Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

M.Sc. Risky Oktriani

Born in Semarang, Indonesia

Oral examination: 11.11.2021

The discovery of genes that exhibit strong effects on metastatic pancreatic cancer

Referees: Prof. Dr. Michael Boutros

Prof. Dr. Stefan Wölfl

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

5FU	: 5-Fluorouracil
APS	: Ammoniumpersulfat
ASR	: Age-standardized rate
ATCC	: American Type Culture Collection
ATP	: adenosine triphosphate
BCA assay	: Bicinchoninic acid assay
Cas9	: CRISPR associated protein 9
CCNB1	: Cyclin B1
CHR	: Cell cycle genes homology region
CRISPR	: Clustered regularly interspaced short palindromic repeats
CRISPRi	: CRISPR-interference
crRNA	: CRISPR RNA
DAVID	: The Database for Annotation, Visualization and Integrated Discovery
DMEM	: Dulbecco's Modified Eagle Medium
DNA	: Deoxyribonucleic acid
dNTP	: deoxyribonucleotide triphosphate
dTMP	: deoxythymidine monophosphate
DTT	: Dithiothreitol
dUTP	: deoxyuridine triphosphate
E. coli	: Escherichia coli
EDTA	: Ethylenediaminetetraacetic acid
EMT	: Epithelial-mesenchymal-transition
EPCAM	: Epithelial Cell Adhesion Molecule
FACS	: Fluorescence-activated cell sorting
FBS	: Fetal Bovine Serum
FITC	: Fluorescein isothiocyanate
FOXM1	: Forkhead box M1
GAPDH	: Glyceraldehyde 3-phosphate dehydrogenase
GCO	: Global Cancer Observatory
GFP	: Green Fluorescence Protein
GLOBOCAN	: Global Cancer Incidence, Mortality and Prevelance
GO	: Gene Ontology

HCl	: Hydrogen chloride
HRP	: Horseradish peroxidase
IARC	: International Agency for Research on Cancer
IMDM	: Iscove's Modified Dulbecco's Medium
JCRB	: Japanese Collection of Research Bioresources Cell Bank
Kb	: kilobase
kDa	: kiloDalton
KEGG	: Kyoto Encyclopedia of Genes and Genomes
LB	: Luria-Bertani
MAGeCK	: Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout
MEN1	: Menin1
MET	: Mesenchymal-epithelial transition
MMB complex	: *B-Myb*-MuvB complex
MOI	: Multiplicity of Infection
MWCO	: Molecular weight cut-off
MYBL2	: MYB Proto-Oncogene Like 2
NaCl	: Sodium chloride
NaOH	: Sodium hydroxide
NGS	: Next Generation Sequencing
NTC	: No target control
PanINs	: Pancreatic intraepithelial neoplasias
PBS	: Phosphate-buffered saline
PCI	: Phenol Chloroform Isoamyl Alcohol
PCR	: Polymerase chain reaction
PDAC	: Pancreatic ductal adenocarcinoma
PI	: Propidium Iodide
Polyester	: PET
QC	: quality control
qPCR	: quantitative Polymerase chain reaction
RNA	: Ribonucleic acid
RPMI	: Roswell Park Memorial Institute
SDS-PAGE	: Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
SDS	: Sodium dodecyl sulfate

sgRNA	: single guide RNA
shRNA	: short hairpin RNA
TBST	: Tris-buffered saline-Tween 20
TEMED	: Tetramethylethylenediamine
tracrRNA	: trans-activating crRNA
Tris- HCL	: Tris-Hydrochloride
TU	: Transduction unit
TYMS	: Thymidylate Synthetase
UBE2C	: Ubiquitin conjugating enzyme E2 C
WHO	: World Health Organization
α-tubulin	: alpha-tubulin

SUMMARY

Pancreatic cancer is one of the deadliest cancers. Alterations in gene expression play an essential role during disease development. Also, variations in gene expression are involved in metastasis. Due to the complexity of gene expression alterations, however, it is still challenging to pinpoint and thus possibly target the functional reason behind the aggressiveness of pancreatic cancer. A high-throughput screening was conducted using the CRISPR-Cas9 whole-genome screening strategy to discover genes that might be essential for metastasis, to define processes that may have the potential to affect the metastatic progression of pancreatic cancer. Whole-genome CRISPR-Cas9 screening was conducted in two cell lines from the same parental cells but with different metastatic potential: S2-007 (highly metastatic) and S2-028 (lowly metastatic). The screening was performed, and samples were analyzed by MAGeCK analysis to select candidate genes. *TYMS*, *MEN1*, *MYBL2* were selected based on their gene essentiality. In vitro validation was then conducted to confirm the gene essentiality in both cell lines. Cell proliferation, cell viability, soft agar colony formation assay was performed to narrow down the candidate. In vitro validation, *MYBL2* showed the most promising potential as an essential gene in the metastatic pancreatic cancer cell. Wound healing assay and invasion assay were done to further validate *MYBL2*. Next, KEGG pathway analysis and Ingenuity analysis showed that *MYBL2* interactions mainly in cell cycle network.

Further experiments were performed which focused more on the effect of *MYBL2* knockout in the cell cycle. *FOXM1* gene knockout was then added to the experiments since *FOXM1* plays a role in the G2/M phase. The double transduction was done, followed by cell sorting. Sorted cells were then cultured, and cell viability assay and cell cycle assay were performed. All knockouts were validated using Western blot.

MYBL2 knockout showed inhibition in cell proliferation, cell viability, and soft agar colony formation assay in the S2-007 cell line. In addition, *MYBL2* loss also showed invasion inhibition in S2-007. The double knockout *FOXM1-MYBL2* affects cell viability more than the *FOXM1* or *MYBL2* single knockout. The cell cycle assay results demonstrated that the *FOXM1-MYBL2*, *FOXM1*, and *MYBL2* knockout showed cell cycle arrest at G2/M in S2-007. Altogether the results showed that *MYBL2* showed potential as an essential gene in the high metastatic pancreatic cancer cell line. However, the mechanism and the reason why *MYBL2* showed essentiality only in the S2-007 cell line is still unclear. Therefore, more experiments still need to be conducted to achieve the goal of further elucidating the role of *MYBL2* in metastatic pancreatic cancer.

ZUSAMMENFASSUNG

Pankreaskrebs gehört zu den tödlichsten Krebsarten. Veränderungen der Gene oder ihrer Expression tragen zur Krankheitsentstehung und Metastasierung bei. Die Komplexität der veränderten Genexpressionsmuster erschwert dabei, die Grundlagen der Aggressivität des Pankreaskrebs zu verstehen und gezielt zu therapieren. Im Rahmen eines High-Throughput Screens wurde hier ein genomweiter CRISPR/Cas9 Screen dazu benutzt, neue und potenziell essenzielle Gene zu finden, die die Metastasierung vorantreiben und dadurch die zellulären Signalwege in metastasierendem Pankreaskrebs besser zu verstehen.

Ein genomweiter CRISPR/Cas9 Screen wurde in zwei Zelllinien durchgeführt. Beide Zelllinien stammen von derselben Mutterzelle ab, wobei sich die Linien S2-007 (hochgradig metastasierend) und S2-028 (wenig metastasierend) in ihrem Metastasierungspotential unterscheiden. Der Screen wurde durchgeführt und mit MAGeCK analysiert, um Kandidatengene zu finden. *TYMS, MEN1* und *MYBL2* wurden als essenzielle Gene identifiziert. Die Kandidaten wurden anschließend *in vitro* mittels Zellproliferations-, Zellviabilitäts- und Soft Agar Colony Formation Assays validiert. Dabei zeigte *MYBL2* das stärkste Potential, als ein essenzielles Gen in Pankreaskrebszellen zu funktionieren. Wunheilungs- and Invasionassays wurden durchgeführt, um MYBL2 weiter zu validieren. Die Analyse mit Hilfe von KEGG Pathway und eine Ingenuity Analyse zeigten, dass MYBL2 hauptsächlich im Zellzyklus involviert ist.

Weitere Experimente zeigten die Effekte eines *MYBL2* Knockouts auf den Zellzyklus. Zusätzlich wurden Zellen mit einem *FOXM1* Knockout wurden hergestellt, da *FOXM1* eine Rolle in der G2/M Phase spielt. Die doppelt transduzierten Zellen wurden mit einem Cellsorter identifiziert, weiter kultiviert und dann mittels Zellviabilitäts- und Zellzyklusassays untersucht. Alle Knockouts wurden mittels Western blot validiert.

In der S2-007 Zelllinie führte der Knockout von *MYBL2* zu verminderter Proliferation und Zellviabilität und die Zellen bildeten weniger Kolonien im Soft Agar Assay. Darüber hinaus verringerte der *MYBL2* Knockout die Invasionsfähigkeit der Zellen. Der gleichzeitige Doppel-Knockout von *FOXM1* und *MYBL2* reduzierte die Zellviabilität stärker als jeder einzelne Knockout. Mittels Zellzyklusassay zeigte sich, dass S2-007 Zellen mit entweder einem Einzel-Knockout von *FOXM1* oder *MYBL2* oder dem Doppel-Knockout in der G2/M Phase akkumulieren. Die Ergebnisse deuten darauf hin, dass *MYBL2* ein essenzielles Gen in der hier untersuchten hochgradig metastasierenden Pankreaskrebszelllinie ist. Allerdings ist der Grund dafür, dass *MYBL2* nur in der S2-007 Zelllinie essenziell ist, unklar. Weitere Experimente sind daher nötig, um die Rolle von *MYBL2* in Pankreaskrebs und während der Metastasierung aufzuklären.

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 The incidence of pancreatic cancer

Amongst all types of cancer, pancreatic cancer has one of the highest mortalities followed by lung, colorectum, prostate, and breast cancer (Fig. 1.1). Based on global comparison by regions with regards to the incidence and mortality of pancreatic cancer, Europe (western, central, eastern, northern, and southern) positioned in the top five. In Germany itself, the incidence of pancreatic cancer is the highest in Europe. Observations on all the calculated mortality rates in several regions in Europe revealed that the rate is high, more than 90%. Moreover, the 5-year survival rate of pancreatic cancer in Germany is only 9%. (Fig. 1.2).

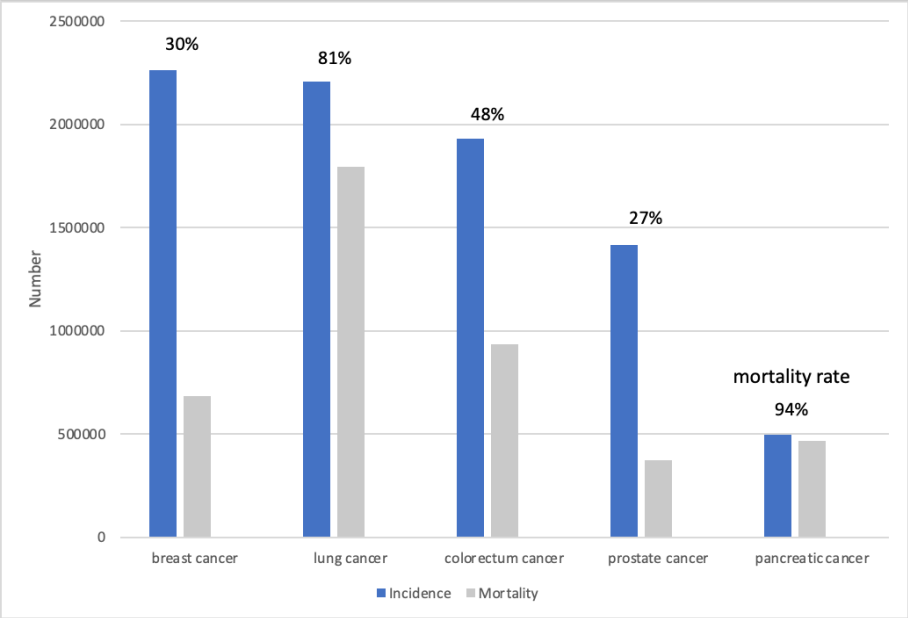


Figure 1.1. Pancreatic cancer has the highest mortality rate compared to breast, lung, colorectum, and prostate cancer. Incidence and mortality numbers express the absolute number of cases within the population in 2020. The mortality rate ratio (%) is calculated by dividing the incidence by the mortality number. The figure was copied from GLOBOCAN (WHO-IARC, 2020).

The reason behind this phenomenon is the lack of a screening method. Early screening methods and/or biomarkers for pancreatic cancer are not clinically implemented, because there has yet been an adequately proven early screening method. Most patients are admitted to a clinic solely due to symptoms typically observed at the advanced stages of the disease. Another reason is the lack of significant progress in terms of treatment for this type of cancer (Kamisawa et al., 2016; Robert Koch Institut, 2020; Zhang et al., 2018).

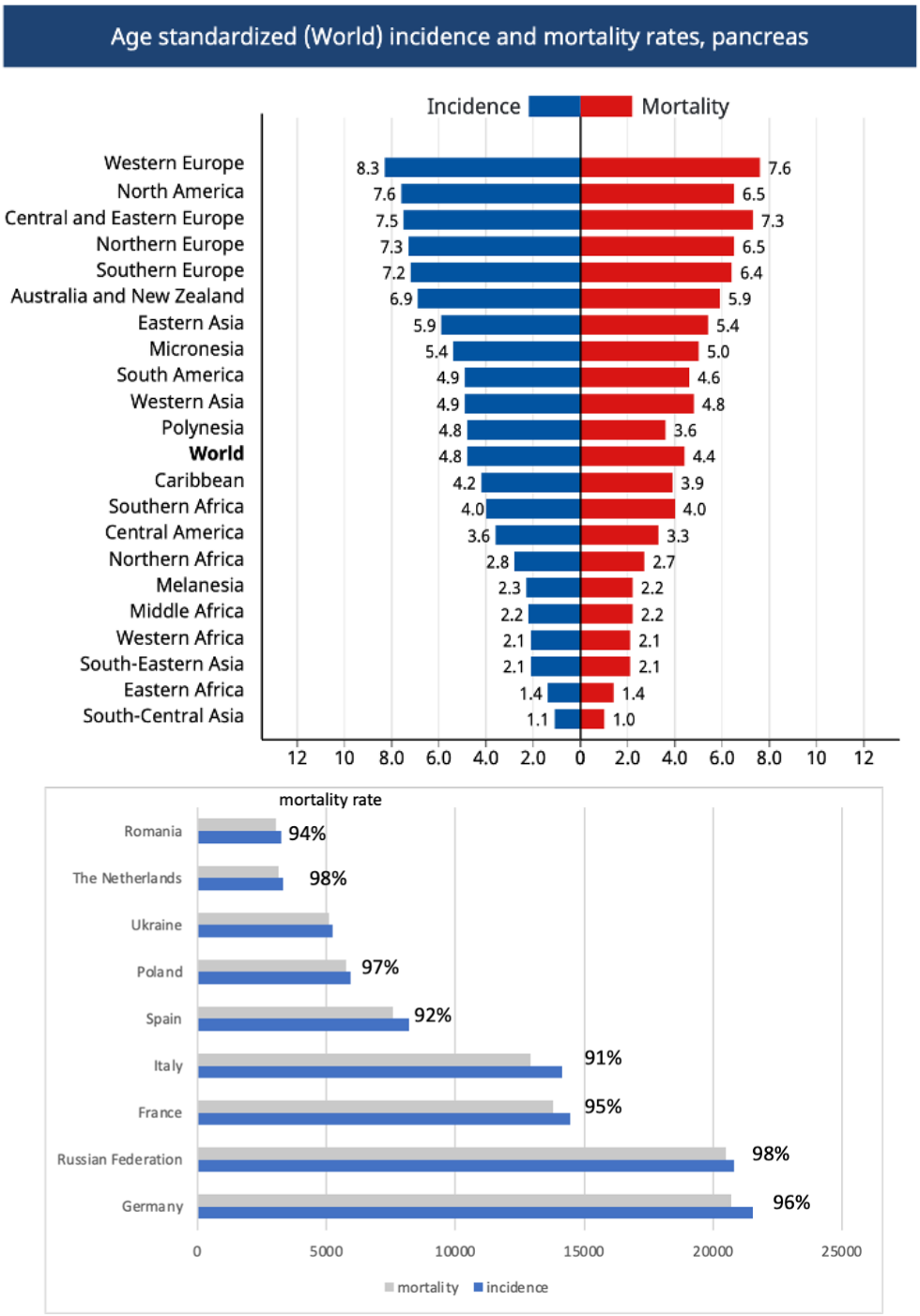


Figure 1.2. The incidence and mortality of pancreatic cancer in the world. The age-standardized rate (ASR), incidence and mortality numbers were collected from GLOBOCAN (WHO-IARC, 2020). ASR is the rate of cancer incidence that would have been observed in a standard age structure. (Top panel) The graph shows the incidence and mortality rate of pancreatic cancer in the world. (Bottom panel) The incidence and mortality rate of pancreatic cancer in Europe are shown. The numbers represent the mortality and incidence rate both sexes all ages in the population. The mortality rate ratio (%) was calculated by dividing the mortality number by the incidence number.

1.2 Biological hallmark of pancreatic cancer

Exocrine (acinar), epithelial (ductal), and endocrine (α, β, δ, ε) cells are three major cell types, which are present in the pancreas. Pancreatic ductal adenocarcinoma (PDAC) is the most common pancreatic cancer (Kamisawa et al., 2016). PDAC could arise from two distinct lineages, namely acinar and ductal cells. The acinar cells undergo an acinar to ductal metaplasia when the cells differentiate to have a more ductal-like phenotype. There is the effect of the inducements of micro- and macroenvironments. During the process, the cells are prone to pro-oncogenic hits, such as a mutation in KRAS. Having this mutation, cells transform into pancreatic intraepithelial neoplasias (PanINs). After this initial step, several genetic alterations and mutations occur in the genetic progression to PDAC (Fig. 1.3) (Orth et al. 2019).

Involvement of driver mutations in KRAS, CDKN2A, TP53, and SMAD4 has shown to be the most frequent cause (Jones et al., 2008). KRAS mutations occur in >95% of human PDAC causing the protein to remain active and affecting the downstream signaling from growth factor receptors (Kamisawa et al., 2016; Morris et al., 2010). CDKN2A encodes an essential cell-cycle regulator. Mutations in this gene were found in more than 90% of PDAC. Two tumor suppressor genes, TP53 and SMAD4, appear in the late progression of PDAC and have different types of involvement: TP53 affects the cellular stress response whereas SMAD4 mediates signaling downstream of transforming growth factor β (Kamisawa et al., 2016).

In terms of treatment, pancreatic cancer can be divided into several groups, namely resectable, locally advanced or un-resectable, as well as metastatic pancreatic cancer (D Li et al., 2004). The standard treatment for resectable pancreatic cancer is surgery (Kamisawa et al. 2016). Chemotherapy is still the best options for un-resectable pancreatic cancer. Initially, gemcitabine was applied as the first line of chemotherapy(Burris & Storniolo, 1997). However, the very high degree of chemoresistance to gemcitabine in pancreatic cancer due to several mechanisms of actions lead to the necessity to not only rely on gemcitabine monotherapy alone (Bafna et al., 2009; Carpenter et al., 2020; Kuramitsu et al., 2012; Weadick et al., 2021; Zhou et al., 2020)

Several combination therapies are being used in the clinical environment to treat metastatic cancer as alternatives to gemcitabine monotherapy, such as FOLFORINOX (combination of oxaliplatin, irinotecan, leucovorin, fluorouracil), gemcitabine plus nab-paclitaxel, and gemcitabine plus erlotinib. While the median overall survival for these combinations is higher than that of gemcitabine alone for metastatic pancreatic cancer, there is the issue of toxicity

(Conroy et al., 2011; Moore et al., 2007; von Hoff et al., 2011). Therefore, combinations of chemotherapy would only be given to patients with relatively good physical conditions (Springfeld et al., 2019). Based on all the information above, there is still a need to find an effective therapy for pancreatic cancer, specifically for metastatic pancreatic cancer.

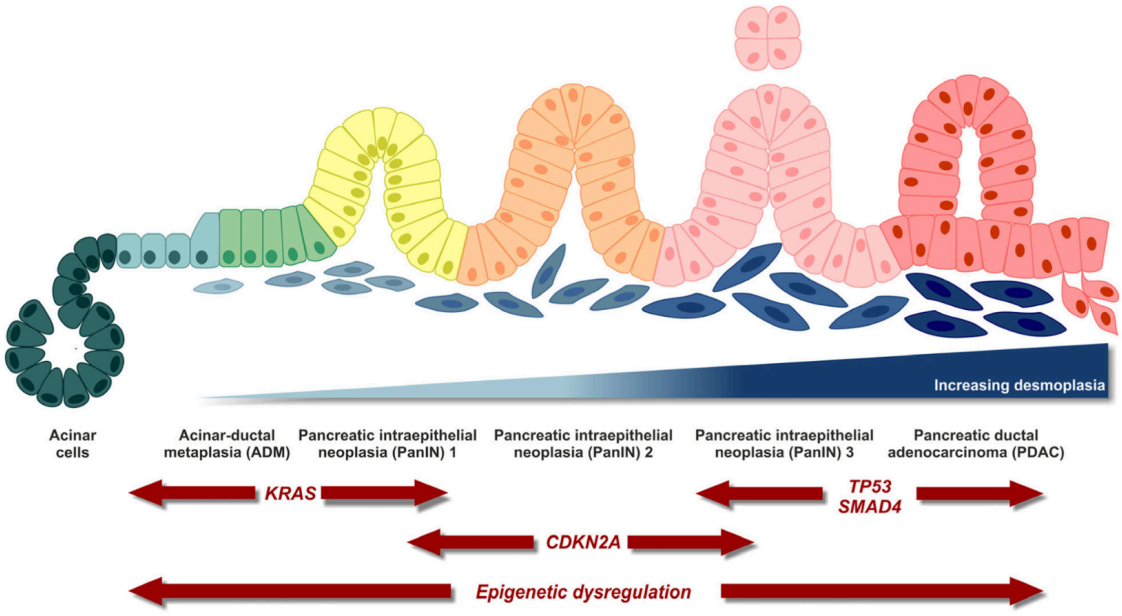


Figure 1.3. Genetic progression of PDAC. The progression begins with the acinar cells transforming into ductal-like cells in a process called acinar-ductal metaplasia, during which acinar cells gain inducements from micro- and macroenvironments, making them prone to activation of proto-oncogene mutations, such as in KRAS. Thereby, the cells transform to pancreatic intraepithelial neoplasias (PanINs). Afterward, several mutations and alterations in CDKN2A, TP53, and SMAD4 and epigenetic dysregulation are involved in the sequential PDAC progression from PanIN1 to PDAC. (Figure copied from Orth et al., 2019).

1.3 Metastatic pancreatic cancer

Pancreatic cancer reaches the metastatic stage when the cancer cells travel through the lymphatic system or blood stream to distant organs, such as the liver and lung (Orth et al., 2019). In general, the metastatic process involves several steps, from the detachment of the primary mass to the cells' colonization in the new organ (Fig. 1.4). Moreover, during the cancer cell dissemination and clonal expansion, which causes metastatic progression, the gene alteration is even more complex (Yachida et al., 2010).

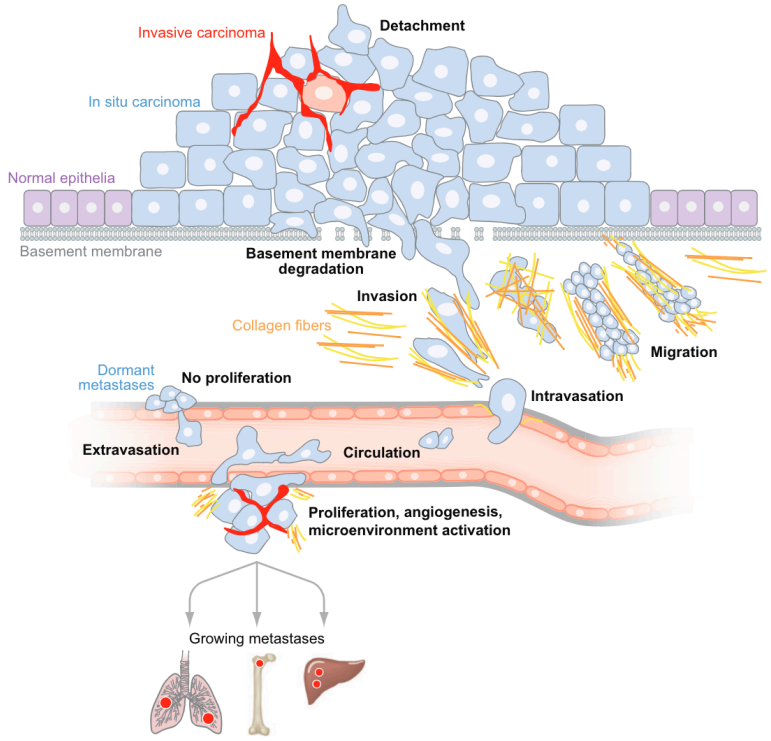


Figure 1.4. The principal steps in metastasis. The main steps in metastasis involve detachment from the primary mass, invasion of the local host tissue stroma, penetration to the local lymphatic and blood vessels, survival of the cells within the circulation, arrest of cells in the capillaries or venules of other organs, penetration to the parenchyma, adaptation to the newly colonized microenvironment, and then division to form a new tumor. (Figure copied from Bacac & Stamenkovic, 2008).

A study regarding clonal relationships in 13 pancreatic cancer patients showed that metastatic seeding cells in the early stage required more driver mutations than the primary cancer cells. The phylogenetic trees study also showed evidence of organ-specific genetic rearrangements in the metastatic process (Campbell et al., 2010).

A critical step for cells from the primary site to move to the distant organ is epithelial-mesenchymal-transition (EMT). EMT is controlled by EMT transcription factors, such as SNAIL, ZEB1, and TWIST. The transition of epithelial cells to mesenchymal cells is necessary for cell invasion. Traditionally defined EMT mechanisms suggested that cells migrate as single cells in blood or lymphatic circulation, before later on during the extravasation the process, turning into mesenchymal-epithelial transition (MET). Recent developments showed an alternative process. To adapt to the changing environment, cancer cells undergo partial EMT, where cells still have both epithelial and mesenchymal properties. Consequently, during the migration, cells move as clusters instead of single cells (Aiello et al., 2018; Brabletz et al., 2018).

1.4 CRISPR-Cas9 system

CRISPR-Cas9 system (Clustered regularly interspaced short palindromic repeats – CRISPR associated protein 9) is a system found firstly in microbial organisms as an adaptive immune system. Overall, there are three types of CRISPR systems identified in different bacteria and archaea. CRISPR-Cas9 is the type II CRISPR system consisting of Cas9, CRISPR RNA (crRNA) array (encodes the guide RNAs), and trans-activating crRNA (tracrRNA) that process the crRNA array into separate units. Furthermore, the crRNA contains 20 nucleotides, which function is to lead Cas9 to recognize the target sequence. The CRISPR-Cas9 originates from *Streptococcus pyogenes*. The target sequence must be followed by the protospacer adjacent motif (PAM) 5'-NGG-3'. Different Cas9s found in different species have different PAM sequences. In addition, the crRNA and tracrRNA are fused in mammalian cells, forming a single guide RNA (sgRNA). Thus, controlling the CRISPR-Cas9 system in mammalian cells can be adjusted by changing the 20 nucleotides within the sgRNA (Doudna & Charpentier, 2014; Hsu et al., 2014a; Ran et al., 2013). The mechanism on how CRISPR-Cas9 works is explained in Fig. 1.5.

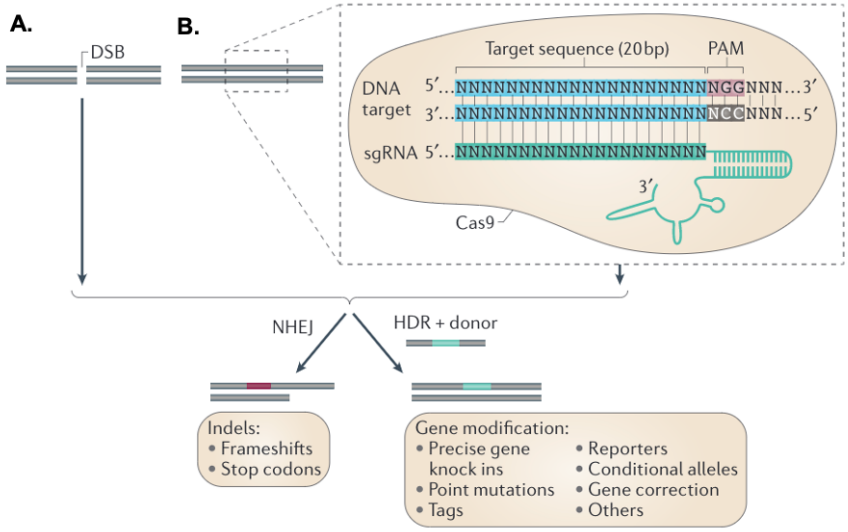


Figure 1.5. Gene editing in the CRISPR-Cas9 system. (A) When a double-strand break (DSBs) occurs, there are two possibilities in the DNA repair pathways: the non-homologous end joining (NHEJ) or the homology-direct repair (HDR) pathway. The NHEJ pathway makes the DNA prone to insertion and deletion mutations (indels) leading to frameshifts or introducing premature stop codons. In contrast, the HDR pathway creates specific modifications by using the donor DNA template. **(B)** In the CRISPR-Cas9 system, the sgRNA is located next to the DNA sequence target complementary to its sequence and adjacent to the PAM sequence (NGG or NAG). Cas9-mediated double-strand break induce the DNA repair pathways in the same manner as the information explained above (Figure copied from Sánchez-Rivera & Jacks, 2015).

1.5 Thesis scope

High-throughput whole-genome screening has become an important approach to study biological process. Complementary, whole-genome CRISPR-Cas9 screening has proven to be a robust method to identify genes that essential for cell viability in cancer (Shalem et al., 2014a; Wang et al., 2015, 2017). An essential gene is a gene whose loss of function reduces the viability of the living system(Bartha et al., 2018; Rancati et al., 2018).

Moreover, a study showed that the CRISPR system is better than the short hairpin RNA (shRNA) and CRISPR-interference (CRISPRi) based technologies for identifying essential and non-essential genes. There are several reasons for this, such as less variation in the data, no off-target effects that could produce a low false-discovery rate (Evers et al., 2016).

S2-028 and S2-007 cell lines are two different cell lines that originated from the same parental SUIT-2 cell line, which in turn represents a pancreatic cancer metastasis to the liver. These cell lines have been studied and showed a low and very high metastatic potential for S2-028 and S2-007, respectively. In vitro study showed that cells attachment to Matrigel, the type IV collagenolytic activity, and chemotactic activity toward Matrigel of S2-028 was lower than S2-007. Moreover in vivo experiments proved that S2-007 triggered pulmonary metastasis after being injected into mice. On the contrary, there was no sign of a metastatic site within S2-028 (Buchholz et al., 2003; Katsuki & Ide, 1987; Michl et al., 2003; Taniguchi et al., 1992, 1994).

In this study, the CRISPR-Cas9 whole-genome screening approach was applied in order to discover genes that are essential for metastatic pancreatic cancer. The genetic hits discovered from the comparative screening of S2-028 and S2-007 were validated and analyzed further to verify the gene essentiality only in a metastatic cancer cell.

2. MATERIALS AND METHODS

2.1 List of reagents

Table 2.1. List of common reagents.

Reagents	Catalogue number	Manufacturer
Yeast extract	1133.0500	Serbu, USA
Peptone from casein, tryptic digest	70172-500G	Sigma Aldrich, USA
NaCl (sodium chloride)	BP358-1	Fischer Chemical, Germany
Agar-agar Kobe I	5210.3	Roth, Germany
Trizma Base	T1503-1KG	Sigma Aldrich, USA
Glycine	23390.04	Serva, Germany
SDS pellets	CN 30.3	Roth, Germany
HCL 37%	20252.290	VWR Chemicals, USA
NaOH (sodium hydroxide, pellets)	SC-203387	Chem Cruz, USA
Methanol	32213-2.5L-M	Sigma Aldrich, USA
6-aminocaproic acid	A7824-100G	Sigma Aldrich, USA
Tris HCL (Trizma-Hydrochloride)	T3253-1KG	Sigma Aldrich, USA
Tween 20	P2287-500mL	Sigma Aldrich, USA
Skim milk powder	1602.0500	Serbu, USA
Ethanol absolute	64-17-5	VWR Chemicals, USA
2 mercaptoethanol	M3148	Sigma, USA
Acetic acid	A0820, 2500 PE	Applichem, Germany
EDTA	1.08418	Merck, USA
Acrylamide	3029.1	Roth, Germany
TEMED	2367.3	Roth, Germany
Ammoniumpersulfat (APS)	A1142,0250	AppliChem, Germany
Isopropanol	2474.3	Roth, Germany

2.2 Cell culture

For the optimization of CRISPR lentivirus production, BxPC3 ((American Type Culture Collection (ATCC) CRL-1687), Capan-1 (ATCC HTB-79), and SUIT-2 (Japanese Collection of Research Bioresources Cell Bank (JCRB) 1094) were used. For CRISPR screening, S2-028 and S2-007, which are SUIT-2 clones, were used. S2-028 and S2-007 were kindly provided by Prof. Malte Buchholz, University of Ulm, Germany). S2-028 and S2-007 are two pancreatic cancer cell lines derived from the same parental cells (SUIT-2) but having different metastasis characteristics. Based on the previous studies, S2-028 has less metastatic potential compared to S2-007 (Taniguchi et al., 1992, 1994). HEK 293T (ATCC CRL-11268) cells have also been used in this study for lentivirus production.

SUIT-2, S2-028, and S2-007 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) complete medium consisting of DMEM (Cat. No. 41965062, Gibco, USA) supplemented with 10% Fetal Bovine Serum (FBS) (Cat. No. 1050006, Gibco, USA), and 1% Penicillin Streptomycin (Cat. No. 15140122, Gibco, USA). BxPC3 and HEK 293T were cultured in Iscove's Modified Dulbecco Medium (IMDM) (Cat. No. 21056023, Gibco, USA) supplemented with 10% FBS and 1% Penicillin Streptomycin. Roswell Park Memorial Institute (RPMI) 1640 (Cat. No. 21875091, Gibco, USA) supplemented with 10% FBS and 1% Penicillin Streptomycin was then used to culture Capan-1. Cells were cultured and maintained in the 5% CO2 incubator at 37°C.

Several kinds of flasks and plates have been used to culture the cells, such as 25 cm^2 (Cat No. 13640, Greiner Bio-One, Germany), 75cm^2 flask (Cat No. 12667, Greiner), 175 cm^2 (Cat. No. 12649, Greiner), 100 mm plates (Cat. No. CLS430167, Corning, USA), 150 mm plates (Cat. No. CLS430599, Corning, USA), 6-well plate (Cat No. 657160, Greiner Bio-One, Germany), 12 well plate (Cat No. 665180, Greiner Bio-One, Germany), and 96 well plate (Cat No. 655101, Greiner Bio-One, Germany). Cells were observed and pictured using Zeiss microscope (Inverted Zeiss Cell Observer.Z.1, Germany). Cell line authentication was performed at the DKFZ Genomics and Proteomics Core Facility. S2-028, and S2-007 cells were reauthenticated by examining the expression level of Serpine 2 and Claudin 4 expression by quantitative polymerase chain reaction (qPCR) (Buchholz et al., 2003; Michl et al., 2003). Mycoplasma was also checked regularly by polymerase chain reaction (PCR) to confirm that all cells were mycoplasma-free.

2.3 Blasticidin and puromycin treatment

3000 and 6000 cells were plated in a 6-well plate for blasticidin and puromycin treatment, respectively, on the day before the treatment. After overnight incubation, blasticidin and puromycin concentrations were prepared within the range of 0 - 7.5 µg/mL for blasticidin and 0 - 6.5 µg/mL for puromycin in complete medium. Cells were washed with Phosphate-buffered saline (PBS) (Cat. No. 10010056, Gibco, USA) once, and then 100 uL of each drug dilution was pipetted into each well. Cells were then incubated for 3 days and 6 days for puromycin (Cat. No. A1113803, Gibco, USA) and blasticidin (Cat. No. ant-bl-05, InvivoGen, USA), respectively. At the end of the incubation time, the medium containing the drug was discarded, and cells were washed with PBS once. Resazurin (5 mg/mL) dilution in DMEM complete medium (Cat. No. 189900010, Acros Organics, USA) was added before one hour incubation

of the plate. The reduction of resazurin (blue) to resazurin (fluorescence pink) by mitochondrial enzymes in the viable cells were then detected (O'Brien et al., 2000). Fluorescence microplate reader at excitation 544 nm and emission 590 nm (Fluostar Galaxy, BMG Labtech, Germany) was used for the measurement.

The blasticidin and puromycin kill curves were calculated by comparing the number of the control (without drug) and the treatment. The ratio between those groups was then plotted against the drug concentrations.

2.4 sgRNA cloning, colony PCR, and sequencing

2.4.1 sgRNA cloning into MP-783

The plasmid used for cloning of the sgRNA-targeting Epithelial Cell Adhesion Molecule (*EPCAM*) and validation has the same backbone with the plasmid used for the CRISPR library (Figure 2.1).

First, the empty plasmid was digested by the AaRI restriction enzyme (Cat. No. ER1581, Thermo Fischer, USA). Digestion was done according to the instruction manual but with modifications. The process was initiated by mixing 0.75 µg DNA plasmid, 1x Buffer, 20x Oligonucleotide, and 4U AaRI in a total volume of 20 µL. Each reaction was then incubated at 37°C for 2 hours, followed by heat inactivation at 65°C for 20 minutes.

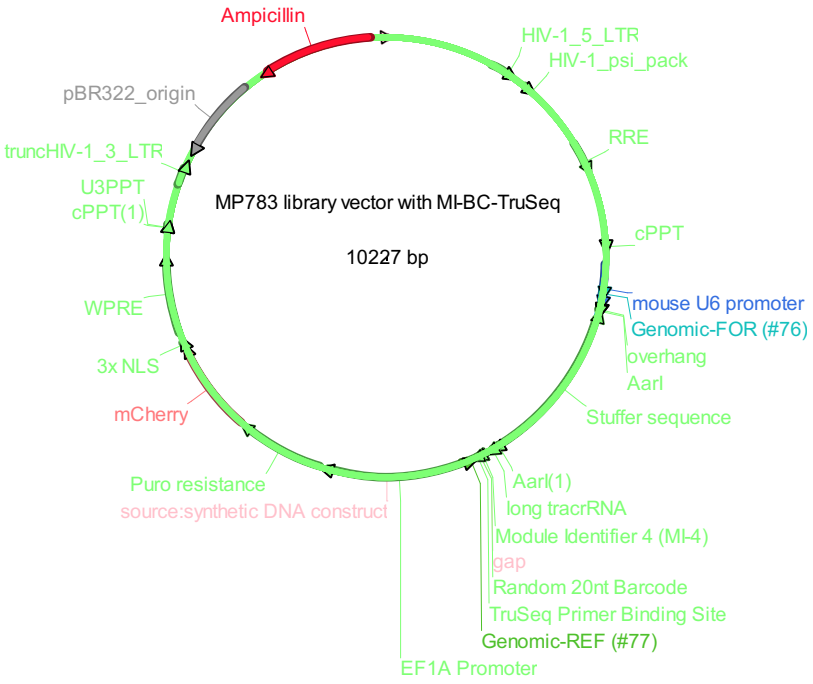


Figure 2.1. MP-783 library vector. (Figure copied from Boettcher et al., 2018).

Digestion of the plasmid was analyzed by running gel electrophoresis. 1.5% agarose gel (Cat. No. 3810.4, Roth, Germany) was prepared in 1x Tris Acetic Acid EDTA Buffer (Appendix 1) supplemented with Peq Green (Cat. No. ACT-1DMG04-1ML, PeqLab, Germany).

4 µL of Peq Green was added for every 100 µL of agarose gel that was made. The solidified gel was placed in the electrophoresis chamber containing 1x TAE Buffer. Digested product was mixed with 6x loading dye to a final concentration of 1x dye, and then loaded into the gel. Electrophoresis was performed at 100 Voltage for 1 hour, with GeneRuler 1 kilobase (kb) DNA Ladder (Cat. No. SM0311, Thermo Fischer Scientific, USA) used as a size marker. The gel was visualized using gel documentation C200 (Azure Biosystem, USA).

Two fragments appeared after the restriction enzyme digestion: 8934 base pairs (bp) and 1293 bp. The 8934 bp fragment was then cut and extracted using NucleoSpin Gel and PCR Clean-up (Cat. No. 740609.50, Macherey-Nagel, Germany), following the manufacturer's protocol. Plasmid DNA was eluted in nuclease-free water (Cat. No. AM9930, Invitrogen, USA), and the concentration was measured using NanoDrop (NanoDrop Spectrophotometer ND-1000). This plasmid was either used directly or stored in -20°C for long term storage.

The next step was sgRNA phosphorylation. The list of sgRNA sequences was shown in Table 2.2. To insert the sgRNA into the plasmid, the 5'-TTGG-3' sequence was added at the beginning of the forward sequence and 5'- AAAC - 3' for the reverse sequence. Phosphorylation of sgRNA primers was conducted using T4 Polynucleotide Kinase (Cat. No. M0201S, New England Biolabs, UK) by mixing 100 µM sgRNA forward primer, 100 µM sgRNA reverse primer, 1x T4 DNA Ligase Buffer (Cat. No. B0202S, NEB, UK), and T4 PNK in a total volume of 10 µL.

The mixture was then incubated at 37°C for 45 minutes, 95°C for 20 minutes and 30 seconds, followed by an annealing process by cooling down the mix at 0.1°C per second to a final temperature of 22°C.

Phosphorylated sgRNAs and digested plasmids were then ligated together. Before proceeding with the ligation, the annealed and phosphorylated sgRNAs were diluted 1:100 in nuclease-free water. Ligation was then done by mixing 50 ng digested plasmid, 1 µL annealed-phosphorylated sgRNAs, 1x Quick ligase reaction buffer, and quick ligase enzyme (Cat. No. M2200S, NEB, UK) in a total volume of 5.5 µL.

The reaction was then gently mixed by pipetting and a brief spin. The ligation was continued by incubating the reaction mixture at 25°C for 5 minutes, after which it was ready to be used for transformation either directly or following storage at -20°C.

Table 2.2. sgRNA list.

Construct name	sgRNA sequence
EPCAM	Forward 5'- GATCCTGACTGCGATGAGAG - 3'
	Reverse 5'- CTCTCATCGCAGTCAGGATC - 3'
*TYMS*_sg1	Forward 5'- GTCAGGAAGGACGACCGCAC - 3'
	Reverse 5'- GTGCGGTCGTCCTTCCTGAC - 3'
*TYMS*_sg2	Forward 5'- CACGGGCACCGGCACCCTGT - 3'
	Reverse 5'- ACAGGGTGCCGGTGCCCGTG - 3'
*TYMS*_sg3	Forward 5'- ACCGGCACCCTGTCGGTATT - 3'
	Reverse 5'- AATACCGACAGGGTGCCGGT - 3'
*TYMS*_sg4	Forward 5'- CTGCATGCCGAATACCGACA - 3'
	Reverse 5'- TGTCGGTATTCGGCATGCAG - 3'
*TYMS*_sg5	Forward 5'- CCTGTCGGTATTCGGCATGC - 3'
	Reverse 5'- GCATGCCGAATACCGACAGG - 3'
*MEN1*_sg1	Forward 5'- CACACCGGAGCTGTCCAATT - 3'
	Reverse 5'- AATTGGACAGCTCCGGTGTG - 3'
*MEN1*_sg2	Forward 5'- CACCCCCTTCTCGAGGATAG - 3'
	Reverse 5'- CTATCCTCGAGAAGGGGGTG - 3'
*MEN1*_sg3	Forward 5'- CTGTCCCTCTATCCTCGAGA - 3'
	Reverse 5'- TCTCGAGGATAGAGGGACAG - 3'
*MEN1*_sg4	Forward 5'- CAGGTCGGCCACGGGAAAGT - 3'
	Reverse 5'- ACTTTCCCGTGGCCGACCTG - 3'
*MEN1*_sg5	Forward 5'- CAGGCGCACCACGTCGTCGA - 3'
	Reverse 5'- TCGACGACGTGGTGCGCCTG - 3'
*MYBL2*_sg1	Forward 5'- CTGACCAGCAATGCCAGTAC - 3'
	Reverse 5'- GTACTGGCATTGCTGGTCAG - 3'
*MYBL2*_sg2	Forward 5'- TGAATCCAGACCTTGTCAAG - 3'
	Reverse 5'- CTTGACAAGGTCTGGATTCA - 3'
*MYBL2*_sg3	Forward 5'- ATCGAGCTGGTTAAGAAGTA - 3'
	Reverse 5'- TACTTCTTAACCAGCTCGAT - 3'
*MYBL2*_sg4	Forward 5'- GGTGGTTGTGCCAGCGTTCA - 3'
	Reverse 5'- TGAACGCTGGCACAACCACC - 3'
*MYBL2*_sg5	Forward 5'- CAGGCGCACCACGTCGTCGA - 3'
	Reverse 5'- CGGTCCAGCAAGACTTCTTC - 3'
No target control 1 (NTC1)	Forward 5'- AACTTCTTAAGCCAACTGAC - 3'
	Reverse 5'- GTCAGTTGGCTTAAGAAGTT - 3'
No target control 2 (NTC2)	Forward 5'- CGGTAAAGCCTACCGCAGAA - 3'
	Reverse 5'- TTCTGCGGTAGGCTTTACCG - 3'

The transformation was conducted using One-Shot STBL3 Escherichia coli (E. coli) (Cat. No. C737303, Invitrogen, USA) according to the manufacturer's protocol with some modifications. One vial of STBL3 cells was thawed and combined with 1 µL of ligation mix, and then gently mixed and incubated on ice for 30 minutes. After the incubation, the cells were incubated at 42°C for 45 seconds without shaking and then quickly moved and kept on ice for 2 minutes.

250 µL pre-warmed SOC medium was then added into the tube before the process continued to incubation at 37°C for 1 hour in a shaking incubator. After incubation, the cells were centrifuged at 1000 g (Eppendorf centrifuge 5415D, Germany) for 10 minutes. The supernatant was discarded, and the pellet was resuspended by pipetting. Luria-Bertani (LB)-Ampicillin (100 µg/mL) agar was prepared previously as a growth medium (Table 2.3). The pellet was then spread onto the medium and then incubated at 37°C overnight.

Table 2.3. *Escherichia coli* growth media.

Luria-Bertani (LB) Broth	
Component	Concentration
Yeast extract	5 g/L
Peptone from casein, tryptic digest	10 g/L
NaCl	5 g/L
Luria-Bertani (LB) Agar	
Agar Kobe I	20 g/L
Yeast extract	5 g/L
Peptone from casein, tryptic digest	10 g/L
NaCl	5 g/L

2.4.2 Colony PCR and sequencing for MP-783 cloning validation

Colony PCR and sequencing were done next to screen the colonies growing on the LB-Ampicillin agar for the success of sgRNA insertion. Firstly, the crude DNA was obtained by heating colonies during the process. Then, the PCR mixture was made by mixing 1x PCR buffer, 2 mM dNTP (Cat. No. M3015.4500, Genaxxon Bioscience, Germany), 100 nM forward primer (Biomers, Germany), 100 nM reverse primer (Biomers, Germany), (2.5 Units) Taq Polymerase (Cat. No. 201207, Qiagen, Germany), and nuclease-free water to a total volume of 25 µL. The colony was selected using a sterile tip and then immediately placed and mixed into the PCR mixture. For every transformation, five colonies were chosen for colony PCR. The forward primer was the forward sequence of sgRNA, and the reverse primer was 5'-CAGTGACTGGAGTTCAGACG-3'. The reaction mixed was incubated in the thermal cycler using touchdown PCR as follows: pre denaturation at 94°C for 3 minutes, and then 16 cycles consisting of denaturation at 94°C for 30 seconds, annealing at 69-54°C (depending on the annealing temperature of each primer) with a gradual decrease of 1°C per cycle, elongation at 72°C for 30 seconds, and then 19 cycles consisting of denaturation at 94°C for 30 seconds, annealing at 54°C (depending on the last annealing temperature of the first cycle), elongation

at 72°C for 30 seconds, and final elongation at 72°C for 10 minutes. The next step was electrophoresis, used to visualize a 200 bp PCR product in an agarose gel. O'Gene Ruler Low Range DNA Ladder (Cat. No. SM 1203, Thermo Fischer Scientific, USA) was used as a marker. Samples of the colony showing the expected band were inoculated into 5 mL LB-Ampicillin broth using a sterile loop, and then incubated at 37°C in a shaking incubator overnight.

Fresh LB-Ampicillin broth was prepared to transfer a small volume of bacteria to prepare plasmid stock. The remaining bacterial culture was used for plasmid isolation using Mini Kit plasmid isolation (Cat. No 12123, Qiagen, Germany), and was prepared for sequencing using reverse primer 5'- CAGTGACTGGAGTTCAGACG – 3' (Eurofins, Luxemburg). The results were analyzed using the BioEdit software.

Once the sequencing result matched the colony PCR, the bacterial stock previously prepared was then isolated following overnight incubation using Midiprep (Cat. No. 12143, Qiagen, Germany) or Maxiprep plasmid isolation (Cat. No. 12163, Qiagen, Germany) depending on the amount of stock needed for further experiments. These plasmid stocks were either used directly or stored at -20°C until further use.

2.4.3 sgRNA cloning into pAW12 and pAW13 vectors

pAW12.lentiguide.GFP (Addgene plasmid No. 104374) and pAW13.lentiguide. mCherry (Addgene plasmid No.104375) were used for the double transduction study. These plasmids contain different fluorescence markers, Green Fluorescence Protein (GFP) marker in pAW12 and mCherry marker in pAW13, which would be beneficial for positively transduced cell selection (Weintraub et al., 2017). The map of the plasmid was shown in Appendix 2. To clone sgRNA into pAW12 and pAW13 plasmids, a protocol was followed as described by Sanjana et al., 2014 and Shalem et al., 2014b, with modifications.

First, pAW12 and pAW13 were digested by Esp3I (FD0454, Fermentas, USA). This step was initiated by mixing 5 µg of plasmid, 3 µL Fast Digest Esp3I, 3 µL FastAP buffer, 1x Fast Digest Buffer, 1mM Dithiothreitol (DTT) (Cat. No. D9779-250MG, Sigma Aldrich, USA) in a total volume of 60 µL. Second, the reaction mixture was incubated at 37°C for 30 minutes. After the incubation, the mixture was then analyzed by electrophoresis as previously described in section 2.4.1. When visible, ~2 kilo base (kb) filler fragment and larger fragment showed indicating that the restriction was successful. A larger piece was then subjected to gel

purification, followed by a protocol described in section 2.4.1. The plasmids can then be used directly or stored at -20°C for long-term storage.

sgRNA phosphorylation was then performed using the protocol as explained in section 2.4.1. The list of sgRNA sequences was shown in Table 2.4. The 5'-CACCG-3' sequence was added at the beginning of the forward sequence and 5'-AAAC-sgRNA's sequence-C'3'.

Ligation and transformation of ligated plasmid plus sgRNA into E. coli was done using the protocol as described previously in section 2.4.1. Each sgRNA targeting *MYBL2*, *FOXM1*, or NTC were ligated with selected different plasmids as described in Table 2.4.

Table 2.4. sgRNA sequences.

Genes	sgRNA sequence	Plasmid
MYBL2	Forward 5'- CTGACCAGCAATGCCAGTAC - 3'	pAW12.lentiguide.GFP
	Reverse 5'- GTACTGGCATTGCTGGTCAG - 3'	
FOXM1	Forward 5'-CCTGGAGCAGCGACAGGTTA- 3'	pAW13.lentiguide.mCherry
	Reverse 5'-TAACCTGTCGCTGCTCCAGG- 3'	
no target control	Forward 5'-AACTTCTTAAGCCAACTGAC - 3'	pAW12.lentiguide.GFP and pAW13.lentiguide.mCherry
	Reverse 5'- GTCAGTTGGCTTAAGAAGTT - 3'	

2.4.4 Colony PCR and sequencing for pAW12 and pAW13 cloning validation

Colony PCR and sequencing were then performed to screen and validate which colonies carry the right insert. The protocol was followed as described previously in section 2.4.2. PCR was performed by using forward sequence of sgRNA as the forward primer and 5'-GCCAAAGTGGATCTCTGCTGT-3' sequence (Eurofins, Luxemburg) as reverse primer. The second validation is sequencing which was conducted by using the reverse primer. As described before, once the sequencing result matched with the PCR result, the plasmid isolation was performed as described in section 2.4.2. The plasmid stocks were either used directly or stored at -20°C until further use.

2.5 Lentivirus production

2.5.1 PEI

The PEI lentivirus production was performed as previously described (Pirona et al., 2021). 600,000 HEK-293T cells were seeded in 10 cm2 culture dishes in 10 mL, and then incubated overnight. A DNA solution consisting of 8 µg of CRISPR library, 4 µg of psPAX2 (Addgene

plasmid no. 12260), 4 µg of pMD2.G (Addgene plasmid no. 12259), and Opti-MEM (Cat. No. 31985062, Gibco, USA) in a total volume of 250 µL was prepared for transfection. For this method, the ratio between PEI: DNA was 3:1, hence 48 µL of 1 µg/ µL PEI solution (Cat. No. Polysciences, Germany) was added to 202 µL Opti-MEM. The DNA solution was added to the PEI solution, homogenized well using a pipette, and then incubated at room temperature for 20 minutes. The mixed solution was added dropwise to HEK-293T cells. The petri dish was gently moved during the addition to make sure that the solution was evenly distributed. The cells were then incubated overnight, and the medium was replaced with IMDM complete medium as previously described. 72 hours after transfection, the supernatant was collected and filtered through a 0.45 µm filter (Cat. No. 514-0063, VWR, USA). The filtered supernatant containing lentivirus was then aliquoted and stored at -80°C until further use.

2.5.2 Lipofectamine 3000 and Lipofectamine LTX

The lipofectamine 3000 and lipofectamine LTX lentivirus production was made as previously described (Pirona et al., 2021). $10x10^6$ cells were seeded in a 15 cm dish in an IMDM medium, and then incubated overnight. DNA solution and lipofectamine mixed solution were made. 17 µg of CRISPR library, 17 µg of psPAX2, and 5.7 µg of pMD2.G plasmids were mixed with 2100 µL Opti-MEM I reduced serum medium and 140 µL P3000 solution. Lipofectamine solution was made by mixing 40 µL lipofectamine 3000 (Cat. No. L3000015, Invitrogen, USA), 160 µL lipofectamine LTX (Cat. No. 15338030, Invitrogen, USA) and 2100 µL Opti-MEM I reduced serum medium. The DNA solution was then added to the lipofectamine solution and then homogenized using a pipette. The mixed solution was then incubated for 30 minutes at room temperature. At the end of the incubation time, the mixed solution was then added dropwise to the cells while gently swirling the petri dish to evenly distribute the solution. After overnight incubation, the medium was changed to a fresh medium, and then the incubation was continued for 48 hours. 72 hours after transfection, the supernatant was collected and filtered using a 0.45 µm filter. Filtered supernatant containing medium was then transferred into Vivaspin tube 30,000 Molecular weight cut-off (MWCO) (Cat. No. Z614629-12EA, Sartorius, Germany), and centrifuged at 1000g for 40 minutes to concentrate the virus approximately ten times. The filtered and concentrated supernatant containing lentivirus was then aliquoted and stored at -80°C until further use.

2.6 Establishment of Cas9 stable cells

Lentivirus containing lentiCas9-Blast (Addgene plasmid No. 52962), psPAX2 (Addgene plasmid No. 12260), and pMD2.G (Addgene plasmid No. 12259) were firstly prepared following lentivirus production as described previously in section 2.5.2. 60,000 cells were plated in a 6-well plate and then incubated overnight. Transduction was performed using 121 µL of the previously made lentivirus in DMEM medium supplemented with 1% Penicillin-Streptomycin containing 8 µg/mL polybrene (Cat. No. TR-1003-G, Merck Millipore, USA). The next day, the medium was refreshed with a complete DMEM medium. 72 hours after transduction, cells were treated with 6 µg/mL and 5 µg/mL blasticidin for S2-028 and S2-007, respectively, for 6 days. The blasticidin concentration was previously determined in section 2.3. After blasticidin selection, cells were cultured until they reached confluency, and then expanded and stored for further analysis.

2.7 Protein isolation and Western blot

Protein isolation was initiated by discarding the medium in the culture dishes. Cells were then washed with cold PBS and treated with 80 µL pre-prepared lysis buffer for 6-well plate petri dish. Lysis buffer was prepared by mixing RIPA buffer (Cat. No. 89900, Thermo Fisher Scientific, USA), 0.25 U/µL Benzonase® Nuclease, Purity > 99% (Cat. No. 70664-3, Merck Millipore, USA), and 1x Protease/phosphatase inhibitor single use (Cat. No. 5872S, Cell Signaling Technology, USA). Cells were then scraped using a plastic scraper and then collected in microcentrifuge tubes. Next, cells were agitated for 30 minutes at 4°C and then homogenized using a 20 Gauge syringe. The homogenized cells were centrifuged at 16000 g for 20 minutes at 4°C (Sigma 2K15 centrifuge, B Braun, Germany). The protein-containing supernatant was collected and transferred into a new microcentrifuge tube. Protein concentration was then measured following the manufacturer's protocol from Pierce BCA assay kit (Cat. No. 23227, Thermo Fischer Scientific, USA).

10% SDS page gel was prepared for Western blot by mixing 2 mL ddH2O, 1.5 mL Trenn buffer, 1.67 mL acrylamide, 50 µL SDS 10 %, 50 µL APS 10%, and 5 µL TEMED. The mixture was directly transferred into the gel chamber, and then incubated for 30 minutes or until the gel solidified. The stacking gel was prepared next by mixing 2.9 mL ddH2O, 1.2 mL Sammel buffer, 0.83 mL acrylamide, 50 µL SDS 10%, 50 µL APS 10%, five µL TEMED. The stacking gel mixed was then added on top of the previous solidified gel. Carefully, the gel comb was

then placed in the gel chamber. Thirty minutes of incubation was then added to make sure that the gel was solidified. All buffer was pre-made based on the recipe shown in Appendix 1.

Western blot was initiated by mixing 20 µg protein samples and pre-prepared sample buffer containing 2x Laemli Buffer (Cat. No. 1610737, Bio-Rad, USA) in 20 µL lysis buffer. The gel was placed in the Western blot apparatus (Mini Protean TetraCell, Biorad, USA) and immersed in 1x running buffer (Appendix 1). Samples were mixed, spun down, and incubated for 5 minutes at 95°C. 20 µL denatured protein samples were loaded into wells of the pre-prepared gel together with the protein ladder, Spectra multicolor broad range protein (Cat. No. 26634, Thermo Fischer Scientific, USA). The samples were run at 75 Voltage for 10 minutes or until they reached the bottom of the stacking gel, after which they were run at 135 Voltage for one hour or until the ladder has adequately separated.

Table 2.5. List of antibodies for Western blots

Primary antibody			
Name	**Dilution**	**Catalogue number**	**Manufacturer**
Cas 9 antibody	1:1000	61577	Active Motif, USA
Anti Thymidylate synthase (TS) antibody	1:750	ab108995	Abcam, UK
Anti MYB Proto-Oncogene Like 2 (MYBL2) protein clone LX015.1 antibody	1:750	MABE886	Merck Millipore, USA
Anti Menin antibody	1:500	ab92443	Abcam, UK
Anti FOXM1 protein antibody	1:1000	5436S	Cell Signaling Technology, USA
Anti GAPDH antibody	1:1000	CB1001-500UG	Merck Millipore, USA
Anti alpha-tubulin (α-tubulin) antibody	1:1000	ab4074	Abcam, UK
Secondary antibody			
Name	**Dilution**	**Catalogue number**	**Manufacturer**
2nd antibody mouse Horse anti-mouse Ig antibody (H+L), peroxidase	1:5000 - 1:10000	PI-2000-1	Vector Laboratories, USA
2nd antibody rabbit Goat anti-rabbit Ig Antibody (H+L), peroxidase	1:5000 – 1:10000	PI-1000-1	Vector Laboratories, USA

The gel was then transferred to a nitrocellulose membrane (Cat. No. 10600007, GE Life Science, USA) using a semi-dry method in the Transfer chamber (Trans-Blot Turbo Transfer System, BioRad, USA). After the transfer, the membrane was washed using washing buffer, and then 5% skim milk in 1x Tris-buffered saline-Tween 20 (TBST) blocking solution (Appendix 1) was used for the 1-hour blocking. The membrane was washed and incubated with the primary antibody overnight at 4°C. The next day, the membrane was incubated for 1 hour in the secondary antibody (Table 2.5) before being prepared for signal detection, using Immobilon Western horseradish peroxidase (HRP) substrate luminol reagent (Cat. No. WBKL S0500, Merck Millipore, USA) done based on the manufacturer's protocol. The images were then captured by placing the membrane inside the chemiluminescence imaging system (Intas ECL ChemoStar, Intas science imaging, Germany).

The Western blot work, especially for detection of FOXM1 and MYBL2 proteins after a double knockout experiment were done as part of the bachelor thesis of Lili Kalmar.

2.8. Transduction and flow cytometry to count transduction efficiency

For each lentivirus that will be used for screening and validation, the transduction efficiency was calculated to examine the quality of the produced virus and then predict how much virus that will be used for the transduction.

70,000 cells were plated in 12-well plates, and then incubated overnight. The next day, cells were transduced with varying volumes of lentivirus, ranging from 20 µL to 0.5 µL, in 500 µL DMEM medium supplemented with 10% FBS containing 8 µg/mL polybrene. Cells were then incubated overnight before the medium was changed with a complete DMEM medium the following day. 96 hours after transduction, the cells were harvested, transferred to microcentrifuge tubes, and centrifuged at 1400 rpm for 4 minutes (Megafuge 1.0R Heraeus Sepatech, Germany). Cells were washed with PBS once, and then the supernatant was discarded. 1 mL PBS was added before cells were filtered. 5000 cells were then analyzed using Luminex Guava EasyCyte HT (Luminex Corporation, USA).

Positively transduced cells or mCherry positive cells were then used to count the virus titer using formula: $T= (P*N)/(D*V)$, where T is the virus titer (transduction unit/mL or TU/mL), P is the positive-transduced cells (mCherry positive cells), N is the number of cells, D is the virus dilution factor, and V is the final volume of transduction (mL). Once the virus titer was calculated, Multiplicity of Infection (MOI) can be determined. MOI represents the number of

viruses that are added per cell during infection. This formula can calculate the MOI for every volume of lentivirus: MOI = (T*V)/N, where T is the virus titer (TU/mL), V is the final volume of transduction (mL), and N is the number of cells. The information about MOI and percentage of mCherry positive cells were then used to assemble a MOI chart, which later can be used to calculate the number of cells to be used at the starting point of the CRISPR screening according to the MOI number.

2.9 Flow cytometry to detect EPCAM expression

Flow cytometry to detect EPCAM expression was performed to analyze the Cas9 activity by calculating the difference of EPCAM expression between wild type cells and Cas9 cells (Figure 2.2).

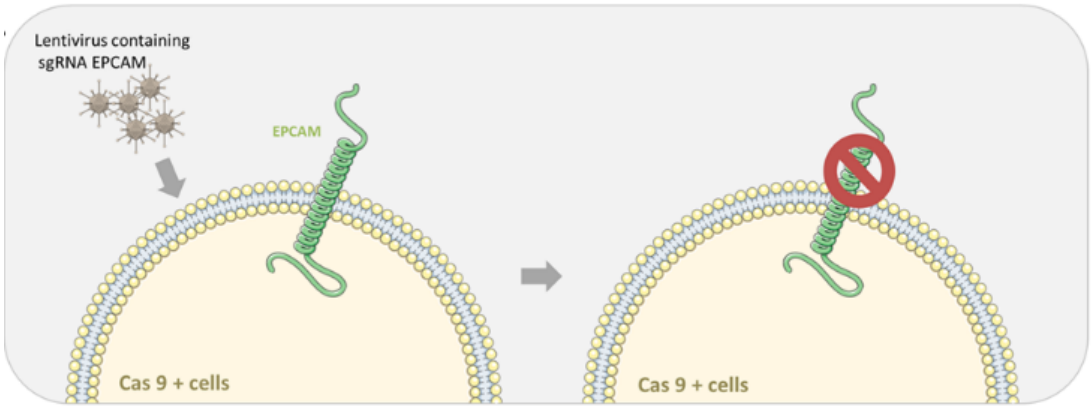


Figure 2.2. Strategy to examine Cas9 activity in Cas9-positive cells. Lentivirus contains EPCAM protein-targeting sgRNA was transduced in the Cas9-positive cells. The EPCAM protein expression was measured by flow cytometry. When the cas9 in the cas9-positive cells was active, the EPCAM expression was lower than the wild-type.

70,000 cells were plated in a 12-well plate and then incubated overnight. The transduction efficiency of previously produced lentivirus containing sgRNA targeting *EPCAM* was examined using the protocol described in section 2.8. Next, the transduction S2-028 and S2-007 cells were performed by adding 1.5 µL and 1.05 µL of the lentivirus, respectively. Transduction was conducted in 400 µL DMEM medium containing 8 µg/mL polybrene. One day after transduction, the medium was refreshed with a complete DMEM medium. Cells were split as it was necessary to maintain cell viability.

Ten days after transduction, cells were harvested and calculated. $1x10^6$ cells were then plated into a V-shaped microplate well and then centrifuged for 2 minutes, at 250g and 4°C. The supernatant was discarded by flipping the plate as quickly as possible. 20 µL of anti- EPCAM

Fluorescein isothiocyanate (FITC) antibody (Cat. No. 347197, BD Biosciences, USA) or FITC mouse IgG Isotype control (Cat. No. 555748, BD Biosciences, USA) antibody was mixed with 30 µL PBS and added into the cell pellet. Cells were resuspended with the antibody dilution, and then incubated for 30 minutes at 4°C. The plate was centrifuged at 250 g for 2 minutes at 4°C (Eppendorf Centrifuge 58120 R, Germany), the supernatant was discarded, and then the pellet was washed twice with cold PBS for 2 minutes, at 250g, and 4°C. After the second washing step, cells were resuspended in 500 µL PBS containing 5.4 µg/mL Actinomycin D, 7-Amino (Cat. No. 129935, Sigma Aldrich, USA) as a live/dead marker. Cell suspension was then transferred to a Fluorescence-activated cell sorting (FACS) cell filter (Cat. No. 10585801, Corning, USA) for flow cytometry analysis, where 10,000 cells were analyzed using Luminex Guava EasyCyte HT (Luminex Corporation, USA).

2.10 Cell counting to determine doubling time

Cells with $1x10^5$ density were plated in 6 well plates (n=2). Nine plates were prepared each for counting day. One plate was harvested per day, stained with Trypan blue (Cat. No. 15250061, Gibco, USA) as the live and dead marker, and counted using a hemacytometer. The cell counting was done according to the manufacturer's protocol. The formula then was used to measure the cell concentration (cells/mL), which was cell number after counting multiplied by 10^4 and the dilution factor. The final number of cells was the average between two cell numbers from two wells.

Once cell numbers were counted, a growth curve was made by plotting incubation time on the x-axis and cell counts on the y-axis. By observing the curve, the exponential phase was then determined. The cell counts of the exponential phase were then used to calculate the doubling time following this equation: DT = Tln2/ln (Xe/Xb), where DT is doubling time, T is the incubation time (hours), Xb is the cell number at the beginning of the exponential phase, and Xe is the cell number at the end of the exponential phase.

2.11 CRISPR screening

Figure 2.3 shows the screening preparation, which consists of the establishment of Cas9-stable cells and preparing the library lentivirus. Most of the steps on establishing Cas9-stable cells were described in section 2.6. CRISPR screening was performed using a library containing 267,109 sgRNAs (12 sgRNAs/gene) as previously published (Boettcher et al., 2018). A total

of 24 days were needed for screening, from transduction to 18 days of the screening incubation and then sample collection. MOI 0.3 and 100x coverage were used at the beginning of the transduction (Figure 2.4). The samples collected at t_{zero} and t_{end} were subjected to DNA isolation and later on, for Next Generation Sequencing (NGS). The detailed methods of each step will be explained in the following subsections.

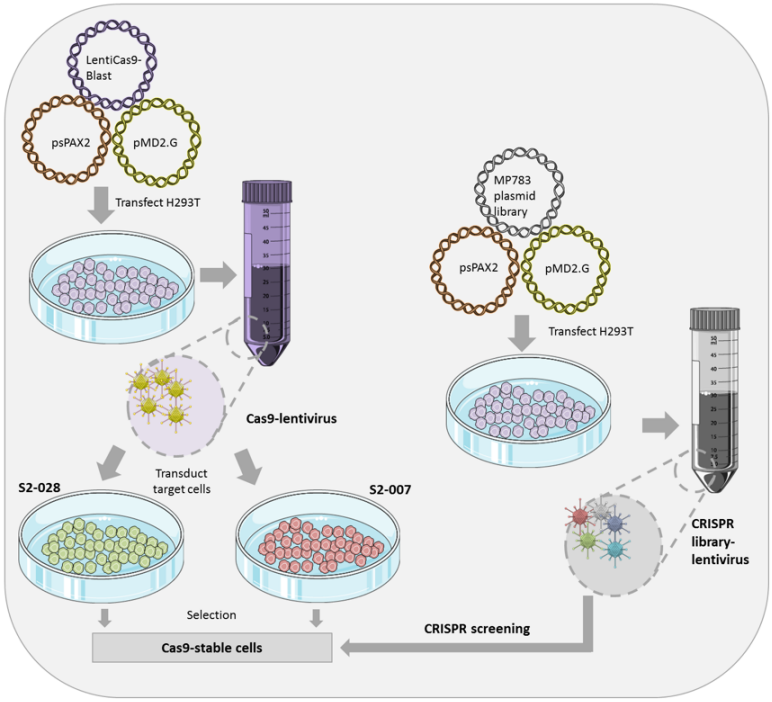


Figure 2.3. An overview of the CRISPR screening preparation. On the left side, an overview of the establishment of Cas9-stable cells is shown. Plasmids LentiCas9-Blast, psPAX2, and pMD2.G were used to produce lentivirus, which was later used to establish Cas9-stable cells. On the right side, an overview of the lentivirus production for CRISPR screening is shown. The MP783 plasmid library plasmid psPAX2 and plasmid pMD2.G were used for that purpose.

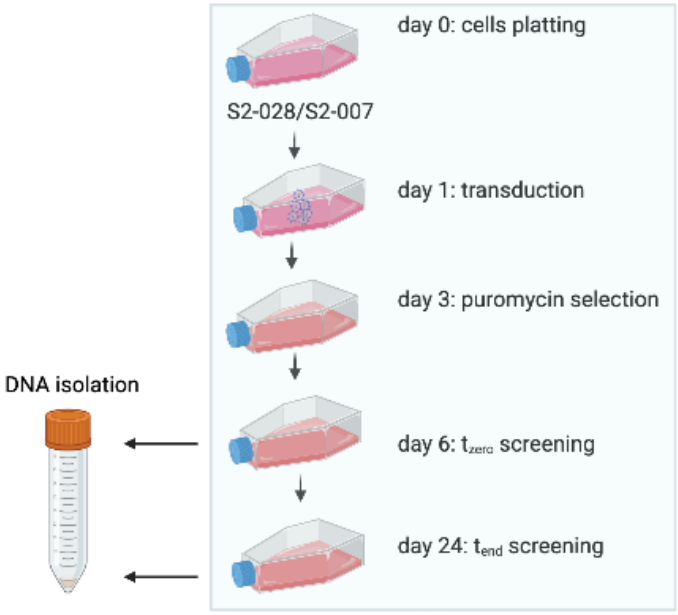


Figure 2.4. Overview of the CRISPR screening protocol. The timeline for CRISPR screening: starting from day 0 with cell plating, followed by transduction, puromycin selection, and two-time points of DNA collection.

2.11.1 CRISPR library expansion and Lentivirus production

CRISPR library preparation was done prior to lentivirus production to ensure sufficient plasmid viability during lentivirus production. Firstly, one tube of ElectroMAX™ DH5α-E Competent Cells (Cat. No. 11319019, Invitrogen, USA) was mixed with 100 ng CRISPR library by tapping the tube gently (Boettcher et al., 2018). Secondly, a cuvette from the electroporator (BioRad Gene Pulser II, USA) was prepared and place onto the ice. The mix was carefully pipetted into the cuvette. The electroporator was set and run with 2kV, 200 Ω, and 25 µF. Cells were moved into a tube containing 5 mL pre-warmed SOC media and then incubated for one hour in a shaker at 37°C. After overnight incubation, cells were transferred to LB-broth medium, and then subjected to plasmid isolation using Maxiprep (Cat. No. 12163, Qiagen, Germany). Several cell dilutions (1:100 and $1:10^6$) were prepared in parallel. From each dilution, 100 µL was taken and plated onto the LB-Ampicillin agar plate, which were then incubated overnight at 37°C. Next, colony counting, following the manufacturer's protocol, was performed to determine the transformation efficiency, which was calculated to ensure an even representation of all library sgRNA sequences (70x clones per sgRNA) (Boettcher et al., 2018).

Lentivirus production was then prepared using the protocol described previously in section 2.5.1, using Lipofectamine 3000 and LTX. The transduction efficiency of the produced lentivirus was then calculated by using protocol in section 2.8.

2.11.2 Screening transduction, puromycin selection and cells maintenance

CRISPR screening was initiated by platting the cells in a 15 cm petri dish. $3x10^6$ cells were plated in 50 petri dishes, and then incubated overnight. Cells in 34 plates were transduced with MOI 0.3 in DMEM medium supplemented with 10% FBS containing 8 µg/mL polybrene. Two additional plates were then used as a negative control without virus transduction. Overnight incubation was done, and the next day, the plates were replenished with fresh complete DMEM medium supplemented with 10% FBS and 1% Penicillin-Streptomycin. Forty-eight hours after transduction, cells were then subjected to puromycin selection. Cells were harvested and plated in a T175 flask, to prevent the cells from being over confluent and losing the screening coverage. From 34 plates, cells were transferred to 170 T175 flasks and treated with 5 µg/mL and 4 µg/mL puromycin concentrations for S2-007 and S2-028. Cells were incubated for 72 hours in the puromycin selection, after which they were harvested, pooled, and counted. To preserve the coverage, cells were then maintained by culturing in 1000x coverage. $300x10^6$

cells were then platted into 150 T175 flasks, with each flask contained $2x10^6$ cell density. Cell pellets with the same number of cells were collected at time point zero (t_{zero}) and was washed and stored at -20°C for further analysis. Screening cells were maintained for 18 days. Every three days, the medium in each flask was refreshed. Cells were harvested, pooled, and plated in a new T175 flask every time the cells reached confluency (± 6 days) in the same number of cells ($300x10^6$ cells) to maintain screening coverage. At the end of the screening (day 18), cells were then harvested, pooled, and counted once again. Finally, $300x10^6$ cells were collected for DNA isolation as the last time point (t_{end}).

2.11.3 DNA isolation

Collected DNA samples were subjected to DNA isolation. Cell pellets were resuspended in 20 mL PI buffer (Cat. No. 19051, Qiagen, Germany), 100 µg/mL RNAse (Cat. No. 19101, Qiagen, Germany), and 0.5% SDS, and then incubated for 30 minutes at 37°C. 100 µg/mL Proteinase K (Cat. No. AM2546, Thermo Fischer Scientific, USA) was pipetted in before the pellet was subjected to another incubation at 55°C for 30 minutes. Homogenization was then performed to the cell pellets using 18 Gauge and 22 Gauge needles, three times each. The homogenized cell pellets were mixed with 20 mL Phenol: Chloroform: Isoamyl Alcohol (PCI) (Cat. No. 15593049, Thermo Fischer Scientific, USA), and then 50 mL mixture was transferred to Max tract tubes (Cat. No. 129056, Qiagen, Germany) and centrifuged for 5 minutes, at 1500 g. The aqueous phase was moved into a new falcon and mixed with 2 mL 3M sodium acetate (Cat. No. R1181, Thermo Fischer Scientific, USA) and 16 mL isopropanol before the step continued to centrifugation at 15,000 g, for 15 minutes at room temperature. The pellet containing the DNA was washed with 10 mL 70% ethanol and centrifuged as done in the previous step. The cell pellet was dried at 37°C and resuspended in nuclease-free water. The DNA concentration was then measured using NanoDrop Spectrophotometer ND-1000 (Thermo Fischer Scientific, USA), adjusted to 1 µg/µL, and then stored at -20°C until further use.

2.11.4 PCR recovery of sgRNA sequences

The DNA pellet was subjected to fragmentation by boiling at the 95°C. Then, the fragmentation was run alongside with Track 1kb Plus DNA ladder (Cat. No. A10264, Invitrogen, USA) via 1% agarose electrophoresis at 100 V.

Two PCRs were performed for NGS preparation to amplify sgRNA from the DNA, which had already been isolated. The first PCR was done by mixing 5 µg shared DNA, 25 µL NEB Next Ultra II Q5 master mix (Cat. No. M0544L, NEB, UK), and 0.5 µM forward (5'-GGCTTGGATTTCTATAACTTCGTATAGCA-3') and reverse (5'- CGGGGACTGTGGGC GATGTG-3') primers in a 50 µL reaction. The PCR program was as follows: 1x (98°C - 30 seconds), 16x (98°C - 30 seconds, 65°C - 75 seconds), 1x (65°C - 5 minutes). All first PCRs were then pooled. For the second PCR, 50 µL total volume was mixed, which consisted of 25 µL NEB Next Ultra II Q5 master mix, 0.5 µM forward (5'-AATGATACGGCGACC ACCGAGATCCACAAAAGGAAACTCACCCTAAC-3') and reverse (5'-CAAGCAGAA GACGGCATACGAGAT-(N)6-GTGACTGGAGTTCAGACGTG-3') primers, where (N)6 is a 6 nucleotides index for sequencing on the Illumina HiSeq platform (Table 2.6), and 2 µL pooled first PCR. The 344 base pair amplicon was extracted from 1.75% agarose using NucleoSpin Gel and PCR Clean. The extracted amplicon was mixed with Qubit dsDNA HS Assay Kit (Cat. No. Q32851, Invitrogen, USA) following the manufacturer's protocol, and then measured using Qubit 3 fluorometer (Thermo Fischer Scientific, USA). Next, the extracted amplicon was subjected to sequencing using Hiseq 2000 v4 single-read 50 bp with the sequencing primer: 5'-GAGACTATAAGTATCCCTTGGAGAACCACCTTGTTGG-3'.

Table 2.6. Index's list for Illumina HiSeq platform.

Index sequence	samples
ATCCTC	S2-007 t_{zero}
ATCACG	S2-007 t_{end}
ACTTGA	S2-028 t_{zero}
TAGCTT	S2-028 t_{end}
GGCTAC	CRISPR library

2.11.5 Data analysis and candidate selection

Total read counts of sgRNA sequences from each sample were analyzed using MAGeCK (Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout) analysis. The sgRNA counts from t_{zero} and t_{end} from each cell line were further analyzed. The gene essentiality score is represented as a beta-score. The MAGeCK algorithm was allowed to examine the cell-type-specific genes identified by two conditions such as S2-007 vs. S2-028. Candidate selection was determined based on the cell-type-specific genes. The cell-type-specific genes are genes considered essential in one condition, but not the other. Comparisons of beta-score from each

cell was performed by showing graph comparison analyzed by R software. The cell-type-specific genes were shown to be located close to 0 on the x-axis or y-axis (Fig. 2.5).

The Gene Ontology (GO) annotation of negatively or positively selected significant genes ($p<0.05$) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway study was performed using The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 (https://david.ncifcrf.gov/). Network gene analysis was performed using Ingenuity Pathway Analysis (IPA) (Qiagen, Germany). Pathview in R software version 1.4.1106 was used to plot the affected genes on the KEGG pathway.

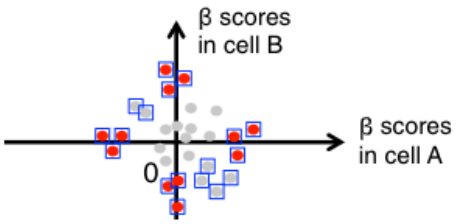


Figure 2.5. An illustration of differentially expressed genes identified by two different conditions. The blue squares represent the differentially expressed genes in a comparison of two conditions. Red dots show the cell type-specific genes (Figure copied from Li et al., 2015).

2.12 Cell proliferation using resazurin assay

190,000 cells were plated in a 6-well plate and then incubated overnight. Cells were transduced with lentivirus containing targeted sgRNA candidate. Each candidate has five sgRNAs that target different locations across the gene. Two no target control (NTC) and one empty plasmid were prepared as well, as controls. 5.5 µL and 3.6 µL previously prepared lentivirus were used for transduction in S2-028 and S2-007 cells, respectively. These volumes were used to get ~30% positive transduced cells. In a total volume of 1120 µL, virus was mixed with DMEM medium containing 8 µg/mL polybrene. The next day, the medium was then refreshed with DMEM complete medium and incubated overnight.

Two days after the incubation, the antibiotic selection was performed similarly to the screening selection protocol as previously described in section 2.11.2. After antibiotic selection, 2000 cells were plated in 7 of 96-well plates with eight replications. One plate was used to measure cell viability per day. The resazurin protocol was similarly performed as previously described in section 2.3.

2.13 Cell viability assay using flow cytometry

The cell viability measurement using flow cytometry was done to quantify mCherry positive cells (positively transduced cells) over time. Transduced and selected cells were harvested, and then counted. Cells were mixed with the non-transduced cells or wild type in a ratio of 4:1 (Fig. 2.6). Cells were then plated in 6-well plate in triplicate with a density of $0.024x10^6$ cells per well, followed by incubation for 7 days. Every 7 days, cells were harvested and subjected to flow cytometry measurement. $1x10^6$ cells were transferred to a microcentrifuge tube, and then centrifuged at 1400 rpm for 4 minutes. Pelleted cells were washed with PBS and filtered. 10,000 cells were then analyzed using Luminex Guava EasyCyte HT (Luminex Corporation, USA). For the following quantification, the same cell density was plated in a new 6-well plate. Three time-points of flow cytometry measurement (t_{zero}, t7, t14) were performed for *TYMS*, while four time-points of measurement (t_{zero}, t7, t14, t21) was done for *MEN1* and *MYBL2*.

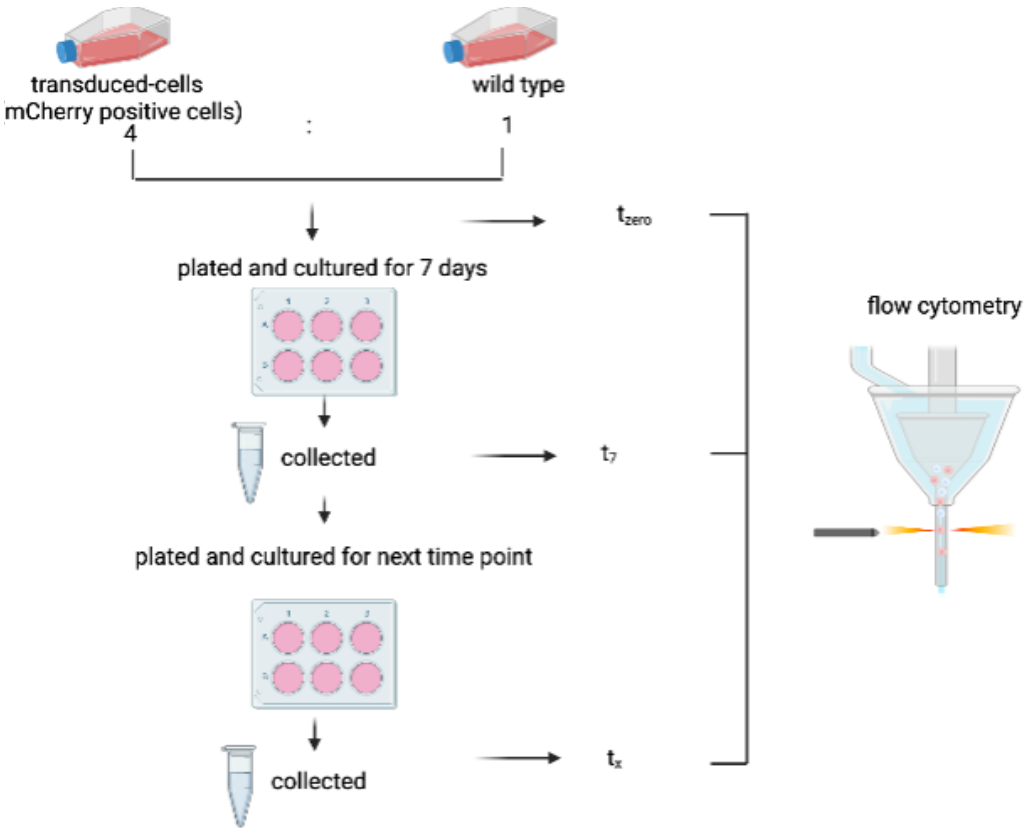


Figure 2.6. An illustration of cell viability assay using flow cytometry. Transduced cells (mCherry positive cells) were mixed with wild type in a 4:1 ratio. The cells were then plated and cultured for 7 days and continued until tx. Some of the cells were analyzed to count for mCherry positive cells at t_{zero}, t_7, or t_x.

2.14 Soft agar assay

A soft agar assay was performed to observe the ability of cells to grow in an anchorage-independent manner. 2% SeaKem GTG Agarose (Cat. No. 733-1541, Lonza, Switzerland) was previously prepared in PBS and autoclaved. For base agar, 0.5% agar that was made from the 2% agar stock and diluted in PBS was pre-warmed, and 2 mL immediately transferred to a 6-well plate. The plate was then incubated at room temperature for 30 minutes or until the agar has solidified. The top agar (0.05% agar) was prepared by mixing 3.2 mL pre-warmed culture medium, 0.4 mL FBS and 0.4 mL 2% agar. 0.05% agar was warmed in a 37°C incubator. Transduced and selected cells were harvested and then counted. 1mL of 0.05% agar was mixed with the 2500 cells and immediately transferred onto the base agar. Each sample was plated in triplicates. Plates were incubated at room temperature for 30 minutes or until the top agar was solidified, and then moved to the 37°C incubator for 30 minutes. An additional 20 minutes of incubation at room temperature was done before 1 mL complete DMEM medium was pipetted into the well. Cells were then incubated for 15 days until a visible colony appeared.

After 15 days of incubation, the remaining medium was discarded, and 0.5 mL crystal violet 0.005% (Cat. No. V5265-250mL, Sigma Aldrich, USA) was added, before 1-hour incubation at room temperature. Cells were then washed with PBS, and the colony was counted using the Open CFU 3.8 software.

2.15 Wound healing assay

A wound-healing assay was performed to observe the ability of the cells to migrate, and then close the gap. This assay was performed following the manufacturer's protocol from the Ibidi µ-Dish 35 mm, high (Cat. No. 81156, Ibidi, Germany). Previously, cells were transduced and puromycin selected. Next, 35,000 cells in 70 µL DMEM complete medium were plated into the chamber in triplicates and incubated overnight. The next day, the chamber was then lifted off, and the cells were washed with PBS to remove any remaining medium. DMEM medium supplemented with 1% Penicillin-Streptomycin was then added to the well. Several images of the cell gaps were taken at the t_{zero} and then after 3 days of incubation. The observation and imaging were carried out using an inverted microscope (Inverted Zeiss Cell Observer.Z.1, Germany). The gap between the two cell-populations was then calculated using the ImageJ software.

2.16 Invasion assay

An invasion assay was done to study the ability of the cells to migrate through extracellular matrix such as matrigel. This invasion assay was performed following the manufacturer's protocol from Corning® BioCoat™ Control Inserts with 8.0 µm Polyester (PET) Membrane (Cat. No. 354578, Corning, USA). Cells were transduced and puromycin selected before changing the culture medium to DMEM medium supplemented with 1% Penicillin-Streptomycin, and then incubated. A 10,000 cells density in 100 µL DMEM medium without any serum supplement was then seeded into the insert, and 650 µL complete DMEM medium was added into the reservoir triplicate. The plate was then incubated for 18 hours at 37°C. After incubation, the medium was removed from the insert, and the remaining cells on the top of the insert were cleaned and washed with PBS. Cells below the insert were fixed with 70% ethanol for 15 minutes at room temperature to restore the cells, while the insert was left to air-dry. Fixed cells were subjected to staining by 0.5% crystal violet for 30 minutes. The excess dye was washed off with ultrapure water. The insert was then dried overnight, and stored until observation, which was performed using an inverted microscope, with images taken for colony observation. Colony counting was done using the Open CFU 3.8 software.

2.17 Double transduction and sorting

3,500,000 S2-028 and S2-007 cells each were plated in a T175 flask and then incubated overnight. Based on the protocol described in section 2.8, the transduction efficiency of previously prepared lentiviruses was examined. The next day after cells plating, double transduction was conducted by following the protocol in Table 2.7. These volumes were estimated to get ~75% positively transduced cells. The next day after transduction, the medium was refreshed with DMEM complete medium and then incubated overnight. Two days after incubation, the cells were harvested and then prepared for cell sorting.

Table 2.7. Double transduction protocol.

S2-028		**S2-007**	
Group 1	***FOXM1-MYBL2* double transduction**	**Group 1**	***FOXM1-MYBL2* double transduction**
Transduction component	**Volume (uL)**	**Transduction component**	**Volume (uL)**
DMEM containing 8 µg/mL polybrene	20,000	DMEM containing 8 µg/mL polybrene	20,000

DMEM medium	4806.7	DMEM medium	4801.51
FOXM1-pAW13.lentiguide.mCherry lentivirus	107.26	*FOXM1*-pAW13.lentiguide.mCherry lentivirus	92.73
MYBL2-pAW12.lentiguide.GFP lentivirus	86.04	*MYBL2*-pAW12.lentiguide.GFP lentivirus	105.76
Total volume	**25,000**	**Total volume**	**25,000**

Group 2	***FOXM1*-NTC double transduction**	**Group 2**	***FOXM1*-NTC double transduction**
Transduction component	**Volume (uL)**	**Transduction component**	**Volume (uL)**
DMEM containing 8 µg/mL polybrene	20,000	DMEM containing 8 µg/mL polybrene	20,000
DMEM medium	4752.47	DMEM medium	4794.76
FOXM1-pAW13.lentiguide.mCherry lentivirus	86.04	*FOXM1*-pAW13.lentiguide.mCherry lentivirus	92.73
NTC-pAW12.lentiguide.GFP lentivirus	161.49	NTC-pAW12.lentiguide.GFP lentivirus	112.51
Total volume	**25,000**	**Total volume**	**25,000**

Group 3	***MYBL2*-NTC double transduction**	**Group 3**	***MYBL2*-NTC double transduction**
Transduction component	**Volume (uL)**	**Transduction component**	**Volume (uL)**
DMEM containing 8 µg/mL polybrene	20,000	DMEM containing 8 µg/mL polybrene	20,000
DMEM medium	4767.38	DMEM medium	4800,53
MYBL2-pAW12.lentiguide.GFP lentivirus	107.26	*MYBL2*-pAW12.lentiguide.GFP lentivirus	105,76
NTC-pAW13.lentiguide.mCherry lentivirus	125.36	NTC-pAW13.lentiguide.mCherry lentivirus	93,70
Total volume	**25,000**	**Total volume**	**25,000**

Group 4	**NTC-NTC double transduction**	**Group 4**	**NTC-NTC double transduction**
Transduction component	**Volume (uL)**	**Transduction component**	**Volume (uL)**
DMEM containing 8 µg/mL polybrene	20,000	DMEM containing 8 µg/mL polybrene	20,000
DMEM medium	4713.15	DMEM medium	4793.79
NTC-pAW12.lentiguide.GFP lentivirus	161.49	NTC-pAW12.lentiguide.GFP lentivirus	112.51
NTC-pAW13.lentiguide.mCherry lentivirus	125.36	NTC-pAW13.lentiguide.mCherry lentivirus	93.70
Total volume	**25,000**	**Total volume**	**25,000**

Cells were diluted in 1 mL BD FACS Pre-Sort Buffer (Cat. No. 563503, BD Biosciences, USA) containing 5.4 µg/mL Actinomycin D, 7-Amino and then filtered before sorted using FACS Aria I (BD Biosciences, USA). Double transduction-cells population was shown in Figure 2.7. The double transduction population was located on quadrant 2 (Q2) where the GFP expression on the y-axis and mCherry expression on the x-axis were positive. The number of S2-028 cells collected after sorting in every group was as follows 0.564 million, 0.939 million, 0.866 million, and 1.16 million for groups 1, 2, 3, and 4, respectively. As in the S2-028 cell, the result from S2-007 cell sorting was also not different in every group 1.5 million, 1.2 million, 1.3 million, 1.1 million for groups 1, 2, 3, and 4 in S2-007, respectively.

After sorting, cells were collected in DMEM complete medium and then subjected to washing with PBS, and then cells were plated and incubated for 3 days. Cells were then harvested after 3 days of incubation and ready to be used for further studies.

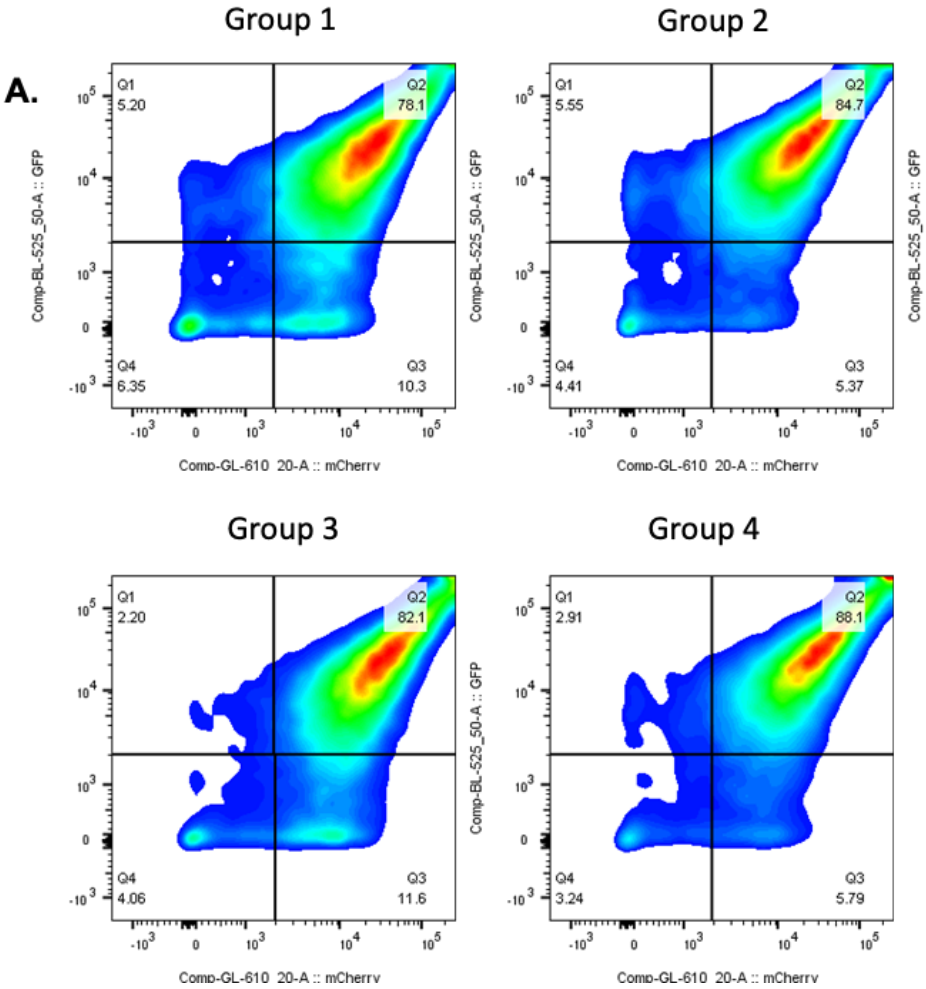


Figure 2.7. Double transduction-cells population. For details, see next page.

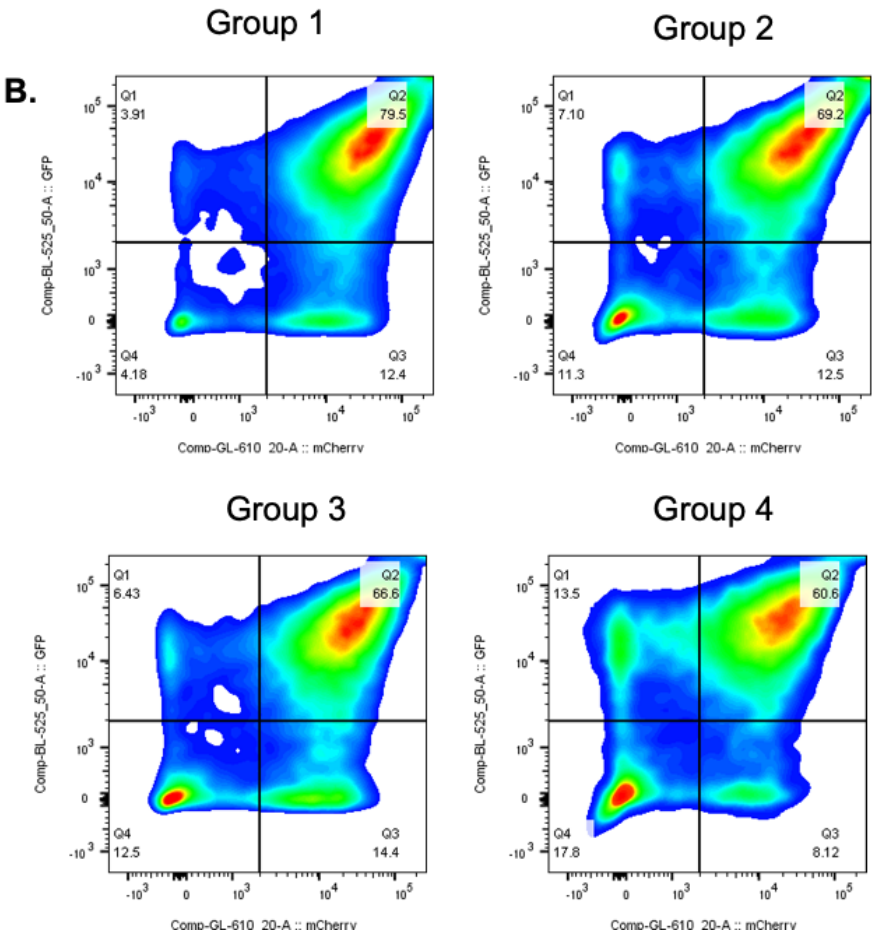


Figure 2.7. Double transduction-cells population. Cells were analyzed using FACS Aria I. mCherry expression level was shown on the x-axis, and the GFP expression level was shown on the y-axis. **(A, B)** Each chart was divided into four quadrants based on the expression level of each fluorescence for each group (group 1, 2, 3, and 4) in S2-028 (A) and S2-007 (B) cells. Group 1 = *FOXM1-MYBL2* transductions, Group2 = *FOXM1* transduction, Group 3 = *MYBL2* transduction, and Group 4 = NTC.

2.18 Cell viability assay using flow cytometry for double transduction

The cell viability measurement using flow cytometry was done to quantify GFP and mCherry positive cells (positively double transduced cells) over time. The cell viability was performed as described previously in section 2.13 except for the additional GFP expression level and the combination ratio between the positively transduced and wild-type cells. Instead of using 4:1, a 1:1 ratio was used due to the limited cells number. 10,000 cells were then analyzed using Luminex Guava EasyCyte HT (Luminex Corporation, USA). Four time-points measurements (t_{zero}, t_7, t_{14}, and t21) were performed for each group (1, 2, 3, and 4) in S2-028 and S2-007 cells.

2.19 Cell cycle assay

Three days after sorting, cells were harvested and then plated in a 6-well plate for cell cycle assay. After 3 days incubation, cells were trypsinized, collected in a 15 mL falcon tube, and centrifuged at 1000 rpm for 5 minutes. Next, PBS was used to wash the cells two times by centrifugation at 1000 rpm for 5 minutes. PBS was then discarded. Pre-chilled 70% ethanol was then used to fix the cells for one hour or more (cells could be stored for weeks in 4°C). The fixed cells were transferred into a 96-well "V" bottom plate, at 150 µL/well. The plate was then centrifuged at 900g for 5 minutes in 4°C. Ethanol was quickly discarded, and then cells were washed two times using PBS. Cells were centrifuged at 850g for 5 minutes at 4°C. 150 µL PBS was used to resuspend the cells and then 150 µL Propidium Iodide (PI) (Cat. No. P4864-10ML, Sigma Aldrich, USA) and 5 µL of a 1 µg/ µL RNAse (Cat. No. 8003088, Thermo Fischer, USA) were added. The unstained sample was prepared by not mixing PI in the cell dilution. The total volume for each sample was 300 µL. Cells were then analyzed using BD FACSCanto (BD Biosciences, USA).

2.20 Data analysis and statistics

Flow cytometry data was analyzed using the FlowJo software (BD Biosciences, USA). Statistical analysis was performed using GraphPad Prism 8 software. Data normality was tested before the t-test was performed to verify that the data follows the normal distribution. Unpaired t-test was used if the data followed a normal distribution, while the Mann-Whitney test was used if otherwise. On the graphs, the p value of each test was shown as an asterisk (*). The meaning of the asterisk are as follows: * = p value ≤ 0.05, ** = p value ≤ 0.01, *** = p value ≤ 0.001, **** = p value ≤ 0.0001.

3. RESULTS

3.1 S2-007 and S2-028 as cell samples for screening and validation

In this study, the CRISPR-Cas9 screening was used to identify genes that can eliminate metastatic pancreatic cancer. Two cell lines with different metastatic characteristics were used, namely S2-028 and S2-007. These two cell lines were derived from one parental cell line, SUIT-2, liver metastatic pancreatic cancer. Based on the previous studies, S2-028 has less metastatic potential compared to S2-007 (Taniguchi et al., 1992, 1994).

Cas9 stable cells need to be established before the screening. Therefore, lentivirus containing lentiCas9-Blast plasmid was initially produced in the preparation process. Since the plasmid contains the blasticidin resistance gene, blasticidin was used to select the positively transduced cells. Blasticidin kill curve was then used to determine the proper antibiotic concentration for selection. Based on the blasticidin kill curve and cells observed under the microscope, 6 µg/mL and 5 µg/mL blasticidin were used to select S2-028 and S2-007 cells (Fig. 3.1). After the cells were transduced with the lentivirus containing lentiCas9-Blast plasmid, cells were then subjected to the blasticidin selection for six days.

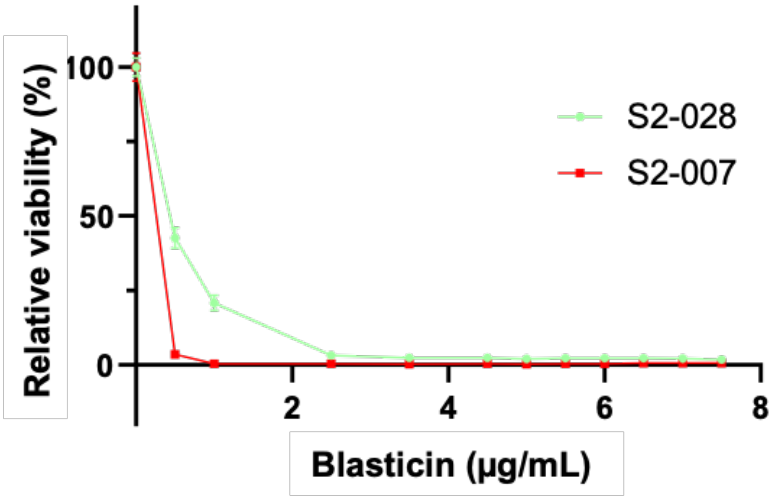


Figure 3.1. Blasticidin kill curve. Cells were cultured in 96 well plates and incubated overnight. The following day cells were treated with various concentrations of blasticidin ranging from 0 to 7.5 µg/mL for six days. After the incubation period, the cells were subjected to viability assay by adding resazurin. The fluorescence, which appeared after one hour of incubation, was then measured using a microplate reader. Percentage of relative viability ratio was then calculated by dividing the fluorescence amount in the treated samples to the control (0 µg/mL of blasticidin). A graph was then made by plotting the relative viability on the y-axis and blasticidin concentration on the x-axis (n=3). The green line represents S2-028 relative viability, and the red line represents S2-007.

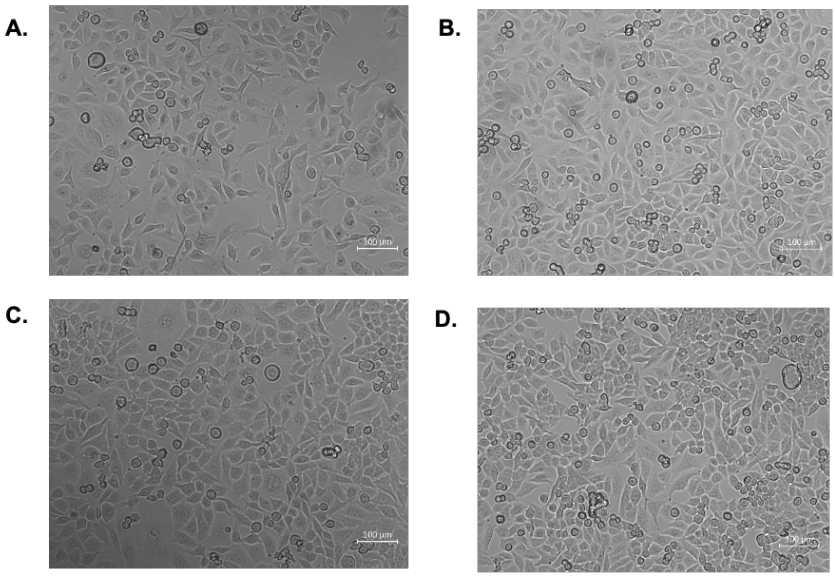


Figure 3.2. Similarity in the morphology of wild type and Cas9 cells in S2-028 and S2-007. Cells were cultured and then observed using an inverted microscope with 100x magnification. Scale bars show unit of 100 µm. **(A, B)** The pictures show S2-028 cell line wild-type (A) and Cas9 positive cells (B). **(C, D)** The pictures show S2-007 cell line wild-type (C) and Cas9 positive cells (D).

After blasticidin selection, cells were cultured. Positively selected cells were then observed to determine any morphology alteration by monitoring the cells under the microscope (Fig. 3.2). According to the observation, there was no difference in the morphology of the cells between wild-type and Cas9 cells in S2-028 and S2-007. Cells then were further analyzed to determine the Cas9 protein expression level by Western blot. For this purpose, the protein was first isolated from the cells, quantified, and then subjected to Western blot. According to the result, Cas9 protein was expressed in Cas9 cells in S2-028 and S2-007 (Fig. 3.3).

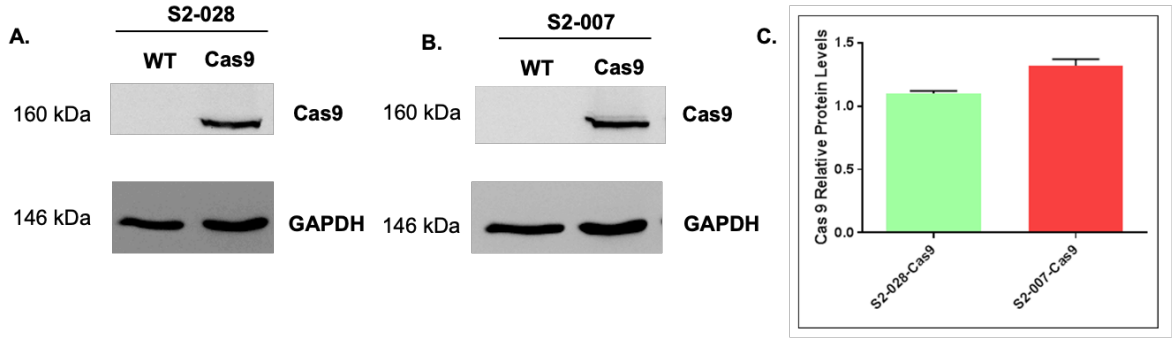


Figure 3.3. S2-028 and S2-007 express Cas9 protein. (A, B) Western blot results of Cas9 protein (160 kiloDalton (kDa)) and GAPDH (146 kDa) as a control protein in S2-028 (A) and S2-007 (B). **(C)** Cas 9 relative protein levels were plotted on the y-axis and sample names were plotted on the x-axis. Quantification of Cas9 protein in S2-028 and S2-007 was a ratio calculated relative to GAPDH expression. The quantification levels were plotted in the bar graph (n=2).

The functionality of the Cas9 protein was then examined by transducing the cells with the lentivirus containing sgRNA, which targets *EPCAM*. EPCAM protein is a surface protein expressed in both cells. Therefore, the expression of EPCAM was used as an indicator of Cas9 functionality because the expression level of EPCAM will be depleted when the Cas9 has its function and successfully caused perturbation on *EPCAM*. Observation of the expression level was then done by flow cytometry. Based on the expression level of EPCAM in Cas9 positive cells and wild-type cells, the cas9 activity was then calculated (Fig. 3.4). The graphs from the flow cytometry showed that the expression level of EPCAM protein reduces in Cas9 positive cells transduced with *EPCAM* sgRNA compared to the controls, indicating the functionality of the Cas9 in both S2-028 cell lines and S2-007. The quantification of the Cas9 activity was also shown to have been around 71.53% and 81.17% for S2-028 and S2-007, respectively.

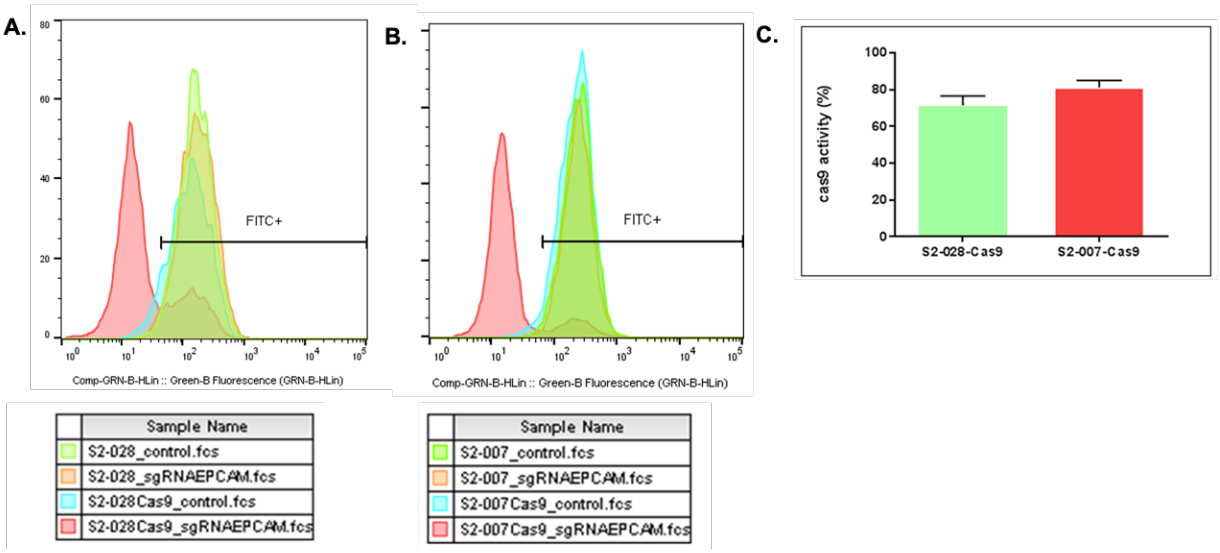


Figure 3.4. S2-028 and S2-007 express functional Cas9 protein. Flow cytometry was used to measure the EPCAM expression in positively transduced cells (mCherry positive cells) with or without the presence of sgRNA targeting *EPCAM*. EPCAM expression is represented by a FITC fluorescence signal since EPCAM proteins on the cells were stained with FITC. The functionality of Cas9 was determined by subtracting the FITC expression level of treated cells (cells transduced with the plasmid containing sgRNA targeting *EPCAM*) with control cells (cells transduced with empty plasmid). **(A, B)** Graphs show the expression level of EPCAM plotted on the x-axis and cell counts on the y-axis for both S2-028 (A) and S2-007 cells (B). The additional gating was added to separate the positive FITC population from autofluorescence by comparing the fluorescence levels between treated and control cells. The green peak is the expression level of *EPCAM* protein in wild-type cells transduced with empty plasmid (without *EPCAM* sgRNA). The orange peak is the expression level of EPCAM protein in wild-type cells transduced with *EPCAM* sgRNA containing the plasmid. The blue peak is the expression level of EPCAM protein in Cas9 positive cells transduced with empty plasmid (without *EPCAM* sgRNA). The red peak is the expression level of EPCAM protein in Cas9 positive cells transduced with *EPCAM* sgRNA-containing plasmid. **(C)** The quantification is shown of Cas9 activity measured by subtracting EPCAM protein levels of treated cells with that of control cells.

In preparation for the screening, the doubling time of the cells also needs to be counted to estimate the period of screening. The growth curve was used to visualize the growth speed of the cell in culture (Fig. 3.5). Based on the growth curve, the exponential phase of the growth could be determined, starting roughly from 96 – 168 hours of incubation time. The formula DT= Tln2/ln(Xe/Xb) was then used to calculate the doubling time using cell counts at particular time points of interest and the duration of the exponential phase estimate. The doubling time for both cells was successfully counted 28.31 hours for S2-028 and 30.31 hours for S2-007. Based on these numbers, we can estimate the screening period, which is 18 days (± 14 doubling time).

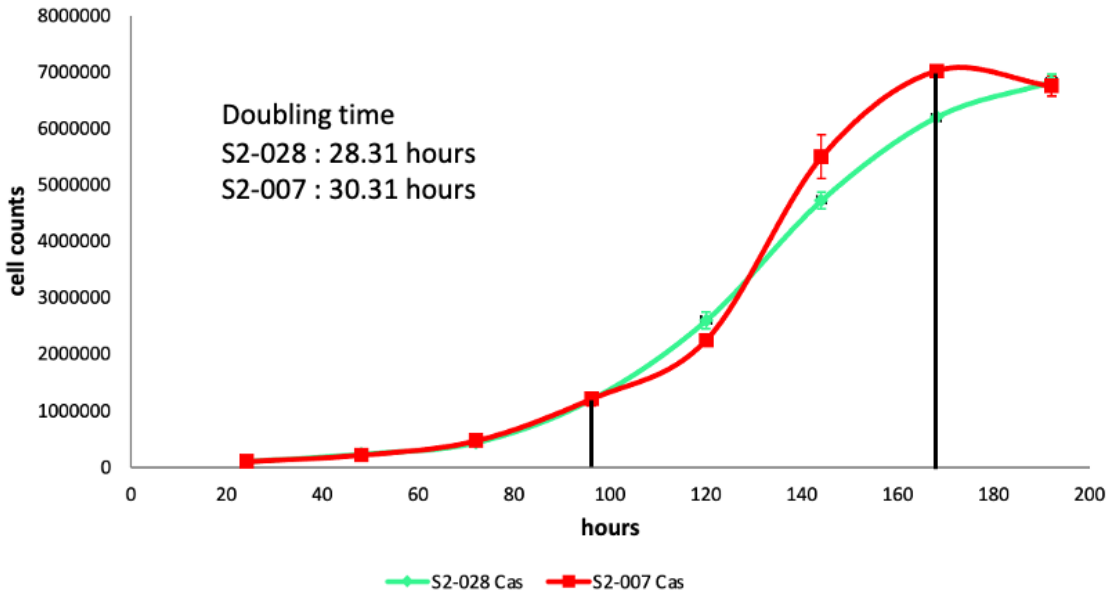


Figure 3.5. Growth curve for calculating the doubling time of S2-028 and S2-007. Cells were cultured and maintained in the 6-well plates for 192 hours (eight days). Every 24 hours (one day), cells from one well were harvested and counted. The cell counts were then plotted on the y-axis against incubation time in hours (n=2). The doubling time was calculated using the cell number in the exponential phase of the curve, which is 96- 168 hours. The formula DT= Tln2/ln(Xe/Xb) was used to calculate the doubling time. DT is doubling time, T is the incubation time (hours), Xb is the cell number at 96 hours, and Xe is the cell number at 168 hours. 28.31 hours and 30.31 hours are the doubling time for S2-028 and S2-007, respectively.

3.2 CRISPR-Cas9 library lentivirus production and transduction efficiency

Lentivirus production is a critical step for determining the transduction efficiency so that a successful CRISPR-Cas9 screening can be obtained. The transduction efficiency of the lentivirus can be examined by calculating the positively transduced cells or mCherry-positive

cells. Optimization of lentivirus production using PEI and Lipofectamine 3000 and LTX was previously done and tested in several different pancreatic cancer cells (SUIT-2, BxPC3, and Capan-1).

In SUIT-2 and BxPC3, the Lipofectamine 3000+LTX shows a significantly higher transduction efficiency than PEI (Fig. 3.6), but not in Capan-1 (Pirona et al., 2021). SUIT-2 is the parental cell for S2-028 and S2-007. Therefore, for the CRISPR screening, the Lipofectamine 3000+LTX method was used for lentivirus production.

After the lentivirus production for screening, the virus titer and the multiplicity of infection (MOI) were determined. The MOI is the ratio of the number of lentivirus particles to the number of cells. The cells need to be subjected to transduction with different volumes of lentivirus particles to calculate the virus titer. After the transduction, the positively transduced cells should be expressing mCherry, quantifiable via flow cytometry (Fig 3.7). The lentivirus titer was then calculated after obtaining the number of positively transduced cells or mCherry-positive cells. The virus titer was 7.5 million TU/mL and 8.2 million TU/mL for S2-028 and S2-007, respectively.

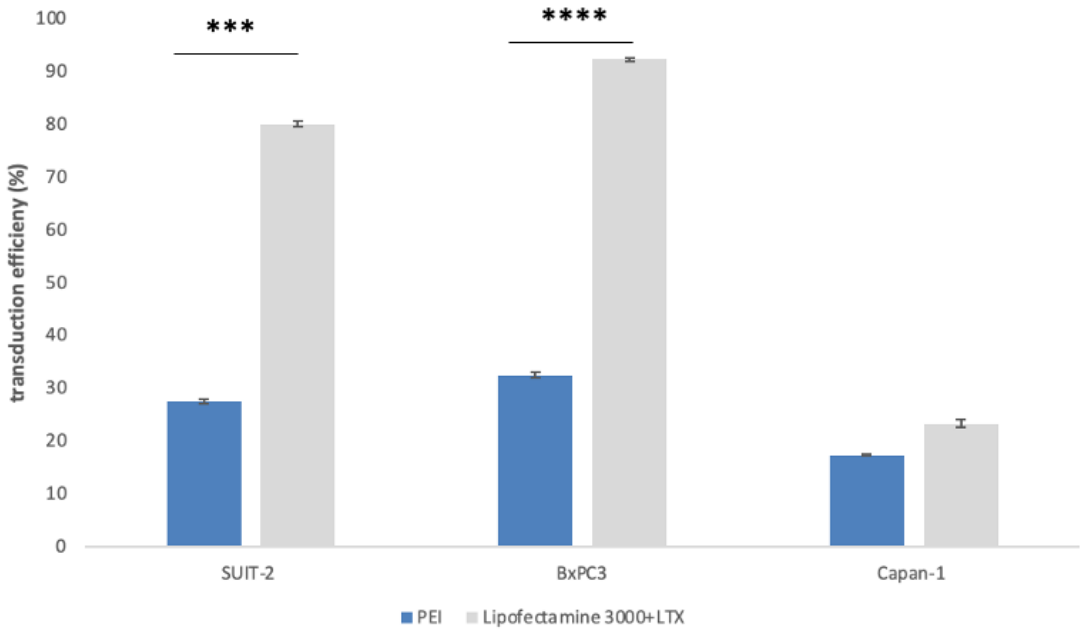


Figure 3.6. Lipofectamine 3000+LTX shows higher transduction efficiency compared to PEI in SUIT-2 and BxPC3 cells. SUIT-2, BxPC3, and Capan-1 were transduced with lentivirus produced from either PEI or Lipofectamine 3000 +LTX. Cells were then harvested and then subjected to flow cytometry to measure the percentage of the cells expressing mCherry. mCherry expressed cells represented positively transduced cells. The graph showed the percentage of transduction efficiency (mCherry + cells) on the y-axis and the group of the cells on the x-axis. * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

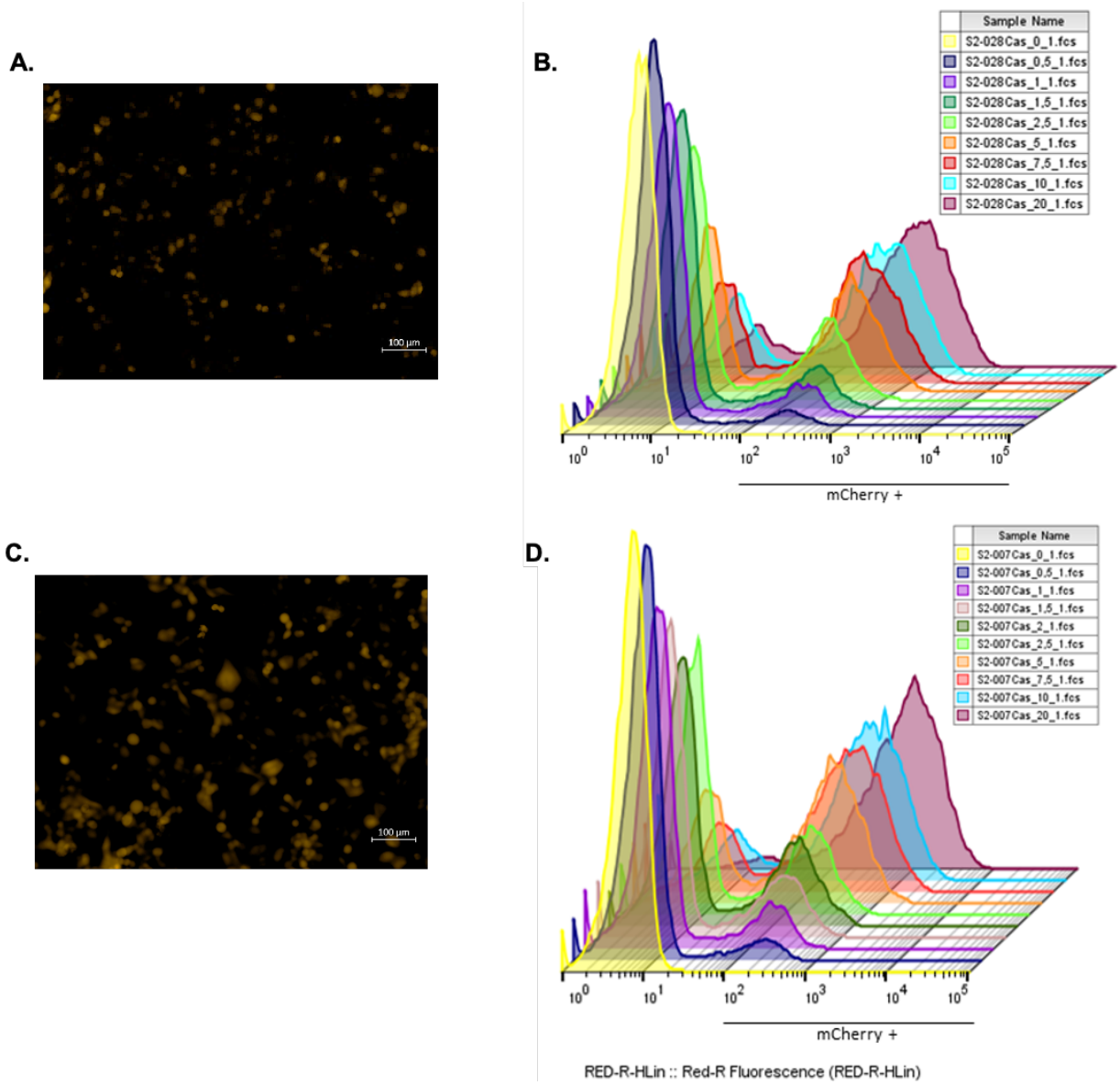


Figure 3.7. Lentivirus containing the CRISPR library is efficient enough to infect the S2-028 and S2-007. S2-028 and S2-007 cells were transduced with different amounts of lentivirus containing CRISPR library ranging from 0 to 20 µL. **(A, C)** Two days after transduction, cells were observed under an inverted microscope; the pictures show the expression of mCherry from the positively transduced cells with 100x magnification for S2-028 (A) and S2-007 (C), respectively. Scale bars show units of 100 µm. The orange-colored cells indicated that the cells were successfully transduced by expressing mCherry. **(B, D)** The graph showed the Red-R plotted on the x-axis. The additional gating was added (below the chart) to separate the positive mCherry population (Red-R positive) from the autofluorescence. The analysis was done by comparing the fluorescence level between transduced cells and control cells (transduced with 0 µL lentivirus). In addition, the graph showed the overview of mCherry expression level in S2-028 (B) and S2-007 (D) after transduction with 0 to 20 µL lentivirus containing CRISPR library. The increasing level of mCherry expressions corresponds to the amount of lentivirus transduced.

After calculating the lentivirus titer, MOI can then be determined. The number of mCherry-positive cells and MOI was then plotted to get the MOI chart (Fig. 3.8). The MOI chart is a tool to predict the efficiency of the CRISPR screening. The equation can be generated from the chart, which can be used to control virus volume to get specific efficiency in the screening in a particular cell number.

The screening was then conducted in 0.3 MOI to ensure the presence of only one sgRNA in one cell. The coverage number also has to be considered when planning the screening. The coverage number is the estimated number of sgRNA present during transduction. Several factors need to be considered, such as the number of sgRNAs in the library and the resources. The more the coverage, the more cells need to be prepared for the screening. The screening was then conducted using the CRISPR library consisting of 267,109 sgRNAs, 100x coverage, and 0.3 MOI based on all the considerations.

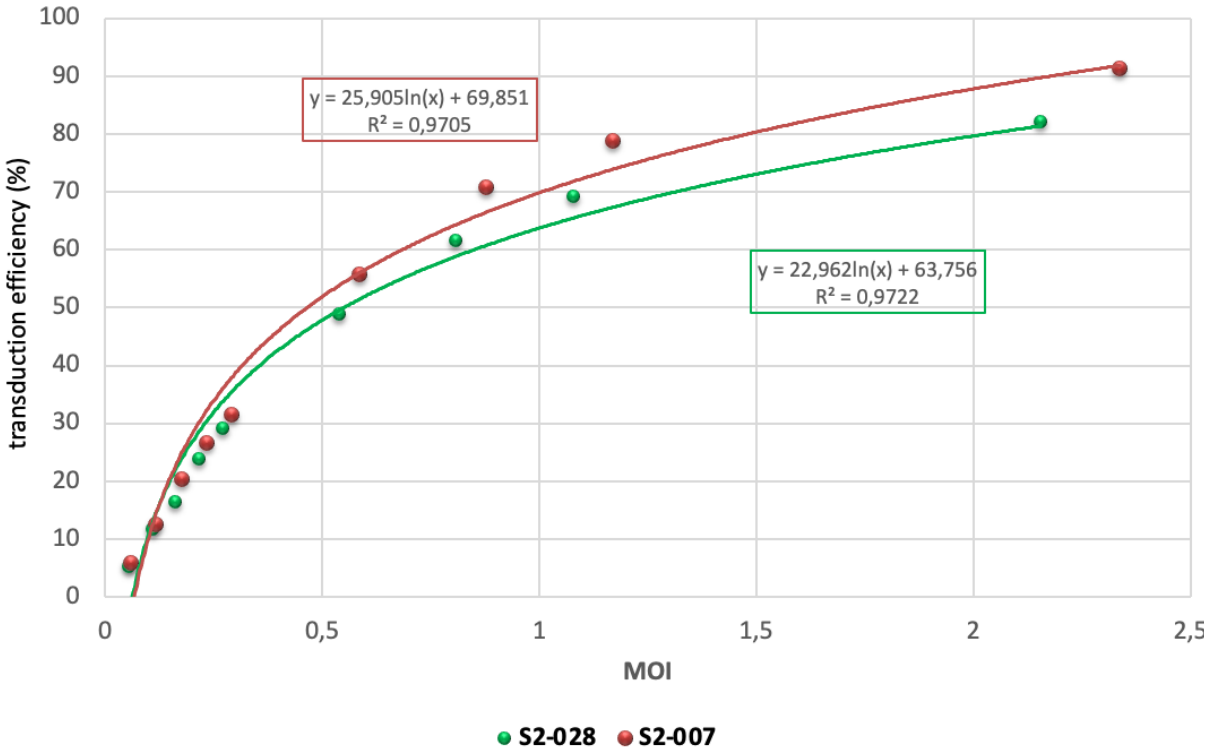


Figure 3.8. MOI chart as a tool to predict the efficiency of CRISPR screening. After the cells were transduced with gradual volumes of lentivirus and the transduction efficiency (%) (mCherry-positive cells) were assessed, the virus titer was then determined. The virus titer was then used to calculate the multiplicity of infection (MOI). Both transduction efficiency (%) and the MOI values were then plotted to form a nonlinear regression curve. The R squared (R^2) indicates how close the data is to the fitted regression line. The equation was shown to calculate the volume of the virus that corresponds to a particular MOI. The efficiency of the CRISPR screening can be controlled in both S2-028 and S2-007 cell lines based on the equation.

The final preparation before screening was to determine the puromycin concentration for antibiotic selection after transduction, which could be predicted using a puromycin kill curve (Fig 3.9). Based on the puromycin kill curve and cell observation under the microscope, 4 µg/mL and 5 µg/mL puromycin were used to select S2-028 and S2-007 cells.

After all the preparation above, the CRISPR screening was then conducted in S2-028 and S2-007 cells. The screening was performed in 100x coverage, 0.3 MOI, with the library of 267,109 sgRNAs, followed by puromycin selection to select only the positively transduced cells. After puromycin selection, the time point was considered as zero. The incubation time for the screening was 18 days, and the cells were continuously cultured and maintained during this period. During the screening, the cells needed to be cultured in good condition. Therefore, the cells needed to be periodically split into a new flask. In this process, the number of the cells had to be calculated to ensure that the screening coverage was still preserved by splitting the cells, resulting in more than the actual coverage (1000x). After the incubation time, at the end of day 18, the cells were subjected to DNA isolation. Isolated DNA was kept for further analysis.

The next step was the NGS preparation, which consists of two steps of PCRs aimed at recovering the sgRNA sequence from the DNA, and then followed by the NGS on Illumina HiSeq 2000 V4 using single read 50 base pairs (bp).

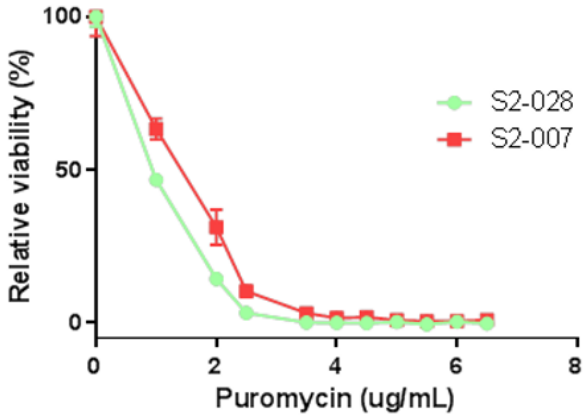


Figure 3.9. Puromycin kill curve. Cells were cultured in 96-well plates and incubated overnight. The following day, cells were treated with various concentrations of puromycin ranging from 0 to 6.5 µg/mL for three days. After the incubation period, the cells were subjected to a viability assay by adding resazurin. The fluorescence, which appeared after one hour of incubation, was then measured using a microplate reader. Percentage of relative viability ratio was then calculated by dividing the fluorescence amount of the treated samples with the control (0 µg/mL of puromycin). A graph was made by plotting the relative viability (%) on the y-axis and puromycin concentration on the x-axis (n=3). The green line represents S2-028 relative viability, and the red line represents S2-007.

3.3 Profiles of negatively and positively selected gene after MAGeCK data analysis

The NGS results were then subjected to analysis using MAGeCK. Several steps of quality control (QC) were firstly conducted to evaluate the result of the CRISPR screening before discovering essential genes. sgRNA distributions and Pearson correlation were two QCs that have been used. sgRNA counts distribution between samples were then observed using the box plot. Based on the box plot, the 1st quartile, median, and 3rd quartile of all the samples tested showed narrow distribution between samples. The correlation between samples' log sgRNA counts was then analyzed using Pearson correlation analysis and then visualized using a heatmap. The correlation between S2-028 and S2-007 samples at the t_{zero} was 0.97, indicating that the number of sgRNA at the beginning of the screening for both cells were similar (Fig. 3.10).

S2-007 and S2-028 were grown in their normal condition for 14 doubling times in the CRISPR screening. When the CRISPR-Cas9 is activated, a negative selection occurred. The essential genes were discovered by comparing the sgRNAs population before and after the screening. The absence of sgRNAs at the end of the screening indicated that gene perturbations caused the sgRNAs and CRISPR-Cas9 activation resulted in cell death.

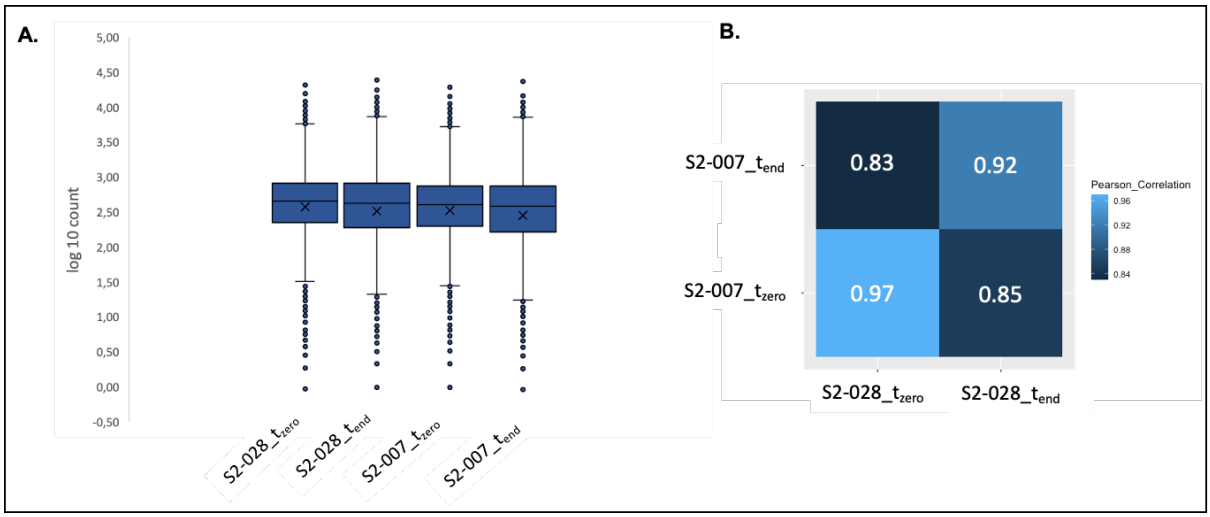


Figure 3.10. sgRNA counts distribution and Pearson correlation of the CRISPR screening results. CRISPR-Cas9 results were then analyzed using MAGeCK. The sgRNA counts were collected from MAGeCK analysis. **(A)** The box plot was made by plotting log 10 sgRNA count on the y-axis and the sample name (S2-028_t_{zero}, S2-028_t_{end}, S2-007_t_{zero}, and S2-007_t_{end}) on the x-axis. **(B)** The pairwise Pearson correlations of samples' log sgRNA counts were then calculated. Shown on the heatmap, the value of the correlation (p value < 2.2e-16) and the color represent the value.

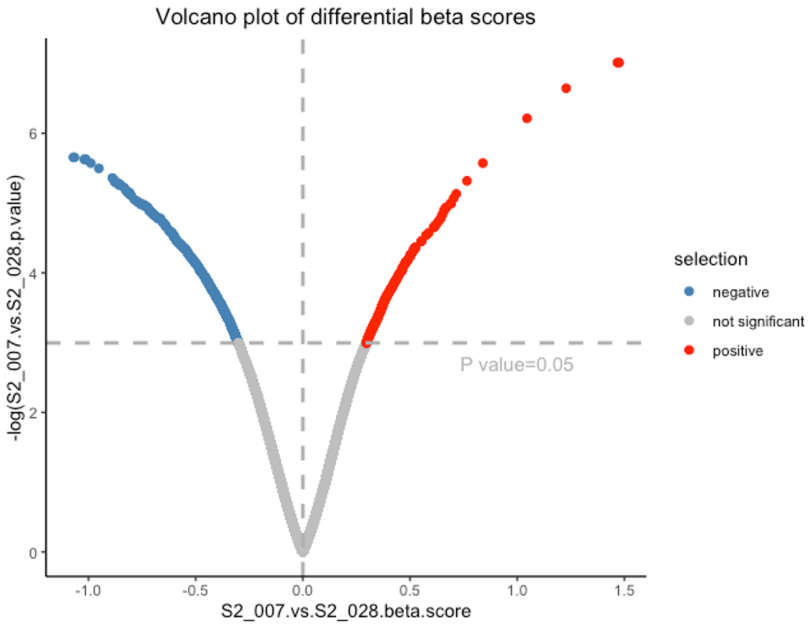


Figure 3.11. Negative and positive beta-scores indicate two different types of selection. CRISPR-Cas9 results were then analyzed using MAGeCK. The algorithm of MAGeCK calculated the beta-score by comparing the screening results at t_{end} between S2-007 and S2-028 (S2-007 vs. S2-028). The beta-score after the comparison of S2-007 and S2-028 screening results was then plotted against -log p value. A negative beta-score (left side; blue color) indicates negatively selected genes meaning that these genes were depleted in the S2-007 after the screening. A positive beta-score (right side; red color) indicates positively selected genes meaning that these genes were enriched in the S2-007 after the screening. Genes in blue and red colors are genes that have a significant p value (p value < 0.05).

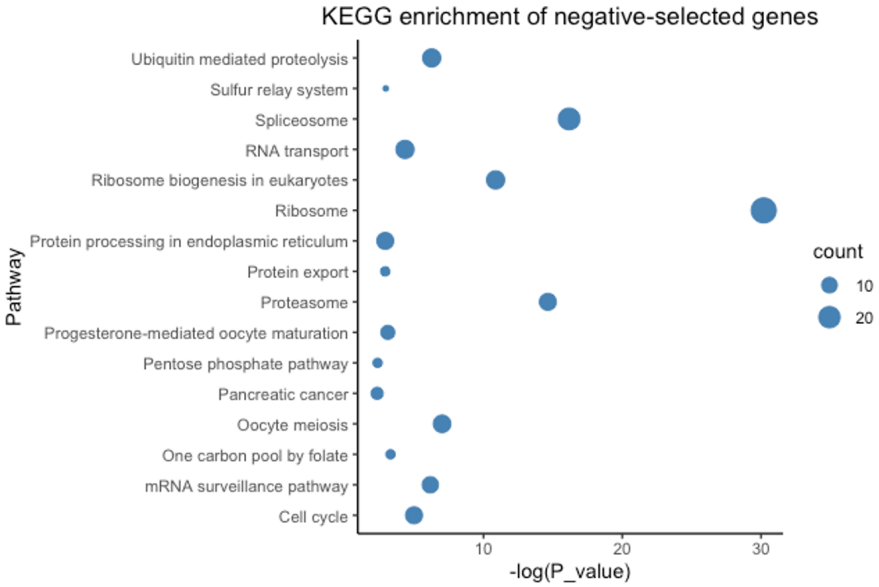


Figure 3.12. Ribosome pathway in the KEGG enrichment analysis of negatively selected genes. The significantly negative-selected genes were then analyzed using DAVID to know the KEGG enrichment pathway. The result was then plotted using R software to visualize the pathway's name on the y-axis and -log (p value) on the x-axis. The size of the dots represents the number of genes involved in a particular pathway.

The gene essentiality score is the result of the analysis using the MAGeCK algorithm. The beta-score or gene essentiality score is a value to show the level of selection as a consequence of gene perturbation in two different groups, which are compared. In this case, S2-007 and S2-028 were compared and the negative and positive beta-scores are shown (Fig. 3.11).

Further analysis was then conducted as part of the quality control. The significantly negative beta-score was then subjected to KEGG enrichment analysis using DAVID. The result showed that the ribosome pathway was the most prominent pathway among several other pathways, such as spliceosome, proteasome, and ribosome biogenesis in eukaryotes (Fig. 3.12). The complete KEGG enrichment analysis is shown in Appendix 3.

Both negatively and positively selected genes were then analyzed using DAVID to know their gene ontology (GO) in cellular component, molecular function, and biological processes. The cellular component enrichment analysis for negatively selected genes has shown extracellular exosome, membrane, nucleolus, focal adhesion, cytoplasm, cytosol, nucleus, and nucleoplasm. In addition, the most cellular components GO for positively selected genes were nucleus, mitochondrion, and nucleoplasm. The complete top 20 GO cellular component analysis is shown in Appendix 4. Nucleoplasm and nucleus are the two prominent locations for negative and positive selections (Fig. 3.13).

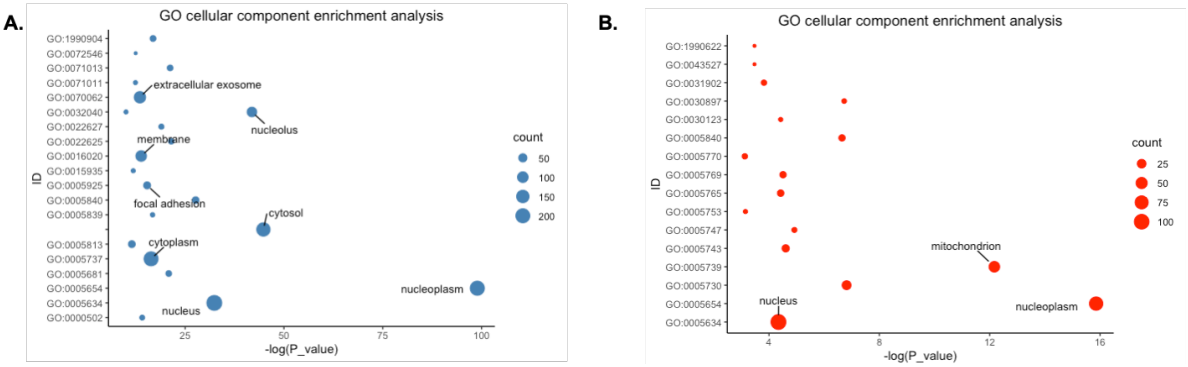


Figure 3.13. Nucleoplasm and nucleus are two prominent locations where the function of the genes from negatively and positively selected are located. The significantly negative- and positive-selected genes were then analyzed using DAVID to know the genes' gene ontology (GO). **(A, B)** The GO term's identification number (ID) was shown on the y-axis, and the value of -log (p value) was shown on the x-axis. The size of the dots corresponds to the number of the gene involves in a particular GO. The name of the specific GO IDs was shown if the count >30. The colors of the dots show two different types of selection; blue is negatively selected genes (A), and red is positively selected genes (B). The cellular component enrichment analysis on the negatively (A) and the positively (B) selected groups showed that the gene product functions in the nucleoplasm and nucleus.

Furthermore, the molecular function gene ontology analysis showed that adenosine triphosphate (ATP) binding, RNA binding, protein binding, and poly(A) RNA binding were found in the negatively selected genes. Poly(A) binding and protein binding were also found among GO terms for positively selected genes (Fig 3.14). The complete top 20 GO molecular function analysis is shown in Appendix 5. The top molecular function GO that both negatively and positively selected genes shared in molecular function gene ontology analysis are protein binding and poly(A) binding.

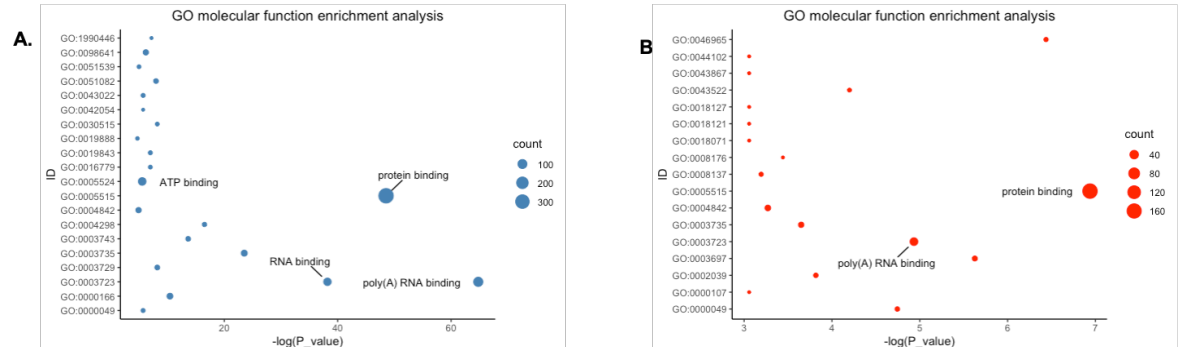


Figure 3.14. Protein binding and poly(A) binding are the molecular functions which both negatively and positively selected genes share based on the GO molecular function enrichment analysis. The genes were analyzed using DAVID to know the genes' gene ontology (GO). **(A, B)** The GO term's identification number (ID) was shown on the y-axis, and the value of -log (p value) was shown on the x-axis. The size of the dots corresponds to the number of the gene involves in particular GO terms. The name of the specific GO IDs was shown if the count >30 for negative selected genes (A), and count >20 for the positive selected genes (B). The colors of the dots show two different types of selection; blue is negatively selected genes (A), and red is positively selected genes (B).

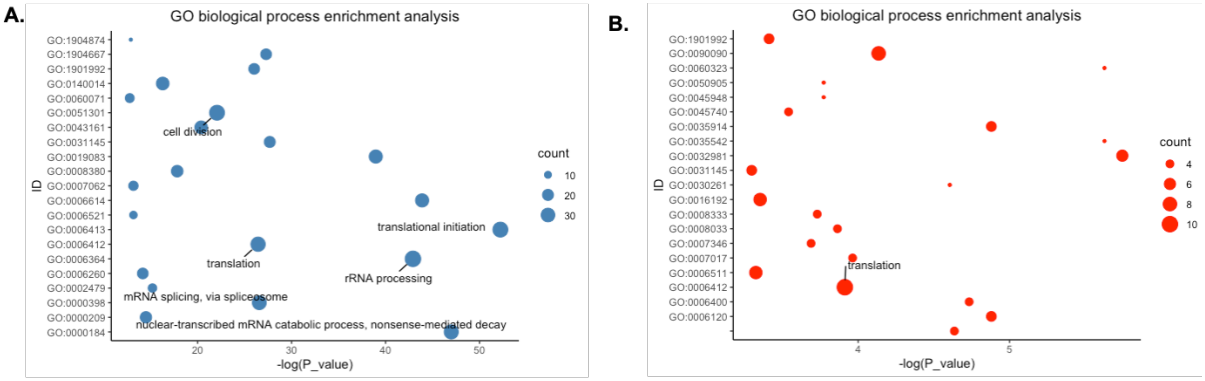


Figure 3.15. Translation is the biological process which all selected genes shared based on the GO biological process enrichment analysis. Analysis was done as in Fig. 3.14. The name of the specific GO IDs was shown if the count >30 for negative selected genes **(A)**, and count >9 for the positive selected genes **(B)**. The colors of the dots show two different types of selection; blue is negatively selected genes (A), and red is positively selected genes (B). The biological process enrichment analysis on the negative (A) and the positive (B) selected groups showed that the gene products perform in translation.

The GO biological process enrichment analysis showed that cell division, translation, mRNA splicing via spliceosome, rRNA processing, and translational initiation were among several terms that resulted in the analysis for negatively selected genes. However, translation also showed as most prominent in the positively selected genes GO analysis (Fig. 3.15). The complete top 20 GO biological process analysis is presented in Appendix 6. Translation is the top molecular function GO that both negatively and positively selected genes shared.

3.4 Selection of candidate genes

The beta-scores that were already obtained from the MAGeCK analysis distinctly distinguish the beta-score from negative and positive selection. There are 590 genes in the significantly positive-selected group and 348 genes in the significantly negative-selected group. Furthermore, the clustering of the genes based on the beta-score showed no overlapping gene between the negative and positive selection (Fig. 3.16).

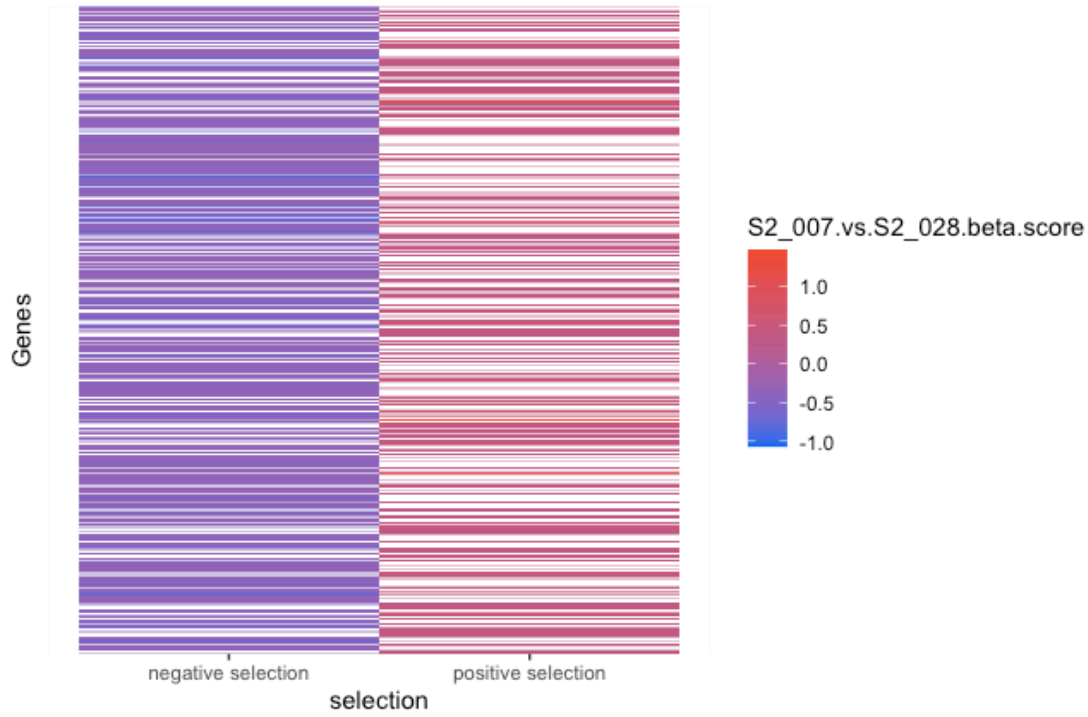


Figure 3.16. Clustering of significantly negative- and positive-selected genes based on the beta-score. A heatmap was made by plotting genes on the x-axis and the type of selections (negative and positive selections) on the y-axis. The color represents beta-score from the comparison between S2-007 and S2-028 using its sgRNAs counts at t_{end}. Thus, red represents a positive score, and blue represents a negative score.

A plot was then made to visualize the comparison of the beta-score from both samples (S2-028 and S2-007) (Fig 3.17). The plot was then overlayed by the beta-score comparison of S2-007 and S2-028, resulting in two different color overlays. The blue dots represent the negative or depleted genes after the negative screening in S2-007, while the red dots represent the positive or enriched genes after the negative screening in S2-007. To discover the essential genes in S2-007 but not in S2-028, the position showing the negative beta-score needs to be explored more. Based on the MAGeCK analysis, the dots close to the 0 represents gene essential in one condition but not in another. *TYMS*, *MEN1*, and *MYBL2* were chosen because these genes are negatively selected genes (depleted genes in S2-007) and close to 0.

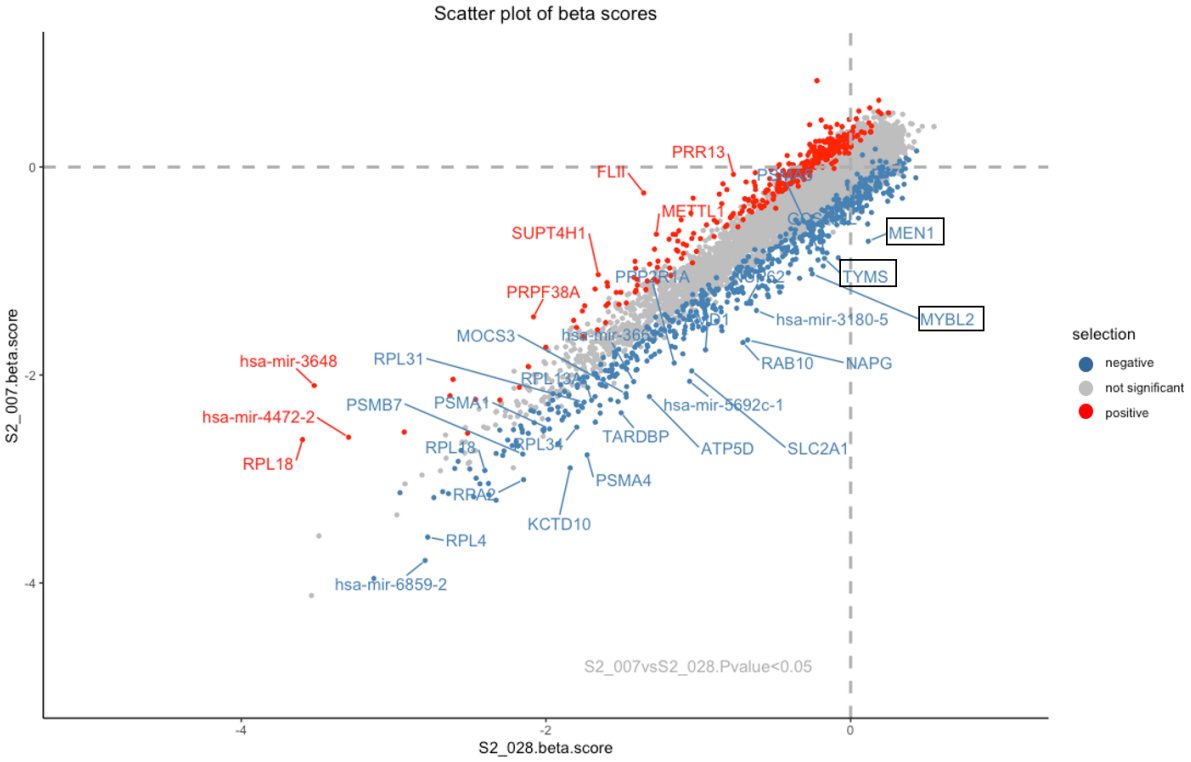


Figure 3.17. *TYMS*, *MEN1*, and *MYBL2* were selected as candidates based on these essentiality scores (beta-scores). The beta-score plot was constructed by plotting the beta-score of S2-007 and S2-028. Each beta-score of each cell line was measured by comparing the group of sgRNA at t_{zero} to t_{end}. The plot then overlays by the comparison between S2-007 and S2-028 using its sgRNAs counts at t_{end}. The red dots represent a positive selection, and the blue dots represent a negative selection. Genes colored red, or blue are significantly selected genes (p value < 0.05).

After selecting three candidate genes based on the beta-score, all sgRNA detected of each candidate gene was then observed to verify the gene essentiality of these candidates. The sgRNA counts of each gene were then subjected to the drop-out value calculation. Drop-out value is the ratio between sgRNA counts at t_{end} and t_{zero}. The lower the value of the sgRNA,

the more cell death caused by the corresponding sgRNA; therefore, the more we observed a low drop-out value in one sample, the more the gene essentiality of that gene in a particular sample. Finally, a heatmap showing all the drop-out values for all sgRNA in one gene was made (Fig. 3.18). Based on the overview of the drop-out value, *TYMS*, *MEN1*, and *MYBL2* are considered essential genes in S2-007 but not in S2-028.

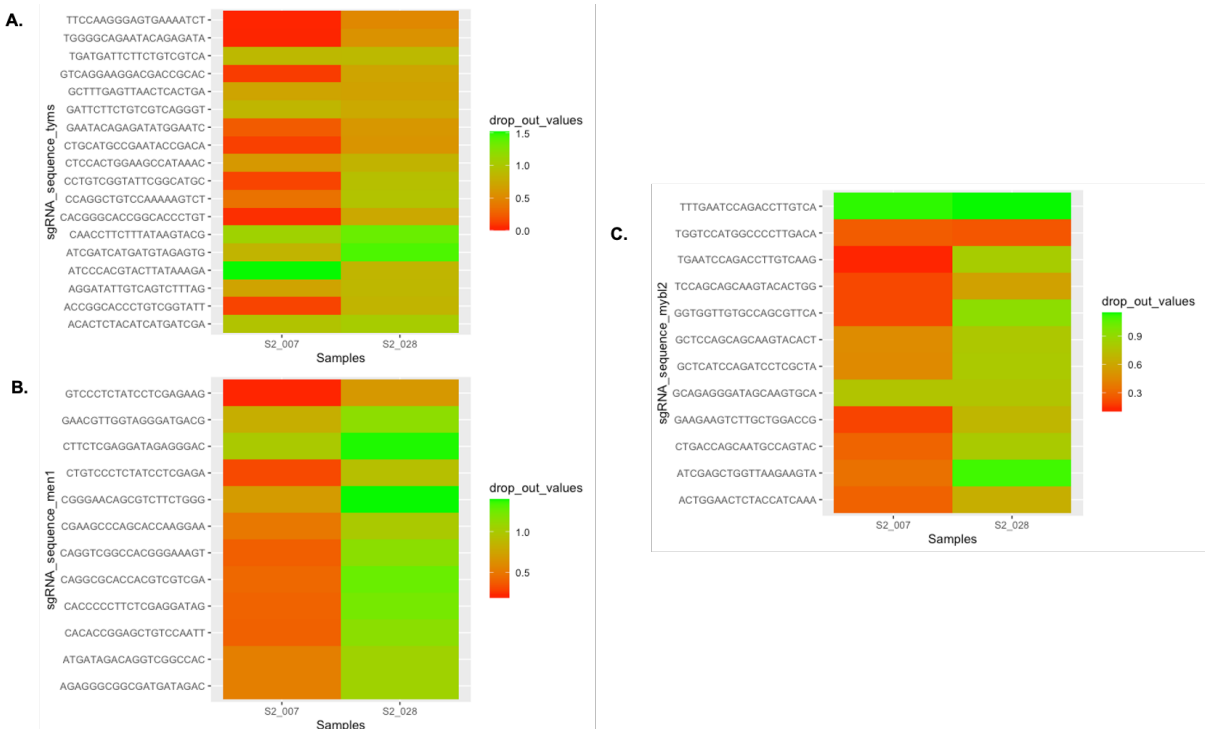


Figure 3.18. Drop-out values of each sgRNA detected for candidate genes (*TYMS*, *MEN1*, *MYBL2*) showed gene essentiality in S2-007. Drop out values for each cell are the ratio between sgRNA counts of each sgRNA at t_{end} and t_{zero}. **(A, B, C)** Heatmap of drop-out value of each sgRNA for *TYMS* (A), *MEN1* (B), and *MYBL2* (C). The heatmap was constructed by plotting sgRNA_sequence for each candidate gene on the y-axis and sample name (S2-007 and S2-028) on the x-axis. The color represents the value of the drop-out value. The green color indicates high, and red shows a low drop-out value.

3.5 Effect of *TYMS*, *MEN1*, *MYBL2* knockout in cell proliferation, cell viability, and colony formation

In vitro validation such as cell proliferation, cell viability, and colony formation for each candidate was then conducted. Five sgRNAs sequences from each candidate genes and two no target control (NTC) and empty vector were used. Cells were transduced with lentivirus containing sgRNAs or NTC or only vector. Two days after transduction, cells were in puromycin selection for three days to select for positively transduced cells. At the end of the selection, cells were cultured in a 96-well plate for seven days, and the cell viability was

measured daily by adding resazurin followed by incubation. After the incubation, the fluorescence from the resorufin formed due to the metabolic activity of live cells. The fluorescence was then measured using a fluorometer. The viability test from each day was then plotted to make a cell proliferation plot.

In addition to the resazurin assay for cell viability, cell viability based on the mCherry positive cells was also performed. By using this approach, the viability detection is more sensitive only to positively transduced cells. In this assay, at the end of antibiotic selection, cells were mixed with the non-transduced cells with a ratio of 1:4. Cells were then plated in a 6-well plate, and then incubated and maintained for seven days. On day seven, cells were harvested and collected for flow cytometry analysis. Cells were then also maintained for the next time point.

The colony formation assay was also conducted on the soft colony agar. In this assay, cells were maintained for almost 15 days until the colonies become visible. These three assays were performed to validate the screening result of whether *TYMS*, *MEN1*, and *MYBL2* are essential genes in S2-007 but not in S2-028.

3.5.1 *TYMS* knockout decreases cell proliferation, cell viability, and soft agar colony formation in S2-028 and S2-007 cell lines

Cells were transduced with lentivirus containing either *TYMS*_sg1 (*TYMS*_sgRNA1), *TYMS*_sg2 (*TYMS*_sgRNA2), *TYMS*_sg3 (*TYMS*_sgRNA3), *TYMS*_sg4 (*TYMS*_sgRNA4), *TYMS*_sg5 (*TYMS*_sgRNA5), NTC1 (no target control 1), NTC2 (no target control 2), or vector (empty vector). Two days after transduction, cells were subjected to puromycin selection for three days to single out positively transduced cells. At the end of the selection, cells were cultured in a 96-well plate for seven days, and the cell viability was measured daily using resazurin assay. Based on the fluorescence measured from the assay, a cell proliferation plot was made (Fig. 3.19). In both cell lines, *TYMS* knockout shows a decrease in cell proliferation compared to the controls (NTC1 and NTC1) and the vector. The cell viability on day seven was then extracted and plotted as a bar graph to verify the knockout effect at the end of the incubation time. Both graphs show that all the *TYMS* knockout groups are significantly lower in cell viability on day seven than the controls in S2-028 and S2-007. Subsequently, *TYMS*_sg2 and *TYMS*_sg4 were selected for further analysis because the level of reductions in cell viability for both sgRNAs in both cell lines was more than the other sgRNAs.

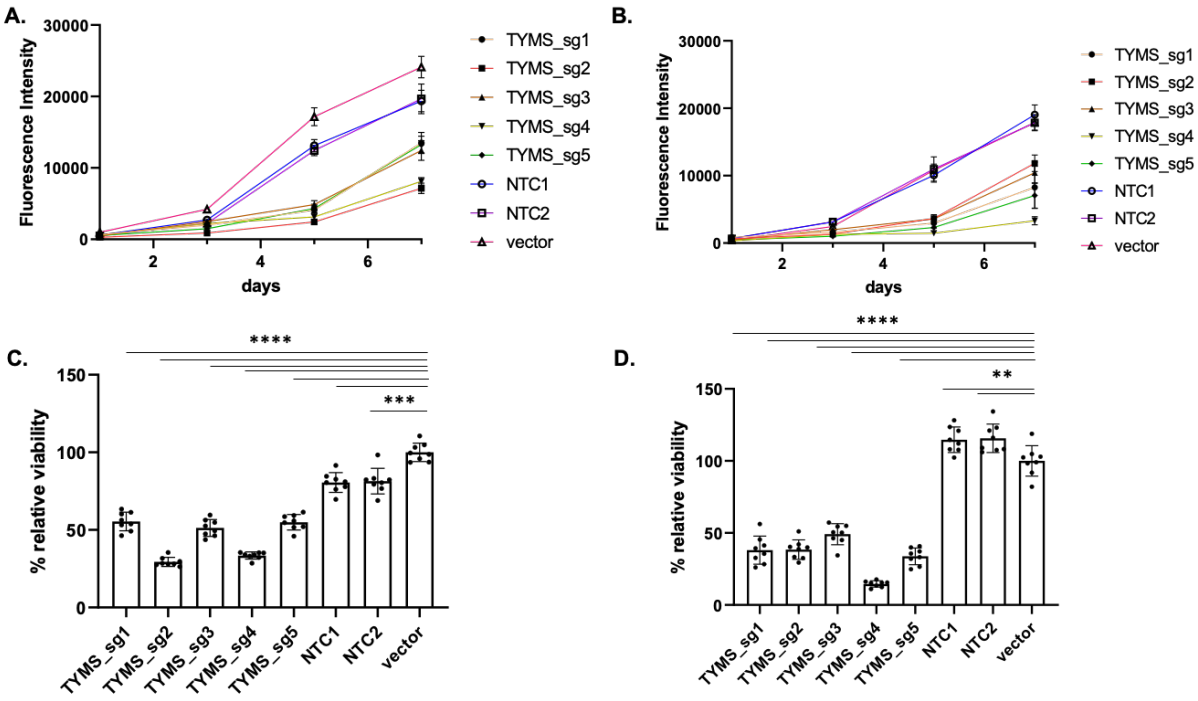


Figure 3.19. *TYMS* knockout decreases cell proliferation and cell viability in S2-028 and S2-007 based on resazurin assay. Cells were transduced, selected, and cultured for seven days and subjected for viability assay using resazurin every 24 hours. Fluorescence signals after resazurin assay were measured. **(A, B)** Fluorescence values from the resazurin assay were collected and then plotted against the day of data collection to make a cell proliferation plot. S2-028 cell proliferation plot is shown on the left side (A), and S2-007 on the right side (B). Each plot shows different groups tested which are *TYMS*_sg1, *TYMS*_sg2, *TYMS*_sg3, *TYMS*_sg4, *TYMS*_sg5, NTC1, NTC2, and vector (n=8). **(C, D)** The cell viability on day seven was then analyzed further by normalizing the fluorescence values of the *TYMS* knockout groups (*TYMS*_sg1- *TYMS*_sg5) and controls (NTC1 and NTC2) to the vector. The bar graph was then constructed with the groups' names on the x-axis and percentage of relative viability on the y-axis. The bar graphs above are shown for S2-028 (C) and S2-007 (D) samples. * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

Furthermore, to confirm the knockout of the *TYMS*, Western blot was then performed. Cells were transduced with lentivirus containing either *TYMS*_sg2 (*TYMS*_sgRNA2) plasmid construct, *TYMS*_sg4 (*TYMS*_sgRNA4) plasmid construct, NTC1(no target control 1), or empty vector (vector). Two days after transduction, cells were subjected to puromycin selection for three days to single out positively transduced cells. At the end of the selection, the cells were then plated in a 6-well plate and incubated for three days or seven days before their protein was isolated. Protein was then measured with a BCA assay and subjected to Western blot.

Based on the Western blot, on day 3 the bands of the thymidylate synthase (TS) protein (encoded by *TYMS* gene) only appeared in NTC1 and vector, and not in the *TYMS*_sg2 and

*TYMS*_sg4 samples, indicating complete knockout of *TYMS* in S2-028 and S2-007. On day 7 the smear band appeared in the *TYMS*_sg2 and *TYMS*_sg4, but intact in NTC1 and vector, indicating that the *TYMS* protein was present in S2-028 and S2-007 (Fig. 3.20).

Next, further confirmation on the cell viability was observed using flow cytometry. For this purpose, at the end of the puromycin selection, cells were mixed with the non-transduced cells with a ratio of 4:1. Cells were then plated in a 6-well plate, and then incubated and maintained for seven days. On day 7, cells were harvested and collected for flow cytometry analysis. The leftover cells were then plated in a new 6-well plate for another 7 days incubation until day 14. Flow cytometry was conducted on day 0, 7, and 14 to monitor the percentage of transduced cells that contain mCherry protein.

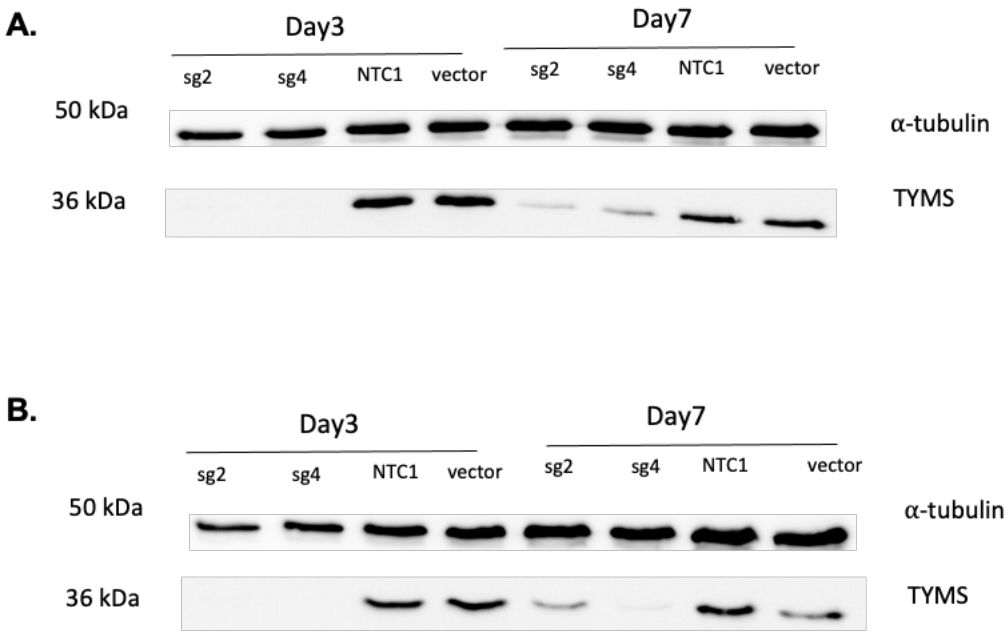


Figure 3.20. *TYMS* sgRNA2 and sgRNA4 plasmid constructs successfully knockout the *TYMS* protein on day 3 in S2-028 and S2-007. 20 µg of protein of each plasmid construct group (*TYMS*_sg2 (*TYMS*_sgRNA2), *TYMS*_sg4 (*TYMS*_sgRNA4), NTC1 (no target control 1), or empty vector (vector) was loaded to a 10% sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE)gel. The gel was run in parallel with Spectra Multicolor Broad Range Protein marker, 135 voltage. The protein on the SDS-PAGE gel was then transferred to a nitrocellulose membrane using a semi-dry transfer method. The membrane was washed and incubated with the primary and secondary antibodies for *TYMS* and alpha-tubulin (α-tubulin). The Western blot bands were then visualized by adding HRP substrate, and an image was taken directly using the Intas ECL ChemoStar imager. **(A, B)** The Western blot picture shows the bands of α-tubulin (50 kDa) and the thymidylate synthase (TS) protein (36kDa) proteins in S2-028 (A) and S2-007 (B) on day 3 and day 7, respectively.

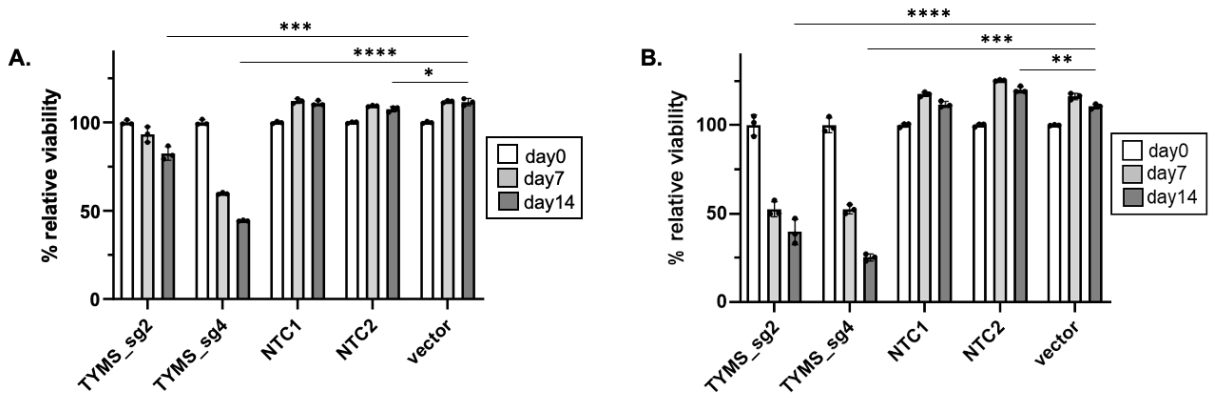


Figure 3.21. *TYMS* knockout decreases cell viability in S2-028 and S2-007 based on flow cytometry assay. Cells were transduced, selected, and cultured for 14 days in total and subjected for viability assay using flow cytometry every 7 days. Flow cytometry was conducted on day 0, 7, and 14 to monitor the percentage of transduced cells that contain mCherry protein. The number of mCherry positive cells represents the viability of transduced cells over the incubation period. The percentage of each time point were then normalized with the value on day 0. **(A, B)** The bar graph was then constructed by plotting the groups' names on the x-axis to the percentage of relative viability on the y-axis (n=3) for both samples, S2-028 (A) and S2-007 (B). * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

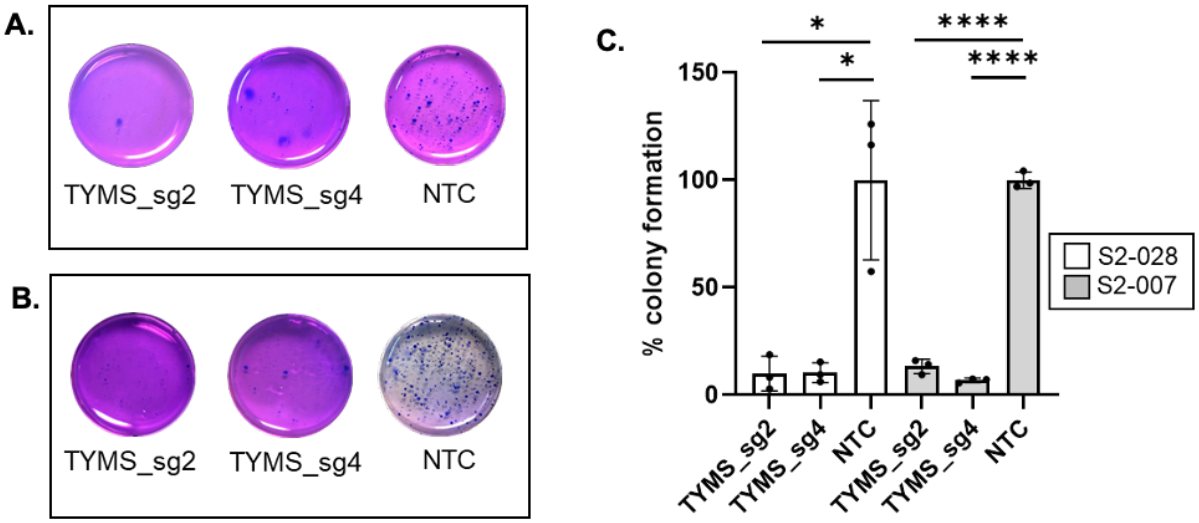


Figure 3.22. *TYMS* knockout decreased colony formation in S2-028 and S2-007. Cells were transduced, selected, and cultured on soft agar for 15 days. On day 15, cells were washed with PBS and then stained with crystal violet. **(A, B)** The picture of each well was then taken for S2-028 (A) and S2-007 (B), respectively. The cells, stained by crystal violet, appear darker than the background. The picture was then subjected for analysis using Open CFU software to count the number of cells, which was normalized with control to obtain the percentage of colony formation. (C) A bar graph was then made by plotting colony formation (%) on the y-axis against the group's name on the x-axis (n=3). * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

The bar graph was then made to observe the difference in relative ability (%) based on mCherry expression (Fig. 3.21). *TYMS*_sg2 and *TYMS*_sg4 significantly decreased cell viability in both S2-028 and S2-007 when compared to the controls (NTC1 and NTC2) and vector. However, the decrease of cell viability in S2-007 was relatively higher than in S2-028.

Colony formation assay was initiated by cell transduction and then puromycin selection. At the end of the selection, cells were then plated in a 6-well plate on soft agar. Cells were then maintained in the incubator for 15 days. Visible cells were stained and counted (Fig. 3.22). Based on the observation, the colony formation of *TYMS* knockout *TYMS*_sg2 and *TYMS*_sg4 were significantly lower than the control (NTC) in S2-028 and S2-007. Furthermore, a bar graph was composed to compare percentage of colony formation between sample groups. Based on the graph, the *TYMS* knockouts (*TYMS*_sg2 and *TYMS*_sg4) had significantly reduced the colony formation in S2-028 and S2-007.

3.5.2 *MEN1* knockout decreased cell proliferation and cell viability in S2-007 cell lines, as well as soft agar colony formation in S2-028 and S2-007 cell lines.

For the cell proliferation assay, cells were firstly transduced with lentivirus containing either *MEN1*_sg1 (*MEN1*_sgRNA1), *MEN1*_sg2 (*MEN1*_sgRNA2), *MEN1*_sg3 (*MEN1*_sgRNA3), *MEN1*_sg4 (*MEN1*_sgRNA4), *MEN1*_sg5 (*MEN1*_sgRNA5), NTC1 (no target control 1), NTC2 (no target control 2), or vector (empty vector). Then, two days after transduction cells were subjected to puromycin selection for three days to single out positively transduced cells. At the end of the selection, cells were cultured in a 96-well plate for 7 days, and cell viability was measured daily using a resazurin assay.

A cell proliferation plot was made based on fluorescence measured from the assay (Fig. 4.23). In S2-007, *MEN1* knockout showed decreased cell proliferation compared to the controls (NTC1 and NTC1) and the vector. Cell viability on day 7 was then extracted and plotted as a bar graph to verify the knockout effect. Based on the observation, *MEN1* knockout groups were significantly lower in cell viability on day 7 than the controls in S2-007, but not in S2-028. *MEN1*_sg3 and *MEN1*_sg4 were then selected for further analysis because the level of reductions in cell viability for both sgRNAs in both cell lines was more than the other sgRNAs.

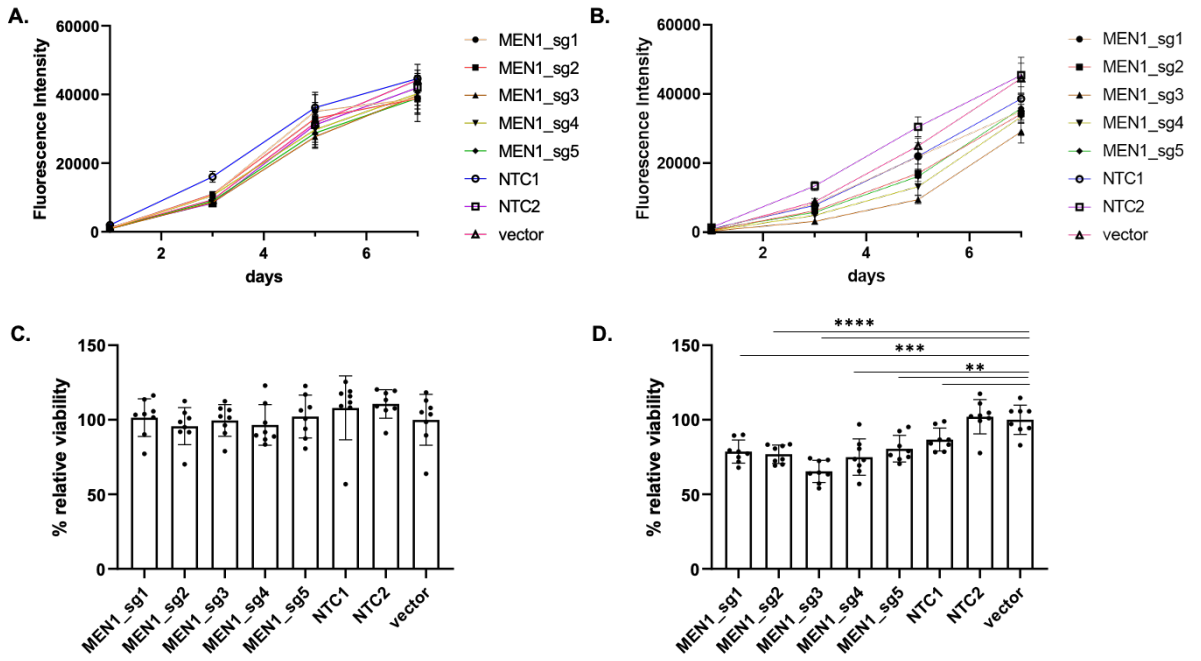


Figure 3.23. *MEN1* knockout decreased cell proliferation and cell viability in S2-007 based on resazurin assay. Cells were transduced, selected, and cultured for 7 days and subjected for viability assay using resazurin every 24 hours. Fluorescence signals produced after resazurin assay were measured. **(A, B)** Fluorescence's values were then plotted against the day of data collection to generate a cell proliferation plot. S2-028 cell proliferation plot is shown on the left side (A), and S2-007 on the right side (B). Each plot shows different groups tested, which are *MEN1*_sg1, *MEN1*_sg2, *MEN1*_sg3, *MEN1*_sg4, *MEN1*_sg5, NTC1, NTC2, and vector (n=8). **(C, D)** The cell viability on day 7 was then analyzed further by normalizing the fluorescence values of the *MEN1* knockout groups (*MEN1*_sg1-*MEN1*_sg5) and controls (NTC1 and NTC2) to the vector. The bar graph was then constructed with the groups' names on the x-axis and % relative viability on the y-axis for S2-028 (C) and S2-007 (D). The significant value was tested using an unpaired student t-test or Mann-Whitney test * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

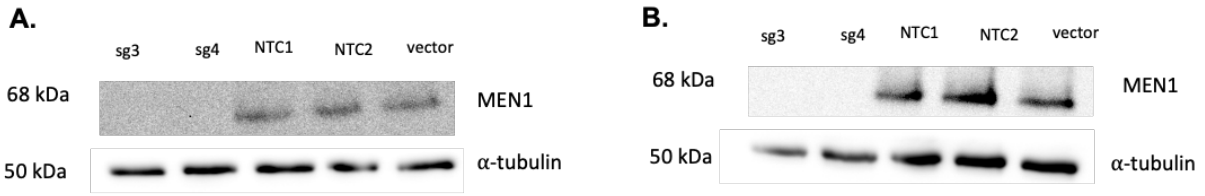


Figure 3.24. *MEN1* sgRNA3 and sgRNA4 plasmid constructs successfully knocked out the *MEN1* protein in S2-028 and S2-007. 20 µg of protein of each group were loaded to a 10% SDS-PAGE gel. The gel was run in parallel with Spectra Multicolor Broad Range Protein marker, 135 Voltage. The protein on the SDS-PAGE gel was then transferred to a nitrocellulose membrane using a semi-dry transfer method. The membrane was then washed and incubated with the primary and secondary antibodies for *MEN1* and alpha-tubulin (α-tubulin). The Western blot bands were visualized by adding HRP substrate, and then an image was taken directly using the Intas ECL ChemoStar imager. **(A, B)** The Western blot picture shows the bands of α-tubulin (50 kDa) and Menin (68 kDa) proteins in S2-028 (A) and S2-007 (B) on day 7, respectively.

Furthermore, to confirm the knockout of the *MEN1*, Western blot was performed. Cells were transduced with lentivirus containing either *MEN1*_sg3 (*MEN1*_sgRNA3) plasmid construct, *MEN1*_sg4 (*MEN1*_sgRNA4) plasmid construct, NTC1(no target control 1), NTC2 (no target control 2), or empty vector (vector). Two days after transduction, cells were subjected to puromycin selection for three days to single out positively transduced cells. At the end of the selection, the cells were plated in a 6-well plate and incubated for 7 days before their protein was isolated. Protein was then measured with a BCA assay and subjected to Western blot. Based on the Western blot, Menin protein (encoded by *MEN1* gene) bands only appear in NTC1, NTC2, and vector, and not in the *MEN1*_sg3 and *MEN1*_sg4 samples, indicating complete knockout of *MEN1* in S2-028 and S2-007. (Fig. 3.24).

Cell viability measurement based on the mCherry positive cells was then conducted. The assay was initiated by transducing cells with lentivirus and then cells were subjected to puromycin selection. At the end of the selection, cells were mixed with the non-transduced cells with a ratio of 4:1. Cells were then plated in a 6-well plate, and then incubated and maintained for 7 days. On day 7, cells were harvested and collected for flow cytometry analysis. The leftover cells were then plated into a new 6-well plate for another 7-day incubation until day 14 and day 21. Flow cytometry was conducted on day 0, 7, 14, and 21 to monitor the percentage of transduced cells that contain mCherry protein.

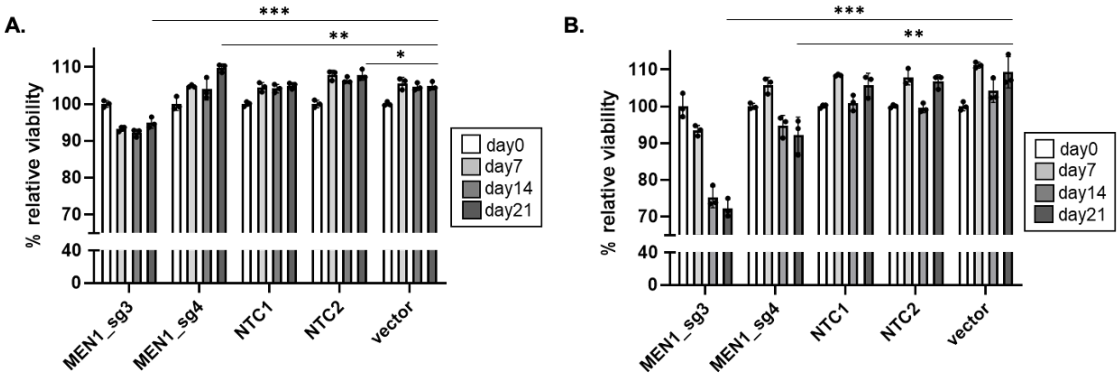


Figure 3.25. *MEN1* knockout decreased cell viability in S2-028 and S2-007based on flow cytometry assay. Cells were transduced, selected, and cultured for 21 days in total and subjected to viability assay using flow cytometry every 7 days. Flow cytometry was conducted on day 0, 7, 14, and 21 to monitor the percentage of transduced cells containing mCherry. The number of mCherry positive cells represents the viability of transduced cells over the incubation period. The percentage of each time point were then normalized with the value on day 0. (A, B) The bar graph was then constructed by plotting the groups' names on the x-axis to the % relative viability on the y-axis (n=3), for both sample S2-028 (A) and S2-007 (B). * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

Bar graphs were made for both samples to observe the difference in cell viability between different groups (Fig 3.25). Based on the observation, the *MEN1* knockouts (*MEN1*_sg3 and *MEN1*_sg4) significantly decreased the cell viability only in S2-007 (B), and not in S2-028 (A), when compared to the controls (NTC1 and NTC2) and vector. *MEN1*_sg3 showed more reduction compared to the *MEN1*_sg4 knockout in S2-007. In S2-028, *MEN1*_sg3 significantly decreased the cell viability, whereas *MEN1*_sg4 showed a significant increase. However, the decrease of cell viability in S2-007 relatively higher than in S2-028.

For colony formation assay, cells were transduced with lentivirus containing either *MEN1*_sg3 (*MEN1*_sgRNA3) plasmid construct, *MEN1*_sg4 (*MEN1*_sgRNA4) plasmid construct, or NTC (no target control). Two days after transduction, cells were subjected to puromycin selection for three days to single out positively transduced cells. At the end of the selection, cells were then plated in a 6-well plate on soft agar and maintained in the incubator for 15 days. Visible cells were stained and counted (Fig. 3.26). The colony formations of *MEN1* knockouts (*MEN1*_sg3 and *MEN1*_sg4) were lower than the control (NTC) in S2-028 and S2-007. Furthermore, the bar graph was composed to observe the percentage of colony formation between sample groups. Based on the graph, the *MEN1* knockouts (*MEN1*_sg3 and *MEN1*_sg4) had significantly reduced the colony formation assay in S2-028 and S2-007.

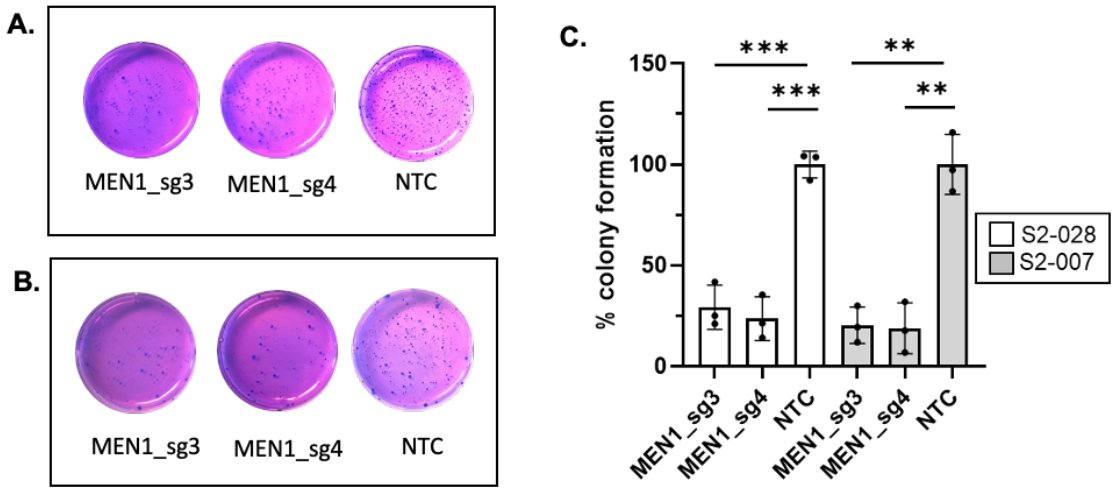


Figure 3.26. *MEN1* knockout decreases colony formation in S2-028 and S2-007. Cells were transduced, selected, and cultured on soft agar for 15 days of incubation. On day 15, cells were washed with PBS and then stained with crystal violet. **(A, B)** The picture of each well was then taken for S2-028 (A) and S2-007 (B), respectively. The cells, stained by crystal violet, appear darker than the background. The picture was subjected for analysis using Open CFU software to count the number of cells, which was then normalized with control to obtain the percentage of colony formation. **(C)** A bar graph was then made by plotting colony formation (%) on the y-axis against the group's name on the x-axis (n=3). * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

3.5.3 *MYBL2* knockout decreased cell proliferation, cell viability and soft agar colony formation in S2-007 cell line

For the cell proliferation assay, cells were transduced with lentivirus containing either *MYBL2*_sg1 (*MYBL2*_sgRNA1), *MYBL2*_sg2 (*MYBL2*_sgRNA2), *MYBL2*_sg3 (*MYBL2*_sgRNA3), *MYBL2*_sg4 (*MYBL2*_sgRNA4), *MYBL2*_sg5 (*MYBL2*_sgRNA5), NTC1 (no target control 1), NTC2 (no target control 2), or vector (empty vector). Two days after transduction, cells were subjected to puromycin selection for three days to single out positive transduction cells. At the end of the selection, cells were cultured in a 96-well plate for 7 days, and cell viability was measured daily using a resazurin assay.

A cell proliferation plot was made based on the fluorescence measured from the assay (Fig. 3.27). In S2-028 and S2-007, *MEN1* knockout showed no effect on cell proliferation compared to the controls (NTC1 and NTC1) and the vector. In addition, the cell viability on day 7 was extracted and plotted as a bar graph to validate the knockout effect on cell viability at the end of the incubation time. *MYBL2* knockouts showed no effect on cell viability on day 7 compared to the controls in S2-028 and S2-007.

Based on the viability result using resazurin, even though *MYBL2*_sg1 and *MYBL2*_sg5 showed a significant decrease in the cell viability, the difference was not too distinct. Western blot was then performed to verify the knockout from each group.

Cells were transduced with lentivirus containing either *MYBL2*_sg1 (*MYBL2*_sgRNA1) plasmid construct, *MYBL2*_sg4 (*MYBL2*_sgRNA4) plasmid construct, NTC1(no target control 1), NTC2 (no target control 2), or empty vector (vector). Two days after transduction, cells were subjected to puromycin selection for three days to single out the positively transduced. At the end of the selection, the cells were then plated in a 6-well plate and incubated for 7 days before their protein was isolated. Proteins were measured with a BCA assay and then subjected to Western blot.

Based on the Western blot result, in S2-028, the bands of MYB Proto-Oncogene Like 2 or MYBL2 protein (encoded by *MYBL2* gene) only appeared in NTC1, NTC2, and vector, and not in the *MYBL2* knockouts samples (*MYBL2*_sg1 to *MYBL2*_sg5), indicating complete knockout of *MYBL2*. In S2-007, the control bands (α-tubulin) appeared in all samples and controls. MYBL2 protein band did not that appear in the *MYBL2*_sg1, indicating complete knockout. The rest of the *MYBL2* knockout showed smear bands even though the intensity of the bands appeared lower than the controls. *MYBL2*_sg1 and MYBL_sg4 were then selected for further analysis based on the Western blot result (Fig. 3.28).

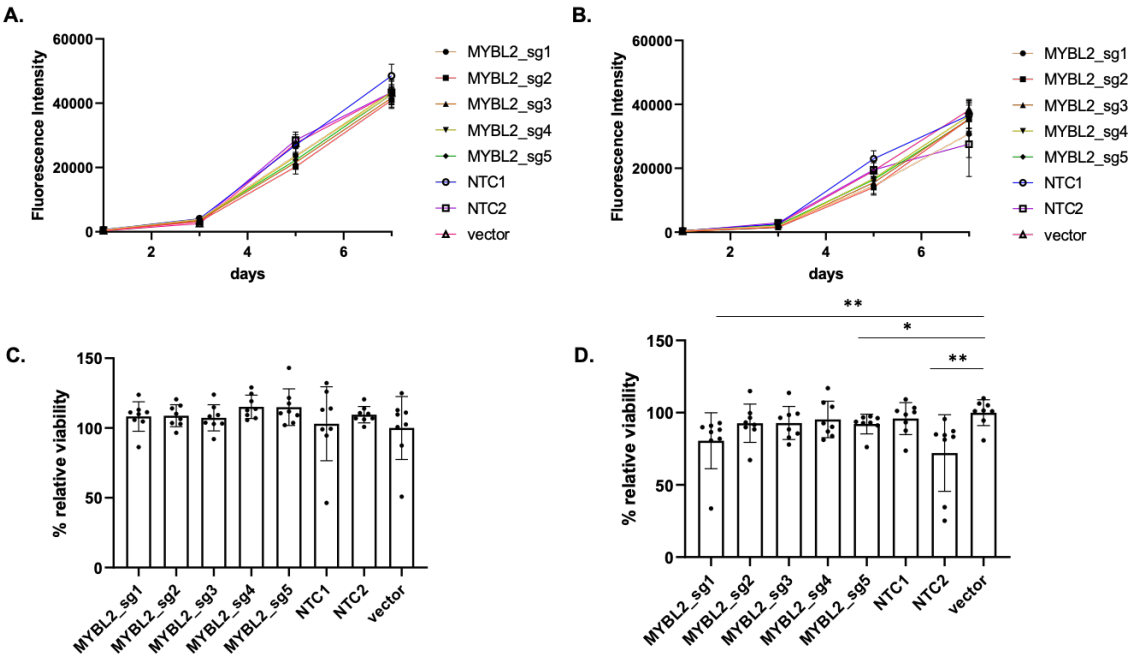


Figure 3.27. *MYBL2* knockout decreased cell proliferation and cell viability in S2-007 based on resazurin assay. Cells were transduced, selected, and cultured for seven days and subjected for viability assay using resazurin every 24 hours. Fluorescence signals following resazurin assay were measured. **(A, B)** Fluorescence values from the resazurin assay were collected and then plotted against the day of data collection to generate a cell proliferation plot. S2-028 cell proliferation plot is shown on the left side (A), and S2-007 on the right side (B). Each plot shows different groups tested, which are *MYBL2*_sg1, *MYBL2*_sg2, *MYBL2*_sg3, *MYBL2*_sg4, *MYBL2*_sg5, NTC1, NTC2, and vector (n=8). **(C, D)** Cell viability on day 7 was analyzed further by normalizing the fluorescence values of the *MYBL2* knockout groups (*MYBL2*_sg1- *MYBL2*_sg5) and controls (NTC1 and NTC2) to the vector. The bar graph was then constructed with the groups' names on the x-axis and relative viability (%) on the y-axis. The bar graphs above are shown for S2-028 (C) and S2-007 (D) samples. * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

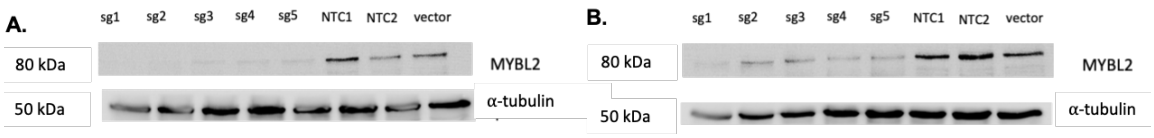


Figure 3.28. *MYBL2* sgRNA1 and sgRNA4 plasmid constructs successfully knocked out the *MYBL2* protein in S2-028 and S2-007. 20 µg of protein of each group were loaded to the 10% SDS-PAGE gel. The gel was run in parallel with Spectra Multicolor Broad Range Protein marker, 135 Voltage. The protein on the SDS-PAGE gel was then transferred to a nitrocellulose membrane using a semi-dry transfer method. The membrane was washed and incubated with the primary and secondary antibodies for *MYBL2* and alpha-tubulin (α-tubulin). The Western blot bands were then visualized by adding HRP substrate, and an image was taken directly using the Intas ECL ChemoStar imager. **(A, B)** The Western blot picture shows the bands of α-tubulin (50 kDa) and MYBL2 protein (80 kDA) on day 7 in S2-028 (A) and S2-007 (B).

Cell viability was then analyzed using flow cytometry. Cells were transduced and were subjected to puromycin selection. At the end of the selection, cells were mixed with the non-transduced cells with a ratio of 4:1. Cells were plated in a 6-well plate, and incubated and maintained for 7 days, after which cells were harvested and collected for flow cytometry analysis. The leftover cells were then plated in a new 6-well plate for another 7-day incubation until day 14 and 21. Flow cytometry was conducted on day 0, 7, 14, and 21 to monitor the percentage of transduced cells that contain mCherry protein.

Bar graphs were then made for both samples to observe the difference in cell viability between different groups (Fig 3.29). Based on the result, the *MYBL2* knockouts (*MYBL2*_sg1 and *MYBL2*_sg4) significantly decreased the cell viability compared to the controls (NTC1 and NTC2) and vector, in S2-007, and in S2-028. However, the decrease of cell viability in S2-007 relatively higher than in S2-028.

Furthermore, to investigate the ability of the *MYBL2* knockout cells to grow colonies on soft agar, a colony formation assay was conducted. Cells were firstly transduced with lentivirus containing either *MYBL2*_sg1 (*MYBL2*_sgRNA1) plasmid construct, *MYBL2*_sg4 (*MYBL2*_sgRNA4) plasmid construct, or NTC (no target control). Two days after transduction cells were subjected to puromycin selection for three days to single out positively transduced cells. At the end of the selection, cells were then plated in a 6-well plate on soft agar. Cells were then maintained in the incubator for 15 days.

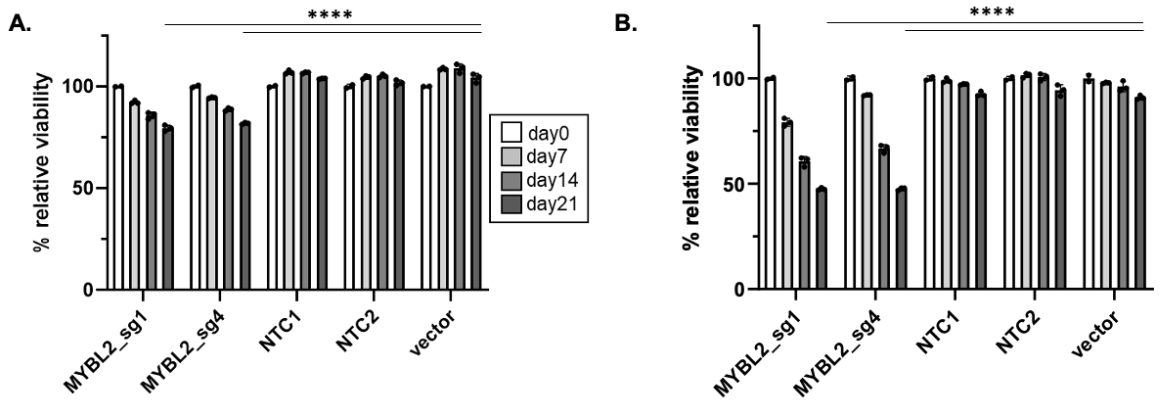


Figure 3.29. *MYBL2* knockout decreased cell viability in S2-007 and S2-028 based on a flow cytometry assay. Cells were transduced, selected, and cultured for 21 days in total and subjected for viability assay using flow cytometry every 7 days. Flow cytometry was conducted on day 0, 7, 14, and 21 to monitor the percentage of transduced cells that contain mCherry. The number of mCherry positive cells represents the viability of transduced cells over the incubation period. The percentage of each time point were then normalized with the value on day 0. **(A, B)** The bar graph was then constructed by plotting the groups' names on the x-axis to the percentage of relative viability on the y-axis (n=3) for samples S2-028 (A) and S2-007 (B). * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

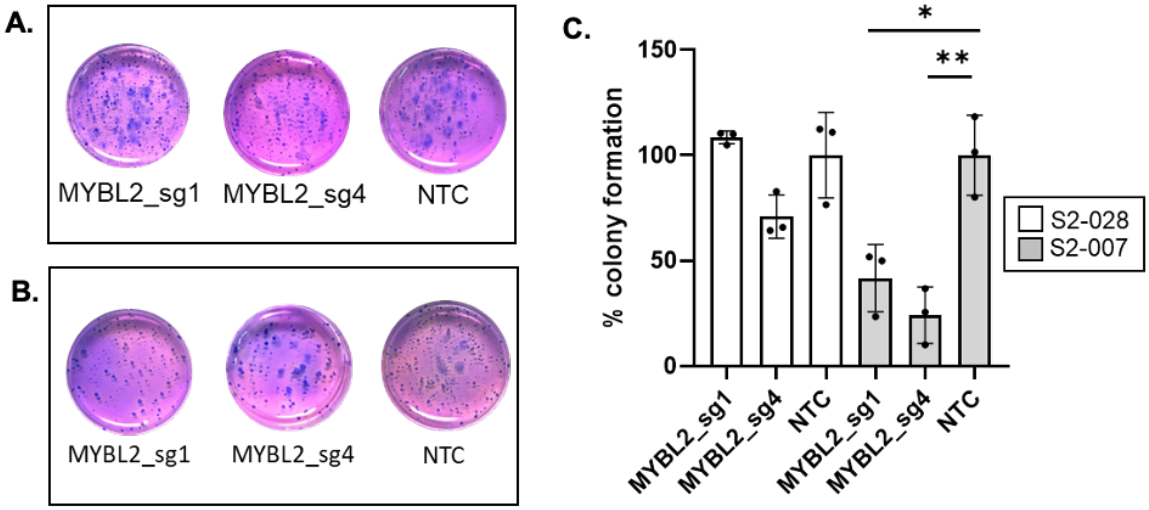


Figure 3.30. *MYBL2* knockout decreased colony formation in S2-007. Cells were transduced, selected, and cultured on soft agar for 15 days of incubation. On day 15, cells were washed with PBS and then stained with crystal violet. **(A, B)** The picture of each well was taken for S2-028 (A) and S2-007 (B). The cells, stained with crystal violet, appear darker than the background. The picture was then subjected for analysis using Open CFU software to count the number of cells. The number of cells of each group was then normalized with control to obtain the percentage of colony formation ratio. **(C)** A bar graph was made by plotting colony formation (%) on the y-axis against the group's name on the x-axis (n=3). * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

Visible cells were stained and counted (Fig. 3.30). Based on the result, the colony formation of *MYBL2* knockouts (*MYBL2*_sg1 and *MYBL2*_sg4) were lower than the control (NTC) in S2-028 and S2-007. Next, a bar graph was composed to compare the percentage of colony formation between sample groups. Based on the graph, the *MYBL2* knockouts (*MYBL2*_sg1 and *MYBL2*_sg4) significantly reduced the colony formation in S2-007.

3.6 *MYBL2* knockout inhibited invasion in metastatic cell line (S2-007)

Based on the cell proliferation, cell viability, and colony formation assays, *MYBL2* was the most promising candidate gene to show gene essentiality only in S2-007, and not in S2-028. Further investigation was then conducted to observe the ability of *MYBL2* in the migration and invasion process.

Cells were transduced and were subjected to puromycin selection for three days to single out positively transduced cells. At the end of the selection, cells were plated into the Ibidi wound healing chamber and followed by overnight incubation. After incubation, the chamber was removed, making the gap between attached cells visible. Cells were then cultured in the medium containing 1% FBS. Images of the gap were taken at three different locations on each

wound healing chamber on day 0 before cells were incubated for 3 days. After incubation, an image was retaken at the exact location.

Based on the pictures on day 0 and day 3, the gap closing was observed not only in S2-028, but also in S2-007. The quantification of the gap closing activity was then measured by Image J, and then the result was used to calculate the percentage of wound closure, which was plotted as a bar chart (Fig. 3.31). According to the result, there is no significant difference between *MYBL2* knockout samples (*MYBL2*_sg1 and *MYBL2*_sg4) and NTC in S2-028 and S2-007.

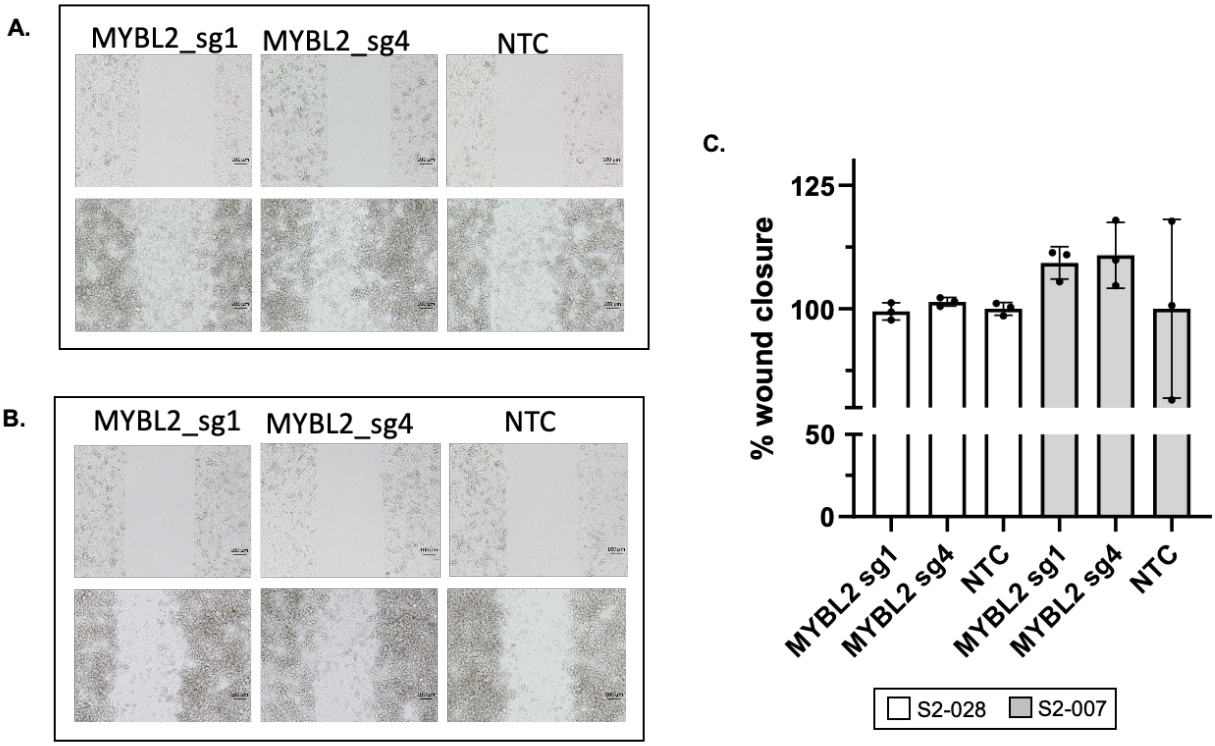


Figure 3.31. No effect of *MYBL2* knockout in cell migration in S2-028 and S2-007. Cells were transduced, selected, and plated into the Ibidi wound healing chamber. 1% FBS containing medium was then used during the incubation. Next, cells were incubated for migration assay for 3 days. Several pictures were then taken on day 0 and after 3 days of incubation at different places on the chamber to observe the gap closing. **(A, B)** Images were taken using an inverted microscope with a magnification of 100x. The scale bar was added, showing 100 µm. The image of *MYBL2*_sg1, *MYBL2*_sg4, and NTC from S2-028 (A) and S2-007 (B) on day 0 (upper images) and day 3 (below images). ImageJ was used to analyze the gap to calculate the area without any cells (free cells area) from the images. Wound closure ratio (%) was further calculated using those values by subtracting the free cells area on day 0 to day 3 and then divided by value on day 0 (n=3). **(C)** A graph was then made by plotting the percentage of wound closure on the y-axis against the group's name on the x-axis. * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

Next, knockout cells were subjected to invasion assay. Cells were transduced with lentivirus containing either *MYBL2*_sg1 (*MYBL2*_sgRNA1) plasmid construct, *MYBL2*_sg4 (*MYBL2*_sgRNA4) plasmid construct, or NTC (no target control). Two days after transduction, cells were subjected to puromycin selection for three days to single out positively transduced cells. At the end of the selection, cells were then cultured with the medium without FBS overnight, followed by harvesting the next day and plating into the upper layer of cell culture insert with the permeable membrane (8 µm membrane) coated with Matrigel. Medium containing FBS was then placed below the cell-permeable membrane. Cells were incubated for 18 hours, and then the insert was washed with PBS. The upper layer of the insert was then cleaned to remove all the remaining cells. The invasive cells, located below the insert, were then washed, stained with 0.5% crystal violet, and observed under an inverted microscope.

The images showed a clear cell number between the knockout groups and the control in S2-007, but not in S2-028. The cell numbers were then observed, counted, and then used to calculate the percentage of relative invasion. The result was then plotted into a bar graph to show the various invasions' ability in tested groups. Based on this observation, the *MYBL2* knockouts (*MYBL2*_sg1 and *MYBL2*_sg4) significantly inhibited the invasion in S2-007 (p value < 0.05) (Fig. 3.32).

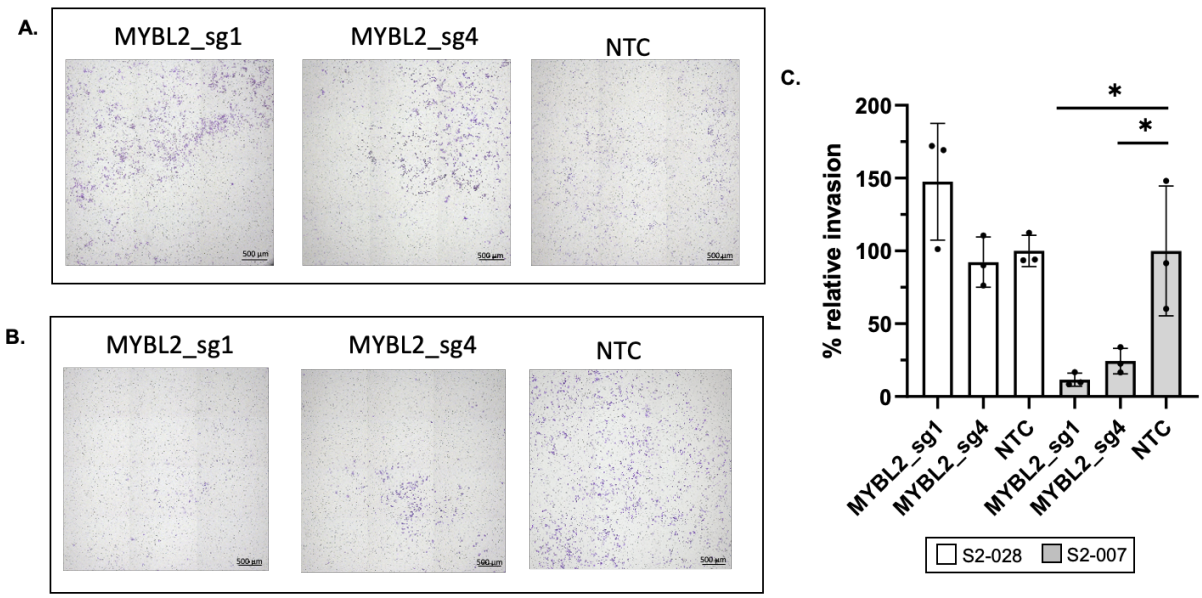


Figure 3.32. *MYBL2* knockout inhibited invasion in S2-007. For legend, see next page.

Figure 3.32. *MYBL2* knockout inhibited invasion in S2-007. Cells were transduced, selected, and then incubated with the medium without FBS. Then, cells were harvested and plated into the upper layer of cell culture insert with the permeable membrane (8 µm membrane) coated with Matrigel. Medium containing FBS was then placed below the cell-permeable membrane. Next, cells were incubated, and after 18 hours, cells below the membrane were washed and stained for further observation. (A, B) Images were taken using an inverted microscope with a magnification of 100x. Several images were taken to cover most of the insert surface area. These images were then stitched into one image using the software. The scale bar was added to the final combined image, showing 500 µm. Images of *MYBL2*_sg1, *MYBL2*_sg4, and NTC groups from S2-028 (A) and S2-007 (B) were taken. The picture was then subjected for analysis using Open CFU software to count the number of cells. The number of cells of each group was normalized with the control (NTC) to obtain a percentage of relative invasion ratio. (C) A bar graph was then made by plotting relative invasion (%) on the y-axis against the group's name on the x-axis (n=3). * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001.

3.7 *MYBL2* involves in cell cycle process in the pathway analysis

Based on the KEGG pathway analysis in Fig. 3.11, the cell cycle was among the pathways discovered. Further analysis was then done using Pathview in R to visualize which gene was altered in this pathway based on the negatively selected genes result. The Pathview showed the KEGG cell cycle pathway overlay by the beta-score (Fig. 3.33). Several genes that were considered negatively selected genes in the CRISPR screening, such as *E2F1,2,3*, *Wee*, *CycB*, and *APC/C*, were shown in the KEGG cell cycle pathway.

In addition, further analysis was conducted using Ingenuity, showing that *MYBL2* is involved in the cell cycle network. Based on the interaction in the network, *MYBL2* has either direct or indirect interaction with such as *CCNB1*, *UBE2C*, *Cbp/p300*, *E2f,* Cyclin E, Cyclin A, *FOXM1*, and *Cdc2* (Fig. 3.34). The network shows that *E2f* directly activates *MYBL2*, whereas Cyclin A and Cyclin E have indirect activation. Moreover, *MYBL2* directly activates *UBE2C*, *Cdc2*, *CCNB1*, and *FOXM1*, according to the data. Four out of eight molecules interacting with *MYBL2* showed green on the graph, meaning that these molecules have a negative beta score with a significant p value (p value < 0.05); they are *CCNB1*, *Cbp/p300*, *E2f*, and *UBE2C*.

To know more about the role of *MYBL2* in cell networks, further study was conducted mainly to understand the effect of *MYBL2* knockout in the cell cycle. Based on the network, a strong connection was shown between MYBL2 and FOXM1 proteins. MYBL2 activates FOXM1 and *CCNB1*, and FOXM1 activates *CCNB1*. Moreover, Cyclin B1 protein (encoded by *CCNB1*) is essential in the G2/M phase (Otto & Sicinski, 2017).

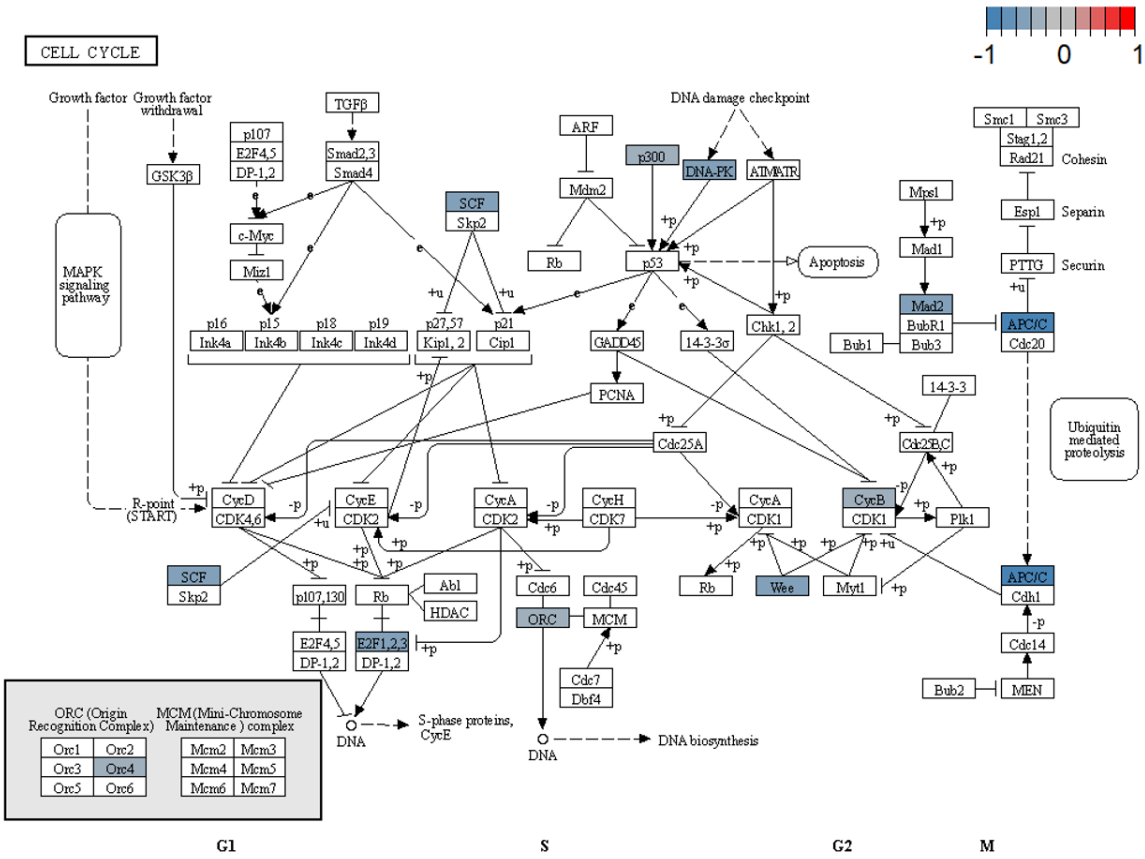


Figure 3.33. Negatively selected genes shown within the KEGG cell cycle pathway. The significantly negative genes with their beta-score were then analyzed using Pathview in R software to visualize the KEGG cell cycle pathway plotted with the corresponding beta-score on the genes. The color represents the beta score. The blue and red color gradient represents a -1 to 1 scale, respectively, on the legend above the graph.

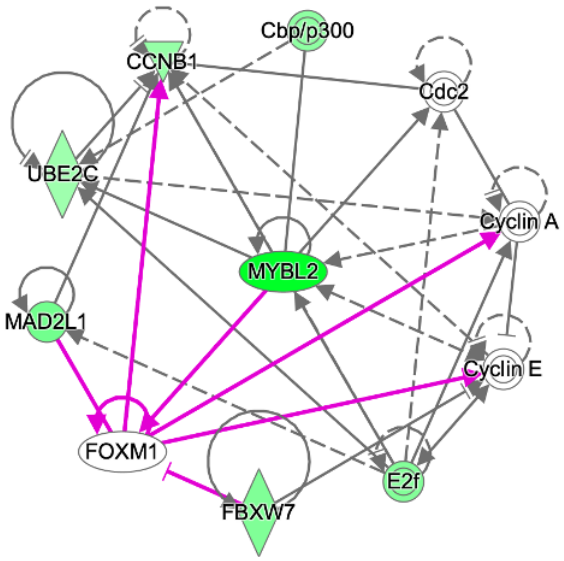


Figure 3.34. *MYBL2* interaction based on Ingenuity's network analysis. The CRISPR screening results of *MYBL2* were analyzed using Ingenuity. The network shows the interactions of the molecules. Pink lines highlight the interactions of *FOXM1*. Green coloring indicates decreased measurement; solid lines stand for direct, broken lines for indirect interaction.

3.8 *FOXM1-MYBL2* knockouts induce cell cycle arrest at G2/M phase in S2-007

The study was followed by in vitro study to know the role of *MYBL2* in the cell cycle. The additional gene, *FOXM1*, was also included in this study because, based on the network, *FOXM1* has several direct connections in cell cycle network, such as to Cyclin A, Cyclin E, and *CCBN1*.

Cells were transduced with lentivirus containing sgRNA targeting *FOXM1-MYBL2* (Group 1), *FOXM1* (Group 2), *MYBL2* (Group 3), and controls NTC (Group 4). For this purpose, the plasmid pAW12 and pAW13 were prepared with each sgRNA combination which carries different fluorescence. pAW12 has GFP, and pAW13 carries mCherry. Two days after transduction, cells were sorted to obtain a double-positive transduction, meaning cells that expressed both GFP and mCherry. After the sorting, cells were cultured for three days and were subjected to various downstream experiments such as Western blot, viability assay, and cell cycle.

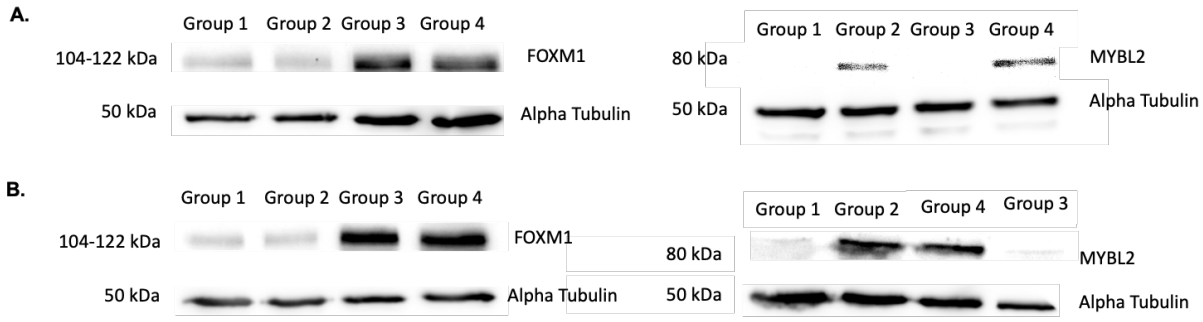


Figure 3.35. *MYBL2* and *FOXM1* knockout validation in S2-028 and S2-007. 20 µg of protein of each group were loaded to a 10% SDS-PAGE gel. The gel was run in parallel with Spectra Multicolor Broad Range Protein marker, 135 Voltage. The protein on the SDS-PAGE gel was then transferred to a nitrocellulose membrane using a semi-dry transfer method. The membrane was then washed and incubated with the primary and secondary antibodies for FOXM1, MYBL2, and alpha-tubulin (α-tubulin) proteins. The Western blot bands were visualized by adding HRP substrate, and then an image was taken directly using the Intas ECL ChemoStar imager. **(A, B)** The Western blot picture shows the bands of α-tubulin (50 kDa), FOXM1 (104-122kDa), and MYBL2 (80 kDa) proteins for each groups in S2-028 (A) and S2-007 (B) respectively. Group 1 = *FOXM1-MYBL2* knockouts, Group2 = *FOXM1* knockout, Group 3 = *MYBL2* knockout, and Group 4 = NTC.

To confirm the knockout of *FOXM1*, *MYBL2*, or both, Western blot was performed. Proteins were isolated from a 6-well plate and then measured with a BCA assay before being subjected to a Western blot. Based on the FOXM1 protein Western blot, a smear or weak bands appear in Group 1 and Group 2 but are intact in Group 3 and Group 4 (104-122 kDa), indicating *FOXM1* knockout in Group 1 and Group2 but not Group 3 and Group 4, as expected. The

second validation is the *MYBL2* knockout. No bands or smear bands appeared in Group 1 and Group 3. Meanwhile, strong bands were shown in Group 2 and Group 4, indicating *MYBL2* knockout (Fig. 3.35).

Next, the cell viability was examined by flow cytometry. The bar graph was made to observe the relative viability (%) based on the GFP and mCherry expression. Based on the chart, the cells viability in Group 1, 2, and 3 were significantly reduced compared to Group 4 (control) in both S2-028 and S2-007 (Fig 3.36. A-B). Moreover, the observation of the cell viability on day 21 showed that the double knockout *FOXM1-MYBL2* reduced the cell viability significantly in S2-028 and S2-007 compared to the single knockout of *FOXM1* or *MYBL2* (Fig. 3.36C).

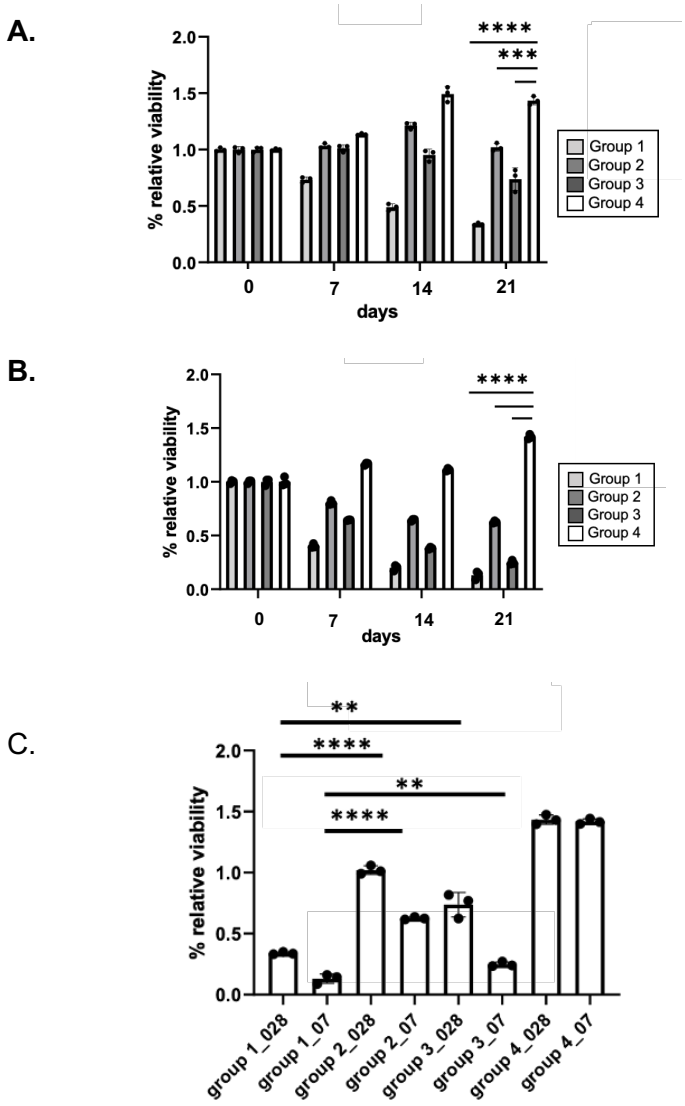


Figure 3.36. *FOXM1-MYBL2* knockouts decrease cell viability in S2-028 and S2-007 based on flow cytometry assay. Cells were transduced, sorted, and cultured for 21 days in total. A viability assay by flow cytometry was conducted on day 0, 7, 14, and 21 to monitor the percentage of transduced cells. The number of mCherry-positive cells represents the viability of transduced cells over the incubation period. The percentage of each time point was normalized with the value on day 0. **(A, B)** The bar graph was constructed by plotting the groups' names on the x-axis to the percentage of relative viability on the y-axis (n=3) for S2-028 (A) and S2-007 (B). **(C)** Relative viability bar graph to compare relative cell viability on day 21 between S2-028 and S2-007. * = p value ≤ 0.05; ** = p value ≤ 0.01; *** = p value ≤ 0.001; **** = p value ≤ 0.0001. Group 1 = *FOXM1-MYBL2* knockouts, Group2 = *FOXM1* knockout, Group 3 = *MYBL2* knockout, and Group 4 = NTC.

The cell cycle assay was then performed to know the impact of the knockouts during a corresponding cell cycle phase. After the sorting, cells were plated in a 6-well plate and then subjected to cell cycle assay preparation and stained with Propidium Iodide (PI). PI intercalates DNA generating a highly fluorescent signal. By observing the signal, the DNA content of a cell can be known (Krishan, 1975). The results showed no difference in cell cycle phase distribution between each group in S2-028 (Fig. 3.37. A-B). In contrast, there was a difference in S2-007. Group 1, 2, and 3 showed more G2/M than Group 4 (NTC) (Fig. 3.37 C-D).

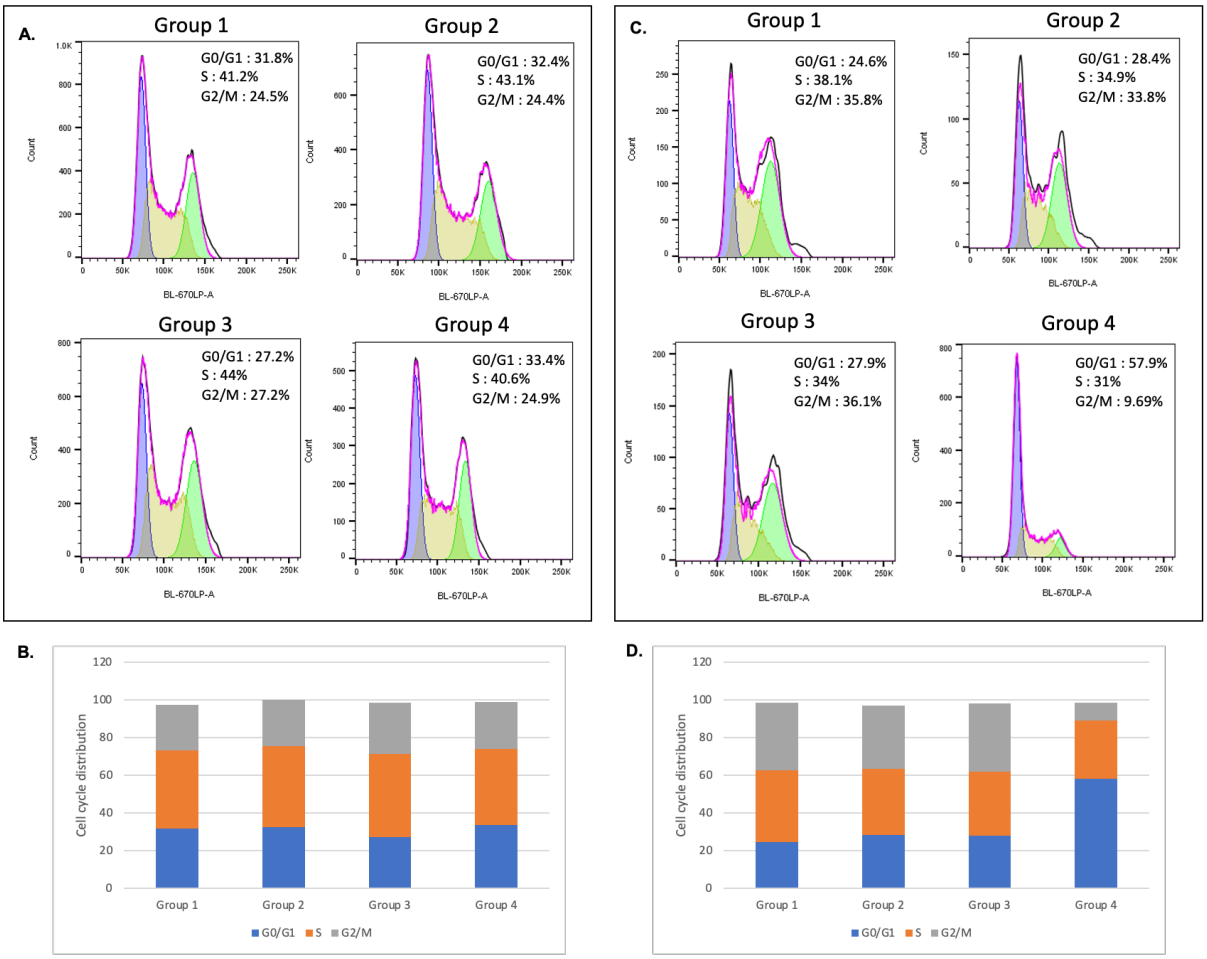


Figure 3.37. Cell cycle phase distribution in S2-028 and S2-007. Cells were transduced, sorted, and cultured for 3 days and subjected to cell cycle assay using flow cytometry. Propidium iodide staining of S2-028 and S2-007 cells are presented. Distribution of S2-028 **(A,B)** and S2-007 **(C,D)** cells among G0/G1, S, and G2/M are shown. Group 1 = *FOXM1-MYBL2* knockouts, Group2 = *FOXM1* knockout, Group 3 = *MYBL2* knockout, and Group 4 = NTC.

4. DISCUSSION

The complex genetic alteration in metastatic pancreatic cancer causes several challenges in obtaining a clear understanding of this disease progression and setting treatment approaches to eliminate it. CRISPR-Cas9 technology begins to open new ways to discover a comprehensive way to unravel the complex biological processes by providing a possibility for making precise cellular and animal model perform functional genomic screens (Hsu et al., 2014). Specifically, the whole genome screening approach using the CRISPR-Cas9 system showed its capability as a powerful tool to discover novel essential genes or new gene targets for drug susceptibility in several types of cancer (Bakke et al., 2019; K. Huang et al., 2019; Shalem et al., 2014a; Shi et al., 2015; Zheng et al., 2019).

Essential genes are genes that affect cell survival and proliferation. In the context of this study, an essential gene in metastatic pancreatic cancer is a gene that, when depleted, affects cell viability of the metastatic cells. Since the beginning of the disease progression, metastatic cells required more genetic alterations than the primary tumors, indicating differences in gene profile between metastatic and primary tumors cells. Moreover, the complex mechanism involves the metastatic process, such as tumor microenvironment adaptation showing there is cell plasticity in cancer cells. Cell plasticity is the source of tumor heterogeneity. There are two types of tumor heterogeneity, namely inter-tumor heterogeneity and intra-tumor heterogeneity (Alizadeh et al., 2015; Liu et al., 2018; Lüönd et al., 2021; McGranahan & Swanton, 2017). The two cell lines, S2-028, and S2-007 have proven to be intra-tumor heterogeneity in metastatic pancreatic cancer. These cells originated from metastatic pancreatic cancer to the liver, SUIT-2 cell line. Based on the previous study, these two cell lines have different metastatic potential, S2-028 has low metastatic potential compared to S2-007 during in vivo experiments (Taniguchi et al., 1994). These two cells are helpful as a model to discover the essential genes in metastatic pancreatic cancer by comparing genes with two different metastatic characteristics that originated from the same metastatic cell line. By doing this, the genetic background of these two cells will be more similar compared to two cell lines with different metastatic characteristics. Thus, the genetic background bias remains minimum.

In this study, whole-genome CRISPR-Cas9 screening was done to discover the essential gene in metastatic pancreatic cancer. Analysis of gene essentiality was then conducted using the MAGeCK algorithm. The output of the analysis, a beta-score or gene essentiality is a score used to show the level of gene essentiality in metastatic versus non-metastatic cells. Beta-score results lower than zero would indicate that the gene is negatively selected, and vice versa for

the positive score. Negatively selected genes mean that these genes are low in abundance in the metastatic cells (W.Li et al., 2014, 2015). Thus, these genes could be essential because when these genes lose function, the cells do not survive until the end of the screening, resulting in low cell abundance at the end of the screening.

The KEGG pathway analysis of the negatively selected genes showed that these genes are related to ribosome, spliceosome, and ribosome biogenesis. According to the previous study, these pathways were also revealed in screenings to discover essential genes. Moreover, the finding of ribosomal gene pathways in negatively selected gene group indicated good quality screening (Blomen et al., 2015; H. Y. Li et al., 2015). GO analysis from the results also confirmed that the genes in both negatively and positively selected genes are located mostly in nucleoplasms and nucleus (cellular component- GO analysis), protein binding and poly (A) binding (molecular function- GO analysis), and translational (biological process-GO analysis). Based on the GO results, these results confirmed the KEGG pathway analysis.

The three candidates of essential genes, which are *TYMS*, *MEN1*, and *MYBL2*, were selected based on analysis after the screening was conducted. According to the Online Gene Essentiality (OGEE) database, *TYMS*, *MEN1*, *MYBL2* are listed as essential genes in humans. Furthermore, the database also shows that *MYBL2* and *MEN1* are the genes listed as essential genes in mice and humans (W. H. Chen et al., 2012; R. Zhang et al., 2004). The essentiality of genes in several systems such as cell lines, mice, or humans could be different. There are several factors to be considered when observing the gene essentiality across these systems, such as variable ploidy in cell lines, the viability of heterozygous animals when used as a model, and the presence of haploinsufficiency in the human population associated with heterozygous genotypes (Bartha et al., 2018).

TYMS encodes thymidylate synthase (TS) and plays a role in DNA synthesis. Based on previous studies, *TYMS* was associated with poor overall survival in several different cancer, such as pancreatic cancer and retroperitoneal liposarcoma (Fu et al., 2019; S. Zhang et al., 2020). In metastatic pancreatic cancer, 5-Fluorouracil (5-FU) was used with oxaliplatin, irinotecan, and leucovorin (Conroy et al., 2011). 5-FU active metabolite binds to TS and inhibits deoxythymidine monophosphate (dTMP) synthesis, which causes deoxyribonucleotide triphosphate (dNTP) disproportion and increased deoxyuridine triphosphate (dUTP). Both events lead to DNA damage in the cells (Longley et al., 2003). Therefore, the mRNA expression and protein level of TS, as well as the variation of *TYMS* gene based on several studies, correlated with 5-FU sensitivity in cancer (Chao & Anders, 2018; Cho et al., 2011;

Luo et al., 2014). In this study, *TYMS* knockout decreased cell proliferation, cell viability, and colony formation in S2-028 and S2-007 cell lines, non-metastatic and metastatic cell lines, respectively. Even though the effect of *TYMS* knockout in S2-007 cells was more than in the S2-028, the comparison between the control showed the significant difference between the knockouts and control in both cell lines. The results indicated that *TYMS* is an essential gene in both cell lines, which is nonspecific only for the metastatic cell line. According to previous studies, the loss of function of *TYMS* causes DNA damage, which can trigger a DNA damage response, one of which is apoptosis potentially leading to cell death (Matt & Hofmann, 2016; Roos et al., 2016). The effect of cell death caused by the *TYMS* knockout was immediate. The Western blot confirmed it at day 3, when the complete knockout of *TYMS* protein was shown. However, on day 7, the *TYMS* protein appeared once again, caused by the complete *TYMS* knockout cells not surviving and being replaced by the wild-type cells.

The second candidate is *MEN1*. Menin, the product of the *MEN1* gene, is a transcription factor to histone-modifying protein complexes. Protein-protein interactions study showed that most of the proteins which interact with Menin are transcriptional regulators and chromatin modifiers, such as Jun, SMAD, Forkhead (FOX2), nuclear receptor, and beta-catenin. Menin works as a bridge that connects transcription factors and chromatin-modifying protein complexes during transcription initiation. This connection helps to activate initiation when the transcription factors bind to the enhancers located far from the transcription start sites (TSS). For instance, Menin interacts with MLL1 and MLL2, which have H3K4me methyltransferase activity, so that the transcription of the target genes is initiated (Dreijerink et al., 2017; J. Huang et al., 2012; Jiang et al., 2014). The study showed that *MEN1* knockout decreased cell proliferation and cell viability in S2-007 cell lines and colony formation in S2-028 and S2-007. The cell proliferation and cell viability confirmed the screening result that *MEN1* is essential to S2-007 growth. However, the cell colony formation in soft agar showed that the colony formation after *MEN1* knockout was reduced in both cell lines. Soft agar colony formation assay aims to evaluate the ability of the cells to grow in an an-anchorage-independent manner, which indicates carcinogenic properties (Borowicz et al., 2014). Moreover, the colony formation assay evaluates the phenotypic effects that usually require time and several cell divisions (Franken et al., 2006). According to those results, the *MEN1* knockout affected S2-007 and S2-028 cells.

MYBL2 or *B-Myb* gene encodes MYBL2 protein. MYBL2 is a transcription factor of the MYB transcription factor family and has a role in cell cycle progression, cell survival, and cell

differentiation (Musa et al., 2017). In the cell cycle, MYBL2 is related to the DREAM complex (LIN9, LIN37, LIN52, LIN54, and RBBp4 protein), which is the master regulator in the cell cycle progression. During the early phase of the cell cycle (G0 and G1), the DREAM complex binds E2F4-5/DP and p107/p130, CHR (cell cycle genes homology region), CDE (cell cycle-dependent element), and CLE (CHR-like element) to repress transcription. LIN54 is the complex that binds to the CHR. In late G1 and early S, the E2F4-5/DP and p107/p130 dissociate from the complex, leading to the binding of MYBL2 protein, forming the *B-Myb*-MuvB (MMB) complex. The formation of the MMB complex activates the transcription (Cicirò & Sala, 2021; Engeland, 2018; Musa et al., 2017). The balance between the DREAM and MMB complexes is an important aspect of cell cycle regulation. The overexpression of each component of the complex causes alterations in the process that leads to abnormality, therefore the disparity between those two often occurs in cancer. In several different cancer, such as breast cancer, renal cell carcinoma, and lung cancer, the expression of MYBL2 protein correlates with metastasis (Bayley et al., 2020; Jin et al., 2017; Sun et al., 2020). The result shows that *MYBL2* knockout decreased cell proliferation, cell viability, and colony formation in the S2-007 cell line, thus confirming the essentiality of *MYBL2* in metastatic cells.

According to the result above, the most promising candidate is *MYBL2*. The inhibition of cell proliferation, cell viability, and colony formation is specific only in S2-007, the metastatic cell line. A previous study in lung cancer showed that the knockdown of *MYBL2* inhibits cell proliferation, migration, and invasion (Jin et al., 2017; Xiong et al., 2020). Further study was then conducted to determine whether the loss of function of *MYBL2* would affect the wound healing and invasion of the cells. Based on the results, the knockout of *MYBL2* showed invasion inhabitation in S2-007, but no difference was observed in wound healing assay. In the wound healing assay, cells were grown on an uncoated plastic chamber and without serum, which kept the proliferation speed of the cells to a minimum. According to a previous study, there was no significant difference in terms of the ability of the cells to be attached to the uncoated dish. On the contrary, when the cells were plated on the Matrigel coated dish, the S2-007 cells were more able to attach compared to the S2-028 cells. Moreover, to explain the invasion potential, the S2-007 has more type IV collagenolytic ability than S2-028. Therefore, the S2-007 can better pass the Matrigel-coated membrane than S2-028 (Taniguchi et al., 1994).

Based on the interaction of *CCNB1*, *UBE2C, Cbp/p300*, *E2f,* Cyclin E, Cyclin A, *FOXM1*, and Cdc2, *MYBL2* is mainly involved in the cell cycle network. To validate the effect of *MYBL2* knockout in the cell cycle, further experiments were done. According to the cell cycle network

pathway analysis, MYBL2 activates *CCNB1*. *CCNB1* plays a role in the cell cycle, especially in the G2/M phase as a protein checkpoint (Engeland, 2018; Musa et al., 2017). To check the role of *MYBL2* in the cell cycle, especially in G2/M transition, the additional candidate was then added. *FOXM1* (Forkhead Box M1) gene encodes FOXM1 protein. FOXM1 is a transcription factor that regulates several genes in the G2/M transition (Chen et al., 2013; Liao et al., 2018). The *FOXM1-MYBL2* double knockout experiments induce cell cycle arrest at G2/M transition in S2-007 cells. Moreover, the cell viability assay confirmed that the double knockout decreases the cell viability more than the single knockout of *FOXM1* and *MYBL2*. The viability assay also confirmed that the knockout of *MYBL2* caused more cell death than the knockout of *FOXM1* alone. The result verified previous studies that the binding of FOXM1 protein to G2/M gene promoters is dependent upon MYBL2. FOXM1 binds the promoter via CHR (cell cycle genes homology region) elements via protein-protein interactions with the MMB transcriptional activator complex-forming FOXM1-MMB. This complex occurs during the late S or early G2 phase. Later in the G2/M phase, MYBL2 dissociates from the complex and leaves only the FOXM1-MuVB complex. MMB complex, *FOXM1*-MMB, and *FOXM1*-MuVB complexes are the activators of late-cycle genes (Chen et al., 2013; Down et al., 2012; Engeland, 2018).

In this study, it is still unclear why the cell cycle arrest at G2/M by knockout of *MYBL2* was only observed in S2-007. A previous study showed that the downregulation of *FOXM1* and *MYBL2* induced cell cycle arrest and EMT in glioma cells. This study showed that the *MYBL2* is affected by Akt/FOXM1 signaling (X. Zhang et al., 2017). Moreover, a study in breast cancer discovered that the knockdown of *MYBL2* suppressed the cell invasion, anchorage-independent growth, and tumor formation by acting as a regulator of SNAIL, an important EMT key protein (Tao et al., 2015). Further experiments still need to be conducted to discover the process and which pathway is affecting the cell cycle regulation of *MYBL2* in metastatic pancreatic cancer.

5. CONCLUSIONS AND FUTURE WORK

This study aimed at discovering genes that are essential for metastasis of pancreatic cancer by taking advantage of CRISPR-Cas9 whole-genome screening in two pancreatic cancer cell lines with different metastatic potential: S2-028 – low – and S2-007 – high. The following conclusions can be drawn from the study:

1. Using CRISPR-Cas9 whole-genome screening comparing a pancreatic cancer cell line of high metastatic potential (S2-007) and a cell line of low metastatic potential (S2-028), three essential genes were discovered: *TYMS*, *MEN1*, and *MYBL2*.
2. In vitro validation, which consists of cell proliferation, cell viability, and soft agar colony formation assay, showed that *MYBL2* knockout decreased cell proliferation, cell viability, and colony formation of S2-007. In addition, it also inhibited the invasion capacity of S2-007.
3. The *FOXM1-MYBL2* double knockout affects cell viability more than the *FOXM1* or *MYBL2* single knockout.
4. The *FOXM1-MYBL2*, *FOXM1*, and *MYBL2* knockout in S2-007 caused cell cycle arrest at the G2/M.

Even though the G2/M arrest was confirmed to be caused by the *MYBL2* knockout, the reason why the effect was only observed in S2-007 still needs to be elucidated.

ACKNOWLEDGMENT

I would like to express my most profound appreciation to Dr. Jörg Hoheisel. Thank you for your encouragement and patience throughout this project; it would not be possible to reach this point without that support.

To Prof. Dr. Michael Boutros and Prof. Dr. Stefan Wölfl, I would like to extend my deepest gratitude for your guidance and direction in every TAC meeting and in my defense committee. Thank you for your support.

I am also grateful to Prof. Dr. Stefan Wiemann and Prof. Dr. Aurelio Teleman, who are willing to join my defense committee and be part of my journey.

I am very grateful for the scholarship from DAAD (The German Academic Exchange Service). Thank you for providing me the help to continue my study.

Avery, special thanks to Dr. Michael Boettcher; without his advice and help, I could not have finished my project. Thank you so much for all the discussion.

Also, I would like to say thank you to Prof. Malte Buchholz, the DKFZ FACS core facility, and the DKFZ LMF core facility for supporting my project.

I would like to extend my gratitude to my friends and colleagues at the Faculty of Medicine, Universitas Gadjah Mada, especially the Department of Biochemistry. To Prof. Dr. Dra. Sunarti, M. Kes, dr. Arta Fatmawati., Ph.D, and dr. Hamim Sadewa., Ph.D., thank you for your support.

To Prof. dr. Sofia Mubarika., M.Med.Sc., Prof. Dr. dr. Teguh Aryandono., Sp. B (K) Onk, and Dr. med. dr. Indwiani Astuti., thank you for your unparalleled support since the beginning of my career.

Thank you to Anna, Yenan, Beiping, Haoqi, Andrea, Anke, Daniel, Chaoyang, Liang, Everardo, Henning, Melanie, Sandra, and my former students (Angga, Ariani, Noah, and Lili) for helping and supporting me during my work in the laboratory and or during the writing of my dissertation. I appreciate the support and great atmosphere during my study from all of the crew members of the Division of Functional Genome Analysis (B070), DKFZ.

I would not be able to accomplish this milestone without constant support from my family and friends. Your encouragement has given me comfort throughout this journey. I am grateful and appreciative.

Last but not least, without the grace from the almighty Allah S.W.T., this journey would not have been possible. Thank you for giving me the courage to reach my dreams and letting me evolve as a person throughout my life.

"Omnia mutantur, nihil interit"

— Ovid, Metamorphoses

BIBLIOGRAPHY

Aiello, N. M., Maddipati, R., Norgard, R. J., Balli, D., Li, J., Yuan, S., Yamazoe, T., Black, T., Sahmoud, A., Furth, E. E., Bar-Sagi, D., & Stanger, B. Z. (2018). EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration. Developmental Cell, 45(6), 681-695.e4. https://doi.org/10.1016/j.devcel.2018.05.027

Alizadeh, A. A., Aranda, V., Bardelli, A., Blanpain, C., Bock, C., Borowski, C., Caldas, C., Califano, A., Doherty, M., Elsner, M., Esteller, M., Fitzgerald, R., Korbel, J. O., Lichter, P., Mason, C. E., Navin, N., Pe'Er, D., Polyak, K., Roberts, C. W. M., … Zucman-Rossi, J. (2015). Toward understanding and exploiting tumor heterogeneity. In Nature Medicine (Vol. 21, Issue 8, pp. 846–853). Nature Publishing Group. https://doi.org/10.1038/nm.3915

Bacac, M., & Stamenkovic, I. (2008). Metastatic cancer cell. In Annual Review of Pathology: Mechanisms of Disease (Vol. 3, pp. 221–247). https://doi.org/10.1146/annurev.pathmechdis.3.121806.151523

Bafna, S., Kaur, S., Momi, N., & Batra, S. K. (2009). Pancreatic cancer cells resistance to gemcitabine: The role of MUC4 mucin. British Journal of Cancer, 101(7), 1155–1161. https://doi.org/10.1038/sj.bjc.6605285

Bakke, J., Wright, W. C., Zamora, A. E., Oladimeji, P., Crawford, J. C., Brewer, C. T., Autry, R. J., Evans, W. E., Thomas, P. G., & Chen, T. (2019). Genome-wide CRISPR screen reveals PSMA6 to be an essential gene in pancreatic cancer cells. BMC Cancer, 19(253). https://doi.org/10.1186/s12885-019-5455-1

Bartha, I., di Iulio, J., Venter, J. C., & Telenti, A. (2018). Human gene essentiality. Nature Reviews Genetics, 19(1), 51–62. https://doi.org/10.1038/nrg.2017.75

Bayley, R., Ward, C., & Garcia, P. (2020). *MYBL2* amplification in breast cancer: Molecular mechanisms and therapeutic potential. In Biochimica et Biophysica Acta - Reviews on Cancer (Vol. 1874, Issue 2). Elsevier B.V. https://doi.org/10.1016/j.bbcan.2020.188407

Blomen, V., Majek, P., Jae, L. T., & Bihenzahn, J. W. (2015). Gene essentiality and syntheticlethality in haploid human cells. Science, 350(6264), 1089–1092. https://doi.org/10.1126/science.aad3318

Boettcher, M., Tian, R., Blau, J. A., Markegard, E., Wagner, R. T., Wu, D., Mo, X., Biton, A., Zaitlen, N., Fu, H., McCormick, F., Kampmann, M., & McManus, M. T. (2018). Dual gene activation and knockout screen reveals directional dependencies in genetic networks. Nature Biotechnology, 36(2), 170–178. https://doi.org/10.1038/nbt.4062

Borowicz, S., van Scoyk, M., Avasarala, S., Karuppusamy Rathinam, M. K., Tauler, J., Bikkavilli, R. K., & Winn, R. A. (2014). The soft agar colony formation assay. Journal of Visualized Experiments, 92. https://doi.org/10.3791/51998

Brabletz, T., Kalluri, R., Nieto, M. A., & Weinberg, R. A. (2018). EMT in cancer. In Nature Reviews Cancer (Vol. 18, Issue 2, pp. 128–134). Nature Publishing Group. https://doi.org/10.1038/nrc.2017.118

Buchholz, M., Biebl, A., Neee, A., Wagner, M., Iwamura, T., Leder, G., Adler, G., & Gress, T. M. (2003). SERPINE2 (Protease Nexin I) Promotes Extracellular Matrix Production and Local Invasion of Pancreatic Tumors in Vivo 1. In CANCER RESEARCH (Vol. 63).

Burris, H., & Storniolo, A. M. (1997). Assessing Clinical Benefit in the Treatment of Pancreas Cancer: Gemcitabine Compared to 5-Fluorouracil. In European Journal ojCancer (Vol. 33).

Campbell, P. J., Yachida, S., Mudie, L. J., Stephens, P. J., Pleasance, E. D., Stebbings, L. A., Morsberger, L. A., Latimer, C., McLaren, S., Lin, M. L., McBride, D. J., Varela, I., Nik-Zainal, S. A., Leroy, C., Jia, M., Menzies, A., Butler, A. P., Teague, J. W., Griffin, C. A., … Futreal, P. A. (2010). The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature, 467(7319), 1109–1113. https://doi.org/10.1038/nature09460

Carpenter, E. S., Steele, N. G., & di Magliano, M. P. (2020). Targeting the microenvironment to overcome gemcitabine resistance in pancreatic cancer. Cancer Research, 80(15), 3070–3071. https://doi.org/10.1158/0008-5472.CAN-20-1692

Chao, Y. L., & Anders, C. K. (2018). *TYMS* Gene Polymorphisms in Breast Cancer Patients Receiving 5–Fluorouracil-Based Chemotherapy. Clinical Breast Cancer, 18(3), e301–e304. https://doi.org/10.1016/j.clbc.2017.08.006

Chen, W. H., Minguez, P., Lercher, M. J., & Bork, P. (2012). OGEE: An online gene essentiality database. Nucleic Acids Research, 40(D1). https://doi.org/10.1093/nar/gkr986

Chen, X., Muller, G. A., Quaas, M., Fischer, M., Han, N., Stutchbury, B., Sharrocks, A. D., & Engeland, K. (2013). The Forkhead Transcription Factor *FOXM1* Controls Cell Cycle-Dependent Gene Expression through an Atypical Chromatin Binding Mechanism. Molecular and Cellular Biology, 33(2), 227–236. https://doi.org/10.1128/mcb.00881-12

Cho, Y. B., Chung, H. J., Lee, W. Y., Choi, S. H., & Kim, H. C. (2011). Relationship between *TYMS* and ERCC1 mRNA Expression and In Vitro Chemosensitivity in Colorectal Cancer. AntiCancer Research, 31, 3843–3850.

Cicirò, Y., & Sala, A. (2021). MYB oncoproteins: emerging players and potential therapeutic targets in human cancer. In Oncogenesis (Vol. 10, Issue 19). Springer Nature. https://doi.org/10.1038/s41389-021-00309-y

Conroy, T., Ychou, M., Bouché, O., Guimbaud, R., Bécouarn, Y., Adenis, A., Raoul, J.-L., Gourgou-Bourgade, S., de La Fouchardière, C., Bennouna, J., Bachet, J.-B., Khemissa-Akouz, F., Péré-Vergé, D., Delbaldo, C., Assenat, E., Chauffert, B., Michel, P., Montoto-Grillot, C., Chem, M., & Ducreux, M. (2011). FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. In n engl j med (Vol. 364).

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. http://science.sciencemag.org/

Down, C. F., Millour, J., Lam, E. W. F., & Watson, R. J. (2012). Binding of *FOXM1* to G2/M gene promoters is dependent upon *B-Myb*. Biochimica et Biophysica Acta - Gene Regulatory Mechanisms, 1819(8), 855–862. https://doi.org/10.1016/j.bbagrm.2012.03.008

Dreijerink, K. M. A., Timmers, H. T. M., & Brown, M. (2017). Twenty years of menin: emerging opportunities for restoration of transcriptional regulation in *MEN1*. In Endocrine-related cancer (Vol. 24, Issue 10, pp. T135–T145). https://doi.org/10.1530/ERC-17-0281

Engeland, K. (2018). Cell cycle arrest through indirect transcriptional repression by p53: I have a DREAM. In Cell Death and Differentiation (Vol. 25, Issue 1, pp. 114–132). Nature Publishing Group. https://doi.org/10.1038/cdd.2017.172

Evers, B., Jastrzebski, K., Heijmans, J. P. M., Grernrum, W., Beijersbergen, R. L., & Bernards, R. (2016). CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. Nature Biotechnology, 34(6), 631–633. https://doi.org/10.1038/nbt.3536

Franken, N. A. P., Rodermond, H. M., Stap, J., Haveman, J., & van Bree, C. (2006). Clonogenic assay of cells in vitro. Nature Protocols, 1(5), 2315–2319. https://doi.org/10.1038/nprot.2006.339

Fu, Z., Jiao, Y., Li, Y., Ji, B., Jia, B., & Liu, B. (2019). *TYMS* presents a novel biomarker for diagnosis and prognosis in patients with pancreatic cancer. Medicine , 98(51). https://doi.org/10.1097/MD.0000000000018487

Hsu, P. D., Lander, E. S., & Zhang, F. (2014a). Development and applications of CRISPR-Cas9 for genome engineering. In Cell (Vol. 157, Issue 6, pp. 1262–1278). Elsevier B.V. https://doi.org/10.1016/j.cell.2014.05.010

Hsu, P. D., Lander, E. S., & Zhang, F. (2014b). Development and applications of CRISPR-Cas9 for genome engineering. In Cell (Vol. 157, Issue 6, pp. 1262–1278). Elsevier B.V. https://doi.org/10.1016/j.cell.2014.05.010

Huang, J., Gurung, B., Wan, B., Matkar, S., Veniaminova, N. A., Wan, K., Merchant, J. L., Hua, X., & Lei, M. (2012). The same pocket in menin binds both MLL and JUND but has opposite effects on transcription. Nature, 482(7386), 542–546. https://doi.org/10.1038/nature10806

Huang, K., Liu, X., Li, Y., Wang, Q., Zhou, J., Wang, Y., Dong, F., Yang, C., Sun, Z., Fang, C., Liu, C., Tan, Y., Wu, X., Jiang, T., & Kang, C. (2019). Genome-Wide CRISPR-Cas9 Screening Identifies NF-κB/E2F6 Responsible for EGFRvIII-Associated Temozolomide Resistance in Glioblastoma. Advanced Science, 6(17). https://doi.org/10.1002/advs.201900782

Jiang, X., Cao, Y., Li, F., Su, Y., Li, Y., Peng, Y., Cheng, Y., Zhang, C., Wang, W., & Ning, G. (2014). Targeting beta-catenin signaling for therapeutic intervention in *MEN1*-deficient pancreatic neuroendocrine tumours. Nature Communications, 5. https://doi.org/10.1038/ncomms6809

Jin, Y., Zhu, H., Cai, W., Fan, X., Wang, Y., Niu, Y., Song, F., & Bu, Y. (2017). *B-Myb* is up-regulated and promotes cell growth and motility in non-small cell lung cancer. International Journal of Molecular Sciences, 18(6). https://doi.org/10.3390/ijms18060860

Kamisawa, T., Wood, L. D., Itoi, T., & Takaori, K. (2016). Pancreatic cancer. In The Lancet (Vol. 388, pp. 73–85). Lancet Publishing Group. https://doi.org/10.1016/S0140-6736(16)00141-0

Katsuki, T., & Ide, K. (1987). Establishment and Characterization of a Human Pancreatic Cancer Cell Line (SUIT-2) Producing Carcinoembryonic Antigen and Carbohydrate Antigen 19-9 Takeshi IWAMURA. In Jpn. J. Cancer Res. (Gann) (Vol. 78).

Krishan, A. (1975). RAPID FLOW CYTOFLUOROMETRIC ANALYSIS OF MAMMALIAN CELL CYCLE BY PROPIDIUM IODIDE STAINING. http://rupress.org/jcb/article-pdf/66/1/188/1071572/188.pdf

Kuramitsu, Y., Wang, Y., Taba, K., Suenaga, S., Ryozawa, S., Kaino, S., Sakaida, I., & Nakamura, K. (2012). Heat-shock Protein 27 Plays the Key Role in Gemcitabine-resistance of Pancreatic Cancer Cells.

Li, (D, Xie, K., Wolff, R., Abbruzzese, J. L., Li, D., Xie, K., Wolff, R., & Abbruzzese, J. L. (2004). Pancreatic cancer. In THE LANCET • (Vol. 363). www.thelancet.com1049

Li, H. Y., Cui, Z. M., Chen, J., Guo, X. Z., & Li, Y. Y. (2015). Pancreatic cancer: diagnosis and treatments. In Tumor Biology (Vol. 36, Issue 3, pp. 1375–1384). Kluwer Academic Publishers. https://doi.org/10.1007/s13277-015-3223-7

Li, W., Köster, J., Xu, H., Chen, C. H., Xiao, T., Liu, J. S., Brown, M., & Liu, X. S. (2015). Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. Genome Biology, 16(1). https://doi.org/10.1186/s13059-015-0843-6

Li, W., Xu, H., Xiao, T., Cong, L., Love, M. I., Zhang, F., Irizarry, R. A., Liu, J. S., Brown, M., & Liu, X. S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biology, 15(12), 554. https://doi.org/10.1186/s13059-014-0554-4

Liao, G. bin, Li, X. Z., Zeng, S., Liu, C., Yang, S. M., Yang, L., Hu, C. J., & Bai, J. Y. (2018). Regulation of the master regulator *FOXM1* in cancer. In Cell Communication and Signaling (Vol. 16, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12964-018-0266-6

Liu, J., Dang, H., & Wang, X. W. (2018). The significance of intertumor and intratumor heterogeneity in liver cancer. In Experimental and Molecular Medicine (Vol. 50, Issue 1). Nature Publishing Group. https://doi.org/10.1038/emm.2017.165

Longley, D. B., Harkin, D. P., & Johnston, P. G. (2003). 5-Fluorouracil: Mechanisms of action and clinical strategies. In Nature Reviews Cancer (Vol. 3, Issue 5, pp. 330–338). https://doi.org/10.1038/nrc1074

Luo, Y., Li, Z., Cui, S., Shen, C., Zhao, J., Wu, M., Li, Y., Wang, M., Chen, R., Liu, Z., & Ri-Li, G. (2014). Joint detection of ERCC1, TUBB3, and *TYMS* guidance selection of docetaxel, 5-fluorouracil and cisplatin (DDP) individual chemotherapy in advanced gastric cancer patients. European Journal of Medical Research, 19(1). https://doi.org/10.1186/s40001-014-0050-z

Lüönd, F., Tiede, S., & Christofori, G. (2021). Breast cancer as an example of tumour heterogeneity and tumour cell plasticity during malignant progression. In British Journal of Cancer (Vol. 125, pp. 164–175). Springer Nature. https://doi.org/10.1038/s41416-021-01328-7

Matt, S., & Hofmann, T. G. (2016). The DNA damage-induced cell death response: a roadmap to kill cancer cells. In Cellular and Molecular Life Sciences (Vol. 73, Issue 15, pp. 2829–2850). Birkhauser Verlag AG. https://doi.org/10.1007/s00018-016-2130-4

McGranahan, N., & Swanton, C. (2017). Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. In Cell (Vol. 168, Issue 4, pp. 613–628). Cell Press. https://doi.org/10.1016/j.cell.2017.01.018

Michl, P., Barth, C., Buchholz, M., Lerch, M. M., Rolke, M., Holzmann, K.-H., Menke, A., Fensterer, H., Giehl, K., Löhr, M., Leder, G., Iwamura, T., Adler, G., & Gress, T. M. (2003). Claudin-4 Expression Decreases Invasiveness and Metastatic Potential of Pancreatic Cancer 1. In CANCER RESEARCH (Vol. 63).

Moore, M. J., Goldstein, D., Hamm, J., Figer, A., Hecht, J. R., Gallinger, S., Au, H. J., Murawa, P., Walde, D., Wolff, R. A., Campos, D., Lim, R., Ding, K., Clark, G., Voskoglou-Nomikos, T., Ptasynski, M., & Parulekar, W. (2007). Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: A phase III trial of the National Cancer Institute of Canada Clinical Trials Group. Journal of Clinical Oncology, 25(15), 1960–1966. https://doi.org/10.1200/JCO.2006.07.9525

Morris, J. P., Wang, S. C., & Hebrok, M. (2010). KRAS, Hedgehog, Wnt and the twisted developmental biology of pancreatic ductal adenocarcinoma. In Nature Reviews Cancer (Vol. 10, Issue 10, pp. 683–695). https://doi.org/10.1038/nrc2899

Musa, J., Aynaud, M. M., Mirabeau, O., Delattre, O., & Grünewald, T. G. (2017). *MYBL2* (*B-Myb*): a central regulator of cell proliferation, cell survival and differentiation involved in tumorigenesis. In Cell death & disease (Vol. 8, Issue 6, p. e2895). https://doi.org/10.1038/cddis.2017.244

O'Brien, J., Wilson, I., Orton, T., & Pognan, F. (2000). Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity. European Journal of Biochemistry, 267(17), 5421–5426. https://doi.org/10.1046/j.1432-1327.2000.01606.x

Orth, M., Metzger, P., Gerum, S., Mayerle, J., Schneider, G., Belka, C., Schnurr, M., & Lauber, K. (2019). Pancreatic ductal adenocarcinoma: Biological hallmarks, current status, and future perspectives of combined modality treatment approaches. In Radiation Oncology (Vol. 14, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13014-019-1345-6

Pirona, A. C., Oktriani, R., Boettcher, M., & Hoheisel, J. D. (2021). Process for an efficient lentiviral cell transduction. Biology Methods and Protocols, 5(1), 1–8. https://doi.org/10.1093/biomethods/bpaa005

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nature Protocols, 8(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143

Rancati, G., Moffat, J., Typas, A., & Pavelka, N. (2018). Emerging and evolving concepts in gene essentiality. In Nature Reviews Genetics (Vol. 19, Issue 1, pp. 34–49). Nature Publishing Group. https://doi.org/10.1038/nrg.2017.74

Robert Koch Institut. (2020). Cancer in Germany. https://www.krebsdaten.de/

Roos, W. P., Thomas, A. D., & Kaina, B. (2016). DNA damage and the balance between survival and death in cancer biology. In Nature Reviews Cancer (Vol. 16, Issue 1, pp. 20–33). Nature Publishing Group. https://doi.org/10.1038/nrc.2015.2

Sánchez-Rivera, F. J., & Jacks, T. (2015). Applications of the CRISPR-Cas9 system in cancer biology. Nature Reviews Cancer, 15(7), 387–395. https://doi.org/10.1038/nrc3950

Sanjana, N. E., Shalem, O., & Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. In Nature Methods (Vol. 11, Issue 8, pp. 783–784). Nature Publishing Group. https://doi.org/10.1038/nmeth.3047

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G., & Zhang, F. (2014a). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science, 343(6166), 84–87. https://doi.org/10.1126/science.1247005

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G., & Zhang, F. (2014b). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science, 343(6166), 84–87. https://doi.org/10.1126/science.1247005

Shi, J., Wang, E., Milazzo, J. P., Wang, Z., Kinney, J. B., & Vakoc, C. R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nature Biotechnology, 33(6), 661–667. https://doi.org/10.1038/nbt.3235

Springfeld, C., Jäger, D., Büchler, M. W., Strobel, O., Hackert, T., Palmer, D. H., & Neoptolemos, J. P. (2019). Chemotherapy for pancreatic cancer. In Presse Medicale (Vol. 48, Issue 3P2, pp. e159–e174). Elsevier Masson SAS. https://doi.org/10.1016/j.lpm.2019.02.025

Sun, S. S., Fu, Y., & Lin, J. Y. (2020). Upregulation of *MYBL2* independently predicts a poorer prognosis in patients with clear cell renal cell carcinoma. Oncology Letters, 19(4), 2765–2772. https://doi.org/10.3892/ol.2020.11408

Taniguchi, S., Iwamura, T., & Katsuki, T. (1992). Correlation between spontaneous metastatic potential and type I collagenolytic activity in a human pancreatic cancer cell line (SUIT-2) and sublines. In Clin. Exp. Metastasis (Vol. 10).

Taniguchi, S., Iwamura, T., Kitamura, N., Yamanari, H., & Setoguchi, T. (1994). Heterogeneities of attachment, chemotaxis, and protease production among clones with different metastatic potentials from a human pancreatic cancer cell line. In Clin. Exp. Metastasis (Vol. 12).

Tao, D., Pan, Y., Jiang, G., Lu, H., Zheng, S., Lin, H., & Cao, F. (2015). *B-Myb* regulates snail expression to promote epithelial-to-mesenchymal transition and invasion of breast cancer cell. Medical Oncology, 32(1), 1–6. https://doi.org/10.1007/s12032-014-0412-y

von Hoff, D. D., Ramanathan, R. K., Borad, M. J., Laheru, D. A., Smith, L. S., Wood, T. E., Korn, R. L., Desai, N., Trieu, V., Iglesias, J. L., Zhang, H., Soon-Shiong, P., Shi, T., Rajeshkumar, N. v., Maitra, A., & Hidalgo, M. (2011). Gemcitabine plus nab-paclitaxel is an active regimen in patients with advanced pancreatic cancer: A phase I/II trial. Journal of Clinical Oncology, 29(34), 4548–4554. https://doi.org/10.1200/JCO.2011.36.5742

Wang, T., Birsoy, K., Hughes, N. W., Krupczak, K. M., Post, Y., Wei, J. J., Lander, E. S., & Sabatini, D. M. (2015). Identification and characterization of essential genes in the human genome. Science, 350(6264), 1096–1101. https://doi.org/10.1126/science.aac7041

Wang, T., Yu, H., Hughes, N. W., Liu, B., Kendirli, A., Klein, K., Chen, W. W., Lander, E. S., & Sabatini, D. M. (2017). Gene Essentiality Profiling Reveals Gene Networks and Synthetic Lethal Interactions with Oncogenic Ras. Cell, 168(5), 890-903.e15. https://doi.org/10.1016/j.cell.2017.01.013

Weadick, B., Nayak, D., Persaud, A. K., Hung, S. W., Raj, R., Campbell, M. J., Chen, W., Li, J., Williams, T. M., & Govindarajan, R. (2021). EMT-induced gemcitabine resistance in pancreatic cancer involves the functional loss of equilibrative nucleoside transporter 1. Molecular Cancer Therapeutics, 20(2), 410–422. https://doi.org/10.1158/1535-7163.MCT-20-0316

Weintraub, A. S., Li, C. H., Zamudio, A. v., Sigova, A. A., Hannett, N. M., Day, D. S., Abraham, B. J., Cohen, M. A., Nabet, B., Buckley, D. L., Guo, Y. E., Hnisz, D., Jaenisch, R., Bradner, J. E., Gray, N. S., & Young, R. A. (2017). YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell, 171(7), 1573-1588.e28. https://doi.org/10.1016/j.cell.2017.11.008

World Health Organization-International Agency for Research on Cancer (IARC). (2020). GLOBOCAN 2020. https://gco.iarc.fr/today

Xiong, Y. C., Wang, J., Cheng, Y., Zhang, X. Y., & Ye, X. Q. (2020). Overexpression of *MYBL2* promotes proliferation and migration of non-small-cell lung cancer via

upregulating NCAPH. Molecular and Cellular Biochemistry, 468(1–2), 185–193. https://doi.org/10.1007/s11010-020-03721-x

Yachida, S., Jones, S., Bozic, I., Antal, T., Leary, R., Fu, B., Kamiyama, M., Hruban, R. H., Eshleman, J. R., Nowak, M. A., Velculescu, V. E., Kinzler, K. W., Vogelstein, B., & Iacobuzio-Donahue, C. A. (2010). Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature, 467(7319), 1114–1117. https://doi.org/10.1038/nature09515

Zhang, L., Sanagapalli, S., & Stoita, A. (2018). Challenges in diagnosis of pancreatic cancer. In World Journal of Gastroenterology (Vol. 24, Issue 19, pp. 2047–2060). Baishideng Publishing Group Co. https://doi.org/10.3748/wjg.v24.i19.2047

Zhang, R., Ou, H. Y., & Zhang, C. T. (2004). DEG: A database of essential genes. Nucleic Acids Research, 32(DATABASE ISS.). https://doi.org/10.1093/nar/gkh024

Zhang, S., Yan, L., Cui, C., Wang, Z., Wu, J., Zhao, M., Dong, B., Guan, X., Tian, X., & Hao, C. (2020). Identification of *TYMS* as a promoting factor of retroperitoneal liposarcoma progression: Bioinformatics analysis and biological evidence. Oncology Reports, 44(2), 565–576. https://doi.org/10.3892/or.2020.7635

Zhang, X., Lv, Q. L., Huang, Y. T., Zhang, L. H., & Zhou, H. H. (2017). Akt/*FOXM1* signaling pathway-mediated upregulation of *MYBL2* promotes progression of human glioma. Journal of Experimental and Clinical Cancer Research, 36(1). https://doi.org/10.1186/s13046-017-0573-6

Zheng, A., Chevalier, N., Calderoni, M., Dubuis, G., Dormond, O., Ziros, P. G., Sykiotis, G. P., & Widmann, C. (2019). CRISPR/Cas9 genome-wide screening identifies KEAP1 as a sorafenib, lenvatinib, and regorafenib sensitivity gene in hepatocellular carcinoma. In Oncotarget (Vol. 10, Issue 66). www.oncotarget.com

Zhou, C., Yi, C., Yi, Y., Qin, W., Yan, Y., Dong, X., Zhang, X., Huang, Y., Zhang, R., Wei, J., Ali, D. W., Michalak, M., Chen, X. Z., & Tang, J. (2020). LncRNA PVT1 promotes gemcitabine resistance of pancreatic cancer via activating Wnt/β-catenin and autophagy pathway through modulating the miR-619-5p/Pygo2 and miR-619-5p/ATG14 axes. Molecular Cancer, 19(1). https://doi.org/10.1186/s12943-020-01237-y

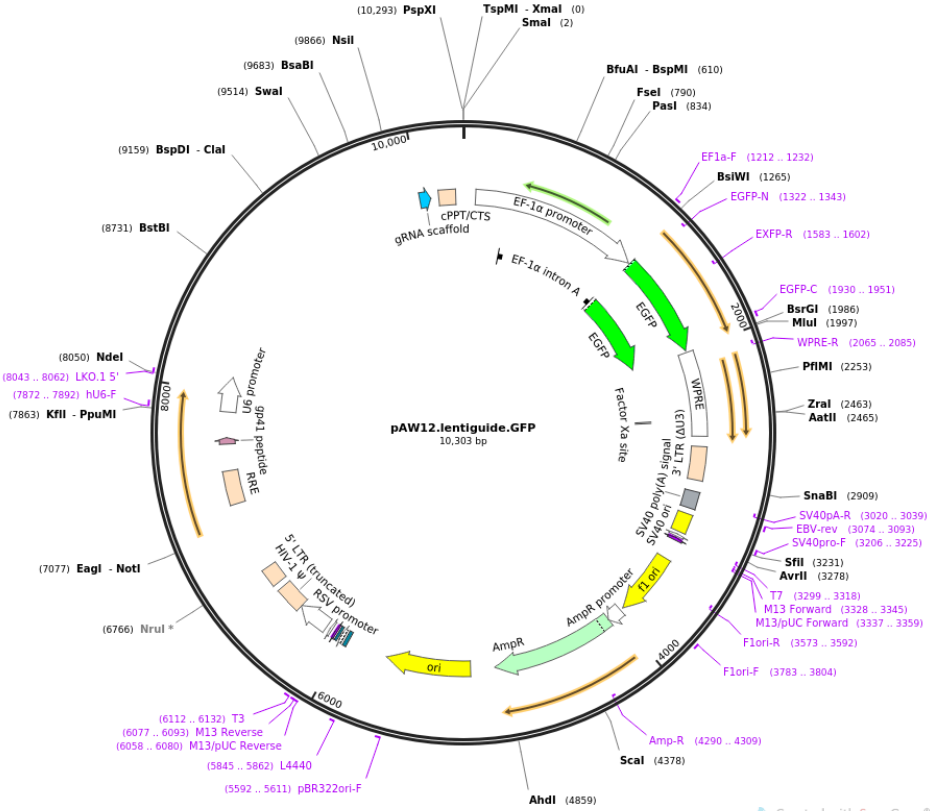
APPENDIXES

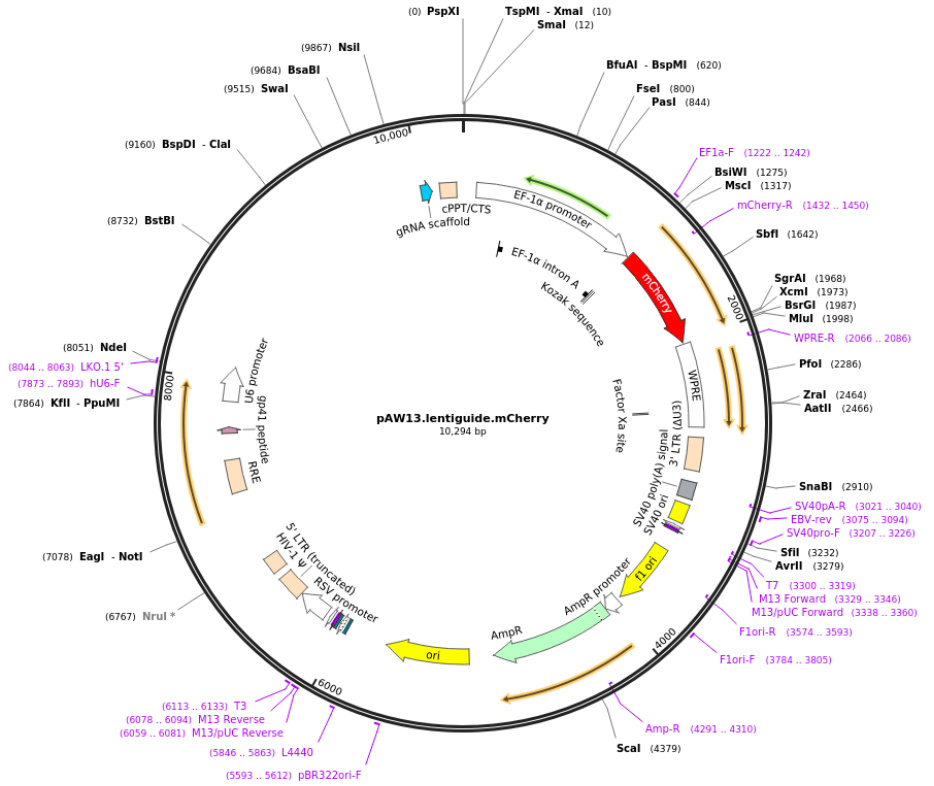
Appendix 1. Component of common buffers

Name of Buffer	Component	Volume
Tris Acetic Acid		
	242 g tris base	
	57.1 mL glacial acetic acid	
	100 mL 0.5 M EDTA	
Running Buffer 10x		
	30 g tris base	In 1 liter ddH_2O
	144 g glycine	
	10 g SDS	
Trenn Buffer (1x)		
	36.33 g tris base	In 200 mL
	adjust pH to 8.8 with HCl	
Sammel Buffer (1x)		
	47.28 g Tris HCL	In 200 mL
	adjust pH to 6.6 with NaOH	
Anode Buffer I		
	36.4 g Tris Base	Fill up with ddH_2O to 1 L
	200 mL methanol	
Anode Buffer II		
	3 g tris base	Fill up with ddH_2O to 1 L
	200 mL methanol	
Cathode Buffer		
	5.2 g 6-aminocaproic acid	Fill up with ddH_2O to 1 L
	200 mL methanol	
TBS 10 x		
	24 g tris HCl	Dissolved in 900 mL distilled water, adjust to Ph 7.6 and then add ddH_2O to the final volume 1 L , sterilized
	5.6 g tris Base	
	88 g NaCl	
1x TBST plus 0.1% Tween 20 (Washing Buffer)		
	100 mL 10x TBS	Fill up with ddH_2O to 1 L
	1 mLTween 20	
Blocking Buffer		
5% Skim milk powder	5 g skim milk powder	In 100 mL1x TBST

Appendix 2. The map of pAW12.lentiguide.GFP and pAW13.lentiguide. mCherry

(Figure copied from Addgene)





Appendix 3. KEGG pathway negative-selected genes

KEGG	Genes	Count	p value
Ribosome	MRPL30, MRPL33, RSL24D1, RPL11, RPL12, RPL13A,RPL14, RPL17, RPL18, RPL21, RPL31, RPL34, RPL35A, RPL4, RPL5, RPL7, RPS11, RPS14, RPS15, RPS16, RPS19, RPS2, RPS28, RPS3A, RPS4X, RPS5, RPS6, RPS9	29	7,70E-14
Spliceosome	THOC3, AQR, CRNKL1, EFTUD2, HNRNPC, MAGOH, NCBP2, PLRG1, PRPF31, SRSF1, SRSF6, SNRPD2, SNRPD3, SNRNP70, SNRPA, SNRPB, SF3A2, SF3B1, SF3B6, TRA2A, USP39	21	9,50E-08
Proteasome	PSMD1, PSMD8, PSMA1, PSMA3, PSMA4, PSMA5, PSMB1, PSMB4, PSMB5, PSMB5, PSMB7, SHFM1	12	4,40E-07
Ribosome biogenesis in eukaryotes	XRN2, FCF1, IMP4, NHP2, NOP56, NOP58, RBM28, SBDS, WDR3, WDR43, WDR75, EIF6, FBL, TBL3	14	1,90E-05
Oocyte meiosis	FBXW11, FBXO5, MAD2L1, ANAPC1, ANAPC10, ANAPC2, CALML3, CDC16, CCNB1, PPP1CB, PPP2CA, PPP2R1A, RBX1	13	9,00E-04
Ubiquitin mediated proteolysis	FBXW11, FBXW7, SAE1, ANAPC1, ANAPC10, ANAPC2, CDC16, CUL3, KEAP1, RBX1, UBE2C, UBE2D3, UBE2H, UBR5	14	1,90E-03
mRNA surveillance pathway	DAZAP1, RNPS1, CPSF3, MAGOH, NCBP2, PELO, PPP1CB, PPP2CA, PPP2R2A, PPP2R1A	11	2,10E-03
Cell cycle	EP300, E2F1, MAD2L1, WEE1, ANAPC1, ANAPC10, ANAPC2, CDC16, CCNB1, ORC4, PRKDC, RBX1	12	6,80E-03
RNA transport	RNPS1, THOC3, CLNS1A, EEF1A1, EIF2S1, EIF3B, EIF3B, EIF4G2, EIF4A1, EIF4B, EIF5, MAGOH, NCBP2, NUP62, NUP93	14	1,30E-02
One carbon pool by folate	ATIC; DHFR, MTHFD1, *TYMS*	4	3,70E-02
Progesterone-mediated oocyte maturation	AKT3, MAD2L1, RAF1, ANAPC1. ANAPC10, ANAPC2, CDC16, CCNB1	8	4,50E-02
Sulfur relay system	CTU1, CTU2, MOCS3	3	5,20E-02
Protein export	HSPA5, SPCS2, SRP14, SRP9	4	5,40E-02
Protein processing in endoplasmic reticulum	STT3A, EIF2AK4, EIF2S1, GANAB, HSP90B1, HSPA5, MOGS, MBTPS2, PRKCSH, RNP2, RBX1, UBE2D3	12	5,40E-02
Pentose phosphate pathway	PGLS, ALDO5, G6PD, RPE	4	9,40E-02
Pancreatic cancer	AKT3, BCL2L1, E2F1, RAD51, RAF1, TGFBR1	6	9,70E-02

Appendix 4. Top 20 Cellular Component Gene Ontology of negative and positive selected genes.

Table A.4A. Cellular Component Gene Ontology of negative-selected genes

ID	GO_Term	count	P_value
GO:0005654	nucleoplasm	206	1,10E-43
GO:0005829	cytosol	179	3,50E-20
GO:0032040	nucleolus	76	6,40E-19
GO:0005634	nucleus	234	8,70E-15
GO:0005840	ribosome	27	9,70E-13
GO:0022625	cytosolic large ribosomal subunit	16	4,70E-10
GO:0071013	catalytic step 2 spliceosome	18	6,20E-10
GO:0005681	spliceosomal complex	18	8,80E-10
GO:0022627	cytosolic small ribosomal subunit	13	5,70E-09
GO:1990904	intracellular ribonucleoprotein complex	19	4,70E-08
GO:0005839	proteasome core complex	9	5,30E-08
GO:0005737	cytoplasm	202	7,70E-08
GO:0005925	focal adhesion	32	2,10E-07
GO:0000502	proteasome complex	12	7,30E-07
GO:0016020	membrane	100	9,40E-07
GO:0070062	extracellular exosome	120	1,30E-06
GO:0072546	ER membrane protein complex	6	3,80E-06
GO:0071011	precatalytic spliceosome	8	4,00E-06
GO:0015935	small ribosomal subunit	8	7,00E-06
GO:0005813	centrosome	30	1,00E-05
GO:0032040	small-subunit processome	8	4,40E-05

Table A.4B. Cellular Component Gene Ontology of positive-selected genes

ID	GO_Term	count	P_value
GO:0005654	nucleoplasm	79	1,30E-07
GO:0005739	mitochondrion	44	5,20E-06
GO:0005730	nucleolus	27	1,10E-03
GO:0030897	HOPS complex	4	1,20E-03
GO:0005840	ribosome	10	1,30E-03
GO:0005747	mitochondrial respiratory chain complex I	5	7,30E-03
GO:0005743	mitochondrial inner membrane	15	1,00E-02
GO:0005769	early endosome	10	1,10E-02
GO:0005765	lysosomal membrane	11	1,20E-02
GO:0030123	AP-3 adaptor complex	3	1,20E-02
GO:0005634	nucleus	104	1,30E-02
GO:0031902	late endosome membrane	6	2,20E-02
GO:0043527	tRNA methyltransferase complex	2	3,10E-02
GO:1990622	CHOP-ATF3 complex	2	3,10E-02
GO:0005753	mitochondrial proton-transporting ATP synthase complex	3	4,30E-02
GO:0005770	late endosome	6	4,40E-02

Appendix 5. Top 20 Molecular Function Gene Ontology of negative and positive selected genes.

Table A5A. Molecular Function Gene Ontology of negative-selected genes

ID	GO_Term	Count	p value
GO:0003723	poly(A) RNA binding	108	7,30E-29
GO:0005515	protein binding	357	8,40E-22
GO:0003723	RNA binding	58	2,60E-17
GO:0003735	structural constituent of ribosome	29	6,10E-11
GO:0004298	threonine-type endopeptidase activity	9	6,80E-08
GO:0003743	translation initiation factor activity	12	1,20E-06
GO:0000166	nucleotide binding	26	3,00E-05
GO:0030515	snoRNA binding	6	2,80E-04
GO:0003729	mRNA binding	13	2,80E-04
GO:0051082	unfolded protein binding	12	3,50E-04
GO:1990446	U1 snRNP binding	4	7,80E-04
GO:0016779	nucleotidyltransferase activity	6	9,50E-04
GO:0019843	rRNA binding	7	9,50E-04
GO:0098641	cadherin binding involved in cell-cell adhesion	19	2,10E-03
GO:0042054	histone methyltransferase activity	4	3,40E-03
GO:0000049	tRNA binding	7	3,40E-03
GO:0043022	ribosome binding	7	3,40E-03
GO:0005524	ATP binding	62	4,00E-03
GO:0051539	4 iron, 4 sulfur cluster binding	6	7,00E-03
GO:0004842	ubiquitin-protein transferase activity	19	7,60E-03
GO:0019888	protein phosphatase regulator activity	5	9,30E-03

Table A5B. Molecular Function Gene Ontology of negative-selected genes

ID	GO_Term	Count	P_value
GO:0005515	protein binding	166	9,70E-04
GO:0046965	retinoid X receptor binding	4	1,60E-03
GO:0003697	single-stranded DNA binding	7	3,60E-03
GO:0003723	poly(A) RNA binding	30	7,20E-03
GO:0000049	tRNA binding	5	8,70E-03
GO:0043522	leucine zipper domain binding	3	1,50E-02
GO:0002039	p53 binding	5	2,20E-02
GO:0003735	structural constituent of ribosome	9	2,60E-02
GO:0008176	tRNA (guanine-N7-)-methyltransferase activity	2	3,20E-02
GO:0004842	ubiquitin-protein transferase activity	11	3,80E-02
GO:0008137	NADH dehydrogenase (ubiquinone) activity	4	4,10E-02
GO:0000107	imidazoleglycerol-phosphate synthase activity	2	4,70E-02
GO:0018127	NAD(P)-serine ADP-ribosyltransferase activity	2	4,70E-02
GO:0044102	purine deoxyribosyltransferase activity	2	4,70E-02
GO:0018121	NAD(P)-asparagine ADP-ribosyltransferase activity	2	4,70E-02
GO:0043867	7-cyano-7-deazaguanine tRNA-ribosyltransferase activity	2	4,70E-02
GO:0018071	NAD(P)-cysteine ADP-ribosyltransferase activity	2	4,70E-02

Appendix 6. Top 20 Biological Process Gene Ontology of negative and positive selected genes.

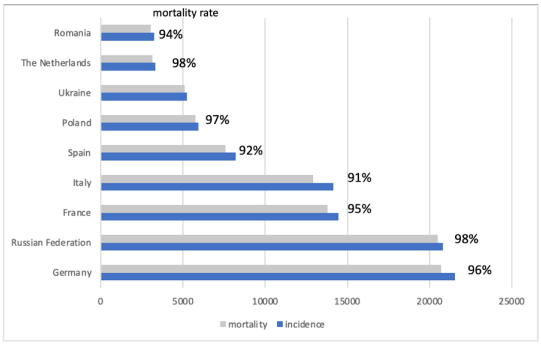
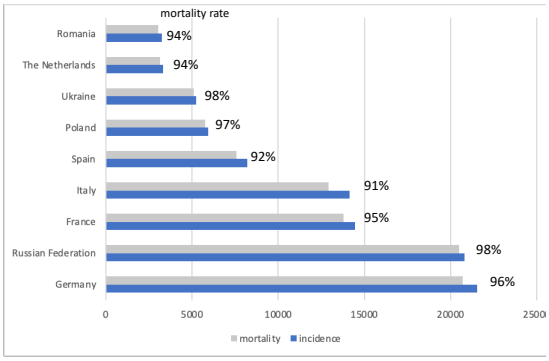
Table A6A. Biological Process Gene Ontology of negative-selected gene

ID	GO_Term	Count	p value
GO:0006413	translational initiation	36	2,10E-23
GO:0000184	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	32	3,90E-21
GO:0006614	SRP-dependent cotranslational protein targeting to membrane	28	8,70E-20
GO:0006364	rRNA processing	39	2,30E-19
GO:0019083	viral transcription	28	1,20E-17
GO:0031145	anaphase-promoting complex-dependent catabolic process	20	9,50E-13
GO:1904667	negative regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle	19	1,40E-12
GO:0000398	mRNA splicing, via spliceosome	31	2,90E-12
GO:0006412	translation	33	3,30E-12
GO:1901992	positive regulation of ubiquitin-protein ligase activity involved in regulation of mitotic cell cycle transition	19	5,00E-12
GO:0051301	cell division	36	2,60E-10
GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process	26	1,40E-09
GO:0008380	RNA splicing	22	1,80E-08
GO:0140014	mitotic nuclear division	26	8,40E-08
GO:0002479	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	13	2,50E-07
GO:0000209	protein polyubiquitination	21	5,00E-07
GO:0006260	DNA replication	19	7,00E-07
GO:0007062	sister chromatid cohesion	15	1,90E-06
GO:0006521	regulation of cellular amino acid metabolic process	11	1,90E-06
GO:1904874	positive regulation of telomerase RNA localization to Cajal body	7	2,50E-06
GO:0060071	Wnt signaling pathway, planar cell polarity pathway	14	2,80E-06

Table A6B. Biological Process Gene Ontology of positive-selected gene

ID	GO_Term	Count	p value
GO:0032981	mitochondrial respiratory chain complex I assembly	6	3,20E-03
GO:0060323	head morphogenesis	3	3,60E-03
GO:0035542	regulation of SNARE complex assembly	3	3,60E-03
GO:0006120	mitochondrial electron transport, NADH to ubiquinone	5	7,60E-03
GO:0035914	skeletal muscle cell differentiation	5	7,60E-03
GO:0006400	tRNA modification	4	8,80E-03
GO:0003281	ventricular septum development	4	9,70E-03
GO:0030261	chromosome condensation	3	1,00E-02
GO:0090090	negative regulation of canonical Wnt signaling pathway	8	1,60E-02
GO:0007017	microtubule-based process	4	1,90E-02
GO:0006412	translation	10	2,00E-02
GO:0008033	tRNA processing	4	2,10E-02
GO:0050905	neuromuscular process	3	2,30E-02
GO:0045948	positive regulation of translational initiation	3	2,30E-02
GO:0008333	endosome to lysosome transport	4	2,40E-02
GO:0007346	regulation of mitotic cell cycle	4	2,50E-02
GO:0045740	positive regulation of DNA replication	4	2,90E-02
GO:1901992	positive regulation of ubiquitin-protein ligase activity involved in regulation of mitotic cell cycle transition	5	3,30E-02
GO:0016192	vesicle-mediated transport	7	3,50E-02
GO:0006511	protein ubiquitination involved in ubiquitin-dependent protein catabolic process	7	3,60E-02
GO:0031145	anaphase-promoting complex-dependent catabolic process	5	3,70E-02

CORRIGENDUM

Page	Original	Correction
2		
6	CRISPR-Cas9 system (Clustered regularly interspaced short palindromic repeats – CRISPR associated protein 9) is a system found firstly in microbial organisms as an adaptive immune system.	CRISPR-Cas system (Clustered regularly interspaced short palindromic repeats – CRISPR associated) is a system found firstly in microbial organisms as an adaptive immune system.
12	The transformation was conducted using One-Shot STBL3 Escherichia coli (E. coli) (Cat. No. C737303, Invitrogen, USA) according to the manufacturer's protocol with some modifications.	The transformation was conducted using One-Shot STBL3 *Escherichia coli* (*E. coli*) (Cat. No. C737303, Invitrogen, USA) according to the manufacturer's protocol with some modifications.
15	The 5'-CACCG-3' sequence was added at the beginning of the forward sequence and 5'-AAAC-sgRNA's sequence-C'3'.	The 5'-CACCG-3' sequence was added at the beginning of the forward sequence. The reverse sequence was modified by adding the following bases: 5'-AAAC-sgRNA's reverse sequence-C-3'.
15	The PEI lentivirus production was performed as previously described (Pirona et al., 2021). 600,000 HEK-293T cells were seeded in 10 cm2 culture dishes in 10 mL, and then incubated overnight.	The PEI lentivirus production was performed as previously described (Pirona et al., 2021). 600,000 HEK-293T cells were seeded in 10 cm^2 culture dishes in 10 mL, and then incubated overnight.
22	**Figure 2.4. Overview of the CRISPR screening protocol.** The timeline for CRISPR screening: starting from day 0 with cell plating, followed by transduction, puromycin selection, and two-time points of DNA collection.	**Figure 2.4. Overview of the CRISPR screening protocol.** The timeline for CRISPR screening: starting from day 0 with cell plating, followed by transduction, puromycin selection, and two-time points of DNA collection (Figure created with BioRender.com).
27	**Figure 2.6. An illustration of cell viability assay using flow cytometry.** Transduced cells	**Figure 2.6. An illustration of cell viability assay using flow cytometry.** Transduced

	(mCherry positive cells) were mixed with wild type in a 4:1 ratio. The cells were then plated and cultured for 7 days and continued until tx. Some of the cells were analyzed to count for mCherry positive cells at t_{zero}, t_7, or t_x.	cells (mCherry positive cells) were mixed with wild type in a 4:1 ratio. The cells were then plated and cultured for 7 days and continued until tx. Some of the cells were analyzed to count for mCherry positive cells at t_{zero}, t_7, or t_x (Figure created with BioRender.com).
57	In S2-028 and S2-007, *MEN1* knockout showed no effect on cell proliferation compared to the controls (NTC1 and NTC1) and the vector.	In S2-028 and S2-007, *MYBL2* knockout showed no effect on cell proliferation compared to the controls (NTC1 and NTC1) and the vector.
71	Based on the results, the knockout of *MYBL2* showed invasion inhabitation in S2-007, but no difference was observed in wound healing assay.	Based on the results, the knockout of *MYBL2* showed invasion inhibition in S2-007, but no difference was observed in wound healing assay.
74	To Prof. Dr. Dra. Sunarti, M. Kes, dr. Arta Fatmawati., Ph.D, and dr. Hamim Sadewa., Ph.D., thank you for your support.	To Prof. Dr. Dra. Sunarti, M. Kes, dr. Arta Fatmawati, Ph.D, and dr. Hamim Sadewa, Ph.D., thank you for your support.
74	To Prof. dr. Sofia Mubarika, M.Med.Sc., Prof. Dr. dr. Teguh Aryandono., Sp. B (K) Onk, and Dr. med. dr. Indwiani Astuti., thank you for your unparalleled support since the beginning of my career.	To Prof. dr. Sofia Mubarika, M.Med.Sc. Ph.D., Prof. Dr. dr. Teguh Aryandono, Sp. B (K) Onk, and Dr. med. dr. Indwiani Astuti, thank you for your unparalleled support since the beginning of my career.