	0.028104	0.21176
HIST1H2BE	6	0.0071638	0.031849	0.92485	0.98326	0.97459	0.9936	0.5351	2.3993	0.00018252	0.030075	0.016427	0.16877
NonTargetingControlGuideForHuman_0006	2	0.24236	0.84439	0.21766	0.68623	0.39845	0.78873	0.52706	1.8365	0.032991	0.40963	0.066279	0.3051
NonTargetingControlGuideForHuman_0048	2	0.37132	1.2919	0.06224	0.45573	0.1964	0.64328	0.52645	1.8318	0.033311	0.41054	0.066988	0.30625
GLTSCR1	6	0.25588	0.82907	0.025827	0.32289	0.40707	0.79349	0.52543	1.7078	0.00022815	0.034483	0.087673	0.34585
NonTargetingControlGuideForHuman_0528	2	0.4544	1.429	0.025873	0.32289	0.153	0.60586	0.52278	1.6443	0.034725	0.41975	0.10011	0.36575
NonTargetingControlGuideForHuman_0856	2	0.62678	1.6602	0.0037417	0.13099	0.096884	0.531	0.52128	1.3808	0.03509	0.42045	0.16733	0.46096
NonTargetingControlGuideForHuman_0918	2	0.25459	0.95974	0.19366	0.6623	0.33718	0.74823	0.51609	1.9458	0.037372	0.42769	0.051674	0.27159
NonTargetingControlGuideForHuman_0719	2	0.24311	0.99162	0.21597	0.68445	0.32138	0.73785	0.5156	2.1031	0.037509	0.42769	0.035454	0.23149
NonTargetingControlGuideForHuman_0948	2	0.46686	1.4631	0.022131	0.30599	0.14344	0.59209	0.51496	1.6142	0.037737	0.42769	0.10648	0.37607
NonTargetingControlGuideForHuman_0157	2	0.44404	1.5087	0.029249	0.33879	0.13137	0.57538	0.51366	1.7455	0.038239	0.42974	0.080893	0.33217
NonTargetingControlGuideForHuman_0337	2	0.32234	1.2152	0.10358	0.55466	0.22428	0.67074	0.5108	1.9258	0.039151	0.43412	0.054131	0.27697
SRD5A1	6	0.36706	1.5517	0.0016427	0.081633	0.12072	0.56269	0.50957	2.1563	0.00022815	0.034483	0.031058	0.21948
NPL	6	0.37469	1.8366	0.0014146	0.073986	0.066272	0.47278	0.50941	2.4983	0.00022815	0.034483	0.01248	0.15075

Supplementary Table - CB-839 vs DMSO D21 data generated by MAGeCK. Genes are organised according to decreasing beta-scores in CB-839 samples on day 21.

Gene	sgRNA	DMSO|beta	DMSO|z	DMSO|p-value	DMSO|fdr	DMSO|wald-p-value	DMSO|wald-fdr	CB-839|beta	CB-839|z	CB-839|p-value	CB-839|fdr	CB-839|wald-p-value	CB-839|wald-fdr
NonTargetingControlGuideForHuman_0431	2	0.28086	0.17265	0.20817	0.73379	0.86293	0.97474	1.5889	0.97674	0	0	0.3287	0.63371
hsa-mir-4679-2	1	1.3576	0.63951	0	0	0.52249	0.89033	1.4402	0.67796	0.0010039	0.10233	0.4978	0.76161
hsa-mir-4679-1	1	1.2892	0.66835	0	0	0.50391	0.88301	1.3552	0.70215	0.002966	0.16582	0.48259	0.75263
ZNF540	6	0.31154	0.5329	0.014967	0.27586	0.5941	0.91369	0.99354	1.7042	0	0	0.088337	0.35574
NonTargetingControlGuideForHuman_0048	2	0.40917	0.72037	0.068401	0.5194	0.4713	0.873	0.93644	1.6486	0.0019165	0.13505	0.09922	0.37536
NonTargetingControlGuideForHuman_0418	2	0.44325	0.74904	0.049966	0.45946	0.45383	0.86511	0.89488	1.5121	0.0028291	0.16402	0.1305	0.42301
hsa-mir-5582	2	0.61241	0.94857	0.0079854	0.2042	0.34284	0.81717	0.86922	1.3479	0.0034223	0.1784	0.17768	0.4841
NonTargetingControlGuideForHuman_0365	2	0.37668	0.84833	0.092585	0.58574	0.39625	0.84418	0.80426	1.8112	0.0065252	0.24074	0.070115	0.32315
GZMB	6	0.035207	0.074187	0.80844	0.96868	0.94086	0.99106	0.79629	1.6968	0	0	0.089729	0.35774
NonTargetingControlGuideForHuman_0575	2	0.67285	1.2783	0.0036961	0.14161	0.20113	0.72997	0.77431	1.4711	0.0087155	0.27721	0.14127	0.43689
NonTargetingControlGuideForHuman_0947	2	0.34228	0.71828	0.1258	0.64221	0.47259	0.87341	0.76009	1.595	0.010084	0.29865	0.1107	0.39428
NonTargetingControlGuideForHuman_0217	2	0.86507	1.4734	0.00054757	0.054795	0.14064	0.67287	0.7389	1.2584	0.01191	0.32266	0.20823	0.51624
hsa-mir-449c	4	-0.19206	-0.27569	0.24472	0.76142	0.78279	0.95821	0.73561	1.068	0.00036505	0.053333	0.2855	0.59438
NonTargetingControlGuideForHuman_0607	2	0.030241	0.082299	0.91859	0.98982	0.93441	0.98911	0.719	1.9566	0.014419	0.3495	0.050396	0.28297
NonTargetingControlGuideForHuman_0842	2	0.26442	0.72594	0.23874	0.75729	0.46788	0.87144	0.71822	1.9718	0.014556	0.34973	0.048636	0.2788
RPLP2	6	0.051936	0.087894	0.71015	0.95362	0.92996	0.98822	0.71792	1.2279	0	0	0.21947	0.52769
NonTargetingControlGuideForHuman_0598	2	0.64447	1.2082	0.0056126	0.17107	0.22697	0.75127	0.71176	1.3342	0.015241	0.35532	0.18213	0.49052
hsa-mir-4280	4	0.12305	0.34782	0.45284	0.87419	0.72797	0.94783	0.70489	2.0016	0.00068446	0.083333	0.045327	0.26951
HIST1H2BE	6	-0.079532	-0.14353	0.53689	0.90837	0.88587	0.97889	0.70407	1.291	0	0	0.19669	0.50603
NonTargetingControlGuideForHuman_0576	2	0.27415	0.73849	0.2198	0.74065	0.46022	0.86753	0.69372	1.8682	0.017066	0.37166	0.061729	0.30625
NonTargetingControlGuideForHuman_0144	2	0.35644	0.94519	0.11075	0.61803	0.34456	0.818	0.68562	1.8181	0.018298	0.37795	0.069049	0.32086
NonTargetingControlGuideForHuman_0426	2	0.43384	1.1245	0.054346	0.47745	0.26081	0.77078	0.68157	1.7664	0.0188	0.38383	0.07733	0.33631
NonTargetingControlGuideForHuman_0626	2	0.23143	0.69707	0.30431	0.80272	0.48576	0.87887	0.67869	2.0442	0.019119	0.38721	0.040929	0.25925
TSC2	6	0.49791	1.9359	0.00022815	0.032051	0.052884	0.53324	0.67844	2.6384	0	0	0.0083302	0.12858
hsa-mir-1249	4	0.32021	0.95595	0.044809	0.44115	0.3391	0.81466	0.67788	2.0269	0.001232	0.10976	0.042677	0.26295
NonTargetingControlGuideForHuman_0246	2	0.59504	1.4083	0.0098106	0.22349	0.15904	0.69238	0.67623	1.6005	0.019758	0.38859	0.1095	0.39274
NonTargetingControlGuideForHuman_0244	2	0.40503	1.1402	0.070773	0.52809	0.25418	0.76631	0.65816	1.8529	0.023044	0.40111	0.063903	0.3102
hsa-mir-483	4	-0.29552	-0.4399	0.087201	0.57284	0.66001	0.93126	0.65195	0.98408	0.0015971	0.12635	0.32508	0.63079
NonTargetingControlGuideForHuman_0743	2	0.41506	1.0674	0.064248	0.5116	0.28579	0.78477	0.64883	1.6686	0.024686	0.41172	0.095195	0.36722
NonTargetingControlGuideForHuman_0306	2	0.1409	0.45767	0.54328	0.91126	0.64719	0.92759	0.63795	2.0722	0.026785	0.42079	0.038244	0.25189
NonTargetingControlGuideForHuman_0250	2	0.45372	1.2765	0.04449	0.43959	0.20178	0.731	0.63484	1.786	0.027287	0.42302	0.074097	0.33052
NonTargetingControlGuideForHuman_0824	2	0.14366	0.47479	0.53393	0.90804	0.63494	0.92309	0.62879	2.0781	0.028611	0.43271	0.037701	0.25121
ALDH18A1	6	0.25371	0.9768	0.048232	0.45396	0.32867	0.80951	0.62691	2.4176	4.56E-05	0.023256	0.015625	0.16898
NonTargetingControlGuideForHuman_0783	2	0.40523	0.97158	0.070545	0.52693	0.33126	0.81125	0.61915	1.484	0.03071	0.44718	0.13782	0.43302
NonTargetingControlGuideForHuman_0073	2	-0.018736	-0.058887	0.90573	0.98755	0.95304	0.99336	0.61827	1.9432	0.031075	0.44862	0.051992	0.28628
FAM162A	6	0.27132	1.1912	0.034679	0.39574	0.23358	0.75376	0.61264	2.6904	9.13E-05	0.023256	0.0071369	0.12149
NonTargetingControlGuideForHuman_0954	2	0.67551	1.5584	0.0035592	0.13879	0.11914	0.65102	0.61078	1.409	0.032945	0.45725	0.15883	0.46075
NonTargetingControlGuideForHuman_0965	2	0.69153	1.4366	0.0031029	0.1323	0.15082	0.68408	0.61061	1.2684	0.032945	0.45725	0.20467	0.51183
NonTargetingControlGuideForHuman_0600	2	0.057364	0.19881	0.82359	0.9699	0.84241	0.96811	0.60345	2.0913	0.034862	0.46585	0.036503	0.24779
NonTargetingControlGuideForHuman_0158	2	0.23138	0.74207	0.30449	0.80291	0.45804	0.86706	0.59545	1.9097	0.036687	0.47772	0.05617	0.29439
FOXO3	6	0.24667	1.1453	0.055305	0.48019	0.25207	0.76519	0.59514	2.7662	0.0022815	0.04065	0.0056707	0.10999
NonTargetingControlGuideForHuman_0815	2	-0.30521	-0.84664	0.20388	0.72865	0.39608	0.84416	0.59295	1.6486	0.037965	0.48129	0.099222	0.37536
NonTargetingControlGuideForHuman_0190	2	0.50857	1.4018	0.025645	0.35096	0.16099	0.69287	0.59085	1.6285	0.038604	0.48317	0.10342	0.38317
UBTD1	6	-0.37847	-0.75501	0.0091718	0.21824	0.45024	0.86477	0.58637	1.1756	0.00027379	0.044118	0.23975	0.54886
BIRC8	6	0.2529	1.0294	0.049236	0.45743	0.30327	0.79483	0.58297	2.3758	0.00031942	0.05036	0.01751	0.17873
NonTargetingControlGuideForHuman_0152	2	0.18557	0.66035	0.4141	0.85825	0.50903	0.88563	0.57585	2.0492	0.043212	0.49764	0.040445	0.25833
CCDC146	6	0.18297	0.65935	0.16222	0.68734	0.50967	0.88598	0.57423	2.0734	0.00045631	0.062112	0.038135	0.25189
NonTargetingControlGuideForHuman_0221	2	0.4858	1.3994	0.032444	0.38495	0.16171	0.69319	0.57116	1.6448	0.04449	0.50623	0.1	0.37719
NonTargetingControlGuideForHuman_0425	2	0.16076	0.42476	0.48287	0.88962	0.67101	0.93446	0.57038	1.5068	0.044855	0.50854	0.13186	0.42445

Supplementary data- Empiria Studio analysis example- FOXO3

Project: FOXO3

Project Description
Not specified

Researcher's Name
Not specified

Date
6 Aug 2019, 10:05:03 BST

Principal Investigator Name
Not specified

Lab Mailing Address
Not specified

Analysis Software
Empiria Studio Software v1.0/1.1

Data Location
Not specified

Experiment: FOXO3 new

Description
FOXO3 total 786-0 gel 2
FOXO3 786-0 gel 2

Experiment Type
Analysis: TPS + Targets

Date Created
16 Nov 2020, 13:15:45 GMT

Date Edited
16 Nov 2020, 13:18:20 GMT

Instrument Model
Odyssey CLx

Instrument Serial Number
CLX-0169

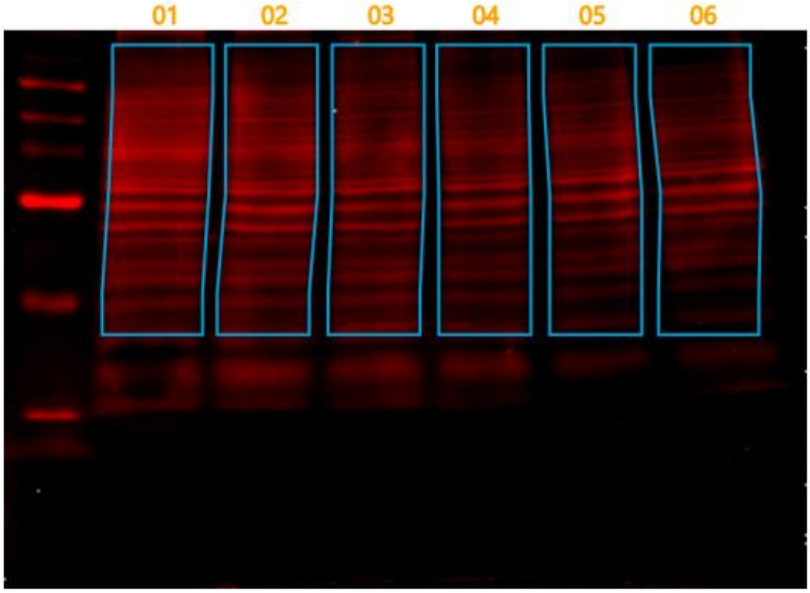
Experimental Analysis

700 Channel

Image Name: 0000661_02_700

Automatic Lane Detection

Protein: FOXO3



Total Protein Stain Analysis Table

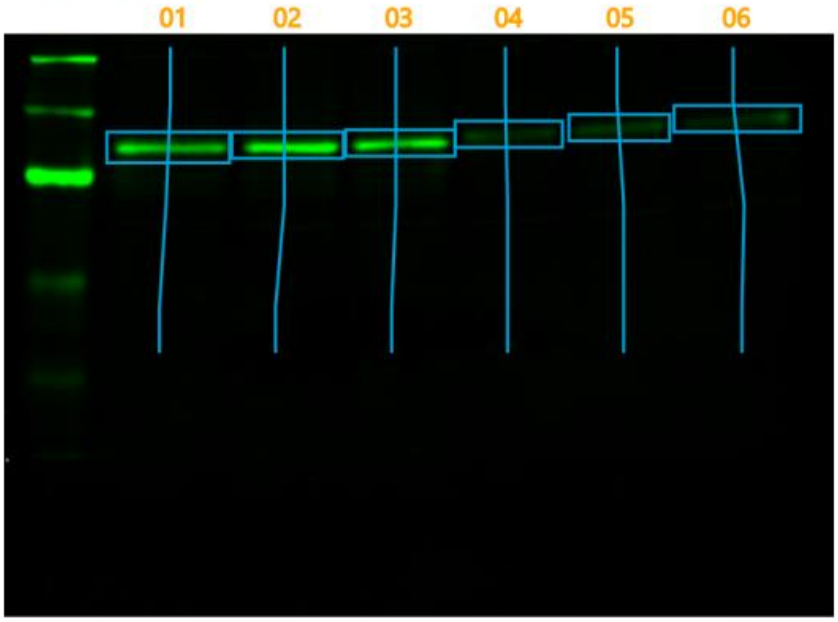
Lane	Name	Signal	SNR	Replicate	^ Avg. Signal	Avg. SNR	Std. Dev.	% CV	Treatment (µl)	Type
				786-0-sgFOXO3	21,200,000	3.23	311,000	1.47	0	Sample
				786-0-sgScr	32,200,000	4.89	6,350,000	19.7	0	Sample

800 Channel

Image Name: 0000667_02_800

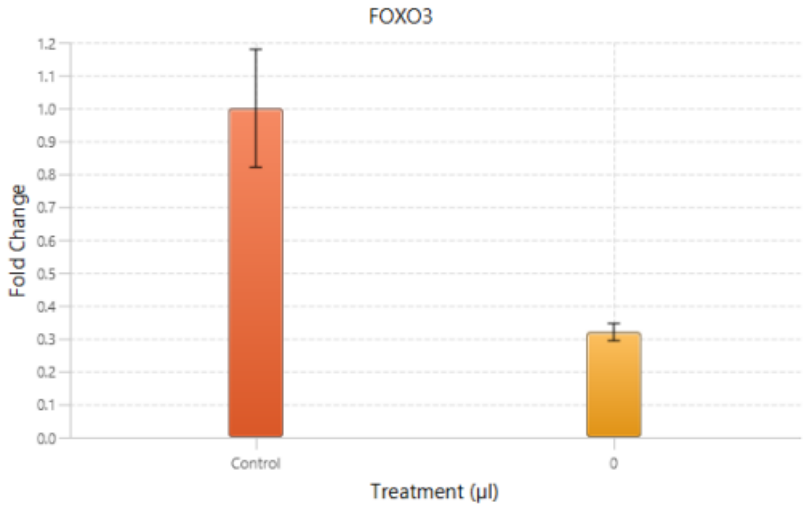
Automatic Lane Detection

Protein: FOXO3



Protein Analysis Table

Lane	Name	MW	Normalized Signal	SNR	Replicate	* Avg. Norm. Signal	Avg. SNR	Std. Dev.	% CV	Treatment (µl)	Type
					786-0-sgFOXO3	519,000	19.6	42,400	8.17	0	Sample
					786-0-sgScr	1,620,000	69.9	290,000	17.9	0	Sample

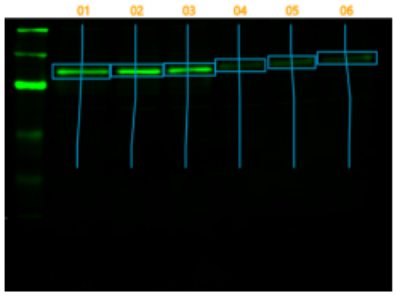


Error bars indicate the standard deviation for each replicate group.

Fold Change Table

Lane	Treatment (X-Axis)	Treatment (µl)	Fold Change
786-0-sgScr	Control	0	1
786-0-sgFOXO3	0	0	0.32

Summary



Experimental Observations
Not specified

Next Steps
Not specified

Summary
Not specified