An investigation of the impact of the B cell lymphoma factor Activation Induced Cytidine Deaminase (AICDA/AID) on epigenetic plasticity

By

Lincon Kgasha Makofane

MKFLIN002

MSc (Med) in Haematology

Faculty of Health Sciences

University of Cape Town

May 2025

Supervisor: A/Prof Shaheen Mowla

Co-Supervisor: Dr Beatrice Ramorola

Division of Haematology, Department of Pathology, Faculty of Health Sciences, University of Cape Town

Declaration

I, Lincon Kgasha Makofane, hereby declare that the work on which this dissertation (MSc Med) is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other University.

I empower the University to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Date: May 2025

Dedication

I dedicate this to all the courageous cancer patients who face their battles with resilience and strength.

Acknowledgments

First, I would like to extend my deepest gratitude to **God** for granting me the strength, wisdom, and perseverance to navigate this journey.

To my beloved mother, your constant support and unwavering belief in my dreams have been my anchor during the toughest times. Your love and encouragement have given me the strength to keep pushing forward.

To my two beautiful sisters, **Manku** and **Lucia**, and my dear boy, **Kganya**, this one is for you. Your love and encouragement have been a source of immense strength and inspiration. To the rest of the Makofane family, thank you for your unwavering support and for always being there for me.

A heartfelt thank you to my fellow lab members and friends: Lungile, Babalwa, Thando, and Maxine. The memories, laughter, and support we shared have made this journey truly special. To the Division of Haematology staff, A/Prof Jessica Opie, A/Prof Karen Shires, Jolene, Cylene, and Zhaheed, thank you for your unwavering support and for creating such a welcoming and nurturing environment.

I would like to extend special thanks to the following individuals for their invaluable contributions to this work:

- **Drs Zahra Latib** and **Aaliyah Saferdien**, thank you for your assistance during the final weeks of writing. It has been a pleasure working and sharing the lab with you. I wish you the best of luck in your career.
- **Dr. Leonardo Alves De Souza Rios**, thank you for your assistance with the lentiviral transduction system work.
- **Prof. Sharon Prince Lab**, thank you for kindly donating the 5-FU experimental drug.
- **Dr. Karabo Serala**, thank you for your assistance with the generation of the CRC cell line model.

I am deeply grateful to the funders: The National Research Foundation and the Fogarty International Centre of the National Institute for their support through the Fogarty Research Training Program.

To my co-supervisor, **Dr Beatrice Ramorala**, thank you for making it easy to adapt to the lab environment and for sharing your extensive knowledge with me. I am immensely appreciative of your guidance and wish you all the best in your career.

Lastly, I would like to express my heartfelt thanks to my supervisor, **A/Prof Shaheen Mowla**, for always keeping her door open and for providing invaluable guidance throughout this journey. It is an absolute pleasure working under your mentorship. Thank you so much.

Table of Contents

Abbreviations

µL	Microliter
AICDA	Activation-induced cytidine deaminase
ANOVA	Analysis of variance
APE-1	apurinic/apyrimidinic endonuclease
APOBEC	Apolipoprotein B mRNA Editing Catalytic Polpeptide-like
ART	Antiretroviral therapy
BCL2	B-cell lymphoma 2
BCR	B cell receptor
BER	Base Excision Repair
BL	Burkitt lymphoma
CRC	Colorectal cancer
CSR	Class switch recombination
DSB	Double Strand Break
DHL	Double hit lymphoma
DLBCL	Diffuse large b cell lymphoma
DNMT family	DNA methyltransferase
EBV	Epstein-Barr virus
EMT	Epithelial-mesenchymal transition

ERK	Extracellular signal-regulated kinase
FANCA	Fanconi anaemia complementation group A
GC	Germinal Center
GLOBOCAN	Global Cancer Observatory
HBV	Hepatitis B/C virus
HCC	Hepatocellular carcinoma
HIV	Human Immunodeficiency virus
HPV	Human papillomavirus
IGH	Immunoglobulin heavy locus
Mg	Milligram
MMR	Mismatch repair
mRNA	messenger RNA
MYC	Myelocytoma
MyD88	Myeloid differentiation primary response 88
NHL	non-Hodgkin lymphoma
^{o}C	degrees celsius
PBS	Phosphate Buffered Saline
PDL-1	Programmed Cell Death Ligand 1
PIM1	PIM-1 proto-oncogene, serine/threonine kinase
PLWH	People Living with HIV
pRb	Retinoblastoma protein
PTEN	Phosphatase and tensin homolog
RAS	Rat sarcoma

SHM	Somatic hypermutation
SMUG1	single-strand selective monofunctional uracil DNA glycosylase
TCGA	The Cancer Genome Atlas
TET2	Ten-Eleven Translocation 2
TGF-β	Transforming Growth Factor beta
TIR	Translocated intimin receptor
TLR4	Toll-like receptor 4
TNF-α	Tumour Necrotic Factor alpha
TP53	tumour protein p53
Trp53	transformation-related protein 53
UNG	Uracil DNA glycosylase
ZEB family	Zinc Finger E-Box Binding Homeobox 1

Abstract

Cancer is a major cause of both mortality and morbidity globally, with 10 million deaths reported in 2020 by the Global Cancer Observatory (GLOBOCAN). To address this burden, it is crucial to deepen our understanding of the biology of the disease to create opportunities for addressing therapeutic challenges. Epigenetic regulation has recently emerged as a prominent factor in cancer progression, and aberrant epigenetic alterations have been reported in many key cancer-driving genes. The overexpression of the DNA modifying enzyme Activation-induced cytidine deaminase (AICDA) is a factor linked to cancer progression, particularly B-cell derived non-Hodgkin lymphomas, but has also been associated with other inflammation-associated cancers, including colorectal cancer (CRC). The widely recognised normal role of AICDA is in antibody-producing B cells, where its cytosine deaminating ability within DNA is essential for the production of a diverse array of antibodies. In high-grade B cell lymphomas, such as diffuse large B cell lymphoma (DLBCL) and Burkitt lymphoma (BL), overexpression of *AICDA* is particularly associated with oncogenic genomic translocations, thought to be driven by aberrant DNA repair, following the deamination events. More recently, the oncogenic function of AICDA has been expanded to include modulation of the epigenetic landscape, through the enzyme's ability to alter CpG sites/islands, because of its deamination function. Additionally, a few studies have shown that AICDA can partner with other regulatory proteins to modulate gene methylation. However, such a direct involvement of AICDA in the alteration of gene expression by disrupting DNA methylation remains uncertain due to a lack of scientific evidence. The current study therefore sought to investigate the impact of AICDA on the transcription of key oncogenes, with the view to focus on the role of AICDA in their methylation status in future studies.

The human AICDA open reading frame (ORF) was cloned into a lentiviral backbone vector, and AICDA-overexpression was verified in HEK293 cells using western blotting. A third-generation Lentiviral transduction system was used to produce lentiviral particles (AICDA expressing and empty-vector control) to transduce a non-cancerous lymphoblastoid cell line (PB-B95-8H) and a DLBCL cell line (HBL-1). Flow cytometry was used to sort successfully transduced cells. However, the AICDA-overexpressing cells did not survive in culture for longer than 72 hours post-sorting. This is likely due to oncogene-induced death, as these cells already express AICDA at relatively high levels. An alternate CRC model was chosen and was justified by the published literature as well as in silico analyses using the Cancer

Genome Atlas (TCGA) CRC dataset which indicated that AICDA expression in CRC correlated with poor survival.

Two CRC cell lines stably expressing AICDA were developed and were found to tolerate high levels of AICDA. Thereafter, proliferation was assessed using growth curve analysis, and WST-1 assay. Growth curves showed that expressing AICDA significantly enhanced cell proliferation. In addition, the WST-1 cell proliferation assay also showed similar results, indicating that AICDA expression promotes cell proliferation. A wound healing assay was performed to investigate the impact of AICDA expression on the migration, and the results indicated that AICDA expression in CRC cells enhances the migratory ability if the cells, relative to the control.

CRC cells expressing AICDA had higher IC_{50} values compared to control cells, when treated with increasing doses of the chemotherapeutic drug 5-Fluorouracil (5-FU), indicating a resistance to cell death. However, this was not reflected in western blot analyses where the expressions of apoptotic markers PARP-1 and Caspase-3 were assessed. The cell cycle profiling showed that AICDA provides mild protection from cell death induced by 5-FU, and delays S-phase, in CRC cells. Quantitative PCR (qPCR) was performed to investigate the impact of AICDA expression on the transcriptional regulation of seven key oncogenes, namely *PIM1, FANCA, DNMT1, DNMT3A, ZEB1, CD274*, and *MYC*. The results showed that AICDA expression in CRC significantly impacted the expression level of the selected oncogenes. Variations could be observed between the two CRC cell models, which imply that the cellular context is important in mediating the impact of AICDA on changes in gene transcription. Additionally, both upregulation and downregulation of genes were observed, indicating diversity in the mechanism via which AICDA modulates gene expression. The findings of this study therefore demonstrate that AICDA functions as an oncogene in CRC and provides the cells with a significant proliferative advantage. Crucially, it shows that AICDA expression impacts the transcription of key oncogenes. Future work will focus on investigating the role of methylation in the transcriptional changes observed, and how AICDA is involved in this process.

Keywords: AICDA, Cancer, epigenetics.

Chapter 1

1.1 Introduction

Cancer is a major cause of both mortality and morbidity globally, with 10 million deaths reported in 2020 by the Global Cancer Observatory (GLOBOCAN) [1]. Infectious agents, including the Human Immunodeficiency virus (HIV), Helicobacter pylori (H. pylori), Hepatitis B (HBV) and C viruses, and Human papillomavirus (HPV), continue to play a significant role in oncogenesis [2]. For example, Kaposi sarcoma, B-cell non-Hodgkin lymphomas (B-NHLs), and cervical cancer are strongly associated with HIV infection [2]. This is very pertinent to the countries with high rates of HIV infection, like South Africa, where approximately 7.8 million people (12.7% of the total population) were reported to be infected with the virus in 2024 [3]. Although access to Antiretroviral therapy (ART) has greatly improved in South Africa, there are still ART access-related barriers for HIV-infected patients, such as negative perceptions, lack of family and community support, and status disclosure issues [4]. Epidemiological studies also show that people living with HIV (PLWH) are at greater risk of developing NHL compared to the general population [5, 6]. Additionally, chronic inflammatory signals driven by these infectious agents are also known risk factors for various cancers, including colorectal cancer (CRC), hepatocellular carcinoma (HCC), breast cancer, lymphoma, and liver cancer [7, 8]. While cancer treatment has advanced significantly, it still faces considerable challenges. Traditional approaches, such as chemotherapy, often lead to a therapeutic plateau. Alternative and more targeted therapies are needed, which underpins cancer research.

Cancer is a genetic disease characterised by the uncontrolled proliferation and abnormal growth of cells, which can invade other tissues and organs through the bloodstream and lymphatic system. Cancer can manifest in any of the cell types, with some being more prevalent than others. For example, the most common cancer among women worldwide remains cancer of the breast, while prostate cancer is most common among men [9]. The hallmarks of cancer cells have been well studied and reported, as a set of characteristics that form the basis of research to define the complexities of cancer cells, their development, and drug response. The term “hallmarks of cancer” was first coined in 2000 by Douglas Hanahan and Robert Weinberg, who originally proposed six major biological processes [10]. In 2022, four more hallmarks were added to the existing categories, (**Figure 1.1**) [9]. It is now well recognised that cancer is a genetic disease resulting from the accumulation of mutations that disrupt the homeostatic balance, progressively transforming normal human cells into their malignant derivatives. The causes and nature of these genetic alterations are extremely diverse, rendering cancer

a highly complex disease, with multiple phenotypically and molecularly distinct subtypes within the same cancer type, hence the challenges of accurate diagnosis and treatment.

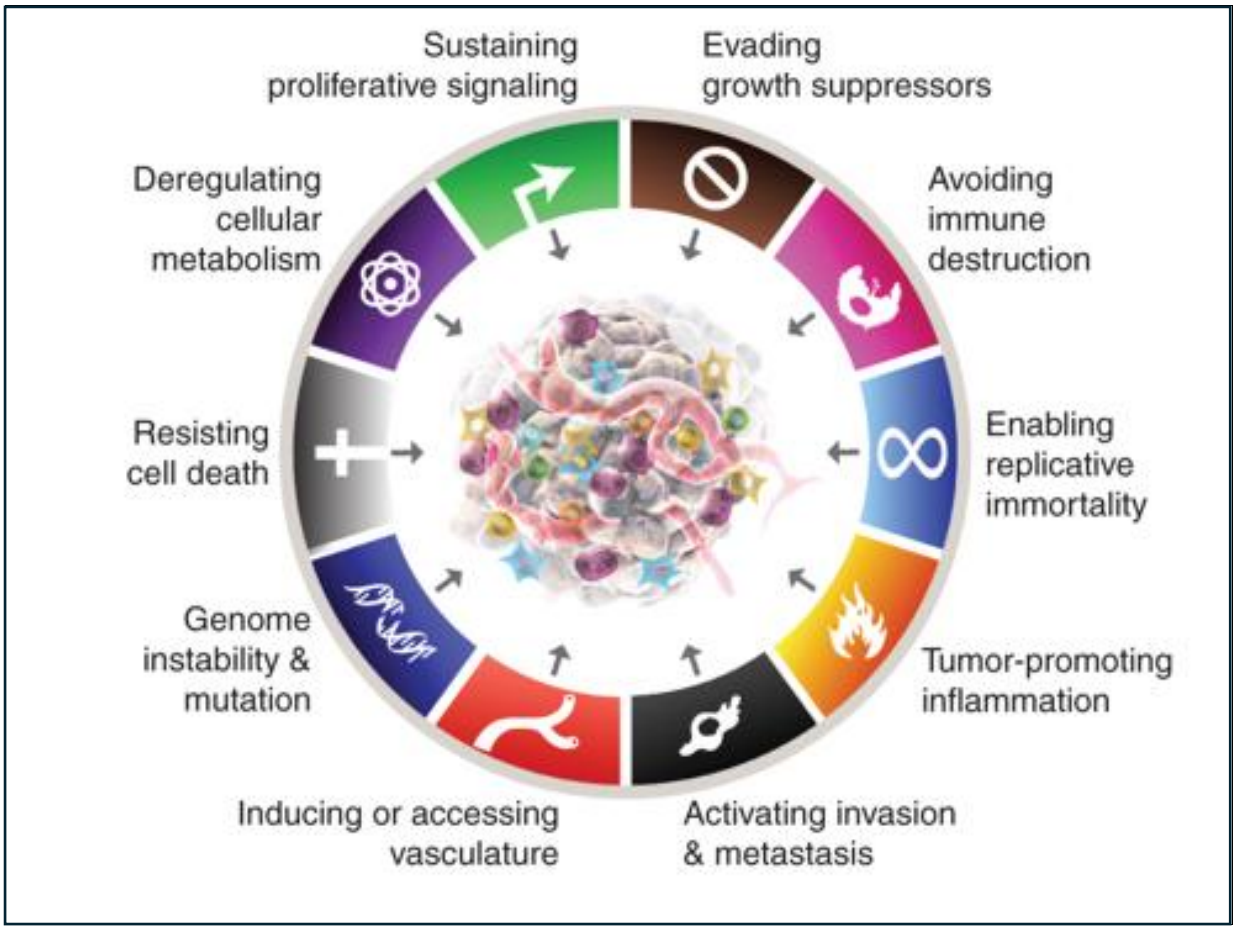


Figure 1.1: The ten hallmarks of cancer. From Hanahan, 2022 [9]

Cancer cells have several distinct characteristics that differentiate them from normal, healthy cells. The key traits are uncontrolled proliferation, primarily as a result of bypassing cell cycle checkpoints; and avoidance of cell death through disruption of death signalling pathways, allowing survival even when the cell has accumulated DNA damage and other abnormalities [11]. Cancer cells can evade mechanisms that suppress excessive growth by inactivating tumour suppressor genes, which are crucial for regulating cell division. Like normal tissues, cancer cells require sustenance in the form of nutrients and oxygen, they therefore stimulate the growth of new blood vessels, inducing vasculature, and providing themselves with the necessary nutrients and oxygen. Cancer cells also have migratory abilities by breaking down the extracellular matrix and navigating through the lymphatic system to invade surrounding tissues and spread to distant sites in the body, leading to secondary tumours. These cancer cells often hijack the body's inflammatory pathway to enhance their growth and survival.

For instance, when necrotic cells burst and release their contents into the surrounding tissue. This release may attract inflammatory immune cells. Numerous studies suggest that these inflammatory immune cells may be actually implicated in promoting cancer progression [12]. Tumours, in contrast to normal tissues, have the ability to continue dividing because they preserve their telomeres. Telomeres primarily function to protect the ends of chromosomes and limit cellular proliferation. Consequently, cancer cells are often described as replicative immortals because they can proliferate without exhibiting signs of senescence or crisis. Cancer cells are known to evade detection and destruction by the immune system in several ways. They may suppress immune responses or create an immunosuppressive microenvironment that protects them. In addition, cancer cells often have increased genetic instability that results in a high mutation rate. Many of them exhibit defects in DNA repair mechanisms, leading to the accumulation of mutations. These genetic alterations drive the progression and adaptation of cancer [9, 13, 14].

1.2 Cancer driver genes

It is now well known that cancer is a genetic disorder characterised by dynamic changes in the genome caused by alterations in genes that regulate the hallmarks of the disease. Some of the key mechanisms driving oncogenesis include mutations, epigenetic changes, and chromosomal rearrangement. For example, cancer-driver genes, unlike those non-cancer driver genes, undergo mutations that alter their normal regulation which gives cells a proliferative advantage. These can be proto-oncogenes or tumour suppressor genes. Genes such as *MYC*, *RAS,* and *ERK* are proto-oncogenes that acquire gain-of-function mutations to promote the oncogenic process; while genes such as *TP53*, *PTEN,* and *pRb* are tumour suppressor genes that become inactivated [15, 16]. Genetic alterations in these genes are generally associated with unfavourable clinical outcomes, such as metastatic disease, poor response to treatment and poor survival. The overexpression of an enzyme called activation-induced cytidine deaminase (AICDA) has been described as an enhancer mutational load within cells, contributing to the tumour-related gene mutations.

1.3 The DNA modifying enzyme Activation Induced Cytidine Deaminase (AICDA)

In 1999, Tasuku Honje's group at Kyoto University first cloned a gene sequence associated with class switch recombination (CSR). Subsequent analyses revealed that this gene exhibited a significant homology of 34% in amino acid identity with the apolipoprotein B (apo B) mRNA editing enzyme, catalytic polypeptide-like (APOBEC) family of cytosine deaminases [17]. The gene in question was

activation-induced cytidine deaminase (*AICDA*), which encodes a unique nucleic acid-modifying factor that is the only member of the APOBEC family with DNA-modifying activities.

Interestingly, the group also found that AICDA is specifically induced in the germinal center (GC) of secondary lymphoid tissues (lymph nodes, spleen, and tonsils), implicating its function in genetic events of GC B cells [18]. The mechanisms of antibody diversity that occur in GC, somatic hypermutation (SHM), and CSR were later demonstrated to be critically dependent on AICDA activity in GC B cells [18, 19]. As a deaminase, AICDA converts deoxycytidine (dC) to deoxyuridine (dU) within the DNA sequence, which creates a U: G mismatch, triggering DNA repair pathways that ultimately lead to changes in the repaired DNA sequence, via the activity of low-fidelity polymerase [20]. This is crucial in promoting the secretion of diverse antibodies by B cells. It is therefore unsurprising that the dysregulation of this factor is strongly associated with cancer, particularly aggressive B-cell-derived non-Hodgkin lymphomas (NHLs), but also inflammation-associated malignancies such as gastric and liver cancer [21, 22, 23].

1.4 Regulation of SHM by AICDA

The hot spot motifs of SHM have a high density of RGYW/WRCY (W=A/T, R=A/G, and Y=C/T) at the G-rich regions compared to the rest of the genome. AICDA preferentially targets the WRCY sequence region to facilitate induction of SHM, which leads to a single point mutation in the antibody V(D)J regions [24, 25]. After binding AICDA naturally slides along the single-strand regions of DNA formed during transcription to search for the hotspot motifs. Once found, it deaminates cytosine [26]. Uracil lesions induced by AICDA within the DNA sequence can be effectively repaired by free error-prone DNA repair, leading to no SHM (**Figure 1.2**) [27].

However, in the case of SHM, downstream error-prone DNA repair pathways, non-canonical base-excision repair (ncBER), and mismatch repair (ncMMR) – may be activated following the AICDA deamination process. As established, the introduced mismatch thus bypasses detection for restoration into the original DNA sequence due to the ncBER pathway [28]. It is also established that after AICDA-induced deamination, the uracil base is often recognised by uracil DNA glycosylase (UNG), leading to uracil excision, and generating the abasic site. Apyrimidinic endonuclease 1 (APE-1) then nick the DNA at the abasic site, DNA polymerase β or other polymerase seal in the nick gap, which is joined together by DNA ligase 3 [29, 30]. Therefore, the uracil that bypasses for detection by UNG or single-strand-selective monofunctional uracil-DNA glycosylase 1 (SMUG1) can be copied to replicate over abasic sites by the error-prone translesion synthesis polymerase (TLS) [29], which results in single-point mutations.

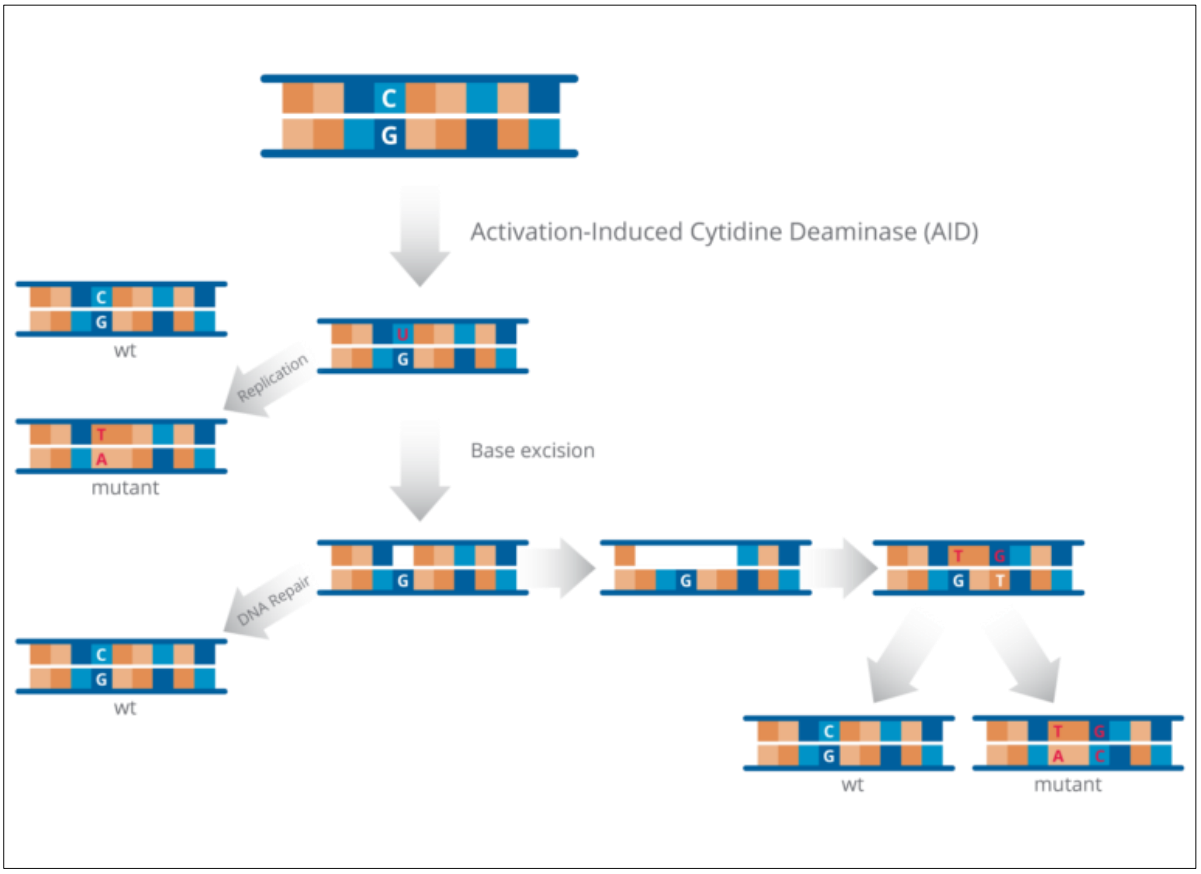


Figure 1.2 Diagram showing the involvement of AICDA resulting in mutagenesis during the early stages of SHM. AID converts cytosine (**C**) into uracil (**U**), resulting in mismatches. These mismatches proceed before correction due to error-prone repair pathways, such as BER or MMR, initiating a process described in the text above. In short, this process leads to SHM, which introduces mutations in the variable regions of BCR. From Rios's PhD thesis, 2020 [31]

1.5 Regulation of CSR by AICDA

The process of CSR involves site-specific DNA deletional recombination. In CSR, the DNA sequence in the switch (S) region DNA upstream of the exons encoding for the genes of the heavy chain locus of Ig-M expressing B cells is replaced with a downstream C_H region. This results in a switch of the antibody isotype from IgM to IgG, IgE, or IgA without changing the antigen-binding specificity [32, 33]. The process of CSR is initiated by AICDA, which deaminates cytosine nucleotides in the S regions. As a result, these mismatches elicit the action of DNA repair machinery, leading to double-strand breaks (DSBs). CSR occurs within 4-7 kb S regions, which contain numerous AICDA hot spots – (AGCT) motifs, which are palindromic variants of the WRCY motif located upstream and downstream of each C_H region

at the transcription start site. Due to its nature, AICDA targets S regions with the help of 14-3-3 adaptor proteins, which specifically bind to the (5'-AGCT-3') motifs and convert cytosine to uracil [34]. The uracil residues introduced by AICDA could be removed by UNG, a process stabilised by the scaffold proteins, such as Rev1 and, 14-3-3 adaptors, and replication protein A (RPA). These factors contribute to ssDNA breaks, while DSBs are formed when ssDNA breaks occur close to each other on opposite DNA strands or by MMR repair pathway on the S region (**Figure 1.3**) [35, 36]. The resulting DSBs in CSR primarily rely on canonical non-homologous end-joining (NHEJ) to join blunt DSBs, whereas staggered DSBs are rejoined by alternative end-joining (A-EJ) [37].

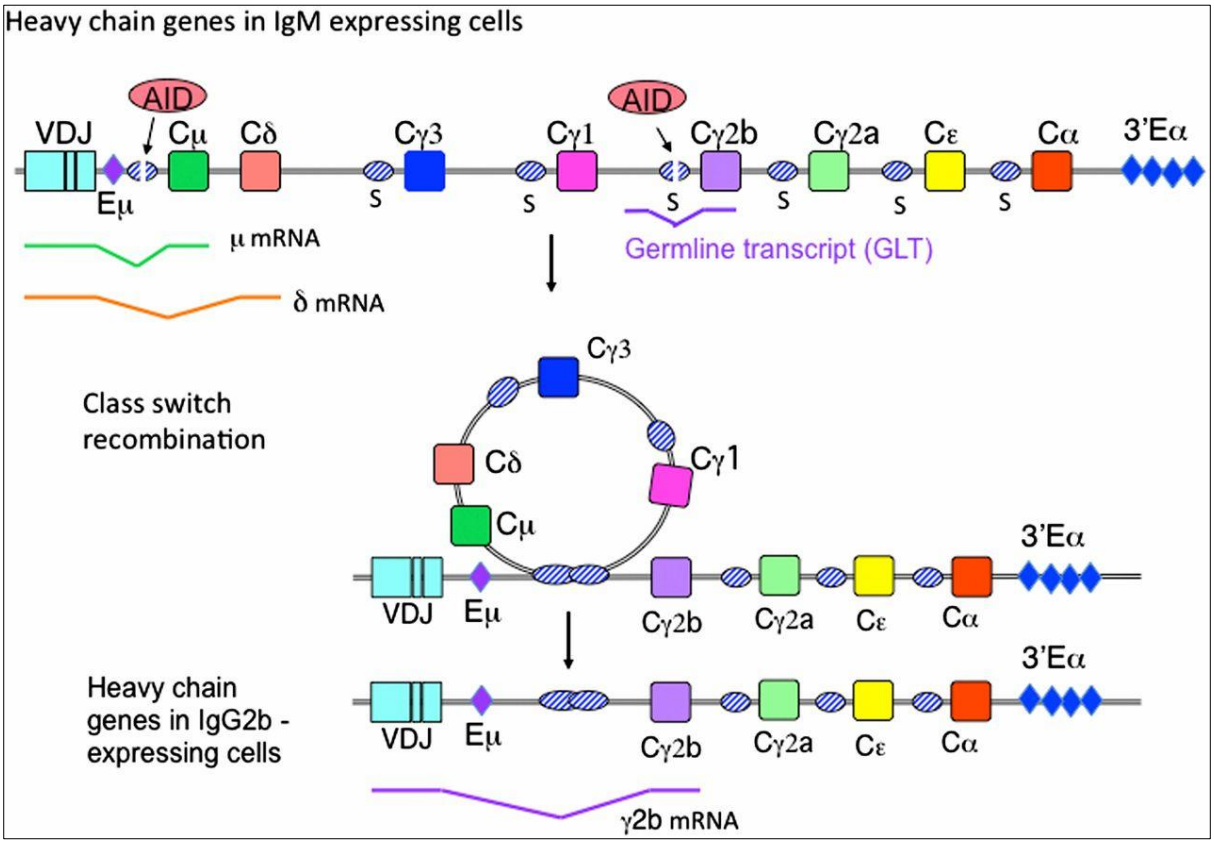


Figure 1.3 Diagram shows the involvement of AICDA in mediating DSB during CSR. During CSR, AID binds to the motif regions upstream and downstream of constant heavy chain genes in IgM-expressing cells. This mediates the DSB at the active transcriptional S region, with mechanisms of NHEJ and A-EJ rejoining the variable regions. From Stavnezer et al., 2014 [36]

1.6 The regulation of AICDA Expression

The regulation of AICDA expression occurs at multiple levels, i.e., at the transcriptional, post-transcriptional, and post-translational levels, which, when disrupted lead to cellular dysfunctions that underpin pathological conditions. While an inflammatory milieux is known to enhance AICDA expression, the precise mechanisms leading to aberrant expression of this enzyme is not well

described. In the case of HIV infection, elevated expression of *AICDA* mRNA expression has been shown in the peripheral blood mononuclear cells (PBMCs) of HIV-infected individuals, relative to uninfected ones [38]. Notably, the expression was highest in those patients who went on to develop lymphoma. Studies from our lab (Prof Mowla's laboratory at the University of Cape Town) and that of others have demonstrated that HIV-1 auxiliary proteins, particularly transactivator of transcription (Tat) [39], and negative factor (Nef) can enhance AICDA expression in B cells [40]. The dysregulated AICDA leads to off-target mutations in some key oncogenes, thereby promoting cell proliferation, and migration, and enhancing chemoresistance in B cell lymphoma (**Figure 1.4a**) [41].

AICDA is also implicated as an oncogenic factor in inflammation-associated cancers. These include HCC, CRC, and other cancers linked to infectious agents such as *H. pylori* and the *HBV C* [22, 42]. The dysregulation of pro-inflammatory cytokine signaling, probably in combination with other not-yet-identified factors, creates a supportive microenvironment that triggers aberrant AICDA expression in a non-B cell background. It has been reported that cytokines, including interleukin-4 (IL-4), IL-13, tumour necrosis factor-alpha (TNF-α), and transforming growth factor-beta (TGF-β), are involved in the aberrant regulation of AICDA [21, 42, 43, 44]. The dysregulated AICDA leads to off-target activity in the epithelial cells, and these changes affect the oncogenes involved in the progression of CRC, skin, breast, pancreatic cancer, and bladder urothelial carcinoma (**Figure 1.4b**) [31].

Importantly, a pioneering study by Yoko Endo et al (2008) demonstrated that pro-inflammatory cytokines such as IL-4 and IL-13 could induce inappropriate AICDA activity, resulting in the generation of genetic aberrations in tumour-related genes such as TP53 [43]. Consistent with this hypothesis, Takai et al (2012) revealed in mice models through sequencing analysis that knockout of AICDA significantly lowers the incidence of somatic mutations in the Trp53 gene [45]. Taken together, these findings strongly suggest that the aberrant expression of AICDA is not limited to B cells, but also a variety of other cell types, where it promotes oncogenic events [7, 46].

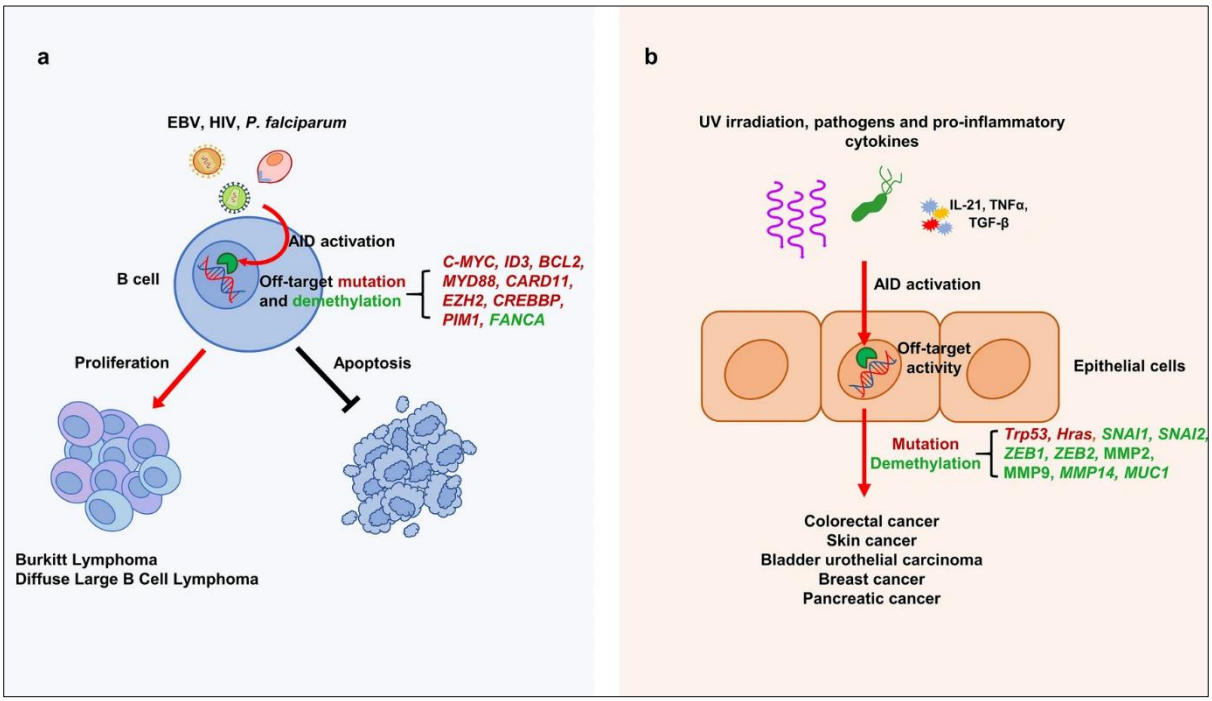


Figure 1.4 A schematic illustration of the AICDA dysregulation in various cancers, affecting downstream genes. **a** Dysregulation of AICDA by EBV, HIV infection, and *P. falciparum* affects key oncogenes and tumour-suppressors via demethylation or off-target mutations, promoting proliferation and inhibiting apoptosis. **b** Dysregulation of AICDA by UV irradiation, pathogens, and chronic inflammation impacts key genes implicated in EMT and proliferation via its ability to induce off-target mutations or demethylation, contributing to the development of carcinomas. From Rios et al. (2020) [47]

1.7 The role of AICDA in cancer - Oncogenic translocations

Early studies using mouse models showed a strong association between AICDA overexpression and oncogenic translocations, particularly the *c-MYC-IGH* translocation, a hallmark of the highly aggressive Burkitt lymphoma, an HIV-associated cancer [48]. Further analyses found AICDA to be the main driver required to catalyse chromosomal breaks of the *c-MYC* and *IGH* loci resulting in translocation events, which include t(8;14), t(8;22) or (2;8) in BL and t(12;15) in mouse plasmacytomas [49, 50, 51, 52]. Interestingly, Duquette et al (2005) demonstrated that a 'G-loop' structure formed during the transcription of the immunoglobulin S region, and *c-MYC* genes allow specific binding of AICDA to this region. Additionally, Pasqualucci et al (2008) demonstrated that mice engineered to overexpress *BCL6* (a major oncogenic factor of B-cell lymphoma), and in which *AICDA* was silenced, were unable to produce *c-MYC-IGH* translocations. Conversely, mice overexpressing *BCL6* and with endogenous *AICDA* showed an increase of *c-MYC-IGH* translocations of over tenfold [53]. These translocation events associated with AICDA and c-MYC activity and stability have been demonstrated in some cases of DLBCL [52, 54, 55], and this was often associated with poor survival and cellular proliferation [56]. Importantly, Zhang et al (2020) showed that AICDA is an independent risk factor in the progression of

Double Hit Lymphoma (DHL) in vivo, which subsequently results in *c-MYC* translocation via the advancement of the CSR process [57].

1.8 The role of AICDA in cancer – Epigenetic regulation

Over the past decade, reports have drawn a debate that although the activity of AICDA is classically confined to the Ig locus, it has been plausibly identified as one of the factors implicated in the disturbance of methylation at non-Ig genes [58]. This is thought to occur via its fundamental DNA-modifying function, which leads to disruptions of CpG methylation sites, thereby orchestrating epigenetic plasticity [15, 16]. This is manifested as abnormal gene expression patterns, which may include the activation of oncogenes directly and the indirect repression of tumour suppressor genes.

1.9 AICDA as a co-factor and DNA demethylator

A study published by Teater and colleagues, using a murine model, showed that murine lymphoma model mice with ectopic expression of *AICDA* had higher levels of methylation heterogeneity compared to *AICDA*$^{-/-}$ mice, which in turn was associated with increased tumour development and decreased survival rate [61]. In another study by Jiao et al (2019) AICDA was shown to recruit Tet methylcytidine dioxygenase 2 (TET2), a DNA demethylator, to the promoter region of the Fanconi anaemia complementation group A (*FANCA)* gene, leading to its active demethylation and overexpression, linked to DLBCL pathogenesis [62]. Conversely, in research published the following year, the same research group demonstrated that AICDA enhances the recruitment of DNMT1, a DNA methylating enzyme, to the promoter of oncogenic *BCL6* in DLBCL cells leading to its downregulation [63]. Taken together, these two studies demonstrate the dual role that AICDA has in gene regulation via methylation/demethylation.

Most recently, Wang et al (2023) showed, using cell line models, that TET2 cooperates with AICDA to form a complex that directly binds to the promoter region of the proviral insertion site of Moloney murine leukaemia virus *(PIM1)*, a known DLBCL oncogene, inducing its expression. The inverse effect of *PIM1* expression is observed when AICDA recruits DNMT1 to the promoter [64]. Once again, these studies demonstrate the dual role of AICDA in gene regulation via either preventing or promoting promoter methylation.

However, it is important to note that a few studies have dismissed AICDA as a modulator of gene regulation via methylation/demethylation. For instance, a study by Fritz et al (2013) showed using RNA-sequencing and reduced-representation bisulfite sequencing (RRBS) on AICDA^{-}/ $^{-}$, wild type and

AICDA overexpressing activated B cells that the levels of AICDA had little impact on the methylation status of genes within these cells [65]. Nevertheless, there exists enough evidence that AICDA is involved in epigenetic plasticity to warrant further studies. This forms the basis of the current project, the specific aims of which are described below.

1.11 Study Rationale

The oncogenic role of AICDA is well established, as it directly promotes DNA mutations and chromosomal rearrangements due to its DNA-modifying enzymatic function. However, its role in epigenetic regulation, particularly in the modulation of DNA methylation, is not yet fully understood and requires further investigation. Understanding this aspect could provide new insights into its contribution to cancer progression and reveal potential therapeutic targets.

1.12 Aim and Objectives of the Study

1.12.1 Aim

The study aimed to investigate the impact of AICDA on the methylation status of key cancer-driving genes via the use of an AICDA overexpressing cell model.

1.12.2 Objectives

I. Development of an AICDA overexpression cell model.
II. Characterize the model by investigating key hallmarks of cancer, namely proliferation, migration, and response to chemotherapy.
III. Assessment of the impact of AICDA overexpression on the transcriptional regulation of oncogenes, including cancer-driving genes.

Chapter 2

Materials and methods

2.1 The TCGA data analysis

Publicly available patient datasets from The Cancer Genome Atlas (TCGA) for the Colon adenocarcinoma (TCGA-COAD, n=275) and Diffuse Large B Cell (TCGA-DLBC, n=47) were accessed. The gene expression data was also analysed using the Gene Expression Profiling Interactive Analysis (GEPIA) platform to evaluate the expression of the AICDA. The datasets were stratified based on AICDA expression, and the Kaplan-Meier method [66] was used to compare overall survival and disease-free survival rates between the high AICDA and low AICDA patient groups (GEPIA (Gene Expression Profiling Interactive Analysis).

2.2 Cell culture

Various cell lines were used in this research project. The human colorectal cancer (CRC) cell lines DLD-1 and Caco-2 were kindly donated by Professor Stoecklein (Heinrich-Heine University Düsseldorf, Germany) and the DLBCL cell line HBL-1 was a kind gift from Professor Sandeep Dave (Duke University, USA). Additionally, the non-cancerous lymphoblastoid cell line PB-B95-8H was donated by Professor Pankaj Trivedi (La Sapienza University, Italy). The DLD-1, and HBL-1 cells were cultured in complete growth media containing Roswell Park Memorial Institute (RPMI)-1640 medium (ThermoFisher Scientific, USA) supplemented with 10% fetal bovine serum (FBS) (ThermoFisher Scientific, USA), and 1% penicillin-streptomycin solution (Sigma Aldrich, USA). The PB-B95-8H was cultured in complete media supplemented with 20% FBS. Caco-2 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (ThermoFisher Scientific, USA) with 10% FBS and 1% penicillin-streptomycin solution. The cells were incubated at 37°C with 5% CO_2 and 65% humidity, and the media were replaced every 2-3 days. All cell lines were routinely tested for mycoplasma contamination. The process involved culturing the cells in antibiotic-free media for at least two days, then fixation, DNA staining with Hoechst stain solution (Cat; 23491-45-4: Sigma Aldrich, USA), and visualization using an inverted fluorescence microscope (ZEISS Axiovert, Germany). For long-term storage, the cells were preserved in liquid nitrogen. Briefly, the cells were resuspended in freezing media supplemented with 10% dimethyl sulfoxide (DMSO) and 10% FBS in either RPMI or DMEM, depending on the specific conditions of the

cells. The cells were then added into cryovials and transferred into a freezing container (Mr. Frosty, USA) for overnight storage at -80^{0}C before being stored in a liquid nitrogen tank. Only mycoplasma-negative cells were used in the experiments.

2.3 Plasmids construct, and cloning of plasmid vectors

The pEGFP-N3-AICDA and pEGFP-N3-Empty were kindly donated by Professor Bernard Luscher (RWTH Aachen University, Germany). The pcDNA-3.1-3xHA-AICDA and pcDNA-3.1-3xHA-Empty were kindly donated by Professor Reuben Harris (University of Minnesota, USA). The lentiviral vectors, pCDH-CMV-MCS-EF1-copGFP, pMSCV-VSV-G, pRSV.REV and pMDL-gPRRE were kindly donated by Prof Sharon Prince (University of Cape Town, South Africa).

2.3.1 Restriction enzyme digest

The human AICDA open reading frame (ORF) from the pEGFP-N3-AICDA plasmid was cloned into the multiple cloning site (MCS) of the lentiviral vector pCDH-CMV-MSC-EF1-copGFP using the restriction enzymes NheI-HF (G/CTAGC) (New England Biolabs, USA) and EcoRI (G/AATTC) (New England Biolabs, USA). These enzymes cut to leave the ends with single-stranded overhangs. The reactions were prepared as outlined in Tables 1 and 2. The digest reactions for the pEGFP-N3-AICDA plasmid were incubated at 37°C for 4 hours, while the reaction for the pCDH-CMV-MSC-EF1-copGFP required a sequential double digest due to the close proximity of the NheI-HF and EcoRI restriction sites within the MCS (see Appendix B1 for the plasmid backbone map). The restriction sites of the two enzymes are separated by only 6 bases, which hinders the activity of the NheI-HF enzyme, as it does not perform optimally when the sites are flanked by such a small number of bases. Initially, the mixture (with EcoRI excluded) was incubated at 37°C for 3 hours; then EcoRI was added, and incubation continued for an additional 3 hours.

Table 1: Double digest of pEGFP-N3-AICDA to remove AICDA insert.

Components	Volume per reaction (µl)
DNA plasmid (5 ug)	*
NheI-HF	1
EcoR I	1
10x buffer E	2
Nuclease free water	Up to 20

** Indicates variable volume*

Table 2: Sequential double digest of pCDH-CMV-MSC-EF1-copGFP for AID insertion.

Components	Volume per reaction (µl)
DNA plasmid (5 ug)	*
NheI-HF	1
EcoR I	1 (added after 3 hours)
10x buffer E	2
Nuclease free water	Up to 20

** Indicates variable volume*

2.4 Agarose gel electrophoresis

For this experiment, a 1.5% agarose gel was prepared (Appendix A). After the incubation, the digest reaction mixtures were mixed with a DNA-loading dye and loaded into wells of the gel. A 100-base pair DNA ladder (cat; 07-11-00060: Soli BioDyne, Estonia) was loaded and used as the molecular weight size reference. The gel was electrophoresed at 80 volts for 50 minutes. The gel was visualised under UV light to confirm successful digestion.

2.5 Extraction and purification

After separating the pEGFP-N3-AICDA digestion mixture on the agarose gel, two distinct bands were observed: one corresponding to the *AICDA* insert and the other representing the backbone of the DNA plasmid that was carrying the gene. The band of AICDA insert was carefully excised from the agarose gel using a clean, sharp scalpel under the UV in the dark room. The MinElute Gel Extraction Kit was used to purify the AICDA inserts following the manufacturer's instructions (Qiagen, Germany). To purify the linearised pCDH-CMV-MSC-EF1-copGFP digestion mixture, the GeneJET PCR purification was performed according to the manufacturer's instructions (ThermoFisher Scientific, USA). The extracted DNA of interest was quantified using a Nanodrop® ND-100 spectrophotometer (ThermoFisher Scientific, USA), which determines the concentration of nucleotides as well as the level of contamination using the 260/280 and 260/230 ratios. Generally, nucleotides are detected at 260 nm while protein and organic (phenol) contamination can be detected at 280 nm and 230 nm, respectively.

2.6 DNA Ligation

AICDA inserts were ligated within the linearized pCDH-CMV-MSC-EF1-copGFP backbone vector using T4 ligase DNA ligation following the manufacturer's instructions (Promega, USA). A molar ratio of 4:1

(insert to vector) was used in the ligation and the reaction was set up using the formula and Table 3 below. The reaction was incubated overnight at 4°C.

$$\frac{(\text{ng of vector x kb size of insert})}{(\text{kb size of vector})} \text{ x molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng of inserts}$$

Table 3: The following components were mixed into a sterile tube

Components	Volume per reaction (µl)
AICDA insert (31 ng)	*
Backbone vector (100 ng)	*
Ligase 10x buffer	1
T4 DNA ligase	1
Nuclease free water	Up 10

** Indicates variable volume*

2.7 Bacterial transformation and colony formation

NEB 10-beta *E. coli* competent cells (New England Biolabs, USA) were used for high transformation efficiency due to the size of the backbone vector following the manufacturer's instructions with some modifications. Briefly, the DNA ligation mixture was added to the NEB 10-beta competent cells mixture and incubated on ice for 45 minutes. This step allows the incorporation of the ligated DNA into the cells. Thereafter, the mixture was heat shocked at 42°C for 30 seconds and placed back on the ice for 5 minutes. Next, 950 µl of the RT NEB 10-beta outgrowth medium (New England Biolabs, USA) was added to the mixture and placed at 37°C for 2 hours and 30 minutes with gentle agitation. A plate containing Luria-Bertani (LB) agar (BDH Chemicals Ltd Poole, United Kingdom) supplemented with ampicillin (100 ug/ml) (Glentham Life Sciences, United Kingdom) was prepared (Appendix A). The plate was incubated to warm up at 37°C. Thereafter, the mixture was spread onto the selection plate and incubated overnight at 37°C to allow the formation of the colonies, which are competent cells that took the backbone vector potential ligated with AICDA inserts. The following day colonies were selected and inoculated into a 5 ml LB medium supplemented with ampicillin (100 ug/ml). Briefly, the colonies were gently picked up with the tip of the pipette and added into the LB medium and thereafter incubated overnight at 37°C with gentle agitation.

2.8 DNA Plasmid Extraction

Plasmid DNA was extracted from overnight growth of the LB medium inoculated with colonies, using the PureYield™ plasmid Miniprep system (Promega, USA) following the manufacturer's instructions. Thereafter, the plasmid was quantified using a Nanodrop® ND-100 spectrophotometer (ThermoFisher Scientific, USA).

2.9 Sanger sequencing

To further verify successful cloning the same restriction enzymes were used to re-digest the extracted plasmid DNA with the same conditions as used before. Thereafter, the digestion mixture was run on the agarose gel. The clones that released the size of the AICDA insert were sent to Inqaba Biotec (South Africa) for sequencing, and the forward and reverse primers for cytomegalovirus (CMV) were used as provided by the company.

2.10 Transfection

The transient transfections were performed using Xtreme-Gene HP DNA transfection reagent (Roche, USA) following the manufacturer's instructions. Briefly, for each experiment cells ranging from $2X10^3$ to $1X10^5$ cells/wells (depending on the type of experiment) were counted using a hemocytometer and then seeded in the 6-well plate, 12-well plate or 6 cm dish to achieve 70-80% confluence on the day of transfection. The HEK293 cells were transfected with 500 ng of either the pCDH-CMV-MSC-EF1-copEFP-AICDA or the pCDH-CMV-MSC-EF1-copEFP-Empty vector. The plasmids were each mixed with 197 µl of serum-free media and 3 µl Xtreme-Gene HP, incubated for 15 minutes at room temperature (RT), and then added to cells in a dropwise manner. The cells were subsequently incubated in a cell culture incubator for 24 or 48 hours before proceeding with extracting the protein for western blotting.

2.11 Lentiviral production and transduction

2.11.1 Lentiviral production

The third-generation lentiviral transduction system was used in this study to overexpress AICDA in a non-cancerous lymphoblastoid cell line (PB-B95-8H) and a DLBCL cell line (HBL-1). This protocol allows the integration of genes into the genome, offering long-term and stable gene expression compared to the transient transfection system, which provides short-term expression. This system is useful for cancer B cells, which are challenging to transfect using traditional methods. This type of system uses three different plasmids, pMSCV-VSV-G, pRSV.REV and pMDL-gPRRE to generate viral particles (with

plasmids containing genes of interest) to be used in transduction experiments. The experiment was conducted in the BCL2 hood. Briefly, HEK293FT cells were plated at a density of 1.5 x 10^6 cells/wells in 60mm dishes in complete growth media containing 500 ug/ml Geneticin (G-418) antibiotic (Glentham Life Sciences, United Kingdom). Before plating the cells for transfection, the dishes were coated with 1 mg/ml of poly-l-lysine and incubated at 37°C for 1 hour to prevent HEK293FT cells from lifting after adding lipofectamine. Thereafter the cells were plated and placed in the tissue culture incubator overnight. HEK293FT cells were used due to their high transfection efficiency, stable SV40 large T antigen expression (which is necessary for the replication of lentiviral vectors in the cells), and rapid growth (which allows for faster production of lentiviral vector) [67].

The next day, the Lipofectamine LTX reagent (ThermoFisher Scientific, USA) was used for transfection. The three plasmids were added together with the newly generated pCDH-CMV-MSC-EF1-copEFP-AICDA (the cells that express pCDH-CMV-MSC-EF1-copEFP-Empty vector were already generated and available in the laboratory) in 500 µl serum-free media. The reactions were set up according to **Table 4**. Then, plus reagent (15µl) was added to the diluted DNA mixtures (15 µg) and incubated for 15 minutes at RT. Thereafter, the Lipofectamine LTX was added to each tube containing DNA + Plus reagent + Serum-free medium dilution and the reactions were incubated for 30 minutes at RT. During the incubation, the medium from the cells was removed, the cells washed with 1 ml of complete medium (to remove G-418 antibiotic) and replaced with 5 ml fresh medium. After 30 minutes of incubation, the transfection mixture was added dropwise to HEK293FT cells and incubated at 37°C with 5% CO_2 and 65% humidity.

After the incubation time of 24 hours, the medium was removed and replaced with fresh complete media, to enhance virus production. The media containing lentivirus particles was collected at the 48-hour and 72-hour time points. At this stage, 5 ml lentivirus medium was harvested from the cell culture into a 50 ml tube (totaling 10 ml lentivirus medium) and centrifuged for 5 minutes at 2000 rpm to remove cell debris. The supernatant medium was filtered through a 0.45 µm filter into a new 50 ml tube and 2 ml aliquots were made and stored at - 80°C (the tubes were covered in the surgical bag).

Table 4: Lentiviral plasmid mixture per dish

DNA plasmids	Volume per reaction (µl)
pCDH-CMV-MSC-EF1-copEFP-AICDA (6 µg)	*
pMSCV-VSV-G (3 µg)	*

pRSV.REV (3 µg)	*****
pRSV.REV (3 µg)	*****
Free-serum medium	**500**

* *Indicate variable values as dependent on the concentration of the plasmids*

2.11.2 Lentiviral transduction

The next step was lentiviral transduction, where the collected lentiviral medium was introduced to PB-B95-8H and HBL-1 cell lines. Cells ($3X10^5$ cells/ml) were plated in a 6-well plate with 1 ml complete media (including an un-transduced control well). A tube of 2 ml viral media (2 ml complete media in the control well) and 3 µl Polybrene stock solution (10 mg/ml) were mixed gently. The mixture was added to the cells gently by pipetting. The plate was wrapped with parafilm and centrifuged at 3000 g for 90 minutes at RT. The cells were incubated overnight at 37°C with 5% CO_2 and 65% humidity. The following day, the cells were checked for polybrene toxicity by inspecting cell morphology under a light microscope. The cells were then expanded and checked regularly. To confirm successful transduction, the cells were visualised under the inverted fluorescence microscope to detect the presence of GFP expression. The GFP-expression transduced cells were sorted using the FACS Aria (BD, United Kingdom) flow cytometer, located at the Institute of Infectious Disease and Molecular Medicine (IDM) Flow Cytometry Core Facility, with the assistance of Mr Ronnie Dreyer. Prior to sorting, the cells were washed with prewarmed HBSS at least three times, which acts as a balanced salt buffer, to maintain the proper pH and osmotic balance for cells when transporting them for sorting. The sorted cells were collected in 20% FBS media and transported back to the tissue culture laboratory. The cells were gently spun down and resuspended in 50-100 µl (depending on the percentage of sorted cells) of the 20% FBS media and plated in a 96-well plate. The cells were monitored daily and fed as needed. After expansion, GFP expression was confirmed once again using fluorescence microscopy, and the cells were frozen down.

2.12 Generation of stable ectopic expression of AICDA in Colorectal cancer cell lines

To generate a stable CRC cell line that expresses the AICDA gene, the cells were transfected with pCDNA-3.1-3xHA-AICDA and pCDNA-3.1-3xHA-Empty using XtremeGene reagent (ThermoFisher Scientific). The backbone of the pCDNA plasmid contained a gene conferring Neomycin (Neo) resistance, an analog of aminoglycoside antibiotic Geneticin (G-418) (Roche, USA), and as such this antibiotic was used for the selection of the clones that stably expressed the AICDA gene. Firstly, the

lowest concentration of G-418 antibiotic needed to kill the untransfected cells for 14 days was established. Caco-2 and DLD-1 cells were plated in 24-well plates at $2X10^4$ cells/well density. A concentration range of 0-1000 µg/ml G-418 antibiotic was added to each well at various time points for 14 days. The selection medium was frequently changed every 3 days to maintain the selective pressure of the antibiotic. The cells were also visualized under the microscope to observe morphological changes, and the Dino-eye eyepieces camera (Dino-Lite, Netherlands) was used to capture the images.

Secondly, AICDA was overexpressed in these CRC cells via transient transfection. Caco-2 and DLD-1 cells were plated at a density of $2x10^5$ cells/well in a 6-well plate and transfected with either pCDNA-3.1-3xHA-AICDA or pCDNA-3.1-3xHA-Empty using XtremeGene reagent (ThermoFisher Scientific, USA). After 48 hours of transfection, a selection media containing complete media and the concentration of G-418 determined by the standard kill curve was added to the cells and incubated at 37°C with 5% CO2 and 65% humidity for 14 days. The selection media were replaced after every 3 days to maintain selective pressure.

Lastly, the single-resistant clone cells (which took up the plasmid vectors) were selected. After the single resistant clones were formed, the media were removed, and approximately 3 ml of 1XPBS was added to wash the clones. Pipette tips were cut to create cloning rings, and rings were coated with sterile Vaseline and gently placed on each single resistant clone. Then, the 1XPBS within the cloning ring were removed and 50 µl of trypsin (Appendix A) was gently added to collect the clones. The cell clones were collected after detaching and plated in the 24-well plate. The cells were expanded, and Mycoplasma testing was performed to assess any potential contamination.

2.13 Western blotting analysis

2.13.1 Protein extraction using 2X boiling blue buffer

Boiling blue was used as one of the methods to extract protein from the cells. Briefly, media was removed, and the cells were washed twice with cold 1X PBS. In a 6-well plate, approximately 100 µl of boiling 2X blue Buffer (depending on the viscosity of the cells) (Appendix A) was added, and the cells scraped and transferred into a 1.5 ml tube. Thereafter, the solution was incubated at 95 °C for 10 minutes, pulse spun, and either used directly on the SDS-PAGE or cooled to room temperature and stored at -20 °C.

2.13.2 Protein extraction using RIPA buffer

Radio-immunoprecipitation assay (RIPA) buffer was also used for protein extraction. This assay is effective as it uses three non-ionic and ionic detergents, which lyse and permeabilise the cells. The RIPA solution was prepared by mixing RIPA buffer with a 1X cOmplete™, Mini, EDTA-free Protease inhibitor Cocktail (Roche, Germany) (Appendix A), which prevents the degradation of the extracted protein from proteases in the cell lysates. The plated cells were collected into 15 ml tubes, pelleted, and washed three times with ice-cold sterile 1X PBS. Approximately 100 µl of RIPA solution was added to the cells, transferred into a 1.5 ml tube, and stored at -80°C overnight for optimal lysis. The next day, the cell lysates were thawed on ice and centrifuged at 4°C (12 00 rpm) for 20 minutes to remove cell debris, and the supernatant, containing protein, was then collected, aliquoted, and stored at -80°C.

2.13.3 Protein quantification

The Bicinchoninic Acid (BCA) reagent was used to determine the concentrations of extracted protein, using the Pierce™ BCA assay kit (ThermoFisher Scientific, USA). The principle behind this assay is that it uses a sensitive and selective colorimetric system. Generally, it relies on the reduction of Cu^{2+} to Cu^{1+} by protein in an alkaline medium. The Cu^{1+} ions then react with BCA to form a complex that absorbs light at 562 nm [68]. A standard curve was generated using bovine serum albumin (BSA) at different concentrations (2000 µg/ml, 1000 µg/ml, 500 µg/ml, 125 µg/ml and 0 µg/ml). Briefly, a working reagent (WR) was made by mixing Reagent A and B in a ratio of 50:1, respectively. The protein samples were diluted with RIPA buffer (with ratios of 1:2, 1:4, and 1:6). Then, 10 µl of the diluted samples were added into a 96-well plate in duplicate, as well as different BSA standards, followed by the addition of 200 µl of WR per well. The plate was incubated for 30 minutes at 37°C, followed by the measurement of colour change using Glo-Max-Multi+ multi-plate reader (Promega, USA) at a wavelength of 560 nm. Microsoft Excel software (Microsoft Office Professional Plus 2019, version 1808) was used to plot the BSA standard curve, and then the protein sample concentration was calculated by extrapolating from the graph.

2.13.4 SDS-PAGE and western blotting

The extracted protein samples were then separated by molecular weight using SDS-PAGE (sodium dodecyl sulphate-polyacrylamide gel electrophoresis), on resolving gels (8%, 10%, or 15%, Appendix A) and a stacking gel (5%, Appendix A), using the Mini-PROTEAN 3 casting apparatus (Bio-Rad, California, USA). All the experiment components were prepared as per Table 5 below and denatured at 95°C for 10 mins. The proteintech® pre-stained 1 kb Protein Ladder (Cat; #Pl0001) was used as the

molecular weight marker. The gel was electrophoresed at 100 volts for 2 to 3 hours (depending on the size of the protein) in 1x Running buffer (Appendix A).

Table 5: Components of the protein sample preparation for SDS-PAGE.

Components	Volume (µl)
Protein sample (15 µg)	*
100 mM Dithiothreitol (DTT)	1
5x SDS loading dye	6
RIPA buffer	Up to 30

** Indicates variable volume*

After electrophoresis, the separated proteins in the gel (resolving portion only) were transferred to a nitrocellulose membrane. The gel was assembled with the nitrocellulose membrane (Advansta, USA), filter paper, and sponge pads rinsed under ice-cold 1x Transfer buffer (Appendix A) in a gel holder cassette. The cassette was inserted into the Bio-Rad Mini-PROTEAN 3 transblotting apparatus (Bio-Rad, California, USA), and covered with ice-cold 1x Transfer buffer. It was then connected to a power pack to initiate the transfer process at 100 volts for 75 minutes.

2.13.5 Antibody incubation and protein detection

Next, the nitrocellulose membrane containing the transferred protein was washed with either 1X PBS Tween 20 or 1X TBS-Tween 20 buffers (Appendix A). After washing, Ponceaus S solution (Appendix A) was added to the membranes, to verify successful transfer. The stain was washed off with either 1X PBS-Tween 20 or 1X TBS-Tween 20 washing buffers. The membrane was then incubated in either 5% fat-free milk or 5% BSA blocking buffer (Appendix A) for 2 hours at room temperature, with gentle agitation. This was followed by incubation with primary antibody diluted in specific blocking buffers overnight at 4^{o}C, with gentle agitation. Unless otherwise stated, the following primary antibodies were used at a 1:1000 dilution: AID (#39-2500) (Invitrogen, USA). HA-tag (#3724: 1:2000) (Cell Signalling Technology, USA). β-actin (sc-47778; 1:10000), Zeb 1 (sc-81428; 1:200), Zeb2/SIP1 (sc-271984; 1:2000), c-Myc (sc-764), PARP-1 (sc-8007; 1:2000), and Caspase-3 (sc-56053) (Santa Crus Biotechnology Inc, Texas, USA).

After overnight incubation, the membrane was washed with one of the washing buffers, depending on the antibody's conditions (2 x 5 minutes and 2 x 10 minutes) with gentle agitation at room temperature. This was followed by incubation with either horseradish peroxidase (HRP)-conjugated

goat anti-rabbit (1:5000; Cell Signalling Technology, USA) or horse anti-mouse (1:5000; Cell Signaling Technology, USA), for 2 hours at room temperature. After incubation, the membrane was washed as before and visualised through enhanced chemiluminescence using Clarity™ Western ECL substrate as per the manufacturer's instructions (Bio-Rad, California, USA). The two reagents in the kit were mixed in a 1:1 ratio and the membranes were incubated in that solution for 5 minutes. The membrane was then exposed to X-ray film and the resulting chemiluminescent signal was captured by developing and fixing the film. The signal intensity of the bands was analysed using Fiji ImageJ software (Version 2.9.0), normalised to β-actin, and relative to control.

2.13.6 Membrane stripping

After visualising the proteins, the membrane was rinsed with one of the washing buffers. The stripping buffer (Appendix A) was pre-heated at 50°C. The stripping buffer was added to the membrane in a small container and incubated at 50°C for minutes, with gentle agitation after every 10 minutes. Thereafter the membrane was washed with washing buffer (2 X 5 minutes and 2 X 10 minutes). The membrane was then blocked again and reused to detect the protein of interest (usually the internal loading control β-actin).

2.14 Quantitative Real-Time PCR

2.14.1 Primer design

To design the primer sets, the complete sequences of genes were obtained from the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/). The obtained sequences of the genes were imported into the online software Primer3Plus program (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/) to design both the forward and reverse primers. The design was based on specific parameters, including primer length (18-25 bp), melting temperature (48-55^{0}C), and GC content (40-60%). Oligoanalyzer (Integrated DNA Technologies, Iowa, USA) was used to analyse the possibility of the primers forming stable secondary structures. It was ensured that the ΔG for potential secondary structures for the primer pair was kept lower than 3 kcal/mole. Subsequently, the chosen primers were examined for complementarity to the specific genes using the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The primer information is shown in **Table 6.**

Table 6: Information on the selected primers

Primer	Sequence	Melting temperature (Tm) (^{0}C)
ZEB1_F	5'-GCCTGAAATCCTCTCTGAATG-3'	59.8
ZEB1_R	5'-CACCTCTTGTCAAAC-3'	58.6
PIM1_F	5'-CTGGGGAGAGCTGCCTAATG-3'	57.6
PIM1_R	5'-GCTCCCCTTTCCGTGATGAA-3'	57.4
FANCA_F	5'-GAGGTTCACTGGCAGAGAGT-3'	56.5
FANCA_R	5'-TATGACAGGAACGCAGAGGG-3'	56.6
c-MYC_F	5'-CTGAGACAGATCAGCAACAACC-3'	55.7
c-MYC_R	5'-TTGTGTGTTCGCCTCTTGAC-3'	55.6
CD274_F	5'-GGAGATTAGATCCTGAGGAAAACCA-3'	56.0
CD274_R	5'-AACGGAAGATGAATGTCAGTGCTA-3'	56.3
DNMT1_F	5'-CGTGGTGGTGGATGACAAGAA-3'	57.4
DNMT1_R	5'-CGTTGTAGGAGATCTCCAGTGC-3'	57.1
DNMT3A_F	5'-TATYGATGAGCGCACAAGAGAGC-3'	57.4
DNMT3A_R	5'-GGGTGTTCCAGGGTAACATTCAG-3'	57.7

2.14.2 RNA Isolation

Prior to RNA extraction, all plasticware to be used was treated with water containing 0.1% diethylpyrocarbonate (DEPC), an effective nuclease inhibitor, and autoclaved. The dedicated RNA bench was also wiped thoroughly with 10% diluted bleach, followed by 70% ethanol. To isolate RNA from cultured cells, the Roche High Pure RNA isolation kit (Roche, USA) was used. The manufacturer's instructions were followed to isolate high pure RNA from treated or untreated cells. The DLD-1 and Caco-2 cells were trypsinised and counted. The cells were then pelleted and washed with ice-cold 1X PBS (Appendix A). The cells were then resuspended in 200 µl ice-cold 1xPBS, per the manufacturer's instructions. Briefly, 400 µl of Lysis buffer was added to resuspended cells and vortexed for 15 seconds. The lysates were transferred to High Pure Filter tubes and centrifuged at 8000 x g for 15 seconds. After discarding the flow-through, DNase I solution was added, and the samples were incubated at room temperature for 15 minutes. The filter was washed once with 500 µl of Wash Buffer I and twice with 500 µl of Wash Buffer II, followed by a final wash with 200 µl of buffer at maximum speed for 2 minutes.

The RNA product was eluted and quantified using a Nanodrop® ND-100 spectrophotometer (ThermoFisher Scientific, USA).

2.14.3 Gel electrophoresis for RNA

To evaluate the integrity of the eluted RNA, gel electrophoresis was performed. The gel tray, comb, and chamber were soaked in DEPC-treated water with 10% sodium hypochlorite overnight to prevent RNA degradation. 1% or 1.5% agarose gel (depending on the size of the gene of interest) was prepared using 1X TBE buffer with Ethidium bromide. The RNA samples (200 ng) were prepared by mixing 2X RNA loading buffer (Appendix A) and topping up with DEPC-treated water to make 20 µl of total volume in clean DEPC-treated 1.5 ml centrifuge tubes. The RNA was then denatured by incubating the samples at 55°C for 5 minutes. Afterward, the samples were briefly centrifuged and loaded into the wells of the gel. The gel was electrophoresed at 100 volts for 45 minutes and visualized using a UV transilluminator (Uvitec, United Kingdom).

2.14.4 cDNA synthesis

The iScript™ cDNA Synthesis Kit (Bio-Rad, USA) was used for reverse transcription. The reverse transcription technique is used for synthesizing complementary DNA (cDNA) from RNA using reverse transcriptase. The manufacturer's instructions were followed to set up each reaction of the samples as indicated in **Table 7**. The tubes were mixed and centrifuged briefly, after which they were loaded onto the PCR machine (Bio-Rad, USA), and the cycling conditions used are shown in **Table 8**. The product of cDNA samples was then sored at -20°C.

Table 7: Reagents for cDNA synthesis

Reagents	Volume (µl)
5X iScriptTM reaction mix	**4**
IScript Reverse Transcriptase	**1**
Nuclease-free water	*****
RNA template (1µg)	*****
Total volume	**20µl**

** Indicates variable volume*

Table 8: Cycling conditions for reverse transcription

Step type	Time (minutes)	Temperature (^{O}C)
Priming	**5**	**16**
Reverse transcription	**30**	**42**
RT inactivation	**5**	**85**
Hold	**∞**	**4**

2.14.5 Quantitative real-time PCR

Quantitative real-time PCR (qPCR) is the laboratory technique used to amplify and quantify specific DNA sequences. The KAPA SYBR® FAST qPCR Master Mix Kit (Merck, USA) was used following the manufacturer's instructions, and the reaction for each sample was set up according to **Table 9**. Briefly, the master mixes (excluding cDNA, accounting for 3 technical repeats for each cDNA sample and a no-template control) were set up and aliquoted into each PCR tube, and 1µl of cDNA template was added to the labelled tubes. The PCR tubes were mixed well and loaded onto the Rotor-Gene Q (Qiagen, Hilden, Germany) machine. The cycling parameters (three steps with MELT) were set up as shown in **Table 10**. The data was exported and analysed using the $2^{-\Delta\Delta Ct}$ method.

Table 9: Components of KAPA SYBR® FAST qPCR Master Mix

Reagents	Volume (µl)
PCR-grade water	**8.2**
2X KAPA SYBR® FAST qPCR Master Mix	**10**
10 µM Forward primer	**0.4**
10 µM Reverse primer	**0.4**
Template cDNA product	**1**

Table 10: The conditions of the KAPA SYBR® FAST qPCR Master Mix

Program Name	Target (°C)	Acquisition Mode	Duration	Cycles
Pre-Incubation	**95**	**None**	**3 minutes**	**Hold 1**
	95	**None**	**10 seconds**	

Amplification	**Primer dependent**	**None**	**20 seconds**	**40**
	72	**Single**	**1 second**	
Melting Curve	**95**	**None**	**5 seconds**	
	60	**None**	**1 minute**	
	97	**Continuous**	**5 – 1 seconds**	
Cooling	**40**	**None**	**10 seconds**	**Hold 2**

2.15 Cell treatment

The fluorouracil (5-FU) drug was kindly donated by Professor Sharon Prince (University of Cape Town, South Africa). The drug was diluted in growth media to achieve the desired concentration(s). The cells were counted and plated the day before to allow them to grow to 60-70% confluency. The cells were treated with different concentrations of 5-FU drug (depending on the experiment). To evaluate the effects of inflammation associated agents on AICDA expression, parental Caco-2 and DLD-1 cells were also treated with *Escherichia coli* Lipopolysaccharide (LPS). Briefly, the cells were seeded at a density of $1X10^5$ cells/wells in a 6-well plate and treated with LPS the following day for 24 hours in culture.

2.16 Cell proliferation assays

To evaluate the impact of AICDA expression on the proliferation of colon adenocarcinoma cells, the growth rate was measured by growth curve analysis. Specifically, DLD-1 and Caco-2 cells ($2x10^4$/well) were plated in a 12-well plate in triplicate. The cells were collected by trypsinisation and counted using a haemocytometer every 2 days for 6 days. To evaluate the anti-tumour effects of the 5-FU drug on CRC cell lines, DLD-1 and Caco-2 cells were treated with 79.86 µg/ml and 72.5 µg/ml of the drugs, respectively for 4 days in culture.

2.17 WST-1 Cell proliferation assay

The WST-1 (2-(-4-iodophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H-tetrazolium) assay (Roche, Germany) was also used to evaluate the effect of AICDA expression on cell proliferation. This assay is based on the action of cellular mitochondrial dehydrogenase, which cleaves the tetrazolium salt WST-1 to produce formazan. This reaction results in a colorimetric change (from a slightly red colour to dark red colour) that can be measured using an absorbance reader [69]. The principle behind the assay is

that the more viable the cells, the higher the activity of the mitochondrial enzymes, and in turn the greater the amount of dye formed. Briefly, Caco-2 and DLD-1 cells ($4X10^3$/wells) were plated in a 96-well plate in triplicate and incubated at 37°C with 5% CO_2 and 65% humidity. On the day of reading the plate, 10 µl of WST-1 reagent was added to the cells, followed by 2 hours of incubation. The absorbance was read using a spectrophotometer (Glo-Max-Multi+ multi-plate reader, Promega, USA) at 450nm every 24 hours for a period of 48 hours and 72 hours. The data were analysed and plotted on Microsoft Excel software (Microsoft Office Professional Plus 2019, version 1808).

2.18 Wounding healing assays

To evaluate the impact of AICDA expression on CRC cell migration, Wound healing was performed to measure the percentage of migration rate. Cells ($5X10^5$/wells) were plated in a 12-well plate to grow until forming a 100% confluent monolayer. A vertical wound was created on the cell monolayer using a 200 µl pipette tip. After the wound was created the cells were treated with 5 µg/ml of Mitomycin C stock solution (10 mg/ml) (Glentham Life Sciences, United Kingdom) to inhibit cell proliferation. To determine the anti-migratory effect of the 5-FU drug, cells were treated with the drug plus Mitomycin C immediately after the wound was created. The wound areas were monitored, and images were taken at 3-hour intervals using a Dino-eye eyepieces camera. The wound area and closure rate was analysed using ImageJ software version 1.8.0 (National Institute of Health, USA). The data were quantified using GraphPad Prism version 8.0 Software (USA).

2.19 Cell Viability

To determine the half-maximal inhibitory concentration (IC_{50}) of the 5-FU drug on Caco-2 and DLD-1, the WST-1 (2-(-4-iodophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H-tetrazolium) assay (Roche, Germany) was used as described above. Briefly, Caco-2 and DLD-1 cells were plated in triplicate at a density of $6X10^3$cells/wells in a 96-well plate to grow until they reached 60-70% confluence. The cells were then treated with a concentration range of 5-FU drug (0-500 µg/ml) and incubated for 72 hours at 37°C with 5% CO_2 and 65% humidity. After the incubation, 10µl of WST-1 reagent was added to the cells, followed by another incubation for 2 hours. The absorbance was read using a spectrophotometer (Glo-Max-Multi+ multi-plate reader, Promega, USA) at 450nm. The data were analysed and plotted using GraphPad Prism version 8.0 Software (GraphPad Prism software, USA).

2.20 Cell cycle analysis

To evaluate the effect of the 5-FU drug on the phases of the cell cycle, flow cytometry was used to measure the changes. DLD-1 and Caco-2 cells (5 X 10^5 cells /wells) were plated in a 6 cm dish well plate in duplicates and treated with ½ IC50 concentrations of 5-FU, followed by incubation for 48 hours. The FxCycleTM PI/RNase staining solution (Invitrogen, USA) was used to prepare the cells for cell cycle profiling, according to the instructions from manufacturers. Briefly, cell samples were collected into a 15 ml tube (Falcon™), pelleted, and washed with 1XPBS. Thereafter, the cells were fixed with ice-cold 70% methanol and incubated on ice for 30 minutes. The cells were then pelleted again, washed with 1 ml 1XPBS, resuspended in 500 µl of PI/RNase staining solution, and incubated for 30 minutes. The samples were analysed using the BD FACSymphony™ A5 flow cytometer (BD Biosciences, USA) with a 488 nm coherent laser. Analysis was done using FlowJo software. The data was analysed using FlowJo software (Version 10, USA).

2.21 Statical analysis

Statistical analysis was conducted using GraphPad Prism version 8.0 (USA) and Microsoft Excel (Microsoft Office Professional Plus 2019, version 1808). Both Student's t-tests and One-way ANOVA were applied, with statistical significance determined at a threshold of *$p<0.05$. In instances where statistical analysis was not performed, representative outcomes were presented to reflect the consistency of the general trends observed in the experimental data.

2.22 Ethics

No ethical considerations were needed for animal or human samples in this study. The use of established cell lines was approved by the UCT Health Sciences Human Research Ethics Committee (Reference: C001/2025).

Chapter 3

Results

3.1 Brief background

The DNA-modifying enzyme AICDA is widely recognised as a B cell specific factor essential in the antibody producing function of these cells [70]. Its overexpression disrupts the integrity of the genome and has been linked to the development of cancer, particularly B-cell-derived non-Hodgkin lymphomas and inflammation-associated solid tumours which have a higher prevalence in HIV endemic regions [71, 72]. More recently, the oncogenic function of AICDA has been expanded to include modulation of the epigenetic landscape, through the enzyme's ability to alter CpG sites, as a consequence of its deamination function. However, the direct involvement of AICDA in the alteration of gene expression by disrupting DNA methylation remains uncertain due to a lack of scientific evidence. The current study sought to develop an AICDA-overexpressing cell model and use it to explore the impact of AICDA on the regulation of key cancer-associated genes with the aim of identifying candidate genes for future analyses.

3.2. AICDA Expression in DLBCL is associated with significantly poorer patient survival.

Abnormally high expression of AICDA is a common feature of the highly aggressive Diffuse Large B Cell lymphoma (DLBCL) [73]. Bioinformatics analyses using publicly available DLBCL patient data from The Cancer Genome Atlas database (TCGA) via the Gene Expression Profiling Interactive Analysis (GEPIA) web tool confirmed that AICDA expression levels were significantly higher in DLBC tissues (n=47) compared to the normal tissue samples (n=337) (**Figure 3.1A**). Furthermore, Kaplan-Meier survival plots showed that high AICDA level in DLBC is correlated with poorer overall survival (OS) (**Figure 3.1B**) and reduced disease-free survival (DFS) (**Figure 3.1C**) relative to patients having lower AICDA levels.

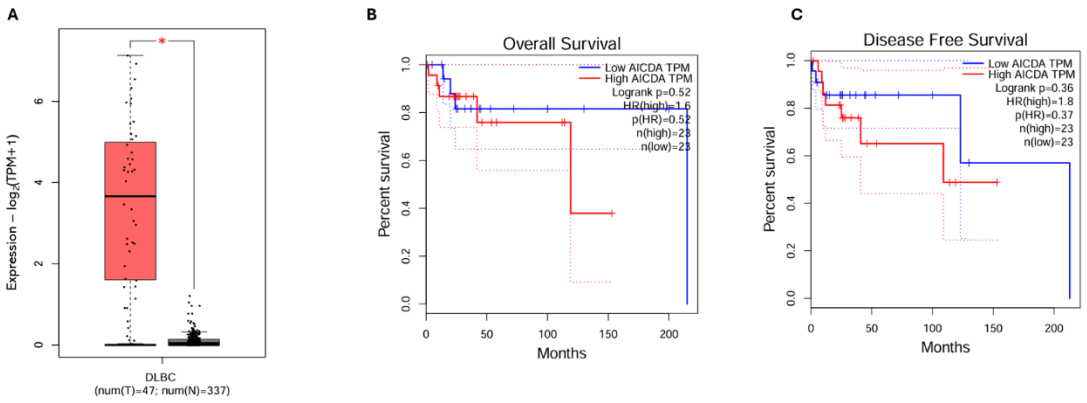


Figure 3.1 The expression of AICDA and patient survival in the TCGA-DLBC database. (A) Boxplots of AICDA expression level in DLBC tissues (n=47) and normal tissue samples (n=337) * denotes a p-value of less than 0.05. **(B, C)** The Kaplan-Meier survival for overall survival and Disease-free survival of the DLBCL cohort using the TCGA classification database. The Logrank p-value are given as inserts.

3.3 Development of an AICDA overexpressing B-cell model.

3.3.1. Cloning of human *AICDA* gene into a lentivector backbone

Two cell lines were chosen, namely a non-malignant EBV-immortalized lymphoblastoid cell line (PB-95H-8H), and a DLBCL cell line (HBL-1). Lentiviral transduction is the most efficient method for introducing foreign DNA into human lymphocytes and therefore this was the method of choice [67]. Previous work in the laboratory involved screening AICDA protein expression in a panel of DLBCL cell lines. Among these, HBL-1 showed lower AICDA expression (Appendix B; Figure B6) and was therefore selected for this study. HBL-1 and PB-95H-8H cells expressing the empty control vector were already available in the laboratory (from previous studies).

The first step was to clone the human *AICDA* mRNA within the open reading frame (ORF) of the pCDH-CMV-MCS-EF1-copGFP lentivector. Among other markers, this construct allows for the expression of Green Fluorescent Protein (GFP) independent of the cloned gene (**Figure 3.2A**) and serves as a marker to measure the efficiency of transduction as well as for capturing transduced cells using Fluorescent-Activated Cell Sorting (FACS). The human *AICDA* ORF was excised from the pEGFP-N3-AICDA plasmid using the restriction enzymes (RE) digest (EcoR I and NheI-HF) and ligated within the multiple cloning site (MCS) of the lentivector linearized using the same enzymes. This therefore involved directional

cloning, and a schematic of the cloning strategy and recombination plasmid amplification is shown in **Figure 3.2B** and the detailed method described under section 2.3.1.

Following verification of successful clones using REs ([**Figure 3.2C**] and Sanger sequencing (Appendix B, Figure B5), ectopic expression of AICDA protein was verified via western blot analysis using an antibody specific to AICDA of total proteins extracted from HEK293 human embryonic kidney cells transfected with the newly developed AICDA-lentivector, and the empty vector as a negative control (section 2.10) showing successful expression of AICDA in cells transfected with the recombinant lentivector, and no AICDA expression in cells transfected with empty vector (**Figure 3.2D**). Together, the data showed that the cloning of AICDA into a lentiviral backbone vector was successful.

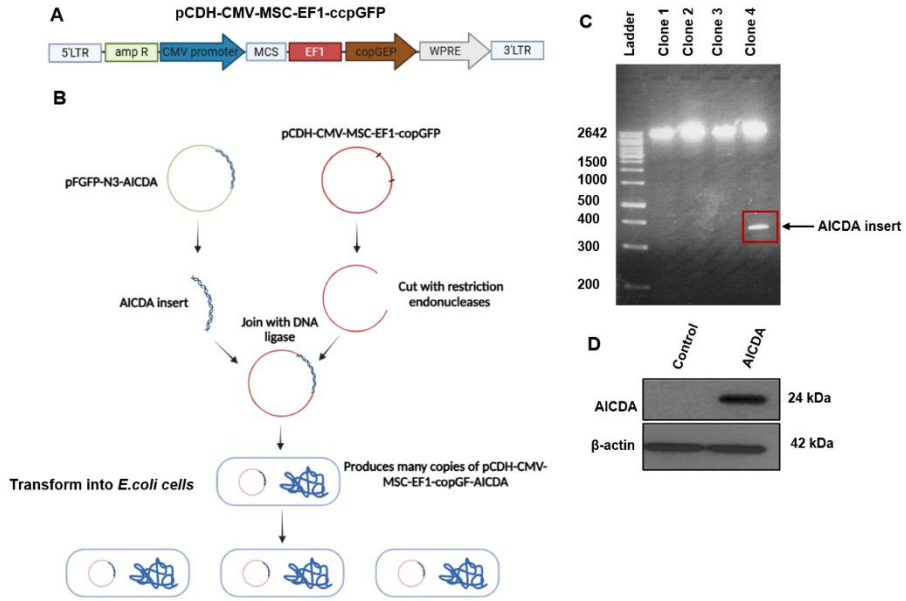


Figure 3.2 The human AICDA gene was successfully cloned into lentivector, pCDH-CMV-MSC-EF1-copGFP. (**A**) The schematic map of the lentivector. (**B**) The schematic diagram of the cloning process. (**C**) Agarose gel image showing the digested plasmid clones 1-4 with EcoR I and NheI-HF to release the insert of the AICDA gene. (**D**) Western blot analysis using an antibody against AICDA to validate the expression level of the protein in an AICDA-transfected HEK293 cell line, relative to cells transfected with an empty vector (control). An antibody to β-actin was used as the loading control.

3.3.2. Overexpression of AICDA in B lymphocytes leads to cell death

Next, PB-B95-8H and HBL-1 cell lines were transduced with a culture medium containing lentivirus generated in HEK293T cells (section 2.11). Over the days following transduction, the cells were visualized daily using an inverted microscope to evaluate the growth and health of the cells. Signs of

cell death, characterized by cell shrinking, were evident for both the PB-B95-8H and HBL-1 cells, albeit worse in the non-malignant PB-B95-8H cell population (Figure 3.3A Right panels). Under the GFP filter, a low percentage of GFP-positive cells was observed in both cell populations (**Figure 3.3A** Left panels). 11 days post-transduction, the cell populations were subjected to FACS (section 2.11) in an attempt to sort GFP-positive cells. During data analysis, the population of the HBL-1 AICDA and parental cells (GFP-negative) were gated and correlated. The populations overlapped, indicating the transduced cells contained background GFP fluorescence, similar to parent cells (**Figure 3.3B**). The yield following FACS was very poor for both HBL-1 and PB-95H-8H transduced cells (~ 2-3% of the total pre-sorted population). Sorted cells were placed in culture but started showing signs of undergoing apoptosis at 48 hours post-sorting. At 72 hours post-sorting, the entire PB-95H-8H cell population looked apoptotic, and attempts to extract protein from these cells failed in producing a yield. The sorted HBL-1 population appeared to recover and thus protein was extracted, quantified, and used in western blot analyses to assess AICDA expression relative to parental cells. No significant difference in AICDA expression could be seen between AICDA-transduced and parental cells (**Figure 3.3C**), which indicated that these cells were false positives, and could be the result of the expansion of a small group of non-AICDA expressing cells which were part of the sorted population. The entire process was repeated with the same result. It would therefore appear that the cells could not tolerate constitutively high expressions of AICDA, over and above their basal levels.

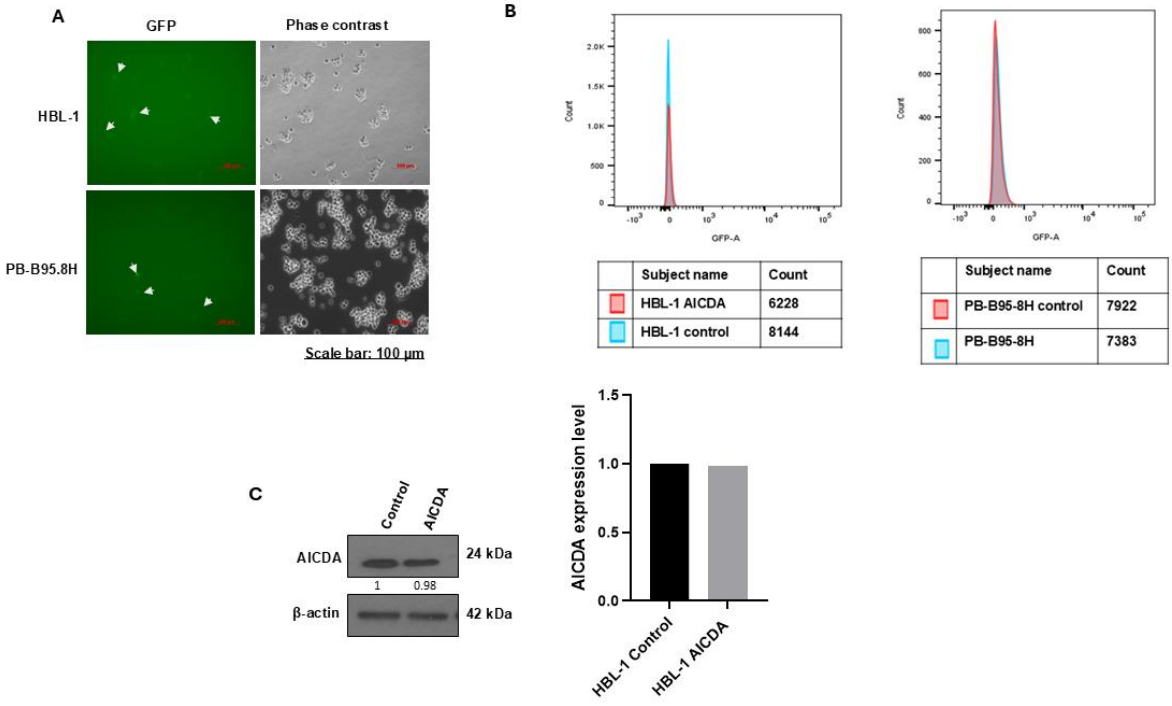


Figure 3.3 Overexpression of AICDA in B lymphocytes via lentiviral transduction was not successful. (**A**) Fluorescence microscopy images (20X magnification; scale bars 100 μm) of transduced HBL-1 and PB-B95.8H cells under GFP filter (Left hand panels) showing GFP-positive cells (represented by the white arrow bars) and under phase-contrast showing apoptotic cells (Right hand panels). (**B**) The histogram graph analysed using FlowJo software (Version 10) shows the distribution of GFP-A fluorescence intensity and the count of events. Both datasets were correlated to show the difference in GFP level. Transduced GFP-positive cells compared with control cells showed a peak at the lower end of the GFP-A axis that indicate low GFP expression. (**C**) Western blotting analysis using an antibody against AICDA to validate the expression level of the total protein extracted from the HBL-1 cells transduced with AICDA or empty vector (control). Bar graphs show quantification, using the ImageJ software (Version 1.5.4)) of AICDA expression relative to the house keeping protein β-actin used as loading control.

3.4. Development of an AICDA overexpression CRC model

3.4.1. The expression of AICDA in CRC tumours samples correlates with poor patient survival.

As an alternate cell model for the study, AICDA overexpressing CRC cells were considered. While several studies using patient samples implicate AICDA as having a role in CRC, these have mostly been correlative, without much insight into the mechanism(s) of oncogenesis. For instance, a study by Li and colleagues (2019) found that the prevalence of elevated AICDA expression was higher in sporadic CRC tumours, and that although AICDA expression was not associated with specific clinical features, there was a significant association between AICDA and the nuclear expression of p53 [74]. To begin validating the potential clinical significance of AICDA in CRC, bioinformatics analyses were performed using a publicly available dataset from the TCGA database. While there was no significant difference in the level of AICDA expression between Colon adenocarcinoma (COAD) tissues (n=349) and normal tissue samples (n=275) (**Figure 3.4A**), the general trend within this group showed that patients whose tumour cells harboured high AICDA expressions had poorer overall survival (**Figure 3.4B**) and reduced disease-free survival (**Figure 3.4C**), relative to patients with low AICDA-expressing tumours.

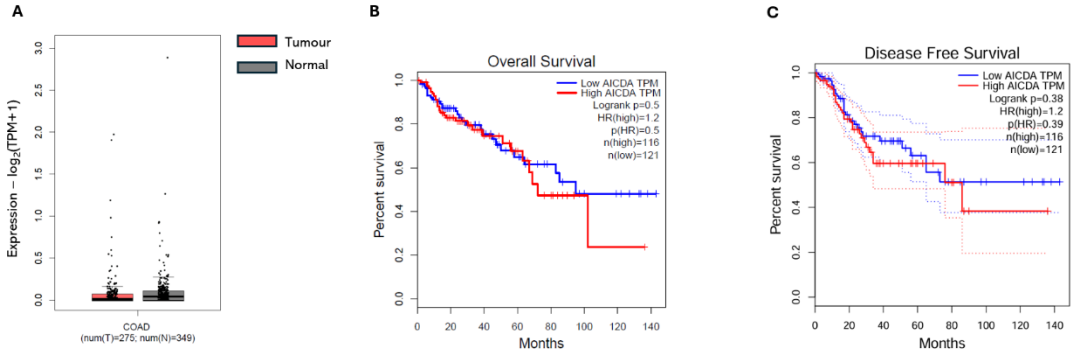


Figure 3.4 Clinical characteristics of CRC patients relative to AICDA expression within the TCGA database. (A) Boxplots shows analysis of the TCGA-COAD database for AICDA levels in COAD tissues (n=275) compared to normal tissues (n= 349). (B, C) Kaplan-Meier analysis of the overall survival and disease-free survival between COAD which express high and low levels of AICDA. The Logrank p-values are given as inserts.

3.4.2. Endogenous AICDA expression is induced by lipopolysaccharide (LPS) in CRC cells.

A previous report demonstrated that the expression of AICDA could be induced in CRC cells upon proinflammatory cytokine stimulation, namely by exposure to TNFα, IL-4 and IL-13, and TGFβ, and this overexpression was linked to enhanced mutagenesis within the TP53 gene [46]. Lipopolysaccharide (LPS) is a component of the outer membrane wall of Gram-negative bacteria and is also known to induce gut barrier dysfunction and inflammation, with a significantly link to the development of CRC [75]. Two CRC cell lines, namely DLD-1 and Caco-2, were cultured and exposed to LPS (10 µg/ml) for 24 hr, following which total proteins were extracted and the expression of AICDA investigated (section 2.16). As shown in **Figure 3.5**, significant induction of endogenous AICDA was detected in LPS-exposed cells, relative to unexposed cells, confirming that a proinflammatory environment enhances AICDA expression.

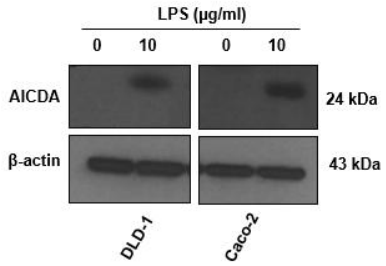


Figure 3.5 Endogenous AICDA expression is stimulated in CRC cells exposed to LPS. Western blotting analysis using antibody against AICDA to validate the expression of the total protein extracted from CRC cells treated and untreated with LPS (10 µg/ml) for 24 hours. β-actin was used as a loading control. At least two independent biological repeats were performed.

3.4.3. CRC cells tolerate constitutive overexpression of AICDA

Having established that CRC is a valid cancer model to study the oncogenic effects of AICDA with a potential role in epigenetic plasticity, we set out to transiently transfect the two selected CRC cell lines (DLD-1 and Caco-2) with a vector that allows for constitutive expression of GFP-tagged AICDA, followed by observing the cells in culture over a period of 72 hours, paying particular attention to morphological characteristics associated with cell death. Live cell imaging at 48- and 72 hours using fluorescence microscopy showed that successful transfection was achieved (**Figure 3.6 A&B**) in both control cells (transfected with vector backbone) and cells transfected with AICDA-expressing recombinant vector.

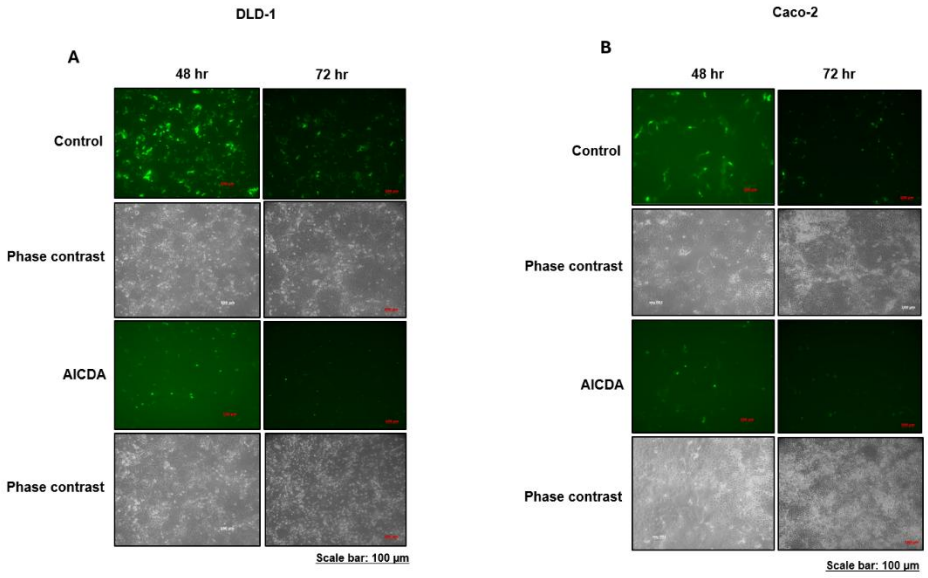


Figure 3.6 CRC cells tolerate AICDA overexpression. (**A, B**) Representative fluorescence microscopy images (40x magnifications; scale bars 100 µm) of DLD-1 and Caco-2 cells transfected with AICDA and control. The GFP signal was detected at 48 hr and 72 hr.

3.4.4. Development of CRC cells stably expressing AICDA

3.4.4.1 Geneticin kill-curve determination.

In order to generate a stable cell line expressing a transgene of interest, the minimum concentration of antibiotic required to kill non-transfected cells should be determined. Resistance to the antibiotic is acquired through antibiotic resistance genes on the plasmid becoming integrated into the nuclear genome of the host cell and thereafter propagated by cell division under selection pressure. The kill curve for both cell lines were developed using the antibiotic Geneticin (G-418) (section 2.12), which confers resistance to its analogue neomycin, the resistance gene present on the pcDNA3.1 plasmid, within which the human AICDA was cloned (plasmid kindly donated by Professor Reuben Harris from University of Minnesota, USA), see plasmid map in Appendix B; Figure B4) (**Figure 3.7A**). In DLD-1 cells, 600 µg/ml of G-418 was chosen as the lowest concentration that kills the cells within 10 days (**Figure 3.7B**). In Caco-2 cells, 1000 µg/ml of G-418 was chosen as the concentration that kills cells in 14 days as compared to the 10 days of the DLD-1 cells (**Figure 3.7C**).

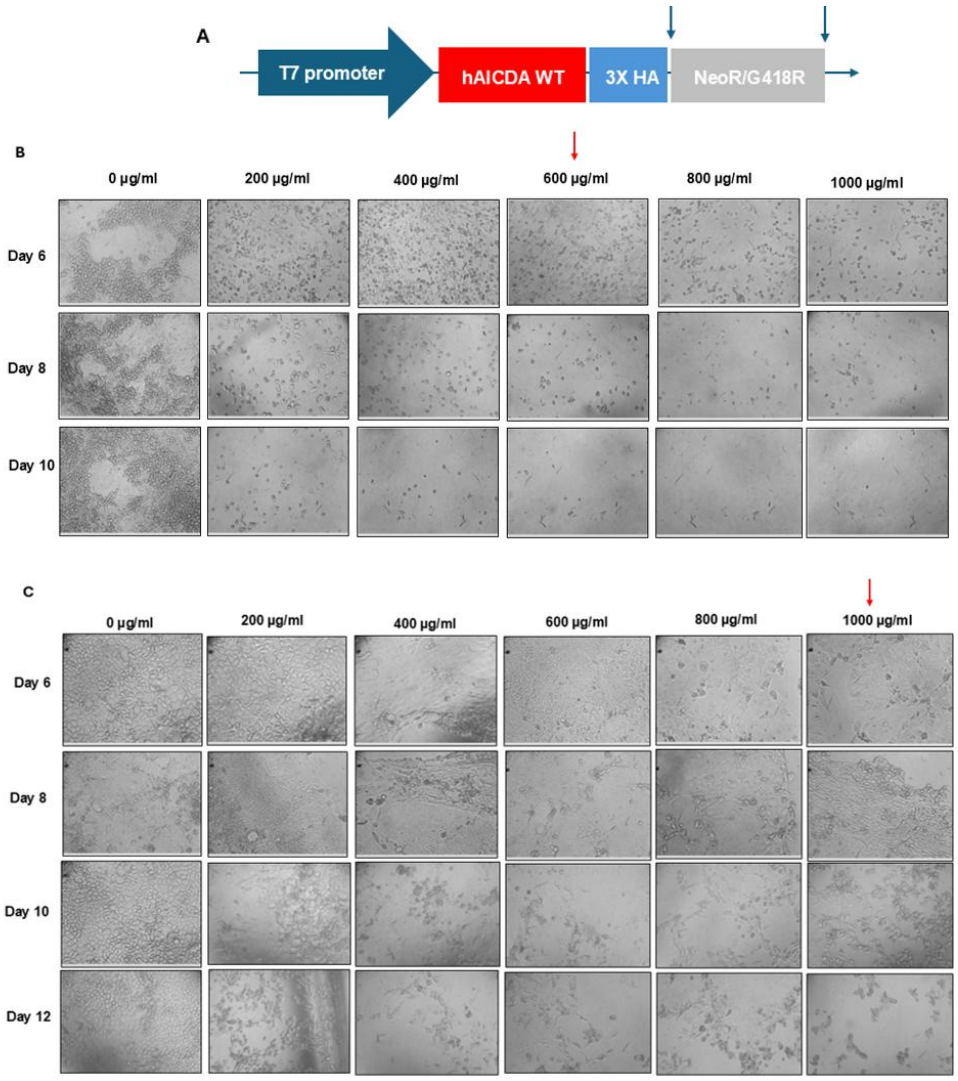


Figure 3.7 Geneticin kill-curve determination. (B&C) Representative of Dino eye image of the DLD-1 (Panel B) and Caco-2 (Panel C), treated with G-418 concentration range (0-1000 ug/ml). The 600 and 1000 ug/ml of G-418 are selected as the lowest concentration required to kill the cells after 10 to 14 days in culture (indicated by Red Arrow), respectively.

3.4.4.2 Selection and expansion of AICDA expressing clones.

Next, the DLD-1 and Caco-2 cells were transfected with the pcDNA-3.1-AICDA or empty vector constructs as controls, and the chosen concentrations of G-418 antibiotic were used as the selective pressure to treat the transfected cells (section 2.12; **Figure 3.8A**). Single resistant clones of the CRC cell lines were formed following 10-14 days of selection pressure, and for each cell line, at least 2-10 independent clones were expanded further, and AICDA expression was validated in DLD-1 (**Figure 3.8B**), Top and Below Panels) and in Caco-2 (**Figure 3.8C**). Therefore, the establishment of CRC cells stably expressing AICDA was successful, and two AICDA-expressing clones per cell line we selected for further analyses (indicated by asterisks).

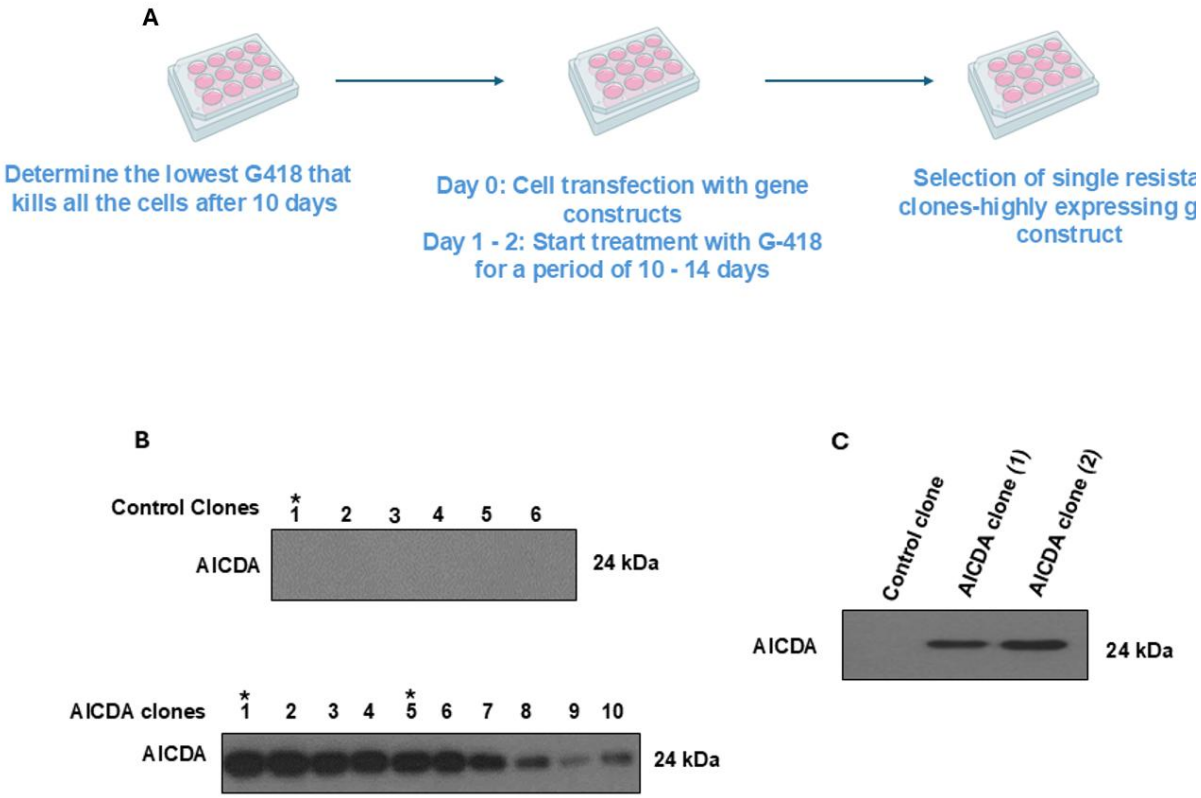


Figure 3.8 **Selection and expansion of AICDA expressing clones**. The schematic illustrates the process of establishing cells that express ectopic and stable AICDA overexpression. The first step involves the determination of the lowest G-418 antibiotic concentration required to kill the cells after 10 days. The second step incorporates the transient transfection of CRC cells with the gene constructs, and treatment with G-418 24 hours later, with the replacement of medium containing G-418 every 3 days. The third step involves the selection of the formed single resistant clones-highly expressing gene constructs. (B&C) Western blotting analysis using an antibody against AICDA to validate the expression of the total protein extracted from the clones of DLD-1 (B – top and lower panel) and Caco-2 (C) with AICDA and empty vector.

3.4.5. Assessment of the impact of AICDA expression on cancer hallmarks in CRC cells.

As previously mentioned, although published reports indicate an association between AICDA expression and the development of colon cancer, scientific studies on the impact of AICDA expression on specific cancer hallmarks are currently lacking. Therefore, to fill this gap, the impact of AICDA on three major cancer hallmarks, namely cell proliferation, migration and response to drug was investigated.

3.4.5.1. AICDA expression enhances proliferation of CRC cells

To assess proliferation, two methods were employed. First, long-term proliferation was investigated by counting cells at 48-hour intervals over a period of 6 days (section 2.17), and the result is shown in **(Figure 3.9A&B)** below. The growth curve shows that AICDA (1) and AICDA (2) clones for both CRC cell

lines proliferated significantly faster relative to their corresponding controls. The difference starts to become noticeable from day 2, and at day 4, AICDA expressing cells had 1.4-fold and 2.1-fold more cells for DLD-1 and Caco-2 cells respectively, relative to controls. At day 6, AICDA expressing cells had 1.3-fold and 1.3-fold more cells for DLD-1 and Caco-2 respectively, relative to controls. Secondly, short-term proliferation was assessed using the WST-1 assay, which is an assay that measures cell viability using mitochondrial activity as a proxy (section 2.18). As shown in **(Figures 3.9C&D),** at 24 hours, AICDA expressing DLD-1 cells had 1.6-fold more cells, while at 48 hours showed 1.3-fold more cells relative to controls. The two clones of Caco-2 behaved differently, and at 24 hours AICDA (1) and AICDA (2) had 1.8-fold and 2.9-fold more cells, respectively. At 48 hours, AICDA (1) had 2.3-fold more cells and AICDA (1) had 2.9-fold for more cells relative to the controls. These results therefore clearly demonstrate that the expression of AICDA provides CRC cells with a proliferative advantage.

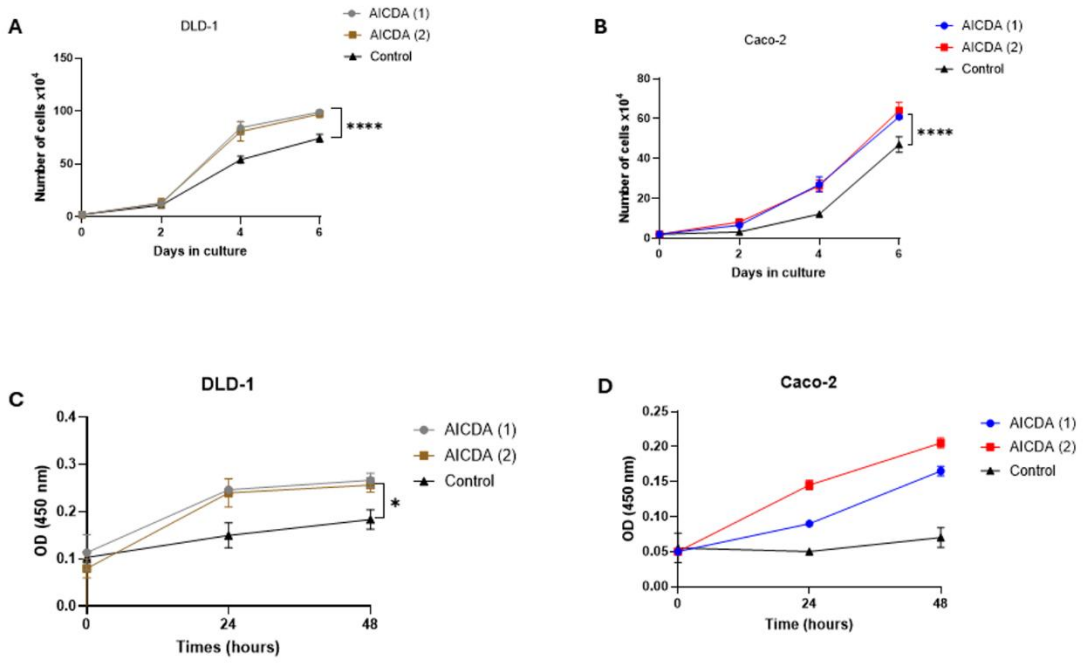
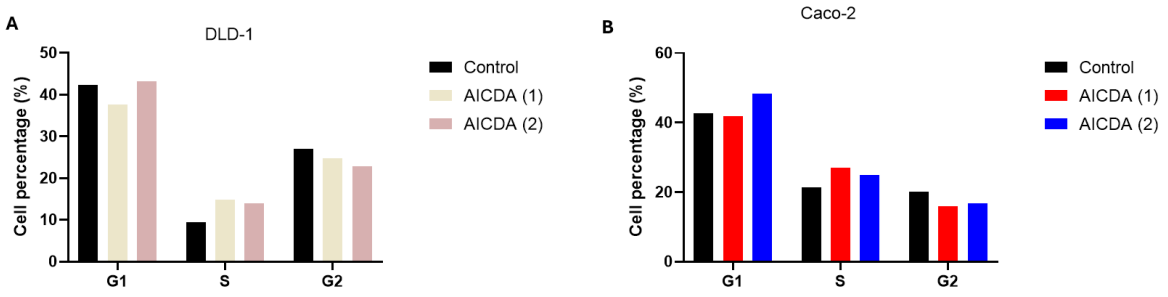


Figure 3.9 AICDA expression promotes CRC cell proliferation in culture. (**A, B**) Growth curve analysis of the CRC cells (control, AICDA (1) and AICDA (2)) was performed over 6 days, and the cells were counted at 2-day intervals. (**C, D**) Representative of the WST-1 assay, Optical density (OD) at 450 nm was measured to quantify the cell proliferation (control, AICDA (1) and AICDA (2)) at 24 and 48 hours. Data was analysed using the parametric unpaired t-test by the GraphPad Prism 8.0 software. P values are indicated as *p<0.5, **p<0.01, ***p<0.001, ****p<0.0001; error bars represent mean ± SEM of the three replicate. At least two independent biological repeats were performed for each experiment.

3.4.5.2. AICDA expressing CRC cells display mild S-phase block

The impact of AICDA expression on cell cycle profiling was evaluated to establish if the protein affected a specific phase of the cell cycle (G1, S, G2/M). As discussed previously, AICDA is a DNA mutagen, and its overexpression is linked to disruption of the genomic integrity [76]. At least three DNA damage checkpoints have been described in the cell cycle, namely within the late G1 phase, S phase, and mid to late G2 phase [77]. Cells were fixed, stained with FxCycleTM PI/RNase, and subjected to flow cytometry as per an established protocol (section 2.21). The data generated was analysed using the FlowJo software (version 10) and represented as bar graphs in (Figure B7, Appendix B; **Figure 3.10** below). A mild block in cells undergoing S-phase can be observed for the AICDA expression clones in both cell lines, with the effect being slightly more pronounced in DLD-1. Specifically, in the representative experiments shown here, in the DLD-1 cell line, S phase cells increased by ~5.5% (~0.6 fold) and ~4,5% (~0.7 fold) in the AICDA (1) and AICDA (2) respectively, relative to the control clone (**Figure 3.10A**). In the Caco-2 cell lines, S phase cells increased by ~5.6% (~0.6 fold) and ~3.4% (~0.8 fold) in the AICDA (1) and AICDA (2) respectively, relative to the control (**Figure 3.10B**). In both cell lines, the mild S-phase block appeared to lead to a slight reduction in G2/M phase cells. These results suggest that AICDA leads to mild DNA replication stress which could be triggering the S phase DNA damage checkpoint.



	G1 (%)	S (%)	G2 (%)
Control	42.3	9.34	27.0
AICDA (1)	37.6	14.8	24.7
AICDA (2)	43.2	13.9	22.8

	G1 (%)	S (%)	G2 (%)
Control	42.8	21.4	20.0
AICDA (1)	41.9	27.0	16.0
AICDA (2)	48.3	24.8	16.7

Figure 3.10 Expression of AICDA promote cell proliferation by influence cell cycle progression. (A, B) The bar graphs represent data analysed using the FlowJo software (version 10). CRC cells (control, AICDA (1) and AICDA (2)) cells stained with the FxCycleTM PI/RNase for cell cycle distribution. The graph represents the percentages of cells in each cell cycle phases. The experiment was done in two biological independent repeats.

3.5 Expression of AICDA promoted cell migration of CRC cells in culture

To investigate whether AICDA expression enhances the migratory abilities of CRC cells, a Wound healing assay was performed (section 2.19). The cells were cultured to form a monolayer, and a vertical wound was created using a pipette tip, and the cells were subsequently treated with 5 ug/ml Mitomycin C, which prevents cell proliferation, therefore ensuring that any changes observed are as a result of migration. The wound area was imaged at 3-hour intervals. The images were quantified using Fiji ImageJ software (Version 2.9.0) and represented as bar graphs in **Figure 3.11**. As shown in Figure **3.11A&B,** in DLD-1 and Caco-2, AICDA (1) and AICDA (2) cells the gap closed significantly faster than the wound in the control dishes. In DLD-1 cells at 24 hours, the gap in the AICDA (1) and AICDA (2) dishes had reduced by ~1.8-fold and ~2.2-fold more, respectively, relative to control. In Caco-2 dishes, it was ~2.6-fold and ~2.3-fold more, respectively. Western blot analyses were performed to assess whether this increase in migratory ability was through enhanced expression of the Epithelial-mesenchymal transition (EMT) promoting transcription factors ZEB1 and 2. As can be seen in **Figure 3.11C**, ZEB1 is upregulated in AICDA (1) within the DLD-1 cell model, and in both AICDA clones within the Caco-2 cell model, relative to the corresponding controls. Zeb 2 was only upregulated in DLD-1 clones and in fact appeared to be downregulated in the Caco-2 AICDA clones. Overall, this data revealed that AICDA expression provides CRC cells with a migratory advantage.

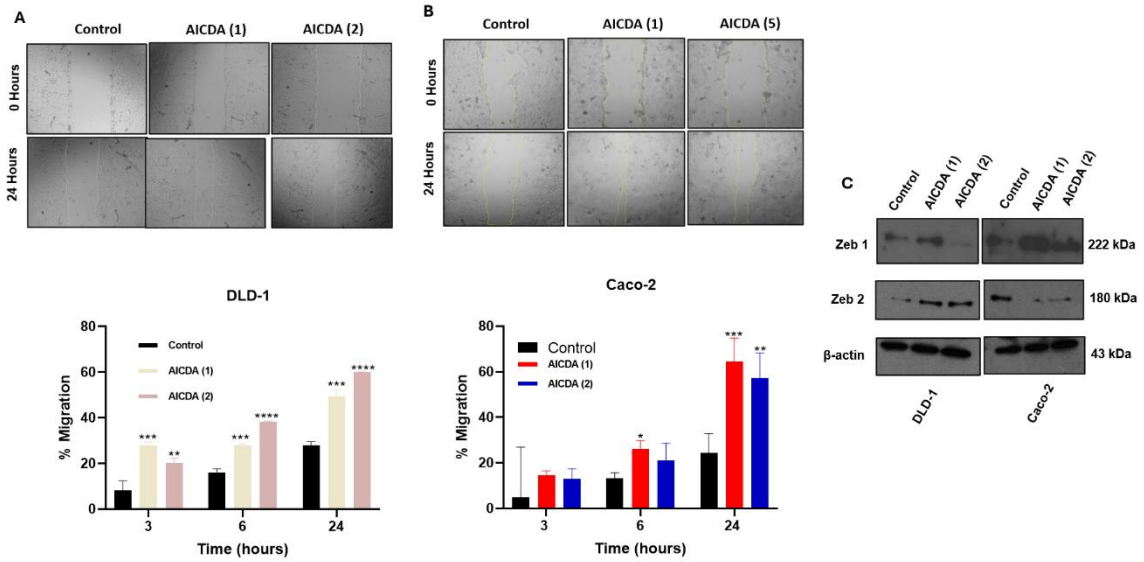


Figure 3.11 AICDA expression promoted CRC cell migration. (A-B) Representative microscopy images (10x magnifications; scale bars 100 µm) and quantification of the migration of empty vector, AICDA (1) and AICDA (2) CRC cells measured using wound healing assay in the presence of anti-proliferative agent Mitomycin C. (**C**)

Western blotting analysis of Zeb 1 and 2 in DLD-1 and Caco-2 cell line with control, AICDA (1) and AICDA (2). β-actin was used as a loading control. Data was analysed using the parametric unpaired t-test by the GraphPad Prism 8.0 software. P values are indicated as *p<0.5, **p<0.01, ***p<0.001, ****p<0.0001; error bars represent mean ± SEM of the four replicates. The experiment was done in triplicate.

3.6 AICDA expression reduces sensitivity to chemotherapeutic drug 5-FU in DLD-1 cells but not in Caco-2 cells.

The chemotherapeutic drug 5-Fluorouracil (5-FU) is an antimetabolite drug that is widely used for the treatment of colorectal cancer [78]. 5-FU is known to exert its anti-tumour effects primarily by integrating into RNA, and inhibiting thymidylate synthase (TS), thereby disrupting intracellular deoxynucleotide pools important for DNA replication and subsequently inhibiting DNA biosynthesis. To determine the effects of AICDA overexpression on cell sensitivity to 5-FU in CRC cells, a WST-1 viability assay was performed (section 2.18). The CRC cells (control, AICDA (1), and AICDA (2)) were treated with a concentration range of 5-FU drug (0-500 µg/ml) for 72 hours. The data was quantified using the GraphPad Prism 8.0 software and shown in Figure 3.12. The IC_{50} values for DLD-1 control cells were 159.72 µg/ml, while that of the AICDA clones were significantly higher at 448.61 µg/ml (AICDA (1) and 447.22 µg/ml (AICDA (2) (**Figure 3.12A& B**). The IC_{50} values for Caco-2 control cells were 145.0 µg/ml, while that of the AICDA clones were 216.67 µg/ml (AICDA (1), and 234.72 µg/ml (AICDA (2) (**Figure 3.12C&D**). Taken together, the results indicated that AICDA overexpression in DLD-1 cells leads to a marked decrease in sensitivity of 5-FU, whereas in Caco-2 cells, the trend is similar but less pronounced.

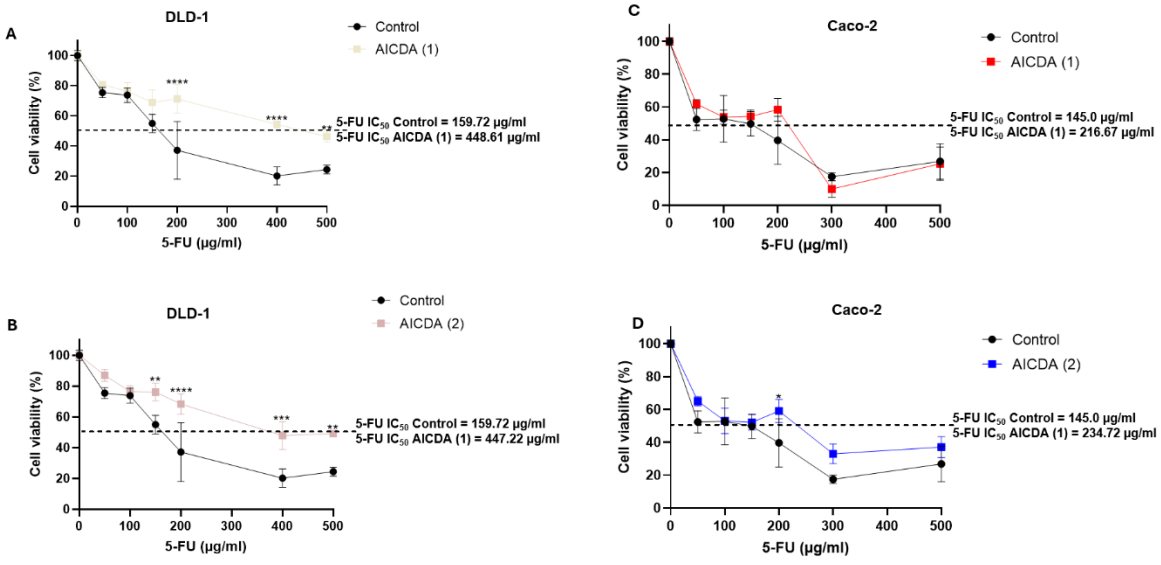


Figure 3.12 CRC cells expressing AICDA exhibit less sensitivity to the 5-FU drug. (**A-D**) The cell viability of the CRC cells (Control, AICDA (1), and AICDA (2)) was evaluated after 72 hours of treatment with the 5-FU drug (0 – 500 µg/ml), using WST-1 assay. Data was analysed using a 2-way ANOVA and Sidak's multiple comparisons test by the GraphPad Prism 8.0 software. P values are indicated as *p<0.5, **p<0.01, ***p<0.001, ****p<0.0001; error bars represent mean ± SEM of the three replicates. The experiment was performed in triplicate.

To confirm the above data, long-term proliferation was investigated by counting cells at 24-hour intervals over a period of 4 days (section 2.17). The DLD-1 clones were treated with 79.86 µg/ml 5-FU, while the Caco-2 clones were treated with 72.5 µg/ml of 5-FU, which represented approximately half of the IC_{50} obtained from the experiment above. The data was quantified using GraphPad Prism 8.0 software and represented as bar graphs in **Figure 3.13** below. While overall the data did not show much statistical significance, a clear trend was observed for the DLD-1 cells, whereby the proliferation rate of the AICDA-expressing cells was higher than that of the control, in the presence of 5-FU (**Figure 3.13A).** On the other hand, no similar trend was seen for the Caco-2 cell lines (**Figure 3.13B**). These findings suggest that AICDA may provide CRC cells with mild resistance to the chemotherapeutic drug 5-FU.

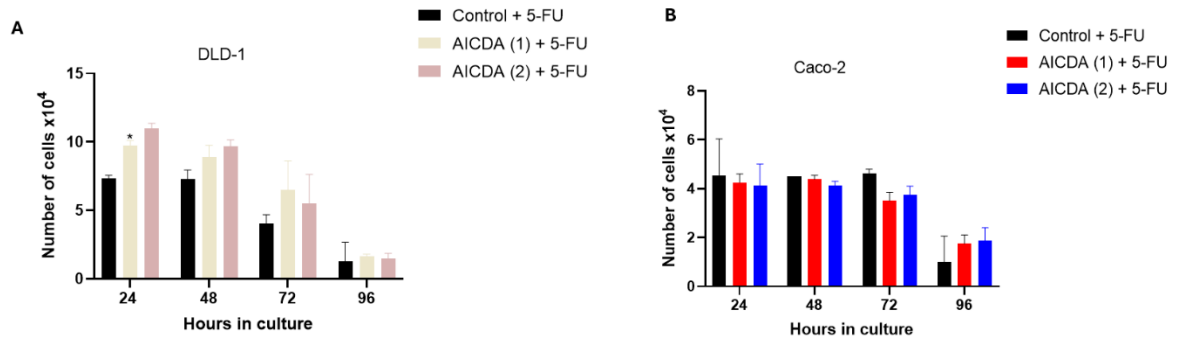
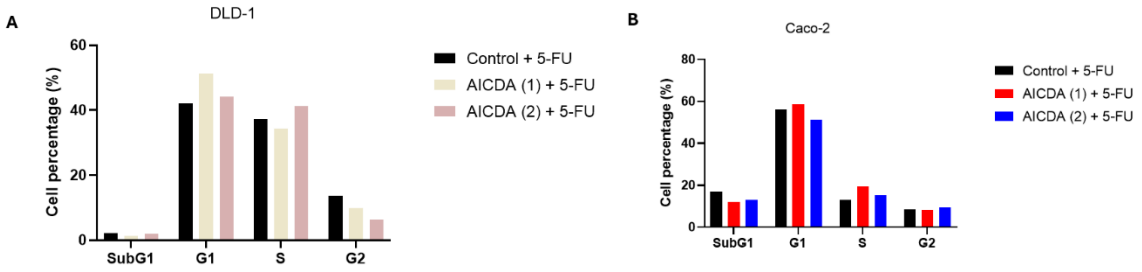


Figure 3.13 AICDA gives a mild proliferative advantage to CRC cells in the presence of 5-FU. (A-B) Growth curve analysis of the CRC with empty vector, AICDA (1) and AICDA (2) treated with half IC50 values for 96 hrs. The cells were trypsinized and counted at 24-hour intervals. Data was analysed using the parametric unpaired t-test by the GraphPad Prism 8.0 software. P values are indicated as *p<0.5, **p<0.01, ***p<0.001, ****p<0.0001; error bars represent mean ± SEM of the two replicates. The experiment was performed in duplicate.

The impact of AICDA expression on cell cycle profiling of the CRC cells in the presence of 5-FU was also investigated. Once again cells were treated with half of the IC_{50}, this time for 48 hours, after which cells were stained with PI and the cell cycle analysed using flow cytometry (section 2.21). The data generated was analysed using the FlowJo software (version 10) and represented as bar graphs in **Figure 3.14** below. A small decrease in the percentage of cells within the Sub/G1 peak was observed for the AICDA expression clones for both cell lines, indicating reduced cell death. Another notable change was an increase in the G1 phase for the DLD-1 AICDA (1) clones, and a decrease in the G2 phase, relative

to controls (**Figure 3.14A**). In Caco-2 cells (**Figure 3.14B**), the S phase increased by ~6.6% (~1.5-fold) and ~ 2.3% (~1.1-fold) in AICDA (1) and AICDA (2) respectively, relative to controls. Overall, these results indicate that AICDA may provide mild protection from cell death by 5-FU (reduced sub-G1 peaks), but at the same time, may delay S-phase in these cells.



	Sub G1 (%)	G1 (%)	S (%)	G2 (%)
Control	2.2	42.0	37.3	13.6
AICDA (1)	1.3	51.2	34.4	9.94
AICDA (2)	2.0	44.1	41.2	6.39

	Sub G1 (%)	G1 (%)	S (%)	G2 (%)
Control	17.0	56.0	13.0	8.6
AICDA (1)	12.1	58.7	19.6	8.2
AICDA (2)	13.1	51.3	15.3	9.6

Figure 3.14 AICDA provides mild protection from cell death induced by 5-FU, and delays S-phase, in CRC cells (A, B) The bar graphs represent data analysed using the FlowJo software (version 10). CRC cells (control, AICDA (1) and AICDA (2)) were treated with the 5-FU drug. The cells were stained with the FxCycleTM PI/RNase for cell cycle distribution. The graph represents the percentages of cells in each cell cycle phase.

3.7 5-FU modulates the expression of apoptotic markers caspase-3 and PARP-1 in AICDA-expressing CRC cell lines

To investigate whether AICDA expression alters the expression of apoptotic markers in CRC cells treated with 5-FU, western blot analyses were performed, probing for the expression of cleaved caspase-3, and its target poly (ADP-ribose) polymerase (PARP). During apoptosis, PARP-1 is targeted and cleaved by caspase-3 and -7 at its caspase-cleavage site, releasing two fragments of 24 kDa and 89 kDa in size. Caspase-3 also gets cleaved upon apoptosis into products between 14 kDa and 21 kDa fragments. Both of these proteins are well-established markers of apoptosis. PARP1 cleavage is also observed during necrosis, where major fragments of 50 kDa are observed.

In this experiment, DLD-1 and Caco-2 cells were treated with IC_{50} of 5-FU drug, respectively, for 48 hours. The corresponding control CRC cells were left untreated for the same period. Western blot was performed on total protein extracts (section 2.13). As shown in **Figure 3.15A-D**, the results indicated that treatment with 5-FU increased the levels of cleaved PARP-1 (89 kDa) in both DLD-1 and Caco-2 cell lines (Figure 3.15A & C) and that no major differences were observed between control and AICDA-

expressing clones (for the major band of 89 kDa). It should be noted that, in the case of the untreated Caco-2 cells, there appears to be a slight increase in cleaved PARP-1 within AICDA (clone 1) (**Figure 3.15C**). Notably, in the DLD-1 cells, 5-FU further induced cleaved fragment sizes of 45, 50 and 74 kDa for PARP-1, relative to the untreated cells, suggesting the potential effect of 5-FU in necrosis induction (**Figure 3.15A**). In contrast, these cleaved PARP-1 fragments were not detectable in Caco-2 cells even after a longer exposure (data not shown). With regards to caspase-3, as was the case for PARP-1, treatment with 5-FU led to the cleavage for all clones (**Figure 3.15B & D**), which is more pronounced in DLD-1 clones, relative to Caco-2 clones. Interestingly, lower expression of cleaved caspase-3 was observed within the AICDA-expressing clones for both cell lines (**Figure 3.15B & D**).

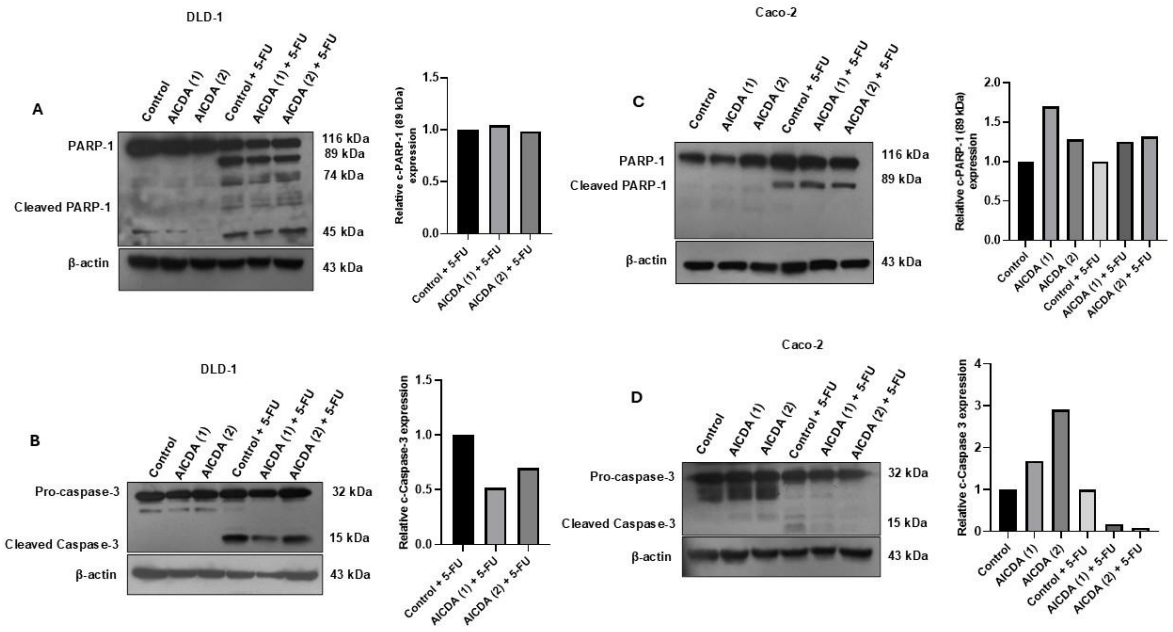


Figure 3.15 The treatment with 5-FU induces apoptosis and necrosis in CRC cells, which was not altered when AICDA was expressed. (A-E) Western blotting analysis using an antibody against c-PARP-1 and c-Caspase-3 to validate the protein expression of c-PARP-1 and c-Caspase in CRC cells expressing AICDA treated with and without 5-FU drug for 48 hours. β-actin was used as a loading control. The signal intensity of the bands was analysed using Fiji ImageJ software (Version 2.9.0), normalised to β-actin, and relative to control. The results are represented as bar graphs.

3.8 The impact of AICDA expression on transcriptional regulation of oncogenes

To investigate the impact of AICDA expression on the transcriptional regulation of a panel of selected oncogenes in CRC, quantitative PCR (qPCR) was performed (section 2.15). A total of seven genes were included in the panel. These included *PIM1* and *FANCA*, which have been previously reported to be regulated by AICDA at their promoter level via epigenetic modulation [62, 64]. The results show that

in DLD1-1, AICDA significantly downregulated *PIM1*, while no significant changes were observed in Caco-2 (**Figure 3.16A&B**). *FANCA* was significantly downregulated relative to the control cells for both cell lines **(Figure 3.16C&D**). The expression of two methyl transferase genes, namely DNMT1 and DNMT3A, were also investigated. Overall, the transcription of these two genes appears to be downregulated in AICDA-expressing cells, relative to the control cells (**Figure 3.16DE-H**). The transcription of *ZEB1*, which is associated with enhanced migratory abilities, was downregulated in AICDA-expressing DLD-1 cells, but not in Caco-2 cells, relative to controls (**Figure 3.16I, J**). For the immunomodulatory molecule PDL-1 (encoded by the *CD274* gene), opposite observations were made for the two CRC cell lines: upregulation in DLD-1 and downregulation in Caco-2 when AICDA is expressed. Lastly, both CRC cell lines showed upregulated *MYC* expression levels in AICDA clones relative to the control (**Figure 3.16M, N**). In conclusion, AICDA expression in CRC cells significantly impacts the transcriptional regulation of the selected genes, however, differential regulation is observed between the two CRC cell lines, for some of the genes.

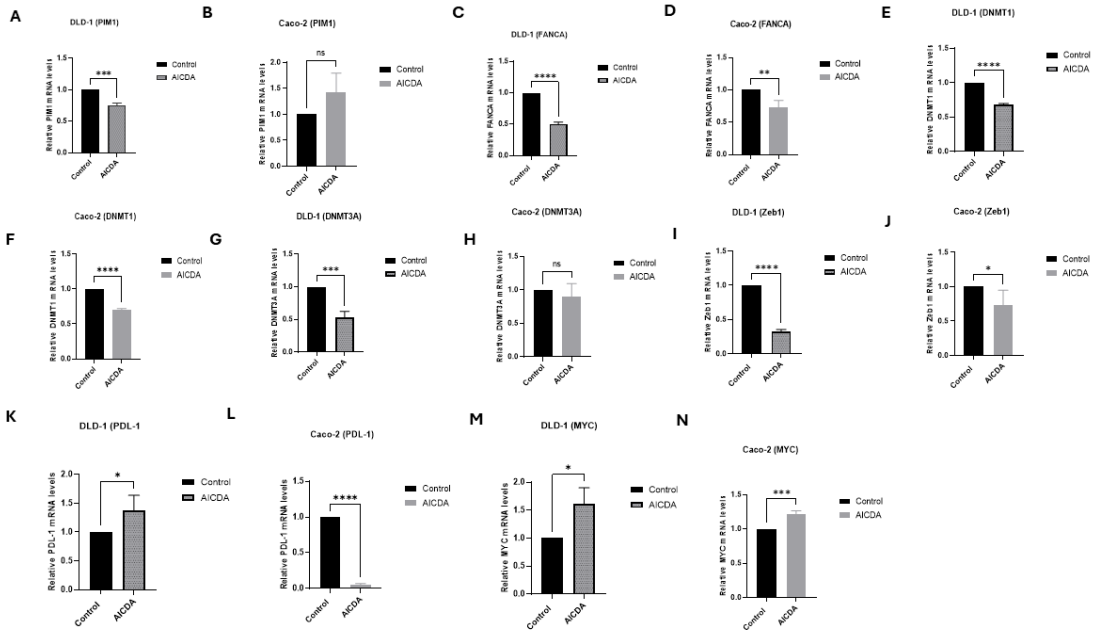


Figure 3.16 Modulation of transcription of selected key oncogenes in CRC cell lines expressing AICDA. (**A-N**) qPCR analysis to evaluate the impact of AICDA on the transcriptional regulation of the selected oncogenes. GAPDH was used as housekeeping gene. Data was analysed using the parametric unpaired t-test by the GraphPad Prism 8.0 software. P values are indicated as *p<0.5, **p<0.01, ***p<0.001, ****p<0.0001; error bars represent mean ± SEM of the three replicates. The experiments were performed at least in duplicate.

3.9 Discussion

Since its discovery in 1999, AICDA has been the focus of numerous studies due to its ability to modify DNA. Researchers have extensively investigated its role in chromosomal rearrangement, particularly the *c-MYC-IGH* translocation in Burkitt lymphoma (BL) [48]. More recently, the oncogenic function of AICDA has been found to extend to the modulation of the epigenetic landscape, attributed to the enzyme's capability to alter CpG sites and islands through its deamination function. Recent studies have described AICDA as a cofactor, in which it partners with epigenetic modulators, such as TET2 and DNMT1. AICDA recruited TET2 to the promoter region of the FANCA gene, and this led to a reduction in methylation within that promoter and was associated with gene expression [62]. AICDA partnered with the DNMT1, and this interaction was demonstrated to be involved in increasing the methylation in BCL2 oncogenes [79]. However, the direct involvement of AICDA in altering gene expression by disrupting DNA methylation remains uncertain due to a lack of scientific evidence. Our findings of the bioinformatics analyses indicated that elevated AICDA expression is associated with poorer patient survival rates in the TCGA-DLBC database. To explore its impact on the epigenetic landscape in B-cell lymphomas, we first needed to generate B-cell models with AICDA overexpression.

The tight membrane structure and low endocytic activity of B cells make these cells difficult to transfect with the traditional chemical methods, possibly due to high sensitivity to physical stress [67]. Therefore, a third-generation lentiviral transduction system was used following the cloning of the human *AICDA* ORF into a relevant lentiviral vector. Lentiviral particles were successfully produced, harvested, and verified in HEK293 cells, however, the development of DLBCL-cells overexpressing AICDA could not be achieved. This was due to rapid cell death, likely due to oncogene-induced death. We hypothesize that since these cells already express relatively high levels of AICDA, pushing this threshold leads to a deadly accumulation of mutations and DNA damage, triggering apoptosis. With that, we propose a different strategy for future experiments, whereby the *AICDA* gene will be knocked down in these cell lines using the siRNA or shRNA.

In the current project, to overcome this potential roadblock and continue to investigate the impact of AICDA on the transcriptional regulation of key cancer genes, an alternative approach was taken by substituting B-cell-derived cells with CRC cell lines. This was supported by the numerous studies indicating that *AICDA* expression is associated with the development of CRC, and is enhanced by pro-inflammatory cytokines, which are associated with tumorigenesis in this cancer type [21, 42, 43, 44]. Our TCGA-COAD bioinformatics analysis indicates that, although AICDA expression was relatively

comparable in CRC tumour and normal tissues, high *AICDA* expression led to poor patient survival rates in the COAD cohort.

LPS, which serves as the potent stimulator of the immune system, and also known as an inflammation driver, could potentially enhance AICDA expression in CRC cells. In general, LPS has diverse functions, one includes its nature to serve as the major structural component of the outer membrane in Gram-negative bacteria [80]. Due to the fact that LPS decorates itself on the surface of many bacterial pathogens, it attracts the host immune system to respond dramatically to its presence. This interaction tends to accelerate the opportunity to induce pro-inflammatory cytokines markers. Previous studies have convincingly demonstrated that exposure to LPS promotes migration, invasion and EMT, contributing significantly to CRC development [81]. LPS binds to the transmembrane receptor TLR4, which is located on the surface of various immune cells and endosomal compartments [75]. TLR4 activation initiates/recruits the MyD88 pathway, an important signalling pathway for TLR4, involving the adaptor protein MyD88 with a death domain [82]. Specifically, MyD88 binds to TLR4 at its TIR domain upon recognition, which leads to the activation of downstream signalling pathways, including a nuclear factor-kappa B (NF-kB), a ubiquitous transcription factor known to upregulate AICDA activity in other cancer contexts. Interestingly, the activation of NF-kβ has also been shown to elevate the production of pro-inflammatory cytokines and chemokines, which are reported to activate AICDA activity [46]. It is, therefore, possible that LPS could be activating and upregulating AICDA expression via the NF-kB signalling pathway. Nevertheless, further research is needed to elucidate the pathway involved in AICDA activation upon LPS treatment.

The CRC cell line models ectopically expressing AICDA did not show signs of stress, contrary to what was experienced for the B cell models. To account for potential variability, two independent AICDA clones were used in assays, for each CRC cell line. Also given that variations could be as a result of different levels of expression of the protein of interest, and also the site of the integration may lead to pleiotropic effects. It was necessary to characterize the AICDA CRC cell lines, and this was done by measuring hallmarks of cancer, namely proliferation, migration, and response to chemotherapeutic drugs. Growth curve analysis was based on counting the cells, demonstrating that AICDA overexpression enhances the proliferation compared to the control in both CRC cell lines. WST-1 assay further confirmed the results of the growth curve, which clearly demonstrates that AICDA promotes cell proliferation.

The p53 transcription factor is one of the tumour suppressors that regulates the cell cycle, which prevents uncontrolled cell growth. This factor achieves its important functions by also activating key genes involved in cell cycle arrest, DNA repair, and apoptosis [83]. Interestingly, previous reports found that the overexpression of *AICDA* in CRC cells enhanced the acquisition of inactivating mutations in the TP53 gene [46, 84]. It is therefore probable that AICDA expression is leading to impaired TP53 protein via its DNA-modifying ability in our cell lines, allowing cells to bypass tumour-suppressing mechanisms. Notably, we found that the transcription of c-MYC, a potent pro-proliferative factor and oncogene, was found to be increased in both CRC cell lines when AICDA was expressed. Although the same extent of increase was not seen for the c-MYC protein (Figure B7, Appendix A), it is likely that the proliferative advantage is acquired, at least in part, by the enhanced MYC expression.

Our findings also show that AICDA expression results in mild DNA replication stress, which could trigger the S phase DNA damage checkpoint. This checkpoint is activated to monitor and respond to any DNA damage that occurs during DNA replication in the cell cycle [85]. Consequently, the results were not surprising due to the fact that AICDA activity is known to induce genomic instability, which may have activated the observed peak in the S phase checkpoint. AICDA expression promoted migration in this CRC cell model. These results were consistent with the upregulation of the EMT transcription factors such as ZEB1 and ZEB2 at the protein level in these cells when AICDA is expressed. It was surprising that CRC clones expressing AICDA exhibited upregulation of Zeb 1 and 2 at the protein level, despite downregulation at the mRNA level. It is speculated that these contrasting results could be due to AICDA activating microRNAs, which targeted and degraded Zeb mRNA post-transcription. The residues of the mRNA might be enhanced by other regulatory mechanisms or potentially by AICDA itself, leading to the observed increase in protein levels.

The chemotherapeutic drug 5-FU is part of the standard regimen for the treatment of CRC [78]. Generally, it exerts its anti-cancer effects by interfering with DNA and RNA synthesis via inhibiting thymidylate synthase (TS) and incorporating it into DNA and RNA [78]. This has been associated with DNA damage and mostly prevents the cancer cells from reproducing and making vital proteins. Taken together, our findings indicate that 5-FU is capable of reducing cell proliferation in CRC cells, and that a mild advantage is seen when AICDA is expressed. However, this could be an artificial observation, due to the rapid proliferating rate seen for cell expression AICDA. 5-FU led to an increased S phase in both cells, suggesting that DNA damage induced by 5-FU halted DNA replication, which is expected. This effect was mildly pronounced in AICDA cells. This could be the fact that AICDA introduces mutations via its deaminase nature,.i.e. conversion of cytosine into uracil, which sometimes leads to

thymine, thus this may perhaps provide genetic diversity in CRC with AICDA expression, allowing them to escape the arrest.

Western blotting was performed to investigate the effect of 5-FU on apoptosis by validating the expression level of apoptotic markers. The findings indicate that 5-FU increased the levels of cleaved PARP-1 (89kDa) in both DLD-1 and Caco-2 cell lines, regardless of AICDA overexpression. Unexpectedly, 45 kDa and 74 kDa product sizes of the cleaved PARP-1 were also detected in DLD-1 cells. These bands have been associated with the state of undergoing necrosis. These bands were not detected in the Caco-2 cells, which further demonstrated the difference in the genetic profile of these cell lines. Another interesting finding was the reduced expression of the 15 kDa cleaved caspase-3 product in both cell lines overexpressing AICDA relative to controls, when treated with 5-FU. This indicates the potential protection against apoptosis by AICDA which should be investigated further.

Quantitative PCR (qPCR) was performed to investigate the impact of AICDA expression on the transcriptional regulation of seven key oncogenes, namely *PIM1, FANCA, DNMT1, DNMT3A, ZEB1, CD274*, and *MYC*. PIM1 and FANCA genes were selected based on published findings indicating that AICDA modulates their expression through methylation/demethylation [62, 86]. Additionally, DNMT1 and DNMT3A were analysed for their roles in enhancing genome-wide methylation frequency [87]. Finally, ZEB1, PDL-1, and MYC – oncogenes linked to critical cancer phenotypes such as migration, immune escape, and proliferation [21, 41, 88] - were also assessed to explore the impact of AICDA regulation. The results showed that AICDA expression in CRC significantly impacted the expression level of the selected oncogenes. The findings of the study showed that AICDA upregulated and downregulated some of the oncogenes, this could be due to the fact this DNA-modifying protein has a dual role, in which it can impact the promoter of the genes via demethylation – which switches on the genes and upregulate their expression level, and methylation – which enhance methyl group frequency in the promoter regions and downregulate the expression level. This protein might have achieved those changes by partnering with either of the known epigenetic regulators, TET2 and DNMT1 and not yet known factors. Variations were observed between the two CRC cell models, which implies that the cellular context is important in mediating the impact of AICDA on changes in gene transcription.

3.10 Overall Conclusion and Future Directions

In conclusion, the findings of this study demonstrate that AICDA functions as an oncogene in CRC and provides the cells with a significant proliferative and metastatic advantage. Crucially, it shows that

AICDA expression impacts the transcription of key oncogenes. These results warrant further investigation to understand how AICDA leads to these transcriptional changes. Future work will include performing pyrosequencing to measure the methylation status of the selected genes, a pull-down assay to identify epigenetic regulator proteins that partner with AICDA at the promoter regions of the genes, and lastly, investigating the effect of AICDA-mediated methylation changes on cancer phenotypes.

3.11 References

[1] H. Sung *et al.*, 'Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries', *CA Cancer J Clin*, vol. 71, no. 3, pp. 209–249, May 2021, doi: 10.3322/caac.21660.

[2] J. Masrour-Roudsari and S. Ebrahimpour, 'Causal role of infectious agents in cancer: An overview', 2017, *Babol University of Medical Sciences*. doi: 10.22088/cjim.8.3.153.

[3] 'www.statssa.gov.za'.

[4] A. Ankomah *et al.*, 'ART access-related barriers faced by HIV-positive persons linked to care in southern Ghana: A mixed method study', *BMC Infect Dis*, vol. 16, no. 1, Dec. 2016, doi: 10.1186/s12879-016-2075-0.

[5] J. T. Navarro, J. Moltó, G. Tapia, and J. M. Ribera, 'Hodgkin lymphoma in people living with hiv', *Cancers (Basel)*, vol. 13, no. 17, Sep. 2021, doi: 10.3390/cancers13174366.

[6] A. Noy, 'Optimizing treatment of HIV-associated lymphoma', 2019. [Online]. Available: http://ashpublications.org/blood/article-pdf/134/17/1385/1502392/bloodbld2018791400c.pdf

[7] T. Chiba and H. Marusawa, 'A novel mechanism for inflammation-associated carcinogenesis; an important role of activation-induced cytidine deaminase (AID) in mutation induction.', 2009. doi: 10.1007/s00109-009-0527-3.

[8] A. Carbone, C. Tripodo, C. Carlo-Stella, A. Santoro, and A. Gloghini, 'The Role of Inflammation in Lymphoma', 2014, pp. 315–333. doi: 10.1007/978-3-0348-0837-8_12.

[9] D. Hanahan, 'Hallmarks of Cancer: New Dimensions', *Cancer Discov*, vol. 12, no. 1, pp. 31–46, Jan. 2022, doi: 10.1158/2159-8290.CD-21-1059.

[10] D. Hanahan and R. A. Weinberg, 'The Hallmarks of Cancer', *Cell*, vol. 100, no. 1, pp. 57–70, Jan. 2000, doi: 10.1016/S0092-8674(00)81683-9.

[11] D. J. Wood and J. A. Endicott, 'Structural insights into the functional diversity of the CDK-cyclin family.', *Open Biol*, vol. 8, no. 9, Sep. 2018, doi: 10.1098/rsob.180112.

[12] F. R. Greten and S. I. Grivennikov, 'Inflammation and Cancer: Triggers, Mechanisms, and Consequences.', *Immunity*, vol. 51, no. 1, pp. 27–41, Jul. 2019, doi: 10.1016/j.immuni.2019.06.025.

[13] D. Hanahan and R. A. Weinberg, 'Hallmarks of cancer: The next generation', Mar. 04, 2011. doi: 10.1016/j.cell.2011.02.013.

[14] D. Hanahan, 'Hallmarks of Cancer: New Dimensions', Jan. 01, 2022, *American Association for Cancer Research Inc.* doi: 10.1158/2159-8290.CD-21-1059.

[15] M. Sinkala, 'Mutational landscape of cancer-driver genes across human cancers', *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-39608-2.

[16] M. Nourbakhsh, K. Degn, A. Saksager, M. Tiberti, and E. Papaleo, 'Prediction of cancer driver genes and mutations: the potential of integrative computational frameworks', *Brief Bioinform*, vol. 25, no. 2, Jan. 2024, doi: 10.1093/bib/bbad519.

[17] M. Muramatsu *et al.*, 'Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells', *Journal of Biological Chemistry*, vol. 274, no. 26, pp. 18470–18476, Jun. 1999, doi: 10.1074/jbc.274.26.18470.

[18] M. Muramatsu, K. Kinoshita, S. Fagarasan, S. Yamada, Y. Shinkai, and T. Honjo, 'Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme', 2000.

[19] P. Revy *et al.*, 'Activation-Induced Cytidine Deaminase (AID) Deficiency Causes the Autosomal Recessive Form of the Hyper-IgM Syndrome (HIGM2) HIGM2 patients (and in AID / mice) demonstrates the absolute requirement for AID in several crucial steps of B cell terminal differentiation necessary for efficient antibody responses. Production of highly efficient neutralizing antibodies re-Kazuo Kinoshita, 2 Tasuku Honjo, 2 quires affinity maturation of antibody responses. This', 2000.

[20] M. Teater *et al.*, 'AICDA drives epigenetic heterogeneity and accelerates germinal center-derived lymphomagenesis', *Nat Commun*, vol. 9, no. 1, Dec. 2018, doi: 10.1038/s41467-017-02595-w.

[21] D. P. Muñoz *et al.*, 'Activation-induced cytidine deaminase (AID) is necessary for the epithelial-mesenchymal transition in mammary epithelial cells', *Proc Natl Acad Sci U S A*, vol. 110, no. 32, Aug. 2013, doi: 10.1073/pnas.1301021110.

[22] Y. Matsumoto *et al.*, 'Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium', *Nat Med*, vol. 13, no. 4, pp. 470–476, Apr. 2007, doi: 10.1038/nm1566.

[23] T. Matsumoto *et al.*, 'Hepatic inflammation facilitates transcription-associated mutagenesis via AID activity and enhances liver tumorigenesis Running title: Inflammation-enhanced mutagenesis and liver cancer'.

[24] C. Tang, D. Bagnara, N. Chiorazzi, M. D. Scharff, and T. MacCarthy, 'AID Overlapping and Polη Hotspots Are Key Features of Evolutionary Variation Within the Human Antibody Heavy Chain (IGHV) Genes', *Front Immunol*, vol. 11, Apr. 2020, doi: 10.3389/fimmu.2020.00788.

[25] I. B. Rogozin and M. Diaz, 'Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process.', *J Immunol*, vol. 172, no. 6, pp. 3382–4, Mar. 2004, doi: 10.4049/jimmunol.172.6.3382.

[26] A. Jaiswal, R. Roy, A. Tamrakar, A. K. Singh, P. Kar, and P. Kodgire, 'Activation-induced cytidine deaminase an antibody diversification enzyme interacts with chromatin modifier UBN1 in B-cells', *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-46448-7.

[27] L. Alves and S. Rios, ‘Understanding the molecular pathogenesis of HIV-associated Burkitt Lymphoma-the impact of HIV-1 protein Tat on lymphoma driver genes’, 2020.

[28] J. A. Hackney *et al.*, ‘DNA targets of AID evolutionary link between antibody somatic hypermutation and class switch recombination.’, *Adv Immunol*, vol. 101, pp. 163–89, 2009, doi: 10.1016/S0065-2776(08)01005-5.

[29] Y. Feng, N. Seija, J. M. Di Noia, and A. Martin, ‘AID in Antibody Diversification: There and Back Again’, Jul. 01, 2020, *Elsevier Ltd*. doi: 10.1016/j.it.2020.04.009.

[30] H. E. Krokan and M. Bjørås, ‘Base excision repair.’, *Cold Spring Harb Perspect Biol*, vol. 5, no. 4, p. a012583, Apr. 2013, doi: 10.1101/cshperspect.a012583.

[31] L. Alves and S. Rios, ‘Understanding the molecular pathogenesis of HIV-associated Burkitt Lymphoma-the impact of HIV-1 protein Tat on lymphoma driver genes’, 2020.

[32] J. A. Roco *et al.*, ‘Class-Switch Recombination Occurs Infrequently in Germinal Centers’, *Immunity*, vol. 51, no. 2, pp. 337-350.e7, Aug. 2019, doi: 10.1016/j.immuni.2019.07.001.

[33] J. Stavnezer, J. E. J. Guikema, and C. E. Schrader, ‘Mechanism and Regulation of Class Switch Recombination’, *Annu Rev Immunol*, vol. 26, no. 1, pp. 261–292, Apr. 2008, doi: 10.1146/annurev.immunol.26.021607.090248.

[34] Z. Xu *et al.*, ‘14-3-3 adaptor proteins recruit AID to 5’-AGCT-3’-rich switch regions for class switch recombination’, *Nat Struct Mol Biol*, vol. 17, no. 9, pp. 1124–1135, Sep. 2010, doi: 10.1038/nsmb.1884.

[35] H. Zan and P. Casali, ‘Regulation of Aicda expression and AID activity’, 2013. doi: 10.3109/08916934.2012.749244.

[36] J. Stavnezer and C. E. Schrader, ‘IgH Chain Class Switch Recombination: Mechanism and Regulation’, *The Journal of Immunology*, vol. 193, no. 11, pp. 5370–5378, Dec. 2014, doi: 10.4049/jimmunol.1401849.

[37] C. T. Yan *et al.*, ‘IgH class switching and translocations use a robust non-classical end-joining pathway’, *Nature*, vol. 449, no. 7161, pp. 478–482, Sep. 2007, doi: 10.1038/nature06020.

[38] M. Epeldegui, E. Crabb Breen, Y. P. Hung, W. J. Boscardin, R. Detels, and O. Martínez-Maza, ‘Elevated expression of activation induced cytidine deaminase in peripheral blood mononuclear cells precedes AIDS-NHL diagnosis’. [Online]. Available: http://journals.lww.com/aidsonline

[39] F. B. Sall *et al.*, ‘HIV-1 Tat protein induces aberrant activation of AICDA in human B-lymphocytes from peripheral blood’, *J Cell Physiol*, vol. 234, no. 9, pp. 15678–15685, Sep. 2019, doi: 10.1002/jcp.28219.

[40] N. Mdletshe, A. Nel, K. Shires, and S. Mowla, ‘HIV Nef enhances the expression of oncogenic c-MYC and activation-induced cytidine deaminase in Burkitt lymphoma cells, promoting genomic instability’, *Infect Agent Cancer*, vol. 15, no. 1, Sep. 2020, doi: 10.1186/s13027-020-00320-9.

[41] G. GODSMARK, L. A. DE SOUZA RIOS, and S. MOWLA, 'Activation-induced cytidine deaminase promotes proliferation and enhances chemoresistance and migration in b-cell lymphoma', *Anticancer Res*, vol. 41, no. 1, pp. 237–247, Jan. 2021, doi: 10.21873/anticanres.14770.

[42] Y. Endo *et al.*, 'Expression of activation-induced cytidine deaminase in human hepatocytes via NF-κB signaling', *Oncogene*, vol. 26, no. 38, pp. 5587–5595, Aug. 2007, doi: 10.1038/sj.onc.1210344.

[43] Y. Endo *et al.*, 'Activation-Induced Cytidine Deaminase Links Between Inflammation and the Development of Colitis-Associated Colorectal Cancers', *Gastroenterology*, vol. 135, no. 3, 2008, doi: 10.1053/j.gastro.2008.06.091.

[44] F. Dedeoglu, B. Horwitz, J. Chaudhuri, F. W. Alt, and R. S. Geha, 'Induction of activation-induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NFκB', *Int Immunol*, vol. 16, no. 3, pp. 395–404, Mar. 2004, doi: 10.1093/intimm/dxh042.

[45] A. Takai *et al.*, 'Targeting activation-induced cytidine deaminase prevents colon cancer development despite persistent colonic inflammation', *Oncogene*, vol. 31, no. 13, pp. 1733–1742, Mar. 2012, doi: 10.1038/onc.2011.352.

[46] Y. Endo, H. Marusawa, and T. Chiba, 'Involvement of activation-induced cytidine deaminase in the development of colitis-associated colorectal cancers', *J Gastroenterol*, vol. 46, no. SUPPL. 1, pp. 6–10, Jan. 2011, doi: 10.1007/s00535-010-0326-1.

[47] L. A. de S. Rios, B. Cloete, and S. Mowla, 'Activation-induced cytidine deaminase: in sickness and in health', Nov. 01, 2020, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s00432-020-03348-x.

[48] A. R. Ramiro *et al.*, 'AID Is Required for c-myc/IgH Chromosome Translocations In Vivo', *Cell*, vol. 118, no. 4, pp. 431–438, Aug. 2004, doi: 10.1016/j.cell.2004.08.006.

[49] S.-R. Park, 'Activation-induced Cytidine Deaminase in B Cell Immunity and Cancers', *Immune Netw*, vol. 12, no. 6, p. 230, 2012, doi: 10.4110/in.2012.12.6.230.

[50] N. Mdletshe, A. Nel, K. Shires, and S. Mowla, 'HIV Nef enhances the expression of oncogenic c-MYC and activation-induced cytidine deaminase in Burkitt lymphoma cells, promoting genomic instability', *Infect Agent Cancer*, vol. 15, no. 1, Sep. 2020, doi: 10.1186/s13027-020-00320-9.

[51] Y. Dorsett, D. F. Robbiani, M. Jankovic, B. Reina-San-Martin, T. R. Eisenreich, and M. C. Nussenzweig, 'A role for AID in chromosome translocations between c-myc and the IgH variable region', *Journal of Experimental Medicine*, vol. 204, no. 9, pp. 2225–2232, Sep. 2007, doi: 10.1084/jem.20070884.

[52] A. R. Ramiro *et al.*, 'Role of genomic instability and p53 in AID-induced c-myc-Igh translocations', *Nature*, vol. 440, no. 7080, pp. 105–109, Mar. 2006, doi: 10.1038/nature04495.

[53] L. Pasqualucci *et al.*, 'AID is required for germinal center–derived lymphomagenesis', *Nat Genet*, vol. 40, no. 1, pp. 108–112, Jan. 2008, doi: 10.1038/ng.2007.35.

[54] P. Dharanipragada and N. Parekh, 'In Silico Identification and Functional Characterization of Genetic Variations across DLBCL Cell Lines', *Cells*, vol. 12, no. 4, Feb. 2023, doi: 10.3390/cells12040596.

[55] S. A. Atallah-Yunes, D. J. Murphy, and A. Noy, 'HIV-associated Burkitt lymphoma', *Lancet Haematol*, vol. 7, no. 8, pp. e594–e600, Aug. 2020, doi: 10.1016/S2352-3026(20)30126-5.

[56] J. William and Y.-H. Chen, '291 C-MYC Expression in Diffuse Large B-Cell Lymphoma and Correlation With Hans Classification and Proliferation index', 2011. [Online]. Available: https://academic.oup.com/ajcp/article/138/suppl_1/A291/1774089

[57] J. Zhang, Y. Shi, M. Zhao, H. Hu, and H. Huang, 'Activation-induced cytidine deaminase overexpression in double-hit lymphoma: potential target for novel anticancer therapy', *Sci Rep*, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-71058-y.

[58] E. Çakan and G. Gunaydin, 'Activation induced cytidine deaminase: An old friend with new faces', Oct. 27, 2022, *Frontiers Media S.A.* doi: 10.3389/fimmu.2022.965312.

[59] R. Kumar, L. J. DiMenna, J. Chaudhuri, and T. Evans, 'Biological function of activation-induced cytidine deaminase (AID)', Sep. 01, 2014, *Chang Gung Medical Journal*. doi: 10.4103/2319-4170.128734.

[60] A. R. Ramiro and V. M. Barreto, 'Activation-induced cytidine deaminase and active cytidine demethylation', Mar. 01, 2015, *Elsevier Ltd*. doi: 10.1016/j.tibs.2015.01.006.

[61] M. Teater *et al.*, 'AICDA drives epigenetic heterogeneity and accelerates germinal center-derived lymphomagenesis', *Nat Commun*, vol. 9, no. 1, Dec. 2018, doi: 10.1038/s41467-017-02595-w.

[62] J. Jiao *et al.*, 'AID and TET2 co-operation modulates FANCA expression by active demethylation in diffuse large B cell lymphoma', *Clin Exp Immunol*, vol. 195, no. 2, pp. 190–201, Feb. 2019, doi: 10.1111/cei.13227.

[63] J. Jiao *et al.*, 'AID assists DNMT1 to attenuate BCL6 expression through DNA methylation in diffuse large B-cell lymphoma cell lines', *Neoplasia (United States)*, vol. 22, no. 3, pp. 142–153, Mar. 2020, doi: 10.1016/j.neo.2020.01.002.

[64] Y. Wang, Y. Miao, W. Zhou, Y. Bi, Y. Ji, and Y. Ma, 'Activation-induced cytidine deaminase displays an alternative co-factor for modulating PIM1 expression in diffuse large B cell lymphoma cell lines', *Cell Mol Biol*, vol. 69, no. 3, pp. 1–7, 2023, doi: 10.14715/cmb/2023.69.3.1.

[65] E. L. Fritz, B. R. Rosenberg, K. Lay, A. Mihailović, T. Tuschl, and F. N. Papavasiliou, 'A comprehensive analysis of the effects of the deaminase AID on the transcriptome and methylome of activated B cells', *Nat Immunol*, vol. 14, no. 7, pp. 749–755, Jul. 2013, doi: 10.1038/ni.2616.

[66] M. Goel, P. Khanna, and J. Kishore, 'Understanding survival analysis: Kaplan-Meier estimate', *Int J Ayurveda Res*, vol. 1, no. 4, p. 274, 2010, doi: 10.4103/0974-7788.76794.

[67] A. Do Minh *et al.*, 'Characterization of Extracellular Vesicles Secreted in Lentiviral Producing HEK293SF Cell Cultures', *Viruses*, vol. 13, no. 5, p. 797, Apr. 2021, doi: 10.3390/v13050797.

[68] J. Cortés-Ríos *et al.*, 'Protein quantification by bicinchoninic acid (BCA) assay follows complex kinetics and can be performed at short incubation times', *Anal Biochem*, vol. 608, p. 113904, Nov. 2020, doi: 10.1016/j.ab.2020.113904.

[69] L.-M. Yin, Y. Wei, Y. Wang, Y.-D. Xu, and Y.-Q. Yang, 'Long term and standard incubations of WST-1 reagent reflect the same inhibitory trend of cell viability in rat airway smooth muscle cells.', *Int J Med Sci*, vol. 10, no. 1, pp. 68–72, 2013, doi: 10.7150/ijms.5256.

[70] J. B. Moroney, D. P. Chupp, Z. Xu, H. Zan, and P. Casali, 'Epigenetics of the antibody and autoantibody response', *Curr Opin Immunol*, vol. 67, pp. 75–86, Dec. 2020, doi: 10.1016/j.coi.2020.09.004.

[71] M. Teater *et al.*, 'AICDA drives epigenetic heterogeneity and accelerates germinal center-derived lymphomagenesis', *Nat Commun*, vol. 9, no. 1, Dec. 2018, doi: 10.1038/s41467-017-02595-w.

[72] G. GODSMARK, L. A. DE SOUZA RIOS, and S. MOWLA, 'Activation-Induced Cytidine Deaminase Promotes Proliferation and Enhances Chemoresistance and Migration in B-cell Lymphoma', *Anticancer Res*, vol. 41, no. 1, pp. 237–247, Jan. 2021, doi: 10.21873/anticanres.14770.

[73] L. A. de S. Rios, B. Cloete, and S. Mowla, 'Activation-induced cytidine deaminase: in sickness and in health', Nov. 01, 2020, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s00432-020-03348-x.

[74] L. Li *et al.*, 'Activation-induced cytidine deaminase expression in colorectal cancer.', *Int J Clin Exp Pathol*, vol. 12, no. 11, pp. 4119–4124, 2019.

[75] Q. Li *et al.*, 'Gut Barrier Dysfunction and Bacterial Lipopolysaccharides in Colorectal Cancer', Jul. 01, 2023, *Springer*. doi: 10.1007/s11605-023-05654-4.

[76] N. Mdletshe, A. Nel, K. Shires, and S. Mowla, 'HIV Nef enhances the expression of oncogenic c-MYC and activation-induced cytidine deaminase in Burkitt lymphoma cells, promoting genomic instability', *Infect Agent Cancer*, vol. 15, no. 1, Sep. 2020, doi: 10.1186/s13027-020-00320-9.

[77] M. Teater *et al.*, 'AICDA drives epigenetic heterogeneity and accelerates germinal center-derived lymphomagenesis', *Nat Commun*, vol. 9, no. 1, Dec. 2018, doi: 10.1038/s41467-017-02595-w.

[78] N. Zhang, Y. Yin, S.-J. Xu, and W.-S. Chen, '5-Fluorouracil: mechanisms of resistance and reversal strategies.', *Molecules*, vol. 13, no. 8, pp. 1551–69, Aug. 2008, doi: 10.3390/molecules13081551.

[79] J. Jiao *et al.*, 'AID assists DNMT1 to attenuate BCL6 expression through DNA methylation in diffuse large B-cell lymphoma cell lines', *Neoplasia (United States)*, vol. 22, no. 3, pp. 142–153, Mar. 2020, doi: 10.1016/j.neo.2020.01.002.

[80] B. Bertani and N. Ruiz, 'Function and Biogenesis of Lipopolysaccharides.', *EcoSal Plus*, vol. 8, no. 1, Aug. 2018, doi: 10.1128/ecosalplus.ESP-0001-2018.

[81] X. Wu *et al.*, 'Lipopolysaccharide promotes metastasis via acceleration of glycolysis by the nuclear factor-κB/snail/hexokinase3 signaling axis in colorectal cancer', *Cancer Metab*, vol. 9, no. 1, Dec. 2021, doi: 10.1186/s40170-021-00260-x.

[82] Y. Yang *et al.*, 'The emerging role of toll-like receptor 4 in myocardial inflammation', 2016, *Nature Publishing Group*. doi: 10.1038/cddis.2016.140.

[83] M. Capuozzo *et al.*, 'p53: From Fundamental Biology to Clinical Applications in Cancer.', *Biology (Basel)*, vol. 11, no. 9, Sep. 2022, doi: 10.3390/biology11091325.

[84] T. Shimizu, H. Marusawa, Y. Endo, and T. Chiba, 'Inflammation-mediated genomic instability: Roles of activation-induced cytidine deaminase in carcinogenesis', Jul. 2012. doi: 10.1111/j.1349-7006.2012.02293.x.

[85] R. A. Weinberg, 'The Biology of Cancer Chapter 8: pRb and Control of the Cell Cycle Clock', 2007.

[86] Y. Wang, Y. Miao, W. Zhou, Y. Bi, Y. Ji, and Y. Ma, 'Activation-induced cytidine deaminase displays an alternative co-factor for modulating PIM1 expression in diffuse large B cell lymphoma cell lines', *Cell Mol Biol*, vol. 69, no. 3, pp. 1–7, 2023, doi: 10.14715/cmb/2023.69.3.1.

[87] H. Meng *et al.*, 'DNA methylation, its mediators and genome integrity', *Int J Biol Sci*, vol. 11, no. 5, pp. 604–617, 2015, doi: 10.7150/ijbs.11218.

[88] M. Soltani, M. Abbaszadeh, H. Fouladseresht, M. J. M. Sullman, and N. Eskandari, 'PD-L1 importance in malignancies comprehensive insights into the role of PD-L1 in malignancies: from molecular mechanisms to therapeutic opportunities', *Clin Exp Med*, vol. 25, no. 1, p. 106, Apr. 2025, doi: 10.1007/s10238-025-01641-y.

Appendix A

Recipes and Reagents

1 % Agarose gel in 50 ml

Agarose 0.5 g plus 1X TBE 50 ml

Pulse heat in a microwave with gentle agitation, before pouring add 2 µl Ethidium Bromide

(10 mg/ml stock, Sigma- Aldrich)

LB (Luria Broth) Agar in 100 mL

Agar 1.5 g

LB powder 2 g

Add 100 mL of distilled water

Mix LB powder, agar will dissolve during autoclaving

Trypsin-EDTA (for TC) in 1 L

8 g NaCL

1.26 g Na_2HPO_4

0.2 g KCL

0.2 g KH_2PO_4

0.5 g Trypsin

0.2 g EDTA

Autoclave: beaker, stirrer bar, measuring cylinder, bottles & distilled water

Measure the required amounts (above), make up to volume & pH to 7.4.

Filter sterilizes in tissue culture hood using 0.2 micro filters into autoclaved bottles.

Store stock at -20^oC & working solution at 4^oC

2x Boiling Blue solution in 10 ml

1.25 ml of 1 M Tris-HCL pH 6.8

4 ml of 10% SDS

1 ml of β-mercaptoethanol

2 ml of Glycerol

1.75 ml of distilled water

Pinch of bromophenol blue

Mix and store at -20^{o}C

RIPA buffer

0.5 g (1%) deoxycholate powder in 40 ml distilled water

1.5 ml 5M NaCl (150 mM)

0.5 ml Triton X100 (1%)

0.25 ml 20% SDS (0.1%)

0.5 ml 1M Tris **(**pH 7.5) (10 mM)

Top up to 50 ml with distilled water

Store at 4°C

RIPA cell lysis buffer solution

423 µl RIPA buffer

71 µl 7X protease inhibitor

Store on ice until use

10 % SDS

10 g SDS

80 ml ddH2O

Dissolve and make up to 100 ml distilled water

5 X SDS Protein loading dye.

10% sodium dodecyl-sulphate (SDS) – 1 g

0.04% bromophenol blue – 0.004 g

4.83 ml of ddH2O

1.5 M Tris (pH 6.8) - 1.67 ml

100% glycerol – 3 ml

0.5 ml β-mercaptoethanol

10X SDS-PAGE running buffer.

10 g SDS

30.3 g Tris

144.1 g glycine

800 ml distilled water

Adjust to 1 L of distilled water and store at room temperature.

1X SDS-PAGE running buffer.

100 ml of 10X SDS-PAGE running buffer stock

900 ml distilled water

Store at room temperature.

10X SDS-PAGE transfer buffer

38 g Tris

144 g glycine

800 ml distilled water.

Adjust volume to 1 L with distilled water and store at room temperatur**e.**

1X SDS-PAGE transfer buffer

100 ml 10X SDS-PAGE transfer buffer stock

700 ml distilled water

200 ml Isopropanol

Make in advance and store at 4°C.

30% acryl-bisacrylamide

29 g acrylamide

1 g N.N'-methylenbisacrylamide

60 ml distilled water

Heat the solution to 37°C.

Adjust volume to 100 ml and cover with foil to protect from light Store at 4°C.

8% Resolving gel

3.3 ml distilled water

30% acryl-bisacrylamide – 2.7 ml

1 M Tris pH 6.8 – 3.9 ml

10% SDS – 0.1 ml

10% APS – 0.1 ml

TEMED – 0.004 ml

12% Resolving gel

2.4 ml distilled water

30% acryl-bisacrylamide – 3 ml

1 M Tris pH 6.8 – 1,9 ml

10% SDS – 0.075 ml

10% APS – 0.075 ml

TEMED – 0.003 ml

5% Fat-free milk in 1X TBS-Tween (0.1%) or 1XPBS-Tween (0.1%)

41 ml Fat-free milk

Up to 100 mL with TBST or PBST

5% BSA in 1X TBS-Tween (0.1%) or 1XPBS-Tween (0.1%) in 50 ml

2.5g BSA

50 ml TBST or PBST

Fixing solution (50 ml)

Glacial acetic acid: Methanol 1:3 (Keep at 4°C)

17 ml Glacial acetic acid

33 ml Methanol

1 X PBS-Tween

100 ml 10X PBS

0.5 ml Tween-20

Make up to 1 L with distilled water

10 X PBS

40 g NaCl

1 g KCl

1 g KH2PO4

5.75 g Na2HPO4.2H2O

2.1 g KH2PO4

Up to 500 ml with distilled water

pH of 7.4, autoclave

Ponceau S stain

0.1 g of Ponceau powder

95 ml of distilled water

5 ml of glacial acetic acid

7X Protease inhibitor

1 protease inhibitor tablet

2.5 ml 1X PBS

Store at -20°C

Stripping Buffer in 100 ml

0.7 ml of B-Merceptoethanol

20 ml of SDS

4.2 ml of Tris-HCL (pH 6.8)

75.1 ml of Distilled water

Store at room temperature

Appendix B

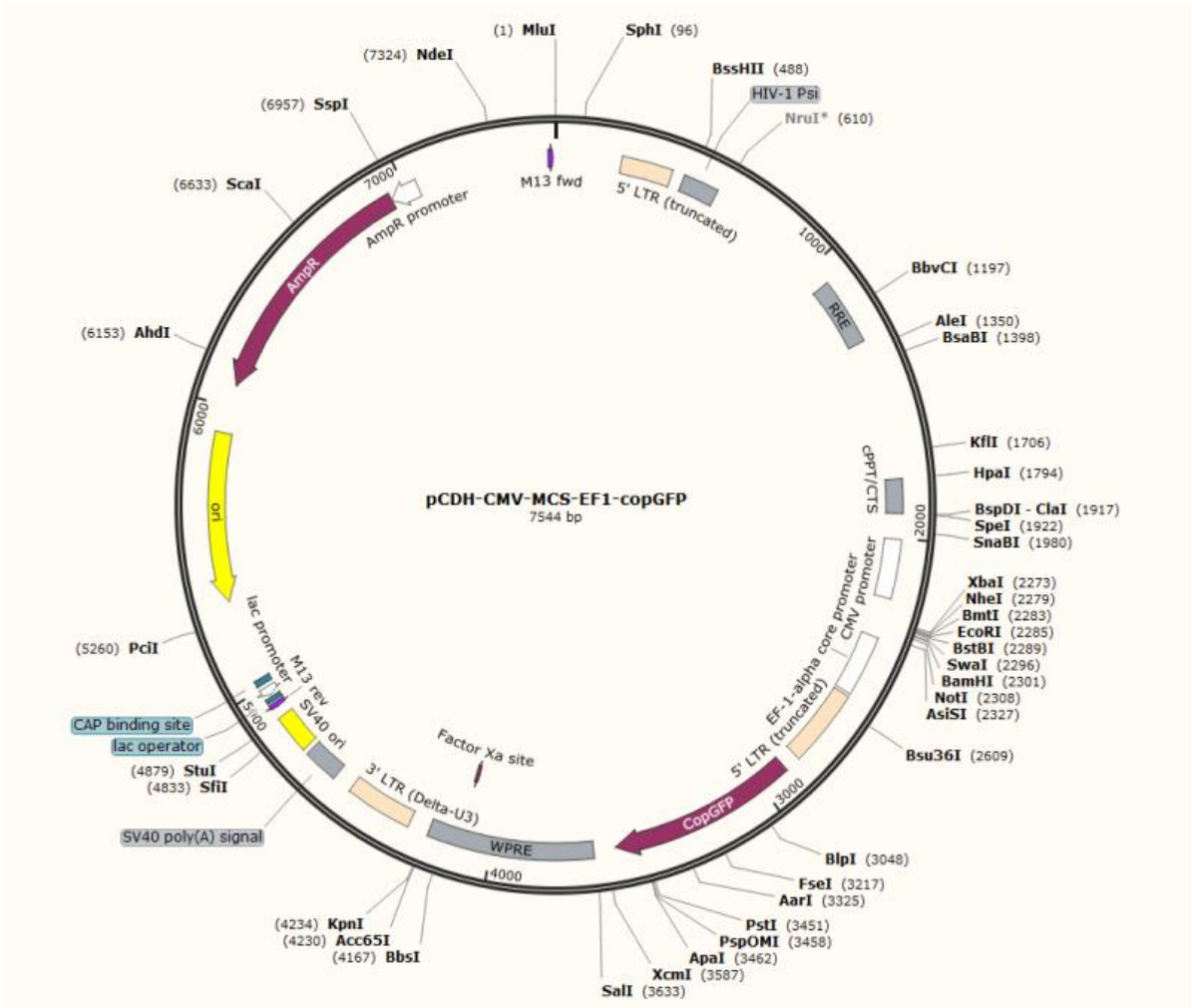


Figure B1: The backbone of pCDH-CMV-EF1-copGFP vector. The human AICDA was cloned into this vector.

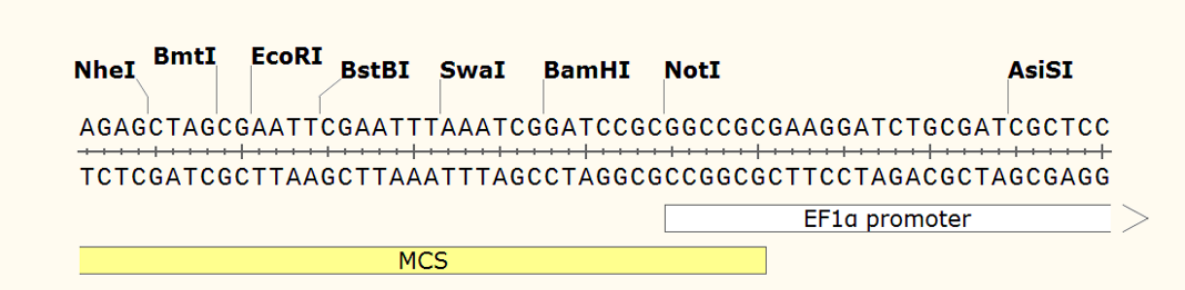


Figure A2: The schematic diagram showing the MSC region of the pCDH-CMV-MSC-EF1-copEFP map. The restriction sites for NheI and EcoRI are separated by 6 base pairs.

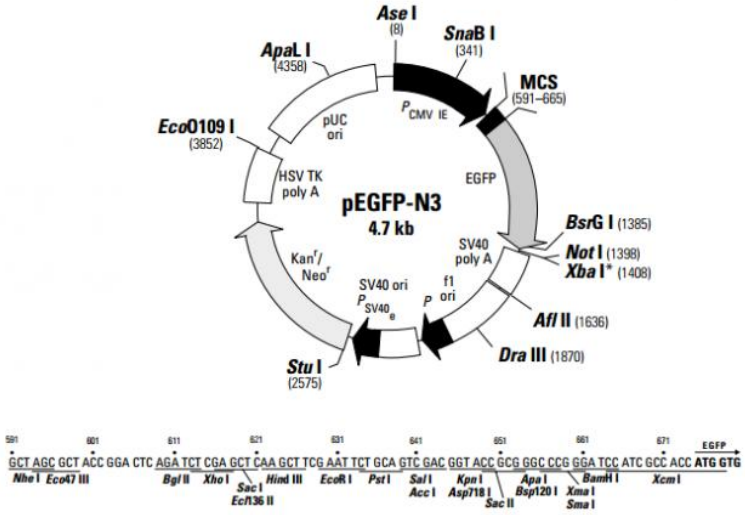


Figure B3: The backbone map of the pEGFP-N3 vector**.**

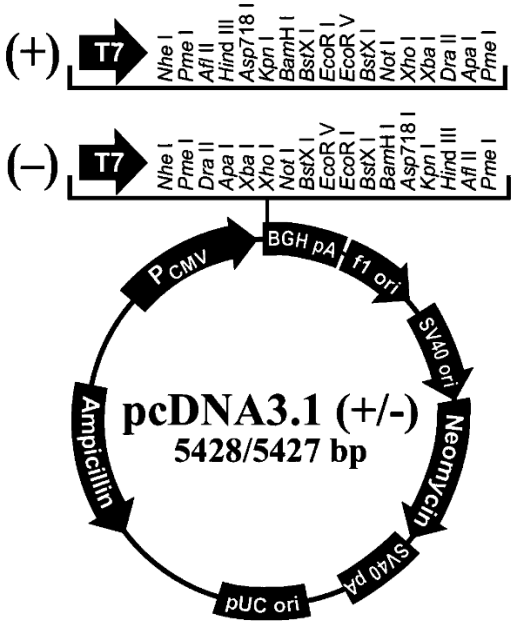


Figure B4: The backbone map of pcDNA 3.1 vector.

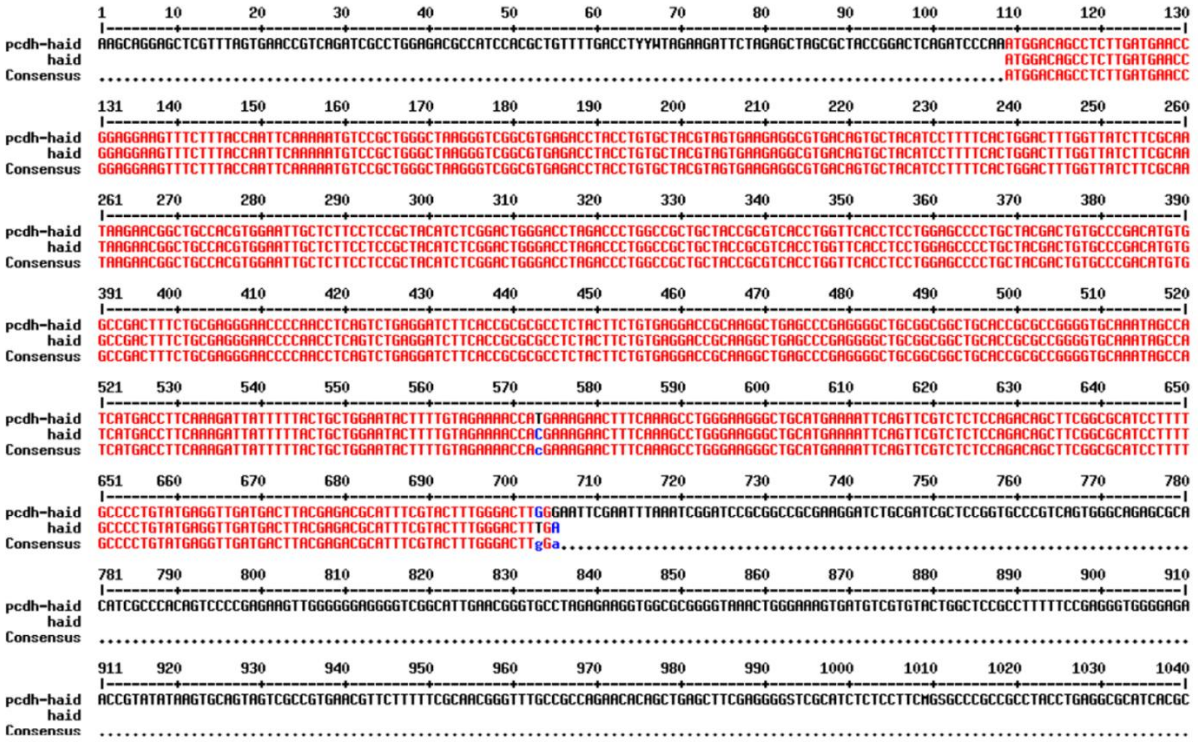


Figure B5: Multiple alignment of the cloned pCDH-CMV-MSC-EF1-copGFP-AICDA. The multiple alignment Sanger sequence data from pCDH-CMV-MSC-EF1-copGFP- AICDA with the reference AICDA sequence.

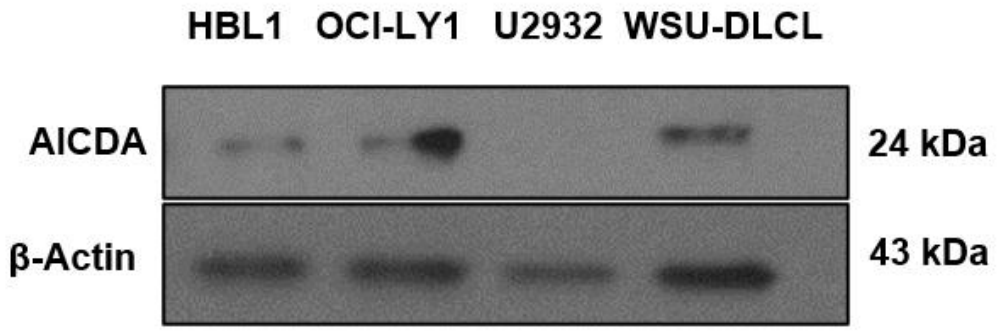


Figure B6: Screening of AICDA expression level in B cell lymphoma cell lines. Western blotting analysis showing AICDA expression level. β-actin was used as the control. (Adopted from Ms T. Khubeka)

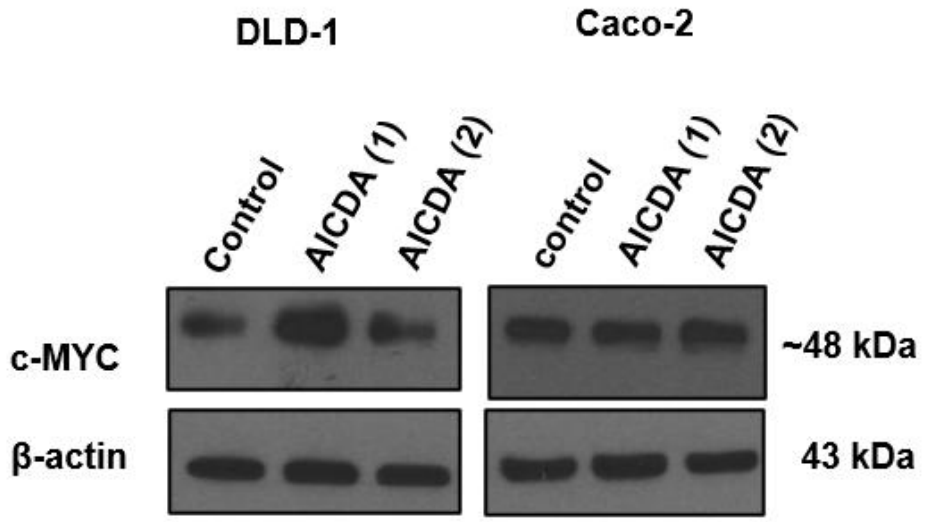


Figure B7. AICDA does not affect the c-MYC protein level in CRC cells. Western blotting analysis using an antibody against c-MYC to validate the protein expression of c-MYC in CRC cells expressing AICDA relative to control. β-actin was used as a loading control.

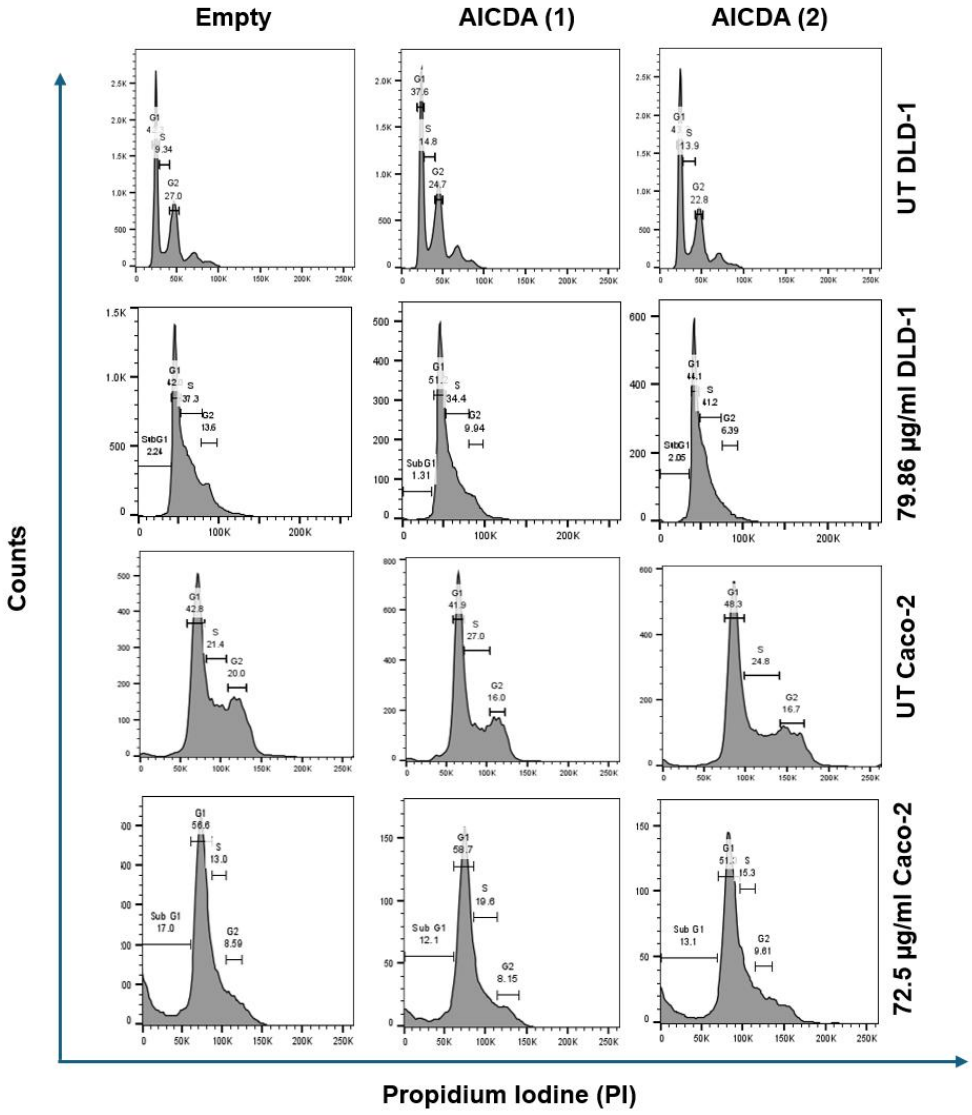


Figure B8: AICDA and 5-FU modify the cell cycle profile. DLD-1 and Caco-2 cells were plated and treated with 5-FU (79.86 ug/ml and 72.5 ug/ml, respectively) for 48 hours. The cells were washed with 1XPBS, fixed with 70% methanol, and stained with the FxCycleTM PI/RNase. The cells were analysed using flow cytometry.

Appendix C

Research Output during the MSc research period

1. **Lincon K. Makofane***., Beatrice Ramorola., Leonardo Alves de Souza Rios., Tarryn Willmer., Shaheen Mowla. The impact of lymphoma factor Activation Induced Cytidine Deaminase (AICDA/AID) on epigenetic plasticity. Postgraduate Academic Training program, 27 June 2023, University of Cape Town, (Oral presentation)
2. **Lincon K. Makofane***., Beatrice Ramorola., Leonardo Alves de Souza Rios., Tarryn Willmer., Shaheen Mowla. The impact of lymphoma factor Activation Induced Cytidine Deaminase (AICDA/AID) on epigenetic plasticity. The South African Society of Biochemistry and Molecular Biology conference, 7 – 10 July 2024, Protea Hotel Ranch Resort, Polokwane (Poster presentation).
3. **Lincon K. Makofane***., Beatrice Ramorola., Leonardo Alves de Souza Rios., Tarryn Willmer., Shaheen Mowla. The impact of lymphoma factor Activation Induced Cytidine Deaminase (AICDA/AID) on epigenetic plasticity. 16th Annual Haematology Oncology Symposium, 13 – 15 September 2024, Century City, Cape Town (Poster presentation). Award for best poster presentation.

Appendix D

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

E53-Room 46 Old Main Building
Groote Schuur Hospital
Observatory 7925
Email: hrec-enquiries@uct.ac.za
Website: https://health.uct.ac.za/home/human-research-ethics

20 January 2025

HREC REF: C001/2025

A/Prof Shaheen Mowla
Department of Haematology
Faculty of Health Sciences, UCT
Email PI: Shaheen.mowla@uct.ac.za
Student email: makofanelincon@gmail.com

Dear A/Prof Mowla

PROJECT TITLE: AN INVESTIGATION OF THE IMPACT OF ACTIVATION INDUCED CYTIDINE DEAMINASE (AICDA/AID) OVEREXPRESSION ON THE EPIGENOME OF AGGRESSIVE B-CELL LYMPHOMAS OF HIGH PREVALENCE AMONG HIV-INFECTED INDIVIDUALS. (MASTER DEGREE – MR LINCOLN MAKOFANE)

Thank you for submitting your FHS036 application to the Faculty of Health Sciences Human Research Ethics Committee for review and approval.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study which includes HEK293T, HBL1, PB.B95-8H, Caco-2 and DLD1 cells.

Approval is granted for one year until the 30 January 2026.

The HREC acknowledges that the Master's degree student: ***Mr Lincoln Makofane*** *will also be involved in the study.*

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the wine alo we oeference Number C001/2025 in all your correspondence.

Yours sincett

PROFESE MAN
CHAIRPERI OF HEALTH SCIENOES HUMAN RESEARCH ETHICS COMMITTEE