Investigating the DNA methylation status of the *PXDN* and *PXDNL* promoter regions in OSCC cell lines

by

Mistral Sebastian

(1852774)

Dissertation

Submitted in fulfilment of the requirements for the degree

Master of Science

in

Molecular and Cell Biology

in the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

Supervisor: Prof. Demetra Mavri-Damelin

July 2024

Declaration

I, Mistral Sebastian (1852774), am a student registered for the degree of Master of Science in the academic year 2024.

I hereby declare the following:

• I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.

• I confirm that the research dissertation submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged.

• I have followed the required conventions in referencing the thoughts and ideas of others.

• I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signature  Date: 10 July 2024

Abstract

Background: Oesophageal squamous cell carcinoma (OSCC) is the most prevalent form of oesophageal cancer in South Africa. Aberrant DNA methylation is a well-established epigenetic mechanism involved in various cancers, including OSCC. This study focuses on the DNA methylation status of the *peroxidasin* (*PXDN*) and *perodixasin like* (*PXDNL*) promoter regions and the expression of PXDN and PXDNL in OSCC cell lines. PXDN consolidates the basement membrane through collagen IV unit oligomerization, influences epithelial-mesenchymal transition and correlates with poor prognosis in various cancers. PXDNL modulates the extracellular matrix (ECM) by antagonising PXDN. Since PXDNL shares domains with PXDN, that allow PXDN to interact with the ECM, it is speculated that PXDNL may possess other ECM modulation roles that require further elucidation. Dysregulated PXDNL expression also correlates with poor cancer prognosis. To date, within the context of South African derived OSCC cell lines, no studies pertaining to the DNA methylation status of the *PXDN* and *PXDNL* promoter regions and the expression of PXDN and PXDNL have been carried out.

Aim: The aim of this project was to investigate the DNA methylation status of the *PXDN* and *PXDNL* promoter regions and observe PXDN and PXDNL expression in the SNO and WHCO5 OSCC cell lines.

Methods: PXDN and PXDNL localisation was observed using immunofluorescence microscopy; expression of PXDN and PXDNL was quantified using western blotting and the DNA methylation status of the *PXDN* and *PXDNL* promoters was assessed using methylation specific PCR and bisulfite sequencing, respectively.

Results: Immunofluorescence microscopy results indicated that both cell lines show varying degrees of PXDN and PXDNL expression. In addition, these results also showed that PXDN and PXDNL localise in the ECM. The western blotting results established that these cell lines express the canonical version of PXDN and possibly a PXDNL isoform (146kDa). Methylation specific PCR has shown that the promoter region of *PXDN* is differentially methylated across both cell lines.

The sequencing results of the bisulfite converted *PXDNL* promoter region were unsuccessful. Hence, bisulfite sequencing requires further optimisation before the DNA methylation status of the *PXDNL* promoter region can be determined.

Conclusion: This study is the first to show the novel finding that PXDN and PXDNL are expressed in South African derived OSCC cell lines. Within the context of OSCC, further investigation is warranted in order to elucidate the underlying mechanisms that these proteins play a role in. In addition, further study may determine whether a correlation exists between *PXDN* and *PXDNL* promoter methylation, protein expression as well as prognosis and whether these aspects could serve as novel markers for diagnosis and therapy. This may subsequently lead to increased OSCC patient survival rates by contributing to early diagnosis of OSCC and efficacious targeted therapeutic intervention.

Dedication

To my parents, Brino Sebastian and Managee Sebastian, who always knew that that I would make them proud.

Acknowledgements

Above all else, I would like to show my utmost gratitude and thanks to God. I would not have been able to complete this degree without the love, wisdom and comfort that comes only from the Father, the Son and the Holy Spirit.

To my parents, Brino Sebastian and Managee Sebastian, I know that I am loved by God because He decided to save the best parents in the world for me. Everything I am, I owe to the two of you. Thank you for your unconditional love and support, for prioritising my education and for constantly seeing the best in me.

To my best friends, Tekeyah, Thando and Keila, you are truly some of the best people that I know. Thank you for the abundance of love and laughter that you have given me throughout the course of my degree. Just like Aaron and Hur held up Moses' hands when he grew tired, you three have kept me going even through the hardest of times.

To my grandmothers who never let me leave home without praying for me and my experiments, I am eternally grateful to you both. To my grandfather who could not be here to see me complete this degree, I know that you would have been proud to see the Sebastian name on a master's dissertation. I am also thankful to Pastor Antonie and Aunty Mpho for their love, guidance, prayers and support.

To my incredible supervisor, Prof. Mavri-Damelin, thank you for all your guidance throughout the years. In a world where there are smart supervisors and kind supervisors, I am honoured to say that I had a supervisor who was the perfect blend of both. Your wisdom, patience and gentleness are traits that have allowed me to become the independent thinker that I am today.

To my GH516 family, Thoko, Jamie, Jem and Kayleen, I am so grateful to be surrounded by incredible women in STEM. The memories and laughs that were shared between us will always be dear to me.

I owe a great deal of thanks to Dr Reddy who always made time to help with any query I had.

Table of contents

List of figures

List of tables

List of abbreviations

Akt	-	protein kinase B
ANOVA	-	analysis of variance
APS	-	ammonium persulfate
ATP	-	adenosine triphosphate
Br^-	-	bromide
BSA	-	bovine serum albumin
CANSA	-	Cancer Association of South Africa
CDK	-	Cyclin dependant kinase
Cl^-	-	chloride
CpG	-	Cytosine-phospho-guanine
CTCF	-	corrected total cell fluorescence
DAPI	-	4′,6-diamidino-2-phenylindole
DMEM	-	Dulbecco's Modified Eagles Medium
DNA	-	deoxyribonucleic acid
DNMT	-	DNA methyltransferases
ECM	-	extracellular matrix
EDTA	-	ethylenediamine tetraacetic acid
EMT	-	epithelial to mesenchymal transition
ERK	-	extracellular signal-regulated kinases
FAK	-	focal adhesion kinase

FITC	-	fluorescein isothiocyanate
gDNA	-	genomic DNA
GST	-	Glutathione S-transferase
H_2O_2	-	hydrogen peroxide
HEPES	-	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HOBr	-	hypobromous acid
HOCl	-	hypochlorous acid
LRP	-	lipoprotein receptor-related protein
LVL	-	Lugol Voiding Lesion
M	-	methylated
MBD	-	Methyl-CpG-binding domain
MeCP2	-	Methyl CpG binding protein 2
		Member A
Mg2+	-	magnesium ions
MLH3	-	MutL Homolog 3
MMR	-	Mismatch repair
mRNA	-	messenger RNA
MSPCR	-	methylation sensitive polymerase chain reaction
NAC	-	No Antibody Control
NADPH	-	nicotinamide adenine dinucleotide phosphate
NEC	-	no enzyme control

NPC	-	No Primary Antibody Control
Nrf2	-	Nuclear factor erythroid 2–related factor 2
OSCC	-	Oesophageal squamous cell carcinoma
P1K3	-	phosphorylated phosphatidylinositol-3 kinase
PBS	-	phosphate-buffered saline
PCR	-	polymerase chain reaction
PGE2	-	Prostaglandin E2
PLB	-	protein loading buffer
PXDN	-	Peroxidasin
PXDNL	-	Peroxidasin like
RASSF10	-	RAS association domain family 10
RIPA	-	radioimmunoprecipitation
ROS	-	reactive oxygen species
RNA	-	ribonucleic acid
RT	-	room temperature
RT-qPCR	-	quantitative reverse transcription polymerase chain reaction
SD	-	standard deviation
SDS	-	sodium dodecyl sulfate
SDS-PAGE	-	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sin3A	-	Swi-independent 3 Transcription Regulator Family

TAE	-	Tris base, acetic acid and EDTA
TBST	-	tris-buffered saline with tween
TCA	-	trichloroacetic acid
TEMED	-	tetramethylethylenediamine
TET2	-	Ten-Eleven-Translocation 2
TRD	-	Transcriptional repression domain
Tris HCL	-	tris hydrochloride
Trypsin-EDTA	-	Trypsin-Ethylenediamine tetra acetic acid
TSS	-	transcription start site
UM	-	unmethylated
vWF	-	von Willebrand Factor

1. Introduction

1.1 Overview of cancer

Cancer is a disease that arises due to an accumulation of genetic and epigenetic abnormalities that instigate malignant tumour formation (Bréchot, 1998; Jones and Baylin, 2007; Knudson, 2002). While genetic abnormalities present as alterations to the DNA sequence, epimutations do not affect the DNA code and are defined as alterations in gene expression that are as a result of aberrant gain or loss of chromatin modifications such as DNA methylation (Hitchins, 2010). Within the context of cancer, the presence of mutated DNA can be due to a plethora of reasons such as inheritance, failure of DNA polymerases' proofreading ability during DNA replication, defects in DNA repair machinery genes such as DNA mismatch repair (MMR) genes and exposure to certain environmental or lifestyle factors (Brown, 2002; English and Armstrong, 1988; Imai and Yamamoto, 2008; Knudson, 1989; Wingo et al., 1999). Epimutations may be acquired due to inheritance, exposure to specific environmental and lifestyle factors as well as aberrations in cis or trans-acting epigenetic factors (Argos et al., 2015; Hitchins et al., 2011; Horsthemke, 2006; Ligtenberg et al., 2009; McCarrey et al., 2016; Zeilinger et al., 2013).

Malignant tumour cells often present with a complex myriad of both DNA mutations and epimutations which subsequently lead to the display of cancer hallmarks such as, anti-growth signal evasion, infinite growth capability, apoptotic signal insensitivity, tissue invasion and metastasis as well as deregulated metabolism (Cooper, 2000a; Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011). A few genes that contribute to cancer manifestation, when presenting with DNA mutations or epimutations, include proto-oncogenes, tumour suppressor genes as well as DNA repair genes (Bréchot, 1998; Cui et al., 1998; Dobrovic and Simpfendorfer, 1997; Knudson, 2002; Wang et al., 2003). These mutations may lead to aberrant gene expression and inappropriate cell cycle continuation despite mutational overload and apoptotic signals (Cooper, 2000b; Hanahan and Weinberg, 2000). Thus, leading to cancer causation and propagation.

Recent studies have shown that in 2020, there were 19.3 million new cancer cases as well as 10 million cancer related deaths reported globally (Sung et al., 2021).

Subsequently, cancer has been deemed the leading cause of mortality and decreased life expectancy in the global population (Bray et al., 2021). The increase in cancer related mortality is thought to be due to the aging of the population as well as the increasing global distribution of cancer associated risk factors (which are often associated with socioeconomic development) (Gersten and Wilmoth, 2002; Omran, 1971). In the global population, cancer shows a similar incidence in both sexes while the mortality rates differ significantly with men displaying 43% higher mortality rate in comparison to women (Sung et al., 2021). This aforementioned gender gap in the mortality rate is attributed to the variation in cancer type distribution (Sung et al., 2021). Recent statistics have shown that oesophageal cancer ranks amongst the most prominent type of cancer for both men and women (Sung et al., 2021).

1.2 Oesophageal cancer

Oesophageal cancer is a malignant neoplasia that localises in the portion of the gastrointestinal tract known as the gullet or oesophagus. Oesophageal cancer ranks amongst the most prominent forms of cancer worldwide with the seventh highest incidence (with an estimated 604100 new cases annually) and ranks sixth with regards to mortality (544076 deaths per year) (Sung et al., 2021). Furthermore, it has a two-fold higher incidence rate and three-fold higher mortality rate for men in comparison to women (Sung et al., 2021). Thus, oesophageal cancer is responsible for one in every 18 cancer related deaths (Recio-Boiles and Babiker, 2021; Sung et al., 2021). The phenotypic manifestations, geographic variation and sex distribution is heavily dependent on the subtype of oesophageal cancer.

1.2.1 Subtypes and symptoms of oesophageal cancer

There are two common subtypes of oesophageal cancer, namely adenocarcinoma and oesophageal squamous cell carcinoma (OSCC) that possess varying aetiologies (Blot et al., 1993). OSCC is a subtype of oesophageal cancer that is characterised by the perpetuation of malignant neoplasia specifically in the squamous cells that

line the oesophagus (Blot et al., 1993). While adenocarcinoma makes up only 30% of total oesophageal cancer cases globally, OSCC accounts for remainder of the cases and is therefore the most prevalent subtype of oesophageal cancer (Cook, 2011; Smyth et al., 2017). Individuals diagnosed with OSCC have shown an average post-onset life expectancy of five years (Recio-Boiles and Babiker, 2021). Although individuals with OSCC do not typically present with symptoms during the early onset of this disease, further development of OSCC can lead to symptoms such as dysphagia, severe chest pain, heartburn, dysphonia as well as persistent coughing (Smyth et al., 2017; Recio-Boiles and Babiker, 2021). These aforementioned clinical manifestations can be attributed to the histological presentation of OSCC.

1.2.2 Clinical presentation, geographical distribution and major risk factors of OSCC

The histological presentation of OSCC comprises of malignant Lugol-voiding lesions (LVLs) in the oesophageal canal (Mori et al., 1993). Although LVLs may manifest anywhere in the oesophagus, they are often localised in the upper and mid-portion of the oesophageal canal (Berry, 2014). The quantity of lesions varies depending on the severity of the oesophageal intraepithelial neoplasia. OSCC severity and the subsequent degree of OSCC malignancy can be elucidated by classifying the degree of primary LVL invasion and by investigating the extent of metastatic invasion (Sobin et al., 2009). Although the hypothesised causative genetic mechanisms and risk factors of OSCC may vary depending on geographic location, the aforementioned clinical presentation of OSCC remains identical in all individuals irrespective of their geographic region.

In developed countries, such as Australia and France, OSCC represents one third of total oesophageal cancer cases (Blot and Tarone, 2018). Major risk factors pertaining to the onset of OSCC in western countries include chronic alcohol consumption as well as exposure to tobacco smoke (Blot and Tarone, 2018). In recent years, first world countries have shown a rapid decline in OSCC manifestation, which has been hypothesised to be associated with the decrease in

exposure to tobacco smoke (Blot and Tarone, 2018). In contrast, OSCC has been shown to be the most prevalent form of oesophageal cancer in low-income regions where major risk factors of OSCC, excluding alcohol consumption and tobacco smoke, vary greatly depending on the geographical region (Alaouna et al., 2019; McCormack et al., 2017). Varying risk factors include insufficient dietary nutrition, consumption of hot beverages and exposure to nitrosamines (Alaouna et al., 2019; Blot and Tarone, 2018; Jakszyn and González, 2006; McCormack et al., 2017). These aforementioned low-income regions include parts of Asia and South Africa (Alaouna et al., 2019; Jakszyn and González, 2006; McCormack et al., 2017).

According to the Cancer Association of South Africa (CANSA), oesophageal cancer ranks eighth in terms of cancer incidence and accounts for 3.1% of total new cancer cases annually in South Africa. In addition, oesophageal cancer has the fifth highest mortality, contributing to 5.6% (3176 cases) of total cancer related deaths annually. As mentioned above, OSCC is the most prevalent type of oesophageal cancer in South Africa and exhibits an incidence of 46.7/100,000 and 19.2/100,000 in males and females, respectively (Loots et al., 2017). Major environmental risk factors of OSCC onset, specifically in the South African population, include a shift in traditional beer derivation and increased maize consumption (Isaacson, 2005; Pillay et al., 2015).

Previously, traditional beer that was brewed and consumed in South Africa was derived from sorghum. However, in recent years it has been found that traditional beer is more often derived from maize (Alaouna et al., 2019). This shift in beer derivation has been deemed a risk factor of OSCC onset as, unlike sorghum, maize is more likely to be contaminated with toxin producing fungi that stimulates the synthesis of carcinogenic N-nitrosamines (Isaacson, 2005; Moretti et al., 1995). Localisation of these carcinogenic N-nitrosamines in the oesophageal canal may elicit the onset of malignant neoplasia associated with OSCC (Lin et al., 2002). In addition to the shift in traditional beer derivation, the staple diet of the South African population has changed from heavily sorghum dependent to heavily maize dependent (Isaacson, 2005). An increase in dietary maize consumption in the South African population has been shown to lead to gastric fluid that is rich in the cell proliferation stimulant, prostaglandin E2 (PGE2) (Pink et al., 2011). Thus, exposure

of the oesophageal cells to this PGE2 rich gastric fluid may cause malignant neoplasia of oesophageal squamous cells (Pink et al., 2011). Although environmental risk factors of OSCC have been found, they are weak carcinogens and need to work in tandem with a genetic component in order to predispose an individual to OSCC. While the exact causative genetic mechanisms have not been elucidated, a few genetic risk factors that exhibit an association with the clinical manifestation of OSCC have been found. In the South African population, examples of these genetic risk factors include polymorphisms that have been noted in *glutathione S-transferase* (*GST*) genes as well as DNA MMR genes (Li et al., 2010; Vogelsang et al., 2012).

GST genes are responsible for the detoxification of carcinogenic compounds. The detoxification and subsequent removal of these carcinogens is achieved by GST proteins catalysing the conjugation of glutathione to these aforementioned electrophilic carcinogenic compounds (Marchewka et al., 2017). *GST* genes can be further divided into three subfamilies, namely *GSTA*, *GSTM*, *GSTT* and *GSTP*. In the South African population, the *GSTM1* null genotype has shown association with increased OSCC development in the Black Xhosa population while the *GSTP1* (341C/T polymorphism) shows an increased risk of OSCC onset in the Black Xhosa as well as the mixed ancestry population (Li et al., 2010). In addition, polymorphisms noted in DNA MMR genes, a group of genes that are responsible for protection against DNA damage induced by carcinogenic environmental factors, have been found to show an association with OSCC onset in the South African population. In the South African population, polymorphisms found in the *MutS Homolog 3*, *Postmeiotic Segregation Increased 1* and *MutL Homolog 3* (*MLH3*) genes have been shown to increase the susceptibility of the mixed ancestry population to OSSC (Vogelsang et al., 2012). In particular, the polymorphism found in the *MLH3* gene is likely to be associated with OSCC onset as the variation in this gene results in an amino acid substitution (Arg797His) that compromises the structure and subsequently the function of the MLH3 protein, and therefore, prevents this protein from adequately participating in DNA repair (Vogelsang et al., 2012). Other modifications found in the global population that may induce OSCC associated malignant neoplasia include epigenetic alterations (Taghavi et al., 2010).

1.3 DNA methylation as an epigenetic mechanism

DNA methylation is an epigenetic mechanism that is able to regulate an individual's phenotypic manifestations by altering the expression of certain genes without changing the DNA sequence (Singal and Ginder, 1999). Investigations of DNA methylation since its initial discovery in 1948 has led to this epigenetic mechanism being the most well studied epigenetic modification in eukaryotes (Kiselev et al., 2021; Singal and Ginder, 1999). This gene silencing associated epigenetic regulation is achieved by the transient addition of a methyl group to the fifth carbon present on the pyrimidine ring known as cytosine that is proceeded by a guanine. This results in the formation of a 5′-methylcytosine and is catalysed by enzymes known as DNA methyltransferases (DNMTs) (Bestor et al., 1988; Bestor and Verdine, 1994; Gujar et al., 2019; Kalousek and Morris, 1969; Shiraishi et al., 2002).

The type of DNMT determines whether the enzyme is able to catalyse de novo methylation and/or maintenance methylation. It has been postulated that DNMT3a and DNMT3b are generally responsible for de novo methylation while DNMT1 is responsible for maintenance methylation. (Leonhardt et al., 1992; Liang et al., 2002; Okano et al., 1999). However, in DNA methylation dense regions of the genome, the maintenance of DNA methyl marks by only DNMT1 is often insufficient (Liu et al., 1998). Thus, the remainder of the maintenance is catalysed by DNMT3a and DNMT3b (Liang et al., 2002). Numerous mechanisms pertaining to how DNA methylation contributes to transcriptional repression and subsequent gene silencing have been elucidated.

One mechanism proposes that the addition of repressive methyl marks on the 5′ carbons of cytosines, found within cytosine-phospho-guanine (CpG) sites, induces gene silencing by sterically hindering transcription factors from binding to these regions (Comb and Goodman, 1990; Prendergast et al., 1991). In addition, when a CpG site is methylated, methyl CpG binding protein 2 (MeCP2) is recruited to this region (Bird and Wolffe, 1999; Lewis et al., 1992). This protein consists of various domains including, a methyl-CpG-binding domain (MBD) and transcriptional repression domain (TRD) (Nan et al., 1993; Nan et al., 1997). The TRD portion of

MeCP2 is able to confer transcriptional repression by associating the Swi-independent 3 Transcription Regulator Family Member A (Sin3A) (Jones et al., 1998). Sin3a is a transcriptional regulator that is known to interact with histone deacetylases, an epigenetic modulator that contributes to increased chromatin compaction, inaccessibility to transcription factors and subsequent gene silencing (Allfrey et al., 1964; Furdas et al., 2012; Jones et al., 1998) Thus, the co-existence of these various mechanisms impedes genomic regions from being poised for transcriptional activation. DNA methylation marks, which are catalysed by the aforementioned mechanisms, are often found in cytosine-guanine rich sites known as CpG islands, which are prevalent within the genome (Bird, 1986).

Due to the high cytosine and guanine content found within CpG islands, these regions are known as hotspots for DNA methylation. It is important to note that DNA methylation status of a CpG site, irrespective of its location within the genome, ultimately depends on numerous factors such as tissue type, physiological factors and developmental stage (Cedar, 1988; Feinberg, 2007; Suzuki and Bird, 2008). In general, methylated CpG sites found in the genome are localised in genetically inactive regions that make up condensed chromatin known as heterochromatin (Henikoff, 2000; Saksouk et al., 2014). In contrast, CpG sites found within gene promoters and the first exon of a gene often lack DNA methylation marks (Ehrlich, 2019). This lack of promoter methylation allows for transcription factors to access these regions and therefore, allows for a gene to remain transcriptionally active. Thus, due to the aforementioned gene regulating capabilities of this epigenetic mechanism, DNA methylation related epimutations within the genome may result in aberrant protein expression and subsequently contribute to the onset and propagation of diseases such as cancer.

1.4 Techniques utilised to analyse DNA methylation

Several protocols are available to investigate the DNA methylation status of genomic regions, each with their own advantages and limitations. The technique chosen to investigate the epigenetic status of a particular region depends on various factors such as the location of the genomic region, the CpG richness of the region,

resource availability and cost. Some examples of these techniques include bisulfite sequencing, methylation sensitive polymerase chain reaction (MSPCR) and nanopore sequencing.

Bisulfite sequencing has been deemed the gold standard of DNA methylation investigation techniques as, unlike many other techniques, it is an efficient approach that can determine the number and exact location of methylated CpG sites within the region of interest (Li and Tollefsbol, 2011). This technique consists of three major steps, namely: sodium bisulfite mediated conversion and alkaline desulfonation, PCR amplification and sequencing. A summary of this technique can be seen below in **Figure 1**.

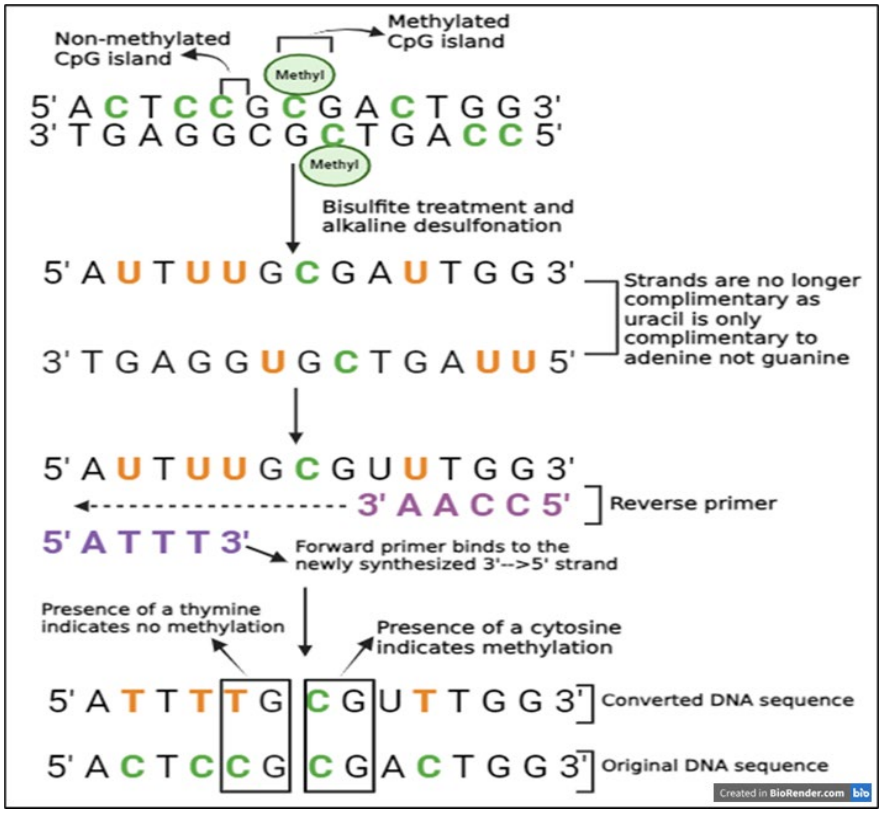


Figure 1: **Summary of the bisulfite sequencing technique.** The basic principle of bisulfite sequencing relies on the methylation dependent conversion of cytosine to uracil. When an unmethylated cytosine is exposed to sodium bisulfite, this results in the deamination of the nitrogenous base and subsequent alkaline desulfonation chemically converts the base to uracil (Hayatsu et al., 1970; Shapiro et al., 1970). In contrast, when a cytosine is methylated, the cytosine remains unreactive when

exposed to sodium bisulfite and alkaline desulfonation (Hayatsu et al., 1970; Shapiro et al., 1970). After this conversion, the forward and reverse strands of DNA are no longer complimentary to each other and hence, one set of PCR primers is needed to amplify the forward strand while another set is needed to amplify the reverse strand. Since DNA methylation, even within the context of cancer, is deemed a symmetrical process, some studies only investigate the forward strand for DNA methylation (Komori et al., 2011; Shao et al., 2009). Amplification of the forward strand is achieved by synthesising a reverse primer (3′ →5′) that binds to the converted forward strand, thereby synthesising a complimentary reverse strand (Frommer et al., 1992). In addition, the PCR reaction is run with a forward primer (5′→3′) that binds to the newly synthesised reverse strand thereby synthesising a forward strand (Frommer et al., 1992). When the converted DNA undergoes PCR, the converted cytosines (now uracils) will be recognised as thymine while unconverted cytosines will be recognised as cytosine (Frommer et al., 1992; Susan et al., 1994). Following sequencing, the amplicon sequence can be compared to the original unconverted DNA sequence. After the DNA has been fully converted, if any cytosines that precede a guanine are present within the amplicon sequence, it can be inferred that these cytosines were methylated (Frommer et al., 1992; Susan et al., 1994). It is important to note that cytosines that are immediately preceded by a guanine on the same strand of DNA, thereby forming a CpG site, are the only cytosines that may be methylated and protected from conversion. All other cytosines that are not followed by a guanine will be converted to uracil. Thus, if there are thymines in the amplicon sequence at locations where a cytosine preceded a guanine in the original DNA sequence then it is clear this CpG site was not methylated (Frommer et al., 1992; Susan et al., 1994).

Although bisulfite sequencing is a promising technique, this approach is slightly more expensive as it requires specialised kits and amplicon sequencing. In addition, since the DNA sequence is almost devoid of cytosine, primer design for this technique is rather challenging. Primers designed for this technique need to be long (26- 30 bases) to ensure that the primers will amplify the correct region. In addition, the primers should ideally not contain any CpG sites (if necessary, one CpG site

may be included at the 5′-end of the primer). Lastly, since the DNA is often damaged after bisulfite conversion, amplicon size has shown an inverse correlation with yield and the insertion of uracil into the DNA after conversion has been speculated to tamper with the processivity of Taq Polymerase during PCR mediated amplification, it is recommended that primers should amplify a short region of the DNA (Darst et al., 2010) (Zymo Research, 2024). These primer design pre-requisites are often difficult to adhere to when investigating CpG rich portions of the genome. Hence, other techniques such as MSPCR may be employed to investigate these CpG rich regions.

MSPCR is a relatively inexpensive technique that is composed of two major steps, restriction enzyme DNA digestion and PCR amplification. The first step consists of digesting genomic DNA with a methylation sensitive restriction enzyme such as HpaII. HpaII is a restriction enzyme that recognises and subsequently restricts the palindromic sequence 5′…CCGG…3′ (cut site: C/CGG). HpaII is deemed as methylation sensitive as the presence of a methylated CpG site within its DNA recognition site sterically hinders the enzyme from cleaving this region. In contrast, when the CpG site within the recognition site is not methylated, the restriction enzyme is able to cut the DNA. The final step of this protocol consists of amplifying the digested genomic DNA with primer sets that amplify regions of the promoter that consist of the aforementioned restriction enzyme recognition sites. If any portion of the DNA that the primers amplify is not methylated and therefore cleaved by HpaII, the region will not be amplified during PCR and no PCR bands will be visualised on an agarose gel. In contrast, if the region of interest that is amplified by the primers is methylated within the enzyme recognition sites, and thus remains intact, there will be amplification and subsequently PCR bands present on an agarose gel. In addition, another sample of genomic DNA should be digested with MspI, an isoschizomer of HpaII that is able to cleave the aforementioned restriction site irrespective of its methylation status. The sample digested with MspI should always be cut at the region of interest and show no amplification after PCR. The lack of PCR amplification indicates that complete digestion occurred and confirms that the resulting banding pattern produced by the digestion with the methylation sensitive enzyme (HpaII) after PCR is because of the DNA's methylation status and

not due to incomplete digestion. If a CpG within a single recognition site is not methylated and subsequently cleaved, there will be no amplicon produced after PCR with digested HpaII genomic DNA as only a single cut is needed to inhibit PCR. Hence, the lack of an amplicon after PCR with HpaII digested DNA does not indicate that every CpG that forms part of a recognition site within the amplicon region is not methylated but rather signifies that at least one CpG site within a restriction enzyme recognition site is not methylated. In contrast, all CpG sites within every enzyme recognition site located in the region that the primer amplifies needs to be methylated and subsequently left intact in order to produce an amplicon. In addition to this technique not being able to decipher which of the CpG containing recognition sites within an amplicon is not methylated, another limitation stems from the fact that, this technique cannot determine the DNA methylation status of CpG sites that are not within the aforementioned enzyme recognition site. Hence, this technique can only partially decipher the overall DNA methylation status of the region of interest. A technique that may be able to overcome this limitation is nanopore sequencing.

Nanopore sequencing is a technique that is able to sequence a region of the DNA without prior sodium bisulfite treatment or restriction enzyme digestion. This technique involves translocating single stranded DNA through a protein nanopore (Gombert et al., 2023). Since each of the four DNA bases alter the current in their own unique way, the DNA sequence is read by monitoring the change in current as the DNA translocates (Gombert et al., 2023). This technique is able to determine the methylation status of a CpG site as methylated cytosine and unmethylated cytosine produce different current changes (Laszlo et al., 2013). Although this technique is efficient, accurate and requires no prior treatment or digestion, this technique is quite expensive (Gombert et al., 2023). Some of these techniques have been utilised when analysing aberrant DNA methylation of genes in various forms of cancer.

1.5 Aberrant DNA methylation in OSCC

Within the context of cancer, DNA methylation epimutations may present as either hypermethylation or hypomethylation. It has been determined that a dichotomy of methylation patterns exist, with global hypomethylation and regional hypermethylation both co-existing in various cancers, including OSCC (Ehrlich et al., 2006; Hoshimoto et al., 2015; Huang et al., 1999; Soares et al., 1999). The presence of global hypomethylation in a cancerous genomic landscape has been proposed to be as a result of decreased DNA methylation activity as well as inadequate DNA methylation maintenance and subsequent passive DNA demethylation (Auclair and Weber, 2012; Szyf, 2001). Whereas the presence of regional hypermethylation has be speculated to be as a result of increased DMNT1 expression and activity in cancer (Agoston et al., 2005; Furst and Barton, 2015). In addition, site-specific DNA hypermethylation can also result from mutations in *Ten-Eleven-Translocation 2* (*TET2*), a gene that codes for the TET2 protein that is able to oxidise methylated cytosines into either 5-formylcytosine or 5-carboxycytosine (Ito et al., 2011; Tahiliani et al., 2009; Tulstrup et al., 2021). The 5-formylcytosine and 5-carboxycytosine intermediates can undergo thymine-DNA glycosylase mediated excision. This recently excised site is then able to acquire a unmethylated status through base excision repair (He et al., 2011; Maiti and Drohat, 2011). Thus, mutations of *TET2*, may render its protein unable to demethylate appropriate regions which would contribute to regional genomic hypermethylation.

Hypermethylation of genes, that are active in healthy tissue such as tumour suppressor genes, could lead to aberrant protein expression would result in transcriptional inactivity and subsequent lack of essential protein expression, a phenomenon that has been noted in OSCC (Baba et al., 2013; Hoshimoto et al., 2015; Ma et al., 2016). Conversely, in OSCC, hypomethylation of regions that are silenced in healthy tissue such as proto-oncogenes lead to inappropriate protein expression, thereby contributing to tumorigenesis (Murata et al., 2014). In addition, hypomethylation of heterochromatic regions and subsequent transcriptional activation has shown to associate with genomic instability in cancer (Estécio et al., 2007; Hoshimoto et al., 2015). Thus, although DNA hypomethylation and

hypermethylation have both been noted in cancer, hypomethylation is not as well characterised (Ehrlich, 2002; Grady, 2021).

Repressed gene expression due to aberrant DNA hypermethylation is often implicated in a plethora of cancers, such as OSCC, due to its ability to downregulate many regulatory cells cycle proteins. In individuals with OSCC, under expression of integral cell cycle proteins that halt the cell cycle consequently contribute to the development of malignant neoplasia that presents itself as LVLs (Baba et al., 2013).

The under expression of cell cycle halting proteins that has been noted in OSCC LVLs has been attributed to decreased expression of genes that code for these aforementioned proteins. This decreased gene expression is as a result of aberrant DNA methylation of cell cycle regulatory gene promoters, which leads to decreased gene expression and subsequently decreased expression of cell cycle proteins that elicit the appropriate halting of the cell cycle (Ma et al., 2016). Aberrant DNA methylation of the *cyclin dependent kinase* (*CDK*) *2A* promoter region that codes for p16 has been observed in OSCC. The p16 protein is involved in the regulation of D type cyclin dependent kinases that promote progression of the cell cycle from the G1 phase to the S phase (Baba et al., 2013; Guo et al., 2006; Ma et al., 2016).

Under conventional conditions, when there are events that trigger cell cycle halting signals such as DNA damage or erroneous oncogene activity during cell proliferation, transcription factors will interact with the unmethylated and therefore uncondensed *CDK2A* promoter region (Goel et al., 2018). This interaction results in the expression of p16, a tumour suppressor protein that inhibits cell cycle progression from the G1 phase to the S phase until the aspect that is causing the cell halting signal is amended (**Figure 2**) (Rayess et al., 2012). If the erroneous aspect cannot be amended, the cell undergoes apoptosis. This lack of cell cycle progression in response to cell halting signals ensures that no mutated DNA, that may induce oncogenesis, is allowed to propagate into daughter cells. Although there may be cell halting signals in the malignant oesophageal squamous cells, these cancerous cells often exhibit methylation of the *CDK2A* promoter and the subsequent condensation of this region does not allow transcription factors to induce expression of the p16 protein, which results in continuous cell cycle progression from the G1

phase to the S phase (**Figure 2**) (Baba et al., 2013; Guo et al., 2006; Rayess et al., 2012).

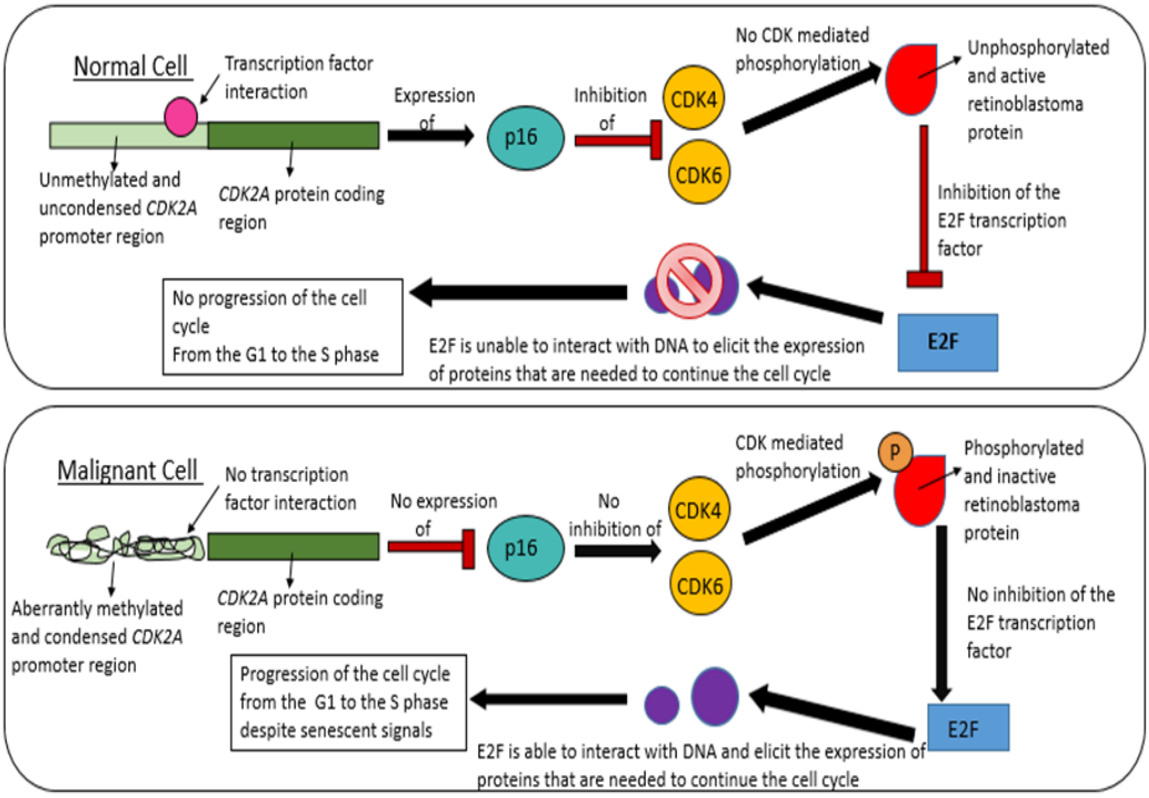


Figure 2: Diagrammatical representation of the p16 pathway under conventional conditions and during aberrant methylation of the *CDK2A* promoter region. An unmethylated *CDK2A* promoter region is essential to regulating the pathway that is involved in cell cycle arrest in response to cell cycle halting signals.

In addition to the aberrant *CDK2A* promoter methylation in OSCC, the promoter region of the *RAS association domain family 10* (*RASSF10*) gene has been found to be methylated in approximately 44.3% of OSCC lesions (Lu et al., 2014). The RASSF10 protein that is encoded by this gene is involved in inhibiting cell cycle progression from the G2 phase to the M phase in response to DNA damage (**Figure 3**) (Lakshmi Ch et al., 2021). Aberrant methylation of the *RASSF10* promoter in malignant OSCC cells causes condensation of this region which prevents

transcription factors from interacting with the promoter, which leads to the progression of the cell cycle from the G2 phase to the M phase (**Figure 3**) (Lakshmi Ch et al., 2021).

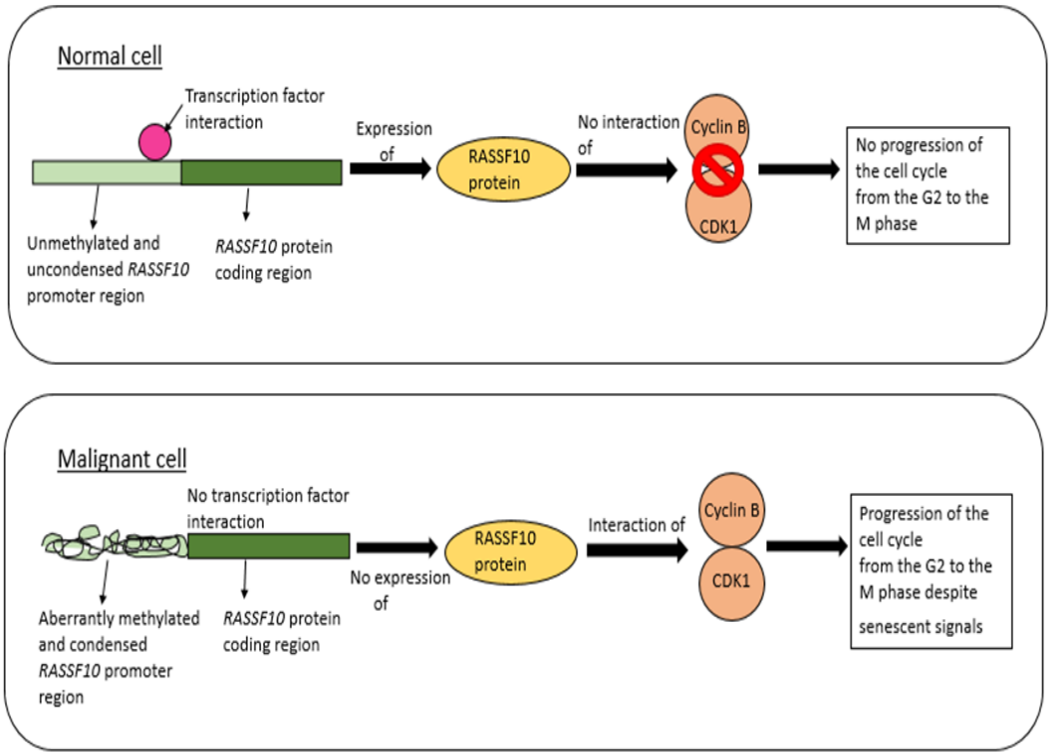


Figure 3: Diagrammatical representation of the RASSF10 pathway under normal conditions and during aberrant methylation of the *RASSF10* promoter region. An unmethylated *RASSF10* promoter region is essential to propagate the pathway that is involved in cell cycle arrest in response to cell cycle halting signals.

The subsequent lack of p16 and RASSF10 expression resulting from the promoter hypermethylation of *CDK2A* and *RASSF10,* respectively, contributes to unregulated cell cycle continuation as well as cell division. This inappropriate cell division leads to the overproliferation of cells with altered genomes that render the cells unable to arrest the cell cycle in response to cell cycle halting signals. This subsequently leads to the propagation of malignant neoplasia that presents itself as LVLs in the oesophageal canal. Thus, it is evident that the erroneous methylation is highly correlated with the development of OSCC. Due to the implication of these epimutations in OSCC, various genes have come under investigation in order to

determine the roles that the methylation of these gene promoters may play in oesophageal squamous cell carcinogenesis. In our study we looked at the expression and methylation status of the *peroxidasin* (*PXDN*) and *perodixasin like* (*PXDNL*) promoter regions in OSCC cell lines.

1.6 PXDN

1.6.1 Structure of PXDN

Peroxidasin, previously referred to as vascular peroxidase 1, is a homotrimeric multidomain protein that is localised in the extracellular matrix (ECM) and is encoded for on the antisense strand of chromosome 2p25.3 (Cheng et al., 2008; NCBI, 2022; Paumann-Page et al., 2017; Péterfi and Geiszt, 2014). PXDN is the human orthologue of the PXDN protein that was initially discovered in Drosophila melanogaster (Nelson et al., 1994). PXDN has been classified as part of the peroxidase protein family due to its conserved haem-containing peroxidase domain (Cheng et al., 2011). In addition to its characteristic motif, this protein has domains that are non-characteristic of the peroxidase containing family that allows it to interact with ECM proteins (Péterfi and Geiszt, 2014). These domains include leucine rich repeats, C2 immunoglobulin-like domain and von Willebrand Factor (vWF) C-type domains. Although the domains of PXDN have been elucidated, some functions associated with these regions remain unclear.

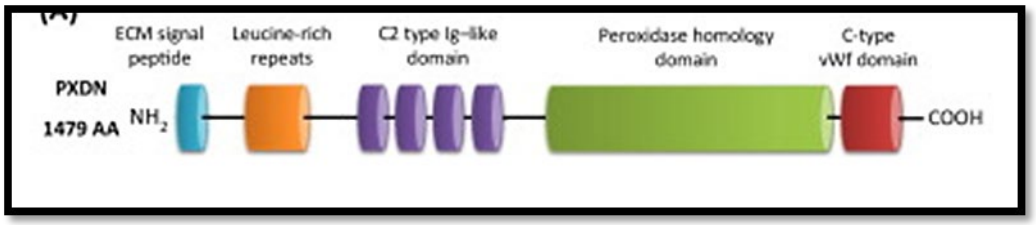


Figure 4: Diagrammatical representation of the linear human PXDN protein. Domains including the leucine rich repeat region, C2 immunoglobulin-like domain and vWF C-type region are unique to this peroxidase. The only conserved region of this protein, that is characteristic of the haem-containing peroxidase family, is the peroxidase homology domain (Péterfi and Geiszt, 2014).

1.6.2 Functions of PXDN

The conserved and unique domains of PXDN are associated with various functions. In accordance with other members that form part of the peroxidase containing family, PXDN is able to carry out the catalysis of substrate oxidation in the presence of hydrogen peroxide (H_2O_2) using its catalytic domain (Cheng et al., 2011). The ability of PXDN to function as an oxidoreductase allows PXDN to consolidate the specialised ECM like structure that is located beneath epithelial cells known as the basement membrane, a fibrous matrix consisting of glycoproteins, collagen (type IV) and laminin (Hellmark and Segelmark, 2007). This membrane mechanically supports the growth of epithelial and endothelial cells and plays a role in cell signalling (Paulsson, 1992). PXDN mediated ECM strengthening is carried out in a plethora of ways.

The C2 immunoglobulin-like domain and peroxidase catalytic domain of PXDN play a role in sulfilimine crosslinks by contributing to the colocalization of collagen IV (Ero-Tolliver et al., 2015). PXDN utilises H_2O_2 and halide ions, such as chloride (Cl^-) or bromide (Br^-), to invoke the formation of a reactive intermediates known as hypochlorous acid (HOCl) and hypobromous acid (HOBr), respectively (Bhave et al., 2012; Ero-Tolliver et al., 2015; McCall et al., 2014). The halide of either newly formed hypohalous acid is transferred to the sulphur atom of methionine to form a halosulfonium cation (Bhave et al., 2012; McCall et al., 2014; Ronsein et al., 2014). The cation, that is located on a collagen IV protomer, consolidates the collagen IV network by crosslinking with another collagen IV protomer's lysine to form a stable sulfilimine bond (Bhave et al., 2012; Ronsein et al., 2014). Once the sulfilimine bond has been catalysed, the bromide ions are excreted back into the ECM. This allows for PXDN to further enrich the basement membrane by utilising this halogen ion to brominate collagen IV tyrosines (Cheng et al., 2011; DiMarco and Giulivi, 2007). This results in the formation of bromo-tyrosines which contributes to the crosslinking of collagen IV (Bathish et al., 2020; Cheng et al., 2011; DiMarco and Giulivi, 2007; He et al., 2020). In addition, the leucine rich repeats of PXDN bind to laminin in the basement membrane, which might suggest that PXDN contributes to ECM structural integrity (Sevcnikar et al., 2020). Moreover, in myofibroblasts, PXDN may modulate the ECM by co-localising with fibronectin, a key component

of the ECM (Péterfi et al., 2009). Lastly, the regulation of PXDN by Nuclear factor erythroid 2–related factor 2 (Nrf2), redox sensitive transcription factor that is involved in combating oxidative stress, suggests that PXDN may be involved in ECM related redox modifications (Hanmer and Mavri-Damelin, 2018). Interestingly, PXDN has also been hypothesised to play a role in innate immunity and host defence.

In the lungs, leucine-rich repeats and C2 immunoglobulin-like domains allow PXDN interact with liposaccharides, a constituent that is found in the membrane of gram negative bacteria. The catalytic domain allows PXDN to mediate bacterial activity by generating hypohalous acids known to play a role in host defence (Shi et al., 2018). In addition, PXDN localised in the vascular system has been found to generate hypochlorous acid in the presence of H_2O_2 and Cl^- (Li et al., 2012). Hypochlorous acid has been found to possess potent antibacterial properties (Boecker et al., 2023; Li et al., 2012). Further study of PXDN has also led to the discovery of PXDN's involvement in angiogenesis.

PXDN has been found to possess pro-angiogenic properties, which is thought to be mediated by its catalytic activity (Medfai et al., 2019). It has been proposed that PXDN is able to incite this pro-angiogenic response by modulating extracellular signal-regulated kinases (ERK1/2), protein kinase B (Akt), and focal adhesion kinase (FAK) signalling (Medfai et al., 2019). Due to the various putative functions that PXDN plays in integral cell function, such as ECM consolidation, this protein has been hypothesised to play a role in the development of diseases that present with a disrupted ECM such as cancer (Venning et al., 2015).

1.6.3 PXDN expression in cancer

Aberrant expression of PXDN has been associated with pathological phenotypes that exhibit abnormal ECM formation and epithelial to mesenchymal transition (EMT), such as cancer. Although associations between PXDN and varying forms of cancer have been found, some of the underlying mechanisms that PXDN participates in require further elucidation. PXDN expression has been noted in

various patients presenting with head and neck cancer, thyroid cancer, liver cancer, pancreatic cancer, prostate cancer, melanoma, breast cancer and ovarian cancer. Although the expression of PXDN in patients with oesophageal cancer has not previously been determined, its expression in the TE-4 and YSE-270 OSCC cell lines has been established (Uhlen et al., 2017; Zhou et al., 2022). However, the role that PXDN plays within these OSCC cell lines requires further elucidation. In addition, a few cancer cell lines have been investigated to elucidate the role that PXDN may play in malignancy. Varying expression of PXDN has been specifically noted in cell lines derived from ovarian, oral, breast and bladder cancers (Di et al., 2019; Kurihara-Shimomura et al., 2020; Young et al., 2015; Zheng and Liang, 2018).

In ovarian cancer cell lines, particularly in the HEY cell line, cells with increased expression of PXDN also exhibited higher levels of phosphorylated phosphatidylinositol-3 kinase (P1K3) and Akt, proteins that are key for the induction of the PIK3/Akt pathway (Zheng and Liang, 2018). When proteins in the PIK/Akt pathway (such as PIK3 and Akt) are phosphorylated, this signalling pathway is able to induce cell proliferation, migration as well as invasion (Zheng and Liang, 2018). In addition, upon silencing of PXDN, there was a severe downregulation of phosphorylated PIK3 and Akt which led to decreased PIK3/Akt pathway activation (Zheng and Liang, 2018). Individuals with ovarian cancer who exhibit increased PXDN expression have shown poorer prognosis than individuals who have downregulated PXDN expression (Zheng and Liang, 2018). Thus, it has been hypothesised that PXDN may induce common metastatic cancer characteristics such as hyperplasia, cell migration and invasion by upregulating the expression of the PIK3/Akt pathway (Zheng and Liang, 2018). In agreement with the findings of PXDN expression in ovarian cancer, similar correlations were found with oral squamous cell carcinoma in that expression patterns of PXDN are associated with poor prognosis (Kurihara-Shimomura et al., 2020).

Increased levels of PXDN expression in oral squamous cell carcinoma has been associated with poor prognosis (Kurihara-Shimomura et al., 2020). This poor prognosis has been hypothesised to be induced by a hallmark of cancer known as the Warburg effect. The Warburg effect, also known as aerobic glycolysis, can be

defined as alterations in the source of adenosine triphosphate (ATP) in cancer cells (Kim and Dang, 2006; Kurihara-Shimomura et al., 2020). While non-cancerous cells derive their ATP from mitochondrial oxidative phosphorylation of pyruvate in the presence of oxygen and only derive ATP from glycolysis under anaerobic conditions, cancer cells are thought to derive their ATP exclusively from glycolysis despite adequate oxygen availability (Kim and Dang, 2006). Although oxidative phosphorylation in the mitochondria leads to a higher amount of ATP, glycolysis occurs at a faster rate and is able to yield a higher amount of ATP in a shorter amount of time (Kim and Dang, 2006). Therefore, the Warburg effect leads to a higher rate of ATP production which is thought to subsidise the energy requirements needed for the constant proliferation that cancer cells undergo (Kim and Dang, 2006). Since PXDN expression was found to be upregulated in cancerous oral squamous cells that exhibited Warburg effect characteristics, such as increased glycolytic enzymes expression, lactate and ATP expression, PXDN could potentially be implicated in inducing tumorigenesis by playing a role in the Warburg effect. However, the exact role that PXDN plays in the Warburg effect requires further elucidation (Kurihara-Shimomura et al., 2020). Moreover, the expression of PXDN in prostate cancer has been implicated in the prevention of cancer cell apoptosis.

In prostate cancer, the expression of PXDN is associated with decreased reactive oxygen species (ROS) levels. This indicates that under cancerous conditions, PXDN maintains its role as a ROS scavenger (Dougan et al., 2019). Since the presence of ROS is essential to increasing mitochondrial membrane permeability, which subsequently leads to the release of pro-apoptotic proteins from the mitochondria, the expression of PXDN is implicated in preventing apoptosis of cancerous cells by decreasing the ROS levels in prostate cancer (Stojnev et al., 2013). In addition, further investigation of the degree of PXDN expression has revealed that PXDN may act as a marker for the staging and potential prognosis of cancer.

Utilising PXDN expression as a diagnostic marker can be implemented for cancers such as breast cancer, where the degree of PXDN expression fluctuates depending on the stage of breast cancer (Young et al., 2015). In addition, PXDN expression

has shown to be statistically significant with staging and prognosis of bladder cancer (Di et al., 2019). Recently, the investigation of the *PXDN* promoter region has been investigated to determine its association with cancer.

1.6.4 *PXDN* promoter methylation in cancer

In skin cutaneous melanoma and lung squamous cell carcinoma, hypermethylation of the *PXDN* promoter region and subsequent decreased PXDN expression has shown an association with longer patient survival times (Zhou et al., 2022). This suggests that the methylation status of the *PXDN* promoter region may possesses prognostic value. Interestingly, within the context of pancreatic cancer, the methylation frequency of the *PXDN* promoter region was stable throughout stages I-IV (Ying et al., 2021). In addition, the *PXDN* promoter region of pancreatic intraductal papillary mucinous neoplasms, showed that an increased frequency of methylation within this region is associated with higher disease severity. This may suggest that the methylation status of the *PXDN* promoter region could assist with the detection of premalignant and early-stage cancerous pancreatic lesions (Ying et al., 2021). Thus, indicating that the methylation status of the *PXDN* promoter region shows promising prognostic and diagnostic value.

Although a few cancer studies have investigated the correlation between PXDN expression and *PXDN* promoter methylation in relation to cancer manifestation, characterisation of PXDNL requires further elucidation.

1.7 PXDNL

PXDNL is a novel homologue protein of PXDN that has been shown to localise exclusively in cardiac muscle, particularly in the ECM of its intercalating discs (gap junctions) (Péterfi et al., 2014). This localisation of PXDNL is thought to be due to the abundance of ROS in this region that stems from the high mechanical and subsequent metabolic stress that this organ undergoes (Gu et al., 2012; Péterfi et al., 2014). Similarly to PXDN, this peroxidase homologue belongs to the aforementioned peroxidase protein family (Péterfi et al., 2014). Analogous to the

structure of PXDN, PXDNL is classified as a hybrid protein as it possesses multiple domains including the leucine rich repeat region, C2 immunoglobulin-like domain, vWF C-type region and a peroxidase domain (Péterfi et al., 2014).

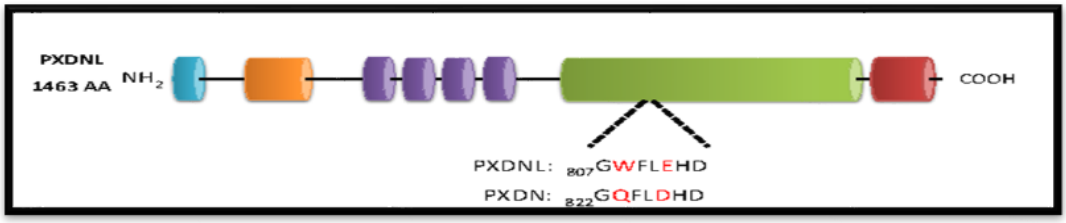


Figure 5: Diagrammatical representation of the linear human PXDNL protein. PXDNL possesses multiple domains including the leucine rich repeat region, C2 immunoglobulin-like domain and vWF C-type region. The only conserved region of this protein, that is characteristic of the peroxidase containing family, is the peroxidase homology domain. The PXDNL peroxidase homology domain is devoid of any catalytic ability (Péterfi et al., 2014; Péterfi and Geiszt, 2014).

Due to the fact that PXDNL possesses these domains it is speculated that the protein, similarly to PXDN, may play a role in ECM consolidation and cell signalling (Péterfi and Geiszt, 2014). As mentioned above, PXDNL localises in the ECM of intercalating discs of the heart which may suggest that this protein plays a role as an adhesive ECM component in cardiac muscle (Péterfi et al., 2014) However, unlike other peroxidases in this protein family, PXDNL contains a peroxidase domain that possesses no detectable enzymatic ability (Gu et al., 2012; Péterfi and Geiszt, 2014). This idiosyncratic lack of catalytic function is thought to be due to a two amino acid residue change within its conventionally catalytic peroxidase domain (Soudi et al., 2012; Péterfi and Geiszt, 2014). Interestingly, when PXDNL forms a hetero-oligomer with PXDN, it acts as an antagonist that inhibits the peroxidase activity of PXDN (Péterfi et al., 2014; Péterfi and Geiszt, 2014). Increased PXDNL expression is particularly noted in failing cardiac muscle and is thought to be induced by angiotensin II, a protein that plays a significant role in cardiac remodelling (Péterfi et al., 2014; Péterfi and Geiszt, 2014). This

remodelling is thought to be as a result of PXDNL inhibiting PXDN from synthesising collagen IV which in turn may alter the ECM of cardiomyocytes (Péterfi et al., 2014; Péterfi and Geiszt, 2014). Further study is needed to determine whether PXDNL is able to inhibit PXDN during cardiac failure. In addition, further work is needed to elucidate whether PXDNL acts as a contributory or compensation mechanism during cardiac disease (Péterfi et al., 2014; Péterfi and Geiszt, 2014). Within the context of other diseases, such as cancer, the expression of PXDNL has been thought to be implicated in contributory mechanisms and has also been speculated to possess promising biomarker capabilities.

1.7.1 PXDNL expression in cancer

Although PXDNL expression has not been previously established in patient samples that are derived from oesophageal cancer, its expression has been noted in the TE-4 OSCC cell line (Uhlen et al., 2017). However, further investigation of this OSCC cell line is required to elucidate the mechanisms that PXDNL is involved in. It has been noted that in urothelial bladder cancer, PXDNL plays a role in inducing EMT (Lu et al., 2023;). It has been proposed that PXDNL contributes to EMT by participating in the Wnt/β-catenin pathway, a well-established modulatory pathway of EMT (Lu et al., 2023; Marcucci et al., 2016; Thiery et al., 2009). The canonical Wnt/β-catenin pathway is initiated once the Wnt molecule binds to its membrane receptors, frizzled and lipoprotein receptor-related protein (LRP) (Bhanot et al., 1996; Wehrli et al., 2000). This ultimately leads to the relocation or disruption of the protein complex that is responsible for the degradation of β-catenin. Subsequently, β-Catenin translocates to the nucleus where it interacts with T cell factor and lymphocyte enhancer factor-1 to incite the transcription of genes involved in cell proliferation and migration (Behrens et al., 1996; Brunner et al., 1997; Molenaar et al., 1996). This study noted that an increase in PXDNL was associated with an increase in phospho (Active) β-Catenin (Ser33/37/Thr41) and a decrease in phosphorylated β-catenin (Ser33/37/Thr41) while the opposite was noted after stable PXDNL knockdown (Lu et al., 2023). This finding suggests that PXDNL may play a role in EMT facilitation by activating the Wnt/ β-Catenin

pathway that could potentially lead to cancer cell invasiveness and migration (Lu et al., 2023; Thiery et al., 2009). In addition, PXDNL has been found to be a promising biomarker for breast cancer (Li et al., 2019).

Increased PXDNL expression was shown to correlate with lower survival rates of breast cancer patients (Li et al., 2019). This correlation has been attributed to the fact that PXDNL has been shown to play a role in the migration of MCF-7 cells, a human breast cancer cell line (Gu et al., 2016). Thus, it has been speculated that the correlation between increased PXDNL expression and low patient survival rate is due to PXDNL contributing to migration that is characteristic of tumour metastatic mechanisms (Gu et al., 2016; Li et al., 2019).

Although PXDN and PXDNL expression has been investigated in some cancers, no studies pertaining to their expression in South African derived OSCC cell lines have been carried out. In addition, the DNA methylation status of the *PXDN* and *PXDNL* promoter regions and the subsequent expression of these proteins in South African derived OSCC cell lines has not yet been determined. Due to the integral roles that PXDN plays in the ECM as well as the novelty and lack of studies pertaining to PXDNL, further investigation of the expression of these proteins and their respective gene promoter methylation statuses within OSCC may reveal promising diagnostic, prognostic, and biomarker capabilities.

2. Aims and objectives

2.1 Aim:

2.1.1 To investigate the DNA methylation status of the *PXDN* and *PXDNL* promoter regions and observe PXDN and PXDNL expression in the SNO and WHCO5 OSCC cell lines.

2.2 Objectives:

2.2.1 Observe PXDN and PXDNL protein localisation using immunofluorescence microscopy.

2.2.2 Quantify the expression of PXDN and PXDNL in the OSCC cell lines using western blotting.

2.2.3 Determine the DNA methylation status of the *PXDN* promoter region in OSCC cell lines using MSPCR.

2.2.4 Determine the DNA methylation status of the *PXDNL* promoter region in OSCC cell lines using bisulfite sequencing.

3. Methods and materials

3.1 Cell lines

The two human OSCC cell lines used in this project were the WHCO5 and SNO cell lines (gifted by Prof. Veale). Both cell lines were derived from OSCC tumours present in South African males (Bey et al., 1976; Veale & Thornley, 1989). Since the highest rates of OSCC related incidence and death primarily affect the Black South African population, the aforementioned cell lines were chosen due to their genetic relevance within the context of the South African population (Alaouna et al., 2019). In addition, the HEK293 cell line (gifted by Prof. Penny) was used as a control since PXDN expression has previously been established in this cell line (Hanmer and Mavri-Damelin, 2018).

3.2 Cell culture maintenance conditions

All cell culture reagents were purchased from Gibco Life Technologies (Thermo Fisher Scientific, Waltham, Massachusetts, USA) unless otherwise stated. All three cell lines were cultured in media that constituted of Dulbecco's Modified Eagles Medium (DMEM) and Hams F12 (1:1) supplemented with 10% foetal bovine serum, 1% penicillin streptomycin and 1% 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES). Cell lines were incubated at a constant temperature of 37°C with 5% CO_2. Cells were maintained in 100 mm culture dishes (NEST Biotechnology, Wixu, China) and routinely passaged when 70-80% confluency was reached. During the passaging process, the cells were rinsed three times with 1X Phosphate-buffered saline (PBS) at pH 7.4. Cells were detached by adding 300-1000 µL of 0.25% Trypsin-Ethylenediamine tetra acetic acid (Trypsin-EDTA) and 700-1000 µL of 1X PBS to the dish. After the addition of Trypsin-EDTA and 1X PBS, the cells were incubated for 2-5mins at 37°C. The cell suspension was pipetted gently to remove cell clumps and 600-1400 µL of cell suspension was placed into the dish along with 8 ml of fresh cell media.

3.3 Immunofluorescence (IF) microscopy

Cells were cultured on a 10 mm X 10 mm coverslip in a 35 mm six well cell culture dish (NEST Biotechnology, Wixu, China) with 3 ml of fresh media. The cells were allowed to grow for three days with no media change until they reached approximately 80% confluency. A total of three wells at a time were seeded and were labelled as follows: No Antibody Control (NAC), No Primary Antibody Control (NPC) and Experiment. The NAC depicts the autofluorescence of the cells as these cells are not incubated with any primary or secondary antibody. This control is implemented to distinguish the degree of autofluorescence that is a result of natively expressed biological components of the cell such as nicotinamide adenine dinucleotide phosphate (NADPH) and flavins or ECM components such as collagen and elastin and as such the degree of autofluorescence may fluctuate depending on the cell type and degree of these biological components (Monici, 2005). The NPC consists of cells that have been incubated with only the secondary fluorescein isothiocyanate (FITC) conjugated antibody and thus, indicates how much of the fluorescent signal may be attributed to the non-specific binding of the secondary FITC conjugated antibody to biological components of the cell (Burry, 2011). This non-specific binding is as a result of the ionic FITC forming covalent bonds with positively charged biological components including aldehydes or histones (Burry, 2011). The degree of experiment fluorescent signal was analysed relative to the NAC and NPC controls in order to determine the actual degree of fluorescence that may be attributed specifically to protein expression.

When the cells had reached approximately 80% confluency, the cells were rinsed three times with 1X PBS and fixed at room temperature (RT) for 10mins using 2 ml of 10% formaldehyde solution (Sigma-Aldrich, St. Louis, Missouri, United States). Cells were rinsed three times with 1X PBS and permeabilised at RT with 2 ml of 0.1% Triton X-100 (Sigma-Aldrich, St. Louis, Missouri, United States) for 7mins. The cells were rinsed three times with 1X PBS and blocking was done at RT with 2 ml of 1% bovine serum albumin (BSA) (Glentham Life Sciences, United Kingdom) in 1 X PBS. After blocking, the coverslip was placed on a tip box lid covered with parafilm (with the cells facing upwards). For the cells on the

Experiment coverslip, immunolabelling of PXDN and PXDNL expression was done by adding 200 µL of 1% BSA in 1X PBS with 4 µL of the respective primary antibody (**Table 1**) directly onto the coverslip and incubating at 37℃ for 1hr. For the No Antibody Control and The No Primary Antibody Control, 200 µL of 1% BSA in 1X PBS was added directly onto the coverslips, incubated at 37℃ for 1hr and then the cells were rinsed with 1X PBS five times. The No Primary Antibody Control as well as the Experiment were incubated with 1 µL of the corresponding secondary antibody (**Table 1**) in 200 µL of 1% BSA in 1X PBS at 37℃ for 1hr. For the No Antibody Control, only 200 µL of 1% BSA in 1X PBS was added directly onto the coverslips and incubation was done at 37℃ for 1hr. The cells were rinsed five times with 1X PBS and the nuclei was stained by incubating the cells with 4′,6-diamidino-2-phenylindole (DAPI) (0.1µL/mL) for 5mins at room temperature. The coverslips were rinsed three times with 1X PBS and were mounted on microscope slides using Invitrogen™ ProLong™ Gold Antifade Mountant Technologies (Thermo Fisher Scientific, Waltham, Massachusetts, USA). The coverslips were sealed using clear nail polish and stored at 4℃ until they were viewed with the Olympus BX63. Images were captured using a 60X oil immersion lens on the fluorescence setting. The images and the degree of fluorescent signal was detected using ImageJ software.

Table 1: Antibody details for IF microscopy.

PXDN primary antibody details	
Supplier	Santa Cruz Biotechnology
Catalogue Number	PXDN (2C11): sc-293408
Primary Species	Mouse
Type of antibody	Monoclonal
Protein residues recognized by antibody	1369-1409
Residue sequence	STSAFSTRSDASGTNDFREFVLEMQKTITDLRTQIKKLESR
PXDN secondary antibody details	
Supplier	SouthernBiotech
Catalogue Number	1050-02
Primary Species	Goat
Recognised species	Mouse (Kappa light chain)
PXDNL primary antibody details	
Supplier	Santa Cruz Biotechnology
Catalogue Number	PXDNL (N-18): sc-98093
Primary Species	Goat
Type of antibody	Polyclonal
Conjugate	FITC
Amino acid residues recognized by antibody	26-56
Recognised residue sequence	SRCLCFKSTVRCMHLMLDHIPQVPQQTTVLD
PXDNL secondary antibody details	
Supplier	Santa Cruz Biotechnology
Catalogue Number	SC 2356
Primary Species	Mouse
Recognised species	Goat
Conjugate	FITC

3.4 Statistical analysis

Data pertaining to the corrected total cell fluorescence (CTCF) is expressed as mean ± standard deviation (SD). All statistical analysis was performed using Prism Graph Pad 8. Statistical significance ($p < 0.05$) was determined using a one-way analysis of variance (ANOVA) and subsequently a Tukey's Honest Significant Difference test. In all cases, p values are reported up to four decimal places and p star values are denoted as follows:

$p > 0.05$	non-significant
$p \leq 0.05$	*
$p \leq 0.01$	**
$p \leq 0.001$	***
$p \leq 0.0001$	****

3.5 DNA extraction

All genomic DNA (gDNA) utilised in this study for MSPCR and bisulfite sequencing was extracted using the Thermo Fisher Scientific™ GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA). This kit was chosen as high concentrations of genomic DNA could be extracted quickly while remaining intact (high molecular weight) with little to no contamination. Although alternate methods such as phenol-chloroform DNA extraction yield high concentrations of DNA, this method is time consuming and any phenol-chloroform contamination results in DNA that possesses a lower purity than DNA extracted with column-based methods which subsequently inhibits downstream applications such as restriction enzyme digests (Evans et al., 2001; New England BioLabs, 2024).

Once the cells (SNO, WHCO5 and HEK293) reached approximately 90% confluency, the media was removed from the dish and the cells were rinsed three times with 1X PBS. 500 µL of fresh 1X PBS was added and cells were scraped with

a cell scraper. The solution was transferred to a 1.5 ml centrifuge tube and the cells were pelleted by centrifuging at 2500rpm for 5mins. The supernatant was discarded and the cells were resuspended in 200 µL of 1X PBS and 20 µL of Proteinase K (20 mg/mL). The cell suspension was then vortexed thoroughly to remove any cell clumps. 200 µL of lysis buffer was added and the cell suspension was vortexed to ensure that the lysis buffer was evenly distributed. The cell suspension was placed in a water bath at 56°C for 10mins with occasional vortexing every 2-3mins. After the samples were removed from the water bath, 20 µL of RNase A (10 mg/mL) was added and the cell lysate was vortexed before allowing the samples to incubate at RT for 10mins. 400 µL of 50% molecular grade ethanol was added to the cell lysate and was vortexed. The lysate was then transferred into a GeneJET Genomic DNA Purification Column™ inserted in a collection tube and centrifuged at 10 500rpm for 1min. The flow through was discarded and 500 µL of the first wash buffer was added to the column prior to centrifuging the sample at 10 500rpm for 1min. The flow through was discarded and the second wash buffer was added to the column before centrifuging at 13 500rpm for 3mins. The column was then transferred to a sterile 1.5ml centrifuge tube and 200 µL of autoclaved MilliQ water added to the column and incubated in the column at room temperature for 5mins before centrifuging the sample at 10 500rpm for 1min. The flow through was placed back into the column and the centrifugation step was repeated to elute the maximum amount of genomic DNA. The DNA was eluted in autoclaved MilliQ water as the constituents of other ethylenediamine tetraacetic acid (EDTA) containing buffers has been shown to inhibit downstream applications such as MSPCR. The DNA was quantified using a Nanodrop One™ on the double stranded DNA setting. The molecular weight and intactness of the DNA was assessed using a 1% agarose gel that was prepared according to section 3.6 below.

3.6 Agarose gel

Genomic DNA quality was assessed using a 1% agarose gel while the length of MSPCR and bisulfite sequencing amplicons were analysed using a 1.7% agarose gel. Agarose gels were prepared by adding 0.5-0.75 g of agarose powder (Cleaver Scientific, Rugby, United Kingdom) to 50 ml of 1X Tris base, acetic acid and EDTA (TAE) buffer. The agarose mixture was heated for 10-20 second intervals and intermittently swirled to dissolve the agarose powder granules. Once the agarose granules were completely dissolved and the liquid gel was clear, the mixture was left to cool down until ~55°C, 5 µL of ethidium bromide (10 mg/ml) (Merck, Munich, Germany) was added to allow for genomic DNA and amplicon visualisation. The liquid gel was swirled gently to mix in the ethidium bromide and the gel was poured into a casting tray and allowed to solidify at RT for 20-30mins. The solidified gel was placed into an agarose gel electrophoresis tank containing 1X TAE buffer. 5 µL of Genomic DNA was mixed with 1 µL of TriTrack DNA Loading Dye (6X) (Thermo Fisher Scientific, Waltham, Massachusetts, USA) prior to the sample being loaded into a well of the gel and the length of the genomic DNA was determined by loading 5 µL of GeneRuler™ 1kb DNA ladder (which already contained loading dye) (New England Biolabs, Ipswich, Massachusetts, United States) into a well of the gel. MSPCR and bisulfite sequencing amplicons that were produced as a result of PCR were not mixed with any loading dye prior to loading 15 µL of these samples into the wells as all types of Taq DNA polymerase mastermix used in this study already contained loading dye. 5 µl of the GeneRuler™ 100 base pair ladder (New England Biolabs, Ipswich, Massachusetts, United States) was loaded into a well and used to determine the size of MSPCR and bisulfite sequencing amplicons. All gels were electrophoresed at 90V for 30-45mins. The gels were viewed using the BIO RAD The Gel Doc XR+ System on the ethidium bromide setting.

3.7 Methylation specific PCR

Since large portions of the *PXDN* promoter region were not conducive to bisulfite sequencing as these regions were too CpG rich to allow for long primers with short amplicons that consist of no more than one CpG site to be designed, another technique was needed to determine the methylation status of the CpG sites within the *PXDN* promoter. Thus, MSPCR was chosen as, in addition to the above reasoning, MSPCR is inexpensive as it requires no specialised kits, reagents or amplicon sequencing.

Following gDNA extraction from the OSCC cells, three reactions were set up: gDNA that undergoes digestion with MspI (methylation insensitive restriction enzyme), gDNA that undergoes digestion with HpaII (methylation sensitive restriction enzyme) and a no enzyme control (NEC) reaction that consists of gDNA that is incubated with all the other components of the digestion reaction excluding the restriction enzymes. The NEC was set up in order to ensure that digestion of the gDNA was as a result of the restriction enzymes and not due to other digestion reaction constituents.

The restriction enzyme digests were set up as shown in **Table 2**. All restriction enzyme digestion components were purchased from New England Biolabs (New England Biolabs, Ipswich, Massachusetts, United States) unless otherwise stated.

Table 2: Restriction enzyme digest composition.

Component	**Volume**	**Final Concentration**
Genomic DNA	*Volume consisting of 95.2 ng of DNA	1.9 ng/µl
10X Restriction Enzyme Buffer	5 µl	1X
Restriction enzyme (MspI/HpaII)	**1.5 µl	-
Nuclease-free Water	Up to 50 µl	-

*Volume differed depending on the concentration of the DNA

**The NEC consisted of 0 µl of restriction enzyme

All reactions were incubated at 37°C for 1.5hrs to activate the enzymes and provide adequate time for digestion. Subsequently, PCR was performed using the restriction enzyme digest DNA. PCR was performed using the primers shown in **Table 3** and their respective amplicon details are shown in **Table 4**. The primers utilised for MSPCR (**Table 3**) and their respective binding regions within the *PXDN* promoter are shown in **Figure 6** below. All PCR reactions were visualised on a 1.7% agarose gel that was electrophoresed at 90V for 45mins.

Table 3: MSPCR primers designed for *PXDN* promoter amplification.

Primer Set	Primer Sequences (5′ →3′)	Annealing temperature
P2B	Forward primer: GCGCACAGCTGGGACGTG Reverse primer: GCACTCACAGGATGGAGGTC	65°C
P3	Forward primer: CAGACTCCCTTGCTGTGCGCTTTG Reverse primer: AGCTGTGCACATGCGCGAGGCT	58°C
P4	Forward primer: TCTGAATCTGGCACCGTCACCCTG Reverse primer: AAAGCCCCTCATTCCTTTTGAGGA	55°C
Control	Forward primer: TCCCATTCCAGGCTGCTTTC Reverse primer: ATACGCACAAAGGTGGCGTT	58°C

Table 4: Details of amplicons produced by MSPCR primers.

Amplicon name	Amplicon position in relation to the transcription start site	Number of recognition sites within the amplicon	Amplicon size
P2B	-346bp to +267bp	4	304bp
P3	-524bp to -54bp	5	471bp
P4	-935bp to -460bp	4	476bp
Control	-1037bp to -795bp	0	243bp

The control primer set was utilised as this set amplifies a region of the *PXDN* promoter that contains no CCGG recognition sites and should never be cleaved by the restriction enzymes. Thus, after this primer set is used for PCR with the digested genomic DNA, there should always be amplification of this region.

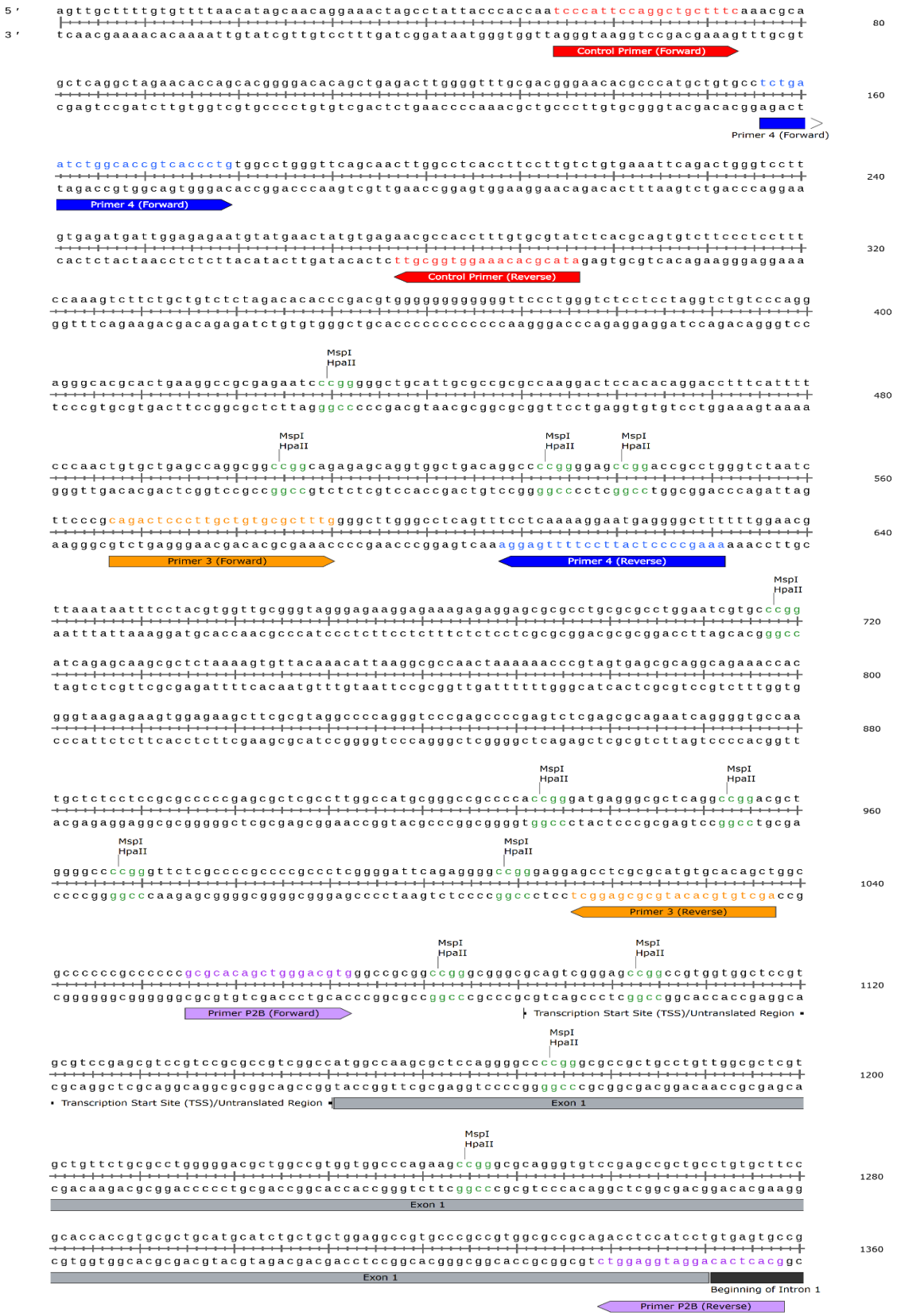


Figure 6: MSPCR primer binding regions and restriction enzyme cut sites within the *PXDN* promoter. The *PXDN* promoter region with the CCGG sites, primer binding regions and transcription start site (TSS) shown.

Table 5: PCR reaction composition for MSPCR.

Component	Volume	Final concentration
Restriction enzyme digested DNA	7.1 µl (13.5 ng of DNA)	0.54 ng/µl
Forward primer [10 µM]	1 µl	[0.4 µM]
Reverse primer [10 µM]	1 µl	[0.4 µM]
***KAPA Taq ReadyMix with loading dye (2X) (Sigma-Aldrich, St. Louis, Missouri, United States)	12.5 µl	1X
Nuclease free water	Up to 25 µl	-

***When performing a PCR with primer set P2B, OneTaq Hot Start Quick-Load 2X Master Mix with Standard Buffer (New England Biolabs, Ipswich, Massachusetts, United States) was used.

A PCR positive control was set up with 13.5 ng of undigested genomic DNA to ensure that other components of the PCR reaction were in working order and to confirm that the lack of HpaII digested DNA amplification was because of the methylation status of the recognition sites and not due to the inability of the PCR components to amplify this region. In addition, a PCR negative control was set up with 0 µl of DNA. All PCR reactions underwent the PCR cycling conditions shown in **Table 6**

Table 6: PCR cycling conditions for MSPCR.

Step	Temperature	Duration	Number of cycles
Initial denaturation	95°C	3mins	1
Denaturation	95°C	0.5mins	35
Annealing	****	0.5mins	
Extension	72°C	1min	
Final Extension	72°C	5mins	1

****See **Table 3**

3.8 Bisulfite sequencing

The methylation status of the *PXDNL* promoter region was investigated by using bisulfite sequencing. This method was chosen as, in contrast to the *PXDN* promoter, the *PXDNL* promoter region is more conducive to bisulfite sequencing analysis as it possesses a moderate number of CpG sites, thus, allowing long primers with no CpG sites and short amplicons to be designed for the analysis of this region. This technique is generally more expensive than other methylation assessment techniques due to the need for specialised kits and amplicon sequencing. However, unlike techniques such as MSPCR, bisulfite sequencing can determine the number and exact location of every CpG sites that is methylated and not just the CpG sites within restriction enzyme recognition sites.

For this project, only the forward strand of the *PXDNL* promoter was inspected for its methylation status. This approach was chosen since methylation of the DNA is generally deemed as a symmetrical process (Komori et al., 2011). Although some cancer studies have found hemimethylated CpG dyads, this phenomenon has been deemed as an intermediate state during cancer associated demethylation rather than as a result of methylation maintenance failure (Shao et al., 2009). In addition, this approach was chosen due to the limitations of the available bisulfite primer seeker

software that is only able to synthesise primer sets that are able to amplify the forward strand.

Bisulfite conversion of all genomic DNA utilised in this project was done by using the EZ DNA Methylation-Lightning™ Kit (Zymo Research Corp, Irvine, California, United States). Once the genomic DNA was extracted according to the protocol above, 130 µl of the Lighting Conversion Reagent was added to 20 µl of DNA (consisting of 300-500ng of genomic DNA). The solution was gently pipetted before incubating the reaction in a PCR machine for 8mins at 98°C and then for 1hr at 54°C. The reaction was placed into a Zymo-Spin IC Column inserted into a collection tube along with 600 µl of M-Binding Buffer. The collection tube was inverted several times to mix the solution and then centrifuged at full speed for 30 seconds. Once the flow through was discarded, 100 µl of the M-Wash Buffer was added to the column before centrifuging at full speed for 30 seconds. The flow through was discarded and 200 µl of L-Desulphonation Buffer was added to the column and allowed incubate at room temperature for 20mins after which the collection tube was centrifuged at full speed for 30 seconds. 200 µl of M-Wash Buffer was then added to the column before centrifuging at full speed for 30 seconds. This wash step was repeated and the flow through was discarded. Lastly, the column was transferred to a 1.5 ml tube before adding 10 µl of autoclaved MilliQ water to the column and then centrifuged at maximum speed for 30 seconds in order to elute the bisulfite converted DNA. The DNA was quantified using a Nanodrop One™ on the single stranded DNA setting since the forward and reverse strands were no longer complimentary. The DNA was then used immediately for PCR with the bisulfite sequencing primers. Primer and amplicon details can be found in **Table 7** and **Table 8**, respectively. The binding regions of the bisulfite sequencing primers used to amplify the converted *PXDNL* promoter are shown in **Figure 8** below. The primer binding regions in relation to the original *PXDNL* promoter region can also be seen in **Figure 7.** Once the primer annealing temperatures had been optimised, PCR mediated amplification of the bisulfite converted *PXDNL* promoter region was conducted. PCR reactions and cycling conditions were set up according to **Table 9** and **Table 10,** respectively. Before sending the PCR reactions to Inqaba for sequencing, 10 µl of the reaction was

loaded onto a 1.7% agarose gel that was electrophoresed at 90V for 45mins in order to confirm that there was amplification of the correct region.

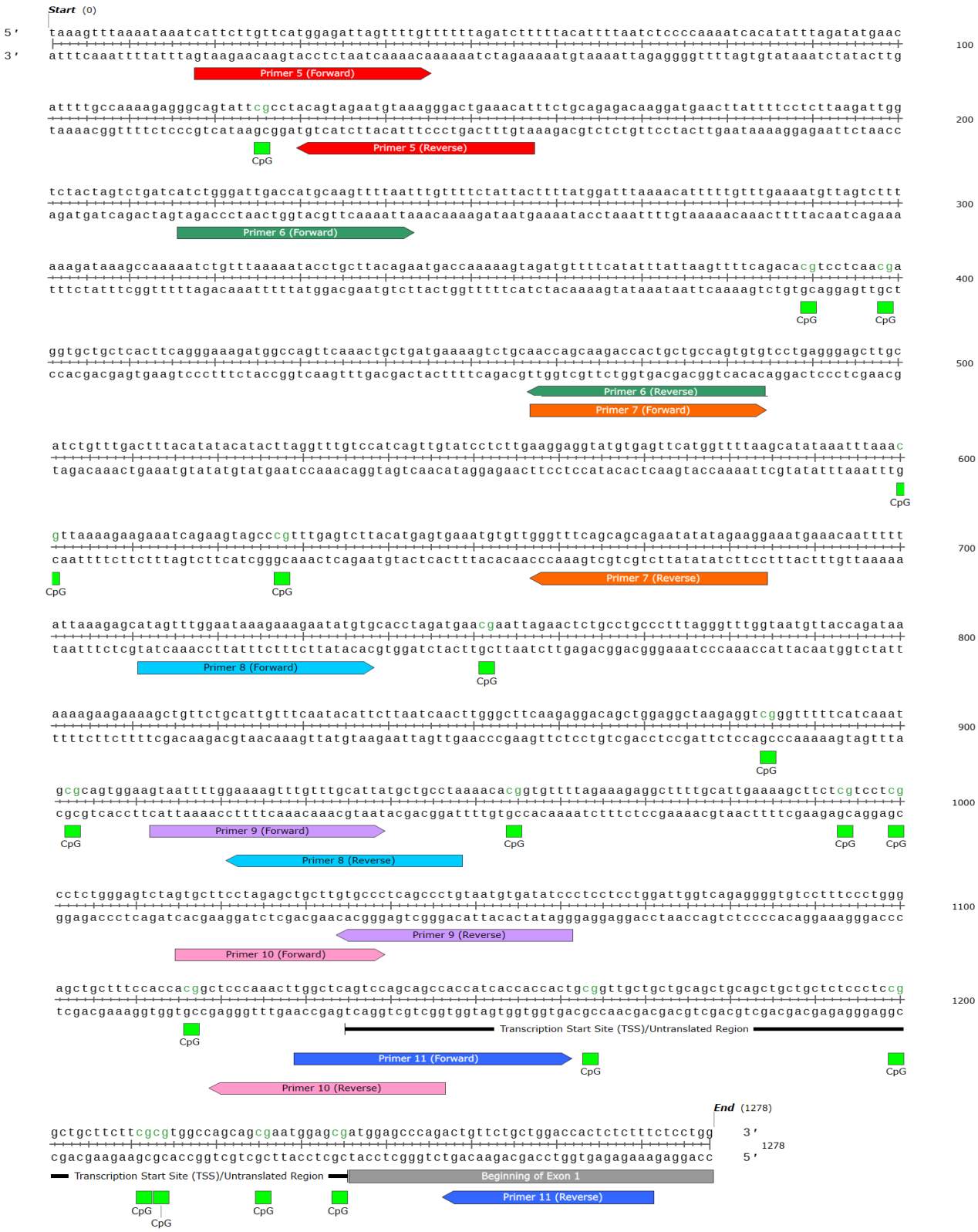


Figure 7: Bisulfite sequencing primer binding regions in relation to the unconverted *PXDNL* promoter. The *PXDNL* promoter region with all the CpG sites, primer binding regions and TSS shown.

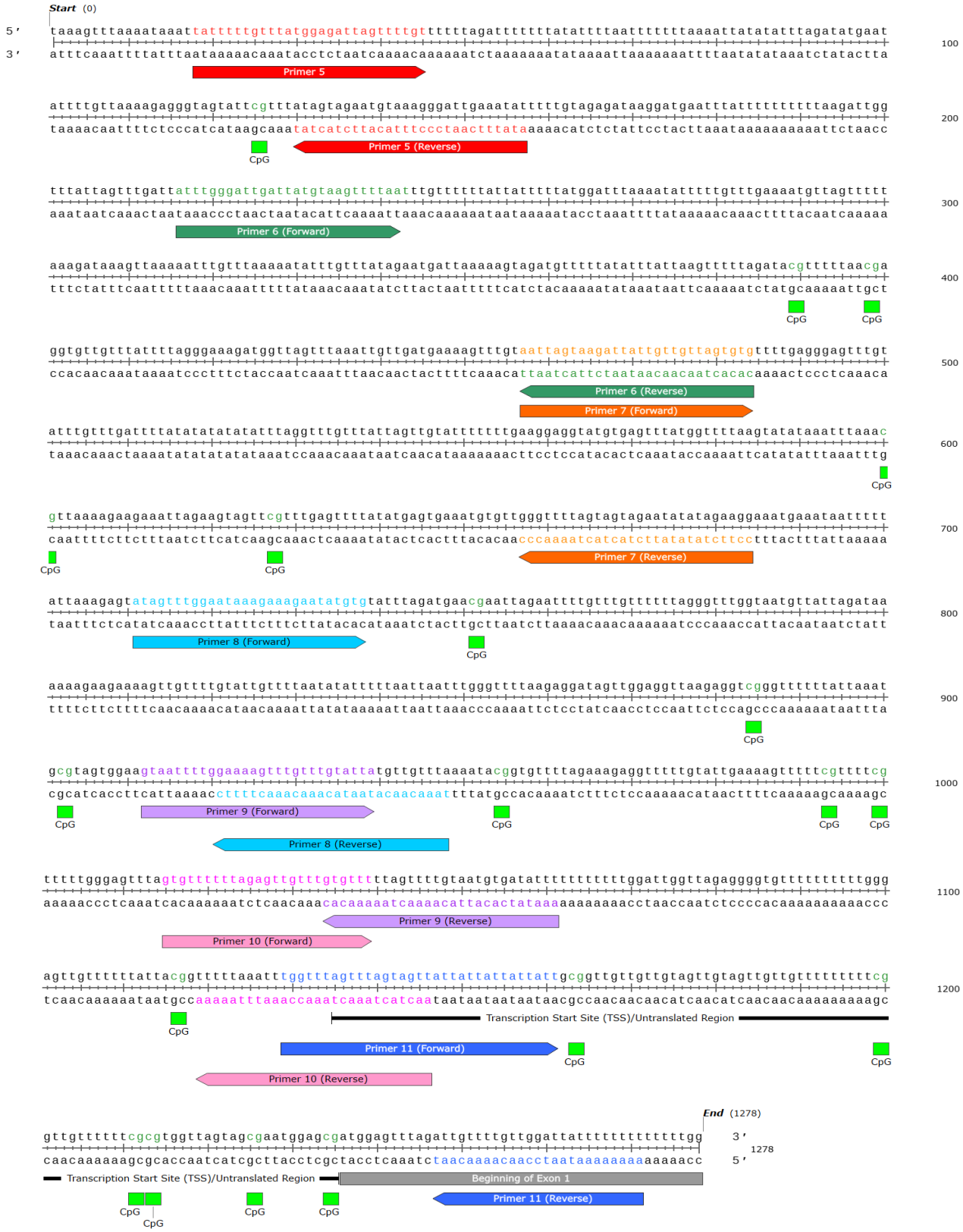


Figure 8: Bisulfite sequencing primer binding regions in relation to the bisulfite converted *PXDNL* promoter. The bisulfuite converted *PXDNL* promoter region with all the CpG sites, primer binding regions and TSS shown.

Table 7: Bisulfite sequencing primer details.

Primer Set	Primer Sequences (5′ →3′)	Annealing temperature
5	Forward primer: TATTTTTGTTTATGGAGATTAGTTTTGT Reverse primer: ATATTTCAATCCCTTTACATTCTACTAT	49°C
6	Forward primer: ATTTGGGATTGATTATGTAAGTTTTAAT Reverse primer: CACACTAACAACAATAATCTTACTAATT	49°C
7	Forward primer: AATTAGTAAGATTATTGTTGTTAGTGTG Reverse primer: CCTTCTATATATTCTACTACTAAAACCC	50°C
8	Forward primer: ATAGTTTGGAATAAAGAAAGAATATGTG Reverse primer: TAAACAACATAATACAAACAAACTTTTC	51°C
9	Forward primer: GTAATTTTGGAAAAGTTTGTTTGTATTA Reverse primer: AAATATCACATTACAAAACTAAAAACAC	49°C
10	Forward primer: GTGTTTTTTAGAGTTGTTTGTGTTT Reverse primer: AACTACTAAACTAAACCAAATTTAAAAA	52°C
11	Forward primer: TGGTTTAGTTTAGTAGTTATTATTATTATTATT Reverse primer: AAAAAAAATAATCCAACAAAACAAT	51°C

Table 8: Details of amplicons produced by bisulfite sequencing primers.

Amplicon	Amplicon position in relation to the transcription start site	Number of CpG sites within the amplicon	Amplicon size
5	-1117bp to -978bp	1	140bp
6	-919bp to -651bp	2	269bp
7	-678bp to -451bp	2	228bp
8	-424bp to -187bp	3	238bp
9	-223bp to -74bp	3	150bp
10	-120bp to +11bp	1	132bp
11	+6bp to +136bp	6	143bp

Table 9: PCR reaction composition for bisulfite sequencing.

Component	Volume	Final concentration
Bisulfite converted DNA	*****Volume consisting of 15 ng of DNA	0.6 ng/µl
Forward primer [10 µM]	1 µl	[0.4 µM]
Reverse primer [10 µM]	1 µl	[0.4 µM]
OneTaq Hot Start Quick-Load 2X Master Mix with Standard Buffer	12.5 µl	1X
Nuclease free water	Up to 25 µl	-

*****A PCR negative control with 0 µl of DNA was set up

Table 10: PCR cycling conditions for amplification of the bisulfite converted *PXDNL* promoter region.

Step	Temperature	Duration	Number of cycles
Initial denaturation	94°C	0.5mins	1
Denaturation	94°C	0.5mins	36
Annealing	******	1min	
Extension	68°C	1min	
Final Extension	68°C	5mins	1

****** Refer to **Table 7**

3.9 Western Blot

3.9.1 Cell preparation and whole cell protein harvesting

Cells were seeded and allowed to grow until 90% confluent. Once this confluency was reached, the cells were rinsed three times with 1X PBS. 1 ml of fresh 1X PBS was placed into the dish and the cells were removed using a cell scraper. The cell suspension was then transferred to an autoclaved 1.5 ml Eppendorf tube before the cells were pelleted by centrifugation at 11 000rpm for 10mins at 4°C. The supernatant was discarded, and the pellet was resuspended in 70-150 µl of 1X radioimmunoprecipitation (RIPA) (containing PMSF and leupeptin) depending on the pellet size. The samples were then placed on a rotator for 1hr at 4°C in order to allow the 1X RIPA to fully lyse the cells. The samples were vortexed briefly and then centrifuged at 12 500rpm for 5mins at 4°C. The supernatant containing the protein was collected and the pellet containing the cell debris was discarded. The supernatant was stored at -80°C prior to use in order to preserve the protein.

3.9.2 Bramhall assay for protein quantification

A Bramhall assay is a colourmetric technique that can be used to quantify protein from a sample of unknown concentration by setting up a standard curve with samples of known protein concentrations (Bramhall et al., 1969).

Whatman filter paper was prepared by soaking the paper with distilled water at 50 rpm on a shaker for 20mins. The paper was then incubated at 50rpm on a shaker with 95% ethanol, 100% ethanol and 100% acetone for 5mins each. Once the paper had dried, a bovine albumin serum (BSA) solution (1 mg/mL) was blotted onto the paper. The solution was blotted on the paper in increments of 1, 3, 6, 12, 16 and 20 µl along with 2 µl of protein that was extracted previously. Once the blotted protein had completely dried, the paper was incubated with 7.5% trichloroacetic acid (TCA) at 50rpm for 45mins to allow for protein fixation. The TCA was removed, and the paper was incubated with Coomassie brilliant blue at 50rpm for 1hr before following up with an incubation with destain for 1hr at 50rpm. The paper was then completely dried before cutting out the protein containing regions and allowing them to incubate individually with 5 ml of elution buffer for 2hrs at room temperature in the dark. The samples were inverted several times before pipetting 160 µl of each protein sample into a 96 well plate in triplicate. The absorbance of the protein samples was measured at 595 nm using the Multiskan GO microplate reader. A standard curve was set up using the concentration (µg/µl) and corresponding absorbance of the BSA standard samples in Microsoft Excel. The unknown protein concentration of the extracted protein was then determined by using its corresponding absorbance and straight-line equation that was derived from the standard curve.

3.9.3 Sodium dodecyl sulfate polyacrylamide gel electrophoresis and protein preparation

Sodium dodecyl sulfate Polyacrylamide gel electrophoresis (SDS-PAGE) was utilised to separate the whole cell protein lysate by molecular weight. Two glass plates with a 1mm spacer were sandwiched together and placed into a cassette. The

stacking and resolving gels were prepared according to **Table 11** below with the exception of the ammonium persulfate (APS) and tetramethylethylenediamine (TEMED) which was added just before the gel was poured between the glass plates to polymerise. After the resolving gel was poured, a layer of 20% Sodium Dodecyl Sulfate (SDS) was used as an overlay to prevent air bubbles which would affect protein migration during electrophoresis. Once the resolving gel had polymerised, the SDS overlay was removed and the stacking gel was poured between the glass plates. The comb was immediately inserted, and the stacking gel was left to polymerise. The gels were then transferred into a gasket within the gel tank along with the running buffer.

Table 11: Gel compositions for SDS-PAGE.

	Gel percentage		
	12% (β-actin)	8% (PXDN/PXDNL)	Stacking Gel (5%)
Reagent	Volume		
Distilled water	3.3 ml	4.6 ml	3.4 ml
29:1 acrylamide: bisacrylamide	4 ml	2.7 ml	830 µl
Tris HCL (pH 6.8)	-	-	630 µl
Trisl HCL (8.8)	2.5 ml	2.5 ml	-
10% SDS	100 µl	100 µl	50 µl
10% APS	100 µl	100 µl	50 µl
TEMED	4 µl	6 µl	4 µl

While the gels were polymerising, equal amounts of protein (20 µg and 80 µg for β-actin and PXDN/PXDNL respectively) were prepared with either 2X or 4X protein loading buffer (PLB) and the samples were boiled in a water bath for 10mins at 70°C. The protein was loaded in triplicate into the wells. In addition, 5 µl of PageRuler™ prestained protein ladder was loaded alongside the samples. The gel was run at 80V until the dye front reached the end of the stacking gel. Once the

stacking gel was reached, the gel was electrophoresed at 100V until the dye front reached the bottom of the gel. The region of the gel corresponding with the molecular weight of interest was cut out of the gel and it was soaked in transfer buffer in preparation for electroblotting. The remainder of the gel was soaked in Coomassie brilliant blue for an hour and destained overnight at 50rpm. This was done to check the quality of the protein and to check that the gel electrophoresis ran adequately.

3.9.4 Electrophoretic transfer

The nitrocellulose membrane was cut to match the dimensions of the gel and this was soaked in transfer buffer along with the other constituents of the transfer sandwich. The transfer sandwich was assembled by placing a sponge on the black side of the cassette followed by the filter paper, gel, nitrocellulose membrane, filter paper and a sponge. The cassette was placed in the transfer tank along with cold transfer buffer and an ice pack. The transfer was run for 3hrs at 300mA with ice pack and buffer changes every half an hour and hour, respectively. The gel was then soaked in Coomassie brilliant blue for an hour and destained overnight at 50rpm to confirm that the transfer was successful.

3.9.5 Antibody incubation and detection

In order to prevent non-specific antibody binding, the nitrocellulose membrane was blocked with 5% BSA in 1X Tris-buffered saline with Tween (TBST) for 1hr at 50rpm. The membrane was rinsed once with 1X TBST at 50rpm for 5mins. The membrane was then incubated with the primary antibody diluted in 5% BSA in 1X TBST according to the **Table 12** and left on the shaker (50rpm) overnight at 4°C. The membrane was washed with 1X TBST five times at 50rpm for 5mins. The membrane was incubated with the secondary antibody diluted in 5% BSA in 1X TBST according to **Table 12** and left on the shaker for 1hr at room temperature.

Table 12: Western blot antibody details.

Protein recognised by primary antibody	Primary antibody's brand and catalogue number	Primary antibody dilution	Species recognised by the secondary antibody	Secondary antibody's brand and catalogue number	Secondary antibody dilution
β-actin	Cell Signalling Technology (CST8547S)	1:2000	Rabbit	Cell Signalling Technology (70745)	1:2000
PXDN	Novus Biologicals (NBP 3-05251)	1:1250	Rabbit		1:2000
PXDNL	Santa Cruz Biotechnology (SC29/293408)	1:1250	Goat	Santa Cruz Biotech-nology (SC 2354)	1:2000

Once the incubation with the secondary antibody was complete, the membrane was washed five times with 1X TBST at 50rpm for 5mins. Prior to detection, the membrane was briefly dipped in equal parts of each western Bright™ ECL plus chemiluminescent reagents (Advansta, San Jose, California, United States). The resulting bands were detected using the BIO RAD The Gel Doc XR+ System on the chemiluminescent setting. Densitometry analysis was performed on the images using ImageJ.

4. Results

4.1 PXDN and PXDNL expression in OSCC Cell Lines

PXDN and PXDNL expression has not previously been characterized in the South African derived OSCC cell lines, SNO and WHCO5. Since it has been previously established in our lab that the HEK293 cells express PXDN and PXDNL, these cells were used as a control. The results of the IF microscopy and CTCF analysis indicated that both the SNO and WHCO5 cell lines express PXDN and PXDNL and, as expected, these proteins localise in the ECM (**Figure 9** and **Figure 10**).

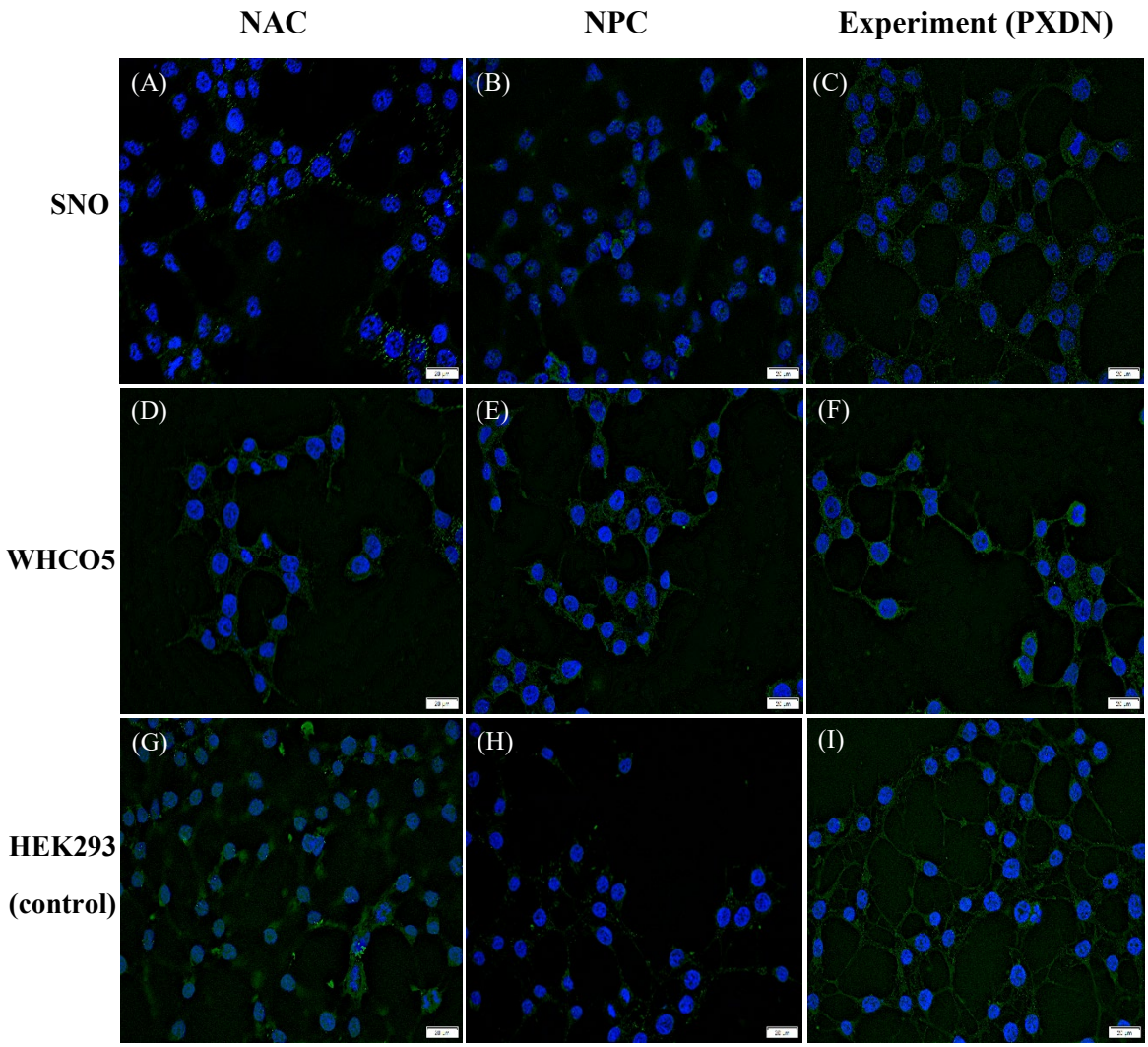


Figure 9: IF microscopy shows that OSCC cell lines express PXDN. IF microscopy images were captured using the Olympus BX63 with a 60X oil immersion lens on the fluorescence setting. The nuclei were stained with DAPI and PXDN was detected using a mouse anti-PXDN primary antibody and a mouse anti-mouse FITC conjugated secondary antibody. FITC and DAPI were excited at a wavelength of 490nm and 359nm, respectively. The ECM of the experiment images shows a higher fluorescence relative to the respective controls. The images denote the No Antibody Control (NAC) (A)(D)(G), No Primary Antibody Control (NPC) (B)(E)(H) and the PXDN experiment (C)(F)(I). The scale bar shows 20 µM.

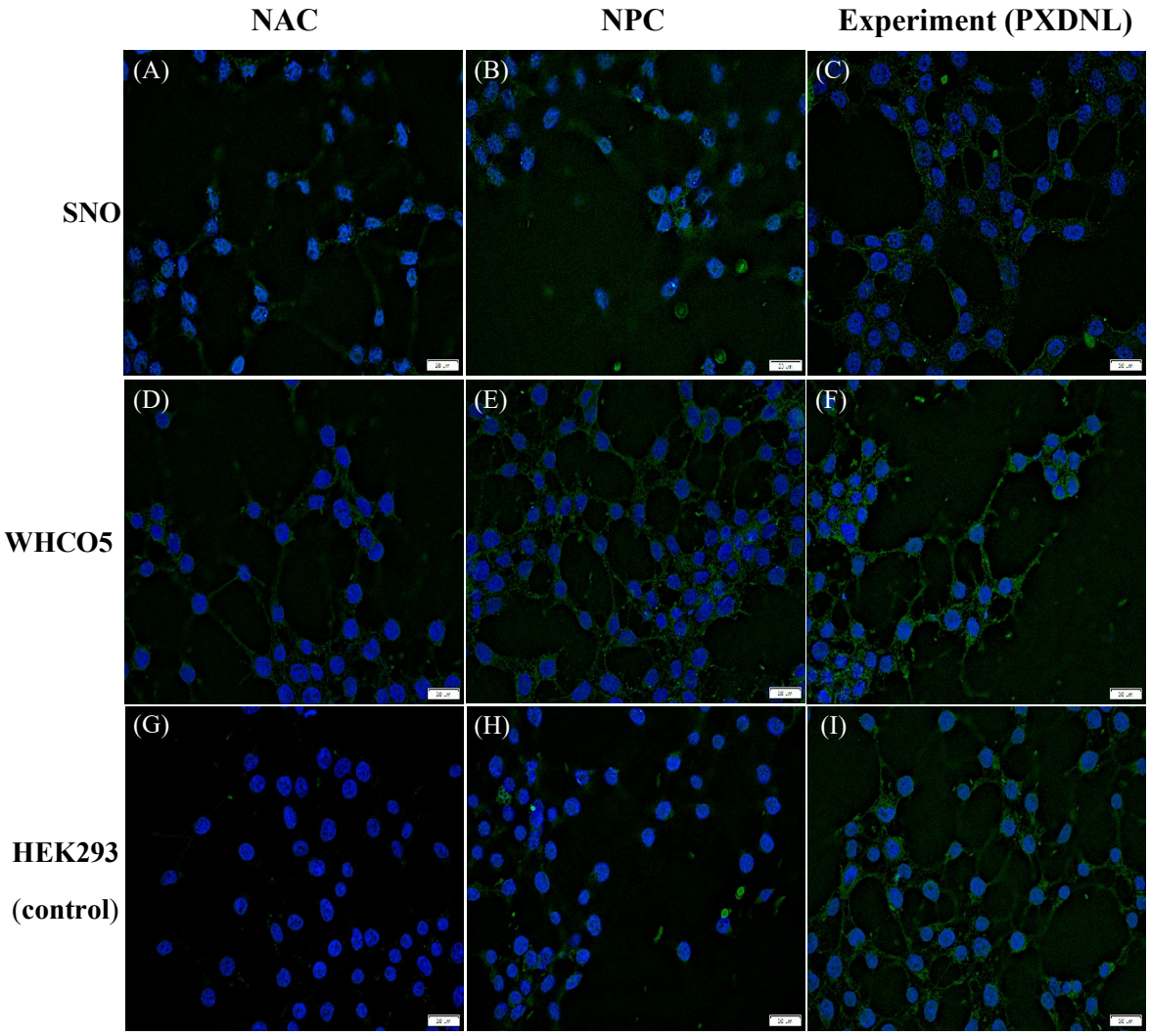


Figure 10: IF microscopy shows that OSCC cell lines express PXDNL. IF microscopy images were captured using the Olympus BX63 with a 60X oil immersion lens on the fluorescence setting. The nuclei were stained with DAPI and PXDNL was detected using a goat anti-PXDNL primary antibody and a mouse anti-goat FITC conjugated secondary antibody. FITC and DAPI were excited at a wavelength of 490nm and 359nm, respectively. The ECM of the experiment images shows a higher FITC fluorescence relative to the respective controls. The images denote the No Antibody Control (NAC) (A)(D)(G), No Primary Antibody Control (NPC) (B)(E)(H) and the PXDNL experiment (C)(F)(I), respectively. The scale bar shows 20 μM.

Although the experiment images (**Figure 9** and **Figure 10**) visually depict a higher fluorescent signal in comparison to their respective controls, it is not possible to visually assess whether the degree of fluorescence portrayed in the experiment images of each cell line is significantly higher than the degree of fluorescence shown in the respective controls. Hence, a CTCF analysis was performed to obtain the mean corrected fluorescence of each cell line's independent groups (NAC, NPC, Experiment) by deducting the background fluorescence from each cell's total fluorescent signal. Since the results of a one-way ANOVA for all cell lines showed statistically significant difference ($p < 0.05$) between the CTCF means of at least two independent groups, a Tukey's Honest Significant Difference test was conducted in order to determine which independent group's mean was statistically significant in relation to the other independent groups.

The mean SNO CTCF values ± SD for the NAC, NPC and Experiment (PXDN) were 121785.43 ± 25672.83, 161255.10 ± 18285.99 and 245659.68 ± 39205.48 arbitrary units, respectively. The one-way ANOVA showed a significant p value of 0.0053. The Tukey's Honest Significant Difference test revealed that there was significant difference between the CTCF means of the NAC and the Experiment (PXDN) as well as the NPC and the Experiment (PXDN) with p values of 0.0047 and 0.0276, respectively.

The mean WHCO5 CTCF values ± SD for the NAC, NPC and Experiment (PXDN) were 230548.42 ± 25401.95, 236692.03 ± 42162.73 and 336747.39 ± 7471.38 arbitrary units, respectively. The one-way ANOVA showed a significant p value of 0.0067. The Tukey's Honest Significant Difference test revealed that there was significant difference between the CTCF means of the NAC and the Experiment (PXDN) as well as the NPC and the Experiment (PXDN) with p values of 0.0095 and 0.0125, respectively.

The mean HEK293 CTCF values ± SD for the NAC, NPC and Experiment (PXDN) were 51889.60 ± 6506.32, 74426.22 ± 6325.11 and 100307.02 ± 4646.61arbitrary units, respectively. The one-way ANOVA showed a significant p value of 0.0002. The Tukey's Honest Significant Difference test revealed that

there was significant difference between the CTCF means of the NAC and the Experiment (PXDN) as well as the NPC and the Experiment (PXDN) with p values of 0.0010 and 0.0040, respectively.

The mean SNO CTCF values ± SD for the NAC, NPC and Experiment (PXDNL) were 166559.39 ± 19767.71, 188635.19 ± 24360.65 and 281715.80 ± 26904.16 arbitrary units, respectively. The one-way ANOVA showed a significant p value of 0.0023. The Tukey's Honest Significant Difference test revealed that there was significant difference between the CTCF means of the NAC and the Experiment (PXDNL) as well as the NPC and the Experiment (PXDNL) with p values of 0.0025 and 0.0073, respectively.

The mean WHCO5 CTCF values ± SD for the NAC, NPC and Experiment (PXDNL) were 111803.64 ± 16998.82, 167569.43 ±15211.69 and 280369.00 ± 54128.80 arbitrary units, respectively. The one-way ANOVA showed a significant p value of 0.0025. The Tukey's Honest Significant Difference test revealed that there was significant difference between the CTCF means of the NAC and the Experiment (PXDNL) as well as the NPC and the Experiment (PXDNL) with p values of 0.0022 and 0.0154, respectively.

The mean HEK293 CTCF values ± SD for the NAC, NPC and Experiment (PXDNL) were 93896.96 ± 4572.72, 169391.37 ± 6794.04 and 229766.82 ± 14690.99 arbitrary units, respectively. The one-way ANOVA showed a significant p value of $7.9312 \text{ X } 10^{-6}$. The Tukey's Honest Significant Difference test revealed that there was significant difference between the CTCF means of the NAC and the Experiment (PXDNL) as well as the NPC and the Experiment (PXDNL) with p values of 0.0010 and 0.001, respectively. The CTCF data as well as the significance established by the Tukey's Honest Significant Difference test can be seen in **Figure 11**.

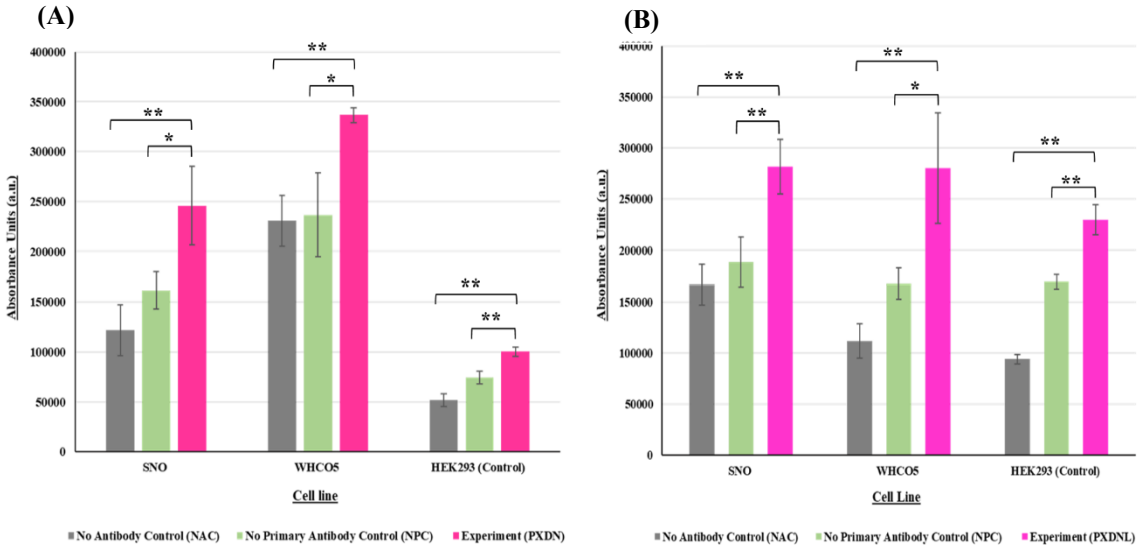


Figure 11: CTCF analysis shows that OSCC cell lines express (A) PXDN and (B) PXDNL. Across all cell lines there was a statistically significant difference between the corrected total cell fluorescence of the Experiment groups in comparison to their respective NAC and NPC groups. CTCF values are expressed as (n=3 ± SD).

After PXDN and PXDNL expression and localisation was established by IF microscopy, the expression of PXDN and PXDNL was further confirmed by performing a western blot using whole cell protein from the OSCC cell lines and the HEK293 control cell line. Subsequently, densitometry was perfomed to compare and contrast the protein abundance between cell lines. Results of the western blot indicated that both OSCC cell lines express PXDN and PXDNL (**Figure 12**). As expected, the HEK293 control cell line expressed PXDN and PXDNL.

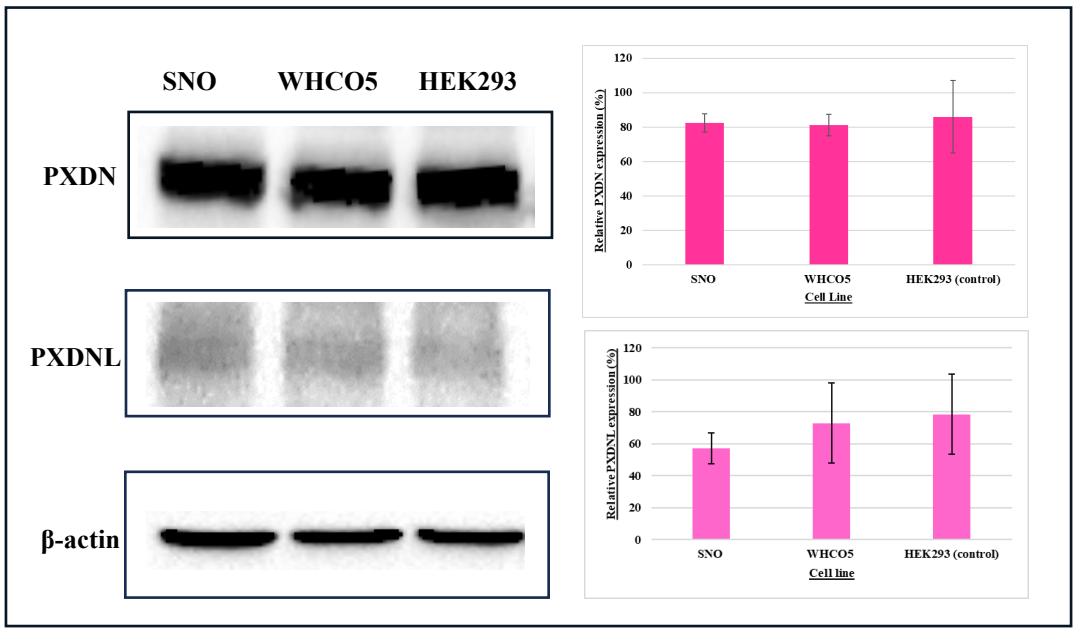


Figure 12: Western blot results indicate that OSCC cell lines express PXDN and PXDNL. Whole cell protein was extracted from OSCC cell lines to detect PXDN and PXDNL expression. PXDN was detected using a rabbit anti-PXDN primary antibody and a goat anti-rabbit HRP conjugated secondary antibody while PXDNL was detected using a goat anti-PXDNL primary antibody in conjunction with a mouse anti-goat HRP. Densitometry analysis is expressed as (n=3 ± SD).

4.2 The methylation status of the *PXDN* and *PXDNL* promoter regions

4.2.1 MSPCR

Once the protein expression of PXDN in the OSCC cell lines had been determined, the methylation status of the *PXDN* promoter region was investigated in order to deduce the relationship between the methylation status of the *PXDN* promoter region and PXDN expression. As seen in **Figure 13**, all MspI digested genomic DNA showed no amplification of the *PXDN* promoter region with all the primer sets (excluding the control primer set), which is expected since this enzyme is not methylation sensitive (lane 2). However, when the DNA was digested with methylation sensitive HpaII, each cell line exhibited various amplification patterns for the *PXDN* promoter region with primer set 2B, primer set 3 and primer set 4

(lane 3). The SNO cell line displays amplification with HpaII digested genomic DNA across all primer sets excluding primer set 2B while the WHCO5 and HEK293 cell lines only exhibit amplification with primer set 4 and the control primer. In addition, as expected, NEC and PCR positive control samples in lanes 4 and 5 respectively, show amplification of the *PXDN* promoter region with all the primer sets, and the PCR negative control shows no amplification (lane 6).

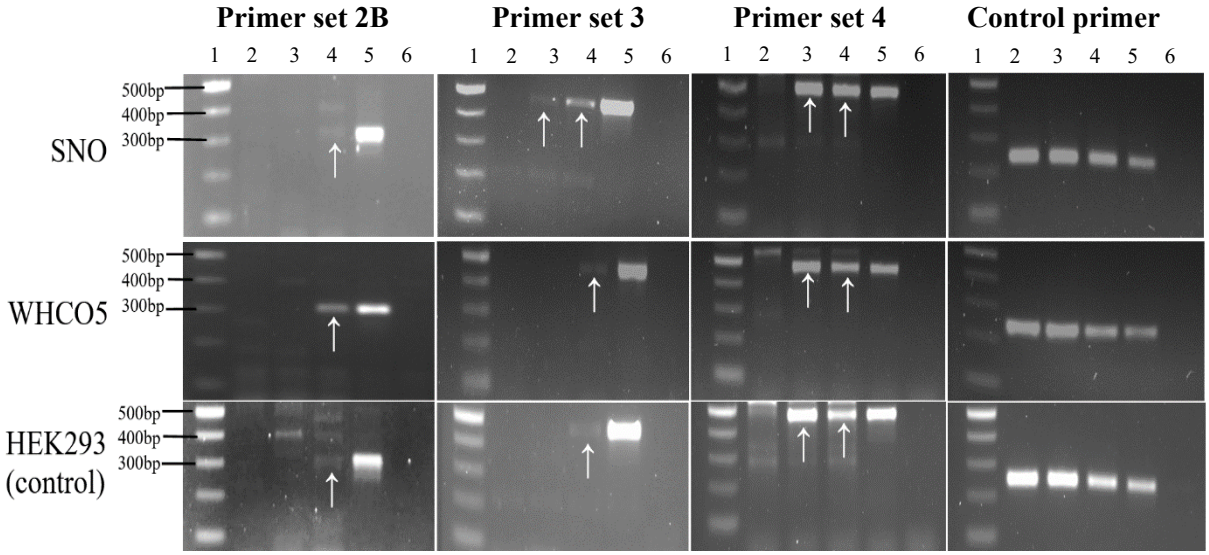


Figure 13: MSPCR results indicate that the *PXDN* promoter region was variably methylated. 1.7% ethidium bromide-stained agarose gel that was electrophoresed for 45mins at 90V with 1X TAE buffer shows a region of the *PXDN* promoter that was amplified during PCR with digested genomic DNA. PCR amplicons produced by primer set 2B, 3, 4 and the control primer set were 304bp, 471bp, 476bp and 243bp, respectively. Lane 1- GeneRuler™ 100bp Ladder; Lane 2- MspI digested genomic DNA; Lane 3- HpaII digested genomic DNA; Lane 4- no enzyme control; Lane 5- PCR positive control (undigested genomic DNA); Lane 6-PCR negative control (no genomic DNA). Arrows indicate the band of interest.

4.2.2 Bisulfite sequencing

Once the bisulfite sequencing primers had been designed, their respective annealing temperatures were optimised by using bisulfite converted HEK293 DNA and a gradient PCR. The gradient PCR temperatures ranged from 49-53°C as all primer annealing temperatures fell within this temperature range. According to **Figure 14**, the optimal annealling temperature for primer sets 5, 6 and 9 was 49°C, for primer sets 8 and 11 it was 51°C, and for primer sets 7 and 10 it was 50°C and 52°C, respectively. The optimal annealing temperature for each primer set was chosen based on which temperature produced the brightest band of the correct amplicon size and minimal non-specific binding.

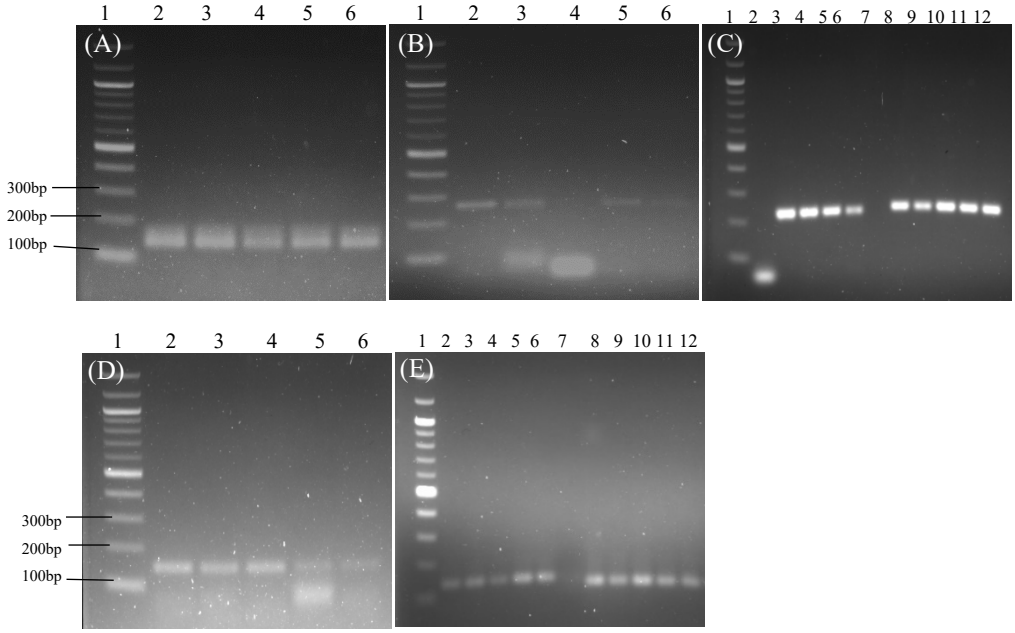


Figure 14: Successful gradient PCR using bisulfite treated HEK293 DNA with bisulfite sequencing primer set 5 (A), primer set 6 (B), primer set 7-8 (C), primer set 9 (D) and primer set 10-11 (E). 1.7% ethidium bromide-stained agarose gel that was electrophoresed for 1hr at 90V with 1X TAE buffer shows the results of a gradient PCR performed with bisulfite converted HEK293 genomic DNA to determine the optimal annealing temperatures for each primer set. Amplicons produced by primers 5-11 were 140bp, 269bp, 228bp, 238bp, 150bp, 132bp and 143bp, respectively. All amplicons were the correct size. Lane 1- GeneRuler™ 100bp Ladder; Lane 2 and 8- 49°C; Lane 3 and 9 - 50°C; Lane 4 and 10- 51°C; Lane 5 and 11- 52°C; Lane 6 and 12- 53°C.

Once the annealing temperatures for the bisulfite sequencing primers had been optimised, the resulting amplicons were sent for sequencing in order to determine the *PXDNL* promoter region methylation status of the HEK293 control cell line. Despite all amplicons being the correct size, the resulting sequences of the aforementioned amplicons did not match the predicted bisulfite converted sequence. **Figure 15** below shows the electropherogram of the amplicon sequence for primer set 5 and **Table 13** shows the amplicon sequence of primer set 5 in comparison to the predicted bisulfite converted *PXDNL* promoter sequence.

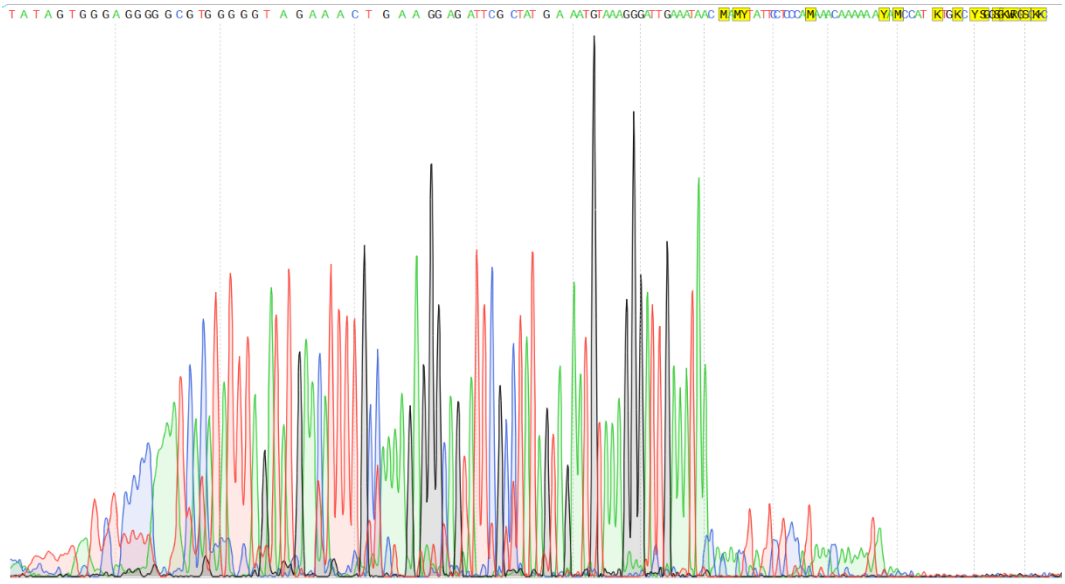


Figure 15: Electropherogram of the amplicon produced by primer set 5 from bisulfite treated DNA. The electropherogram shows a 124bp DNA sequence with noisy peaks and regions of the amplicon that could not be accurately determined and have been deemed ambiguous nucleotides.

Table 13: Comparison of the predicted primer set 5 amplicon sequence produced from the bisulfite converted *PXDNL* promoter region and its sequencing result.

Predicted amplicon sequence:	
TATTTTTGTTTATGGAGATTAGTTTTGTTTTTTAGATTTTTTTATA TTTTAATTTTTTTAAAATTATATATTTAGATATGAATATTTTGTTA AAAGAGGGTAGTATTCGTTTATAGTAGAATGTAAAGGGATTGA AATAT	
Sequencing result:	
TATAGTGGGAGGGGGCGTGGGGGTAGAAACTGAAGGAGATTCG CTATGAAATGTAAAGGGATTGAAATAACMAMYTATTCCTCCCA MAAACAAAAAAYAMCCATKTGKCYSGGSGKWRGSCKKC	
Nucleotide key:	Corresponding base:
A	Adenine
C	Cytosine
G	Guanine
T	Thymine
M	Adenine / Cytosine
Y	Cytosine / Thymine
K	Guanine / Thymine
W	Adenine / Thymine
R	Guanine / Adenine
S	Guanine / Cytosine

5. Discussion

OSCC remains the most prevalent forms of oesophageal cancer both globally and within the South African population (Cook, 2011; Loots et al., 2017; Smyth et al., 2017). Environmental risk factors that may predispose an individual to OSCC, both in the global and South African population, include alcohol consumption and exposure to tobacco smoke (Alaouna et al., 2019; Blot and Tarone, 2018; McCormack et al., 2017). Within the South African population, environmental risk factors that have shown an association with OSCC onset include a shift in traditional beer derivation and increased consumption of maize (Isaacson, 2005; Pillay et al., 2015). In addition, South African specific genetic risk factors that exhibit an association with the clinical manifestation of OSCC have been found. Examples of these genetic risk factors include polymorphisms in *GST* genes and MMR genes (Li et al., 2010; Vogelsang et al., 2012). Although some environmental and genetic risk factors in relation to OSCC onset in the South African population have been explicated, it is well documented that the instigation of cancer requires a myriad of both genetic and epigenetic mutations. Epimutations that contribute to the onset and propagation of OSCC require further elucidation.

One such epigenetic alteration that may contribute to OSCC includes aberrant DNA methylation, which is able to induce aberrant gene expression without altering the DNA sequence (Singal and Ginder, 1999). This is achieved by the addition of a methyl group onto to the fifth carbon present on the nitrogenous base cytosine within a CpG site. This transient addition of a methyl group can sterically hinder transcription factors from binding to these regions and causes the recruitment of further chromatin modifiers resulting in a condensed chromatin state which is inaccessible to transcription machinery and subsequently results in gene silencing (Allfrey et al., 1964; Comb and Goodman, 1990; Furdas et al., 2012; Jones et al., 1998; Prendergast et al., 1991). Aberrant DNA methylation of constitutive gene promoter regions such as *CDK2A* and *RASSF10* promoter regions has been thought to contribute to the development of OSCC (Baba et al., 2013; Ma et al., 2016).

The expression of two ECM modulators, PXDN and PXDNL, has recently come under investigation in diseases that display disrupted ECM environments such as

cancer. In this study, we considered the expression of PXDN and PXDNL in OSCC, and whether DNA methylation could be responsible for differences in their expression.

The aim of this project was to investigate the DNA methylation status of the *PXDN* and *PXDNL* promoter regions as well as PXDN and PXDNL expression. This is the first study to show that the SNO and WHCO5 OSCC cell lines express PXDN and PXDNL. The expression of PXDN and PXDNL in these cell lines is supported by the outcomes of both immunofluorescent microscopy and western blotting.

The immunofluorescence microscopy results reveal that PXDN and PXDNL are localised and expressed in the ECM of both the SNO and WHCO5 cell lines (**Figure 9** and **Figure 10**). Although upon initial visual inspection it may not seem as though the fluorescent signal of the experiment is greater than that of its respective NPC, particularly for the WHCO5 (PXDN and PXDNL) and HEK293 (PXDNL) panels, the similarity between these fluorescent signals may be attributed to the manual setting feature of the Olympus BX60. When using this microscope, the exposure settings are determined by the user's discretion for every individual control and experiment slide and this could introduce slight errors as the images may be slightly underexposed or overexposed and subsequently not visually exhibit the true fluorescent signal of the images. This stands to reason why visually, the fluorescent signal between the NPC and experiment images, may seem similar, as the background level of fluorescence may not be taken into account. The exposure discrepancy of the fluorescent signal could have been avoided by using the Wits Medical school confocal microscope that uses a standard setting across all controls and experiment slides. However, this microscope has been under repair since late 2022, and as such the immunofluorescence analysis of PXDN and PXDNL expression was limited to the usage of the Olympus BX60. Unlike the confocal microscope, the Olympus BX60 has no standard setting that can be utilised across all control and experiment slides as manually employing a single standard exposure setting across all slides will result in the exposure being set too high or too low, and will subsequently produce images that are completely washed out or darkened to the point where the fluorescent signal and the background signal becomes indistinguishable, rendering them unusable for analysis. Due to the limitations of

visually assessing the images for protein expression, a quantitative analysis known as a CTCF analysis was utilised to determine the degree of fluorescent signal and subsequent protein expression.

Unlike a visual evaluation, assessment of fluorescent signal using a CTCF analysis is more reliable as the values produced by this analysis are not affected by underexposure or overexposure. This is due to the fact that the degree of total cell fluorescence is calculated and corrected relative to the background signal. The mean CTCF values for both the PXDN and PXDNL show a statistically significant difference between the CTCF means of the experiment and the respective controls (**Figure 11**). This reveals that there is a significantly higher mean fluorescence from the experiment cells in comparison to the controls that can only be attributed to the binding of the primary and secondary antibodies to their respective targets. Therefore, this indicates that the SNO and WHCO5 cell lines express PXDN and PXDNL as the fluorescent signal that is noted in the experiment cells is not due to auto-fluorescence or non-specific secondary antibody binding.

The standard error bars for the NAC, NPC and the experiment groups for the SNO and WHCO5 cell lines are quite large (**Figure 11**) and this could be for a number of reasons. The degree of autofluorescence from the NAC group depends on the degree of NADPH, flavins, collagen and elastin expression, which emit a green fluorescence (Chance et al., 1979; Deyl et al., 1980; Georgakoudi et al., 2002; Monici, 2005). Similarly, the degree of non-specific secondary antibody binding varies depending on the abundance of positively charged components such as unbound aldehydes from the fixative and histone abundance which are available to covalently bond to the negatively charged FITC ionic fluorochrome (Burry, 2011; Mahmudi-Azer et al., 1998). Thus, the degree of autofluorescence and non-specific antibody binding within the NAC and NPC groups respectively will inherently vary between cells depending on the distribution and expression of the aforementioned biological components (Burry, 2011; Monici, 2005). Since the source of autofluorescence in the NAC is as a result of inherent biological components such as elastin and collagen, which have shown increased expression in some forms of cancer, this signal is unavoidable (Kauppila et al., 1998; Li et al., 2020). The non-specific binding that results in high fluorescence signal in the NPC images could

perhaps be mitigated by increasing the amount of blocking time to allow for adequate masking of the aforementioned positively charged components or by increasing the amount of wash steps to remove any excess secondary antibody (Guerin, 2023). Lastly, degree of fluorescence may differ between cells due a phenomenon, that is associated with FITC, known as self-quenching (Bae et al., 2021). This phenomenon occurs when there is a high density of fluorophores due to an abundance of target protein expression. The high concentration of fluorophores and their subsequent proximity results in the collision between these conjugates that causes self-quenching and lack of fluorescent signal (Bae et al., 2021). The degree of self-quenching is variable as it is dependent on the degree of fluorophore collisions which will vary between cells depending on their protein expression and subsequent FITC proximity. This inherent characteristic of FITC may contribute to the standard error noted within the experiment group as this would lead to inconsistent fluorescent signals between cells despite their similar protein expression. This can be mitigated in future by using a different secondary antibody such as an AlexaFluor, which has shown to possess increased photostability and decreased self-quenching which could allow for a brighter and more accurate depiction of protein expression to be exhibited by the fluorescent signal (Berlier et al., 2003; Panchuk-Voloshina et al., 1999). In addition to the above, the subtle degree of fluorescent signal difference between the HEK293 experiment and NPC could be as a result of low PXDN and PXDNL protein expression in this cell line (**Figure 9**, **Figure 10** and **Figure 11**). Despite the limitations of the Olympus BX60 and non-specific antibody binding, protein expression of both PXDN and PXDNL was confirmed by the CTCF analysis, and the expression of these proteins was further confirmed by western blotting.

Western blotting is a semiquantitative technique that was performed to confirm that the protein expression that was noted in the IF images was due to the expression of PXDN and PXDNL in the OSCC cell lines (**Figure 12**). As expected, the predominant band for PXDN was identified at the correct molecular weight of 165kDa for all cell lines. However, the molecular weight for PXDNL was slightly lower than expected, at 146kDa. This is likely due to the detection of a PXDNL isoform that is also recognised by the anti-PXDNL primary antibody. Since the anti-

PXDNL primary antibody cannot distinguish between the canonical and the 146kDa isoform PXDNL, it is likely that this PXDNL isoform is being expressed in the HEK293 (control) and OSCC cell lines. In addition, since the same primary PXDNL antibody it utilised for IF microscopy, it is likely that the fluorescent signal that is produced in these images is as a result of PXDNL isoform expression. It can also be noted that the CTCF PXDN and PXDNL expression patterns (**Figure 11**) and the densitometry results (**Figure 12**) do not mimic each other. This is likely due to the fact that total cell protein was utilised in the western blot and not purified protein. Thus, despite equal amounts of total cell protein being utilised for β-Actin, the degree to which this loading control is expressed in each cell line may differ across cancer cells, hence, resulting in varying β-Actin concentrations across each cell line's total protein samples (Khan et al., 2014). Since western blot densitometry is semi-quantitative, the densitometry results depict PXDN and PXDNL OSCC expression relative to each cell line's respective β-Actin expression which may vary across cancer cells (Khan et al., 2014; Mahmood and Yang, 2012). Despite seemingly equal amounts of β-Actin that is visually depicted in the western blot results, the β-Actin may be slightly variably expressed in these cell lines in a manner that can only be detected by densitometry. Thus, the relative degree of protein expression cannot be compared across cell lines and the expression pattern noted here will not be synonymous to the CTCF data as fluorescent signal is not a measure of relative protein expression but rather a measure of total PXDN and PXDNL expression. Lastly, the PXDNL blot depicts bands than are not sharp and have high background. Although the lack of sharp bands did not affect the confirmation of protein detection, it may have affected the quantification of relative protein expression that was calculated for PXDNL. Since the other blots exhibited sharp and distinct bands, it is likely that the lack of distinct bands is as a result of specific reagents related to the detection of PXDNL, such as the primary or secondary antibody, rather than due to general reagent failure. Subsequently, sharper bands may perhaps be produced with a different anti-PXDNL primary antibody. Unfortunately, the option of an alternate primary antibody for detecting PXDNL was not available during the course of the study as all primary anti-PXDNL antibodies have been discontinued, including the one that was used in this study. In

addition, more washes or increased blocking time may reduce background signal (Mahmood and Yang, 2012). Despite these challenges, PXDNL isoform and PXDN expression was confirmed by western blotting. Subsequently, the *PXDN* and *PXDNL* promoter regions were investigated to determine if the DNA methylation status of these regions showed an association with the degree of PXDN and PXDNL protein expression, respectively.

MSPCR was performed in order to determine the methylation status of the *PXDN* promoter region. This technique was chosen to investigate the *PXDN* promoter as this region is too CpG rich to allow for the design of long primers with short amplicons (that contain a maximum of one CpG site) that are required for bisulfite sequencing.

Results from MSPCR indicate that the *PXDN* promoter region in both OSCC cell lines were differentially methylated, indicating that there is some degree of inherent variability between the DNA methylation status of the *PXDN* promoter regions found within the SNO and WHCO5 cell lines (**Figure 13**). All the controls produced the expected results, hence confirming the validity of the MSPCR findings regarding the methylation status of the *PXDN* promoter region across all cell lines. As expected, the control primer set (that amplifies a portion of the *PXDN* promoter region with no CCGG recognition sites) amplified both the MspI and HpaII digested OSCC genomic DNA. In addition, all MspI digested OSCC genomic DNA showed no amplification by PCR (excluding the control primer set) and all NEC genomic DNA produced an amplicon at the correct molecular weight. Non-specific binding and subsequent faint banding patterns for the region of interest can be seen within the MspI and HpaII digested DNA reactions as well as the NEC reaction, this is likely due to the restriction enzyme buffer that is present within all these reactions. The restriction enzyme buffer contains Magnesium salt (Mg^{2+}) and once this component has been added to the PCR reaction, it alters the specificity of the DNA polymerase which may result in off target effects (Lorenz, 2012). A clean up step between the restriction enzyme reaction and the PCR could not be implemented as the loss of DNA would be too great to perform a subsequent PCR. Furthermore, all PCR positive controls produced a single band at the correct molecular weight. Lastly, as expected, all PCR negative controls showed no bands. Since all controls

produced the expected outcomes, this indicates that any amplification or lack thereof of the HpaII digested DNA was not due to off target restriction enzyme cleaving, incomplete digestion, restriction buffer constituent interference, PCR reaction component failure or DNA contamination but rather as a result of the methylation status of a CpG site located within the restriction enzyme recognition site.

MSPCR results indicate the *PXDN* promoter region is differentially methylated across all cell lines (**Figure 13**). It is important to note that each primer pair (excluding the control primer) that was utilised in MSPCR, amplified a region of the *PXDN* promoter that consisted of four to five enzyme recognition sites and if one of the CpG sites in single recognition site was not methylated and subsequently cleaved, there would be no amplicon produced after PCR with digested HpaII genomic DNA as only a single cut is needed to prevent PCR amplification of the region. Hence, the lack of an amplicon after PCR with HpaII digested DNA does not indicate that every CpG that forms part of recognition site within the amplicon region is not methylated but rather signifies that at least one CpG site within a restriction enzyme recognition site is not methylated. In contrast, all CpG sites within every enzyme recognition site located in the region that the primer amplifies needs to be methylated and subsequently left intact in order to produce an amplicon. A summary of the MSPCR findings regarding the methylation status of the *PXDN* promoter region can be found in **Table 14** below. The regions amplified by the MSPCR primers using HpaII digested gDNA are either classified as all CpG sites (within the restriction enzyme recognition sites) within the respective amplicon being methylated (M) or at least one CpG site (within the restriction enzyme recognition sites) being unmethylated (UM).

Table 14: Summary of the *PXDN* promoter MSPCR results acquired from HpaII digested DNA.

	Primer set		
Cell Line	2B	P3	P4
SNO	UM	M	M
WHCO5	UM	UM	M
HEK293 (control)	UM	UM	M

According to the SNO MSPCR results, HpaII digested DNA that was amplified with primer pair 3 and primer pair 4, produced an amplicon while the region that was amplified with primer pair 2B did not. This indicates across the regions that primer pair 3 and primer pair 4 amplify, all CpG sites within their respective restriction enzyme recognition sites were methylated. In contrast, the CpG sites within at least one of the recognition sites that primer pair 2B amplifies is not methylated. The WHCO5 and HEK293 cell lines show similar methylation patterns. Only HpaII digested DNA that was amplified with primer pair 4, produced an amplicon while the regions that were amplified with primer pair 2B and primer pair 3 did not. This indicates across the region that primer pair 4 amplifies, all CpG sites within their respective restriction enzyme recognition sites were methylated. In contrast, the CpG site within at least one of the recognition sites that primer pair 2B and primer pair 3 amplifies in their respective regions is not methylated. Since the results of IF microscopy, CTCF analysis and the western blot indicated that PXDN was expressed in the OSCC cell lines, it was initially postulated that the *PXDN* promoter region would completely lack DNA methylation. This was proposed as an unmethylated site and, therefore, uncondensed promoter region is accessible to transcription factors that could induce transcriptional activation and subsequent protein expression. However, the results of MSPCR did not coincide with prediction as, across all cell lines, the *PXDN* promoter region seems to have some degree of methylation. This finding could be for a number of reasons.

A limitation of MSPCR is that this technique is only able to comment on CpG sites within restriction enzyme recognitions sites. In the *PXDN* promoter region (including the transcription start site) there are 93 CpG sites and MSPCR is only able to detect the methylation status of 10 CpG sites as these are the only CpG sites that form part of the CCGG restriction enzyme recognition sites for MspI and HpaII. While the CpG sites within the recognition sites of interest may be methylated (as seen with primer pair 4 for all cell lines and primer pair 3 for the SNO cell line), this may not be a holistic assumption about the methylation status of the *PXDN* promoter as CpG sites are methylated independently of each other. Thus, while the aforementioned CpG sites may be methylated, the methylation status of the remainder of the CpG sites within the *PXDN* promoter may not be methylated which allows for an overall uncondensed promoter region that is conducive to transcription factor binding and subsequent PXDN expression. In addition, the reason for simultaneous PXDN expression and PXDN methylation could be due to the fact that the aforementioned regions, that has been deemed methylated by MSPCR, are not sites where transcription factors bind to elicit gene activation, hence, the methylation status of this region does not affect transcriptional activation and subsequent PXDN expression. Moreover, variable methylation within a promoter region may mask some transcription factor binding sites, however, some sites without CpG site methylation may be left exposed for transcription factor binding and subsequent gene expression. It would be valuable to employ an in silico study to investigate the transcription factor binding regions in the *PXDN* promoter to determine which transcription factors may play a role in modulating PXDN expression within the context of OSCC and whether or not the methylation status of CpG sites within or surrounding the transcription factor binding regions affect the binding of transcription factors that are required to elicit PXDN expression.

Lastly, this methylated status of some portions of the *PXDN* promoter in the OSCC cell lines could also be due to an uncommon occurrence whereby DNA methylation deviates from its canonical association with gene silencing and is implicated in gene activation. This phenomenon has been noted in melanoma and colon cancer cases and the exact mechanism regarding how DNA methylation upregulates gene expression requires further investigation (Gambari et al., 1986; Rauluseviciute et

al., 2020). Moreover, when a primer pair does not produce an amplicon (as seen with primer pair 2B for all cell lines and primer pair 3 for the WHCO5 and HEK293 cell lines), it can be concluded that at least one of the CpG sites that is within a restriction site) is unmethylated. This lack of promoter methylation could contribute to PXDN protein expression that is noted in OSCC cell lines by promoting the formation of an uncondensed promoter region that is accessible to transcription factors. In addition, these regions could be where transcription factors interact and bind to elicit transcriptional activation and subsequent PXDN expression. However, the aforementioned assumptions regarding how the MSPCR results relate to transcription factor binding and subsequent PXDN expression is purely speculative and further research needs to be conducted before the underlying mechanism can be confirmed. As mentioned above, due to the limitations of MSPCR, it is difficult to accurately and holistically comment on the overall methylation status of the entire *PXDN* promoter region that ultimately contributes to the degree of *PXDN* promoter condensation, transcription factor interaction and subsequent protein expression. Therefore, before any association between the methylation status of the *PXDN* promoter region and PXDN expression can be determined, the methylation status of the remaining CpG sites needs to be elucidated. Since the *PXDN* promoter region is not conducive to bisulfite sequencing, alternative techniques such as nanopore sequencing could be employed in future studies to investigate the methylation status of this region. In addition to the *PXDN* promoter region, the DNA methylation status of the *PXDNL* promoter region was also investigated.

The DNA methylation status of the *PXDNL* promoter region was examined using bisulfite sequencing. Bisulfite sequencing was utilised, instead of MSPCR, since this region is more conducive to bisulfite sequencing as it possesses a moderate number of CpG sites, thus, allowing long primers (with a maximum of one CpG site) and short amplicons to be designed for the analysis of this region. In addition, bisulfite sequencing can determine the number and exact location of every CpG site that is methylated and not just the CpG sites within restriction enzyme recognition sites. This allows for a more comprehensive finding regarding the overall methylation status of this region.

Despite bisulfite conversion and subsequent primer optimisation being deemed successful (**Figure 14**), the sequencing data retrieved did not match the predicted sequences for any of the amplicons (**Figure 15** and **Table 13**). Hence, the bisulfite sequencing was ultimately deemed unsuccessful and the DNA methylation status of the *PXDNL* promoter could not be determined. The unsuccessful bisulfite sequencing results could be due to a number of reasons.

Since the primers that only anneal to bisulfite converted DNA produced amplicons with the expected molecular weights lengths, this indicates that the bisulfite mediated conversion of the *PXDNL* promoter region was successful (**Figure 14**). However, despite the primer optimisation indicating that the correct region of the *PXDNL* promoter was amplified, the sequencing data retrieved from the aforementioned amplicons did not match the predicted amplicon sequence for any of the amplicons (**Figure 15** and **Table 13**). Since the PCR reactions that were sent for sequencing only consisted of single bands with no non-specific binding, it is unlikely that the discrepancies between the actual and expected amplicon sequences were as a result of another region of the genome being amplified and subsequently sequenced. Furthermore, since all primer sets that only bind to bisulfite converted DNA produced amplicons, it is unlikely that the discrepancies in the sequencing results is as a result of some DNA remaining unconverted and subsequently sequenced. In addition, since multiple successful PCR reactions for a single primer set were pooled together and a more than an adequate volume of the PCR reaction was sent for sequencing as per Inqaba's pre-requisites, it is unlikely that the poor sequencing data is as a result of low amplicon concentration (Crossley et al., 2020). However, it is highly likely that the discrepancy of the sequencing data and noisy electropherogram peaks stem from the limitations of the Sanger sequencing technique that was utilised in this study (Crossley et al., 2020).

An inherent limitation of direct amplicon Sanger sequencing is its inability to produce readable sequence data within the region that the primer binds to as well as the regions adjacent to the primer binding regions (typically 15-40bp) (Crossley et al., 2020). This is known as an unreadable region as the DNA sequencer cannot distinguish the bases that make up this portion of the amplicon. This is particularly problematic when the amplicons that need to be sequenced, such as the bisulfite

sequencing amplicons in this study, are relatively small (between 132-269bp) and consists of CpG sites within regions adjacent to the primer binding regions (unreadable region). Due to this, bases that made up large portions of the amplicons' sequences could not be distinguished and were unreadable. Since CpG sites fall within this unreadable region, their methylation status cannot be determined. Thus, despite assurance from the manufacturer of the bisulfite sequencing kit that direct PCR amplicon sequencing could be utilised, it is not the most effective strategy to employ for this study. A more apt approach for sequencing the region of interest would be to insert the amplified bisulfite converted *PXDNL* promoter into a cloning vector and sequence this, in order to determine the methylation status of the *PXDNL* promoter region. This approach is more suitable as the starting point and end point of sequencing do not begin at either end of the inserted amplicon. Instead, these points begin within the vector's sequence further upstream and downstream of the vector insertion site, respectively. Since these points are further than 15-40bp upstream and downstream of the amplicon sequence, the unreadable region is confined to the vector's sequence and the bases that constitute the amplicon sequence will be readable.

Another limitation of Sanger sequencing that could contribute to the sequencing data differing from the expected bisulfite converted sequence could be attributed to polymerase slippage that often occurs during the sequencing process at repetitive sequences and homopolymeric regions (regions with stretches of the same nucleotide) (Crossley et al., 2020; Mardis, 2013). Since all unmethylated cytosines are converted to uracil during the bisulfite conversion and subsequently to thymine during PCR amplification, the bisulfite converted DNA is highly adenine and thymine rich (Frommer et al., 1992; Susan et al., 1994). Hence, the adenine and thymine richness of the bisulfite may contribute to the formation of repetitive sequences and homopolymeric regions that will then result in polymerase slippage and subsequently erroneous sequencing results. This limitation may be combatted by utilising another form of sequencing such as the Illumina sequencing-by-synthesis approach. Although this form of sequencing is more expensive than bisulfite sequencing, this approach would combat erroneous DNA sequence reads

as it is much more accurate when sequencing repetitive/homopolymeric regions (Mardis, 2013).

Despite this study being unable to holistically comment on the *PXDN* and *PXDNL* promoter regions, it was able to partially comment on the methylation status of the *PXDN* promoter region and confirm the expression of PXDN and PXDNL in the SNO and WHCO5 OSCC cell lines.

PXDN has been deemed integral ECM modulator as this protein has been shown to consolidate the ECM by crosslinking collagen and interacting with laminin and fibronectin (Bhave et al., 2012; Cheng et al., 2011; Péterfi et al., 2009; Sevcnikar et al., 2020). Since PXDNL possesses the same protein domains that allow PXDN to interact with ECM components, it has been suggested that this protein might play a role in ECM modulation, however, its exact function requires further elucidation (Péterfi et al., 2014). In cardiomyocytes, PXDNL has been shown to antagonise PXDN which results in the prevention of PXDN mediated collagen crosslinking and the remodelling of the ECM. Hence, due to the roles that PXDN and PXDNL are postulated to play in ECM, these proteins have been thought to contribute to the development of diseases that present with a disrupted ECM such as cancer (Péterfi et al., 2009; Péterfi et al., 2014; Venning et al., 2015). Further investigation of PXDN expression could reveal that it plays a similar role in OSCC as it does in other forms of cancer.

Although PXDN expression has been noted in various patients presenting with thyroid cancer, liver cancer, pancreatic cancer, prostate cancer, melanoma, breast cancer and ovarian cancer (Uhlen et al., 2017), no conclusive studies pertaining to PXDN expression in OSCC presenting patients have been carried out. A few OSCC cell lines, such as the TE-4 and YSE-270, have shown PXDN expression. However, further investigation of the OSCC cell lines is required to elucidate the mechanisms that PXDN is involved in (Uhlen et al., 2017). Certainly, this finding in South African derived OSCC cell lines is a novel finding. A few other cancer cell lines have been investigated to elucidate the role that PXDN may play in malignancy.

Perhaps, similarly to ovarian cancer cell lines, the expression of PXDN in OSCC could potentially participate in the PIK3/Akt pathway which contributes to cell

proliferation and cell migration as well as invasion (Zheng and Liang, 2018). In addition, PXDN expression and its potential role in OSCC may be similar to its suggested function in oral squamous cell carcinoma whereby PXDN is thought to participate in the Warburg effect, a phenomenon that subsidises energy demands that cancer cells require for continuous proliferation (Kurihara-Shimomura et al., 2020; Kim and Dang, 2006). Further investigations are required to determine whether PXDN acts through these pathways in OSCC. In prostate cancer, the expression of PXDN led to decreased oxidative stress related ROS levels which contributes to decreased mitochondrial membrane permeability and the subsequent prevention of pro-apoptotic protein release from the mitochondria (Stojnev et al., 2013). Thus, leading to decreased apoptosis of cancerous cells. Since PXDN is expressed in the OSCC cell lines, further investigation may reveal that this protein participates in this aforementioned mechanism within the context of OSCC. In addition, in breast cancer cell lines, PXDN has been noted to increase cell migration and further investigation of its role in OSCC might reveal a similar finding (Sigurdardottir et al., 2021). Lastly, the expression of PXDN in breast cancer and bladder cancer has shown an association with staging as well as subtype and prognosis, respectively. Further study of PXDN and its expression profile within OSCC may reveal its potential as a biomarker and its cancer staging capabilities as well as reveal its prognostic value. Further research is required to confirm the proposed role of PXDN in OSCC as the role PXDN plays in cancer may differ depending on the cancer type, cancer subtype and stage. In addition to confirming the expression of PXDN in OSCC cell lines, the current study was also able to partially comment on the *PXDN* promoter methylation status.

In previous skin cutaneous melanoma and lung squamous cell carcinoma studies, it has been reported that the methylation of the *PXDN* promoter is associated with gene silencing and subsequent decreased PXDN expression (Zhou et al., 2022). Perhaps within the context of OSCC, similarly to these studies, the hypermethylation of the *PXDN* promoter region may show an association with longer patient survival times which may suggest that the methylation status of the *PXDN* promoter region in OSCC possesses prognostic value (Zhou et al., 2022). In contrast, further investigation of the *PXDN* promoter region in the context of OSCC

may reveal that the methylation frequency of premalignant OSCC lesions increases with increased disease severity and, similarly to its proposed role in pancreatic cancer, could potentially assist with the detection of premalignant and early-stage cancerous lesions (Ying et al., 2021). Thus, further investigation of the *PXDN* promoter region methylation status may elucidate whether the extent of methylation possesses diagnostic and prognostic value in OSCC.

Unlike PXDN, there is limited expression data for PXDNL in the context of cancer. PXDNL RNA expression has been confirmed in patients presenting with breast cancer and the expression of PXDNL in OSCC cell lines has been previously established in the TE-4 OSCC cell line (Uhlen et al., 2017). This is the first study to establish PXDNL expression in the SNO and WHCO5 cell lines.

The expression of a PXDNL isoform (146kDa) was confirmed by this study. Further study is needed to confirm if the role this isoform plays in cancer is similar to that of the canonical PXDNL and the other PXDNL isoform (57kDa). The other isoform of PXDNL (57kDA) has shown to be expressed in bladder cancer (Lu et al., 2023). Recently it has been noted that in bladder cancer, PXDNL plays a role in inducing EMT by activating the Wnt/ β-Catenin pathway that could potentially lead to cancer cell invasiveness and migration (Lu et al., 2023; Thiery et al., 2009). The expression of PXDNL in OSCC cell lines could suggest that this protein is implicated in the aforementioned mechanism that subsequently contributes to cancer invasiveness and migration, however, further exploration of this mechanism is required in order to support this notion. In addition to the contributory attributes of PXDNL to cancer onset, further study of PXDNL expression within the context of OSCC may reveal that, as noted in breast cancer cell lines, this protein may contribute to cancer cell migration and lower patient survival, thereby elucidating the potential prognostic value of the degree of PXDNL expression within the context of OSCC (Gu et al., 2016; Li et al., 2019). Although this study was able to confirm PXDNL expression in OSCC cell lines, it was unable to determine the methylation status of the *PXDNL* promoter region. There are no published studies available pertaining to the methylation status of the *PXDNL* promoter region. Hence, speculations regarding this region's role in OSCC to similar studies cannot

be made. Perhaps, similarly to the *PXDN* promoter region, further study of this region could reveal its prognostic and diagnostic value with respect to OSCC.

Despite this study being able to partially comment on the methylation status of the *PXDN* promoter and confirm the expression of PXDN and PXDNL in the SNO and WHCO5 cell lines, there were limitations to this study that could be overcome with further investigation and optimisation. Firstly, this study examined the basal expression of PXDN and PXDNL in OSCC cell lines. The expression of these proteins was not investigated in relation to pathways that disrupt the ECM and contribute to EMT in cancer such as Wnt/ β-Catenin or PIK3/Akt pathways. Hence, the contributary role that PXDN or PXDNL may play in these pathways cannot be confirmed by this study. This limitation may be overcome by investigating the degree of PXDNL and PXDNL expression in relation to the expression of other proteins that are implicated in the pathways of interest in order to elucidate the exact role of PXDN and PXDNL in these mechanisms. In addition, these in vitro findings may not be identical in in vivo studies. Hence, PXDN and PXDNL expression should be investigated in patient tissue samples to elucidate the potential diagnostic and prognostic value that the expression of these proteins may possess. Moreover, only the isoform of PXDNL (146kdA) was found in the SNO and WHCO5 cell lines. Since the primary antibody that was used to identify PXDNL expression recognises an amino acid sequence that is present in both the canonical and 146kDa isoform of PXDNL, the lack of canonical PXDNL expression in the western blot results may be as a result of this primary antibody possessing a higher binding affinity for the PXDNL isoform. Identification of the canonical PXDNL in these cell lines may be confirmed by utilising an anti-PXDNL primary antibody that recognises an amino acid sequence that is unique to the canonical PXDNL. The use of an alternate anti-PXDNL primary antibody was not possible during the course of this study as all commercial primary antibodies that probe for PXDNL are currently discontinued. Furthermore, the technique utilised to investigate the *PXDN* promoter region could not determine the methylation status of the majority of its CpG sites. This limitation may be overcome in future studies by investigating the methylation status of this region with another technique such as nanopore sequencing. In addition, the methylation status of the *PXDNL* promoter region could not be

elucidated due to the limitations of direct sanger sequencing of bisulfite converted amplicons. This could be mitigated by inserting the sequence into a vector prior to sequencing or by using another form of sequencing altogether. Lastly, this study focused on determining how the degree of methylation within the *PXDN* and *PXDNL* promoter regions in the SNO and WHCO5 cell lines affects the expression of these genes by solely investigating PXDN and PXDNL protein expression, respectively. Hence, this study could be furthered strengthened by directly assessing the transcriptional activity of the aforementioned genes in relation to the methylation status of their respective promoter regions by utilising quantitative reverse transcription polymerase chain reaction (RT-qPCR) as this technique allows for the quantification of messenger RNA (mRNA), a type of RNA that correlates to the transcriptional activity of a gene (Adams, 2020). Although the degree of mRNA does not necessarily correspond to the degree of protein expression, as transcribed mRNA is not always translated into protein, the use of RT-qPCR could comment on the transcriptional activity of a gene which may be a valuable insight (Vogel and Marcotte, 2012).

6. Conclusion and future avenues of research

Despite OSCC being one of the most prevalent forms of cancer, both within the global and South African population, it remains one of the most fatal forms of cancer. Hence, it is vital to elucidate potential markers that may possess diagnostic, prognostic or therapeutic benefits. PXDN is a well-established ECM modulator and PXDNL is speculated to modulate the ECM since it antagonizes PXDN and shares domain homology with PXDN. Aberrant DNA methylation is a well-established epigenetic mechanism involved in cancer, including OSCC. To date, no studies pertaining to the DNA methylation status of the *PXDN* and *PXDNL* promoter regions and the subsequent expression of PXDN and PXDNL in South African derived OSCC cell lines have been carried out.

This study has shown PXDN and PXDNL are expressed in South African derived OSCC cell lines, which is a novel finding. In addition, this study reported that the *PXDN* promoter region CpG sites (within MspI/HpaII recognition sites) are differentially methylated across both the SNO and WHCO5 OSCC cell lines.

Further study is required to ascertain the roles that PXDN and PXDNL play in OSCC. Future studies could investigate PXDN and PXDNL expression in relation to pathways that disrupt the ECM and contribute to EMT in cancer such as Wnt/ β-Catenin or PIK3/Akt pathways. In addition, further investigation could also fully elucidate the overall methylation status of the *PXDN* and *PXDNL* promoter region by utilizing nanopore sequencing as well as identify the transcriptional regulators that contribute to modifying PXDN and PXDNL expression.

Moreover, additional investigation may determine if there is a correlation that exists between *PXDN* and *PXDNL* promoter methylation, subsequent protein expression and novel marker capability for diagnosis, prognosis and therapy. This may subsequently lead to increased OSCC patient survival rates by contributing to early diagnosis of OSCC and efficacious targeted therapeutic intervention.

7. References

Adams, G., 2020. A beginner's guide to RT-PCR, qPCR and RT-qPCR. The Biochemist, 42, 48–53. https://doi.org/10.1042/BIO20200034

Agoston, A.T., Argani, P., Yegnasubramanian, S., De Marzo, A.M., Ansari-Lari, M.A., Hicks, J.L., Davidson, N.E., Nelson, W.G., 2005. Increased Protein Stability Causes DNA Methyltransferase 1 Dysregulation in Breast Cancer*. Journal of Biological Chemistry, 280, 18302–18310. https://doi.org/10.1074/jbc.M501675200

Alaouna, M., Hull, R., Penny, C., Dlamini, Z., 2019. Esophageal cancer genetics in South Africa. Clinical and Experimental Gastroenterology, 12, 157–177. https://doi.org/10.2147/CEG.S182000

Allfrey, V.G., Faulkner, R., Mirsky, A.E., 1964. Acetylation and Methylation of Histones and their Possible Role in the Regulation Of RNA Synthesis. Proceedings of the National Academy of Sciences, 51, 786–794. https://doi.org/10.1073/pnas.51.5.786

Argos, M., Chen, L., Jasmine, F., Tong, L., Pierce, B.L., Roy, S., Paul-Brutus, R., Gamble, M.V., Harper, K.N., Parvez, F., Rahman, M., Rakibuz-Zaman, M., Slavkovich, V., Baron, J.A., Graziano, J.H., Kibriya, M.G., Ahsan, H., 2015. Gene-Specific Differential DNA Methylation and Chronic Arsenic Exposure in an Epigenome-Wide Association Study of Adults in Bangladesh. Environmental Health Perspectives, 123, 64–71. https://doi.org/10.1289/ehp.1307884

Auclair, G., Weber, M., 2012. Mechanisms of DNA methylation and demethylation in mammals. Biochimie, Special Section on Epigenetics, 94, 2202–2211. https://doi.org/10.1016/j.biochi.2012.05.016

Baba, Y., Watanabe, M., Baba, H., 2013. A review of the alterations in DNA methylation in esophageal squamous cell carcinoma. Surgery Today, 43, 1355–1364. https://doi.org/10.1007/s00595-012-0451-y

Bae, W., Yoon, T.-Y., Jeong, C., 2021. Direct evaluation of self-quenching behavior of fluorophores at high concentrations using an evanescent field. PLOS ONE, 16, e0247326. https://doi.org/10.1371/journal.pone.0247326

Bathish, B., Paumann-Page, M., Paton, L.N., Kettle, A.J., Winterbourn, C.C., 2020. Peroxidasin mediates bromination of tyrosine residues in the extracellular matrix. Journal of Biological Chemistry, 295, 12697–12705. https://doi.org/10.1074/jbc.RA120.014504

Behrens, J., von Kries, J.P., Kühl, M., Bruhn, L., Wedlich, D., Grosschedl, R., Birchmeier, W., 1996. Functional interaction of β-catenin with the transcription factor LEF-1. Nature, 382, 638–642. https://doi.org/10.1038/382638a0

Berlier, J.E., Rothe, A., Buller, G., Bradford, J., Gray, D.R., Filanoski, B.J., Telford, W.G., Yue, S., Liu, J., Cheung, C.-Y., Chang, W., Hirsch, J.D., Beechem

Rosaria P. Haugland, J.M., Haugland, R.P., 2003. Quantitative Comparison of Long-wavelength Alexa Fluor Dyes to Cy Dyes: Fluorescence of the Dyes and Their Bioconjugates. Journal of Histochemical Society, 51, 1699–1712. https://doi.org/10.1177/002215540305101214

Berry, M.F., 2014. Esophageal cancer: staging system and guidelines for staging and treatment. Journal of Thoracic Disease, 6, 289-297. https://doi.org/10.3978/j.issn.2072-1439.2014.03.11

Bestor, T., Laudano, A., Mattaliano, R., Ingram, V., 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells: The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. Journal of Molecular Biology, 203, 971–983. https://doi.org/10.1016/0022-2836(88)90122-2

Bestor, T.H., Verdine, G.L., 1994. DNA methyltransferases. Current Opinion in Cell Biology, 6, 380–389. https://doi.org/10.1016/0955-0674(94)90030-2

Bey, E., Alexander, J., Whitcutt, J.M., Hunt, J.A., Gear, J.H., 1976. Carcinoma of the esophagus in Africans: establishment of a continuously growing cell line from a tumor specimen. In Vitro, 12, 107–114. https://doi.org/10.1007/BF02796356

Bhanot, P., Brink, M., Samos, C.H., Hsieh, J.-C., Wang, Y., Macke, J.P., Andrew, D., Nathans, J., Nusse, R., 1996. A new member of the frizzled family from Drosophila functions as a Wingless receptor. Nature, 382, 225–230. https://doi.org/10.1038/382225a0

Bhave, G., Cummings, C.F., Vanacore, R.M., Kumagai-Cresse, C., Ero-Tolliver, I.A., Rafi, M., Kang, J.-S., Pedchenko, V., Fessler, L.I., Fessler, J.H., Hudson, B.G., 2012. Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis. Nature Chemical Biology, 8, 784–790. https://doi.org/10.1038/nchembio.1038

Bird, A.P., 1986. CpG-rich sites and the function of DNA methylation. Nature, 321, 209–213. https://doi.org/10.1038/321209a0

Bird, A.P., Wolffe, A.P., 1999. Methylation-Induced Repression— Belts, Braces, and Chromatin. Cell, 99, 451–454. https://doi.org/10.1016/S0092-8674(00)81532-9

Bisulfite Beginners Guide 101 | ZYMO RESEARCH (no date) Zymo Research International. Available at: https://zymoresearch.eu/pages/bisulfite-beginner-guide (Accessed: 23 March 2024).

Blot, W.J. and Tarone, R.E., 2018. Esophageal Cancer In. Schottenfeld and Fraumeni Cancer Epidemiology and Prevention.

Blot, W.J., Devesa, S.S., Fraumeni, J.F., Jr, 1993. Continuing Climb in Rates of Esophageal Adenocarcinoma: An Update. JAMA, 270, 1320. https://doi.org/10.1001/jama.1993.03510110056031

Boecker, D., Zhang, Z., Breves, R., Herth, F., Kramer, A., Bulitta, C., 2023. Antimicrobial efficacy, mode of action and in vivo use of hypochlorous acid (HOCl) for prevention or therapeutic support of infections. GMS Hygiene and Infection Control, 18, Doc07. https://doi.org/10.3205/dgkh000433

Bramhall, S., Noack, N., Wu, M., Loewenberg, J.R., 1969. A simple colorimetric method for determination of protein. Analytical Biochemistry 31, 146–148. https://doi.org/10.1016/0003-2697(69)90251-6

Bray, F., Laversanne, M., Weiderpass, E., Soerjomataram, I., 2021. The ever-increasing importance of cancer as a leading cause of premature death worldwide. Cancer, 127, 3029–3030. https://doi.org/10.1002/cncr.33587

Bréchot, J.M., 1998. [Cell cycle, tumor suppressor genes and lung cancers]. Rev Mal Respir 15, 461–466. PMID: 9805756.

Brown, T.A., 2002. Mutation, Repair and Recombination, in: Genomes. 2nd Edition. Wiley-Liss, Oxford. https://www.ncbi.nlm.nih.gov/books/NBK21114/

Brunner, E., Peter, O., Schweizer, L., Basler, K., 1997. pangolinencodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in Drosophila. Nature, 385, 829–833. https://doi.org/10.1038/385829a0

Burry, R.W., 2011. Controls for Immunocytochemistry. Journal of Histochemistry and Cytochemistry, 59, 6–12. https://doi.org/10.1369/jhc.2010.956920

Cancer Statistics. 2012. CANSA - The Cancer Association of South Africa. [Online] Available at: https://cansa.org.za/south-african-cancer-statistics/ [Accessed 4 April 2022].

Cedar, H., 1988. DNA methylation and gene activity. Cell, 53, 3–4. https://doi.org/10.1016/0092-8674(88)90479-5

Chance, B., Schoener, B., Oshino, R., Itshak, F., Nakase, Y., 1979. Oxidation-reduction ratio studies of mitochondria in freeze-trapped samples. NADH and flavoprotein fluorescence signals. Journal of Biological Chemistry, 254, 4764–4771. https://doi.org/10.1016/S0021-9258(17)30079-0

Cheng, G., Li, H., Cao, Z., Qiu, X., McCormick, S., Thannickal, V.J., Nauseef, W.M., 2011. Vascular peroxidase-1 is rapidly secreted, circulates in plasma, and supports dityrosine cross-linking reactions. Free Radical Biology and Medicine, 51, 1445–1453. https://doi.org/10.1016/j.freeradbiomed.2011.07.002

Cheng, G., Salerno, J.C., Cao, Z., Pagano, P.J., Lambeth, J.D., 2008. Identification and characterization of VPO1, a new animal heme-containing peroxidase. Free Radical Biology and Medicine, 45, 1682–1694. https://doi.org/10.1016/j.freeradbiomed.2008.09.009

Comb, M., Goodman, H.M., 1990. CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. Nucleic Acids Research, 18, 3975–3982. https://doi.org/10.1093/nar/18.13.3975

Cook, M.B., 2011. Non-Acid Reflux: The Missing Link Between Gastric Atrophy and Esophageal Squamous Cell Carcinoma?. Official journal of the American College of Gastroenterology, 106, 1930–1932. https://doi.org/10.1038/ajg.2011.288

Cooper, G.M., 2000. The Development and Causes of Cancer, in: The Cell: A Molecular Approach. 2nd Edition. Sinauer Associates, Sunderland (MA). https://www.ncbi.nlm.nih.gov/books/NBK9963/

Cooper, G.M., 2000. Tumor Suppressor Genes, in: The Cell: A Molecular Approach. 2nd Edition. Sinauer Associates, Sunderland (MA). https://www.ncbi.nlm.nih.gov/books/NBK9963/

Crossley, B.M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M.L., Clement, T., Toohey-Kurth, K., 2020. Guidelines for Sanger sequencing and molecular assay monitoring. Journal of Veterinary Diagnostic Investigation, 32, 767–775. https://doi.org/10.1177/1040638720905833

Cui, H., Horon, I.L., Ohlsson, R., Hamilton, S.R., Feinberg, A.P., 1998. Loss of imprinting in normal tissue of colorectal cancer patients with microsatellite instability. Nature Medicine, 4, 1276–1280. https://doi.org/10.1038/3260

Darst, R.P., Pardo, C.E., Ai, L., Brown, K.D., Kladde, M.P., 2010. Bisulfite Sequencing of DNA. Current Protocols in Molecular Biology, Unit-7.917. https://doi.org/10.1002/0471142727.mb0709s91

Deyl, Z., Macek, K., Adam, M., Vancíková, O., 1980. Studies on the chemical nature of elastin fluorescence. Biochimica et Biophysica Acta (BBA) - Protein Structure, 625, 248–254. https://doi.org/10.1016/0005-2795(80)90288-3

Di, Y., Chen, D., Yu, W., Yan, L., 2019. Bladder cancer stage-associated hub genes revealed by WGCNA co-expression network analysis. Hereditas, 156, 7. https://doi.org/10.1186/s41065-019-0083-y

DiMarco, T., Giulivi, C., 2007. Current analytical methods for the detection of dityrosine, a biomarker of oxidative stress, in biological samples. Mass Spectrometry Reviews, 26, 108–120. https://doi.org/10.1002/mas.20109

Dobrovic, A., Simpfendorfer, D., 1997. Methylation of the BRCA1 Gene in Sporadic Breast Cancer1. Cancer Research, 57, 3347–3350.

Dougan, J., Hawsawi, O., Burton, L.J., Edwards, G., Jones, K., Zou, J., Nagappan, P., Wang, G., Zhang, Q., Danaher, A., Bowen, N., Hinton, C., Odero-Marah, V.A., 2019. Proteomics-Metabolomics Combined Approach Identifies Peroxidasin as a Protector against Metabolic and Oxidative Stress in Prostate

Cancer. International Journal of Molecular Sciences, 20, 3046. https://doi.org/10.3390/ijms20123046

Ehrlich, M., 2002. DNA methylation in cancer: too much, but also too little. Oncogene, 21, 5400–5413. https://doi.org/10.1038/sj.onc.1205651

Ehrlich, M., 2019. DNA hypermethylation in disease: mechanisms and clinical relevance. Epigenetics, 14, 1141–1163. https://doi.org/10.1080/15592294.2019.1638701

Ehrlich, M., Woods, C.B., Yu, M.C., Dubeau, L., Yang, F., Campan, M., Weisenberger, D.J., Long, T., Youn, B., Fiala, E.S., Laird, P.W., 2006. Quantitative analysis of associations between DNA hypermethylation, hypomethylation, and DNMT RNA levels in ovarian tumors. Oncogene, 25, 2636–2645. https://doi.org/10.1038/sj.onc.1209145

English, D.R., Armstrong, B.K., 1988. Identifying people at high risk of cutaneous malignant melanoma: results from a case-control study in Western Australia. Br Med J (Clin Res Ed), 296, 1285–1288. https://doi.org/10.1136/bmj.296.6632.1285

Ero-Tolliver, I.A., Hudson, B.G., Bhave, G., 2015. The Ancient Immunoglobulin Domains of Peroxidasin Are Required to Form Sulfilimine Cross-links in Collagen IV*. Journal of Biological Chemistry, 290, 21741–21748. https://doi.org/10.1074/jbc.M115.673996

Estécio, M.R.H., Gharibyan, V., Shen, L., Ibrahim, A.E.K., Doshi, K., He, R., Jelinek, J., Yang, A.S., Yan, P.S., Huang, T.H.-M., Tajara, E.H., Issa, J.-P.J., 2007. LINE-1 hypomethylation in cancer is highly variable and inversely correlated with microsatellite instability. PLOS One, 2, e399. https://doi.org/10.1371/journal.pone.0000399

Evans, J.P., Skrzynia, C., Burke, W., 2001. The complexities of predictive genetic testing. BMJ, 322, 1052–1056. https://doi.org/10.1136/bmj.322.7293.1052

Feinberg, A.P., 2007. Phenotypic plasticity and the epigenetics of human disease. Nature, 447, 433–440. https://doi.org/10.1038/nature05919

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., Paul, C.L., 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proceedings of the National Academy of Sciences, 89, 1827–1831. https://doi.org/10.1073/pnas.89.5.1827

Furdas, S.D., Kannan, S., Sippl, W., Jung, M., 2012. Small Molecule Inhibitors of Histone Acetyltransferases as Epigenetic Tools and Drug Candidates. Archiv der Pharmazie, 345, 7–21. https://doi.org/10.1002/ardp.201100209

Furst, A.L., Barton, J.K., 2015. DNA Electrochemistry Shows DNMT1 Methyltransferase Hyperactivity in Colorectal Tumors. Chemistry & Biology, 22, 938–945. https://doi.org/10.1016/j.chembiol.2015.05.019

Gambari, R., Del Senno, L., Barbieri, R., Buzzoni, D., Gustafsson, K., Giacomini, P., Natali, P.G., 1986. Lack of correlation between hypomethylation and expression of the HLA-DR alpha gene. European Journal of Immunology, 16, 365–369. https://doi.org/10.1002/eji.1830160409

Georgakoudi, I., Jacobson, B.C., Müller, M.G., Sheets, E.E., Badizadegan, K., Carr-Locke, D.L., Crum, C.P., Boone, C.W., Dasari, R.R., Van Dam, J., Feld, M.S., 2002. NAD(P)H and collagen as in vivo quantitative fluorescent biomarkers of epithelial precancerous changes. Cancer Research, 62, 682–687. PMID: 11830520.

Gersten, O., Wilmoth, J.R., 2002. The Cancer Transition in Japan since 1951. Demographic Research 7, 271–306. https://doi.org/10.4054/DemRes.2002.7.5

Goel, S., DeCristo, M.J., McAllister, S.S., Zhao, J.J., 2018. CDK4/6 Inhibition in Cancer: Beyond Cell Cycle Arrest. Trends in Cell Biology, 28, 911–925. https://doi.org/10.1016/j.tcb.2018.07.002

Gombert, S., Jahn, K., Pathak, H., Burkert, A., Schmidt, G., Wiehlmann, L., Davenport, C., Brändl, B., Müller, F.-J., Leffler, A., Deest, M., Frieling, H., 2023. Comparison of methylation estimates obtained via MinION nanopore sequencing and sanger bisulfite sequencing in the TRPA1 promoter region. BMC Medical Genomics, 16, 257. https://doi.org/10.1186/s12920-023-01694-6

Grady, W.M., 2021. Chapter Eleven - Epigenetic alterations in the gastrointestinal tract: Current and emerging use for biomarkers of cancer, in: Berger, F.G., Boland, C.R. (Eds.), Advances in Cancer Research, Novel Approaches to Colorectal Cancer. Academic Press, 425–468. https://doi.org/10.1016/bs.acr.2021.02.006

Gu, S.-Q., Gallego-Perez, D., McClory, S.P., Shi, J., Han, J., Lee, L.J., Schoenberg, D.R., 2016. The human PMR1 endonuclease stimulates cell motility by down regulating miR-200 family microRNAs. RNA, 44, 5811–5819. https://doi.org/10.1093/nar/gkw497

Guerin, C.J., 2023. Using Antibodies in Microscopy: A Guide to Immunohistochemistry. Part 2: IHC Staining Protocols. Microscopy Today, 31, 34–39. https://doi.org/10.1093/mictod/qaad029

Gujar, H., Weisenberger, D.J., Liang, G., 2019. The Roles of Human DNA Methyltransferases and Their Isoforms in Shaping the Epigenome. Genes (Basel), 10, 172. https://doi.org/10.3390/genes10020172

Guo, M., Ren, J., House, M.G., Qi, Y., Brock, M.V., Herman, J.G., 2006. Accumulation of Promoter Methylation Suggests Epigenetic Progression in

Squamous Cell Carcinoma of the Esophagus. Clinical Cancer Research, 12, 4515–4522. https://doi.org/10.1158/1078-0432.CCR-05-2858

Hanahan, D., Weinberg, R.A., 2000. The Hallmarks of Cancer. Cell, 100, 57–70. https://doi.org/10.1016/S0092-8674(00)81683-9

Hanahan, D., Weinberg, R.A., 2011. Hallmarks of Cancer: The Next Generation. Cell, 144, 646–674. https://doi.org/10.1016/j.cell.2011.02.013

Hanmer, K.L. and Mavri-Damelin, D. (2018) 'Peroxidasin is a novel target of the redox-sensitive transcription factor Nrf2', Gene, 674, 104–114. Available at: https://doi.org/10.1016/j.gene.2018.06.076.

Hayatsu, H., Wataya, Y., Kai, K., Iida, S., 1970. Reaction of sodium bisulfite with uracil, cytosine, and their derivatives. Biochemistry, 9, 2858–2865. https://doi.org/10.1021/bi00816a016

He, C., Song, W., Weston, T.A., Tran, C., Kurtz, I., Zuckerman, J.E., Guagliardo, P., Miner, J.H., Ivanov, S.V., Bougoure, J., Hudson, B.G., Colon, S., Voziyan, P.A., Bhave, G., Fong, L.G., Young, S.G., Jiang, H., 2020. Peroxidasin-mediated bromine enrichment of basement membranes. Proceedings of the National Academy of Sciences, 117, 15827–15836. https://doi.org/10.1073/pnas.2007749117

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C.-X., Zhang, K., He, C., Xu, G.-L., 2011. Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. Science, 333, 1303–1307. https://doi.org/10.1126/science.1210944

Hellmark, T., Segelmark, M., 2007. 68 – Glomerular Basement Membrane Autoantibodies. Autoantibodies, 2, 553–559. https://doi.org/10.1016/B978-044452763-9/50072-X

Henikoff, S., 2000. Heterochromatin function in complex genomes. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 1470, 1–8. https://doi.org/10.1016/S0304-419X(99)00034-7

Hitchins, M.P., 2010. 8 - Inheritance of Epigenetic Aberrations (Constitutional Epimutations) in Cancer Susceptibility, in: Herceg, Z., Ushijima, T. (Eds.), Advances in Genetics, Epigenetics and Cancer, Part A. Academic Press, 201–243. https://doi.org/10.1016/B978-0-12-380866-0.60008-3

Hitchins, M.P., Rapkins, R.W., Kwok, C.-T., Srivastava, S., Wong, J.J.L., Khachigian, L.M., Polly, P., Goldblatt, J., Ward, R.L., 2011. Dominantly Inherited Constitutional Epigenetic Silencing of MLH1 in a Cancer-Affected Family Is Linked to a Single Nucleotide Variant within the 5′UTR. Cancer Cell, 20, 200–213. https://doi.org/10.1016/j.ccr.2011.07.003

Horsthemke, B., 2006. Epimutations in Human Disease, in: Doerfler, W., Böhm, P. (Eds.), DNA Methylation: Development, Genetic Disease and Cancer, Current Topics in Microbiology and Immunology. Springer, Berlin, Heidelberg, 45–59. https://doi.org/10.1007/3-540-31181-5_4

Hoshimoto, S., Takeuchi, H., Ono, S., Sim, M.S., Huynh, J.L., Huang, S.K., Marzese, D.M., Kitagawa, Y., Hoon, D.S.B., 2015. Genome–Wide Hypomethylation and Specific Tumor-Related Gene Hypermethylation are Associated with Esophageal Squamous Cell Carcinoma Outcome. Journal of Thoracic Oncology, 10, 509–517. https://doi.org/10.1097/JTO.0000000000000441

Huang, T.H.-M., Perry, M.R., Laux, D.E., 1999. Methylation Profiling of CpG Sites in Human Breast Cancer Cells. Human Molecular Genetics, 8, 459–470. https://doi.org/10.1093/hmg/8.3.459

Imai, K., Yamamoto, H., 2008. Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics. Carcinogenesis, 29, 673–680. https://doi.org/10.1093/carcin/bgm228

Isaacson, C., 2005. The change of the staple diet of black South Africans from sorghum to maize (corn) is the cause of the epidemic of squamous carcinoma of the oesophagus. Medical Hypotheses, 64, 658–660. https://doi.org/10.1016/j.mehy.2004.09.019

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., Zhang, Y., 2011. Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. Science, 333, 1300–1303. https://doi.org/10.1126/science.1210597

Jakszyn, P., González, C.A., 2006. Nitrosamine and related food intake and gastric and oesophageal cancer risk: A systematic review of the epidemiological evidence. World Journal of Gastroenterology: WJG, 12, 4296–4303. https://doi.org/10.3748/wjg.v12.i27.4296

Jones, P.A., Baylin, S.B., 2007. The Epigenomics of Cancer. Cell, 128, 683–692. https://doi.org/10.1016/j.cell.2007.01.029

Jones, P.L., Veenstra, G.J.C., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., Wolffe, A.P., 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genetics, 19, 187–191. https://doi.org/10.1038/561

Kalousek, F., Morris, N.R., 1969. The Purification and Properties of Deoxyribonucleic Acid Methylase from Rat Spleen. Journal of Biological Chemistry, 244, 1157–1163. https://doi.org/10.1016/S0021-9258(18)91823-5

Kauppila, S., Stenbäck, F., Risteli, J., Jukkola, A., Risteli, L., 1998. Aberrant type I and type III collagen gene expression in human breast cancer in vivo. The

Journal of Pathology, 186, 262–268. https://doi.org/10.1002/(SICI)1096-9896(1998110)186:3<262::AID-PATH191>3.0.CO;2-3

Khan, S.A., Tyagi, M., Sharma, A.K., Barreto, S.G., Sirohi, B., Ramadwar, M., Shrikhande, S.V., Gupta, S., 2014. Cell-type specificity of β-actin expression and its clinicopathological correlation in gastric adenocarcinoma. World Journal of Gastroenterology, 20, 12202–12211. https://doi.org/10.3748/wjg.v20.i34.12202

Kim, J., Dang, C.V., 2006. Cancer's Molecular Sweet Tooth and the Warburg Effect. Cancer Research, 66, 8927–8930. https://doi.org/10.1158/0008-5472.CAN-06-1501

Kiselev, I.S., Kulakova, O.G., Boyko, A.N., Favorova, O.O., 2021. DNA Methylation as an Epigenetic Mechanism in the Development of Multiple Sclerosis. Acta Naturae, 13, 45–57. https://doi.org/10.32607/actanaturae.11043

Knudson, A.G., 1989. Nakahara memorial lecture. Hereditary cancer, oncogenes, and anti-oncogenes. Princess Takamatsu Symposia, 20, 15–29. PMID: 2577335.

Knudson, A.G., 2002. Cancer genetics. American Journal of Medical Genetics, 111, 96–102. https://doi.org/10.1002/ajmg.10320

Komori, H., Lamere, S., Torkamani, A., Hart, T., Kotsopoulos, S., Warner, J., Samuels, M., Olson, J., Head, S., Ordoukhanian, P., Lee, P., Link, D., Salomon, D., 2011. Application of microdroplet PCR for large-scale targeted bisulfite sequencing. Genome research, 21, 1738–45. https://doi.org/10.1101/gr.116863.110

Kurihara-Shimomura, M., Sasahira, T., Shimomura, H., Kirita, T., 2020. Peroxidan Plays a Tumor-Promoting Role in Oral Squamous Cell Carcinoma. International Journal of Molecular Sciences, 21, 5416. https://doi.org/10.3390/ijms21155416

Lakshmi Ch, N.P., Sivagnanam, A., Raja, S., Mahalingam, S., 2021. Molecular basis for RASSF10/NPM/RNF2 feedback cascade–mediated regulation of gastric cancer cell proliferation. The Journal of Biological Chemistry, 297, 100935. https://doi.org/10.1016/j.jbc.2021.100935

Laszlo, A.H., Derrington, I.M., Brinkerhoff, H., Langford, K.W., Nova, I.C., Samson, J.M., Bartlett, J.J., Pavlenok, M., Gundlach, J.H., 2013. Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. Proceedings of the National Academy of Sciences, 110, 18904–18909. https://doi.org/10.1073/pnas.1310240110

Leonhardt, H., Page, A.W., Weier, H.-U., Bestor, T.H., 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell, 71, 865–873. https://doi.org/10.1016/0092-8674(92)90561-P

Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F., Bird, A., 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell, 69, 905–914. https://doi.org/10.1016/0092-8674(92)90610-o

Li, D., Dandara, C., Parker, M.I., 2010. The 341C/T polymorphism in the GSTP1 gene is associated with increased risk of oesophageal cancer. BMC Genetics, 11, 47. https://doi.org/10.1186/1471-2156-11-47

Li, H., Cao, Z., Moore, D.R., Jackson, P.L., Barnes, S., Lambeth, J.D., Thannickal, V.J., Cheng, G., 2012. Microbicidal Activity of Vascular Peroxidase 1 in Human Plasma via Generation of Hypochlorous Acid. Infection and Immunity, 80, 2528–2537. https://doi.org/10.1128/IAI.06337-11

Li, J., Xu, X., Jiang, Y., Hansbro, N.G., Hansbro, P.M., Xu, J., Liu, G., 2020. Elastin is a key factor of tumor development in colorectal cancer. BMC Cancer 20, 217. https://doi.org/10.1186/s12885-020-6686-x

Li, Y., Tollefsbol, T.O., 2011. DNA methylation detection: Bisulfite genomic sequencing analysis. Methods in Molecular Biology, 791, 11–21. https://doi.org/10.1007/978-1-61779-316-5_2

Li, Yanqing, Jiao, Y., Luo, Z., Li, Yang, Liu, Y., 2019. High peroxidasin-like expression is a potential and independent prognostic biomarker in breast cancer. Medicine (Baltimore), 98. https://doi.org/10.1097/MD.0000000000017703

Liang, G., Chan, M.F., Tomigahara, Y., Tsai, Y.C., Gonzales, F.A., Li, E., Laird, P.W., Jones, P.A., 2002. Cooperativity between DNA Methyltransferases in the Maintenance Methylation of Repetitive Elements. Molecular and Cellular Biology, 22, 480–491. https://doi.org/10.1128/MCB.22.2.480-491.2002

Ligtenberg, M.J.L., Kuiper, R.P., Chan, T.L., Goossens, M., Hebeda, K.M., Voorendt, M., Lee, T.Y.H., Bodmer, D., Hoenselaar, E., Hendriks-Cornelissen, S.J.B., Tsui, W.Y., Kong, C.K., Brunner, H.G., van Kessel, A.G., Yuen, S.T., van Krieken, J.H.J.M., Leung, S.Y., Hoogerbrugge, N., 2009. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3′ exons of TACSTD1. Nature Genetics, 41, 112–117. https://doi.org/10.1038/ng.283

Lin, K., Shen, W., Shen, Z., Wu, Y., Lu, S., 2002. Dietary exposure and urinary excretion of total N-nitroso compounds, nitrosamino acids and volatile nitrosamine in inhabitants of high- and low-risk areas for esophageal cancer in southern China. International Journal of Cancer, 102, 207–211. https://doi.org/10.1002/ijc.10698

Liu, Y., Oakeley, E.J., Sun, L., Jost, J.-P., 1998. Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. Nucleic Acids Research, 26, 1038–1045. https://doi.org/10.1093/nar/26.4.1038

Loots, E., Sartorius, B., Madiba, T.E., Mulder, C.J.J., Clarke, D.L., 2017. Is Clinical Research in Oesophageal Cancer in South Africa in Crisis? A Systematic Review. World Journal of Surgery, 41, 810–816. https://doi.org/10.1007/s00268-016-3778-5

Lorenz, T.C., 2012. Polymerase Chain Reaction: Basic Protocol Plus Troubleshooting and Optimization Strategies. J Vis Exp, 3998. https://doi.org/10.3791/3998

Lu, D., Ma, J., Zhan, Q., Li, Y., Qin, J., Guo, M., 2014. Epigenetic Silencing of RASSF10 Promotes Tumor Growth in Esophageal Squamous Cell Carcinoma. Discovery Medicine, 17, 169–178.

Lu, M., Liu, B., Li, D., Gao, Z., Li, W., Zhou, X., Zhan, H., 2023. PXDNL activates the motility of urothelial bladder carcinoma cells through the Wnt/β-catenin pathway and has a prognostic value. Life Sciences, 312, 121270. https://doi.org/10.1016/j.lfs.2022.121270

Ma, K., Cao, B., Guo, M., 2016. The detective, prognostic, and predictive value of DNA methylation in human esophageal squamous cell carcinoma. Clinical Epigenetics, 8, 43. https://doi.org/10.1186/s13148-016-0210-9

Mahmood, T., Yang, P.-C., 2012. Western Blot: Technique, Theory, and Trouble Shooting. N Am J Med Sci, 4, 429–434. 10.4103/1947-2714.100998. https://doi.org/10.4103/1947-2714.100998

Mahmudi-Azer, S., Lacy, P., Bablitz, B., Moqbel, R., 1998. Inhibition of nonspecific binding of fluorescent-labelled antibodies to human eosinophils. Journal of Immunological Methods, 217, 113–119. https://doi.org/10.1016/S0022-1759(98)00105-7

Maiti, A., Drohat, A.C., 2011. Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine: Potential Implications For Active Demethylation Of Cpg Sites. Journal of Biological Chemistry, 286, 35334–35338. https://doi.org/10.1074/jbc.C111.284620

Marchewka, Z., Piwowar, A., Ruzik, S., Długosz, A., 2017. Glutathione S - transferases class Pi and Mi and their significance in oncology. Advances in Hygiene and Experimental Medicine, 71, 541–550. https://doi.org/10.5604/01.3001.0010.3835

Marcucci, F., Stassi, G., De Maria, R., 2016. Epithelial–mesenchymal transition: a new target in anticancer drug discovery. Nature Review Drug Discovery, 15, 311–325. https://doi.org/10.1038/nrd.2015.13

Mardis, E.R., 2013. Next-Generation Sequencing Platforms. Annual Review of Analytical Chemistry, 6, 287–303. https://doi.org/10.1146/annurev-anchem-062012-092628

McCall, A.S., Cummings, C.F., Bhave, G., Vanacore, R., Page-McCaw, A., Hudson, B.G., 2014. Bromine Is an Essential Trace Element for Assembly of Collagen IV Scaffolds in Tissue Development and Architecture. Cell, 157, 1380–1392. https://doi.org/10.1016/j.cell.2014.05.009

McCarrey, J.R., Lehle, J.D., Raju, S.S., Wang, Y., Nilsson, E.E., Skinner, M.K., 2016. Tertiary Epimutations – A Novel Aspect of Epigenetic Transgenerational Inheritance Promoting Genome Instability. PLOS One, 11. https://doi.org/10.1371/journal.pone.0168038

McCormack, V.A., Menya, D., Munishi, M.O., Dzamalala, C., Gasmelseed, N., Leon Roux, M., Assefa, M., Osano, O., Watts, M., Mwasamwaja, A.O., Mmbaga, B.T., Murphy, G., Abnet, C.C., Dawsey, S.M., Schüz, J., 2017. Informing etiologic research priorities for squamous cell esophageal cancer in Africa: A review of setting-specific exposures to known and putative risk factors. International Journal of Cancer, 140, 259–271. https://doi.org/10.1002/ijc.30292

Medfai, H., Khalil, A., Rousseau, A., Nuyens, V., Paumann-Page, M., Sevcnikar, B., Furtmüller, P.G., Obinger, C., Moguilevsky, N., Peulen, O., Herfs, M., Castronovo, V., Amri, M., Van Antwerpen, P., Vanhamme, L., Zouaoui Boudjeltia, K., 2019. Human peroxidasin 1 promotes angiogenesis through ERK1/2, Akt, and FAK pathways. Cardiovascular Research, 115, 463–475. https://doi.org/10.1093/cvr/cvy179

Molenaar, M., Wetering, M. van de, Oosterwegel, M., Peterson-Maduro, J., Godsave, S., Korinek, V., Roose, J., Destrée, O., Clevers, H., 1996. XTcf-3 Transcription Factor Mediates β-Catenin-Induced Axis Formation in Xenopus Embryos. Cell, 86, 391–399. https://doi.org/10.1016/S0092-8674(00)80112-9

Monici, M., 2005. Cell and tissue autofluorescence research and diagnostic applications, in: Biotechnology Annual Review. Elsevier, 227–256. https://doi.org/10.1016/S1387-2656(05)11007-2

Moretti, A., Bennett, G.A., Logrieco, A., Bottalico, A., Beremand, M.N., 1995. Fertility of Fusarium moniliforme from maize and sorghum related to fumonisin production in Italy. Mycopathologia, 131, 25–29. https://doi.org/10.1007/BF01103900

Mori, M., Adachi, Y., Matsushima, T., Matsuda, H., Kuwano, H. and Sugimachi, K., 1993. Lugol staining pattern and histology of esophageal lesions. American Journal of Gastroenterology (Springer Nature), 88(5). PMID: 7683176.

Murata, A., Baba, Y., Watanabe, M., Shigaki, H., Miyake, K., Ishimoto, T., Iwatsuki, M., Iwagami, S., Yoshida, N., Oki, E., Morita, M., Nakao, M., Baba, H., 2014. IGF2 DMR0 Methylation, Loss of Imprinting, and Patient Prognosis in Esophageal Squamous Cell Carcinoma. Annals of Surgical and Oncology, 21, 1166–1174. https://doi.org/10.1245/s10434-013-3414-7

Nan, X., Campoy, F.J., Bird, A., 1997. MeCP2 Is a Transcriptional Repressor with Abundant Binding Sites in Genomic Chromatin. Cell, 88, 471–481. https://doi.org/10.1016/S0092-8674(00)81887-5

Nan, X., Meehan, R.R., Bird, A., 1993. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. Nucleic Acids Research, 21, 4886–4892. https://doi.org/10.1093/nar/21.21.4886

NCBI. 2022. PXDN peroxidasin [Homo sapiens (human)]. [Online] Available at: <https://www.ncbi.nlm.nih.gov/gene/7837> [Accessed 4 April 2022].

Nelson, R.E., Fessler, L.I., Takagi, Y., Blumberg, B., Keene, D.R., Olson, P.F., Parker, C.G., Fessler, J.H., 1994. Peroxidasin: a novel enzyme-matrix protein of Drosophila development. EMBO J 13, 3438–3447. https://doi.org/10.1002%2Fj.1460-2075.1994.tb06649.x

Okano, M., Bell, D.W., Haber, D.A., Li, E., 1999. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. Cell, 99, 247–257. https://doi.org/10.1016/S0092-8674(00)81656-6

Omran, A.R., 1971. The epidemiologic transition: a theory of the epidemiology of population change. The Milbank Memorial Fund Quarterly, 49, 509-538. PMID: 16279965; PMCID: PMC2690264.

Panchuk-Voloshina, N., Haugland, Rosaria P., Bishop-Stewart, J., Bhalgat, M.K., Millard, P.J., Mao, F., Leung, W.-Y., Haugland, Richard P., 1999. Alexa Dyes, a Series of New Fluorescent Dyes that Yield Exceptionally Bright, Photostable Conjugates. Journal of Histochemistry and Cytochemistry, 47, 1179–1188. https://doi.org/10.1177/002215549904700910

Paulsson, M., 1992. Basement membrane proteins: structure, assembly, and cellular interactions. Critical reviews in biochemistry and molecular biology, 27, 93–127. https://doi.org/10.3109/10409239209082560

Paumann-Page, M., Katz, R.-S., Bellei, M., Schwartz, I., Edenhofer, E., Sevcnikar, B., Soudi, M., Hofbauer, S., Battistuzzi, G., Furtmüller, P.G., Obinger, C., 2017. Pre-steady-state Kinetics Reveal the Substrate Specificity and Mechanism of Halide Oxidation of Truncated Human Peroxidasin 1. Journal of Biological Chemistry, 292, 4583–4592. https://doi.org/10.1074/jbc.M117.775213

Péterfi, Z., Donkó, Á., Orient, A., Sum, A., Prókai, Á., Molnár, B., Veréb, Z., Rajnavölgyi, É., Kovács, K.J., Müller, V., Szabó, A.J., Geiszt, M., 2009. Peroxidasin Is Secreted and Incorporated into the Extracellular Matrix of Myofibroblasts and Fibrotic Kidney. The American Journal of Pathology, 175, 725–735. https://doi.org/10.2353/ajpath.2009.080693

Péterfi, Z., Geiszt, M., 2014. Peroxidasins: novel players in tissue genesis. Trends in Biochemical Sciences, 39, 305–307. https://doi.org/10.1016/j.tibs.2014.05.005

Péterfi, Z., Tóth, Z.E., Kovács, H.A., Lázár, E., Sum, A., Donkó, Á., Sirokmány, G., Shah, A.M., Geiszt, M., 2014. Peroxidasin-like protein: a novel peroxidase homologue in the human heart. Cardiovascular Research, 101, 393–399. https://doi.org/10.1093/cvr/cvt256

Pillay, V., Isaacson, C., Mothobi, P., Hale, M., Tomar, L.K., Tyagi, C., Altini, M., Choonara, Y.E., Kumar, P., 2015. Carcinogenic nitrosamines in traditional beer as the cause of oesophageal squamous cell carcinoma in black South Africans. South African Medical Journal, 105, 656–658. https://doi.org/10.7196/samjnew.7935

Pink, R.C., Bailey, T.A., Iputo, J.E., Sammon, A.M., Woodman, A.C., Carter, D.R.F., 2011. Molecular Basis for Maize as a Risk Factor for Esophageal Cancer in a South African Population via a Prostaglandin E2 Positive Feedback Mechanism. Nutrition and Cancer, 63, 714–721. https://doi.org/10.1080/01635581.2011.570893

Prendergast, G.C., Lawe, D., Ziff, E.B., 1991. Association of Myn, the murine homolog of Max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation. Cell, 65, 395–407. https://doi.org/10.1016/0092-8674(91)90457-A

Rauluseviciute, I., Drabløs, F., Rye, M.B., 2020. DNA hypermethylation associated with upregulated gene expression in prostate cancer demonstrates the diversity of epigenetic regulation. BMC Medical Genomics, 13, 6. https://doi.org/10.1186/s12920-020-0657-6

Rayess, H., Wang, M.B., Srivatsan, E.S., 2012. Cellular senescence and tumor suppressor gene p16. International Journal of Cancer, 130, 1715–1725. https://doi.org/10.1002/ijc.27316

Recio-Boiles, A., Babiker, H.M., 2021. Esophageal Cancer, in: StatPearls. StatPearls Publishing, Treasure Site (FL).

Ronsein, G.E., Winterbourn, C.C., Di Mascio, P., Kettle, A.J., 2014. Cross-linking methionine and amine residues with reactive halogen species. Free Radical Biology and Medicine, 70, 278–287. https://doi.org/10.1016/j.freeradbiomed.2014.01.023

Saksouk, N., Barth, T.K., Ziegler-Birling, C., Olova, N., Nowak, A., Rey, E., Mateos-Langerak, J., Urbach, S., Reik, W., Torres-Padilla, M.-E., Imhof, A., Déjardin, J., 2014. Redundant Mechanisms to Form Silent Chromatin at Pericentromeric Regions Rely on BEND3 and DNA Methylation. Molecular Cell, 56, 580–594. https://doi.org/10.1016/j.molcel.2014.10.001

Sevcnikar, B., Schaffner, I., Chuang, C.Y., Gamon, L., Paumann-Page, M., Hofbauer, S., Davies, M.J., Furtmüller, P.G., Obinger, C., 2020. The leucine-rich repeat domain of human peroxidasin 1 promotes binding to laminin in basement membranes. Archives of Biochemistry and Biophysics, 689, 108443. https://doi.org/10.1016/j.abb.2020.108443

Shao, C. et al. (2009) 'Hemimethylation footprints of DNA demethylation in cancer', Epigenetics : official journal of the DNA Methylation Society, 4, 165–75. Available at: https://doi.org/10.4161/epi.4.3.8277.

Shapiro, Robert., Servis, R.E. and Welcher, Marvin. (1970) 'Reactions of Uracil and Cytosine Derivatives with Sodium Bisulfite', Journal of the American Chemical Society, 92(2), 22–424. Available at: https://doi.org/10.1021/ja00705a626.

Shi, R., Cao, Z., Li, H., Graw, J., Zhang, G., Thannickal, V.J., Cheng, G., 2018. Peroxidasin contributes to lung host defense by direct binding and killing of gram-negative bacteria. PLOS Pathogens, 14. https://doi.org/10.1371/journal.ppat.1007026

Shiraishi, M., Oates, A.J., Sekiya, T., 2002. An Overview of the Analysis of DNA Methylation in Mammalian Genomes, 383, 893–906. https://doi.org/10.1515/BC.2002.096

Sigurdardottir, A.K., Jonasdottir, A.S., Asbjarnarson, A., Helgudottir, H.R., Gudjonsson, T., Traustadottir, G.A., 2021. Peroxidasin Enhances Basal Phenotype and Inhibits Branching Morphogenesis in Breast Epithelial Progenitor Cell Line D492. Journal of Mammary Gland Biology and Neoplasia , 26, 321–338. https://doi.org/10.1007/s10911-021-09507-1

Singal, R., Ginder, G.D., 1999. DNA Methylation. Blood, 93, 4059–4070. https://doi.org/10.1182/blood.V93.12.4059

Smyth, E.C., Lagergren, J., Fitzgerald, R.C., Lordick, F., Shah, M.A., Lagergren, P., Cunningham, D., 2017. Oesophageal cancer. Nature reviews. Disease primers, 3, 17048. https://doi.org/10.1038/nrdp.2017.48

Soares, J., Pinto, A.E., Cunha, C.V., André, S., Barão, I., Sousa, J.M., Cravo, M., 1999. Global DNA hypomethylation in breast carcinoma. Cancer, 85, 112–118. https://doi.org/10.1002/(SICI)1097-0142(19990101)85:1<112::AID-CNCR16>3.0.CO;2-T

Sobin, L.H., Gospodarowicz, M.K. and Wittekind, C., 2009. TNM Classification of Malignant Tumours. Wiley-Blackwell, 7, 455.

Soudi, M., Zamocky, M., Jakopitsch, C., Furtmüller, P.G., Obinger, C., 2012. Molecular Evolution, Structure, and Function of Peroxidasins. Chemistry & Biodiversity, 9, 1776–1793. https://doi.org/10.1002/cbdv.201100438

Stojnev, S., Ristić-Petrović, A., Janković-Velicković, L., 2013. Reactive oxygen species, apoptosis and cancer. Vojnosanit Pregl, 70, 675–678. https://doi.org/10.2298/vsp1307675s

Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F., 2021. Global Cancer Statistics 2020: GLOBOCAN Estimates of

Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians, 71, 209–249. https://doi.org/10.3322/caac.21660

Susan, J.Ci., Harrison, J., Paul, C.L., Frommer, M., 1994. High sensitivity mapping of methylated cytosines. Nucleic Acids Research, 22, 2990–2997. https://doi.org/10.1093/nar/22.15.2990

Suzuki, M.M., Bird, A., 2008. DNA methylation landscapes: provocative insights from epigenomics. Nature Review Genetics, 9, 465–476. https://doi.org/10.1038/nrg2341

Szyf, M., 2001. Towards a pharmacology of DNA methylation. Trends in Pharmacological Sciences, 22, 350–354. https://doi.org/10.1016/S0165-6147(00)01713-2

Taghavi, N., Biramijamal, F., Sotoudeh, M., Khademi, H., Malekzadeh, R., Moaven, O., Memar, B., A'rabi, A., Abbaszadegan, M.R., 2010. p16INK4ahypermethylation and p53, p16 and MDM2 protein expression in Esophageal Squamous Cell Carcinoma. BMC Cancer, 10, 138. https://doi.org/10.1186/1471-2407-10-138

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., Rao, A., 2009. Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science 324, 930–935. https://doi.org/10.1126/science.1170116

Thiery, J.P., Acloque, H., Huang, R.Y.J., Nieto, M.A., 2009. Epithelial-Mesenchymal Transitions in Development and Disease. Cell, 139, 871–890. https://doi.org/10.1016/j.cell.2009.11.007

Tulstrup, M., Soerensen, M., Hansen, J.W., Gillberg, L., Needhamsen, M., Kaastrup, K., Helin, K., Christensen, K., Weischenfeldt, J., Grønbæk, K., 2021. TET2 mutations are associated with hypermethylation at key regulatory enhancers in normal and malignant hematopoiesis. Nature Communications, 12, 6061. https://doi.org/10.1038/s41467-021-26093-2

Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., Sanli, K., von Feilitzen, K., Oksvold, P., Lundberg, E., Hober, S., Nilsson, P., Mattsson, J., Schwenk, J.M., Brunnström, H., Glimelius, B., Sjöblom, T., Edqvist, P.-H., Djureinovic, D., Micke, P., Lindskog, C., Mardinoglu, A., Ponten, F., 2017. A pathology atlas of the human cancer transcriptome. Science, 357. https://doi.org/10.1126/science.aan2507

Veale, R.B., Thornley, A.L., Veale, R.B., Thornley, A.L., 1989. Increased single class low affinity egf receptors expressed by human esophageal squamous carcinoma cell lines. South African Journal of Science, 85, 375–379.

Venning, F.A., Wullkopf, L., Erler, J.T., 2015. Targeting ECM Disrupts Cancer Progression. Frontiers in Oncology, 5, 224. https://doi.org/10.3389/fonc.2015.00224

Vogel, C., Marcotte, E.M., 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nature Reviews Genetics, 13, 227–232. https://doi.org/10.1038/nrg3185

Vogelsang, M., Wang, Y., Veber, N., Mwapagha, L.M., Parker, M.I., 2012. The cumulative effects of polymorphisms in the DNA mismatch repair genes and tobacco smoking in oesophageal cancer risk. PLOS One, 7. https://doi.org/10.1371/journal.pone.0036962

Wang, Y.-C., Lu, Y.-P., Tseng, R.-C., Lin, R.-K., Chang, J.-W., Chen, J.-T., Shih, C.-M., Chen, C.-Y., 2003. Inactivation of hMLH1 and hMSH2 by promoter methylation in primary non-small cell lung tumors and matched sputum samples. Journal of Clinical Investigation, 111, 887–895. https://doi.org/10.1172/JCI15475

Wehrli, M., Dougan, S.T., Caldwell, K., O'Keefe, L., Schwartz, S., Vaizel-Ohayon, D., Schejter, E., Tomlinson, A., DiNardo, S., 2000. arrow encodes an LDL-receptor-related protein essential for Wingless signalling. Nature, 407, 527–530. https://doi.org/10.1038/35035110

Wingo, P.A., Ries, L.A.G., Giovino, G.A., Miller, D.S., Rosenberg, H.M., Shopland, D.R., Thun, M.J., Edwards, B.K., 1999. Annual Report to the Nation on the Status of Cancer, 1973-1996, With a Special Section on Lung Cancer and Tobacco Smoking. JNCI: Journal of the National Cancer Institute, 91, 675–690. https://doi.org/10.1093/jnci/91.8.675

Ying, L., Sharma, A., Chhoda, A., Ruzgar, N., Hasan, N., Kwak, R., Wolfgang, C.L., Wang, T.H., Kunstman, J.W., Salem, R.R., Wood, L.D., Iacobuzio-Donahue, C., Schneider, E.B., Farrell, J.J., Ahuja, N., 2021. Methylation-based Cell-free DNA Signature for Early Detection of Pancreatic Cancer. Pancreas, 50, 1267. https://doi.org/10.1097/MPA.0000000000001919

Young, C.D., Zimmerman, L.J., Hoshino, D., Formisano, L., Hanker, A.B., Gatza, M.L., Morrison, M.M., Moore, P.D., Whitwell, C.A., Dave, B. and Stricker, T., 2015. Activating PIK3CA mutations induce an EGFR/ERK paracrine signaling axis in basal-like breast cancer. Molecular & Cellular Proteomics. PMID: 25953087; PMCID: PMC4587316.

Zeilinger, S., Kühnel, B., Klopp, N., Baurecht, H., Kleinschmidt, A., Gieger, C., Weidinger, S., Lattka, E., Adamski, J., Peters, A., Strauch, K., Waldenberger, M., Illig, T., 2013. Tobacco Smoking Leads to Extensive Genome-Wide Changes in DNA Methylation. PLOehrS One, 8. https://doi.org/10.1371/journal.pone.0063812

Zheng, Y.-Z., Liang, L., 2018. High expression of PXDN is associated with poor prognosis and promotes proliferation, invasion as well as migration in ovarian

cancer. Annals of Diagnostic Pathology, 34, 161–165. https://doi.org/10.1016/j.anndiagpath.2018.03.002

Zhou, X., Sun, Q., Xu, C., Zhou, Z., Chen, X., Zhu, X., Huang, Z., Wang, W., Shi, Y., 2022. A systematic pan-cancer analysis of PXDN as a potential target for clinical diagnosis and treatment. Frontiers in Oncology, 12. PMID: 35982948; PMCID: PMC9380648.

Appendix A

Table A1: Recipes for (A) IF microscopy, (B) DNA extraction and (C) western blot.

(A) IF microscopy recipes	
10% formaldehyde	10% (w/v) formalin in 1X PBS
0.1% Triton X-100	0.1% (v/v) Triton X-100 in 1X PBS
1% BSA	1% (w/v) BSA in 1X PBS

(B) DNA extraction recipes	
1X TAE	40 Mm Tris-HCl at pH 8, 0.11% (v/v) glacial acetic acid and 1 Mm EDTA at pH 8.0. Made up to volumes with dh_2O
50% Molecular grade ethanol	50% (v/v) ethanol in ddh_2O

(C) Western blot recipes	
1X RIPA (stock)	50 mM Tris-HCl at pH 8.0, 150 mM NaCl, 0.1% SDS, 1% Triton X-100 and 0.5% Sodium deoxycholate. Made up to volume with ddh2O
1X RIPA (with protease inhibitors)	989 µl 1X RIPA (stock), 1 Mm PMSF and 10 µM leupeptin
10% APS	10% (w/v) APS
10% SDS	10% (w/v) SDS
20% SDS	20% (w/v) SDS
29:1 acrylamide: bisacrylamide	29 units acrylamide and 1 unit bisacrylamide. Made up to volume with dh_2O
2X PLB	120 mM Tris-HCl at pH 6.8, 4% SDS, 20% Glycerol, 0.2% Bromophenol blue and 2% β-mercaptoethanol. Made up to volume with dh_2O
4X PLB	240 mM Tris-HCl at pH 6.8, 8% SDS, 40% Glycerol, 0.04% Bromophenol blue and 5% β-mercaptoethanol. Made up to volume with dh_2O
5% BSA	5% (w/v) BSA in 1X TBST
7.5% TCA	7.5% (w/v) Trichloroacetic acid. Made up to volume with dh_2O
BSA (1 mg/ml)	1 mg of BSA in 1ml of 1X RIPA (stock)
Coomassie brilliant blue	50% (v/v) methanol, 3 mM Coomassie Brilliant Blue G-250 and 10% (v/v) glacial acetic acid
Destain buffer	10% (v/v) methanol and 12% (v/v) Acetic acid
Elution buffer	1% (v/v) 25% ammonia and 66% (v/v) Methanol
1X TBST	Tris-HCl at pH 7.6, 137 mM NaCl and 0.1%(v/v) Tween-2020 mM
Transfer buffer	25 mM Tris, 20% (v/v) methanol and 192 mM Glycine

Appendix B

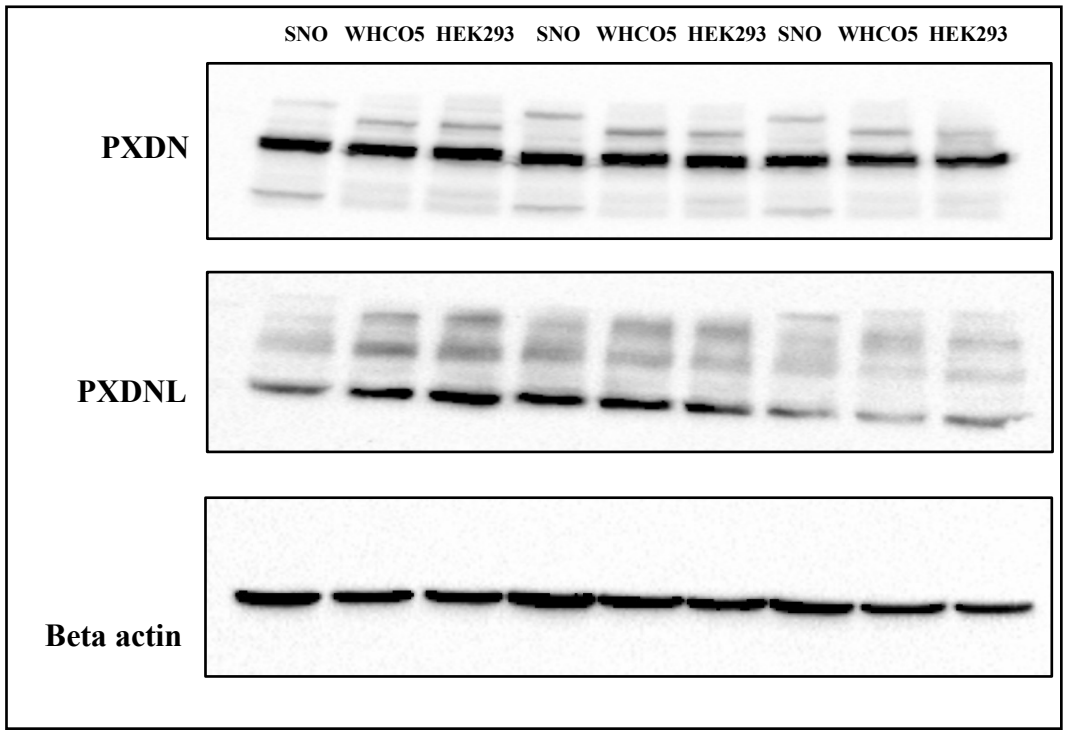
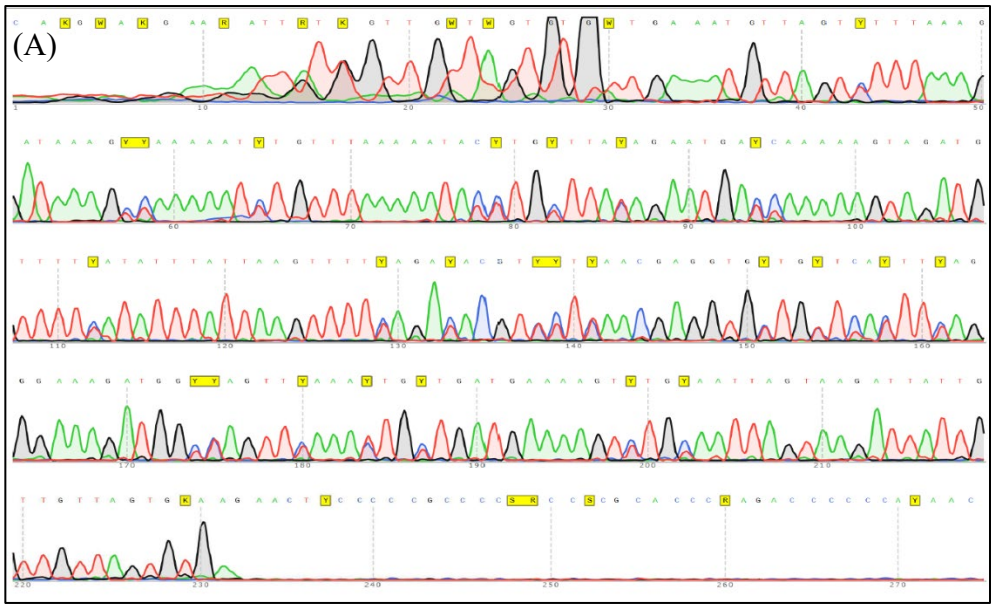
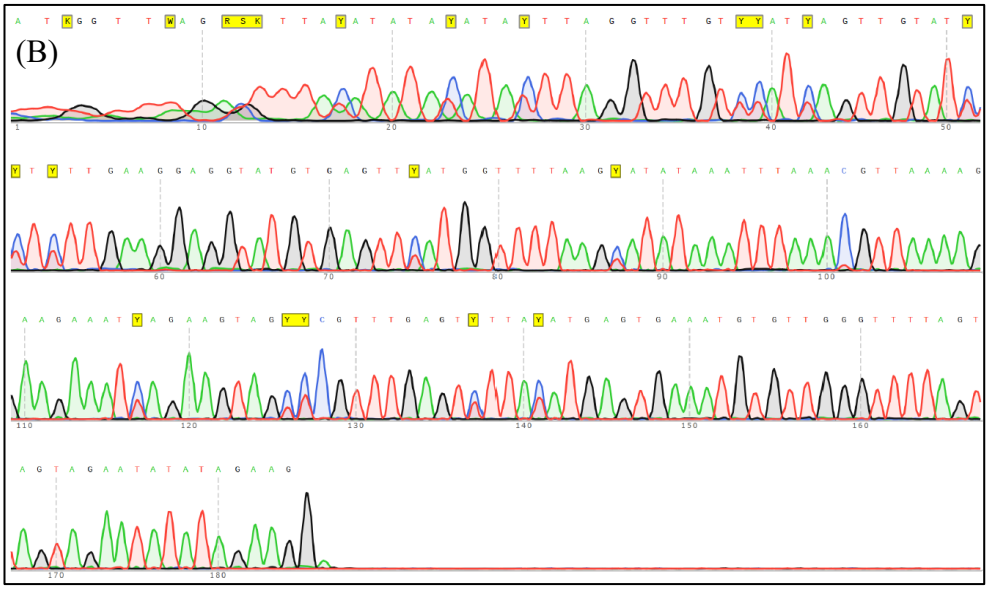


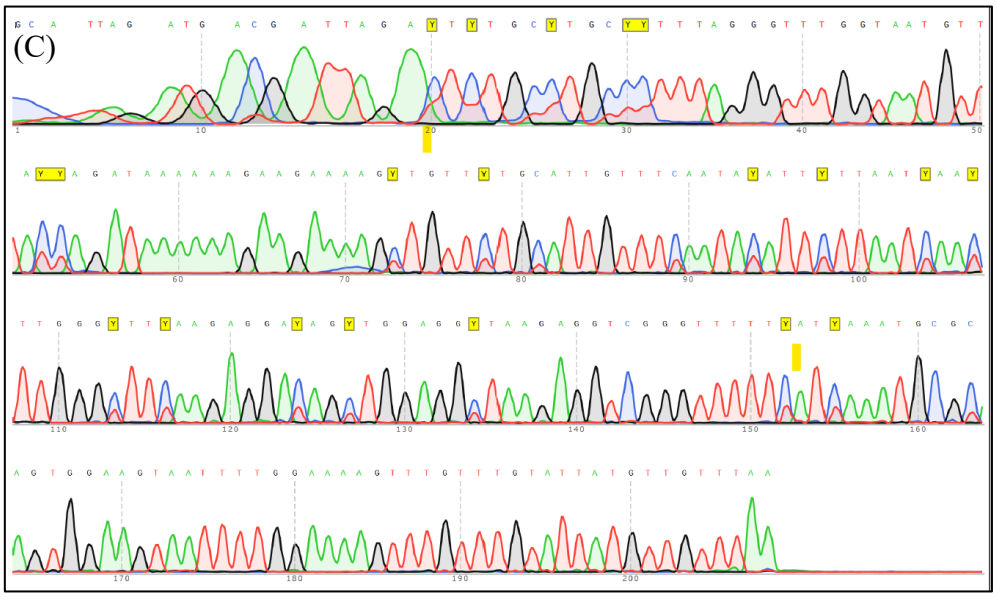
Figure B1: Western blots showing the entire blot and biological triplicate data. The entire western blot shows the proteins of interest as well other proteins that are likely as a result of non-specific primary antibody binding.

Appendix C

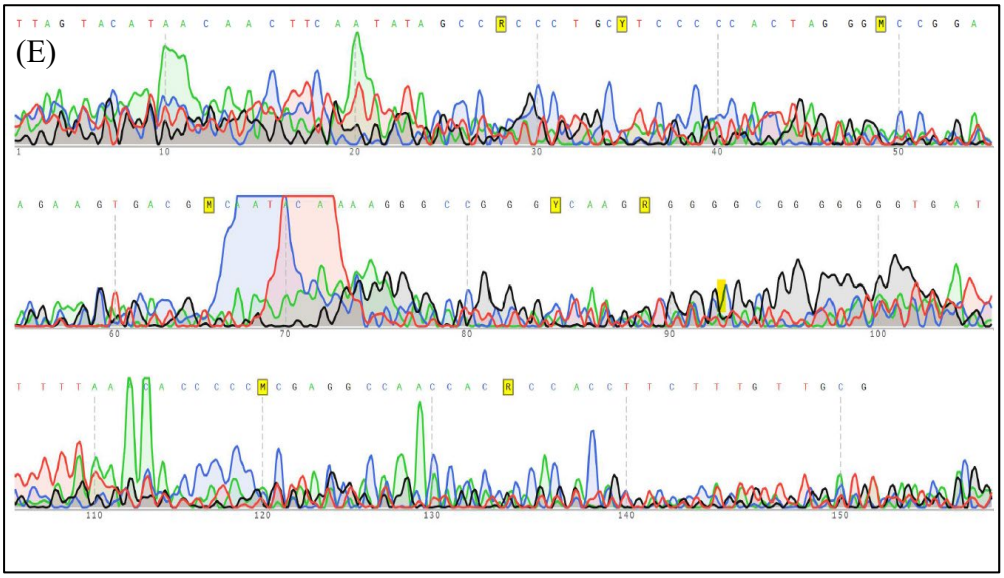
Although bisulfite sequencing primers (primers 5-11) produced amplicons at the correct molecular weight, none of sequencing data matched the predicted bisulfited converted DNA sequence of the *PXDNL* promoter. Many of the bases in the sequencing data were ambiguous and the bases that are shown did not match the predicted bisulfite converted base that should be at that position which rendered the data unusable. All electropherograms showing the forward strand sequencing results of amplicons produced by primers 6-11 are shown below.






(C)


(D)


(E)

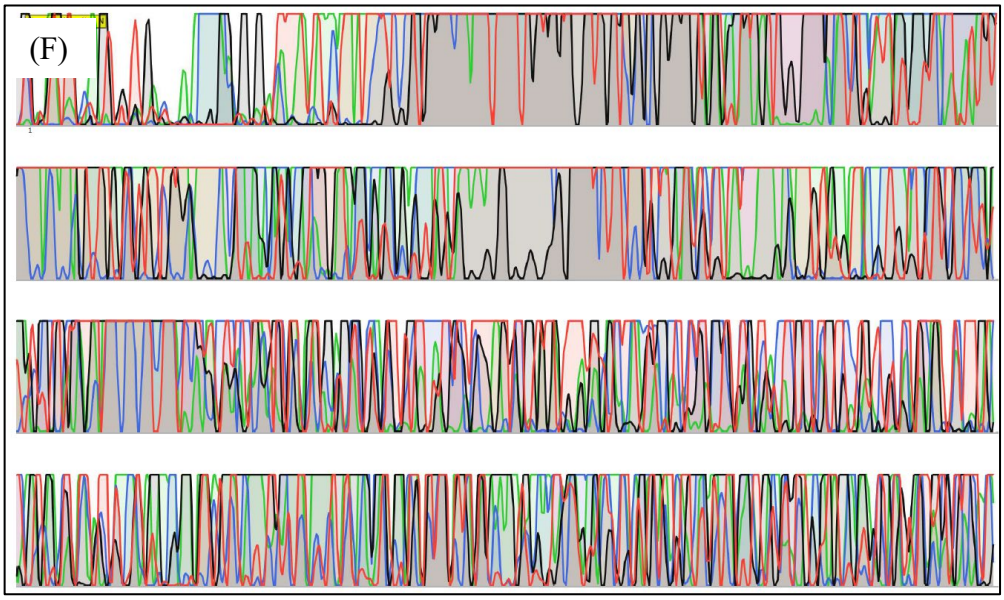


Figure C1: Sequencing results of amplicons produced by primer set 6 (A), 7 (B), 8 (C), 9 (D), 10 (E) and 11(F). Electropherograms show that at least the first 30bp of the sequence presents with overlapping peaks, indicating that the sequencer cannot determine the identity of the base. Bases that were reported by the amplicon sequencing data did not match the predicted sequence of the bisulfite converted *PXDNL* promoter region. Sequencing data for amplicons produced by primer set 10 and 11 were unreadable.

Appendix D

Figure D1: *PXDN* promoter region with highlighted CpG sites. Due to the high amount of CpG sites within the promoter region (including the TSS/untranslated region highlighted in bold), this region was not conducive for bisulfite sequencing. In addition, the MSPCR technique could only investigate the ten CpG sites that form part of the MspI/HpaII recognition sites. Hence, MSPCR cannot provide a holistic understanding of the overall *PXDN* promoter region methylation status. The start of exon 1 is in blue text.