



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e. g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.



**Exploring Genetic Alterations and Protein Marker Expression
in Cancer: Investigating ARID1A and SMARCA4 Knock-Outs
in Colorectal Cancer and Evaluating Histological Markers for
Targeted Therapy in Ovarian Cancer**

Xiaoxi Jin (Cici)

Thesis submitted for the degree of Master of Science by Research:

Cancer Research

The University of Edinburgh

2024

Word count: 14045

Declaration

I, Xiaoxi Jin, declare that this thesis was composed solely by me. All work is my own except when the contributions of others are clearly and appropriately acknowledged. This thesis has not been submitted for any other degree or professional qualification.

Xiaoxi Jin

September 2024

Acknowledgement

My sincerest gratitude to The Myant Lab, Professor Charlie Gourley, Professor Susan Farrington and Dr Robb Hollis for helping with the evolution and completion of this thesis. To Dr Farhat Din for the support she has given me throughout the past year, and to Joanna Porter for working alongside me on my research project.

Abstract

This dissertation presents two distinct yet connected research projects focused on the alteration and evaluation of DNA and protein expression in common cancers, both of which are subjects of ongoing research into their mechanisms and potential therapeutic advancements. Colorectal cancer, one of the most prevalent and deadly cancers globally involves metastasis and therapy resistance driven largely by tumour plasticity, which is affected by changes in the expression of epigenetic modifying proteins. Ovarian cancer, on the other hand, is common among women and exhibits various histotypes that differ in their origin, tumour microenvironment, and genetic and epigenetic expressions. Both cancers stand to benefit from further study to facilitate the development of new treatments.

The first project focuses on producing SWI/SNF CRISPR/Cas9 knockouts in a mouse colon organoid line (Apcfl/fl;KrasLSL-G12D/+;Trp53fl/fl). Knockouts were validated using qPCR and Western blot analysis. Incomplete knockouts were produced, indicating the need for further evaluation using DNA sequencing. Additional knockout experiments are recommended, utilising the existing protocol with different clones picked from the lentiviral transduction pool.

The second project involved the use of TMAs (tumour microarrays) taken from patients with different ovarian cancer histotypes. These TMAs were sectioned and stained for membrane proteins that are potential targets for antibody-drug conjugate (ADC) therapy. The TMAs were then scored by two independent assessors for the expression of these targets, providing a foundation for future studies exploring the suitability of ADC treatments in different patient demographics and ovarian cancer histotypes, as well as the role of target expression levels.

Though these two projects investigate different types of cancer, together they provide a broad foundation in research methods and the diverse skills required for cancer research. Both projects contribute valuable groundwork for future studies on the mechanisms and treatment of colorectal and ovarian cancers.

Lay summary

This dissertation explores two different projects aimed at understanding how DNA and proteins influence two types of cancer—colon cancer and ovarian cancer—so that scientists can find better treatments.

Colon cancer is one of the most common and dangerous types of cancer, and it spreads because certain proteins, which change how DNA works, also change how the tumour behaves. Ovarian cancer affects many women and comes in different types, which not only arise from different tissues, but also behave differently because of the DNA and proteins inside of them.

In the first project, the research focused on colon cancer. A tool called CRISPR/Cas9 was used to “turn off” specific genes in the DNA of mouse colon organoids (mini-colons that are grown in a petri dish). These genes belong to a group of proteins called SWI/SNF, which help control how DNA works. Special tests confirmed whether the genes were successfully turned off, but in some cases, additional testing was needed to confirm the process.

The second project focused on ovarian cancer. Samples from patients were tested for specific proteins found on the surface of cancer cells. These proteins could be targets for a new treatment called Antibody Drug Conjugates (ADC), which delivers medicine directly to the cancer through binding to specific targets expressed on the outside of the tumours. Two assessors measured the levels of expression of 3 different markers (FOLR1, VTCN1/B7-H4, and TACSTD2) in 3 ovarian cancer histotypes (clear cell ovarian carcinoma, endometrioid ovarian carcinoma, and high-grade serous ovarian carcinoma) to determine if the treatment could work for different patients.

Although these projects focus on different cancers, both contribute to understanding how DNA and proteins function in cancer, paving the way for new, life-saving treatments in the future.

Table of Contents

Declaration.....	2
Acknowledgement.....	3
Abstract.....	4
Lay summary.....	5
Table of Contents.....	6
Introduction.....	8
Part 1: Investigation of the role of the epigenetic modifiers Smarca4 and Arid1a in colorectal cancer by generating genetic knock-outs using mouse colon organoids.....	9
1 Background.....	9
1.1 Colorectal cancer, its genetic and epigenetic factors.....	9
1.2 The role of chromatin remodelling in cancer cell plasticity.....	11
1.3 The SWI/SNF complex.....	12
2 Aims and Hypotheses.....	13
3 Methodology.....	13
3.1 Tissue culture.....	13
3.2 Production of knock-outs.....	14
3.2.1 CRISPR-Cas9 and sgRNA design.....	14
3.2.2 Lentiviral vector digestion, oligo annealing and cloning.....	15
3.2.3 Lentiviral transduction.....	17
3.3 Verification of Knock-outs.....	17
3.3.1 qPCR.....	17
3.3.2 Western blot.....	20
4 Results.....	21
4.1 Organoid maintenance.....	21
4.2 Sanger sequencing.....	22
4.2.1 Miniprep results.....	22
4.2.2 Maxiprep results.....	23
4.3 qPCR results.....	24
4.3.1 Baseline mRNA expression.....	24
4.3.2 qPCR results of knock-outs.....	25
3.3 Western blot results.....	27
5 Discussion.....	28
Part 2: Scoring Tumour Microarrays for Protein Marker Expression in Ovarian Cancer: Insights for ADC Targeted Therapy.....	31
1 Background.....	31
1.1 Ovarian cancer.....	31
1.1.1 Patient cohort.....	32
1.2 Tumour Microarrays.....	33
1.3 Antibody-Drug Conjugate (ADC) therapy.....	34
1.3.1 ADC targets.....	35

1.4 Histological scoring.....	36
2 Methodology.....	36
2.1 TMA scoring.....	36
2.2 Inter-assessor agreement evaluation.....	37
3 Results.....	38
3.1 Correlation analysis for inter-assessor agreement.....	38
3.2 Mean target expressions.....	42
4 Discussion.....	46
Appendix 1.....	48
Appendix 2.....	50
Appendix 3.....	52
Appendix 4.....	55
Appendix 5.....	58
Appendix 6.....	60
References.....	63

Introduction

Cancer research is essential for understanding the complexities of different malignancies and developing more effective treatments for cancer. This dissertation focuses on two distinct projects that address key aspects of cancer biology and therapy, specifically in colorectal and ovarian cancers. Although the projects are independent, they both contribute to the broader goal of advancing knowledge in cancer research, particularly in understanding genetic alterations and exploring targeted therapies.

Part 1 of this dissertation will be about the first project, which is centred on colorectal cancer, one of the most common and deadly forms of cancer worldwide (1). This study aims to investigate the role of the *Arid1a* and *Smarca4* genes, which are critical components of the SWI/SNF chromatin remodelling complex. This complex is responsible for regulating gene expression by altering the structure of chromatin, and mutations in these genes are implicated in various cancers, including colorectal cancer (2,3). In this project, CRISPR-Cas9 was used on *Apcf1/fl;KrasLSL-G12D/+;Trp53fl/fl* (AKP) mouse colon tumour organoids, three-dimensional culture systems that closely replicate the tumour environment, to generate *Arid1a* and *Smarca4* knock-outs. Although the functional consequence of these genetic alterations on cancer plasticity and progression have yet to be explored, this foundational work provides a platform for future studies aimed at understanding the role of *Arid1a* and *Smarca4* in colorectal cancer.

Part 2 of this dissertation will be about the second project, which focuses on ovarian cancer. One of the promising avenues in ovarian cancer treatment is the use of antibody-drug conjugates (ADCs), which offer a targeted approach by delivering cytotoxic drugs directly to cancer cells expressing specific protein markers. This project involves the histological analysis of ovarian cancer tissue samples to evaluate the expression of these protein markers. The objective of the project was to assess the concordance between scorers and therefore determine the reliability of such assessment in the event that expression intensity is a biomarker of response. This initial work lays the groundwork for future investigations aimed at identifying correlations between marker expression and potential responsiveness to ADC therapy. Such insights could lead to more personalised treatment strategies for ovarian cancer patients.

In summary, this dissertation presents two projects that contribute to our understanding of cancer biology and the development of targeted therapies. The first project aims to provide a genetic and molecular basis for future studies on colorectal cancer plasticity,

while the second sets the stage for a detailed analysis of protein marker expression in ovarian cancer, which could inform the development of more effective, individualised treatments. Together, these projects underscore the importance of both basic and translational research in the ongoing fight against cancer.

Part 1: Investigation of the role of the epigenetic modifiers Smarca4 and Arid1a in colorectal cancer by generating genetic knock-outs using mouse colon organoids.

1 Background

1.1 Colorectal cancer, its genetic and epigenetic factors

Colorectal cancer (CRC) is a type of cancer that affects the colon (large intestine) and rectum (4) and it is the 4th most prevalent cancer and 2nd most deadliest cancer globally (5,6). It affects about 2 million new patients and kills 1 million each year. High mortality rate is often attributed to the disease's asymptomatic nature in its early phases and the absence of widespread early screening. Adenocarcinoma is the most common type of colorectal cancer (CRC)(7). Colorectal adenocarcinomas evolve through a well-established sequence from adenoma to carcinoma, initiating with the formation of precancerous polyps in the intestinal lining. As these polyps grow, the cellular divisions within them accumulate an increasing number of genetic mutations. These mutations eventually confer the capacity to penetrate the intestinal epithelial layer, extend to lymphatic nodes, and ultimately gain access to the circulatory system, facilitating the invasion of distant organs (8).

CRC may exhibit a variety of genetic mutations, which often vary depending on the location of the tumour (9). In rectal cancers, TP53 is commonly mutated, with mutations present in approximately 70% of cases. These mutations are associated with higher overall cancer stage and lower disease-free survival rates (10). Additionally, FBXW7 mutations, found in about 13% of CRC patients, are also common in rectal tumours and are linked to shorter disease-free survival. NRAS mutations, while less prevalent than KRAS mutations, can influence treatment responses in rectal cancers as well. For

left-sided colon cancers, the SMAD4 gene is the most commonly mutated, with around 12% of CRC patients exhibiting these mutations. SMAD4 plays a crucial role in the TGF-beta signalling pathway, which regulates cell growth and apoptosis. Right-sided colon cancers, on the other hand, frequently harbour mutations in several genes, including KRAS, PIK3CA, BRAF, MET, and PTEN. KRAS mutations, particularly in codons 12 and 13 of exon 2, occur in approximately 40% of CRC tumours and are negative predictors of response to anti-EGFR (epidermal growth factor receptor) antibody treatments (11,12). PIK3CA mutations are also common in right-sided tumours and play a role in the PI3K/AKT signalling pathway. BRAF mutations, especially the V600E mutation, lead to continuous activation of signalling pathways that promote cell growth and division, contributing to tumour development (13). MET mutations are involved in pathways that regulate cell proliferation and survival, while PTEN loss of function mutations can activate the PI3K/AKT pathway, further driving cancer progression. Another significant mutation in CRC is found in the APC gene, with around 60% of CRC patients exhibiting these mutations. APC is a key regulator of the Wnt signalling pathway, and its mutation is often an early event in colorectal carcinogenesis.

In addition to genetic changes, in recent years, an expanding array of research has underscored the significant influence of epigenetic regulatory mechanisms in CRC (14). This includes processes such as DNA methylation, histone modification, chromatin remodelling complexes, and non-coding RNAs, all of which play critical roles not only in the initial occurrence and subsequent development of CRC but also in shaping its response to therapeutic interventions.

Metastatic CRC is highly resistant to therapeutic interventions and is subject to eventual relapse even in patients with initial therapy response. Metastatic cancer cells have the ability to undergo changes to their phenotype without acquiring genetic mutations in response to therapy. This process, termed cellular plasticity, is a major contributor to metastatic progression, therapy resistance and relapse. Studying the underlying biological mechanisms of cell plasticity should not only enable a better understanding of cancer cell phenotypic and functional heterogeneity but also present novel opportunities to prevent the emergence of drug resistance (5). Epigenetic factors play a significant role in enhancing cancer plasticity. During carcinogenesis, genetic lesions can induce alterations in chromatin structure and DNA methylation patterns, which subsequently increase cellular plasticity and promote oncogenic reprogramming. As cancer progresses, epigenetic changes continue to evolve, often triggered by interactions with the tumour microenvironment. These interactions can modulate cancer cell phenotypes and properties, effectively shaping the overall architecture of the tumour (15,16).

Epigenetic changes serve as a crucial integrative mechanism, combining the effects of both cell-intrinsic factors, such as subclonal mutations, and cell-extrinsic influences, including signalling from the surrounding microenvironment. This integration helps establish intratumoral heterogeneity, a key feature that can either promote or inhibit the cancer stem cell (CSC) state, contributing to the dynamic nature of tumour growth. Moreover, epigenetic alterations enhance cellular plasticity, allowing for reversible transitions between different phenotypic states within the tumour (17). This adaptability not only supports cancer cell survival under varying conditions but also contributes to resistance against therapeutic interventions. The dual nature of epigenetic dysregulation—being both heritable and reversible—presents a promising avenue for the development of targeted therapies that focus on modulating epigenetic regulation as a strategic approach to cancer treatment (18).

1.2 The role of chromatin remodelling in cancer cell plasticity

Throughout the advancement of cancer, tumour cells have the ability to transition between different cellular states, largely facilitated by cell plasticity, to evade selective pressures. As a result, cell plasticity significantly contributes to the diversity and adaptability within the tumour (intra-tumour heterogeneity), thereby accelerating tumour expansion, the potential for metastasis, and resistance to treatments (7). Examples of cell plasticity include transdifferentiation (direct fate switching to another differentiated cell type), dedifferentiation (the loss of lineage commitment and reacquisition of stem cell features), and EMT (epithelial-to-mesenchymal transition) (19). One aspect of cellular plasticity is a global rewiring of transcription, driven by epigenetic alterations and transcription factor activity, but the mechanistic basis for these changes and the function of epigenetic modifying enzymes in cancer is understudied.

In eukaryotes, DNA is tightly wound into a complex called chromatin. Thanks to the process of chromatin remodelling, this complex can be “opened” so that specific genes are expressed (20). Chromatin structure affects cellular identity and state transitions by adopting active and repressive states (21). Active states are made accessible to transcription factors (TF) and other regulatory factors, and are enriched for histone modifications such as H3K27ac and H3K4me3. Repressive states are more compact, and characterised by DNA hypermethylation, EZH2-catalysed H3K27me3 and H3K9me3. Chromatin remodelling complexes are protein complexes that use energy from ATP hydrolysis to slide, eject, or reposition nucleosomes along the DNA. They can change the chromatin structure by making DNA more accessible (euchromatin) or less accessible (heterochromatin). SWItch/Sucrose Non-Fermenting (SWI/SNF) is one of

these complexes which mediate cellular plasticity through processes such as interacting with YAP/TAZ (19,22,23).

1.3 The SWI/SNF complex

The SWI/SNF complex is a chromatin remodelling complex that plays a crucial role in various cellular processes, such as regulating gene expression, repairing DNA damage, and facilitating cell differentiation (24–26). Comprising 9-12 subunits, the complex includes a core ATPase subunit, which can be either BRG1 (also SMARCA4) or BRM (also SMARCA2). Through the hydrolysis of ATP, the SWI/SNF complex mobilises nucleosomes and remodels chromatin by destabilising the interactions between DNA and histones. This disruption of chromatin structure opens up transcription-binding sites, allowing transcription factors to access these regions and consequently enhancing transcriptional activity. This complex is highly conserved across eukaryotic organisms and is frequently inactivated in cancers. In fact, mutations in the genes encoding SWI/SNF complex subunits are among the most common in cancer, collectively occurring in nearly 25% of all human cancers (27,28).

Among these, the *AT-rich interacting domain-containing protein 1A* (*ARID1A*, also known as *BAF250a*, *B120*, *C1orf4*, or *Osa1*) that encodes protein ARID1A is the most frequently mutated, and mutations in *ARID1A* have been found in various types of cancer (29). ARID1A is thought to play a significant role both in tumour initiation and in tumour suppression, which is highly dependent upon context. Recent mechanistic research has revealed that ARID1A participates in tumour progression through its effects on control of cell cycle, modulation of cellular functions such as EMT, and regulation of various signalling pathways (29).

Recently, mutations of SWI/SNF complexes arise in multiple kinds of human malignancies, implying the key function of SWI/SNF complexes in tumour suppression. Initially, the SWI/SNF complex was found to be correlated with cancer in 1998 (30). Subsequently, research on cancer whole-exome high-throughput sequencing has revealed frequent mutations of genes encoding multiple subunits of the SWI/SNF complex, among which *ARID1A* is the most frequently mutated (31,32). In addition to ARID1A, mutations in SMARCA4 have also been identified in multiple human malignancies (33–36).



Figure 1: An outline of the mutational status of a subset of epigenetic modifying enzymes in colorectal cancer, showing that ARID1a and SMARCA4 and members of COMPASS complexes (KMT2A, KMT2B, KMT2C, KMT2D, KDM6A and KDM68) are driver genes for colorectal cancer development. ARID1a and SMARCA4 together make up around 17% of colorectal cancer mutations. Image obtained from cBioPortal (37).

2 Aims and Hypotheses

The goal of this project is to produce knockouts of the Arid1a and Smarca4 genes to investigate their role in tumour plasticity. By generating these genetic knockouts, the study aims to uncover the underlying mechanisms by which these genes contribute to the adaptability and progression of tumours, potentially shedding light on their involvement in cancer development and therapeutic resistance.

3 Methodology

3.1 Tissue culture

The APC^{fl/fl};Kras^{G12D/+};Trp53^{fl/fl} (AKP) mouse colon tumour organoid line, originally generated by the Myant lab, was obtained for CRISPR-Cas9 knock-out experiments. The maintenance and splitting of the AKP line followed the “Maintenance of Mouse

Intestinal Organoids” protocol provided by the Myant Lab (Appendix 1). The ADF base medium (stored at 4°C) was prepared using Advanced DMEM/F12 (500ml) from Gibco, TFS stored at 4°C combined with L-glutamine (5ml), 5ml of 10000U/ml Penicillin/Streptomycin (Thermo Fisher Scientific), and 5ml of HEPES (Sigma-Aldrich) which are routinely stored at -20°C. For culturing, EN medium was prepared by mixing 48ml of ADF base medium with B27 neuronal supplement (1ml), N2 neuronal supplement (0.5ml), Noggin BMP inhibitor (0.5ml), and EGF (50µl), which are stored at -20°C. The growth of organoids was monitored daily using an AmScope B100 Series Biological Binocular Compound Microscope (40X-1000X magnification), and splitting was performed every 2-3 days, depending on the confluence and growth of the organoids (Appendix 8).

3.2 Production of knock-outs

3.2.1 CRISPR-Cas9 and sgRNA design

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is a microbial nuclease system that serves as a defence mechanism against invading phages and plasmids. Within microbial hosts, CRISPR loci contain a combination of CRISPR-associated (Cas) genes and non-coding RNA elements, which program the specificity of CRISPR-mediated nucleic acid cleavage. Lentiviral CRISPR/Cas systems can infect a wide range of mammalian cells by co-expressing a mammalian codon-optimised Cas9 nuclease along with a single guide RNA (sgRNA) to enable precise genome editing (38). The lentiCRISPRv2 plasmid used for this purpose includes two expression cassettes: hSpCas9 and a chimeric guide RNA. The plasmid can be digested with BsmBI, allowing for the insertion of annealed oligos into the sgRNA scaffold. These oligos are designed based on a 20-base pair (bp) target sequence, which must be flanked by a 3bp NGG protospacer adjacent motif (PAM) at the 3' end.

For this project, three sgRNA sequences were designed for each target gene, as well as a non-targeting control (NT). After BsmBI digestion, all plasmids have identical overhangs. The sequences were designed using the mouse genome-scale CRISPR knock-out (GeCKO) Library A (MGLibA_05067; MGLibA_05068; MGLibA_05069; MGLibA_49985; MGLibA_49986; MGLibA_49987), and all are presented in the 5'-3' orientation in Table 1. The sgRNAs were prepared to a concentration of 100µM using

double-distilled water (ddH₂O) and vortexed thoroughly. The resuspended sgRNAs were stored at -20°C for future use.

sgArid1a-1 Fwd	CACCGGCCCTGCTGGCCATA CGCAC
sgArid1a-1 Rev	AAACGTGCGTATGGCCAGCA GGGCC
sgArid1a-2 Fwd	CACCGGGTCCCTGTTGTTGC GAGTA
sgArid1a-2 Rev	AAACTACTCGCAACAACAGG GACCC
sgArid1a-3 Fwd	CACCGTAACCCATACTCGCAA CAAC
sgArid1a-3 Rev	AAACGTTGTTGCGAGTATGG GTTAC
sgSmarca4-1 Fwd	CACCGCCACCCTCAGTGTCC GCCAC
sgSmarca4-1 Rev	AAACTGGGCGGACACTGAGG GTTGC
sgSmarca4-2 Fwd	CGCCGAGGCATGTTCAAGAGC CGCCG
sgSmarca4-2 Rev	AAACCGGCGGCTCTGAACAT GCCTC
sgSmarca4-3 Fwd	CACCGTATGGAGTCCATGCA CGAGA
sgSmarca4-3 Rev	AAACTCTCGTGCATGGACTC CATAAC

Table 1: sgRNA sequences designed for CRISPR-Cas9.

3.2.2 Lentiviral vector digestion, oligo annealing and cloning

The lentiviral vector digestion, oligo annealing, cloning, and transformation were conducted following the "lentiCRISPRv2 and lentiGuide-Puro: Lentiviral CRISPR/Cas9 and Single Guide RNA Cloning Protocol" (Appendix 2). The process began by digesting and dephosphorylation of 5 µg of the lentiCRISPRv2 plasmid with BsmBI at 37°C for 30 minutes, using a combination of FastDigest BsmBI, FastAP, 10X FastDigest Buffer, freshly prepared DTT, and double-distilled water (ddH₂O) to make a total reaction volume of 60 µl. Gel purification was performed using an agarose gel prepared by dissolving 2g of agarose powder (Invitrogen by Thermo Fisher Scientific) in 200 ml of TAE buffer, microwaving for 2 minutes, and cooling for 15 minutes. SYBR[™] Safe DNA Gel Stain was added at 1 in 10,000 concentration, and electrophoresis was run at 120V for 45 minutes, loading both a 10,000 bp DNA ladder and the plasmid sample. In a dark room, under UV light, the larger band was extracted, leaving behind a ~2 kb filler band. Purification of the digested plasmid was done using the QIAquick Gel Extraction Kit, and elution was performed in elution buffer (EB) following the QIAquick protocol (Appendix 3).

For oligo annealing, the reaction mixture consisted of 1 µl of Oligo 1 (100 µM), 1 µl of Oligo 2 (100 µM), 1 µl of 10X T4 Ligation Buffer (NEB), 6.5 µl of ddH₂O, and 0.5 µl of T4 Polynucleotide Kinase (PNK) (NEB M0201S), making a total of 10 µl. This mixture was thermocycled at 37°C for 30 minutes, followed by 95°C for 5 minutes, and then cooled to 25°C at a rate of 5°C per minute. The annealed oligos were diluted 1:200 in sterile water.

The initial ligation and transformation were tested using Arid1a-1, Smarca4-1, and a non-targeting control (NT) oligo before expanding to Arid1a-2 and Smarca4-2. The ligation reaction was prepared by mixing 0.5 µl of the BsmBI-digested plasmid, 1 µl of the diluted oligo duplex, 5 µl of 2X Quick Ligase Buffer (NEB), and 3.5 µl of ddH₂O, followed by the addition of 1 µl Quick Ligase (NEB). A negative control, using water in place of oligos, was also set up. The ligated plasmids were then transformed into recombination-deficient Stbl3 bacteria, which were incubated on ice for 30 minutes, followed by a 30-second heat shock at 42°C. Afterward, 250 µl of pre-warmed S.O.C medium was added to each vial, and the vials were shaken at 37°C for 1 hour at 225 rpm. The transformation mixtures were then spread on pre-warmed ampicillin selective plates and incubated overnight at 37°C.

Initially, no bacterial colonies formed, prompting a repeat of the purification, annealing, ligation, and transformation steps. Once colonies formed, three were picked from each plate and cultured in 10 ml of L-Broth at 37°C overnight. Miniprep was performed, and DNA samples were sent for Sanger sequencing.

Colonies were selected based on sequencing results, and prepared for maxiprep using the QIAGEN® Plasmid Mini, Midi, and Maxi Kits (Appendix 5). The DNA was redissolved in 200 µl of ddH₂O and sent for further Sanger sequencing for confirmation.

3.2.3 Lentiviral transduction

The lentivirus used for transduction was produced by the senior research assistant from the Myant Lab. The production process involved co-transfecting the transfer plasmid, lentiCRISPRv2, into HEK293(F)T cells alongside the packaging plasmids pVSVg (AddGene 8454) and psPAX2 (AddGene 12260). Once the lentivirus was ready, AKP organoids were collected and prepared for transduction. The organoids were dissociated into single cells using Accutase, which was inactivated by adding 500 µl of 1% BSA in PBS.

For the transduction, Arid1a-1, Arid1a-2, Smarca4-1, Smarca4-2, and a non-targeting (NT) control lentiviruses were introduced into the AKP line, with a control group that received no virus. The lentiviruses used were designed to confer puromycin resistance, and puromycin was added at a concentration of 2 µg/ml on day 3 of transduction to select for successfully transduced cells. The growth of the organoids was carefully monitored, with passages conducted when necessary to maintain optimal conditions. After 48 hours of puromycin selection, the total number of surviving organoids was counted to assess the transduction efficiency and the success of the selection process. This step was crucial for further downstream analysis of gene knockouts and their effects on organoid growth and behaviour.

3.3 Verification of Knock-outs

3.3.1 qPCR

qPCR (quantitative PCR) allows for quantitative measurement of gene expression by providing precise data on mRNA levels. After CRISPR-Cas9 editing, successful knockouts should result in minimal to no expression of the targeted genes. By

comparing these mRNA levels in edited cells to non-edited controls, qPCR can confirm if the knock-outs have been effective.

For this project, primers were designed using SnapGene (39), with the full mouse genomic sequence for Arid1a and Smarca4 obtained from the UCSC Genome Browser (40). To familiarise with the qPCR protocol, initial tests were conducted to establish the baseline expression levels of Arid1a-1, Arid1a-2, Smarca4-1, and Smarca4-2 in the original AKP organoid line, as well as other organoid lines (APC^{-/-}, AP (APC^{-/-};P53^{-/-}), AKP, KPN (Kras^{G12/D};P53^{-/-}; Rosa^{N1icd}), WT (control)). From each transduction mix, either 7 or 8 clones were selected for testing – 8 clones for each of the two Arid1a sequences and 7 each for the Smarca4 sequences. The primers used for these experiments were ordered from Merck, ensuring specificity to the target genes (Table 2-3).

Name	Sequence (5'-3')	Scale	Purification
mSmarca4-forward1	CAAAGACAAGCATA TCCTAGCCA	25nm	STD
mSmarca4-reverse1	CACGTAGTGTGTGT TAAGGACC	25nm	STD
mSmarca4-forward2	GGTTCTGCCCACAG CATGAT	25nm	STD
mSmarca4-reverse2	GGA CTCCATAGGCT TGTGCAT	25nm	STD
mArid1a-forward1	CTTCCCCAACCACC ACCAGTACAA	25nm	STD
mArid1a-reverse1	CTGTGCGAAGGAC GAAGAC	25nm	STD
mArid1a-forward2	GACCCCTCAGTCAT CCAGTC	25nm	STD
mArid1a-reverse2	GAGTATGGGTAGT CCCACCA	25nm	STD

Table 2: Primer sequences for the initial qPCR reaction for baseline expression levels

Name	Sequence (5'-3')	Scale	Purification
Arid1a-1 (f)	GACCCCTCAGTCATCCAGT C	25nm	STD
Arid1a-1 (r)	TGGCCATACGCACTGGG	25nm	STD
Arid1a-2 (f)	CCGGACCCCTCAGTCATC	25nm	STD
Arid1a-2 (r)	GGTCCCTGTTGTTGCGAG	25nm	STD
Arid1a-3 (f)	GACCCCTCAGTCATCCAGT C	25nm	STD
Arid1a-3 (r)	GAGTATGGGTTAGTCCCAC CA	25nm	STD
Smarca4-1 (f)	ACTACCTCCACCCTCAGT	25nm	STD
Smarca4-1 (r)	ATGGGGTCTACTGTAATTT GG	25nm	STD
Smarca4-2 (f)	ATGCACAAGCCTATGGAGT C	25nm	STD
Smarca4-2 (r)	GACTGGAGGCATGTTCAGA G	25nm	STD
Smarca4-3 (f)	GGTTCTGCCCACAGCATGA T	25nm	STD
Smarca4-3 (r)	GGA CTCCATAGGCTTGTGC AT	25nm	STD

Table 3: Primer sequences for knock-out validation.

To pick and propagate clones, organoids were selected under a microscope, ensuring to avoid those positioned too closely together to maintain the purity of selection. Each clone was centrifuged to obtain a pellet, which was plated with 20µl of BME. The clones were grown and split regularly, with gradual expansion over seven days to prepare for RNA extraction and subsequent qPCR analysis.

For RNA extraction, pellets from the expanded Arid1a, Smarca4, and NT clones were processed. Each pellet was lysed by adding 350µl of buffer RLT and 3.5µl of B-ME (a toxic reagent) under a hood, followed by a 5-minute incubation at room temperature. Afterward, 350µl of 70% ethanol was added to each tube, and 700µl of the mixture was transferred to an RNeasy Mini spin column. The column was centrifuged at 8000 x g for 30 seconds, and the flow-through was discarded. Wash steps followed with 700µl of RW1 and two rounds of 500µl RPE buffer, with centrifugation between steps. The spin column was then dried and the RNA was eluted with 50µl of RNase-free water. For DNA removal, 0.1 volume of 10X DNase I Buffer and 1µl of rDNase I were added to the RNA sample, which was incubated at 37°C for 30 minutes. After this, 5.6µl of resuspended DNase Inactivation Reagent was mixed with the RNA sample and incubated at room temperature for 2 minutes with occasional mixing. The samples were centrifuged at 10,000 x g for 1.5 minutes, and the supernatant was transferred to a clean tube. Next, RNA concentration was quantified using a Thermo Scientific NanoDrop 2000 spectrophotometer. For cDNA synthesis, 500 ng of RNA was reverse transcribed in a reaction made up to 16µl with ddH₂O, followed by the addition of 4µl of reverse transcriptase (RT). This mixture was incubated in a thermocycler for 46 minutes, after which the reaction volume was diluted to 200µl with ddH₂O. For qPCR, primers were prepared by mixing 160µl of RNase-free water with 20µl each of forward and reverse primers. Beta-actin served as the control gene. Sybr Green master mix was prepared for each primer pair, and 7.5µl of master mix was loaded into each well along with 2.5µl of cDNA. The qPCR reaction was performed using a Bio-Rad CFX Connect Thermal Cycler to measure gene expression levels.

3.3.2 Western blot

For protein extraction, four wells from each clone were collected and centrifuged at 300 x g for 3 minutes, after which the supernatant was aspirated. The resulting pellets were lysed using RIPA buffer supplemented with phosphatase and protease inhibitors. The lysates were incubated on ice for 30 minutes, with vortexing every 10 minutes to ensure thorough lysis. To measure protein concentration, BSA standards were prepared using the ThermoFisher Scientific Pierce BCA Protein Assay Kit for a BCA assay. Protein concentration was read at 562 nm.

The Western blot protocol followed the experimental guidelines provided by the Myant Lab (Appendix 6). Primary antibodies used included anti-Arid1a antibody (ab182560) at

a dilution of 1:5000, and anti-BRG1 antibody (ab110641). The secondary antibody was HRP-conjugated anti-rabbit (1:10000), and the membrane was blocked with 5% milk in 0.1% PBST to prevent non-specific binding. The Western blot was repeated twice to optimise the protocol and ensure the validity and reproducibility of the results.

4 Results

4.1 Organoid maintenance

Out of the 8 Arid1a-1 clones that were initially picked, only 4 survived. It is suspected that an error during the splitting process may have caused the remaining 4 clones to undergo necrosis. One possibility is that puromycin may have been accidentally introduced to these clones, leading to their premature death. In contrast, all 8 Arid1a-2 clones, as well as 7 clones from both the Smarca4-1 and Smarca4-2 groups, survived and were thriving. These surviving clones showed robust growth and were successfully propagated for further analysis.

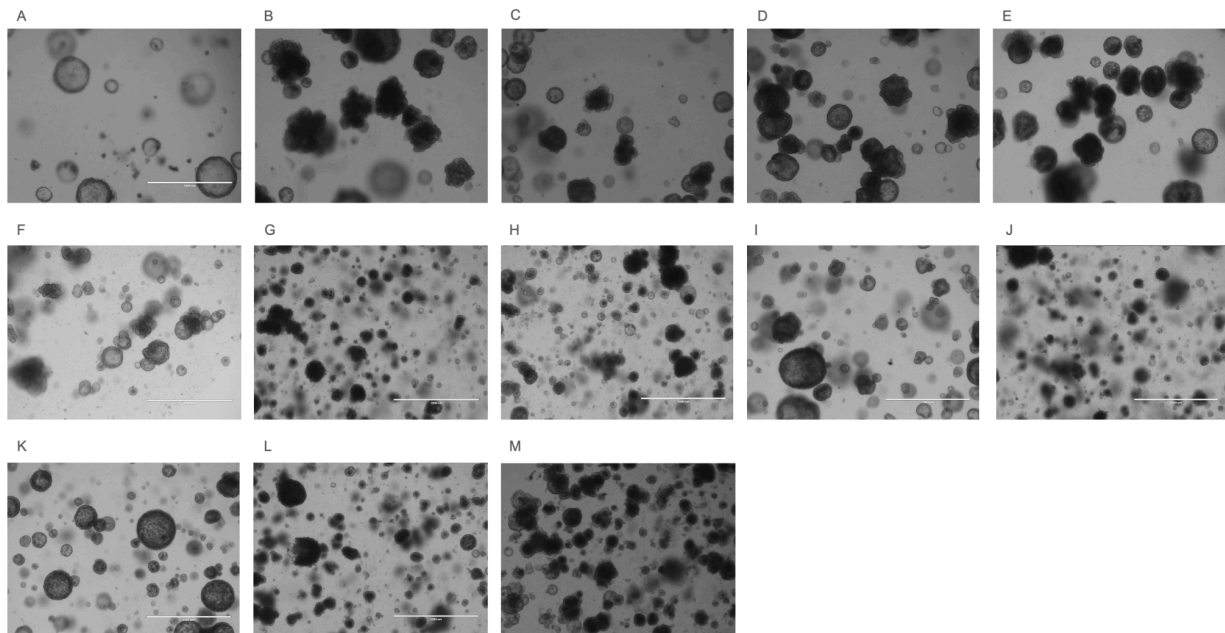


Figure 2: Mouse colorectal cancer organoids. A: AKP original; B: Arid1a-1 clone 1; C: Arid1a-1 clone 2; D: Arid1a-1 clone 3; E: Arid1a-1 clone 4; F: Smarca4-1 clone 1; G: Smarca4-1 clone 2; H: Smarca4-1 clone 3; I: Smarca4-1 clone 4; J: Smarca4-1 clone 5; K: Smarca4-1 clone 6; L: Smarca4-1 clone 7; M: Non-target control.

4.2 Sanger sequencing

4.2.1 Miniprep results

Sanger sequencing analysis of the StbI3 bacterial transformations yielded positive results for several clones. All three colonies from the Arid1a-1 StbI3 transformation exhibited successful miniprep Sanger sequencing results, confirming the accurate incorporation of the intended genetic modifications. Similarly, all three Smarca4-1 and non-targeting (NT) colonies showed positive sequencing outcomes.

In the Smarca4-2 group, two out of three colonies were confirmed to have successful transformations, while the third did not yield the expected results. Likewise, two out of three Smarca4-3 colonies showed positive sequencing data. However, the Arid1a-2 and Arid1a-3 colonies had lower success rates. Only one out of three colonies in both groups produced positive miniprep Sanger sequencing results. Despite these variations, the majority of clones exhibited successful sequencing, allowing for further progression in the validation and experimentation process.

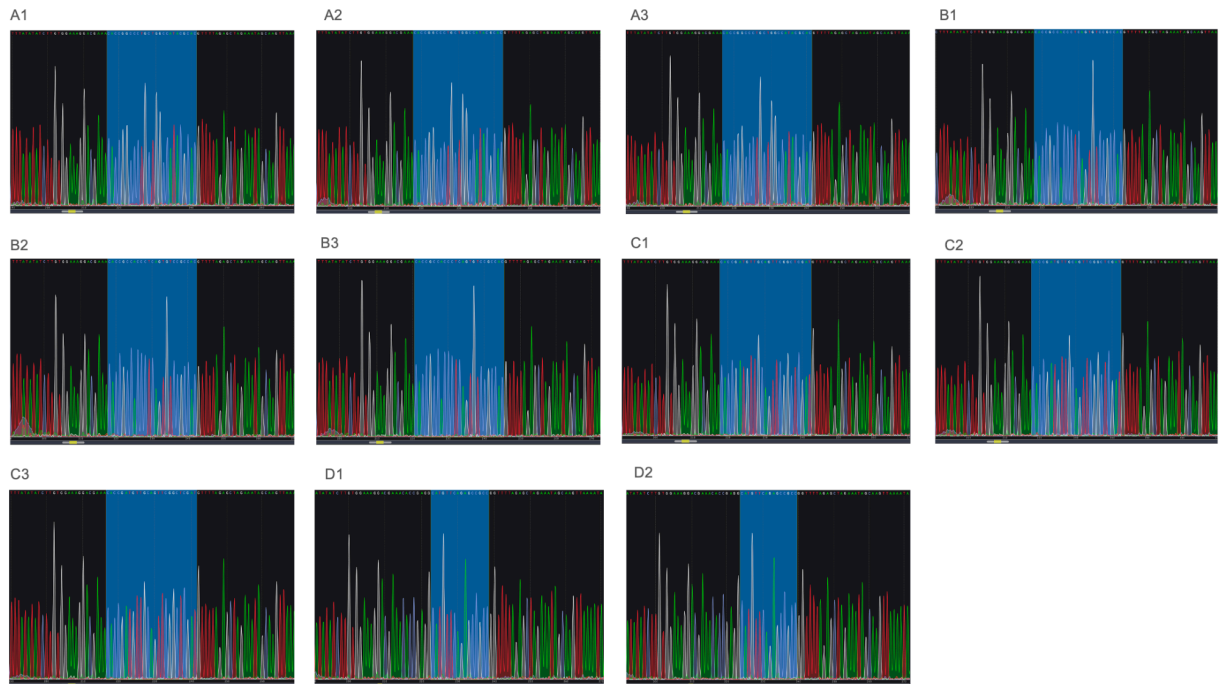


Figure 3a: SWI/SNF transformed StbI3 miniprep sequencing. Successfully transformed plasmid sequence is highlighted blue. A1: Arid1a-1 colony 1; A2: Arid1a-1 colony 2; A3: Arid1a-1 colony 3; B1: Smarca4-1 colony 1; B2: Smarca4-1 colony 2; B3: Smarca4-1 colony 3; C1: NT colony 1; C2: NT colony 2; C3: NT colony 3; D1: Smarca4-2 colony 2; D2: Smarca4-2 colony 3.

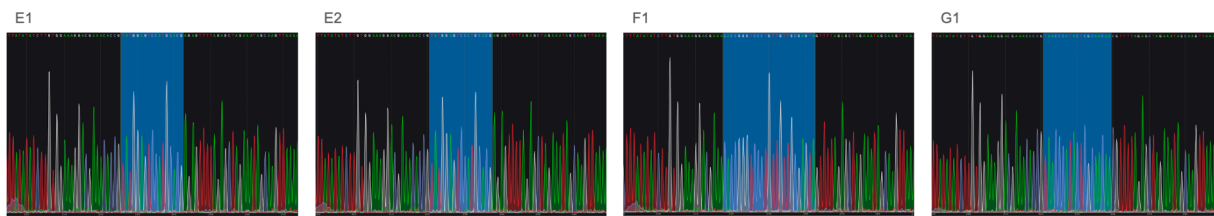


Figure 3b: SWI/SNF transformed Stbl3 miniprep sequencing. Successfully transformed plasmid sequence is highlighted blue. E1: Smarca4-3 colony 1; E2: Smarca4-3 colony 2; F1: Arid1a-2 colony 1; G1: Arid1a-3 colony 1.

4.2.2 Maxiprep results

From each positive miniprep pool, one colony was carefully selected for each oligo targeting Arid1a and Smarca4, and maxiprep was performed to produce sufficient quantities of plasmid DNA for the next stages of the experiment. After DNA extraction, Sanger sequencing was conducted to verify the presence and correct insertion of the guide RNA sequences into the lentiv2 plasmid. All the maxiprep samples returned positive sequencing results, confirming that the plasmid constructs had successfully incorporated the respective guide sequences for targeting Arid1a and Smarca4. These results were crucial, as the accurate assembly of these constructs was a key step toward the knockout of Arid1a and Smarca4 in the AKP mouse colorectal cancer organoid line. With the successful verification of the plasmid constructs, the next steps of the experiment—viral packaging, transfection into the mouse colon organoids, selection under puromycin, and subsequent propagation of the clones—were set in motion. Following the selection of clones that had undergone successful viral transduction, the experiment would proceed to RNA sequencing to assess gene expression changes, verifying the knockouts at the mRNA level.

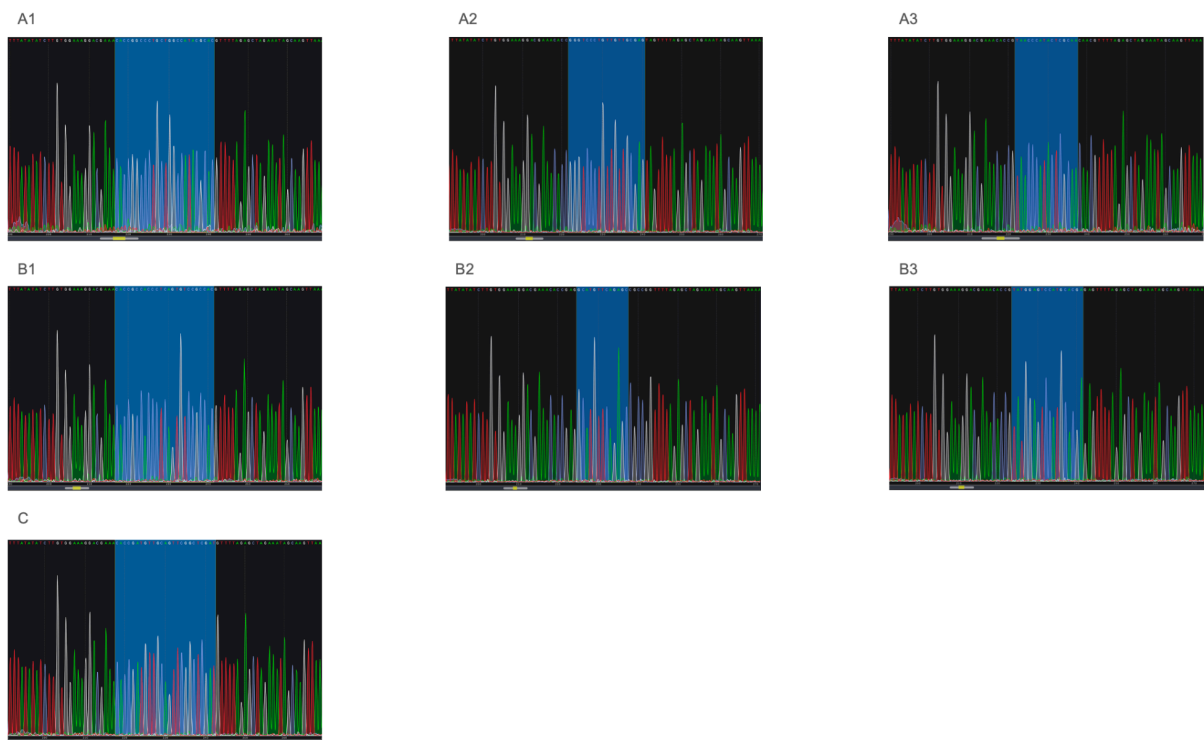


Figure 4: SWI/SNF transformed Stbl3 maxiprep sequencing. Successfully transformed plasmid sequence is highlighted blue. A1: Arid1a-1; A2: Arid1a-2; A3: Arid1a-3; B1: Smarca4-1; B2: Smarca4-2; B3: Smarca4-3; C: Non-targeting control (NT).

4.3 qPCR results

4.3.1 Baseline mRNA expression

Baseline expression of Arid1a and Smarca4 in different mouse colon organoid lines was assessed using qPCR to establish the starting levels of these genes prior to conducting CRISPR-Cas9 knockouts. The organoid lines tested included APC^{-/-}, AP (APC^{-/-}; P53^{-/-}), AKP (APC^{-/-}; KrasG12D; P53^{-/-}), KPN (KrasG12D; P53^{-/-}; RosaN1icd), and WT (wild-type control). The goal of this experiment was to determine whether Arid1a and Smarca4 were expressed at significant levels across these lines, a prerequisite for successful gene knockouts, as the CRISPR-Cas9 system targets actively expressed genes for silencing.

The qPCR data showed that both *Smarca4* and *Arid1a* had relatively high mRNA expression across all the tested organoid lines. This was particularly important in the AKP line, the primary model system used in this project.

For *Smarca4*, the mean Ct (cycle threshold) values, which represent the number of amplification cycles required to detect the gene, were as follows: 25.07 in WT, 24.89 in KPN, 25.85 in AKP, 24.87 in AP, and 25.99 in APC. These values indicate that *Smarca4* expression is fairly consistent across the different organoid lines, with slight variations between them. Similarly, *Arid1a* expression levels showed mean Ct values of 25.01 in WT, 26.05 in KPN, 26.10 in AKP, 25.36 in AP, and 26.66 in APC. These results also revealed relatively stable *Arid1a* expression across the organoid lines, confirming that both target genes were present at detectable levels in all samples.

The significance of these findings is twofold: first, they confirm that *Smarca4* and *Arid1a* are actively expressed in the mouse colon organoids, making them viable candidates for CRISPR-Cas9-mediated knockout; second, the consistent expression across multiple organoid lines ensures that subsequent experiments, including viral transduction, selection, and validation via qPCR and protein blotting, can proceed with confidence that the genes targeted for knockout are indeed expressed in the system. This baseline expression data will serve as a reference point for evaluating the effectiveness of gene knockout in later stages of the study.

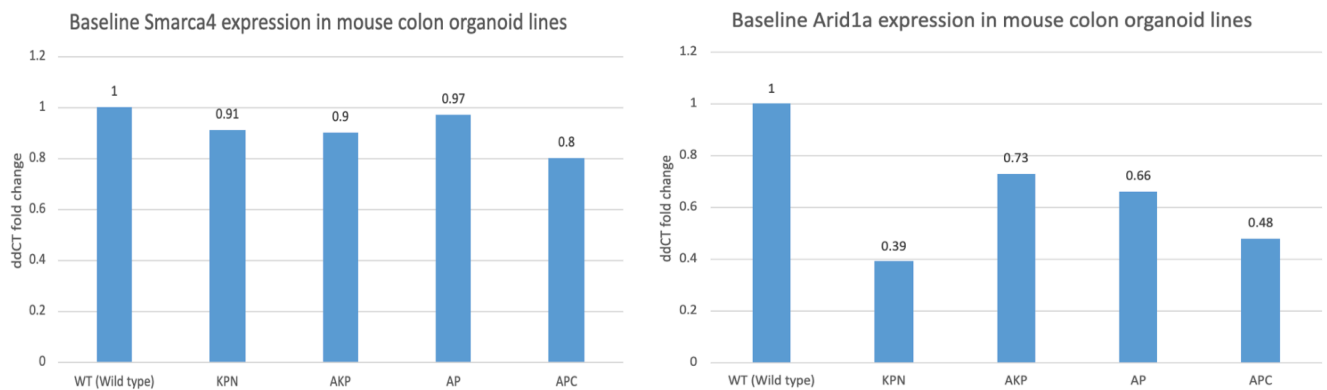


Figure 5: Baseline *Arid1a* and *Smarca4* mRNA expression in various mouse colon organoid lines.

4.3.2 qPCR results of knock-outs

The qPCR results for *Arid1a*-1 clones were highly promising. Clones 1 through 7 showed mean CT (cycle threshold) values above 40, indicating that there was negligible

RNA expression in the samples. To ensure the reliability of these findings, the qPCR was repeated, and similar results were obtained. Unfortunately, clones 5, 6, and 7 were lost due to human error during handling, leaving clones 1 through 4 available for further Western blot validation to assess protein-level knockout.

The qPCR results for the Smarca4-1 clones were more variable. The mean Ct values ranged from 27.34 to 34.00 across the seven clones, indicating varying levels of Smarca4 expression post-CRISPR-Cas9 editing. Clone 6 stood out as the most promising, with an initial mean Ct value of 34.00 with a ddCt fold change of 0.068 relative to the reference (NT control), which suggests a potential reduction in RNA levels compared to the other clones. To further assess the consistency of these results, the qPCR was repeated, and clone 6 showed a slightly reduced mean Ct value of 29.33, with a ddCt fold change of 0.71.

The qPCR of Arid1a-2 clones revealed mean Ct values ranging from 24.50 to 26.72, indicating that these clones still exhibited moderate levels of Arid1a expression. While none of the clones displayed as dramatic a reduction as the Arid1a-1 clones, the consistency of the Ct values suggests that further efforts may be required to optimise the knockout in this group.

For the Smarca4-2 clones, the mean Ct values ranged from 25.16 to 29.40. Clone 3 and clone 5 both showed higher mean Ct values (29.31 and 29.40, respectively). Their ddCt values were 0.19 and 0.22 respectively, which were much lower compared to the NT control reference, indicating potential reductions in Smarca4 expression compared to other clones in the group. As with the other groups, qPCR results were consistent upon repetition, confirming the findings.

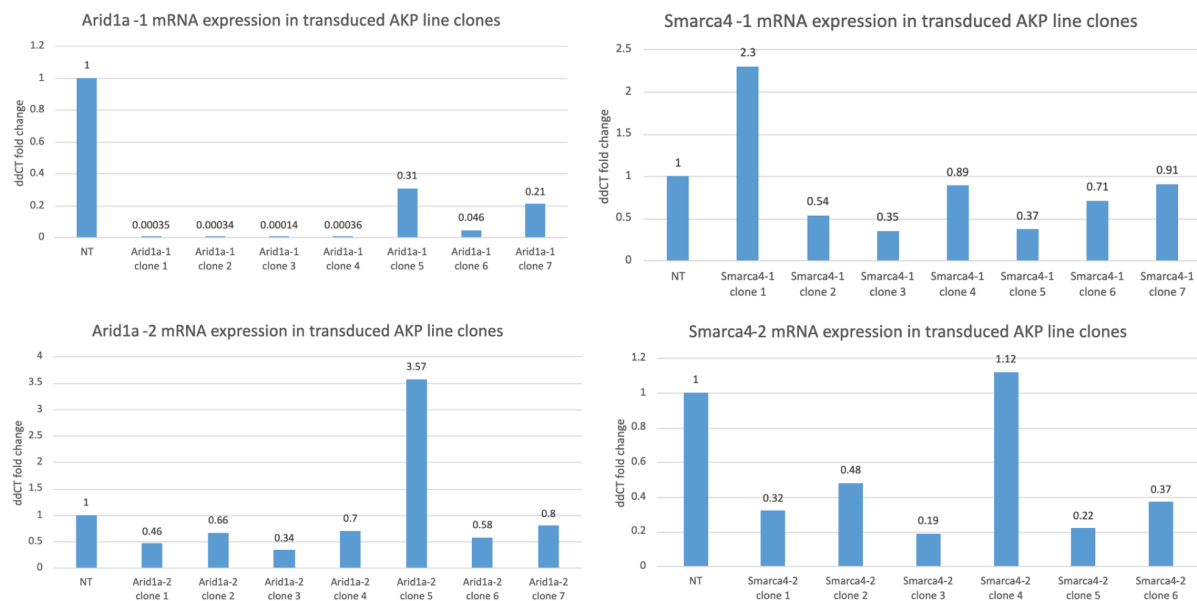


Figure 6: Arid1a and Smarca4 mRNA expression in the AKP organoid line Arid1a-1, Arid1a-2, Smarca4-1 and Smarca4-2 clones.

3.3 Western blot results

Western blot analysis was conducted on Arid1a-1 clones 1 through 4 to assess the presence of the Arid1a protein following CRISPR-Cas9 knockout. The expected molecular weight of the mouse Arid1a protein is approximately 270 kDa. To visualise the protein bands, a 250 kDa molecular weight ladder was used as a reference (Figure 7). Bands slightly above the 250 kDa marker were observed in clones 1, 2, and 4, suggesting the presence of some residual protein, potentially due to incomplete knockout or post-translational modifications. However, clone 3 demonstrated more promising results, with the heaviest band appearing at the 250 kDa marker, indicating a more effective reduction or absence of the Arid1a protein. These findings suggest that clone 3 may represent a more successful knockout, warranting further validation.

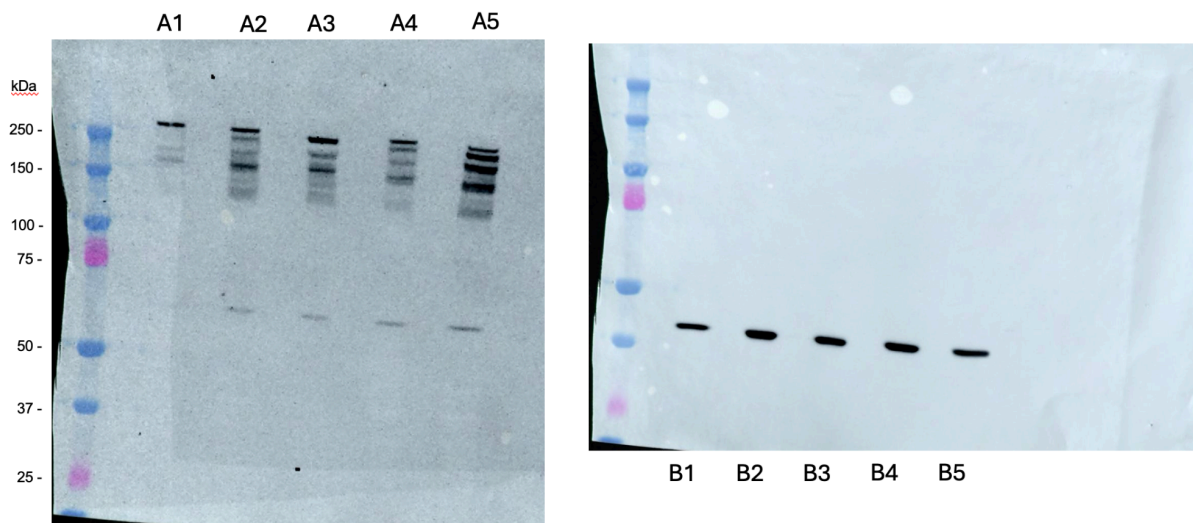


Figure 7: Arid1a-1 clones 1 to 4 western blot protein expression. A1: NT; A2: Arid1a-1 clone 1; A3: Arid1a-1 clone 2; A4: Arid1a-1 clone 3; A5: Arid1a-1 clone 4. B1: NT ActB; B2: Arid1a-1 clone 1 ActB; B3: Arid1a-1 clone 2 ActB; B4: Arid1a-1 clone 3 ActB; B5: Arid1a-1 clone 4 ActB.

5 Discussion

The qPCR results provide valuable insights into the expression levels of Arid1a and Smarca4 genes in various organoid lines. Baseline measurements of Arid1a and Smarca4 expression indicate that there is a substantial level of gene expression in the AKP organoid line, confirming the presence of these genes prior to CRISPR-Cas9 editing. This baseline expression is critical as it establishes that both genes are active in the AKP line, making it a suitable model for studying gene disruption.

In the case of the Arid1a-1 clones, the qPCR results show a near-complete absence of mRNA expression, which strongly suggests successful disruption of the Arid1a gene at the transcriptional level. This indicates that the CRISPR-Cas9-mediated knockout was effective in significantly reducing or eliminating Arid1a expression.

The Western blot results for the Arid1a-1 clones are less definitive compared to the qPCR findings. While clone 3 showed results consistent with a potential full knockout, the other clones produced bands that do not clearly indicate a complete absence of the Arid1a protein. This ambiguity suggests that further optimization of the experimental conditions is necessary to obtain clearer results.

Another factor that could cause complications during the interpretation of the western blot is the fact that although the mouse Arid1a protein has a calculated molecular weight of 270 kDa, the actual observed molecular weight may differ from the predicted value due to factors like post-translational modifications, protein cleavage, and other influences. Several improvements could enhance the accuracy of future Western blot analyses. First, using a protein ladder with a higher molecular weight range that includes 270kDa would allow for more precise identification of the expected protein band. Second, testing a different antibody could increase the specificity and sensitivity of the protein detection. Finally, extending the gel running time could help to improve the separation of protein bands, making it easier to distinguish between individual proteins and better assess the presence or absence of the Arid1a protein.

These adjustments would likely yield more conclusive data, helping to verify whether a full knockout of Arid1a has been achieved in the remaining clones.

The combined results from both qPCR and Western blot suggest that clear knockouts of Arid1a have not been achieved in the clones tested. While the near-absence of mRNA expression in qPCR for some clones suggests successful gene disruption at the transcriptional level, the Western blot results do not strongly support this, with clone 3 being the only one showing potential for a full knockout. The changing band pattern between the clones indicates that the antibody is likely functional, but the inconsistency in protein expression suggests that the knockouts may be incomplete or that only partial deletions of the gene have occurred. These partial deletions may have caused the loss of small DNA segments without fully disrupting the gene's function, resulting in incomplete knockouts. To clarify the extent of gene disruption, further steps are needed. Specifically, performing PCR and DNA sequencing of the targeted regions will help determine if small fragments of the gene remain intact. Based on these findings, it may also be necessary to design new CRISPR guides to target and delete the regions that have not been fully disrupted.

In a comprehensive review published in 2022 (41), 23 studies investigating ARID1A in colorectal cancer (CRC) were compiled, all of which consistently reported that ARID1A functions as a specific tumour suppressor in CRC. The studies highlighted that ARID1A mutations often co-occur with variations in key genes such as TP53, KRAS, APC, FBXW7, and PIK3CA, and that ARID1A plays a role in regulating critical signalling pathways, including WNT, Akt, and MEK/ERK pathways. These findings provide crucial directions for future research, particularly in this project once valid Arid1a knock-outs in the AKP mouse colon organoid line are verified. Following similar approaches, such as genome-wide ARID1A occupancy and epigenomic mapping, may help to decipher

ARID1A-dependent transcriptional regulatory mechanisms. To further explore phenotypic outputs, in vivo models such as subcutaneous and orthotopic transplantation could be employed to assess the effects of ARID1A loss following in vitro validation.

Madhobi et al. conducted a study in 2022 (42) on ARID1A-deficient colorectal cancer using CRISPR/Cas9-mediated gene editing, a technique also used in this project. Their findings demonstrated that ARID1A is essential for Kras-mutated CRC cells, particularly in regulating the MEK/ERK pathway at the transcriptional level. In this study, guide RNAs were designed to target intronic regions flanking exon 5 of ARID1A, inducing a frameshift mutation in the spliced transcript. They confirmed the knock-out (KO) by genotyping PCR, showing a 761-bp product in the KO cells compared to 1205-bp in the parental line, resulting in complete loss of the protein. The combination of CRISPR/Cas9 and these validation techniques allowed the researchers to produce cellular models to study ARID1A loss within the context of additional mutations commonly seen in CRC.

The Madhobi study utilised a range of validation methods, including qPCR, Western blot, and genotyping PCR, to confirm successful knock-outs. While this current project has not yet employed genotyping PCR due to time constraints, it would serve as a logical next step for validating Arid1a-1 clones. Additionally, Madhobi et al. minimised off-target effects by using the MIT CRISPR design tool, an approach that could be useful in this project. However, a notable difference is that Madhobi et al. used electroporation with the pSpCas9(BB)-2A-GFP (PX458) plasmid to deliver CRISPR components into cells, whereas this project used lentiviral transduction. Electroporation offers potential advantages over viral transduction, such as greater flexibility in cargo size and elimination of risks associated with insertional mutagenesis (43). A study from 2023 (44) further explored RNP-based CRISPR approaches, utilising electroporation to target ARID1A in human gut organoids, achieving high KO efficiency (up to 98%). This method allowed for the generation of edited organoids within 26 days, suggesting that electroporation could be a promising alternative for improving KO efficiency in this project.

Partial knock-outs, as observed in this project, could arise from several factors. One possibility is that residual proteins from knock-out genes might partially or fully rescue the gene's function. This can occur through translation reinitiation or alternative splicing mechanisms (45). Additionally, in-frame deletions may have taken place. If the length of the deleted sequence is a multiple of three base pairs, the deletion might not disrupt all gene functions, allowing the protein to still be produced. In such cases, while the protein may have lost or gained a few amino acids, these alterations may not affect a critical

domain and may not have functional significance if the affected region is nonconserved (46). Another potential reason for partial knock-outs is the targeting of a genome-editing tool to an exon that is not present in all transcripts of a gene (47). In this scenario, transcripts that do not contain the targeted exon would remain unaffected, leading to only a partial knock-out.

These insights provide both explanations for the partial knock-outs observed in the current project and potential avenues for refining the CRISPR/Cas9 methodology, including employing genotyping PCR, electroporation, and more precise guide RNA design.

Part 2: Scoring Tumour Microarrays for Protein Marker Expression in Ovarian Cancer: Insights for ADC Targeted Therapy

1 Background

1.1 Ovarian cancer

Worldwide, ovarian cancer ranks as the eighth most common cancer among women, contributing to approximately 3.7% of cancer diagnoses and 4.7% of cancer-related deaths in 2020 (48), accounting for over 200,000 deaths per year worldwide (49). Ovarian cancer is categorised into several distinct types based on the cellular origin of the malignancy. The most prevalent form, epithelial ovarian carcinoma, constitutes approximately 90% of cases (48,50). This subtype includes histotypes such as high-grade and low-grade serous carcinomas, mucinous carcinomas, endometrioid carcinomas, and clear cell carcinomas (51). Less commonly, germ cell tumours arise from the totipotent cells responsible for oogenesis and are more frequently observed in younger females (52). Each of these ovarian cancer types exhibits distinct molecular and histopathological characteristics, which significantly influence clinical management, prognosis, and therapeutic strategies (54). The three histotypes that this project focused on were high grade serous ovarian carcinoma, endometrioid ovarian carcinoma, and clear cell ovarian carcinoma.

1.1.1 Patient cohort

Existing cohorts with detailed clinical and molecular annotation from previous studies performed by Prof. Charlie Gourley's and Dr Robb Hollis' labs were investigated in this project (55,56). The CCOC cohort curation is currently unpublished. The process of the curation of these existing cohorts are described below:

High-grade serous ovarian carcinoma (HGSOC) is the predominant subtype of ovarian cancer, representing approximately 75% of cases with the majority of patients experiencing disease relapse and the subsequent development of chemoresistance, ultimately resulting in treatment failure (57). Genomic and transcriptomic characteristics have been linked to variations in patient outcomes and therapeutic responses (58). An FFPE tumour cohort of HGSOC was interrogated for the ADC target markers. Out of the 539 ovarian cancer patients treated at the Edinburgh Cancer Centre who fulfilled the inclusion criteria— diagnosis of primary ovarian, peritoneal, or fallopian tube carcinoma before 2007, availability of formalin-fixed paraffin-embedded (FFPE) treatment-naïve surgical specimens, administration of first-line platinum-based chemotherapy, and a follow-up period of at least three years— 362 were identified as high-grade serous ovarian carcinoma (HGSOC) following pathology review and were subjected to molecular characterisation.

Endometrioid ovarian carcinoma (EnOC), a type of ovarian cancer that is associated with endometriosis, represents about 10% of all ovarian cancer cases, with the average age of diagnosis falling between 54 and 58 years (59). A significant portion of these cases are detected at an early stage, with roughly half identified at FIGO stage I, 20–40% at stage II, and 10–25% at stages III/IV (60). In terms of pathological grading, the majority are low-grade, with 50–60% classified as grade 1, about 30% as grade 2, and 20% as grade 3. EnOC typically has a better prognosis and responds relatively well to chemotherapy, particularly platinum-based regimens, where the response rate is approximately 60% (61). However, some patients present with more advanced and aggressive forms of the disease, which correlates with poorer outcomes. Despite this, data from well-defined EnOC patient groups remain scarce. 505 patients diagnosed with ovarian cancer between August 1968 and May 2014, and whose pathology reports contained the term “endometrioid”, were identified through the Edinburgh Ovarian Cancer Database. Tumour material was available for 289 cases. 137 cases were included for this project .

Clear cell ovarian carcinoma (CCOC) is classified by the World Health Organization (WHO) 2014 guidelines as one of the five principal histological subtypes of ovarian

tumours (62). Recent data from the Surveillance, Epidemiology, and End Results (SEER) program report that CCOC constitutes 4.8% of epithelial ovarian cancers in Caucasian women, 3.1% in African American women, and 11.1% in Asian women within the United States (63). CCOC is frequently associated with endometriosis, a common gynaecological condition in women of reproductive age, which is widely considered a precursor lesion for many CCOC cases. CCOC is typically diagnosed at a younger age and at an earlier clinical stage compared to high-grade serous ovarian carcinoma (HGSOC). In early-stage disease, CCOC has a prognosis that is comparable to or better than HGSOC (64). However, in advanced-stage disease, patients with CCOC exhibit a poorer prognosis, largely due to the tumour's intrinsic resistance to chemotherapy. For the CCOC cohort, cases diagnosed between 1980 and 2016 were identified from the Edinburgh Ovarian Cancer Database, yielding 85 CCOC cases. DNA sequencing of these tumours revealed mutational frequencies consistent with those reported in published studies from Asian cohorts, with ARID1A and PIK3CA mutations being the most commonly observed genomic alterations.

1.2 Tumour Microarrays

Tissue microarrays (TMAs) consist of tissue cores from multiple tumours that are constructed into blocks and sectioned before being arranged on a single glass slide (65). They serve as a valuable tool for identifying novel diagnostic and prognostic markers, as well as molecular targets in cancer research. Additionally, TMAs can be utilised to identify patient-derived xenograft (PDX) models for cancer types expressing specific targets of interest. TMAs are compatible with a range of analytical techniques, including immunohistochemistry, fluorescence in situ hybridization (FISH), and in situ hybridization assays (66). Because it allows hundreds of different cases to be assembled on a single histologic slide, it allows high throughput analysis of multiple specimens at the same time. Tissue microarrays are paraffin blocks produced by extracting cylindrical tissue cores from different paraffin donor blocks and re-embedding these into a single recipient (microarray) block at defined array coordinates. This versatile technique facilitates retrospective and prospective human tissue studies. It is a practical and effective tool for high-throughput molecular analysis of tissues that is helping to identify new diagnostic and prognostic markers and targets in human cancers, and has a range of potential applications in basic research, prognostic oncology and drug discovery. In this project, tissue microarrays were taken from the patients in the HGSOC, EnOC, and CCOC cohorts, stained for several ADC (antibody

drug conjugate) targets and scanned into digital NPD viewer compatible images. Scoring of these stained TMAs was enabled through this process.

1.3 Antibody-Drug Conjugate (ADC) therapy

Despite considerable progress in the treatment of cervical, ovarian, and uterine cancers, particularly with the integration of checkpoint inhibitors and PARP inhibitors into standard care, patients with recurrent metastatic gynecologic malignancies continue to face poor prognosis, with most experiencing disease relapse (67). Once standard treatment options have been exhausted, available therapies have historically been associated with limited efficacy and significant toxicities. As a result, there is a critical need for novel, effective, and well-tolerated therapies for these patients. Antibody-drug conjugates (ADCs), a class of targeted therapies, have shown efficacy in various cancers, including hematologic malignancies and certain solid tumours (68). Advances in ADC technology and design have improved both their safety and effectiveness, leading to increasing interest in their use for gynecologic cancers. This is exemplified by the recent approvals from the U.S. Food and Drug Administration of tisotumab vedotin for cervical cancer and mirvetuximab soravtansine for ovarian cancer (69).

ADCs are designed to deliver cytotoxic agents directly to tumour cells or the tumour microenvironment by utilising antibody-antigen interactions (70). This approach aims to enhance drug delivery precision, improving clinical efficacy while reducing toxicity. However, ADCs still present toxicity challenges, highlighting the need for further refinement. ADCs consist of three key components: the antibody, the cytotoxic payload, and the linker connecting them. Their mechanism of action involves binding to a tumour antigen, internalisation into the cancer cell, degradation, and release of the cytotoxic payload, which targets microtubules or DNA, leading to cell death. Optimising each of these components is crucial for maximising therapeutic efficacy and minimising adverse effects (69).

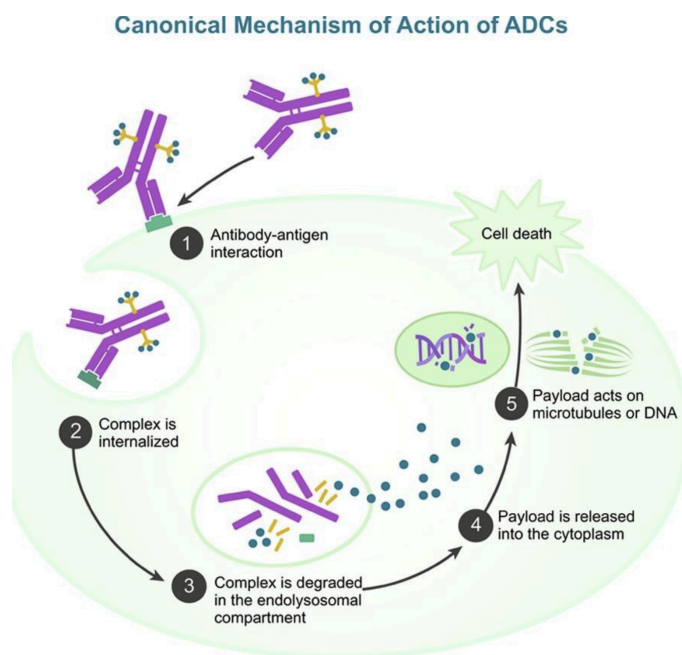


Figure 8: Canonical mechanism of action of ADCs. Image acquired from Tolcher et al. (69)

1.3.1 ADC targets

For many available ADCs, target expression is essential for their clinical activity (71). For example, mirvetuximab soravtansine is significantly more active in FOLR1 expressing ovarian cancers (72). FOLR1, or folate receptor alpha, is a protein that helps regulate the transport of folate, or vitamin B9 into cells (73). FOLR1 is overexpressed in ovarian cancer tissue and serum. It's associated with tumour stage and increases sensitivity to certain chemotherapy agents such as cisplatin treatment in ovarian cancer cells (74). When FOLR1 is present in at least 75% of tumour cells, it can predict a response to treatment with mirvetuximab soravtansine (75).

VTCN1/B7-H4 and Trop2/TACSTD2 are also ADC targets that are highly expressed in endometrial carcinomas that are relatively insensitive to traditional chemotherapies (76). VTCN1, also known as B7-H4, is an immune checkpoint molecule and member of the B7 family of costimulatory proteins (77). It suppresses proinflammatory responses, MHC class I expression, and guides differentiation. It's expressed on the surface of antigen-presenting cells and interacts with ligands on T lymphocytes. High levels of VTCN1/B7-H4 have been correlated with tumour progression (78). Trop2, or otherwise known as TACSTD2 (short for tumour-associated calcium signal transducer 2), is a transmembrane glycoprotein that is overexpressed in many cancers and is associated

with tumour growth, metastasis, and prognosis. It's shown to be ubiquitously expressed in cancer cells with a relatively low expression in most normal tissues (79).

Studies have shown that FOLR1, VTCN1/B7-H4 and Trop2/TACSTD2 are targets that are relevant to ovarian cancer and endometrial cancer (80,81). Furthermore, the expression of multiple ADC targets on the same cancer raises the possibility of improving therapeutic effects by developing bispecific ADCs or combining ADC therapies (82). In this project, potential ADC targets FOLR1, VTCN1/B7-H4, and Trop2/TACSTD2 are assessed in TMAs obtained from patients with three different ovarian cancer histotypes: CCOC, EnOC, and HGSOC, to evaluate the level of expression in the different histotypes and between individual patients who have the same histotype.

1.4 Histological scoring

Histopathologic scoring is a method used to generate semiquantitative data from tissue samples. For pathologists to effectively assess tissues and establish scoring criteria, a comprehensive understanding of the experimental design, study objectives, and methodologies is essential. The development of a reliable scoring system involves several key principles, including careful tissue examination, accurate identification, clear scoring definitions, and consistency in interpretation throughout the process (83).

2 Methodology

2.1 TMA scoring

A total of 11 tissue microarray (TMA) slides were assessed, with the scores recorded using Excel. Of these, six slides represented high-grade serous ovarian carcinoma (HGSOC) in two triplicates, three were endometrioid ovarian carcinoma (EnOC) slides, and two were clear cell ovarian carcinoma (CCOC) slides. The slides were viewed using NDP.view2 software to assess the expression of three antibody-drug conjugate (ADC) targets: FOLR1, VTCN1/B7-H4, and TROP2/TACSTD2. Each TMA core, stained with target-specific dyes, represented a different patient, allowing the visualisation of target protein expression in a range of ovarian cancer subtypes.

To refine the scoring protocol and correlation analysis, the level of expression of an additional membrane protein, P-glycoprotein (p-GP) in the CCOC and EnOC slides was assessed at first. p-GP is a membrane transporter involved in the uptake and efflux of drugs, it plays a crucial role in determining drug concentrations within plasma and tissue. The evaluation of p-GP expression helped improve the approach for scoring other target proteins. Histological scoring was performed based on an established histoscore system adapted for membrane staining. This method utilises a semiquantitative approach that combines expression levels and percentage to generate a single score for each target in each TMA core. The expression levels of the target proteins were categorised into 4 groups: 0= no expression, 1= weak expression, 2= moderate expression, 3= strong expression. Standards for each expression level were established prior to scoring to ensure consistency and minimise errors.

The scoring process involved determining the percentage of the core that exhibited each expression level, with the percentage for each category multiplied by its corresponding level number (0,1,2, or 3). These values were then summed to generate a final score for the core. This scoring procedure was repeated for all cores across the TMA slides, ensuring a comprehensive evaluation of target protein expression across the various ovarian cancer subtypes.

2.2 Inter-assessor agreement evaluation

The inter-assessor agreement was evaluated through correlation analysis using the Intraclass Correlation Coefficient (ICC). The ICC is a statistical measure designed to quantify the degree of similarity or consistency between outcomes provided by different individuals within a group. It is commonly used to assess the reliability of ratings or measurements in clustered data, making it ideal for evaluating the agreement between different scorers (84).

In this case, the correlation coefficient was calculated between the scores assigned by myself and the second scorer, Joanna, to determine the level of concordance in our assessments. This analysis was performed using the Pandas statistical module, a widely utilised open-source library built on the Python programming language (85). Pandas is specifically designed for data manipulation and analysis, offering efficient tools to compute complex statistical measures such as the ICC, ensuring an accurate evaluation of inter-assessor agreement (86). The ICC is chosen over other measures of inter-rater agreement such as Pearson's correlation coefficient because it is flexible and

it incorporates the magnitude of disagreement rather than just an all-or-nothing agreement (87).

3 Results

3.1 Correlation analysis for inter-assessor agreement

The initial analysis and coding pipeline were tested using P-GP expression scores from both endometrioid ovarian carcinoma (EnOC) and clear cell ovarian carcinoma (CCOC) tissue microarrays. The Intraclass Correlation Coefficient (ICC) was employed to assess the level of agreement between the two scorers. A higher ICC score indicates stronger agreement. For the P-GP scores in CCOC, the ICC value was 0.799, signifying good agreement between the scorers (Figure 9). This result aligns with established thresholds for interpreting ICC values (84): scores below 0.50 indicate poor agreement, scores between 0.50 and 0.75 indicate moderate agreement, scores between 0.75 and 0.90 represent good agreement, and scores above 0.90 reflect excellent agreement.

In contrast, the ICC score for P-GP expression in EnOC was 0.351, reflecting poor agreement between the two scorers. Due to this low level of concordance, the data required re-evaluation by a third arbiter to ensure the accuracy and reliability of the scoring process.

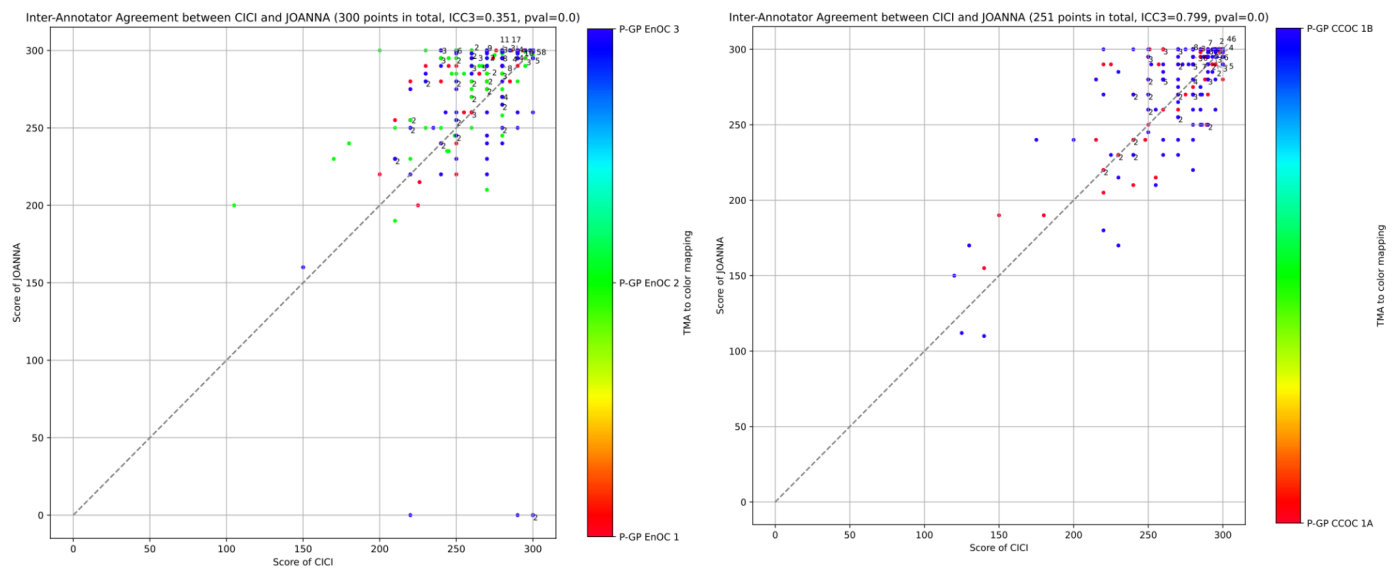
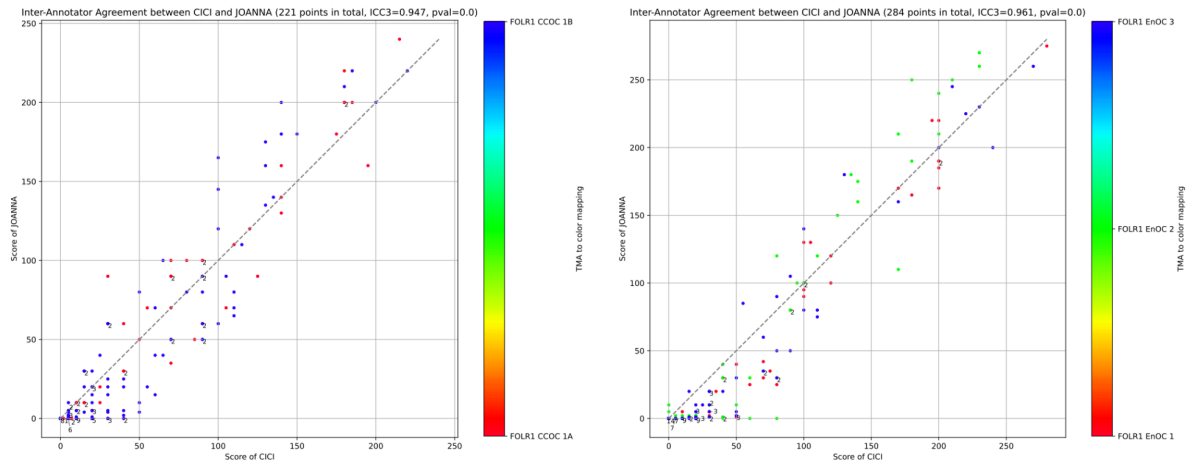


Figure 9: Inter-assessor agreement between scorer 1 (Cici) and scorer 2 (Joanna) for P-GP expression evaluation in TMAs derived from patients with the CCOC and EnOC histotypes. The

Intraclass correlation coefficient (ICC) score is 0.799 for CCOC P-GP and 0.351 for EnOC, with negligible p-values.

The agreement of the target expression scores for FOLR1, TACSTD2, and VTCN1 were subsequently calculated using the optimised pipeline developed with the P-GP data. The Intraclass Correlation Coefficient (ICC) values for FOLR1 expression in CCOC, EnOC, HGSOC 1, and HGSOC 2 were 0.947, 0.961, 0.91, and 0.897, respectively, indicating very high agreement between scorers. These results further support the validity of the FOLR1 scores, as ICC values above 0.90 are considered excellent according to standard interpretive thresholds. For TACSTD2, the agreement values in CCOC, EnOC, HGSOC 1, and HGSOC 2 were slightly lower, at 0.881, 0.843, 0.89, and 0.885, respectively. Although these scores do not all surpass the 0.90 threshold, they still fall within the 0.75–0.90 range, which is indicative of good agreement. The TACSTD2 scores, while not as high as those for FOLR1, remain reliable, as they are very close to the 0.90 mark. Finally, the agreement scores for VTCN1 expression were also calculated, yielding values of 0.902 in CCOC, 0.869 in EnOC, 0.795 in HGSOC 1, and 0.829 in HGSOC 2. These results demonstrate very good agreement, with the majority falling within the good range (0.75–0.90) and the CCOC score exceeding the 0.90 threshold for excellent agreement. Together, these ICC values indicate that the scoring system for all three target proteins—FOLR1, TACSTD2, and VTCN1—shows strong inter-assessor reliability across different histotypes.



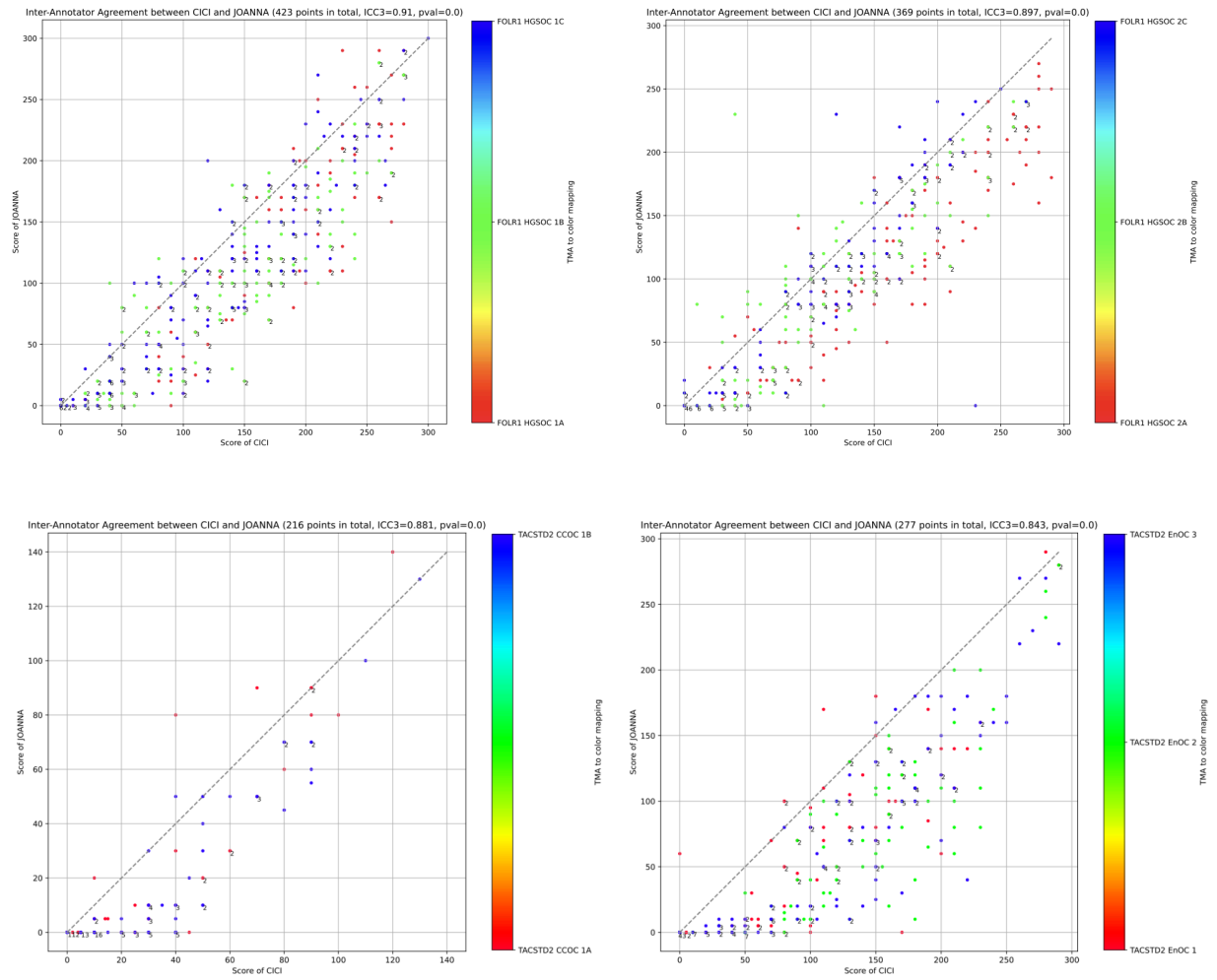


Figure 10a: Inter-assessor agreement between scorer 1 (Cici) and scorer 2 (Joanna) for FOLR1 and TACSTD2 expression evaluation in TMAs derived from patients with the CCOC, EnOC and HGSOc histotypes. FOLR1 CCOC: ICC3=0.947, p=0.0; FOLR1 EnOC: ICC3=0.961, p=0.0; FOLR1 HGSOc 1: ICC3=0.91, p=0.0; FOLR1 HGSOc 2: ICC3=0.897, p=0.0; TACSTD2 CCOC: ICC3=0.881, p=0.0; TACSTD2 EnOC: ICC3=0.843, p=0.0.

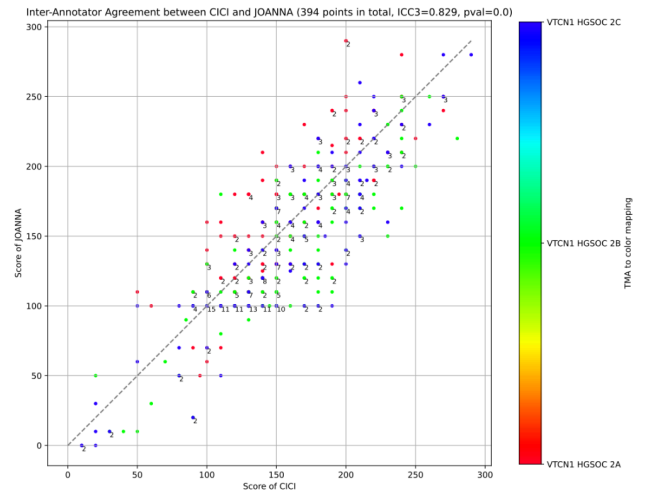
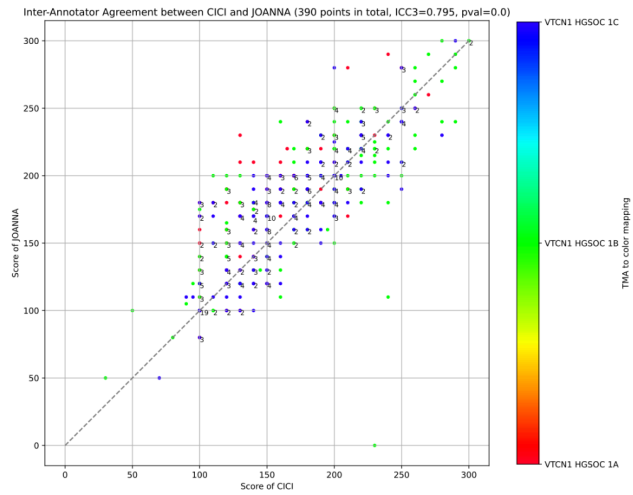
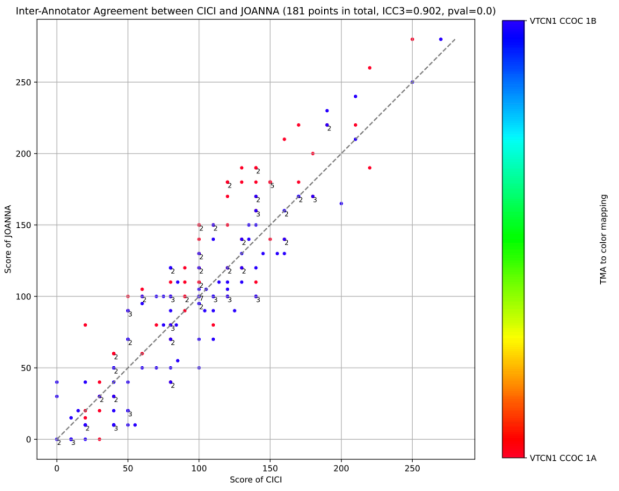
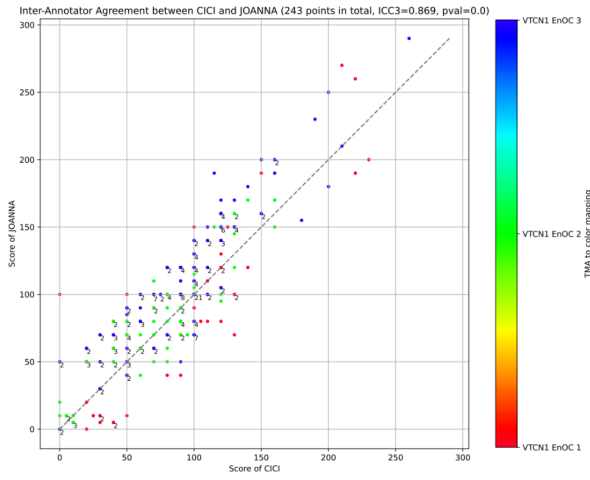
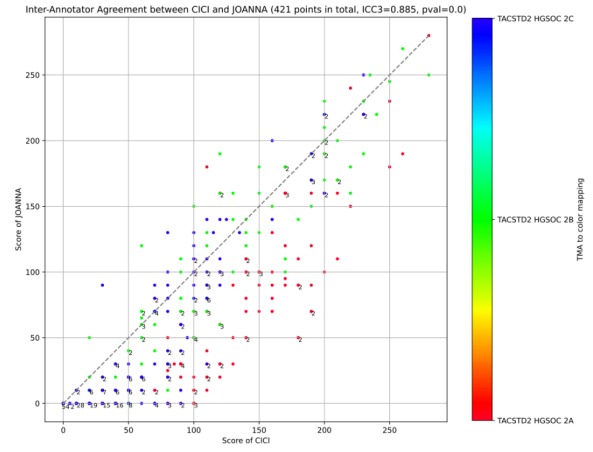
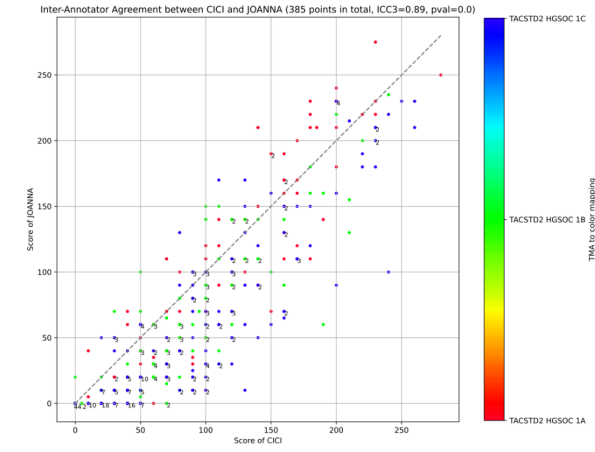


Figure 10b: Inter-assessor agreement between scorer 1 (Cici) and scorer 2 (Joanna) for TACSTD2 and VTCN1 expression evaluation in TMAs derived from patients with the CCOC, EnOC and HGSOC histotypes. TACSTD2 HGSOC 1: ICC3=0.89, p=0.0; TACSTD2 HGSOC 2: ICC3=0.885, p=0.0; VTCN1 CCOC: ICC3=0.902; p=0.0; VTCN1 EnOC: ICC3=0.869, p=0.0; VTCN1 HGSOC 1: ICC3=0.795, p=0.0; VTCN1 HGSOC 2: ICC3=0.829, p=0.0.

3.2 Mean target expressions

The mean scores for each target were calculated by averaging the scores between the two scorers, followed by averaging the scores across duplicates and triplicates within each histotype. The distribution of these scores is illustrated in Figures 11a, 11b, and 11c.

For FOLR1 expression, the distribution in clear cell ovarian carcinoma (CCOC) and endometrioid ovarian carcinoma (EnOC) skews towards the lower end, with the majority of scores (over 50%) falling within the 0–50 range. However, the remaining scores are spread across a broader spectrum, ranging from 50 to over 250. In contrast, the mean scores for FOLR1 in high-grade serous ovarian carcinoma (HGSOC) are notably higher, with most scores (around 60%) in HGSOC 1 ranging between 0 and 150, and in HGSOC 2, between 0 and 170. FOLR1 is significantly differentially expressed in HGSOC compared to EnOC and CCOC.

For TACSTD2 expression, the majority of scores in CCOC are quite low, with more than 50% clustering between 0 and 20. However, in EnOC, HGSOC 1, and HGSOC 2, TACSTD2 expression is considerably higher, with over half of the scores falling within the 0–150 range in EnOC, 0–100 in HGSOC 1, and 0–130 in HGSOC 2.

The expression of VTCN1 in CCOC is significantly elevated compared to the other two targets, with over 50% of the scores ranging between 50 and 120. Similarly, the mean VTCN1 scores in EnOC, HGSOC 1, and HGSOC 2 are generally higher than those of the other targets in these histotypes, with values spanning 50–120 in EnOC, 150–220 in HGSOC 1, and 100–200 in HGSOC 2.

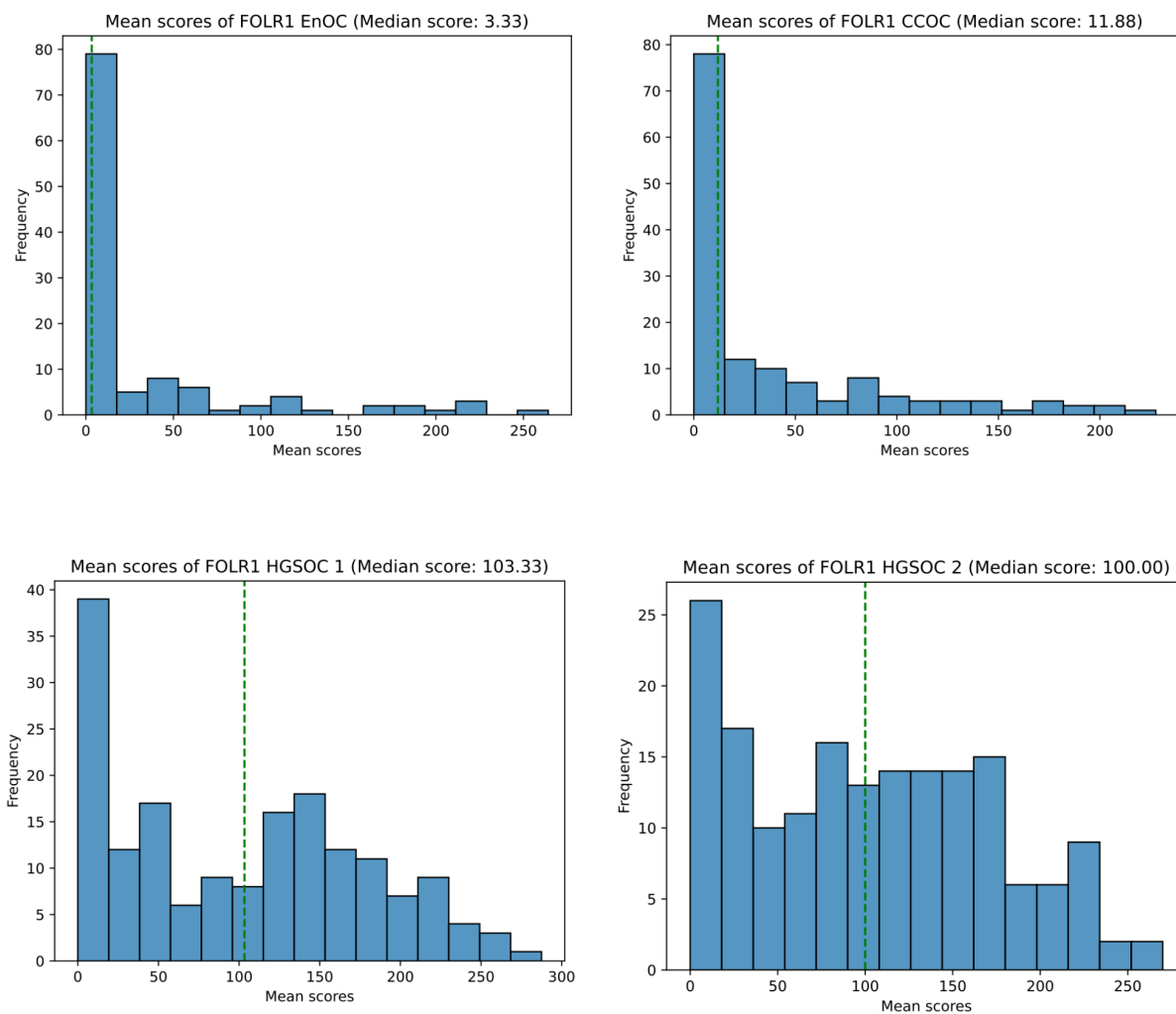


Figure 11a: Histograms of the expressions of each target in various ovarian cancer histotypes.

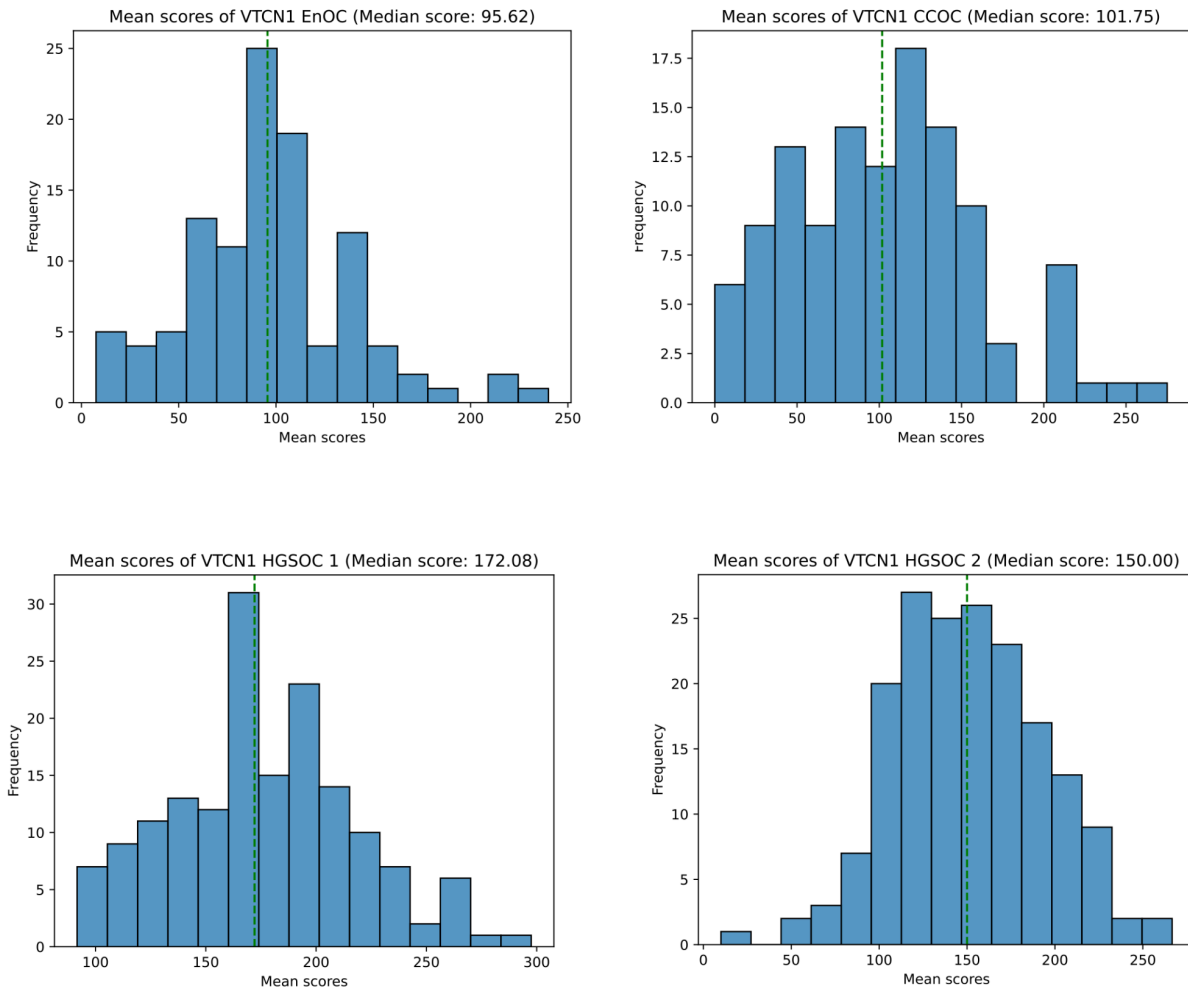


Figure 11b: Histograms of the expressions of each target in various ovarian cancer histotypes.

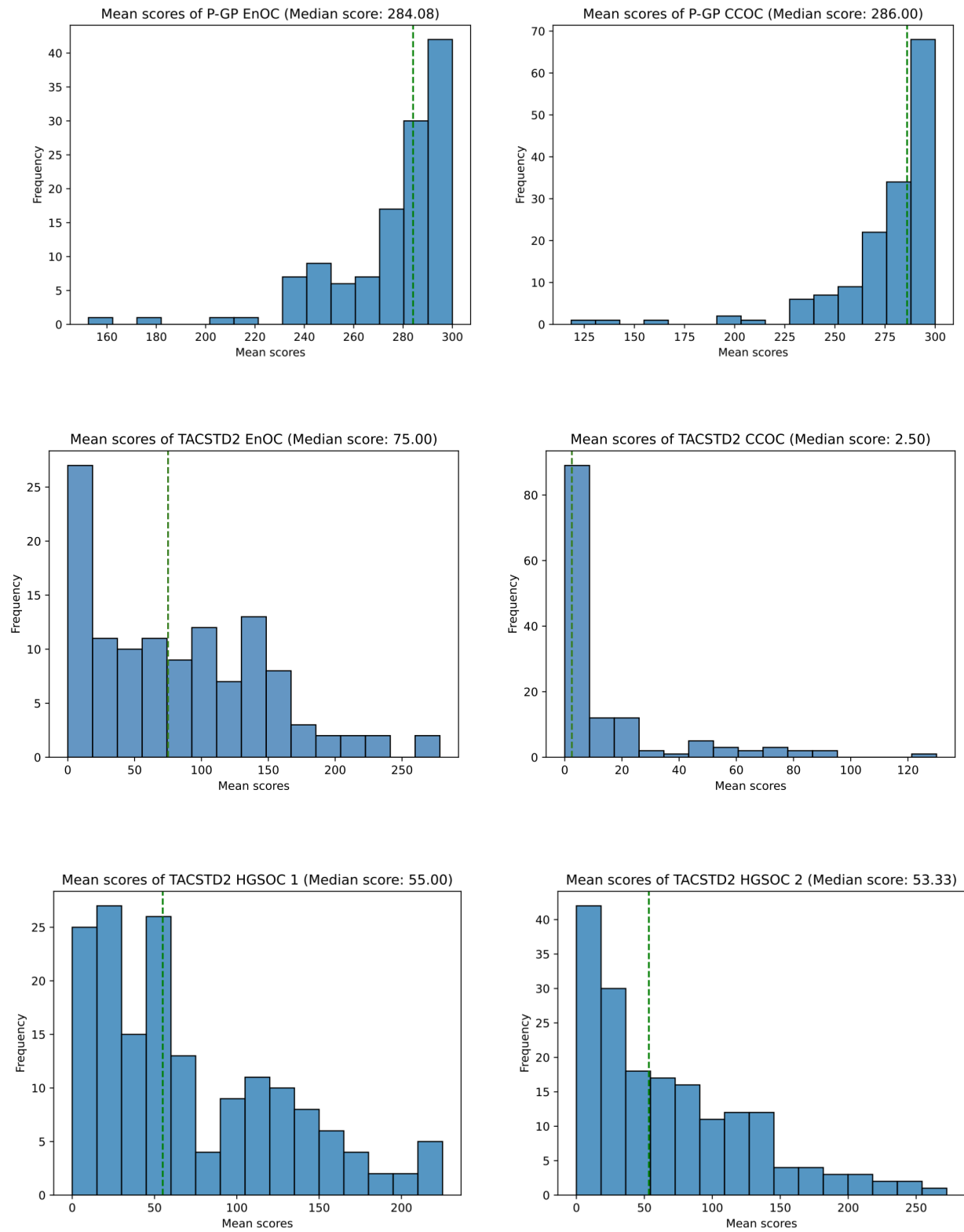


Figure 11c: Histograms of the expressions of each target in various ovarian cancer histotypes.

4 Discussion

The validity of using tissue microarrays (TMAs) for evaluating antigen expression in ovarian cancer has been well-supported by studies, including a notable 2009 investigation, which confirmed the utility of TMAs in assessing invasive ovarian cancer (88). However, the study cautioned against applying TMAs indiscriminately to borderline ovarian tumours, as these non-invasive, low-malignant potential tumours (89)—also referred to as atypical proliferative tumours—were more likely to exhibit missing stains in one or more TMA cores compared to invasive samples. Despite this limitation, the study demonstrated that using a single core specimen per tumour sample provided an accurate representation of the entire tumour, comparable to the assessment of three cores. Furthermore, inter-rater reliability, based on subjective scoring by two independent trained evaluators, was found to be acceptable, which is consistent with the methodology employed in our project. The study also addressed concerns regarding the potential influence of core age on staining outcomes, finding that although some cores were affected by age, the majority of cases across different histotypes were unaffected. Importantly, storage time did not appear to compromise antigen expression, reinforcing the validity of using archived samples in TMA studies.

The intraclass correlation coefficient (ICC) was employed in the 2009 study to assess the reproducibility of replicate measures from the same subject, thereby supporting the robustness of the TMA method. However, a 2018 study (90) examining the use of ICC for inter-rater reliability assessment highlighted the limitations of this approach. It noted that while ICC is commonly used to assess measurement scale reliability, its results can be skewed by statistical assumptions such as normality and stable variance, potentially inflating inter-rater reliability estimates. Therefore, a more comprehensive statistical analysis is needed to account for these assumptions while maintaining interpretability in future studies. In terms of next steps, the scoring of ADC (antibody-drug conjugate) target expressions may help predict patient responses and toxicity to ADC treatment. For instance, a Phase I trial of Mirvetuximab soravtansine-gynx (MIRV), an anti-FOLR1 ADC, assessed FOLR1 expression in both archival and fresh biopsy specimens, confirming that its expression remains stable even after multiple treatment lines and MIRV administration. Higher FOLR1 expression was associated with stronger antitumor responses (91,92). In our project, TMAs from most patients with high-grade serous ovarian cancer (HGSOC), and a smaller proportion of clear cell ovarian cancer (CCOC) and endometrioid ovarian cancer (EnOC) samples, demonstrated high FOLR1 expression. Future studies could involve stratifying these scores and comparing the response to MIRV between low and high expressers. Additionally, generating FOLR1

knock-in ovarian cancer models could further elucidate the role of FOLR1 expression in therapeutic response. A related study by Gray et al. (93) characterised VTCN1/B7-H4 expression across multiple tumour types using immunohistochemistry. They demonstrated that SGN-B7H4V, an investigational vedotin ADC, killed VTCN1/B7-H4-expressing tumour cells in vitro through MMAE-mediated cytotoxicity and antibody-mediated functions, including antibody-dependent cellular cytotoxicity and phagocytosis. In immunocompetent murine models expressing B7-H4, SGN-B7H4V showed robust antitumor activity as a monotherapy, which was further enhanced when combined with an anti-PD-1 agent. This study underscores the therapeutic potential of SGN-B7H4V in tumour subtypes with high VTCN1/B7-H4 expression. Given that our TMAs indicate relatively high VTCN1/B7-H4 expression in HGSOC, EnOC, and CCOC, our models could be valuable for studying the efficacy of SGN-B7H4V ADC therapy. Further studies could also explore whether a threshold value for VTCN1/B7-H4 expression exists, which could help predict therapy response.

Appendix 1

SOP#1 – Maintenance of mouse intestinal organoids

Introduction

- Culturing organoids takes place in 3D matrix called Matrigel/Basement membrane extract (BME).
- They are cultured in cell culture plates:
Small amounts needed: 1x 10-15µl matrix droplets/well; in 6 wells of a 24-well plate Greater amounts needed: ~10x 10-15µl matrix droplets/well; in 1 well of a 6-well plate Droplets are submerged in medium
- General maintenance includes replacing the medium every 2-3 days based on colour change and splitting/passaging of organoids based on growth rate, organoid density and amount of debris detected

Materials

- ADF base medium: keep at 4°C

Reagent	Amount	Company
Advanced DMEM/F12	500ml	Gibco, TFS
L-Glutamine	5ml	?
Pen/Strep	5ml	?
HEPES	5ml	Gibco, TFS

- EN medium (50 ml for AK mice): keep at 4°C

Reagent	Amount	Company
ADF	48ml	see above
B27 neuronal supplement	1ml	
N2 neuronal supplement	0.5ml	
Noggin BMP inhibitor	0.5ml	
EGF	50µl	

- ENR medium (50ml for WT mice): keep at 4°C

Reagent	Amount	Company
ADF	40.5ml	see above
B27 neuronal supplement	1ml	
N2 neuronal supplement	0.5ml	
Noggin BMP inhibitor	0.5ml	
EGF	50µl	
R-Spondin (activator of <u>Wnt signalling</u>)	7.5ml	

Method

1. Preparations:

1. Night before: incubate plates to be used @ 37°C
2. 1-1.5 h before: thaw BME on ice
3. ~30 mins before: warm up EN/ENR medium

2. On the day:

1. Scrape off BME droplets with stripette (6-well) or 1ml pipette tip (24-well)
2. Resuspend/break-up in well with the addition of 0.5-1ml extra ADF medium
3. Transfer to 15ml Falcon tube, resuspend and add up to 6-10ml ADF
4. Spin @ 300g @ 4°C for 5 mins
5. Aspirate supernatant down to ~2ml and resuspend with 1ml extra ADF
6. Add up to 6-10ml ADF then repeat steps 4-6 until all BME is removed
 1. (optional) wash in PBS or spin @ 70g for 3 mins
7. Remove all supernatant (aspirate then P200)
8. Add full volume of BME based on amount of drops intended, and resuspend taking care not to introduce any bubbles.
 1. This step should be done on ice as much as possible to avoid BME solidification
9. Transfer 10-15µl drops to wells on the preheated cell culture plates & incubated for 15- 30 mins based on drop sizes @ 37°C
10. Add EN/ENR medium to the wells in a volume of 0.5ml (24-well plate) or 1.5ml (6-well plate)
11. Maintain in incubator @ 37°C with 5% CO₂

Appendix 2



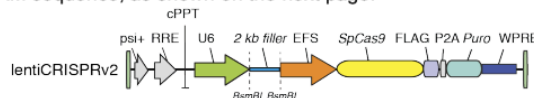
LENTIVIRAL CRISPR TOOLBOX



LentiCRISPRv2 and lentiGuide-Puro: lentiviral CRISPR/Cas9 and single guide RNA

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is a microbial nuclease system involved in defense against invading phages and plasmids. CRISPR loci in microbial hosts contain a combination of CRISPR-associated (Cas) genes as well as non-coding RNA elements capable of programming the specificity of the CRISPR-mediated nucleic acid cleavage. Lentiviral CRISPR/Cas can infect a broad variety of mammalian cells by co-expressing a mammalian codon-optimized Cas9 nuclease along with a single guide RNA (sgRNA) to facilitate genome editing (Shalem*, Sanjana*, et al., Science 2014). Protocols for cloning into the lentiviral transfer plasmid and general considerations for producing lentivirus are described below. **Separate protocols are available for amplifying the genome-scale CRISPR knock-out (GeCKO) libraries. This protocol is for creating individual lentiviral CRISPR plasmids targeting a single genomic locus.**

lentiCRISPRv2 (one vector system): This plasmid contains two expression cassettes, hSpCas9 and the chimeric guide RNA. The vector can be digested using *BsmBI*, and a pair of annealed oligos can be cloned into the single guide RNA scaffold. The oligos are designed based on the target site sequence (20bp) and needs to be flanked on the 3' end by a 3bp NGG PAM sequence, as shown on the next page.



lentiGuide-Puro (two vector system): This plasmid expressed only the chimeric guide RNA. It does *not* contain Cas9. Please use lentiCas9-Blast (a separate lentiviral construct that delivers hSpCas9 and blasticidin resistance) to first integrate Cas9 into your cell line. The lentiGuide-Puro vector can be digested using *BsmBI*, and a pair of annealed oligos can be cloned into the single guide RNA scaffold. The oligos are designed based on the target site sequence (20bp) and needs to be flanked on the 3' end by a 3bp NGG PAM sequence, as shown on the next page.



Which vector to use: lentiCRISPRv2 is identical to the original lentiCRISPRv1 but produces nearly 10X higher titer virus. lentiGuide-Puro produces >100X higher titer virus over lentiCRISPRv1 and should be used in cell lines where Cas9 has already been integrated (e.g. using the separate lentiCas9-Blast lentivirus). For applications where Cas9 cannot first be introduced (e.g. primary cells), lentiCRISPRv2 is recommended. After transduction, use puromycin to select for cells with lentiCRISPRv2 or lentiGuide-Puro.

Lentiviral production: Before starting any lentiviral work, please ensure compliance with your Environmental Health and Safety office and government/organization/university. Briefly, to make lentivirus, a transfer plasmid (e.g. lentiCRISPRv2 or lentiGuide-Puro) must be co-transfected into HEK293(F)T cells with the packaging plasmids pVSVg (AddGene 8454) and psPAX2 (AddGene 12260). As a positive control for viral production, we often use a CMV-EGFP lentiviral transfer plasmid (eg. AddGene 19319).

Target design notes and online resources: For application of Cas9 for site-specific genome editing in eukaryotic cells and organisms, we have computationally identified suitable target sites for the *S. pyogenes* Cas9 and calculated most likely off-targets within the genome. Please visit <http://www.genome-engineering.org> to access these Cas9 target design tools. Complete plasmid sequences, protocols, a discussion forum and additional information can be found at the Zhang Lab GeCKO website: <http://www.genome-engineering.org/gecko/>.

Citation: Please reference the following publications for the use of this material.

Improved lentiviral vectors and genome-wide libraries for CRISPR screening. Sanjana NE*, Shalem O*, Zhang F. *Nature Methods* (2014).

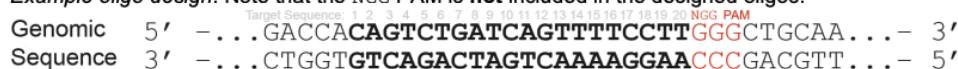
Genome-scale CRISPR-Cas9 knockout screening in human cells. Shalem O*, Sanjana NE*, Hartenian E, Shi X, Scott DA, Mikkelsen T, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F (2014). *Science*, 343, 83-7. DOI: 10.1126/science.1247005

Target Guide Sequence Cloning Protocol

In order to clone the target sequence into the lentiCRISPRv2 or lentiGuide-Puro backbone, synthesize two oligos of the following form. **All plasmids have the same overhangs** after *BsmBI* digestion and the same oligos can be used for cloning into lentiCRISPRv2, lentiGuide-Puro or lentiCRISPRv1.



Example oligo design: Note that the NGG PAM is **not** included in the designed oligos.



Oligonucleotide ordering tips: Standard de-salted oligos (usually the most inexpensive synthesis) are sufficient for cloning. If not already resuspended, dilute each oligo to 100 μ M in sterile water or TE.

* * * * *

Lentiviral vector digestion, oligo annealing and cloning into digested vector:

1. Digest and dephosphorylate 5ug of the lentiviral CRISPR plasmid with *BsmBI* for 30 min at 37C:

5 ug	lentiCRISPRv2 or lentiGuide-Puro
3 ul	FastDigest <i>BsmBI</i> (Fermentas)
3 ul	FastAP (Fermentas)
6 ul	10X FastDigest Buffer
0.6 ul	100 mM DTT (freshly prepared)
X ul	ddH ₂ O
60 ul	total

2. Gel purify digested plasmid using QIAquick Gel Extraction Kit and elute in EB.

If *BsmBI* digested, a ~2kb filler piece should be present on the gel. **Only gel purify the larger band.** Leave the 2kb band.

3. Phosphorylate and anneal each pair of oligos:

1 ul	Oligo 1 (100 μ M)
1 ul	Oligo 2 (100 μ M)
1 ul	10X T4 Ligation Buffer (NEB)
6.5 ul	ddH ₂ O
0.5 ul	T4 PNK (NEB M0201S)
10 ul	total

Please use the T4 **Ligation** Buffer since the buffer supplied with the T4 PNK enzyme does not include ATP (or supplement to 1mM ATP).

Put the phosphorylation/annealing reaction in a thermocycler using the following parameters:

37°C	30 min
95°C	5 min and then ramp down to 25°C at 5°C/min

4. Dilute annealed oligos from **Step 3** at a 1:200 dilution into sterile water or EB.

5. Set up ligation reaction and incubate at room temperature for 10 min:

X ul	<i>BsmBI</i> digested plasmid from Step 2 (50ng)
1 ul	diluted oligo duplex from Step 4
5 ul	2X Quick Ligation Buffer (NEB)
X ul	ddH ₂ O
10 ul	subtotal
1 ul	Quick Ligase (NEB M2200S)
11 ul	total

Also perform a negative control ligation (vector-only with water in place of oligos) and transformation.

6. Transformation into **Stbl3** bacteria. Lentiviral transfer plasmids contain Long-Terminal Repeats (LTRs) and must be transformed into recombination-deficient bacteria. We use homemade Stbl3 (propagated from Invitrogen C7373-03) and get excellent plasmid yields. Although other RecA- strains may work, we have found the most consistent transformations and yields using Stbl3.

July 2018

Quick-Start Protocol

QIAquick® Gel Extraction Kit QIAquick® PCR & Gel Cleanup Kit

The QIAquick Gel Extraction Kit and the QIAquick PCR & Gel Cleanup Kit (cat. nos. 28704, 28706, 28506 and 28115) can be stored at room temperature (15–25°C) for up to 12 months.

Further information

- *QIAquick Spin Handbook*: www.qiagen.com/HB-1196
- Safety Data Sheets: www.qiagen.com/safety
- Technical assistance: support.qiagen.com

Notes before starting

- This protocol is for the purification of up to 10 µg DNA (70 bp to 10 kb).
- The yellow color of Buffer QG indicates a pH ≤7.5. DNA adsorption to the membrane is only efficient at pH ≤7.5.
- Add ethanol (96–100%) to Buffer PE before use (see bottle label for volume).
- Isopropanol (100%) and a heating block or water bath at 50°C are required.
- All centrifugation steps are carried out at 17,900 x g (13,000 rpm) in a conventional table-top microcentrifuge.

-
1. Excise the DNA fragment from the agarose gel with a clean, sharp scalpel.
 2. Weigh the gel slice in a colorless tube. Add 3 volumes Buffer QG to 1 volume gel (100 mg gel ~100 μ l). The maximum amount of gel per spin column is 400 mg. For >2% agarose gels, add 6 volumes Buffer QG.
 3. Incubate at 50°C for 10 min (or until the gel slice has completely dissolved). Vortex the tube every 2–3 min to help dissolve gel. After the gel slice has dissolved completely, check that the color of the mixture is yellow (similar to Buffer QG without dissolved agarose). If the color of the mixture is orange or violet, add 10 μ l 3 M sodium acetate, pH 5.0, and mix. The mixture turns yellow.
 4. Add 1 gel volume isopropanol to the sample and mix.
 5. Place a QIAquick spin column in a provided 2 ml collection tube or into a vacuum manifold. To bind DNA, apply the sample to the QIAquick column and centrifuge for 1 min or apply vacuum to the manifold until all the samples have passed through the column. Discard flow-through and place the QIAquick column back into the same tube. For sample volumes >800 μ l, load and spin/apply vacuum again.
 6. If DNA will subsequently be used for sequencing, in vitro transcription, or microinjection, add 500 μ l Buffer QG to the QIAquick column and centrifuge for 1 min or apply vacuum. Discard flow-through and place the QIAquick column back into the same tube.
 7. To wash, add 750 μ l Buffer PE to QIAquick column and centrifuge for 1 min or apply vacuum. Discard flow-through and place the QIAquick column back into the same tube.
Note: If the DNA will be used for salt-sensitive applications (e.g., sequencing, blunt-ended ligation), let the column stand 2–5 min after addition of Buffer PE.
Centrifuge the QIAquick column in the provided 2 ml collection tube for 1 min to remove residual wash buffer.
 8. Place QIAquick column into a clean 1.5 ml microcentrifuge tube.
 9. To elute DNA, add 50 μ l Buffer EB (10 mM Tris-Cl, pH 8.5) or water to the center of the QIAquick membrane and centrifuge the column for 1 min. For increased DNA concentration, add 30 μ l Buffer EB to the center of the QIAquick membrane, let the
-

column stand for 1 min, and then centrifuge for 1 min. After the addition of Buffer EB to the QIAquick membrane, increasing the incubation time to up to 4 min can increase the yield of purified DNA.

10. If purified DNA is to be analyzed on a gel, add 1 volume of Loading Dye to 5 volumes of purified DNA. Mix the solution by pipetting up and down before loading the gel.

Revision History

Revision no.	Description of change
R3 07/2018	Updated document title and introductory paragraph with additional applicable product QIAquick PCR & Gel Cleanup Kit. Also added additional product's related cat. nos.



Scan QR code for handbook.

For up-to-date licensing information and product-specific disclaimers, see the respective QIAGEN kit handbook or user manual.

Trademarks: QIAGEN®, Sample to Insight®, QIAquick® (QIAGEN Group). 1114358 07/2018 HB-0901-003 © 2018 QIAGEN, all rights reserved.

Ordering www.qiagen.com/contact | Technical Support support.qiagen.com | Website www.qiagen.com



PRODUCT INFORMATION SHEET

One Shot™ Stbl3™ Chemically Competent *E. coli*

Catalog Number C7373-03 (20 reactions)

Doc. Part No. C737303.pps Pub. No. MAN0001497 Rev. A.0



WARNING! This product contains irritants and may be harmful if swallowed. Review the Material Safety Data Sheet before handling. Safety Data Sheets (SDSs) are available from [thermofisher.com/support](https://www.thermofisher.com/support).

Product description

The Stbl3™ *E. coli* strain is derived from the HB101 *E. coli* strain and is recommended for use when cloning unstable inserts such as lentiviral DNA containing direct repeats (e.g. ViraPower™ Lentiviral Expression Kits). The transformation efficiency of Invitrogen™ One Shot™ Stbl3™ chemically competent cells is greater than 1×10^8 cfu/μg DNA.

Genotype

F⁻ *mcrB mrr hsdS20(r_B⁻, m_B⁻) recA13 supE44 ara-14 galK2 lacY1 proA2 rpsL20(Str^R) xyl-5 λ⁻ leu mtl-1*

Note: This strain is *end* A1+.

Contents and storage

Contents	Quantity	Storage conditions
Stbl3™ Cells	21 × 50 μL	Store at -80°C
pUC19 Control DNA (10 pg/μL)	50 μL	
S.O.C. Medium	6 mL	

For Research Use Only. Not for use in diagnostic procedures.

ThermoFisher
SCIENTIFIC

Important guidelines

Handle competent cells gently as they are highly sensitive to changes in temperature or mechanical lysis caused by pipetting. Thaw One Shot™ competent cells on ice, and transform cells immediately following thawing. After adding DNA, mix by swirling or tapping the tube gently. *Do not mix cells by pipetting.*

Note: Cells cannot be used for blue/white screening of plasmid inserts.

Before you begin

Perform the following before starting the transformation procedure:

1. Equilibrate a water bath to 42°C.
2. Warm the vial of S.O.C. Medium (supplied with the kit) and LB Medium (if needed) to room temperature.
3. Warm the selective plates in a 37°C incubator for 30 minutes (use 1 or 2 plates for each transformation). If you are including the pUC19 control, make sure that you have one LB agar plate containing 100 µg/mL ampicillin.

Methods

Transform One Shot™ Stbl3™ chemically competent *E. coli*

We recommend including the pUC19 control plasmid DNA supplied with the kit (10 pg/μL in 5 mM Tris-HCl, 0.5 mM EDTA, pH 8) in your transformation experiment to verify the efficiency of the competent cells. Do not use these cells for electroporation.

1. Thaw, on ice, one vial of One Shot™ Stbl3™ chemically competent cells for each transformation.
2. Add 1 to 5 μL of the DNA (10 pg to 100 ng) into a vial of One Shot™ cells and mix gently. *Do not mix by pipetting up and down.* For the pUC19 control, add 10 pg (1 μL) of DNA into a separate vial of One Shot™ cells and mix gently.
3. Incubate the vial(s) on ice for 30 minutes.
4. Heat-shock the cells for 45 seconds at 42°C without shaking.
5. Remove the vial(s) from the 42°C bath and place them on ice for 2 minutes.
6. Add 250 μL of pre-warmed S.O.C. Medium to each vial.
7. Cap the vial(s) tightly and shake horizontally at 37°C for 1 hour at 225 rpm in a shaking incubator.
8. Spread 25–100 μL from each transformation on a pre-warmed selective plate and incubate overnight at 37°C. We recommend that you plate two different volumes to ensure that at least one plate will have well-spaced colonies. For the pUC19 control, dilute the transformation mix 1:10 into LB Medium (*e.g.* remove 100 μL of the transformation mix and add to 900 μL of LB Medium) and plate 25–100 μL.
9. Store the remaining transformation mix at 4°C. Additional cells may be plated out the next day, if desired.
10. Invert the selective plate(s) and incubate at 37°C overnight.
11. Select colonies and analyze by plasmid isolation, PCR, or sequencing.

Quick-Start Protocol

March 2016

QIAGEN® Plasmid Mini, Midi and Maxi Kits

The QIAGEN Plasmid Mini Kit (cat. nos. 12123 and 12125), the QIAGEN Plasmid Midi Kit (cat. nos. 12143 and 12145), the QIAGEN Plasmid Maxi Kit (cat. nos. 12162, 12163 and 12165) and the Plasmid Buffer Set (cat. no. 19046) can be stored at room temperature (15–25°C) for up to 2 years if not otherwise stated on label.

Further information

- QIAGEN Plasmid Purification Handbook: www.qiagen.com/HB-1193
- Safety Data Sheets: www.qiagen.com/safety
- Technical assistance: support.qiagen.com

Notes before starting

- Add RNase A solution to Buffer P1, mix and store at 2–8°C.
- **Optional:** Add LyseBlue® reagent to Buffer P1 at a ratio of 1:1000.
- Prechill Buffer P3 at 4°C. Check Buffer P2 for SDS precipitation.
- Isopropanol and 70% ethanol are required.
- Symbols: ● QIAGEN Plasmid Mini Kit; ▲ QIAGEN Plasmid Midi Kit; and ■ QIAGEN Plasmid Maxi Kit.

Table 1. Recommended LB culture volumes

Kit	High-copy plasmid	Low-copy plasmid
QIAGEN Plasmid Mini	3 ml	Not recommended
QIAGEN Plasmid Midi	25 ml	100 ml
QIAGEN Plasmid Maxi	100 ml	500 ml

1. Harvest overnight bacterial culture by centrifuging at 6000 x g for 15 min at 4°C.
2. Resuspend the bacterial pellet in ● 0.3 ml, ▲ 4 ml or ■ 10 ml Buffer P1.

— Sample to Insight —



-
3. Add ● 0.3 ml, ▲ 4 ml or ■ 10 ml Buffer P2, mix thoroughly by vigorously inverting 4–6 times and incubate at room temperature (15–25°C) for 5 min. If using LyseBlue reagent, the solution will turn blue.
 4. Add ● 0.3 ml, ▲ 4 ml or ■ 10 ml prechilled Buffer P3, mix thoroughly by vigorously inverting 4–6 times. Incubate on ice for ● 5 min, ▲ 15 min or ■ 20 min. If using LyseBlue reagent, mix the solution until it is colorless.
 5. ●: Centrifuge at 14,000–18,000 x g for 10 min at 4°C. Re-centrifuge if the supernatant is not clear. ▲ and ■: Centrifuge at ≥20,000 x g for 30 min at 4°C. Re-centrifuge the supernatant at ≥20,000 x g for 15 min at 4°C.
 6. Equilibrate a QIAGEN-tip ● 20, ▲ 100 or ■ 500 by applying ● 1 ml, ▲ 4 ml or ■ 10 ml Buffer QBT, and allow column to empty by gravity flow.
 7. Apply the supernatant from step 5 to the QIAGEN-tip and allow it to enter the resin by gravity flow.
 8. Wash the QIAGEN-tip with ● 2 x 2 ml, ▲ 2 x 10 ml or ■ 2 x 30 ml Buffer QC. Allow Buffer QC to move through the QIAGEN-tip by gravity flow.
 9. Elute DNA with ● 0.8 ml, ▲ 5 ml or ■ 15 ml Buffer QF into a clean ● 2 ml, ▲ 15 ml or ■ 50 ml vessel. For constructs larger than 45 kb, prewarming the elution buffer to 65°C may help to increase the yield.
 10. Precipitate DNA by adding ● 0.56 ml, ▲ 3.5 ml or ■ 10.5 ml (0.7 volumes) room-temperature isopropanol to the eluted DNA and mix. Centrifuge at ≥15,000 x g for 30 min at 4°C. Carefully decant the supernatant.
 11. Wash the DNA pellet with ● 1 ml, ▲ 2 ml or ■ 5 ml room-temperature 70% ethanol and centrifuge at ≥15,000 x g for 10 min. Carefully decant supernatant.
 12. Air-dry pellet for 5–10 min and redissolve DNA in a suitable volume of appropriate buffer (e.g., TE buffer, pH 8.0, or 10 mM Tris-Cl, pH 8.5).



Scan QR code for handbook.

For up-to-date licensing information and product-specific disclaimers, see the respective QIAGEN kit handbook or user manual. Trademarks: QIAGEN®, Sample to Insight®, LyseBlue® (QIAGEN Group). 1101291 03/2016 HB-0586-002 © 2016 QIAGEN, all rights reserved.

Appendix 6

SOP

Western Blot

Protein extraction;

Without trypsin method:

- Remove the medium from the cells
- Wash the cells with 1 ml cold PBS
- Aspirate the PBS
- Add 500 cold PBS per well
- Scrape the cells hardly and transfer the cells to labelled 2ml eppendorfs
- Centrifuge the cells at 2000rpm for 5 min at RT
- Remove supernatant
- Place the cells on ice and add 30-50ul RIPA buffer (depends on your pellet volume) (generally 30-75ul RIPA) (containing Phosphatase and Proteinase inhibitor)
- Pipette up-down the mix
- Incubate on ice for 30 min
- Centrifuge at 14,000rpm for 10 min at 4°C.
- Transfer the lysates to fresh and labelled eppendorfs
- Store the samples at -20°C until BCA assay for quantification

**** There is another method which is 'with trypsin':**

- Remove the medium from the cells
- Wash the cells with 1 ml cold PBS
- Aspirate the PBS
- Add 1ml PBS+ trypsin per well
- Transfer the cell suspension to eppendorfs, continue the rest of the protocol the same as above (from 5th step)

**** Preparation of RIPA buffer:** 1 ml RIPA buffer + 10ul Proteinase liquid inhibitor + 10ul Phosphatase liquid inhibitor

After protein extraction;

1. Determine the protein concentration with BCA Assay

- Put 10ul of protein and each standard in 96-well plate – flat well bottom
- Prepare working solution = 1:50 v/v
- Share 200ul of working solution per well
- Incubate the plate at 37°C for 30 min incubator
- Measure the protein concentration at 562nm
- Calculate in the Excel, 15ug/well
- ** Calculate according to 60ul as total volume (= 15ul per well * 4 (3 +1 replicates+extra))**

- Store the remaining of protein at -80°C (thaw protein samples each time you need)

2. Preparation of western blot samples and running the gel

- Make a loading buffer in the hood
= 90ul LDX 4X Sample Buffer + 10ul Sample Reducing Agent 10X
- Label eppendorf tubes
- Mix amount of the samples+ protein loading buffer+ H₂O in eppendorf tubes according to calculation in excel sheet
- Vortex and spin mixture samples
- Denaturate the samples in Thermomixer at 99°C for 5min
- Transfer the tubes on ice
- Prepare 1L Running buffer: 950ml dH₂O + 50ml NuPAGE MOPS SDS Running buffer x20
- Get NuPAGE gel stacked gradient 4-20% - take off white sticker and comb from top
- Insert the gel to the cassette and into the tank
- Insert white apparatus, hold and secure the cassette
- Fill the tank with 1X Running buffer
- Flush out the wells to remove the bubbles inside
- Load 15ul of each samples as well as 10ul (7ul also enough) Precision Plus Protein dual Colour Loading Marker(Ab box)
- Put the lid on
- Run the gel at 150V for 1.30h

3. Blotting

- Prepare 2L of 20X Transfer buffer=
100ml of NuPAGE Transfer Buffer + 400ml Methanol + 1500ml dH₂O
- Cut out 9 by 7 cm (or sponge size) of Whatman paper and Nitrocellulose membrane (Amersham proton membrane)
- Let the materials and the gel wet in transfer buffer in a transfer box
- While preparing the sandwich, do not forget to cut the corner of the membrane
- Do not forget the roll with a falcon tube to make sure removing the bubbles
- Prepare the sandwich models as below;
 - + Transparent Lid +
 - Sponge
 - 2X Whatman Paper
 - Membrane
 - Gel
 - 2X Whatman Paper
 - Sponge
 - - Black hole bottom of cassette-
- Close, push and squeeze the sandwich
- All submerged in transfer buffer flat glass container
- Black part goes at the back

- Carry to cold room green powerpack plug in lid (or put the container into ice bucket, make sure transfer solution passes the line of 4gels)
- Run at 250mA current amps for 1.30h
- Cut right-top corner of the membrane and Rinse the membrane with dH₂O 2 times
- Swirl the membrane with ponceau dye for 1min – pour the ponceau back into its container
- Wash the membrane with dH₂O a few times
- Transfer the membrane to plate containing 0.1% PBST and shake for 15min

** After this step, You can store the membrane in fresh tip box filled cold fresh PBS halfway for a few days

4. **Blocking**

- Block the membrane with 5% milk in 0.1%PBST at RT for 1h on slow shaker
- Cut the membrane from specific marker line and make sure each membrane has little cut on right-top corner

5. **Antibody incubation**

Day1:

- Prepare mix 20ml of blocking solution (5% milk in 0.1%PBST) with an appropriate dilution of primary Ab (can freeze: label name, date, concentration, Ab, how many times used at -20°C) in a 50ml falcon
(*NSUN2 Rabbit primary ab= 1:5000 ; Beta-actin primary ab= 1:5000 dilution)
- Transfer each primary antibody to corresponding labelled box
- Incubate the membrane with Primary Ab overnight on roller in cold room

Day2:

- Wash the membrane with PBST 3X 10min each
- Prepare mix 20ml of blocking solution (5% milk in 0.1%PBST) with Red Cap HRP conjugated anti-Rabbit secondary Ab (1/5000) at RT in box for 1h (*you can use ~~secAb~~ 1:5000-10000)
- Wash the membrane with PBST 3X 10min each
- You can keep the membranes in 1X PBS in cold room until you visualise
- ECL+ chemiluminescence preparation: 975ul solution A + 25ul solution B (for 2 membranes)
- ECL normal 1:1 v/v
- Visualise in ~~ImageQuant~~ LAS 800

References

1. SEER [Internet]. [cited 2024 Sep 3]. Common Cancer Sites - Cancer Stat Facts. Available from: <https://seer.cancer.gov/statfacts/html/common.html>
2. Roberts CWM, Orkin SH. The SWI/SNF complex — chromatin and cancer. *Nat Rev Cancer*. 2004 Feb;4(2):133–42.
3. Tokunaga R, Xiu J, Goldberg RM, Philip PA, Seeber A, Battaglin F, et al. Gene mutations of SWI/SNF complex and molecular profile in colorectal cancer. *J Clin Oncol*. 2019 May 20;37(15_suppl):3600–3600.
4. Colorectal cancer [Internet]. [cited 2024 Aug 21]. Available from: <https://www.who.int/news-room/fact-sheets/detail/colorectal-cancer>
5. Cancer Research UK [Internet]. 2015 [cited 2024 Aug 21]. Cancer mortality for common cancers. Available from: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/mortality/common-cancers-compared>
6. Cancer incidence for common cancers | Cancer Research UK [Internet]. [cited 2024 Aug 21]. Available from: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/common-cancers-compared>
7. DeVita VT, Lawrence TS, Rosenberg SA. DeVita, Hellman, and Rosenberg's Cancer: Principles & Practice of Oncology. Lippincott Williams & Wilkins; 2008. 1748 p.
8. Simon K. Colorectal cancer development and advances in screening. *Clin Interv Aging*. 2016 Jul 19;11:967–76.
9. Ciepiela I, Szczepaniak M, Ciepiela P, Hińcza-Nowak K, Kopczyński J, Macek P, et al. Tumor location matters, next generation sequencing mutation profiling of left-sided, rectal, and right-sided colorectal tumors in 552 patients. *Sci Rep*. 2024 Feb 26;14(1):4619.
10. Lee CS, Song IH, Lee A, Kang J, Lee YS, Lee IK, et al. Enhancing the landscape of colorectal cancer using targeted deep sequencing. *Sci Rep*. 2021 Apr 14;11(1):8154.
11. Berner DA. GeNotes. [cited 2024 Aug 22]. Colorectal cancer — Knowledge Hub. Available from: <https://www.genomicseducation.hee.nhs.uk/genotes/knowledge-hub/colorectal-cancer/>
12. Dinu D, Dobre M, Panaitescu E, Bîrlă R, Iosif C, Hoara P, et al. Prognostic significance of KRAS gene mutations in colorectal cancer - preliminary study. *J Med Life*. 2014;7(4):581–7.
13. BRAF Mutation and Cancer [Internet]. 2023 [cited 2024 Aug 22]. Available from: <https://www.hopkinsmedicine.org/health/conditions-and-diseases/braf-mutation-and-cancer>
14. Cao Q, Tian Y, Deng Z, Yang F, Chen E. Epigenetic Alteration in Colorectal Cancer:

Potential Diagnostic and Prognostic Implications. *Int J Mol Sci.* 2024 Jan;25(6):3358.

15. Wainwright EN, Scaffidi P. Epigenetics and Cancer Stem Cells: Unleashing, Hijacking, and Restricting Cellular Plasticity. *Trends Cancer.* 2017 May 1;3(5):372–86.
16. Epigenetics as a mediator of plasticity in cancer - PMC [Internet]. [cited 2024 Aug 22]. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10249049/>
17. Shen S, Clairambault J. Cell plasticity in cancer cell populations. *F1000Research.* 2020 Jun 22;9:F1000 Faculty Rev-635.
18. Kumar VE, Nambiar R, De Souza C, Nguyen A, Chien J, Lam KS. Targeting Epigenetic Modifiers of Tumor Plasticity and Cancer Stem Cell Behavior. *Cells.* 2022 Apr 21;11(9):1403.
19. Gupta PB, Pastushenko I, Skibinski A, Blanpain C, Kuperwasser C. Phenotypic Plasticity: Driver of Cancer Initiation, Progression, and Therapy Resistance. *Cell Stem Cell.* 2019 Jan 3;24(1):65–78.
20. Chromatin Remodeling in Eukaryotes | Learn Science at Scitable [Internet]. [cited 2024 Aug 22]. Available from: <http://www.nature.com/scitable/topicpage/chromatin-remodeling-in-eukaryotes-1082>
21. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. *Science.* 2017 Jul 21;357(6348):eaal2380.
22. Thrash HL, Pendergast AM. Multi-Functional Regulation by YAP/TAZ Signaling Networks in Tumor Progression and Metastasis. *Cancers.* 2023 Sep 24;15(19):4701.
23. Chang L, Azzolin L, Di Biagio D, Zanconato F, Battilana G, Lucon Xiccato R, et al. The SWI/SNF complex is a mechanoregulated inhibitor of YAP and TAZ. *Nature.* 2018 Nov;563(7730):265–9.
24. The SWI/SNF complex — chromatin and cancer | Nature Reviews Cancer [Internet]. [cited 2024 Aug 22]. Available from: <https://www.nature.com/articles/nrc1273>
25. Mittal P, Roberts CWM. The SWI/SNF complex in cancer — biology, biomarkers and therapy. *Nat Rev Clin Oncol.* 2020 Jul;17(7):435–48.
26. Li K, Wang B, Hu H. Research progress of SWI/SNF complex in breast cancer. *Epigenetics Chromatin.* 2024 Feb 17;17(1):4.
27. Helming KC, Wang X, Roberts CWM. Vulnerabilities of mutant SWI/SNF complexes in cancer. *Cancer Cell.* 2014 Sep 9;26(3):309.
28. Orlando KA, Nguyen V, Raab JR, Walhart T, Weissman BE. Remodeling the cancer epigenome: mutations in the SWI/SNF complex offer new therapeutic opportunities. *Expert Rev Anticancer Ther.* 2019 May;19(5):375–91.
29. Xu S, Tang C. The Role of ARID1A in Tumors: Tumor Initiation or Tumor Suppression? *Front Oncol.* 2021 Oct 4;11:745187.

30. Versteeg I, Sévenet N, Lange J, Rousseau-Merck MF, Ambros P, Handgretinger R, et al. Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature*. 1998 Jul;394(6689):203–6.
31. Jn W, Cw R. ARID1A mutations in cancer: another epigenetic tumor suppressor? *Cancer Discov* [Internet]. 2013 Jan [cited 2024 Sep 3];3(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/23208470/>
32. Bg W, Cw R. SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer* [Internet]. 2011 Sep 6 [cited 2024 Sep 3];11(7). Available from: <https://pubmed.ncbi.nlm.nih.gov/21654818/>
33. Wang X, Lee RS, Alver BH, Haswell JR, Wang S, Mieczkowski J, et al. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat Genet*. 2017 Feb;49(2):289–95.
34. Kohashi K, Oda Y. Oncogenic roles of SMARCB1/INI1 and its deficient tumors. *Cancer Sci*. 2017 Apr;108(4):547–52.
35. Xia W, Nagase S, Montia AG, Kalachikov SM, Keniry M, Su T, et al. BAF180 is a critical regulator of p21 induction and a tumor suppressor mutated in breast cancer. *Cancer Res*. 2008 Mar 15;68(6):1667–74.
36. Pan D, Kobayashi A, Jiang P, Ferrari de Andrade L, Tay RE, Luoma AM, et al. A major chromatin regulator determines resistance of tumor cells to T cell-mediated killing. *Science*. 2018 Feb 16;359(6377):770–5.
37. cBioPortal for Cancer Genomics: ARID1A, SMARCA4 in Colorectal Cancer (MSK, JNCI 2021) [Internet]. [cited 2024 Sep 3]. Available from: https://www.cbioportal.org/results/oncoprint?cancer_study_list=crc_eo_2020&Z_SCORE_THRESHOLD=2.0&RPPA_SCORE_THRESHOLD=2.0&profileFilter=mutations%2Cstructural_variants%2Ccna&case_set_id=crc_eo_2020_cnaseq&gene_list=ARID1A%252C%2520SMARCA4&geneset_list=%20&tab_index=tab_visualize&Action=Submit
38. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen T, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*. 2014 Jan 3;343(6166):84–7.
39. LLC GB. SnapGene | Software for everyday molecular biology [Internet]. [cited 2024 Sep 7]. Available from: <https://www.snapgene.com/>
40. UCSC Genome Browser Home [Internet]. [cited 2024 Sep 7]. Available from: <https://genome.ucsc.edu/>
41. Zhao S, Wu W, Jiang Z, Tang F, Ding L, Xu W, et al. Roles of ARID1A variations in colorectal cancer: a collaborative review. *Mol Med Camb Mass*. 2022 Apr 14;28(1):42.
42. Sen M, Wang X, Hamdan FH, Rapp J, Eggert J, Kosinsky RL, et al. ARID1A facilitates KRAS signaling-regulated enhancer activity in an AP1-dependent manner in colorectal cancer cells. *Clin Epigenetics*. 2019 Jun 19;11(1):92.
43. Fajrial AK, He QQ, Wirusanti NI, Slansky JE, Ding X. A review of emerging physical

- transfection methods for CRISPR/Cas9-mediated gene editing. *Theranostics*. 2020 Apr 15;10(12):5532–49.
44. Skoufou-Papoutsaki N, Adler S, D'Santos P, Mannion L, Mehmed S, Kemp R, et al. Efficient genetic editing of human intestinal organoids using ribonucleoprotein-based CRISPR. *Dis Model Mech*. 2023 Oct 5;16(10):dmm050279.
 45. Wang Y, Zhai Y, Zhang M, Song C, Zhang Y, Zhang G. Escaping from CRISPR–Cas-mediated knockout: the facts, mechanisms, and applications. *Cell Mol Biol Lett*. 2024 Apr 8;29:48.
 46. Musunuru K. Chapter 2 - Choosing a genome editing strategy and target site. In: Musunuru K, editor. *Genome Editing* [Internet]. Academic Press; 2021 [cited 2024 Sep 20]. p. 21–39. Available from: <https://www.sciencedirect.com/science/article/pii/B9780128234846000116>
 47. Gene Knockout - an overview | ScienceDirect Topics [Internet]. [cited 2024 Sep 20]. Available from: <https://www.sciencedirect.com/topics/medicine-and-dentistry/gene-knockout>
 48. Webb PM, Jordan SJ. Global epidemiology of epithelial ovarian cancer. *Nat Rev Clin Oncol*. 2024 May;21(5):389–400.
 49. Hollis RL. Molecular characteristics and clinical behaviour of epithelial ovarian cancers. *Cancer Lett*. 2023 Feb 28;555:216057.
 50. Cho KR, Shih IM. Ovarian Cancer. *Annu Rev Pathol Mech Dis*. 2009 Feb 28;4(Volume 4, 2009):287–313.
 51. Köbel M, Kang EY. The Evolution of Ovarian Carcinoma Subclassification. *Cancers*. 2022 Jan 14;14(2):416.
 52. Pinto MT, Eiras Martins G, Vieira AGS, Galvão JMS, de Pádua Souza C, Macedo CRPD, et al. Molecular Biology of Pediatric and Adult Ovarian Germ Cell Tumors: A Review. *Cancers*. 2023 May 30;15(11):2990.
 53. Pinto MT, Eiras Martins G, Vieira AGS, Galvão JMS, de Pádua Souza C, Macedo CRPD, et al. Molecular Biology of Pediatric and Adult Ovarian Germ Cell Tumors: A Review. *Cancers*. 2023 May 30;15(11):2990.
 54. Tischkowitz M, Huang S, Banerjee S, Hague J, Hendricks WPD, Huntsman DG, et al. Small Cell Carcinoma of the Ovary, Hypercalcaemic Type – genetics, new treatment targets and current management guidelines. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2020 Aug 1;26(15):3908–17.
 55. Hollis RL, Thomson JP, Stanley B, Churchman M, Meynert AM, Rye T, et al. Molecular stratification of endometrioid ovarian carcinoma predicts clinical outcome. *Nat Commun*. 2020 Oct 5;11:4995.
 56. Hollis RL, Meynert AM, Michie CO, Rye T, Churchman M, Hallas-Potts A, et al. Multiomic Characterization of High-Grade Serous Ovarian Carcinoma Enables High-Resolution Patient Stratification. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2022 Aug 15;28(16):3546–56.

57. Lengyel E. Ovarian cancer development and metastasis. *Am J Pathol*. 2010 Sep;177(3):1053–64.
58. Hollis RL, Meynert AM, Michie CO, Rye T, Churchman M, Hallas-Potts A, et al. Multiomic Characterization of High-Grade Serous Ovarian Carcinoma Enables High-Resolution Patient Stratification. *Clin Cancer Res*. 2022 Aug 15;28(16):3546–56.
59. Centini G, Schettini G, Pieri E, Giorgi M, Lazzeri L, Martire FG, et al. Endometriosis-Related Ovarian Cancer: Where Are We Now? A Narrative Review towards a Pragmatic Approach. *J Clin Med*. 2024 Mar 27;13(7):1933.
60. Krämer P, Talhouk A, Brett MA, Chiu DS, Cairns ES, Scheunhage DA, et al. Endometrial Cancer Molecular Risk Stratification is Equally Prognostic for Endometrioid Ovarian Carcinoma. *Clin Cancer Res*. 2020 Oct 15;26(20):5400–10.
61. Scopus preview - Scopus - Document details - Endometrioid epithelial ovarian cancer: 20 years of prospectively collected data from a single center [Internet]. [cited 2024 Sep 11]. Available from: <https://www.scopus.com/record/display.uri?eid=2-s2.0-43049105741&origin=inward&txGid=d0859a7247526564ceb35b42847aa242>
62. Peres LC, Cushing-Haugen KL, Anglesio M, Wicklund K, Bentley R, Berchuck A, et al. Histotype classification of ovarian carcinoma: A comparison of approaches. *Gynecol Oncol*. 2018 Oct;151(1):53–60.
63. Phung MT, Pearce CL, Meza R, Jeon J. Trends of Ovarian Cancer Incidence by Histotype and Race/Ethnicity in the United States 1992–2019. *Cancer Res Commun*. 2023 Jan 3;3(1):1–8.
64. Oliver KE, Brady WE, Birrer M, Gershenson DM, Fleming G, Copeland LJ, et al. An evaluation of progression free survival and overall survival of ovarian cancer patients with clear cell carcinoma versus serous carcinoma treated with platinum therapy: An NRG Oncology/Gynecologic Oncology Group experience. *Gynecol Oncol*. 2017 Nov;147(2):243–9.
65. Tissue Microarray - an overview | ScienceDirect Topics [Internet]. [cited 2024 Sep 11]. Available from: <https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/tissue-microarray>
66. Pollack JR. A Perspective on DNA Microarrays in Pathology Research and Practice. *Am J Pathol*. 2007 Aug;171(2):375–85.
67. Xiao F, Wang Z, Qiao L, Zhang X, Wu N, Wang J, et al. Application of PARP inhibitors combined with immune checkpoint inhibitors in ovarian cancer. *J Transl Med*. 2024 Aug 21;22:778.
68. Antibody drug conjugate: the “biological missile” for targeted cancer therapy | Signal Transduction and Targeted Therapy [Internet]. [cited 2024 Sep 22]. Available from: <https://www.nature.com/articles/s41392-022-00947-7>

69. Tolcher A, Hamilton E, Coleman RL. The evolving landscape of antibody-drug conjugates in gynecologic cancers. *Cancer Treat Rev.* 2023 May;116:102546.
70. Marei HE, Cenciarelli C, Hasan A. Potential of antibody–drug conjugates (ADCs) for cancer therapy. *Cancer Cell Int.* 2022 Aug 13;22:255.
71. Bosi C, Bartha Á, Galbardi B, Notini G, Naldini MM, Licata L, et al. Pan-cancer analysis of antibody-drug conjugate targets and putative predictors of treatment response. *Eur J Cancer.* 2023 Dec 1;195:113379.
72. Persenaire C, Bitler BG, Corr BR. Folate receptor alpha protein expression in ovarian serous cystadenocarcinoma tumors of The Cancer Genome Atlas: exploration beyond single-agent therapy. *medRxiv.* 2024 Apr 14;2024.04.12.24305742.
73. FOLR1 folate receptor alpha [Homo sapiens (human)] - Gene - NCBI [Internet]. [cited 2024 Sep 11]. Available from: <https://www.ncbi.nlm.nih.gov/gene/2348>
74. Folate Receptor 1 - an overview | ScienceDirect Topics [Internet]. [cited 2024 Sep 11]. Available from: <https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/folate-receptor-1>
75. Matulonis UA, Lorusso D, Oaknin A, Pignata S, Dean A, Denys H, et al. Efficacy and Safety of Mirvetuximab Soravtansine in Patients With Platinum-Resistant Ovarian Cancer With High Folate Receptor Alpha Expression: Results From the SORAYA Study. *J Clin Oncol.* 2023 May 1;41(13):2436–45.
76. Gitto SB, Whicker M, Davies G, Kumar S, Kinneer K, Xu H, et al. A B7-H4-Targeting Antibody-Drug Conjugate Shows Antitumor Activity in PARPi and Platinum-Resistant Cancers with B7-H4 Expression. *Clin Cancer Res Off J Am Assoc Cancer Res.* 2024 Apr 15;30(8):1567–81.
77. Zhou J, Tian Y, Qu Y, Williams M, Yuan Y, Karvas RM, et al. The immune checkpoint molecule, VTCN1/B7-H4, guides differentiation and suppresses proinflammatory responses and MHC class I expression in an embryonic stem cell-derived model of human trophoblast. *Front Endocrinol.* 2023 Mar 16;14:1069395.
78. Vtcn1 V-set domain containing T cell activation inhibitor 1 [Mus musculus (house mouse)] - Gene - NCBI [Internet]. [cited 2024 Sep 11]. Available from: <https://www.ncbi.nlm.nih.gov/gene/242122>
79. Lombardi P, Filetti M, Falcone R, Altamura V, Paroni Sterbini F, Bria E, et al. Overview of Trop-2 in Cancer: From Pre-Clinical Studies to Future Directions in Clinical Settings. *Cancers.* 2023 Mar 13;15(6):1744.
80. Wen Y, Ouyang D, Zou Q, Chen Q, Luo N, He H, et al. A literature review of the promising future of TROP2: a potential drug therapy target. *Ann Transl Med [Internet].* 2022 Dec [cited 2024 Sep 11];10(24). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9843409/>
81. Toader D, Fessler SP, Collins SD, Conlon PR, Bollu R, Catcott KC, et al. Discovery and

- Preclinical Characterization of XMT-1660, an Optimized B7-H4-Targeted Antibody–Drug Conjugate for the Treatment of Cancer. *Mol Cancer Ther*. 2023 Sep 5;22(9):999–1012.
82. Zhong H, Xiong Y, Huang H, Pan Y, Cao L, Liu S, et al. Preclinical antitumor activity of an anti-HER2/Trop-2 bispecific antibody-drug conjugate with a new DNA topoisomerase I inhibitor. *J Clin Oncol*. 2023 Jun;41(16_suppl):e15013–e15013.
 83. Gibson-Corley KN, Olivier AK, Meyerholz DK. Principles for valid histopathologic scoring in research. *Vet Pathol*. 2013 Nov;50(6):1007–15.
 84. Liljequist D, Elfving B, Skavberg Roaldsen K. Intraclass correlation – A discussion and demonstration of basic features. *PLoS ONE*. 2019 Jul 22;14(7):e0219854.
 85. Package overview — pandas 2.2.3 documentation [Internet]. [cited 2024 Sep 22]. Available from: https://pandas.pydata.org/docs/getting_started/overview.html
 86. NVIDIA Data Science Glossary [Internet]. [cited 2024 Sep 12]. What is pandas Python? Available from: <https://www.nvidia.com/en-gb/glossary/pandas-python/>
 87. Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol*. 2012;8(1):23.
 88. Hecht JL, Kotsopoulos J, Gates MA, Hankinson SE, Tworoger SS. Validation of tissue microarray technology in ovarian cancer: results from the Nurses' Health Study. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol*. 2008 Nov;17(11):3043–50.
 89. Borderline ovarian tumours [Internet]. [cited 2024 Sep 21]. Available from: <https://www.cancerresearchuk.org/about-cancer/ovarian-cancer/types/borderline>
 90. Bobak CA, Barr PJ, O'Malley AJ. Estimation of an inter-rater intra-class correlation coefficient that overcomes common assumption violations in the assessment of health measurement scales. *BMC Med Res Methodol*. 2018 Sep 12;18(1):93.
 91. Martin LP, Konner JA, Moore KN, Seward SM, Matulonis UA, Perez RP, et al. Characterization of folate receptor alpha (FR α) expression in archival tumor and biopsy samples from relapsed epithelial ovarian cancer patients: A phase I expansion study of the FR α -targeting antibody-drug conjugate mirvetuximab soravtansine. *Gynecol Oncol*. 2017 Nov;147(2):402–7.
 92. Moore KN, Borghaei H, O'Malley DM, Jeong W, Seward SM, Bauer TM, et al. Phase 1 dose-escalation study of mirvetuximab soravtansine (IMGN853), a folate receptor α -targeting antibody-drug conjugate, in patients with solid tumors. *Cancer*. 2017 Aug 15;123(16):3080–7.
 93. Gray E, Ulrich M, Epp A, Younan P, Sahetya D, Hensley K, et al. SGN-B7H4V, an investigational vedotin ADC directed to the immune checkpoint ligand B7-H4, shows promising activity in preclinical models. *J Immunother Cancer*. 2023 Oct 1;11(10):e007572.

